

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE ÓPTICA Y OPTOMETRÍA



TESIS DOCTORAL

**Implementación y precisión de un nuevo método de cálculo
biométrico basado en técnicas de inteligencia artificial**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

David Carmona González

Directora

Nuria Garzón Jiménez

Madrid

©David Carmona González, 2021

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE ÓPTICA Y OPTOMETRÍA

Departamento de Optometría y Visión



TESIS DOCTORAL

IMPLEMENTACIÓN Y PRECISIÓN DE UN NUEVO
MÉTODO DE CÁLCULO BIOMÉTRICO BASADO EN
TÉCNICAS DE INTELIGENCIA ARTIFICIAL

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

David Carmona González

DIRECTORA

Nuria Garzón Jiménez

Madrid, 2021

TESIS DOCTORAL



IMPLEMENTACIÓN Y PRECISIÓN DE UN NUEVO MÉTODO DE CÁLCULO BIOMÉTRICO BASADO EN TÉCNICAS DE INTELIGENCIA ARTIFICIAL

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE ÓPTICA Y OPTOMETRÍA

Departamento de Optometría y Visión

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

David Carmona González

DIRECTORA

Nuria Garzón Jiménez

AGRADECIMIENTOS Y DEDICATORIAS

A mis padres Josea y Maica

A mi mujer Ana

A mis hijas Julia y Marina

Corría 1977 cuando yo tenía 6 años y mi querida abuela Carmen, madre de mi madre, fue operada de cataratas. Le entregaron su cristalino opaco en un pequeño bote con formol para que lo guardara, debía ser costumbre en esa época. No sé por qué misteriosa razón, aquello acabó en casa de mis padres y olvidaron durante algún tiempo en un armario. Cada vez que observaba, y no eran pocas, esa forma lenticular oscura gravitando en el líquido conservante de aquel botecito, me hacía muchas preguntas relacionadas con los ojos.

Cada vez que veía a mi abuela Carmen con aquellas colosales gafas de pasta gruesa y sus ojos multiplicados detrás de ellas, me desconcertaba todavía más y aumentaba mi curiosidad infantil por la cuestión óptica.

En esos años, debido a una anisometropía miópica que derivó temprano en estrabismo intermitente, mi padre se hacía revisiones anuales en el oculista, como le gustaba llamarlo. Era preceptivo acudir con él a estos encuentros que resultaban ser toda una aventura. Aquella consulta en una casa antigua de techos decimonónicos inalcanzables, con la sala colmada y esperas interminables, se convirtió en hábito. Cuando el hastío hacía mella, de repente se abría una puerta y la mujer del Dr. Fernández Díaz avisaba por el nombre, te daba un vuelco el corazón y recobrabas súbitamente la

emoción de volver a entrar en aquella habitación oscura únicamente iluminada por un cartel de letras proyectado sobre la pared.

A ciencia cierta, no sé si estos fueron los motivos de mi actual pasión por mi profesión, pero de lo que sí estoy seguro que algo influyeron subliminalmente en mis decisiones vitales posteriores.

Muchos años más tarde y una vez finalizados mis estudios universitarios de óptica, una tarde de mayo de 1997 cayó en mis manos la tercera edición de *“Oftalmología clínica”* de J. J. Kanski, la biblia. Por aquel entonces llevaba un año trabajando como optometrista clínico en el hospital La Zarzuela y aquel regalo que me habían hecho mis padres desde luego que me permitió avanzar.

La lectura de aquel libro era un acto iniciático, era el santo grial y traía en su anteportada un mensaje manuscrito revelador e inspirador que hoy, 24 años después se culmina con la presentación de esta tesis doctoral.

Mis padres habían dejado escrito *“Estos conocimientos te permitirán prosperar y seguro que algún día consigues el doctorado.”* Visto con perspectiva, fue una osadía *verneriana* por su parte dejar constancia escrita de esto, ya que en aquella época los profesionales de la óptica éramos diplomados universitarios sin posibilidades de acometer otros horizontes académicos, siendo así hasta no hace demasiado. Sin embargo, ahora puedo decir cuánta razón teníais, como siempre, os quiero.

Ana lo es todo para mí desde hace 22 años, que más puedo decir, sobran más palabras. Comprendiste desde el principio la dificultad de emprender la colosal tarea,

quizá debería decir imprudente tarea, de hacer compatible mi vida familiar con la investigación y la ciencia, pero sobre todo la inversión de tiempo que requiere una tesis doctoral y el largo camino recorrido previamente. Me animaste, me incitaste y me has acompañado comenzando el camino juntos hace más de 7 años con la adaptación al grado y aquí estamos, más juntos que nunca. Esta pequeña obra es, sin duda alguna, también gracias a tí y a tus ímprobos esfuerzos, pero sobre todo a tu amor sin límites, te quiero.

Julia y Marina toda nuestra historia se puede explicar gracias a esta explosión de vida que dirige nuestro destino inequívocamente al edén como un imponente viaje al mismo centro de los sentimientos. La inmensa pureza paradigma de ingenuidad que emana de vuestra sonrisa es un auténtico homenaje a nuestra existencia. Sois mi felicidad sin más, sin recovecos ni dobleces, sin ángulos ni estrecheces. Vosotras no lo sabéis todavía, pero cada instante sirve para aprender a ser mejor en el siguiente, y así y solo así, se consigue la plenitud para entender la verdadera razón de nuestra existencia... de vuestra existencia...estoy seguro de que lo váis a conseguir, os quiero mucho.

Me gustaría agradecer eternamente a mi directora de tesis Nuria Garzón Jiménez, en primer lugar por su confianza ciega en este proyecto y en mí, que cuando se lo propuse no dudó un solo segundo en aceptarlo. Por su exorbitante capacidad de trabajo que ha permitido llevar a buen puerto esta tesis, por su comprensión, por su generosidad y por su compromiso. Por último, pero sobre todo y para mí lo más importante, por su amistad y cariño, muchas gracias por todo Nuria, estoy seguro que abordaremos muchos más proyectos juntos, porque es un privilegio y un verdadero placer trabajar a tu lado.

No quiero olvidarme de Carlos Palomino y Alfredo Castillo, los jefes de servicio del departamento de oftalmología en el que desarrollo mi actividad profesional todos los días. Gracias a ellos y a su forma de entender y respetar la optometría como disciplina independiente, a la confianza que me han mostrado continuamente desde que nos conocimos y a la libertad de cátedra otorgada para ejercer la clínica, sigo aprendiendo en estos 15 años compartidos, sigo creciendo como individuo y no merma mi capacidad de asombro día a día.

Quiero mostrar mi enorme gratitud hacia Manuel Rodríguez y Jaime Aramberri por brindarme su ayuda para la consecución de esta investigación. Mención especial a dos grandes profesionales como César Albarrán y Jorge A. Calvo, a los que quiero agradecer su implicación rotunda e incondicional y sus sabios consejos, gracias a los cuales mi caos encontró un sentido.

Por último pero no menos importante, mención especial a mis compañeros del departamento, a los que quiero, admiro y respeto profundamente, y que directa o indirectamente han contribuido aportando valor a esta tesis. Carolina, Margarita, Cristina, Marian, Olivia, Isabel, Ana, Marta, Macarena, Eva y Pedro. A todos, muchas gracias por compartir grandes momentos en el hospital.

RESUMEN

Objetivo: Implementar una fórmula, mediante técnicas de *Machine Learning*, para predecir la potencia refractiva de la lente intraocular que se implantará en una cirugía de cristalino.

Material y método: Se diseña un estudio para estimar la potencia emetropizante de la lente intraocular, modificando por distometría y de forma retrospectiva la potencia real de la lente intraocular implantada. Se implementarán unos modelos regresivos no lineales, que utilicen como variable respuesta la potencia recalculada retrospectivamente y como predictores una serie de *inputs* biométricos.

Se recopiló una muestra de 481 ojos intervenidos de cataratas por 4 cirujanos con distintos tipos de implantes intraoculares. Todos los ojos fueron medidos preoperatoriamente con IOL Master® 700 (Carl Zeiss Meditec AG, Jena, Alemania) y Pentacam® HR (Oculus Optikgeräte GmbH, Wetzlar, Alemania) para la curvatura posterior corneal.

Se preparó y se analizó el *dataset* para la selección de las variables eliminando correlaciones que produjeran colinealidad. La muestra se aleatorizó y posteriormente se dividió en dos partes con una ratio 80/20: entrenamiento (*training*) y validación (*test*) respectivamente.

Se implementaron en crudo modelos regresivos mediante técnicas de *Machine Learning* y se entrenaron con la partición *training*. Posteriormente se optimizaron e hiperparametrizaron para mejorar y potenciar la predictibilidad. Se utilizaron técnicas de mejora (*stacking*) para ensamblar los mejores modelos finales, obteniendo un modelo definitivo al que se llamó Karmona®.

Para la validación del modelo y comprobación del posible sobreajuste, Karmona® se probó sobre la muestra *test* de 95 ojos correspondientes a la partición de validación. En segunda instancia, los resultados sobre esta muestra *test* se compararon con una selección de 12 fórmulas: Hoffer Q, SRK/T, Holladay I, Haigis, Holladay II, Olsen, Barrett Universal II, Ladas Superformula, Kane, Hill-RBF v3.0, EVO v2.0 y PEARL-DGS V2.0, intentando escoger transversalmente las mejores y más utilizadas fórmulas y métodos de cálculo.

Por último, para reforzar la comparación anterior, se optó por conformar dos bases de datos accesorias con muestras de ojos por debajo de 22.00 mm de longitud axial y otra de ojos biométricamente atípicos, por considerarse los puntos críticos en los que las fórmulas fallan. Se comparó el rendimiento en estas muestras entre Karmona®, Haigis, Barrett Universal II, EVO v3.0, Hill-RBF v3.0 y PEARL-DGS.

Resultados: Karmona® mostró el mejor rendimiento sobre la muestra de validación correspondiente al 20% del dataset original (n = 95), obteniendo un MAE ± SD (MedAE) de 0.23 ± 0.20 D (0.19 D) con un 92% de ojos incluidos en el intervalo refractivo postoperatorio de ± 0.50 D, seguida de PEARL-DGS con 0.27 ± 0.25 D (0.22 D) con un 88% de ojos en ± 0.50 D, sin diferencias estadísticamente significativas. Por otro lado, Karmona® consiguió el mejor resultado sobre la muestra accesorias de ojos cortos (n = 33), con 0.38 ± 0.28 D (0.33 D) con un 61% de ojos incluidos en ± 0.50 D, seguida de EVO v2.0 con 0.40 ± 0.29 D (0.34 D) con un 67% de ojos en el rango de ± 0.50 D. No hubo diferencias estadísticamente significativas entre fórmulas en las comparaciones hechas en este grupo. Karmona®, de nuevo obtuvo el mejor resultado en la otra muestra

accesoria de ojos atípicos ($n = 41$) con 0.35 ± 0.47 D (0.20 D) con un 76% en ± 0.50 D, seguida de EVO v2.0 con 0.42 ± 0.52 D (0.22 D) con un 79% de ojos incluidos en el intervalo ± 0.50 D, habiendo diferencias estadísticamente significativas entre Karmona® y Haigis únicamente.

Conclusiones: Karmona® ha mostrado mejores resultados que el resto de las fórmulas con las que se ha comparado sin diferencias estadísticamente significativas, para la población estudiada en este trabajo.

ABSTRACT

Purpose: To implement a formula using Machine Learning techniques to predict the refractive power of the intraocular lens to be implanted in lens surgery.

Material and methods: A study is designed to estimate the refractive power of the intraocular lens, back-calculated the real power of the implanted intraocular lens retrospectively by distometry. Non-linear regression models will be implemented, using the retrospectively back-calculated power as the response variable and biometric inputs as predictors.

A sample of 481 eyes undergoing uneventful cataract surgery by 4 surgeons with different types of intraocular implants was collected. All eyes were measured preoperatively with IOL Master® 700 (Carl Zeiss Meditec AG, Jena, Germany) and Pentacam® HR (Oculus Optikgeräte GmbH, Wetzlar, Germany) for corneal posterior curvature.

The dataset was prepared and analyzed for variable selection by eliminating correlations that produced collinearity. The sample was randomized and subsequently split into two parts with a ratio 80 - 20: training and test.

The regression models were implemented in raw form using Machine Learning techniques and trained with the training partition. Subsequently, they were optimized and hyperparameterized to improve and enhance predictability. Stacking techniques were used to assemble the best final models, obtaining a final model called Karmona®.

To validate the model and check for possible overfitting, Karmona® was tested on the test sample of 95 eyes corresponding to the validation partition. Secondly, the results on this test sample were compared with a selection of 12 formulae: Hoffer Q, SRK/T,

Holladay I, Haigis, Holladay II, Olsen, Barrett Universal II, Ladas Superformula, Kane, Hill-RBF v3.0, EVO v2.0 and PEARL-DGS V2.0, trying to choose cross-sectionally the best and most used formulae and calculation methods.

Finally, to reinforce the previous comparison, we chose to form two accessory databases with samples of eyes below 22.00 mm axial length and another of biometrically atypical eyes, as these are considered the critical points where the formulas fail. Performance on these samples was compared between Karmona[®], Haigis, Barrett Universal II, EVO v3.0, Hill-RBF v3.0 and PEARL-DGS.

Results: Karmona[®] showed the best performance on the validation sample corresponding to 20% of the original dataset (n = 95), obtaining a MAE ± SD (MedAE) of 0.23 ± 0.20 D (0.19 D) and 92% of eyes within ±0.50 D, followed by PEARL-DGS with 0.27 ± 0.25 D (0.22 D) and 88% of eyes within ±0.50 D, with no statistically significant differences. On the other hand, Karmona[®] achieved the best performance on the accessory sample of short eyes (n = 33), with 0.38 ± 0.28 D (0.33 D) and 61% of eyes within ±0.50 D, followed by EVO v2.0 with 0.40 ± 0.29 D (0.34 D) and 67% of eyes within ±0.50 D. Karmona[®] again had the best performance in the other accessory sample of atypical eyes (n = 41) with 0.35 ± 0.47 D (0.20 D) and 76% of eyes within ±0.50 D, followed by EVO v2.0 with 0.42 ± 0.52 D (0.22 D) and 79% within ±0.50 D range. There were no statistically significant differences between formulas.

Conclusions: Karmona[®] has shown better results than the rest of the formulas with which it has been compared, for the population studied in this work.

ÍNDICE

AGRADECIMIENTOS Y DEDICATORIAS	i
RESUMEN.....	vii
ABSTRACT.....	xiii
ÍNDICE	xvii
LISTA DE SÍMBOLOS, ABREVIATURAS Y ACRÓNIMOS	xxiii
LISTA DE FIGURAS	xxix
LISTA DE TABLAS	xlv
1 JUSTIFICACIÓN.....	1
2 HIPÓTESIS Y OBJETIVOS	7
2.1 Hipótesis de trabajo	9
2.2 Objetivo principal	9
2.3 Objetivos secundarios	9
3 INTRODUCCIÓN	11
3.1 Marco histórico	13
3.2 Evolución y cronología de las fórmulas de cálculo	17
3.3 Siglo XXI y nuevos enfoques para el cálculo biométrico.....	35
3.3.1 Inteligencia Artificial y Machine Learning.	36
3.3.2 Inteligencia Artificial en oftalmología	38
3.3.3 Aplicación de Inteligencia Artificial en cálculos biométricos.....	40
3.4 Lenguaje R, nomenclatura básica y abordaje general de problemas de ML....	42
3.4.1 Objetivos y selección del tipo de problema	46
3.4.2 Análisis exploratorio de los datos	47
3.4.3 División de los datos	48
3.4.4 Preprocesado de los datos.....	49
3.4.5 Modelización en crudo	50
3.4.6 Optimización de hiperparámetros	51
3.4.7 Métricas de precisión en modelos de regresión	53
3.4.8 Interpretabilidad de modelos complejos	57
4 ESTADO DEL ARTE.....	61
5 DISEÑO DEL ESTUDIO	77
5.1 Diseño del estudio y flujo de trabajo.....	79
5.2 Diseño del estudio retrospectivo de la Etapa I.....	83
5.2.1 Criterios de inclusión	85

5.2.2	Criterios de exclusión.....	86
5.3	Desarrollo computacional	87
5.4	Análisis estadístico	89
5.5	Cálculo del tamaño muestral.....	90
5.6	Diversidad muestral	93
6	<i>METODOLOGÍA</i>	97
6.1	Etapa I: Recogida, análisis y preparación de los datos.	99
6.1.1	Recogida de la base de datos.....	99
6.1.2	Datos para la predicción de modelos.....	100
6.1.3	Datos para la comparación entre fórmulas.....	110
6.1.4	Correlación de variables y selección de características	113
6.1.5	Transformación de datos	114
6.2	Etapa II: Equivalencias de potencia entre planos	115
6.2.1	Residuo refractivo postoperatorio en plano corneal	116
6.2.2	Obtención del equivalente esférico en plano capsular	123
6.3	Etapa III: Implementación, selección y optimización de modelos ML.....	132
6.4	Etapa IV: Conversión del equivalente esférico capsular a plano corneal	133
7	<i>RESULTADOS</i>	135
7.1	Resultados de la etapa I	137
7.1.1	Exploración descriptiva analítica y gráfica del dataset.....	137
7.1.2	Correlación entre variables y selección de características	140
7.1.3	Transformación de datos	147
7.2	Resultados de la etapa II	148
7.2.1	Obtención de la ELP	148
7.3	Resultados de la etapa III	151
7.3.1	Modelos en bruto de Redes Neuronales	151
7.3.2	Modelos en bruto Bayesianos.....	154
7.3.3	Modelos en bruto con Máquinas de Vector de Soporte.....	156
7.3.4	Modelos en bruto con Splines adaptativos.....	157
7.3.5	Modelos basados en árboles de decisión	159
7.3.6	Selección de los mejores modelos en bruto	162
7.3.7	Hiperparametrización: optimización de los mejores modelos en bruto.	166
7.3.8	Ensamblaje de los mejores modelos optimizados	170

7.3.9	Interpretabilidad de modelos	172
7.4	Resultados de la etapa V	183
7.4.1	Validación del modelo resultante sobre la partición test	183
7.5	Resultados de la etapa VI	186
7.5.1	Comparación del modelo con otras fórmulas y métodos de cálculo.....	186
7.6	Validación del modelo resultante sobre ojos cortos	197
7.7	Validación del modelo resultante sobre ojos atípicos.....	206
7.8	Comparación entre fórmulas teóricas y basadas en IA	211
7.9	Presentación de resultados mediante página web	212
8	<i>DISCUSIÓN</i>	229
9	<i>CONCLUSIONES</i>	241
10	<i>BIBLIOGRAFÍA</i>	245
11	<i>ANEXOS</i>	265
11.1	ANEXO 1.- BASES TEÓRICAS DE LOS ALGORITMOS DE MACHINE LEARNING EMPLEADOS EN LA TESIS.....	267
11.1.1	Funcionamiento de redes neuronales	267
11.1.2	Funcionamiento de redes bayesianas.....	274
11.1.3	Funcionamiento de máquinas de vector de soporte	277
11.1.4	Funcionamiento de regresión multivariante adaptativa con splines	282
11.2	ANEXO 2.- ANÁLISIS CRÍTICO: MÉTRICAS PARA EVALUAR LA PRECISIÓN Y EXACTITUD DE FÓRMULAS.....	287
11.2.1	Actualidad y discrepancias entre trabajos	287
11.2.2	Error predictivo de refracción	291
11.2.3	Error absoluto medio	297
11.2.4	Mediana de errores absolutos	304
11.2.5	Tratamiento matemático de los errores absolutos.....	306
11.2.6	Transformación de la distribución de errores absolutos	308
11.2.7	Estadística robusta y centralidad de errores absolutos	310
11.2.8	Estadística robusta y dispersión de errores absolutos.....	313
11.2.9	Aprobación de las métricas robustas	316
11.2.10	Aplicación de las nuevas métricas sobre un ejemplo real.....	321
11.2.11	Porcentaje de errores predictivos pertenecientes a un intervalo refractivo 324	
11.2.12	Nuevo sistema de clasificación por porcentajes de rangos	326

11.2.13	Nuevo índice IOF: Index Of Formula	332
11.2.14	Esquema de actuación para estudios comparativos entre fórmulas	337
11.3	ANEXO 3.-NORMALIZACIÓN: ANÁLISIS DE POBLACIÓN BIOMÉTRICA	341
11.3.1	Introducción.....	341
11.3.2	Análisis del problema clasificatorio.....	343
11.3.3	Tamaño muestral	346
11.3.4	Metodología de análisis	348
11.3.5	Clasificación de la población biométrica	351
11.3.6	Instauración de un índice de singularidad	387
11.3.7	Aplicación web para clasificar un ojo por infrecuencia.....	393
11.4	ANEXO 4.- APORTACIONES CIENTÍFICAS DERIVADAS DE LA TESIS DOCTORAL.....	397
11.5	ANEXO 5.- APROBACIÓN COMITÉ ÉTICO	399
11.6	ANEXO 6.- INFORME DE ORIGINALIDAD.....	401

LISTA DE SÍMBOLOS, ABREVIATURAS Y ACRÓNIMOS

LISTA DE SÍMBOLOS, ABREVIATURAS Y ACRÓNIMOS

ACD	Profundidad de la cámara anterior (<i>Anterior Chamber Depth</i>)
ANSI	American National Standard Institute
ASD	Profundidad del segmento anterior (<i>Anterior Segment Depth</i>)
ATA	Distancia ángulo a ángulo (<i>Angle to Angle</i>)
AXL	Longitud axial (<i>Axial Length</i>)
BART	Árboles aditivos de regresión bayesianos (<i>Bayesian Additive Regression Trees</i>)
CD	Distancia ciliar (<i>Ciliar Distance</i>)
CMC	Círculo de mínima confusión
CPU	Unidad central de procesamiento (<i>Central Processing Unit</i>)
CV	Validación cruzada (<i>Cross Validation</i>)
D	Dioptrías
DL	Aprendizaje profundo (<i>Deep Learning</i>)
DMAE	Degeneración macular asociada a la edad
EE	Equivalente esférico
ELP	Posición efectiva de la lente (<i>Effective Lens Position</i>)
EVO	Emmetropia Verifying Optical
H	Altura corneal
Hill-RBF	Hill- Radial Basis Function

LISTA DE SÍMBOLOS, ABREVIATURAS Y ACRÓNIMOS

IA	Inteligencia Artificial
ICE	Gráfico de Esperanza Condicional individual (<i>Individual Conditional Expectation</i>)
ICI	Imperial Chemical Industries
IDE	Entorno de Desarrollo (<i>Integrated Development Environment</i>)
JCR	Factor de impacto en publicaciones (<i>Journal Citation Reports</i>)
KM_B	Queratometría media de la córnea posterior (<i>Keratometry Mean_Back</i>)
KM_T	Queratometría media total corneal (<i>Keratometry Mean Total</i>)
KNN	K vecinos más cercanos (<i>k-Nearest Neighbors</i>)
LF	Lens factor
LIO	Lente intraocular
LSF	Ladas super-fórmula
LT	Espesor de cristalino (<i>Lens Thickness</i>)
MAE	Error medio absoluto (<i>Mean Absolute Error</i>)
MARS	Regresión multivariante adaptativa con splines (<i>Multivariate Adaptive Regression Splines</i>)
MedAE	Mediana del error absoluto (<i>Median Absolute Error</i>)
MCMC	Cadenas de Markov Monte Carlo (<i>Markov Chain Monte Carlo</i>)
ML	Aprendizaje automático (<i>Machine Learning</i>)

LISTA DE SÍMBOLOS, ABREVIATURAS Y ACRÓNIMOS

MNNE	Red neuronal multicapa (<i>Multilayer Neural Network Ensemble</i>)
MPE	Error medio de predicción (<i>Mean Predictive Error</i>)
n	Número de observaciones o tamaño muestral
P	Paquimetría
PDP	Gráficos de dependencia parcial (<i>Partial Dependence Plot</i>)
PEARL-DGS	Prediction Enhanced by ARTificial Intelligence and output Linearization- Debellemanière, Gatinel, Saad.
PMMA	Polimetacrilato de metilo
r	Coficiente de correlación de Pearson
R	Coficiente de determinación
R ²	Coficiente de determinación al cuadrado
R_BF	Razón entre curvatura posterior y anterior corneal (<i>Ratio Back/Front</i>)
RBF	Función de base radial (<i>Radial Basis Function</i>)
RC	Radio central corneal
RCP	Radio periférico corneal
RG	Radio globo ocular
RELU	Unidad rectificadora lineal (<i>Rectified Linear Unit</i>)
RF	Bosques aleatorios (<i>Random Forest</i>)

LISTA DE SÍMBOLOS, ABREVIATURAS Y ACRÓNIMOS

RMSE	Raíz del error cuadrático medio (<i>Root Mean Square Error</i>)
RN	Red neuronal
SD	Desviación estándar (<i>Standard Deviation</i>)
SECOIR	Sociedad Española de Cirugía Ocular Implanto Refractiva
SF	Surgeon Factor
SI	Singularity Index
Simk	Queratometría simulada (<i>Simulated keratometry</i>)
SOM	Mapas auto-organizados (<i>Self-Organizing Maps</i>)
SRK/T	Sanders Retzlaff Kraff / Theoretical
SS-OCT	Swept Source – Optical Coherence Tomography
SVM	Vector de soporte de máquinas (<i>Support Vector Machine</i>)
TD	Árboles de decisión (<i>Tree Decision</i>)
VD	Distancia al vértice (<i>Vertex Distance</i>)
WTW	Diámetro horizontal corneal o blanco-blanco (<i>White to White</i>)
ZO	Zona óptica

LISTA DE FIGURAS

LISTA DE FIGURAS

- Figura 1. Dr. Harold Ridley en 1950. (Imagen extraída de Apple, 1996) 13
- Figura 2. Miembros del Club de Implantes Intraoculares reunidos en 1966 en Londres. Fila de arriba de izquierda a derecha: John Pike (Rayner, RU), Robert Murto (EEUU), Michael Roper-Hall (Birmingham, RU), Svyatoslav Fyodorov (Moscú, USSR), Neil Dallas (Bristol, RU), C. A. (Sandy) Brown (Bristol, RU), Alexander Rubinstein (Birmingham, RU), Warren S. Reese (Wills Eye Hospital, Philadelphia, EEUU), Leonard Lurie (Londres, RU). Fila de debajo de izquierda a derecha: Jörn Boberg-Ans (Charlottenlund, Dinamarca), Cees Binkhorst (Terneuzen, Holanda), Peter Choyce (Moorfields, Londres, RU), Harold Ridley (Moorfields, Londres, RU), Benedetto Strampelli (Roma, Italia), Edward Epstein (Johannesburgo, Sudáfrica), Sonja Boberg-Ans (Charlottenlund, Dinamarca). (Imagen extraída de Apple, 1996)..... 15
- Figura 3. Línea de tiempo y años de aparición con los distintos abordajes para el cálculo de la potencia de una lente intraocular y los pioneros en cada etapa. 17
- Figura 4. Esquema de funcionamiento de todos los abordajes que existen para calcular la potencia de una lente intraocular en una cirugía de cataratas. 18
- Figura 5. Vergencias sobre la lente intraocular y sobre la retina. 19
- Figura 6. Fórmula de Barrett universal año 1987: fórmula gaussiana de vergencia teniendo en cuenta el grosor del implante para el cálculo de la ACD_{postop} . (Imagen extraída del artículo de Barrett (46))..... 27
- Figura 7. Variación del posicionamiento de la lente implantada en función de la longitud axial para 4 fórmulas. (Imagen extraída del trabajo de Retzlaff J. et al (49)). . 31
- Figura 8. Artículos publicados relacionados con IA en oftalmología desde el año 1971. 39
- Figura 9. Interfaz de usuario de RStudio® versión 1.1.423 con sus cuatro espacios de trabajo por defecto. El panel superior izquierda es el editor. El panel superior derecha es el entorno de variables. El panel inferior izquierda es la consola de comandos. Panel inferior derecha es el panel de utilidades..... 44
- Figura 10. Histogramas de distribución normal generados con el código exportado de RStudio®..... 45

<i>Figura 11. Datos observados y predichos en un modelo de regresión.</i>	<i>54</i>
<i>Figura 12. Pantalla principal de FullmonteIOL 2.0. (Imagen extraída de https://www.youtube.com/watch?v=p6f8XKJXBY&list=PLm85NWhdte2j1eSzDOLnoVpwW0A2kT3Rw&index=12)..</i>	<i>63</i>
<i>Figura 13. Sistema gráfico de resultados en FullMonte IOL mediante curvas de densidad de probabilidad. (Imagen extraída de https://www.youtube.com/watch?v=T9UwecAw75o).....</i>	<i>64</i>
<i>Figura 14. Calculador Hill-RBF v3.0.</i>	<i>65</i>
<i>Figura 15. Diapositiva del Webinar impartido por W.Hill (07-01-2021) titulado “IOL Power Selection Driven by Artificial Intelligence”. (Imagen extraída de https://www.youtube.com/watch?v=EoZRnSHIZs4).....</i>	<i>66</i>
<i>Figura 16. A) Distribución de datos de entrada. B) En el primer momento las funciones de vecindario no se han aplicado (Time indica el número de iteraciones). C) Tras el entrenamiento inicial de la red el patrón de va dibujando. D) Concluido el entrenamiento o fase de convergencia el mapa final habría reconocido el patrón de entrada. (Imagen extraída del trabajo de Haykin (84)).....</i>	<i>67</i>
<i>Figura 17. Algoritmo genético usado en el diseño de Hill-RBF para selección de características. (Imagen extraída de una conferencia de W. Hill en la ASCRS (82))</i>	<i>68</i>
<i>Figura 18. Esquema de funcionamiento de la red neuronal en la que está basada Hill-RBF. (Imagen extraída de una conferencia de W. Hill en la ASCRS (82)).....</i>	<i>69</i>
<i>Figura 19. Pantalla de cálculo de la fórmula de Ladas.....</i>	<i>70</i>
<i>Figura 20. A) Superposición de “super-superficies” correspondientes a las distintas fórmulas que integran LSF para conformar un único plano. B) Cálculo realizado con Ladas Super Formula y la inclusión de la super-superficie que representa isométricamente la potencia de la lente, la longitud axial y la potencia corneal. La estrella marca la potencia de la lente que hay que implantar y la zona de influencia en el gráfico de la fórmula que subyace en el cálculo.....</i>	<i>71</i>
<i>Figura 21. Esquema simple de la Red Neuronal en la que se basa parte de la fórmula de Ladas. (Imagen extraída de (86))</i>	<i>72</i>
<i>Figura 22. Pantalla de cálculo de la fórmula de Kane.</i>	<i>73</i>
<i>Figura 23. Pantalla principal de la fórmula PEARL-DGS v2.0.....</i>	<i>74</i>

Figura 24. <i>Boxplot de errores predictivos refractivos comparando PEARL-DGS con otras fórmulas. (Imagen extraída del blog del Dr. D. Gatinel: https://www.gatinel.com/recherche-formation/biometrie-oculaire-calcul-dimplant/pearl-dgs-formula-for-iol-power-calculation/)......</i>	74
Figura 25. <i>Pantalla principal del calculador Hoffer-QST.</i>	75
Figura 26. <i>Diagrama de flujo para la tesis.</i>	80
Figura 27. <i>(A) Paralelización: rendimiento de los 8 cores durante el entrenamiento de una Red Neuronal multicapa, el tiempo de computación fue de 19 segundos. (B) Proceso secuencial: sin la activación del procesamiento en paralelo se ve como los núcleos van demandando de forma alternativa, el tiempo de computación en del entrenamiento de la misma Red Neuronal fue de 120 segundos (6 veces más).</i>	89
Figura 28. <i>Diagrama de matriz de correlaciones entre variables predictoras y respuesta.</i>	91
Figura 29. <i>Tamaño muestral en función del tamaño del efecto cuando se comparan dos coeficientes de Pearson.....</i>	92
Figura 30. <i>Tamaño muestral en función del tamaño del efecto en una comparación de coeficientes de correlación de Pearson de dos predictores con la variable respuesta.</i>	92
Figura 31. <i>Poder estadístico en base el tamaño muestral para la partición de validación.</i>	93
Figura 32. <i>Anatomía intra y extraocular con sus variables biométricas.....</i>	101
Figura 33. <i>Conoide de Sturm y distribución de focales astigmáticas para un astigmatismo directo. Localización del círculo de mínima confusión en relación con la queratometría media.....</i>	108
Figura 34. <i>Posicionamiento respecto a la retina del círculo de mínima confusión (CMC) en función de la refracción esferocilíndrica.....</i>	109
Figura 35. <i>Esquema de posibles abordajes en la predicción con ML para una calculadora de potencia emetropizante.</i>	117
Figura 36. <i>Presentación de resultados de diferentes fórmulas (Kane, Barrett Universal II y EVO v2.0) mostrando el gradiente de potencias con precisión de 0.50 D y sus residuos refractivos asociados en plano corneal.....</i>	121

Figura 37. Planos de equivalencia dióptrica entre lente oftálmica, córnea y lente intraocular con sus respectivas distancias: distancia al vértice o Vertex Distance (VD) y posición efectiva de la lente (ELP).	124
Figura 38. Esquema de equivalencia dióptrica entre lente oftálmica y ápex corneal considerando ambos sistemas ópticos como delgados.....	125
Figura 39. Esquema de equivalencia dióptrica entre córnea y lente intraocular considerando ambos sistemas ópticos como delgados y separados por la ELP y sus respectivos índices de refracción: humor acuoso (1.336) y humor vítreo (1.336).....	127
Figura 40. Diagrama de cajas y bigotes que muestra la distribución de todas las variables biométricas exploradas discriminando por tamaños oculares Short, Normal y Long respectivamente.....	139
Figura 41. Histogramas y qqnorm de las variables relevantes de la base de datos.	140
Figura 42. Gráficos de normalidad (Q-Q plot) de todas las variables de modelaje del dataset.....	141
Figura 43. Heatmaps-cluster con las matrices de correlación según los diferentes coeficientes estudiados de todas las variables regresoras y de la variable respuesta (IOL_P).....	142
Figura 44. Heatmap-cluster con matriz de correlación en base al coeficiente r de Pearson de las variables queratométricas.	144
Figura 45. Heatmaps-cluster con matrices de correlación para coeficientes de Pearson, Spearman y Kendall en las variables no queratométricas.....	146
Figura 46. Base de datos en bruto y transformada mediante estandarización.....	147
Figura 47. Hoja de cálculo Excel con los cálculos de ajuste de ELP para la obtención de los mismos resultados refractivos en plano corneal que Barrett Universal II.	149
Figura 48. Pantalla de resultados de Karmona® (fórmula obtenida en esta tesis) con Barrett Universal II y con Hill- RBF v3.0.....	150
Figura 49. Rendimiento de la CPU en pleno proceso de aprendizaje de una RN, con tasas de ocupación de hasta el 100%.....	152
Figura 50. Esquema de red neuronal para modelo 4 con 3 neuronas en capa oculta y una salida para la potencia de la lente intraocular. Funciones de activación SoftReLU.	153

Figura 51. Gráficos resultantes del proceso de regularización mediante cross-validation en el modelo GCV-SPLINE implementado sobre el modelo 1 con predictores corneales KM_F y R_B/F.....	158
Figura 52. Árbol de decisión (CART) para el dataset original de la tesis.....	160
Figura 53. Predicción con XGBOOST sobre la muestra test (20%) poniendo de manifiesto un overfitting.	166
Figura 54. Optimización de parámetros para los cuatro modelos finales seleccionados. A) modelo CUB-TREE: committees = 10. B) modelo BAG-SPLINE: Prune = 6. C) modelo SVM-POLY: Scale = 0.1, degree = 2, C = 0.5. D) modelo BRNN: neurons = 2.....	168
Figura 55. Esquema de modelaje en todos sus niveles de trabajo. El último nivel corresponde al modelo ensamblado que se llamó Karmona®, una combinación de BRNN y CUBIST, que no comete sobreajuste obteniendo el mismo MAE con la muestra de validación (20%).....	171
Figura 56. Importancia de las variables en el modelo de red neuronal bayesiana (BRANN).	173
Figura 57. Importancia de variables en el modelo Cubist (CUB-TREE).	174
Figura 58. Resumen de la optimización para Cubist. Se han empleado 386 ojos para el modelaje, 6 variables predictoras (AXL, ASD, WTW, CTE, SIMK, R_B/F), se han establecido 5 submodelos con 17 reglas por cada uno de ellos (en total 85 reglas).	174
Figura 59. Las 6 reglas establecidas para el submodelo1 con sus regresiones lineales multivariante correspondientes.	175
Figura 60. Gráficos PDP para la variable AXL en los dos modelos finales antes del ensamblaje.....	176
Figura 61. Gráfico PDP para el predictor AXL conjunto de los modelos seleccionados como finales antes del ensamblaje.....	177
Figura 62. Gráficos PDP para la variable ASD en los dos modelos finales antes del ensamblaje.....	177
Figura 63. Heatmap PDP para AXL y KM_F del modelo Cubist.....	178
Figura 64. Heatmap PDP para AXL y R_B/F del modelo Cubist.	179
Figura 65. Heatmap PDP para AXL y ASD del modelo bayesiano.	180

Figura 66. Arriba: gráfico ICE con predictor AXL en modelo bayesiano. Abajo: d-ICE para predictor AXL en el modelo bayesiano.	181
Figura 67. Interacciones entre AXL y el resto de predictores en el modelo bayesiano. .	182
Figura 68. Interacciones entre AXL y el resto de predictores en el modelo Cubist.....	183
Figura 69. Boxplot de todas las fórmulas sin ajuste sobre MPE para el total de la muestra (n = 95).....	190
Figura 70. Boxplot de todas las fórmulas con ajuste 0.00 D sobre MPE para el total de la muestra (n = 95).....	190
Figura 71. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra (n = 95) sin ajustar sobre MPE.....	195
Figura 72. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra (n = 95) con ajuste sobre MPE.....	196
Figura 73. Esquema de flujo de trabajo con la inclusión de las nuevas muestras accesorias (amarillo) de ojos axialmente cortos y anatómicamente atípicos.....	198
Figura 74. A) Boxplot de las 6 fórmulas para la muestra completa (n = 33) sin ajuste sobre MPE. B) Boxplot de las 6 fórmulas para la muestra completa (n = 33) con ajuste sobre MPE.....	202
Figura 75. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra (n = 33) sin aplicar ajuste sobre MPE.....	205
Figura 76. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra (n = 33) con ajuste sobre MPE.....	206
Figura 77. Boxplot para el Mean Absolute Error (D) de todas las fórmulas (n = 41).	209
Figura 78. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra (n = 41).	211
Figura 79. Pantalla principal de la web www.karmona-iol.com	213
Figura 80. Pantalla de resultados de Karmona® en la que se muestra la potencia de la lente emetropizante y 5 potencias redondeadas con precisión de 0.50 D, así como las predicciones refractivas asociados a dichas potencias dióptricas...	216

<i>Figura 81. Mensaje de descargo de responsabilidad en el uso de Karmona® que aparece nada más acceder a la web.....</i>	218
<i>Figura 82. Resultados de Singularity en ACD, LT y WTW (primera parte).</i>	220
<i>Figura 83. Resultados de Singularity en AXL y KM con el cálculo de SI (segunda parte).</i>	221
<i>Figura 84. Pantalla de “Researchers” en la web karmona-iol.com</i>	223
<i>Figura 85. Pantalla “About” de la web karmona-iol.com.....</i>	225
<i>Figura 86. Pantalla “Contact” de la web karmona-iol.com.....</i>	227
<i>Figura 87. Esquema simple de una red neuronal con tres capas.</i>	267
<i>Figura 88. Función sigmoide de activación en una red neuronal.</i>	269
<i>Figura 89. Función ReLu de activación en una red neuronal.....</i>	270
<i>Figura 90. Función SoftReLu de activación en una red neuronal.....</i>	271
<i>Figura 91. Esquema del algoritmo Perceptrón. (Imagen extraída de https://es.wikipedia.org/wiki/Perceptr%C3%B3n).....</i>	273
<i>Figura 92. Esquema del teorema de Bayes para modelaje probabilístico.....</i>	276
<i>Figura 93. Hiperplano de separación en un espacio bidimensional.....</i>	279
<i>Figura 94. Infinitos hiperplanos de separación = infinitas soluciones al problema clasificadorio. El hiperplano óptimo está representado por la recta continua y las rectas discontinua representan los hiperplanos posibles.....</i>	280
<i>Figura 95. Hiperplano óptimo con sus márgenes y vectores soporte (3 datos equidistantes).</i>	281
<i>Figura 96. Mismo conjunto de datos analizado en diferentes dimensiones (R^2 y R^3). Cuando las observaciones pertenecen a un espacio bidimensional no son linealmente separables (diagrama de la izquierda). Cuando los mismos datos pertenecen a un espacio tridimensional con un hiperplano sí se pueden separar (diagrama de la derecha).</i>	282
<i>Figura 97. (A) Regresión lineal normal. (B) Regresión polinómica de segundo grado en la que se capta algo de no linealidad en el dataset. (C) Regresión polinómica de tercer orden en el que se capturó todas las relaciones. (D) Step Function aplicada al mismo dataset con 5 subintervalos. (Imagen extraída de https://bradleyboehmke.github.io/HOML/mars.html)</i>	284

<i>Figura 98. Tres fases de optimización que utiliza MARS para suavizar el spline. (A) Primera fase de step function dónde se imponen los puntos de corte (knots) de la función escalonada. (B) Segunda fase en la que se obliga al segmento polinómico a que los puntos de corte coincidan. (C) Tercera fase o smoothing en la que se suaviza el knot mediante las primeras y segundas derivadas de las funciones contiguas.</i>	285
<i>Figura 99. Diferencias gráficas entre precisión y exactitud tanto para el promedio de errores absolutos como para el promedio de errores predictivos en las fórmulas y métodos de cálculo.....</i>	290
<i>Figura 100. Análisis matemático y gráfico sobre una muestra de 114 observaciones en el que se está contrastando la normalidad o no de la distribución de errores predictivos para dos fórmulas de cálculo.....</i>	293
<i>Figura 101. Centrado de una distribución para que la media de errores predictivos sea 0.00 D.....</i>	296
<i>Figura 102. Demostración gráfica y analítica de que los errores absolutos se distribuyen según una distribución exponencial gamma alejada de la normal gaussiana.</i>	298
<i>Figura 103. Boxplots para los MAE de 11 fórmulas y la explicación detallada en el margen derecho de interpretación de un gráfico de cajas y bigotes.....</i>	300
<i>Figura 104. Métricas clásicas (MAE, SD y MedAE) para clasificar precisión y exactitud de las fórmulas y métodos de cálculo biométrico.....</i>	301
<i>Figura 105. Ejemplo con 3 fórmulas y solo 10 observaciones en valor absoluto para calcular los MAE, SD y MedAE. Ordenación de las fórmulas en base a estas métricas.....</i>	302
<i>Figura 106. Cálculo de la mediana de los errores absolutos como métrica robusta en la clasificación de fórmulas de cálculo biométrico.....</i>	304
<i>Figura 107. Normalización mediante raíz cuadrada de la distribución de errores absolutos.</i>	309

<i>Figura 108. Métricas robustas sobre la población normalizada de errores absolutos en el cálculo real de 11 fórmulas. Métricas paramétricas clásicas MAE y MedAE con sus respectivas anti-transformaciones. Métrica robusta TMAE5 correspondiente a la media acotada al 5% y su anti-transformación. En rojo están calculados cuatro M-estimadores sobre la distribución sin transformar para comparar con los datos normalizados (Huber, Onestep, Pitman y Tukey bponderado).....</i>	<i>312</i>
<i>Figura 109. Errores absolutos de tres fórmulas sin transformar con sus métricas paramétricas clásicas (MAE, SD y MedAE). A la derecha los mismos errores absolutos transformados mediante raíz cuadrada con las nuevas métricas (Anti-MAE, SD, Anti-MedAE y Anti-TMAE20).</i>	<i>318</i>
<i>Figura 110. Contraste de errores absolutos dos a dos.</i>	<i>320</i>
<i>Figura 111. Estadística clásica contra estadística propuesta en la comparativa real de 12 fórmulas de cálculo biométrico para 482 ojos con longitudes axiales entre 19.33 mm y 33.00 mm.</i>	<i>323</i>
<i>Figura 112. Representación gráfica y analítica de rangos dióptricos -0.25/0.00 y 0.01/+0.25 estudiados en la comparación de diferentes fórmulas de cálculo biométrico.....</i>	<i>330</i>
<i>Figura 113. Métricas clásicas aplicadas a 11 fórmulas en una muestra con AXL < 25.00 mm.....</i>	<i>335</i>
<i>Figura 114. Comparativa de 11 fórmulas mediante las métricas propuestas en esta tesis. Se muestran todos los pasos y medidas necesarias para obtener IOF.</i>	<i>336</i>
<i>Figura 115. Histograma de distribución de la AXL dentro de un rango de ACD determinado (2.91 a 3.16 mm).....</i>	<i>343</i>
<i>Figura 116. Biplot representando en dos ejes la correlación entre todas las variables biométricas.</i>	<i>345</i>
<i>Figura 117. Matriz de correlaciones de todas las variables biométricas estudiadas. El diagrama contiene los scatter-plot de relación lineal, los histogramas y los factores de correlación de Pearson.</i>	<i>346</i>

- Figura 118. El gráfico superior representa el histograma de frecuencias con las curvas de densidad poblacional y la de normalidad. El gráfico del medio es un boxplot con la media, la mediana y su intervalo de confianza, así como la nube de puntos. El último gráfico es un QQplot que representa la distribución de los datos comparándolos con una distribución normal. (La figura analiza la distribución de ACD en hombres con ojos cortos)..... 349*
- Figura 119. Diagrama polinómico para representar las frecuencias de una muestra. En este gráfico se ve la distribución de frecuencias del grosor cristalino en ojos considerados axialmente normales en mujeres. 350*
- Figura 120. Tabla de distribución de frecuencias que acompaña a la representación polinómica de la figura anterior (Figura 119). La primera columna (Class limits) se corresponde a los intervalos generados, la segunda (f) es la frecuencia absoluta o número de casos en cada intervalo de clase. La tercera (rf) es la frecuencia relativa. La cuarta columna (fr%) es la frecuencia relativa porcentual. La quinta (cf) es la frecuencia acumulada y la sexta (cf%) es la frecuencia acumulada en porcentaje. Esta tabla de frecuencias representa el espesor de cristalino (LT) en ojos de mujeres axialmente normales. 350*
- Figura 121. Histogramas de densidad para AXL distinguiendo por género de la muestra (n = 10.295)..... 351*
- Figura 122. Distribución de frecuencias en hombres y mujeres para AXL. 353*
- Figura 123. QQ- plot para longitud axial dónde los cuantiles de la distribución de los datos comparada con una normal se alejan de esta en los extremos, sobre todo en longitudes largas por encima de 25.00 mm. 354*
- Figura 124. Histogramas, plots polinómicos de frecuencia y tablas de frecuencia de ACD en ojos cortos para hombres y mujeres. 357*
- Figura 125. Histograma de ACD en ojos cortos para mujeres (inferior) en comparación con una distribución teórica normal (superior). Se puede observar un ejemplo de ACD de 3.00 mm recaería en una zona correspondiente al 2.1% de la población, es decir, muy infrecuente. 359*
- Figura 126. Histogramas, plots polinómicos y tablas de frecuencia de ACD en ojos normales para hombres y mujeres..... 360*

<i>Figura 127. Histogramas, plots polinómicos y tablas de frecuencia de ACD en ojos largos para hombres y mujeres.</i>	<i>361</i>
<i>Figura 128. Histogramas, plots polinómicos y tablas de frecuencia de ACD en ojos extralargos para hombres y mujeres. En hombres hay un sesgo claro en ACD por debajo de 3.50 mm.</i>	<i>362</i>
<i>Figura 129. Representación polinómica de frecuencias absolutas y tabla de distribución de frecuencias del espesor de cristalino en ojos extra cortos para ambos sexos.</i>	<i>364</i>
<i>Figura 130. Histogramas de densidad y diagramas polinómicos de frecuencias del espesor cristalino en ojos axialmente cortos para ambos sexos.</i>	<i>365</i>
<i>Figura 131. Tabla de frecuencias de espesor de cristalino en ojos cortos para hombres y mujeres.</i>	<i>365</i>
<i>Figura 132. Histogramas y curvas polinómicas de frecuencias para grosor de cristalino en ojos normales para ambos sexos.</i>	<i>367</i>
<i>Figura 133. QQplot y tablas de distribución de frecuencias de espesores cristalinos en ojos normales para ambos sexos.</i>	<i>368</i>
<i>Figura 134. Histogramas y QQplot de LT en ojos axialmente largos para los dos sexos.</i>	<i>369</i>
<i>Figura 135. Plots polinómicos y tablas de frecuencia de LT en ojos largos para los dos sexos.</i>	<i>371</i>
<i>Figura 136. Histogramas, diagramas polinómicos y tablas de frecuencia para ojos extralargos en ambos sexos.</i>	<i>372</i>
<i>Figura 137. Histogramas y QQplots de normalidad para WTW en ojos extra cortos en ambos sexos.</i>	<i>373</i>
<i>Figura 138. Diagramas de frecuencia polinómico y las tablas de distribución de frecuencias para WTW en ojos extra cortos en ambos sexos.</i>	<i>374</i>
<i>Figura 139. Histogramas y QQplots para ojos cortos en ambos sexos.</i>	<i>375</i>
<i>Figura 140. Diagramas de frecuencia polinómico y las tablas de distribución de frecuencias para WTW en ojos cortos en ambos sexos.</i>	<i>375</i>
<i>Figura 141. Histogramas de densidad y QQplots de normalidad para WTW en ojos normales en ambos sexos.</i>	<i>376</i>
<i>Figura 142. Diagramas polinómicos y tablas de distribución de frecuencias de WTW en ojos normales en ambos sexos.</i>	<i>377</i>

<i>Figura 143. Histogramas, curvas polinómicas y tablas de frecuencia de WTW en ojos largos para hombres y mujeres.....</i>	<i>379</i>
<i>Figura 144. Histogramas, curvas polinómicas y tablas de frecuencia de WTW en ojos extralargos para hombres y mujeres.</i>	<i>380</i>
<i>Figura 145. Histogramas, diagrama polinómico y tablas de frecuencias de KM en ojos extra cortos para ambos sexos.</i>	<i>381</i>
<i>Figura 146. Histogramas, diagrama polinómico y tablas de frecuencias de KM en ojos cortos para ambos sexos.....</i>	<i>382</i>
<i>Figura 147. Histogramas, diagrama polinómico y tablas de frecuencias de KM en ojos normales para ambos sexos.....</i>	<i>383</i>
<i>Figura 148. Histogramas, diagramas polinómicos, QQplots y tablas de frecuencias de KM en ojos largos para ambos sexos.</i>	<i>384</i>
<i>Figura 149. Histogramas, diagramas polinómicos, QQplots y tablas de frecuencias de KM en ojos extra cortos para ambos sexos.</i>	<i>386</i>
<i>Figura 150. Representación triple (histograma + boxplot + qqnorm) de la ubicación de una AXL de 20.35 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de axialidad.....</i>	<i>388</i>
<i>Figura 151. Representación triple (histograma + boxplot + qqnorm) de la ubicación de una ACD de 2.15 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de profundidad de cámara anterior.</i>	<i>389</i>
<i>Figura 152. Representación triple (histograma + boxplot + qqnorm) de la ubicación de un LT de 3.66 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría espesor de cristalino.....</i>	<i>390</i>
<i>Figura 153. Representación triple (histograma + boxplot + qqnorm) de la ubicación de un WTW de 11.40 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de diámetro horizontal corneal.....</i>	<i>391</i>

- Figura 154. Representación triple (histograma + boxplot + qqnorm) de la ubicación de un KM de 47.41 D sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de queratometría media con índice queratométrico 1.3375. 392*
- Figura 155. Desglose del índice de singularidad (SI) por categorías anatómicas y el sumatorio de todas representando el grado de rareza anatómica del ojo estudiado. 393*
- Figura 156. Captura de pantalla de la aplicación web diseñada para evaluar el índice de singularidad de un ojo (1ª parte). 394*
- Figura 157. Captura de pantalla de la aplicación web diseñada para evaluar el índice de singularidad de un ojo (2ª parte). 395*

LISTA DE TABLAS

LISTA DE TABLAS

<i>Tabla 1. Constantes de las lentes utilizadas en la base de datos extraídas de iolcon.org.</i>	85
<i>Tabla 2. Conjunto de variables incluidas en la base de datos de la tesis y su diferenciación en función del uso al que están destinadas.</i>	99
<i>Tabla 3. Obtención del error predictivo (EP) cometido por las fórmulas mediante del equivalente esférico postoperatorio.</i>	111
<i>Tabla 4. División en dos grupos de todas las fórmulas incluidas en la tesis para la comparación futura entre fórmulas teóricas y basadas en IA.</i>	112
<i>Tabla 5. Variables biométricas predictoras incluidas en los distintos modelos.</i>	132
<i>Tabla 6. Estadística descriptiva de todas las variables biométricas empleadas en el estudio de la muestra total y estratificadas por longitud axial.</i>	138
<i>Tabla 7. Métricas de rendimiento para el modelo 4 implementado en tres redes neuronales: CARET+MxNet (4.72 seg), CARET+Monmlp (Perceptrón) 2 capas (> 6min) y CARET+Monmlp (Perceptrón) 1 capa (1.26 min).</i>	154
<i>Tabla 8. Resultados de rendimiento de modelos bayesianos.</i>	156
<i>Tabla 9. Resultados de los modelos SVM con diferentes kernel (lineal, radial y polinómico).</i>	157
<i>Tabla 10. Resultados para los modelos de regresión adaptativa con splines.</i>	159
<i>Tabla 11. Resultado del modelo de árbol de decisión Cubist (CUB-TREE).</i>	162
<i>Tabla 12. Resultados de todos los modelos en bruto implementados con diferentes técnicas mostrando todas las métricas y el tiempo de computación consumido. Los MAE obtenidos por los modelos en bruto son en el plano óptico de la lente intraocular.</i>	163
<i>Tabla 13. Métricas de los modelos optimizados que rinden por debajo de 0.38 D en plano capsular.</i>	169
<i>Tabla 14. Métricas de rendimiento del modelo Karmona® en la fase de entrenamiento con el 80% del dataset original (columna izquierda) y en la fase de validación con el 20% de los datos (columna central). En la columna de la derecha el MAE se convierte por distometría a plano corneal.</i>	184

<i>Tabla 15. Métricas clásicas de evaluación de precisión y exactitud de una fórmula o método de cálculo biométrico. La fila SIN AJUSTE muestra las métricas sin la corrección de centrado sobre el MPE. La fila CON AJUSTE corresponde a las mismas métricas con la corrección sobre MPE.</i>	<i>185</i>
<i>Tabla 16. Métricas propuestas en esta tesis en relación con Karmona®. IOF (Index Of Formula), ANTI-MAE es el promedio del error absoluto de la distribución transformada. ANTI-TMAE10 es el promedio del error absoluto de la distribución transformada, pero con un acotamiento bilateral del 10%. ANTI-MedAE es la mediana del error absoluto de la distribución transformada y ANTI-MAD es la distancia absoluta entre la mediana y el error de predicción absoluto de la distribución transformada.</i>	<i>185</i>
<i>Tabla 17. Variables biométricas del total de la muestra de validación (n = 95).</i>	<i>187</i>
<i>Tabla 18. Resultados con métricas clásicas para las 13 fórmulas sin ajustar MPE en toda la muestra (n = 95 ojos).</i>	<i>188</i>
<i>Tabla 19. Resultados con métricas clásicas para las 13 fórmulas con ajuste sobre MPE en toda la muestra (n = 95 ojos).</i>	<i>188</i>
<i>Tabla 20. Rendimiento de las 13 fórmulas empleando las métricas descritas en esta tesis. IOF25 Index Of Formula en el rango ± 0.25 D. IOF50 en el rango de ± 0.50 D. MEAN IOF es el promedio de IOF25 e IOF50. ANTI-MAE es el error medio absoluto de la distribución normalizada. ANTI-TMAE10 es el error medio absoluto de la distribución normalizada pero acotada al 10% por ambas colas. ANTI-MedAE es la mediana de los errores absolutos de la distribución normalizada. ANTI-MAD es la distancia promedio entre el error absoluto y la mediana de la distribución normalizada.</i>	<i>191</i>
<i>Tabla 21. Rendimiento de las 13 fórmulas en ojos cortos ($AXL \leq 22.00$ mm) sin y con ajuste de centrado sobre MPE (n = 12 ojos) de la partición test.</i>	<i>192</i>
<i>Tabla 22. Rendimiento de las 13 fórmulas en ojos cortos ($AXL \geq 25.00$ mm) sin y con ajuste de centrado sobre MPE (n = 25 ojos) de la partición test.</i>	<i>193</i>
<i>Tabla 23. Rendimiento de las 13 fórmulas en ojos cortos ($AXL > 22.00$ y $AXL < 25.00$ mm) sin y con ajuste de centrado sobre MPE (n = 25 ojos) en la partición test.</i>	<i>194</i>
<i>Tabla 24. Variables biométricas de la muestra de ojos cortos (n = 33)</i>	<i>200</i>

<i>Tabla 25. Resultados con métricas clásicas para 6 fórmulas sin ajuste sobre MPE en toda la muestra (n = 33 ojos).</i>	200
<i>Tabla 26. Resultados con métricas clásicas para 6 fórmulas con ajuste sobre MPE en toda la muestra (n = 33 ojos).</i>	201
<i>Tabla 27. Rendimiento de las 6 fórmulas empleando las métricas descritas en esta tesis sin ajuste sobre MPE.</i>	203
<i>Tabla 28. Rendimiento de las 6 fórmulas empleando las métricas descritas en esta tesis con ajuste sobre MPE.</i>	204
<i>Tabla 29. Descriptiva de las variables biométricas de la muestra de ojos atípicos (n = 41).</i>	207
<i>Tabla 30. Resultados con métricas clásicas para 6 fórmulas (n = 41 ojos).</i>	208
<i>Tabla 31. Rendimiento de las 6 fórmulas empleando las métricas descritas en esta tesis.</i>	209

1 JUSTIFICACIÓN

1. JUSTIFICACIÓN

Actualmente la cirugía de cristalino se ha convertido en una sinergia multidisciplinar en la que se pueden diferenciar claramente tres fases. La primera es prequirúrgica e interviene, en la mayoría de las ocasiones, un optometrista para la toma de medidas y decisiones biométricas sobre el ojo del paciente. En esta fase se incluye la selección de la fórmula biométrica más adecuada en función de la tipología de ojo, así como la elección de la potencia de la lente intraocular que se tendrá que implantar.

Una segunda etapa intraoperatoria en la que el médico oftalmólogo ejecuta la técnica quirúrgica incluso en ocasiones asistido por un optometrista cuando se trata de técnicas con facofemtosegundo.

Por último la fase postquirúrgica, en la que de nuevo ambos especialistas siguen al paciente para explorar de forma paralela variables puramente médicas y otras solo ópticas. Es en esta última fase dónde se analizará la refracción residual y la agudeza visual asociada, además de verificar si la fórmula empleada resolvió correctamente.

El procedimiento clínico necesario para obtener la potencia dióptrica de una lente intraocular que será implantada en el ojo de un paciente se describe a continuación. En primera instancia, se miden en el ojo del paciente, mediante técnicas biométricas, una serie de variables biológicas importantes que caracterizan matemáticamente la anatomía ocular. Dichas variables corresponden a longitudes y curvaturas, que se introducen posteriormente en unas fórmulas de cálculo produciendo un valor de potencia para la lente intraocular, además de la predicción de un valor de graduación residual

postoperatorio. A priori, el valor dióptrico de la lente que lleve asociado el menor valor refractivo residual es el que se implanta.

La forma en la que estas fórmulas biométricas son capaces de predecir la potencia y la graduación residual es muy diferente, de hecho, no todas operan con las mismas variables biométricas. Son muchos los estudios (1-3) en la literatura oftalmológica que analizan y comparan la precisión y predictibilidad de unas con otras.

Desde el año 1967 en el que el oftalmólogo ruso Svyatoslav Fiodorov propuso la primera fórmula de cálculo biométrico hasta la actualidad, se han pasado por cinco décadas de evoluciones y mejoras en estos algoritmos matemáticos. No obstante, aún hoy en día no hay una fórmula perfecta para todos los casos, ya que dependiendo de las características anatómicas del ojo se escogerá una u otra. Actualmente es aceptado por la comunidad científica (4) como buen resultado, aquel que se encuentra en el rango refractivo postoperatorio de ± 0.50 D.

Ninguna de las nueve fórmulas más utilizadas (5) en la actualidad ofrece la garantía de predecir perfectamente la potencia de la lente en todos los rangos anatómicos (6), ya que unas funcionan mejor que otras dependiendo de la queratometría, de la longitud axial o de la profundidad de cámara anterior que son las variables biométricas básicas con las que operan.

Siendo consciente que este emprendimiento no es original, cuando me planteé el trabajo inicialmente hace seis años, que posteriormente formó parte de mi trabajo fin de máster, solo existían dos métodos de cálculo que operaban de modo distinto a lo que había entonces: Hill-RBF (Dr. Warren Hill) y FullMonte (Dr. Gerald Clarke). El primero basado en un modelo no lineal neuronal con una función gaussiana Radial Basis Function

y que actualmente está siendo sometida a multitud de estudios comparativos con las fórmulas tradicionales (7). El segundo método basado en un modelo probabilístico de *Markov Chain Monte Carlo*, comparado en la literatura oftalmológica publicada mucho menos que el anterior y poco utilizada en la práctica clínica debido a su sistema de acceso privado mediante pago.

Clarke et al (8) en 1997 fue el pionero con su trabajo, en el que comparó la predicción de la potencia de la lente intraocular mediante una red neuronal con una fórmula absolutamente instaurada y fiable en la década de los 90, la Holladay I. Lamentablemente desde entonces hasta la aparición de Hill-RBF no ha habido este tipo de abordajes en el cálculo de lentes intraoculares.

Sin embargo desde 2015, cuando pensé en la manera de mejorar lo que teníamos, hasta hoy en día, han ido apareciendo sucesivamente nuevas fórmulas relacionados con la IA: Ladas Superformula (LSF), Kane, PEARL-DGS y la última y más reciente, Hoffer-QST. Ciertamente, todas estas fórmulas se fundamentan en teoría gaussiana de vergencias con el aporte de IA en ciertos aspectos predictivos. Si bien, la idea general para el desarrollo de esta investigación es crear un sistema basado íntegramente en IA. Por tanto, los retos abordados en esta tesis fueron los siguientes:

- Obtención de datos de la máxima calidad.
- Manipulación, transformación y aprovechamiento máximo de los datos.
- Reinterpretación de ciertos aspectos teóricos de equivalencias.
- Implementación de complejos modelos matemáticos inéditos en la literatura oftalmológica publicada utilizando técnicas avanzadas de ensamblaje en *Machine Learning*.

Y todo ello para intentar romper las tendencias actuales, introduciendo más del 90% de los pacientes en el rango refractivo postoperatorio de ± 0.50 D y mejorando el potencial predictivo de las fórmulas existentes. Con la pretensión de conseguir verdaderamente una fórmula universal, estable y robusta para todos los rangos de axialidad.

Algo más de medio siglo después de la aparición de la primera fórmula de cálculo biométrico, se ha pasado de incluir un 79% de los ojos (9) en el rango postoperatorio de ± 1.00 D a aproximadamente un 80% de los ojos en el rango de ± 0.50 D (10). Por tanto, en estos 50 años se ha mejorado únicamente ± 0.50 D los resultados refractivos postoperatorios en términos absolutos. Con este pequeño margen de mejora y siguiendo esta tendencia, en las próximas dos décadas ¿se incluirán el 80% de los pacientes en el rango ± 0.25 D?

Este es precisamente el nicho dónde tiene cabida este trabajo, esta es la motivación principal de este proyecto, explorar una nueva forma de predecir la potencia de la lente intraocular mediante técnicas actuales de inteligencia artificial (IA), para posteriormente confirmar si este abordaje permitirá mayor precisión y exactitud en la predicción de resultados de la que hasta ahora se ha conseguido en distintos tipos de ojos.

2 HIPÓTESIS Y OBJETIVOS

2.1 Hipótesis de trabajo

Con una pequeña muestra en comparación con otras metodologías ya publicadas, pero el correcto uso de la información, el conocimiento de los datos, la adecuada selección y aplicación de ciertos algoritmos matemáticos, es posible predecir la potencia de una lente intraocular con mayor precisión y exactitud de lo que lo hacen actualmente otras fórmulas o métodos de cálculo.

2.2 Objetivo principal

Predecir la potencia refractiva de la lente intraocular que se implantará en una cirugía de cristalino mediante la creación de una nueva fórmula basada en inteligencia artificial, así como el residuo refractivo en plano corneal asociado al implante colocado.

2.3 Objetivos secundarios

1. Diseñar, desarrollar y validar un método de cálculo de la potencia de una lente intraocular que se implantará en la cirugía de cristalino.
2. Comparación del método generado en la tesis con todas las fórmulas actualizadas empleadas actualmente en la práctica clínica.
3. Comparación de resultados entre fórmulas teóricas y fórmulas basadas en IA.
4. Desarrollar una página web de acceso libre que aloje el algoritmo de IA.

3 INTRODUCCIÓN

3.1 Marco histórico

La primera lente intraocular (LIO) permanente en la edad moderna se implantó en 1949 por Sir Nicholas Harold Ridley (Figura 1) en el hospital St Thomas de Londres (11). Durante la segunda guerra mundial, el Dr. Ridley fue nombrado mayor del ejército británico y en estos años observó que pilotos de las Reales Fuerzas Aéreas Británicas (12) con fragmentos de plástico y vidrio, alojados en el ojo originarios de las carlingas de los aviones, permanecían inertes sin reacción tisular inflamatoria. Por aquel entonces, estas carlingas de los cazas británicos estaban fabricadas en polimetilmetacrilato (PMMA) para aligerar peso y conferir más velocidad al aeroplano. Este plástico se manufacturaba desde 1936 por la empresa británica Imperial Chemical Industries (ICI) y estaba patentado con el nombre comercial de *Perspex*.



Figura 1. Dr. Harold Ridley en 1950. (Imagen extraída de Apple, 1996)

Años más tarde, el Dr. Ridley en asociación con su amigo John Pike de la compañía Rayner que desde 1910 se dedicaba a la fabricación de lentes oftálmicas, con el que ya había colaborado anteriormente en el proyecto de un oftalmoscopio electrónico, idearon sentados en su coche en Cavendish Square en Londres los principios de una nueva cirugía, tal y como relata David Apple en su artículo publicado en enero de 1996 en la revista *Survey of Ophthalmology* (11).

El nuevo proyecto se llamaría "*The artificial lenticulus project*", el material seleccionado sería Perspex por sus características ópticas, estabilidad dimensional y la biocompatibilidad que había observado en los años de guerra el Dr. Ridley en los ojos de sus pacientes heridos. ICI y Rayner propulsaron la fabricación de la primera LIO renunciando a sus derechos económicos para mantener el "espíritu científico" de un proyecto limpio sin beneficios financieros. Téngase en cuenta que había un rechazo frontal de la comunidad científica del momento al implante plástico intraocular.

Para el cálculo de la potencia dióptrica de la primera lente implantada, Ridley y Pike se basaron en los trabajos del oftalmólogo sueco premio novel de medicina Allvar Gullstrand fallecido años antes. Pero debido a un error en los cálculos con los índices de refracción, la estimación del grosor del implante fue fallida, un menisco biconvexo demasiado grueso en este primer paciente intervenido en noviembre de 1949 provocó una alta hipercorrección miópica de 14.00 D.

Durante casi las próximas dos décadas fueron evolucionando las tecnologías que permitieron nuevos diseños de LIOs y nuevas técnicas quirúrgicas. Se introdujeron soportes para la óptica de la lente, llamados hápticos, y se realizaron múltiples diseños de apoyos.

Históricamente este fue un momento convulso en que la oftalmología establecida era hostil al uso de implantes intraoculares, lo que condujo a una invitación de Harold Ridley y Peter Choyce (precursor de la lente de apoyo angular), para que un grupo de oftalmólogos seguidores de estas técnicas se reunieran en la *Royal Society of Medicine* en Londres el 14 de julio de 1966 para discutir sobre sus experiencias médicas. Al final de aquella reunión, se acordó establecer el Club de Implantes Intraoculares (IIC) (Figura 2), que más tarde se convirtió en el Club Internacional de Implantes Intraoculares (IIIC).

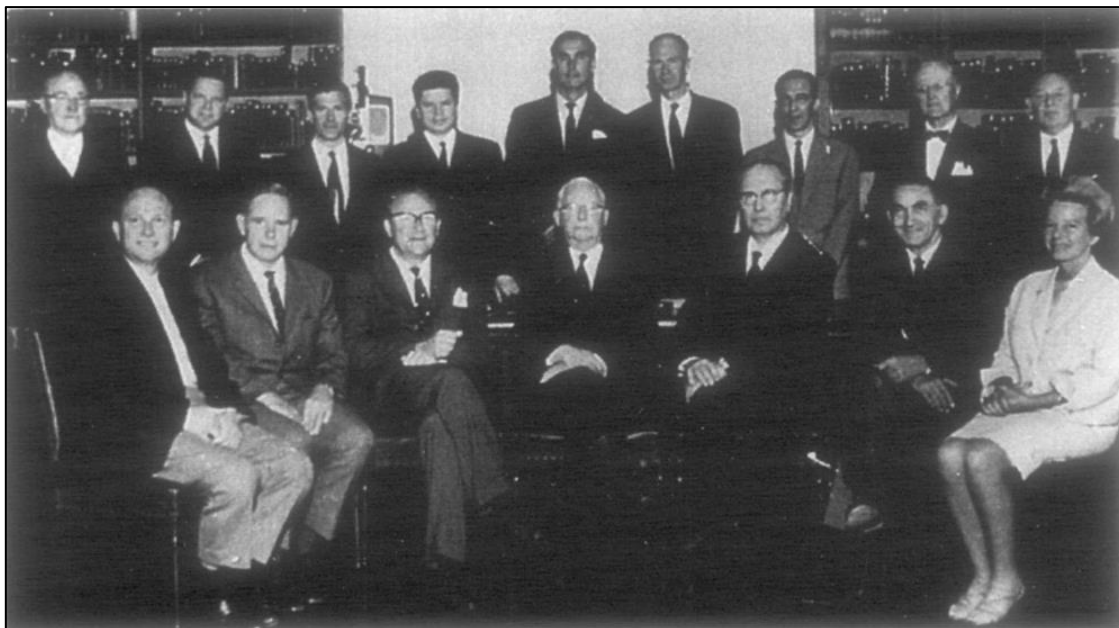


Figura 2. Miembros del Club de Implantes Intraoculares reunidos en 1966 en Londres. Fila de arriba de izquierda a derecha: John Pike (Rayner, RU), Robert Murto (EEUU), Michael Roper-Hall (Birmingham, RU), Svyatoslav Fyodorov (Moscú, USSR), Neil Dallas (Bristol, RU), C. A. (Sandy) Brown (Bristol, RU), Alexander Rubinstein (Birmingham, RU), Warren S. Reese (Wills Eye Hospital, Philadelphia, EEUU), Leonard Lurie (Londres, RU). Fila de debajo de izquierda a derecha: Jörn Boberg-Ans (Charlottenlund, Dinamarca), Cees Binkhorst (Terneuzen, Holanda), Peter Choyce (Moorfields, Londres, RU), Harold Ridley (Moorfields, Londres, RU), Benedetto Strampelli (Roma, Italia), Edward Epstein (Johannesburgo, Sudáfrica), Sonja Boberg-Ans (Charlottenlund, Dinamarca). (Imagen extraída de Apple, 1996)

La potencia refractiva de las lentes que se implantan hasta entonces era de valor fijo, alrededor de 19.00 D, con el objetivo de no dejar mucho más defecto refractivo que el previo. También se aplicaban ciertos valores de corrección en función del estado refractivo del paciente. En ojos anatómicamente normales, con longitudes axiales entre 22.00 mm y 25.00 mm se conseguía aproximadamente este objetivo, pero para ojos extremos las sorpresas refractivas postquirúrgicas eran una constante.

En 1956 Mundt y Hughes (13) fueron los pioneros en la ultrasonografía ocular, pero no fue hasta la segunda mitad de la década de 1960, que las medidas del eje anteroposterior ocular medidas con ecografía ultrasónica se utilizaran como variables biométricas en fórmulas de cálculo para la potencia refractiva de la LIO.

Se consolidaba la tecnología ultrasónica y su aplicación ocular. En este sentido los autores japoneses fueron muy prolíficos en 1961, puesto que Kanki et al (14) reportaron medidas de longitud axial en 694 ojos. Otsuka et al (15) midieron profundidad de la cámara anterior, espesor del cristalino y longitud axial en 46 ojos. Araki (16) publicó su estudio de medidas biométricas con 89 ojos.

Leary et al (17) publican un trabajo en el que muestran la repetibilidad de las medidas biométricas realizadas con tecnología ultrasónica y ciertas consideraciones técnicas en cuanto procesos de estandarización, de calibración y dosificación que bien podrían servir de base a lo que estaba por llegar pocos años más tarde.

3.2 Evolución y cronología de las fórmulas de cálculo

Tal y como sugieren Douglas Koch et al en su editorial (18), es conveniente desterrar la idea generacional para la clasificación de las fórmulas de cálculo relacionada únicamente con su cronología. De hecho, atendiendo a este sistema de ordenación si una fórmula nueva y actual como por ejemplo EVO (<https://www.evoicalculator.com/>) que está basada en óptica geométrica gaussiana de lente gruesa, tendríamos que clasificarla como fórmula de tercera generación, lo que resulta anacrónico. Es más preciso acometer esta labor atendiendo únicamente a la manera que tienen las fórmulas de calcular la potencia de la lente emetropizante. De tal manera que se distinguirán seis tipos de categorías (Figura 3): fórmulas teóricas de lente fina, teóricas de lente gruesa, empíricas, trazado de rayos, basadas en IA y combinadas teóricas con IA.

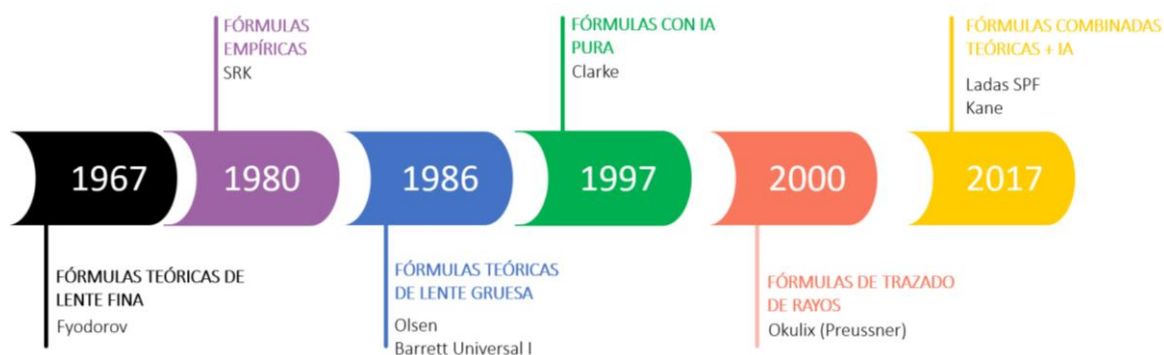


Figura 3. Línea de tiempo y años de aparición con los distintos abordajes para el cálculo de la potencia de una lente intraocular y los pioneros en cada etapa.

Cuando se habla de fórmula de vergencia óptica, se habla de fórmula teórica, son conceptos no idénticos pero muy parecidos (19). Las fórmulas de vergencia se fundamentan en el comportamiento de la luz al encontrar un sistema óptico en su camino y están gobernadas por una expresión elemental, un cociente entre el índice de

refracción del medio en el que la luz viaja y la distancia que recorre. Todo sistema óptico se podría tratar desde el punto de vista vergencial (Figura 4) conocidas las vergencias objeto e imagen así como los índices de refracción de separación de medios, lo que implica que incluso si el enfoque fuera con lente gruesa también se podría realizar con fórmulas de vergencia. De hecho la primera versión de la fórmula de Barrett se basa en un abordaje vergencial contando con el grosor de la lente.

Quizás, el tratamiento matemático con óptica geométrica paraxial o gaussiana, probablemente sería algo más complejo, pero se obtendría el mismo resultado.

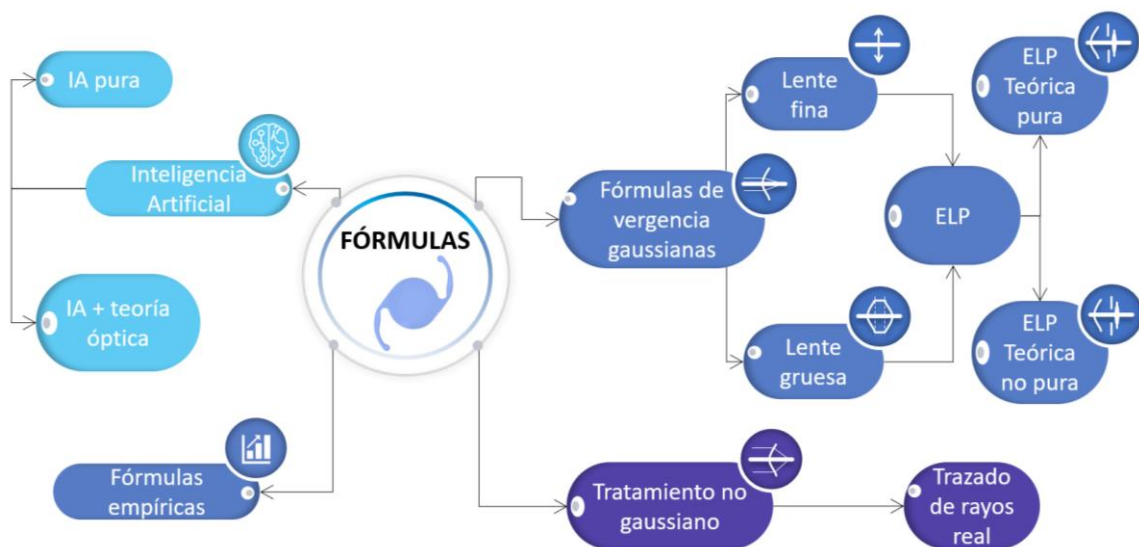


Figura 4. Esquema de funcionamiento de todos los abordajes que existen para calcular la potencia de una lente intraocular en una cirugía de cataratas.

Tal y como se expresó anteriormente, el poder dióptrico de los rayos que vienen del objeto sumado al poder dióptrico de la lente, será igual al poder dióptrico de los rayos que van a la imagen. De manera que todas las fórmulas de vergencia siguen el mismo esquema teórico, la vergencia de la luz en la retina (V_{retina}) es la resultante de la suma de la vergencia en la LIO (V_{LIO}) y la potencia de esta (P_{LIO}) (Figura 5). Por tanto, despejando en la ecuación se obtiene una fórmula de cálculo de la LIO teórica paraxial (3):

$$V_{L10} + P_{L10} = V_{retina} \quad (1)$$

$$\frac{n}{\frac{n}{K} - ACD_{postop}} + P_{L10} = \frac{n}{AXL - ACD_{postop}} \quad (2)$$

$$P_{L10} = \frac{1.336}{AXL - ACD_{postop}} - \frac{1.336}{\frac{1.336}{K} - ACD_{postop}} \quad (3)$$

Donde K es la potencia corneal, AXL la longitud axial, n es el índice de refracción del humor acuoso y del vítreo que se iguala a 1.336 por consenso y ACD_{postop} es la posición en la que ha quedado la lente implantada respecto a la córnea, lo que hoy conocemos como *Effective Lens Position* (ELP), pero en los años 70 todavía no se había acuñado dicho término.

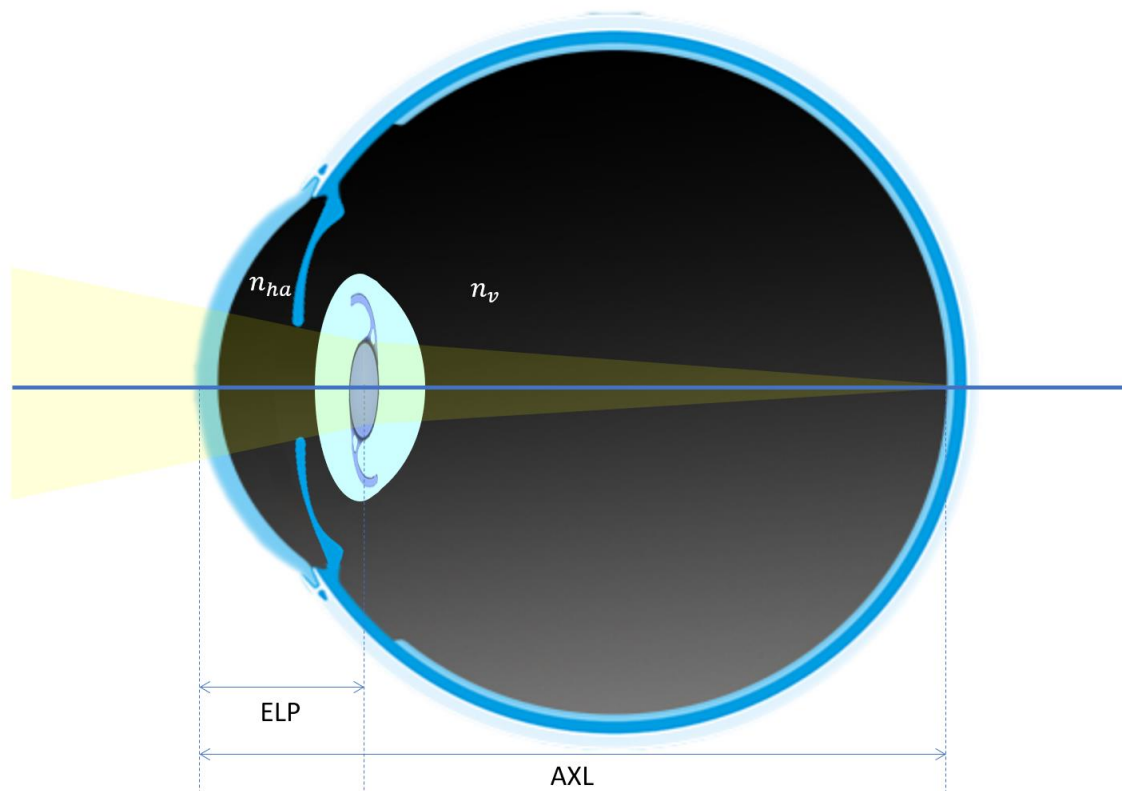


Figura 5. Vergencias sobre la lente intraocular y sobre la retina.

Hay cierta controversia para establecer el origen cronológico de las fórmulas de cálculo, ya que Hoffer (20) en un trabajo publicado en 1982 asegura que la primera fórmula teórica corresponde al oftalmólogo francés R. Dudragne y se publicó en 1955, aunque no se popularizó. Sin embargo, Binkhorst (21) en una publicación de 1972 en el *British Journal Ophthalmology* dice que también autores franceses son los precursores de la formulación teórica atribuyéndosela a Maria, Bonnet y Cochet (1958), Le Grand (1965). Si bien, en toda la literatura oftalmológica revisada publicada, se le atribuye la primera fórmula teórica a Svyatoslav Fyodorov en el año 1967 (22), no obstante resulta complicado establecer el punto de partida aquí.

Fyodorov en su trabajo reportaba los resultados de 150 ojos con implantes diseñados por él mismo en 1960, a los que llamó Sputnik®. El artículo del oftalmólogo ruso ponía de manifiesto la inoperancia de calcular potencias dióptricas de los implantes a partir de la refracción prequirúrgica tal y como se había estado haciendo hasta ese momento. Su fórmula teórica era una fórmula de vergencia sencilla basada en óptica gaussiana asumiendo el implante intraocular de grosor despreciable. Utilizando su misma notación, en esta fórmula (4) hay tres variables preoperatorias medibles: la longitud axial del globo ocular, la potencia promedio de los dos meridianos principales corneales y el diámetro horizontal corneal de limbo a limbo conocido en inglés como *White to White* (WTW), para poder inferir la posición en la que iba a quedar dispuesta la lente implantada en el saco capsular.

$$D_p = \frac{n - aD_c}{(a - k) \left(1 - \frac{kD_c}{n}\right)} \quad (4)$$

Donde D_p es la potencia dióptrica de la lente que se quiere implantar, n es el índice de refracción del humor acuoso y del vítreo (1.336), a es la longitud axial expresada en metros, D_c es el poder refractivo de la córnea medido en dioptrías y k es la distancia a la que quedará la lente intraocular de la córnea medida en metros.

Fyodorov hizo el cálculo de k asumiendo por simplificación geométrica que esta fuera igual a la altura de un segmento esférico, cuyo radio es igual al radio de curvatura de la córnea, y cuyo diámetro interno es de 1.00 mm más que la del diámetro externo visible de la córnea, derivando la siguiente fórmula de ságita corneal:

$$k = r - \sqrt{r^2 - \frac{d^2}{4}} \quad (5)$$

Donde k es la profundidad a la que quedará implantada la lente intraocular desde el apex corneal, r es el radio de curvatura de la córnea y d es un segmento parecido al diámetro horizontal corneal, lo que hoy llamamos distancia blanco a blanco o del inglés WTW, más 1.00 mm.

Este enfoque teórico tan novedoso propuesto por Fyodorov para predecir la posición de la lente intraocular, fue utilizado posteriormente por Olsen o Holladay en sus respectivas primeras fórmulas, al que llamaron ságita o altura corneal.

La década venidera fue muy productiva en el ámbito del cálculo implantacional, ya que se publicaron multitud de fórmulas teóricas basadas en los mismos principios ópticos que había utilizado Fyodorov. Todas ellas pertenecen a una primera generación distribuidas en la década de los 70 y son Gernet et al. (23), Leary (24), Binkhorst (21), Colenbrander (25), Van der Heijde (26), Oguchi and Van Balen (27), Thijssen (28), Binkhorst (29), Epstein et al (30), Kollarits (31) y Drews (32-34).

Todas estas fórmulas tenían el mismo fundamento óptico pero diferían entre sí por pequeños ajustes teóricos y operaban únicamente con dos variables biométricas medidas en clínica: la longitud axial, medida mediante ultrasonografía, y la potencia refractiva corneal. La profundidad de la cámara anterior postoperatoria (ACD_{postop}), que Fyodorov llamó k , era constante y dependía del tipo de implante que se pusiera.

En la prolífica década de los 70, los fracasos conseguidos con las lentes de apoyo angular reportados en la literatura publicada de la época (35), fomentaron el interés de nuevo por los implantes de cámara posterior de PMMA con hápticos de polipropileno (Prolene®) gracias a los trabajos de Choyce, Binkhorst o Fyodorov (36), y la cantidad de diseños y evoluciones mejoradas propuestos durante esta época. En esos años también se conformaron las primeras sociedades de cirugía implantacional a nivel mundial, concretamente en España en 1984 que se llamó inicialmente Club Español de Cirugía ocular Implanto Refractiva (CECIR), actualmente conocida como Sociedad Española de Cirugía Ocular Implanto Refractiva (SECOIR).

Rotundamente se puede establecer otro punto de inflexión histórico en el año 1981, cuando Binkhorst (37) establece una relación entre la ACD_{postop} y la longitud axial. Hito fundamental si se tiene en cuenta que hasta entonces solo se infería teóricamente o era constante en función del diseño de lente implantado. Pocos años más tarde, Hoffer (38) publica una relación matemática obtenida por regresión lineal sobre 153 ojos para ACD_{postop} :

$$ACD_{postop} = 0.292 * AL - 2.93 \quad (6)$$

Por tanto, en estos años, la posición en la que quedaba el implante respecto a la córnea se convirtió en una variable que cambiaba en función del eje anteroposterior del ojo: cuanto mayor era éste, mayor era el posicionamiento postquirúrgico de la lente.

Hasta este momento, la utilización de fórmulas teóricas basadas en mediciones clínicas preoperatorias como la longitud axial y la queratometría, mejoraron mucho la predictibilidad refractiva de la cirugía implantacional.

De forma paralela en estos años, Sanders, Kraff, Retzlaff, y Gills (12, 39, 40) fueron los precursores de las fórmulas empíricas, es decir, fórmulas basadas en datos retrospectivos que no inferían variables preoperatorias no medibles clínicamente como la profundidad de la cámara anterior postoperatoria (ACD_{postop}), sino que directamente aportaban el valor de la potencia de la lente a implantar.

El trabajo de Sanders y Kraff se basaba en una muestra de 923 ojos con cinco tipos de implantes diferentes. Sus resultados los compararon con dos de las fórmulas más empleadas hasta la fecha, Binkhorst y Colenbrander. Hillis (41) en 1981 dio buena cuenta del error técnico que habían cometido Sanders y Kraff en sus trabajos, utilizando la misma muestra para el análisis regresivo que para la comparación. Utilizando la misma notación, la fórmula regresiva que obtuvieron fue:

$$P = A - 2.5L - 0.9K \quad (7)$$

Donde P es la potencia emetropizante de la lente intraocular a implantar, A es una constante que depende del tipo de lente intraocular, L es la longitud axial y K la queratometría media.

Por su parte Retzlaff en un trabajo independiente publicado también en 1980, basó su regresión lineal en una muestra de solo 176 ojos obteniendo la siguiente fórmula:

$$P = A - 2.41L - 0.87K \quad (8)$$

Debido a la coincidencia cronológica, de enfoque y matemática, los tres autores patentaron sus fórmulas empíricas como un acrónimo formado por sus apellidos: SRK™.

En 1981 publicaron un trabajo (42) sobre 654 ojos en el que comparaban los resultados obtenidos con su fórmula SRK™, Binkhorst y Colenbrander. Los autores en su artículo mencionan el editorial crítico de Hillis (41) respecto a la derivación de una fórmula regresiva de una muestra comparando sus resultados con otras fórmulas en la misma muestra, y el error que esto supone. Pero la realidad es que SRK™ fue diseñada así y no hubo cambios en este sentido hasta su versión posterior. Además, debido a que las muestras en las que se basó no contemplaban ojos axialmente extremos, en estos casos la fórmula no funcionaba bien (43).

Durante los siguientes años, coexistieron las primeras fórmulas teóricas con las empíricas más conocidas Gills, SRK™ y la nueva de Thompson et al que en el año 1984 publicaron su propia fórmula a la que llamaron TI-59 (43) para ojos axialmente largos.

Se estaba gestando una nueva revolución en las fórmulas teóricas. Thomas Olsen y Graham Barrett son aproximadamente coincidentes en el tiempo en cuanto a la presentación pública de sus fórmulas, ya que sus respectivos trabajos fueron publicados en la misma revista con solo 4 meses de diferencia. El primero en hacerlo fue el oftalmólogo danés y posteriormente el australiano. La gran diferencia entre ambos trabajos, y no es baladí mencionarlo, es que Olsen elabora una fórmula para un diseño

concreto de lente intraocular planoconvexa, sin embargo Barrett deja abierta su fórmula a cualquier diseño de implante independientemente de su índice de refracción, curvatura y espesor, de hecho la llamó *Universal Theoretical Formula*.

Sobre todo, lo que ambos impulsaron fue la necesidad de implementar mejoras en los cálculos introduciendo una nueva operativa sobre los planos principales del implante intraocular.

En 1986 Olsen (44) publica un trabajo en el que determina de forma teórica la profundidad de la cámara anterior postoperatoria, prolegómenos de su fórmula publicada un año después. Su hipótesis de trabajo se basó en el concepto de altura corneal acuñado por Fyodorov 20 años antes. Con una muestra de 80 pacientes con datos pre y postoperatorios de medición de ACD_{postop} a los 6 meses, mediante una regresión lineal múltiple obtuvo la siguiente ecuación:

$$ACD_{postop} = 0.87 + 0.25 * H + 0.33 * C + 0.17 * L - 0.03 * A \quad (9)$$

Donde H , C , L y A son la altura corneal, la cámara anterior preoperatoria, el espesor del cristalino y la longitud axial respectivamente. H depende del radio de curvatura corneal y del WTW. Era la primera vez que se relacionaba la ACD_{postop} con la queratometría y con la longitud axial. Esta era la primera vez que se relacionaba la ACD_{postop} con otras variables predictoras que no fuera la longitud axial.

Olsen se había preparado el camino para la publicación de su fórmula un año más tarde en 1987 (45). Los tres factores diferenciales respecto a las fórmulas teóricas vistas hasta el momento fueron: considerar la córnea gruesa, la estimación teórica de la profundidad de la cámara anterior postoperatoria y la posición de los planos principales

del sistema óptico completo (LIO + córnea). En primer lugar incide mucho en la posición de los planos principales de la córnea, para ello considera la córnea como un dioptrio grueso de 0.50 mm de espesor y un índice queratométrico basado en mediciones de Gullstrand hechas 90 años antes, de 1.3315, asumiendo que el ratio de curvatura corneal posterior/anterior era de 0.883. Hasta la fecha, la córnea teórica al no poderse medir la curvatura posterior fue una lente fina con un índice queratométrico ficticio de 1.3375. En segundo lugar, asumió un implante planoconvexo de índice de refracción, grosor y curvatura conocidos. Los planos principales de la lente intraocular dadas estas características se calcularon. Y por último la profundidad de la cámara anterior postoperatoria fue estimada por regresión lineal multivariante como se mencionó más arriba, en su trabajo publicado un año antes.

Mediante diferencia de vergencias, teniendo en cuenta la localización de los planos principales de LIO y córnea y su distancia, determinó la potencia del implante. Barrett, (46) en su publicación de 1987, explica la importancia de tener en cuenta el diseño óptico y el material del implante intraocular: no es lo mismo una lente biconvexa que plano convexa, ya que sus planos principales no están en la misma posición y por supuesto el espesor no puede ser despreciable. Hasta ese momento todas las fórmulas teóricas descritas estaban basadas en óptica geométrica gaussiana considerando la lente intraocular fina y a efectos ópticos la posición de los planos principales imagen y objeto compartían espacio. La gran novedad y el aporte que hizo Barrett en este trabajo fue derivar una fórmula teórica gaussiana (Figura 6) de lente gruesa teniendo en cuenta el diseño de la lente.

$$DL = \frac{n_1 \times 1000}{AL - d - t + \frac{n_1}{n_2} \times \frac{DLd - FL_2}{(1 - t \times FL_2)} \times \frac{t}{DLd}} - \frac{n_1 \times 1000}{\frac{n_1}{Rs} - d - \frac{n_1}{n_2} \times \frac{FL_2}{DLd} \times t} \times \frac{1}{k + \frac{1}{1 - S.Rs}}$$

- n_1 = refractive index of aqueous
 n_2 = refractive index of implant
AL = axial length
d = anterior chamber depth
t = lens thickness
DL = dioptric power of implant for desired postoperative refractive result
DLd = designated dioptric power of implant for iterative analysis
 FL_2 = power of posterior surface of implant = $\frac{N_2 - N_1}{R}$
R = posterior radius of curvature of implant
k = preoperative keratometry
Rs = desired postoperative spherical equivalent
S = distance of spectacle plane from corneal vertex in meters

Figura 6. Fórmula de Barrett universal año 1987: fórmula gaussiana de vergencia teniendo en cuenta el grosor del implante para el cálculo de la ACD_{postop} . (Imagen extraída del artículo de Barrett (46)).

Hasta este momento, no hay antecedentes de fórmulas teóricas gaussianas que estimen la posición postquirúrgica del implante a través de un sistema óptico con lente gruesa independientemente del diseño del implante. Todas las fórmulas teóricas descritas eran de lente fina, lo que implica la despreciabilidad del grosor del implante. A esta fórmula la llamó universal porque el modelo teórico era válido para todas las tipologías de ojo así como todos los diseños de lentes intraoculares de la época y venideros.

En este sentido, para que la fórmula fuera operativa para cualquier diseño de lente intraocular, debían conocerse algunas características físicas del implante como el índice de refracción, el espesor y el radio de curvatura de la superficie posterior. Dedujo seis ecuaciones, unificadas en una sola (Figura 6), que se resolvían secuencialmente

mediante una técnica iterativa dando a una de las variables (DLd) un valor de partida para la obtención de la solución final.

En 1988 Holladay et al (47) publican la primera versión de su fórmula en la que introduce su *Surgeon Factor* (SF) como nueva constante para el cálculo de la potencia de la lente intraocular. Este factor representa la distancia que hay desde el plano iridiano anterior al plano principal del implante que sumado a la profundidad de la cámara anterior en un ojo pseudofáquico, conformará la posición efectiva de la lente (ELP). Fue en este momento desde el que a la ACD_{postop} se le comenzó a llamar ELP, así se hará en este trabajo de ahora en adelante. Téngase en cuenta que esta distancia se toma desde los planos ópticos principales, y no desde las superficies ópticas.

La profundidad anatómica pseudofáquica es la distancia que hay desde el apex corneal hasta el plano iridiano anterior. Dicha distancia es más predecible a través de la curvatura corneal y de la longitud axial, tal y como propone Holladay en su trabajo.

Por tanto, para determinar la ELP, Holladay tenía que calcular dos parámetros; SF y profundidad de la cámara anterior pseudofáquica.

$$ELP = sf + ACD_{pseudofáquica} \quad (10)$$

Holladay defendió su factor SF como constante de la lente a implantar, de hecho hoy en día es una de las constantes ofrecidas por los laboratorios. Esta distancia dependerá directamente de las características físico-ópticas del implante: espesor y curvaturas de sus superficies. Para su cálculo, Holladay propone en su trabajo cuatro variables usadas como inputs: la queratometría, la longitud axial, la potencia del implante

y la refracción postoperatoria estabilizada. De este modo y con una muestra significativa (él estimó entre 10 y 25 pacientes), retrospectivamente se podía optimizar el valor de SF.

Al igual que hizo Olsen anteriormente, Holladay también empleó el concepto de altura corneal acuñado por Fyodorov para el cálculo de $ACD_{pseudofáquica}$:

$$ACD_{pseudofáquica} = R - \sqrt{R^2 - \frac{AG^2}{4}} \quad (11)$$

Donde R es el radio de curvatura corneal medio y AG es la anchura corneal. En esta expresión impuso una serie de condicionantes biométricos no explicados en su trabajo, el radio de curvatura medio debía ser superior a 7.00 mm y AG como máximo podía ser 13.50 mm. A su vez la anchura corneal la relacionó con la longitud axial mediante la siguiente expresión:

$$AG = 12.5 * \frac{AXL}{23.45} \quad (12)$$

En este trabajo no se referencia ni a Olsen ni a Barrett, pero sí compara los resultados obtenidos mediante su nueva fórmula (Holladay I) con seis fórmulas teóricas y empíricas del momento, entre las que estaban Fyodorov, Binkhorst y SRK™ II.

Paralelamente las fórmulas empíricas sufrían también mejoras, apareció la evolución de SRK™, la SRK™ II (48). Los autores en esta nueva versión ampliaron la muestra a 2068 ojos en los que se incluyeron más ojos extremos.

En 1990 se publicó SRK/Theoretical o más conocida como SRK/T™ (49). SRK/T™ es una combinación teórico-geométrica con otra empírica. Los autores emplean una

parte teórica en el diseño de su fórmula para predecir la $ACD_{postoperatoria}$ o ELP , y para ello tienen que calcular la altura corneal, como ya hicieran previamente en sus fórmulas teóricas Fyodorov, Olsen y Holladay.

Como se ha visto anteriormente, en la ecuación de altura corneal (11) hay una variable llamada anchura corneal que no se corresponde exactamente con el WTW, por tanto necesitaban inferirla matemáticamente. Aquí es dónde entra en juego la parte empírica, ya que basados en una muestra de 1677 ojos, para determinar la anchura corneal, utilizan una regresión lineal con la longitud axial y la queratometría media como variables predictoras. Utilizando la misma notación que los autores:

$$C_{width} = -5.41 + 0.58412 * L_{cor} + 0.098 * K \quad (13)$$

Donde K es la queratometría media y L_{cor} es la longitud axial corregida. Este factor de corrección para los ojos axialmente largos, concretamente por encima de 24.20 mm, fue este:

$$L_{cor} = 3.446 + 1.715 * L - 0.237 * L^2 \quad (14)$$

Donde L es la longitud axial en milímetros utilizando su nomenclatura.

El argumento que utilizaron al analizar su muestra, es que que los ojos muy miopes tienen una elongación proporcionalmente mayor de la parte posterior que la del segmento anterior. Este fenómeno podría ocurrir gradualmente con ojos cada vez más largos en lugar de abruptamente (Figura 7). Es por ello que corrigieron mediante una ecuación cuadrática la longitud axial para evitar precisamente lo que a fórmulas como

Hoffer y Holladay les pasaba en la predicción de $ACD_{postoperatoria}$ en ojos miopes como se puede observar en la figura:

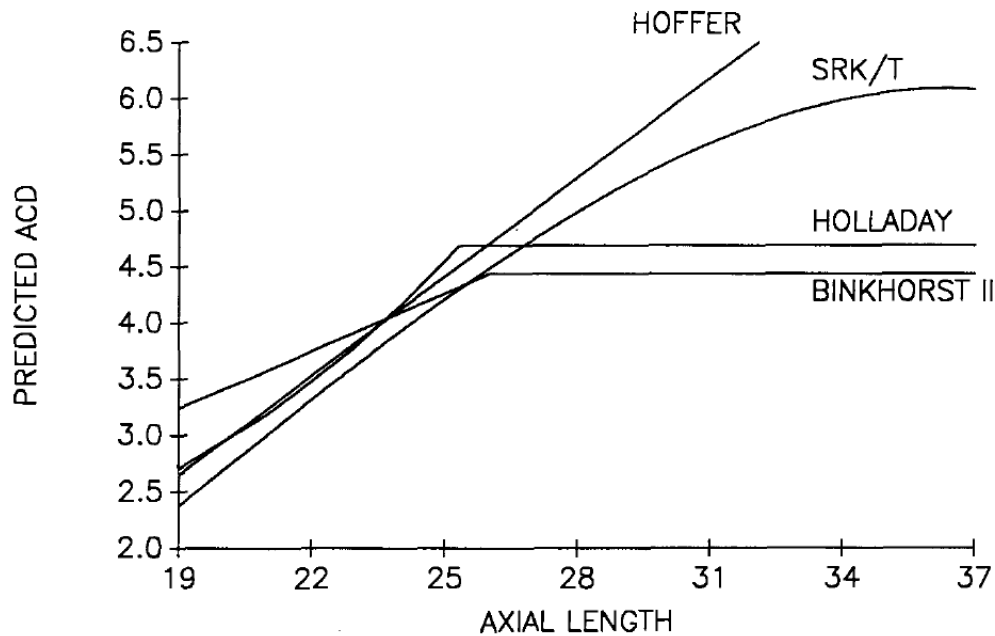


Figura 7. Variación del posicionamiento de la lente implantada en función de la longitud axial para 4 fórmulas. (Imagen extraída del trabajo de Retzlaff J. et al (49)).

Otro factor importante que utilizaron los autores en su fórmula fue la elección del índice queratométrico. De hecho tal y como ellos mismos expresan en su trabajo, hicieron un acercamiento heurístico de esta cuestión, probando varios índices ya testados previamente por otros autores como Binkhorst, Olsen o Holladay, seleccionando finalmente 1.333.

En 1990 Olsen et al (50) propone una nueva regresión multivariante para la predicción de la ELP, lo que hace mejorar la primera versión de su fórmula. En esta ocasión las variables predictoras son longitud axial, profundidad de la cámara anterior preoperatoria y la altura corneal como ya la hemos descrito anteriormente. Por tanto,

elimina de la ecuación el espesor del cristalino respecto a lo publicado cuatro años antes por él mismo.

Tres años más tarde, en 1993, Hoffer publica su fórmula Hoffer Q (51) especialmente considerada para ojos axialmente por debajo de 22.00 mm.

Este mismo año, Barrett (52) publica la evolución de su fórmula universal. Intentó mejorar la fórmula previa de 1987 conservando las expresiones para los planos principales de refracción del implante considerándolo grueso, de tal manera que el usuario no tuviera que conocer los factores de curvaturas, índice de refracción y espesor de la lente intraocular para poder usar la fórmula con precisión.

En el año 1995 Olsen et al (53) publican la segunda versión de su fórmula en la que se propone una nueva ecuación regresiva para la estimación de la posición efectiva, esta vez con cuatro variables:

$$ACD_{postoperatoria} = ACD_{const} + 0.50 * ACD_{pre} + 0.10 * A + 0.15 * H + 0.20 * L - 5.38 \quad (15)$$

Donde ACD_{const} es la constante ACD de la lente, ACD_{pre} es la profundidad de la cámara anterior prequirúrgica, A es la longitud axial, L es el espesor del cristalino y H es la altura corneal.

La segunda versión de la fórmula de Holladay no tiene publicación (54) en revistas *peer-review*, aún así data de 1996. Para Holladay II la estimación de la posición efectiva introduce más predictores, concretamente siete, incluidos la edad y la refracción previa a la cirugía. De hecho es la fórmula teórica que más variables regresoras incluye.

En este mismo año, W. Haigis (55) publicó su fórmula en una revista alemana. Se trata de una fórmula de vergencia de lente fina con la particularidad de que utiliza un índice queratométrico de 1.3315 y una *vertex distance* de 12.00 mm. La estimación de la posición efectiva de la lente mediante regresión lineal, siendo la ecuación resultante:

$$ACD_{postoperatoria} = a_0 + ACD * a_1 + AXL * a_2 \quad (16)$$

La estimación del posicionamiento del implante sólo depende de una combinación lineal de dos variables y ninguna es la queratometría. La constante a_0 está directamente relacionada con la constante ACD del fabricante también conocida como pACD.

Sin embargo los valores de a_1 y a_2 se predeterminaron en $a_1 = 0.4$, $a_2 = 0.1$ (56). Estos parámetros a_0 , a_1 y a_2 pueden optimizarse mediante el análisis postoperatorio de los resultados refractivos, en realidad como con cualquier fórmula. En este caso, para cada paciente, el valor de $ACD_{postoperatoria}$ o ELP que conduce a una refracción postoperatoria perfecta se determina a partir de la ecuación vergencial típica. Las posiciones efectivas de la lente individuales así obtenidas se correlacionan con los datos preoperatorios para la profundidad de la cámara anterior y la longitud axial de acuerdo con la ecuación de regresión. Posteriormente de este análisis de regresión, las constantes optimizadas a_0 , a_1 y a_2 se obtienen fácilmente.

Por último, otra metodología para calcular la potencia de la lente intraocular ha sido el *ray tracing* o trazado de rayos real. Fueron los japoneses Ishikawa et al (57, 58) los primeros en publicar al respecto sobre todo en pacientes intervenidos previamente con

procedimientos quirúrgicos corneales refractivos. Posteriormente Preussner (59) con la primera versión de Okulix® (<http://okulix.de/>) y Norrby (60).

El trazado de rayos se basa en la no paraxialidad de la formulación, es decir en utilizar todas las superficies ópticas involucradas sin restricciones angulares y con todas las implicaciones geométricas de cada sistema.

En la aproximación paraxial la ley óptica fundamental de Snell se reduce, debido a que si los rayos entrantes en un sistema óptico con curvaturas están muy cerca del eje óptico, el seno del ángulo puede equivaler al propio ángulo, lo que se conoce como aproximación de primer orden:

$$n \sin \theta = n' \sin \theta' \quad (17)$$

$$n\theta = n'\theta' \quad (18)$$

La consecuencia más relevante de esta aproximación es que la posición de la imagen no depende del valor del ángulo de incidencia. Así, todos los rayos que salen de un punto objeto se transforman estigmáticamente en un punto imagen, pero es sabido que los sistemas ópticos no son estigmáticos, y mucho menos el ojo.

En general, los rayos de luz que entran en una lente o en un sistema óptico lo hacen formando ángulos de incidencia suficientemente grandes, de tal manera que, si se calculan las marchas de tales rayos a partir de las leyes exactas de la reflexión y refracción, estas no coincidirán con las predichas por la aproximación paraxial. Esta circunstancia implica que la imagen de un punto no sea un punto sino una distribución de luz entorno a la imagen paraxial. La imagen de un objeto extenso dejará de ser nítida e incluso podrá

aparecer distorsionada, dejando de mantener las proporciones entre sus partes. Estas desviaciones del comportamiento paraxial de los rayos reales dan lugar a aberraciones ópticas y los procedimientos por trazado de rayos las tienen en cuenta.

De los dos sistemas ópticos involucrados en estos cálculos, se dispone de la tecnología necesaria para mapear la córnea mediante topógrafos que puedan analizar la cara posterior. Sin embargo, el problema fundamental para que el trazado de rayos sea lo más exacto posible es que es necesario conocer parámetros de la lente a implantar a menudo desconocidos porque suelen estar sometidos a secreto industrial, estos son, el espesor, el índice de refracción y los radios de curvatura de ambas caras.

Actualmente no se utilizan demasiado estos sistemas de cálculo en comparación con las fórmulas teóricas gaussianas. De hecho, hay pocos estudios publicados en la literatura revisada (61, 62) siendo algo más profusa en pacientes intervenidos previamente con técnicas refractivas corneales.

3.3 Siglo XXI y nuevos enfoques para el cálculo biométrico

En los próximos apartados se introducirán los conceptos fundamentales sobre los que se asienta la presente tesis doctoral. Se intentará de manera progresiva explicar la introducción de la IA y sus derivados, así como esta ha irrumpido en el mundo oftalmológico invadiéndolo todo, haciendo hincapié en los sistemas actuales de cálculo de la potencia de lente intraocular relacionados con estas tecnologías.

3.3.1 Inteligencia Artificial y Machine Learning.

El término Inteligencia Artificial se le atribuye a John McCarthy en 1956 (63). En una conferencia organizada por él mismo en la Universidad de Darmouth planteó su definición de IA: "*Este estudio procederá sobre la base de la conjetura de que todos los aspectos del aprendizaje o cualquier otro rasgo de la inteligencia pueden, en principio, ser descritos de una forma tan precisa que se puede crear una máquina que los simule*".

Hoy en día, sin haber alcanzado todavía la simulación fidedigna del pensamiento abstracto humano por una máquina y ni mucho menos a la inteligencia emocional, se puede decir de forma general que la IA es aquella inteligencia que muestran las máquinas, en contraste con la inteligencia natural de los humanos. En la actualidad, las máquinas son capaces de ejecutar acciones *inteligentes* por sí mismas incluso aprendiendo sin supervisión humana.

La IA, desde la década de los 50 hasta la actualidad, ha discurrido paralela a la tecnología de la computación y supercomputación. A medida que estas tecnologías han adquirido más velocidad de procesamiento, más eficiencia energética, más capacidad de almacenamiento y más nanometrización en sus arquitecturas, el interés industrial y financiero por la IA se ha exponencializado. Gracias a la globalización informática y a la democratización del conocimiento científico, se disponen de mejores algoritmos, más depurados, rápidos y eficientes, además de una cantidad ingente de datos (*big data*) a nuestra disposición. Estos son factores determinantes en este despunte de la IA en los últimos tres lustros.

Actualmente la IA convive con nosotros, es el presente: reconocimiento de voz, de textos o de imagen en nuestros terminales telefónicos móviles, algoritmos de análisis comercial, evaluación de riesgos empresariales, predicciones meteorológicas, motores de búsqueda inteligentes en navegadores, ciberseguridad, interpretación de resultados médicos, herramientas de exploración biomédica, medios de transporte autónomos, hiperautomatización industrial, aplicaciones domóticas y un larguísimo etcétera forman parte del día a día.

El Machine Learning (ML) o aprendizaje automático de máquinas es un campo de la IA, al igual que el Deep Learning (DL) o aprendizaje profundo. En esencia, se podría decir que ML consiste en desarrollar algoritmos o secuencias de instrucciones, a las que llamaremos modelo, que permitan a la máquina desempeñar una tarea predictiva y/o clasificatoria, de la que aprenda de manera autónoma sin necesidad de la intervención humana de nuevo gracias a un entrenamiento iterativo con datos reales. En otras palabras, se trata de crear algoritmos capaces de generalizar comportamientos y reconocer patrones a partir de una información suministrada en forma de ejemplos.

Se pueden distinguir dos tipos de aprendizaje en ML, el supervisado y el no supervisado. La distinción entre uno y otro dependerá del tipo de datos reales que se utilicen para el entrenamiento de los modelos implementados. Cuando los datos empleados reportan o incluyen los resultados, se habla de ML supervisado. Sin embargo si los datos no contienen esos resultados el modelo será no supervisado. En el caso de esta tesis doctoral, el *dataset* contendrá los resultados que corresponden a la potencia de la lente implantada y al residuo refractivo resultante por lo que se trata de un aprendizaje supervisado.

Un ejemplo sencillo de modelo ML supervisado es el algoritmo de los k vecinos más cercanos (k -NN o k *Nearest Neighbour*). Es un algoritmo de clasificación y/o regresión supervisado basado en criterios de vecindad, quizás por su sencillez el primero que se aprende y que se aplica en el modelaje de ML. Se basa en la idea de que los nuevos ejemplos serán clasificados según la clase a la cual pertenezca la mayor cantidad de vecinos más cercanos a ellos del conjunto de datos del entrenamiento. La cercanía de los vecinos respecto al nuevo *input* que queramos clasificar o predecir, se determina mediante el cálculo matemático de distancias.

Otro ejemplo sencillo de modelo de ML no supervisado sería el algoritmo de las K -medias o conocido como *K-means* en inglés. Es un método de agrupamiento o *clustering*, que tiene como objetivo la partición de un conjunto de n observaciones en k grupos en el que cada observación pertenece al grupo cuyo valor medio es más cercano.

3.3.2 Inteligencia Artificial en oftalmología

Existen 772 artículos publicados relacionados con IA y sus variantes en oftalmología desde 1971 (Figura 8). Dicha búsqueda avanzada se practicó en Pubmed.gov el 06-02-2020 a las 19:00 horas utilizando MeSH como diccionario de vocabulario terminológico y sinónimos para expresiones como “Machine Learning”, “Deep Learning”, “Artificial Intelligence” y “Neural Network”. En la búsqueda se emplearon operadores booleanos AND y OR sin truncamientos, quedando la secuencia de instrucciones como sigue:

((((((((((((((("Artificial Intelligence"[Mesh] AND ophthalmology))) OR ("Machine Learning"[Mesh] AND ophthalmology))) OR ((Neural Networks, Computer AND ocular)))))) OR ("Neural Networks, Computer"[Mesh] AND ophthalmology))) OR ("Deep Learning"[Mesh] AND "Ophthalmology"[Mesh]))) OR (((inteligence artificial) AND ophthalmology))).

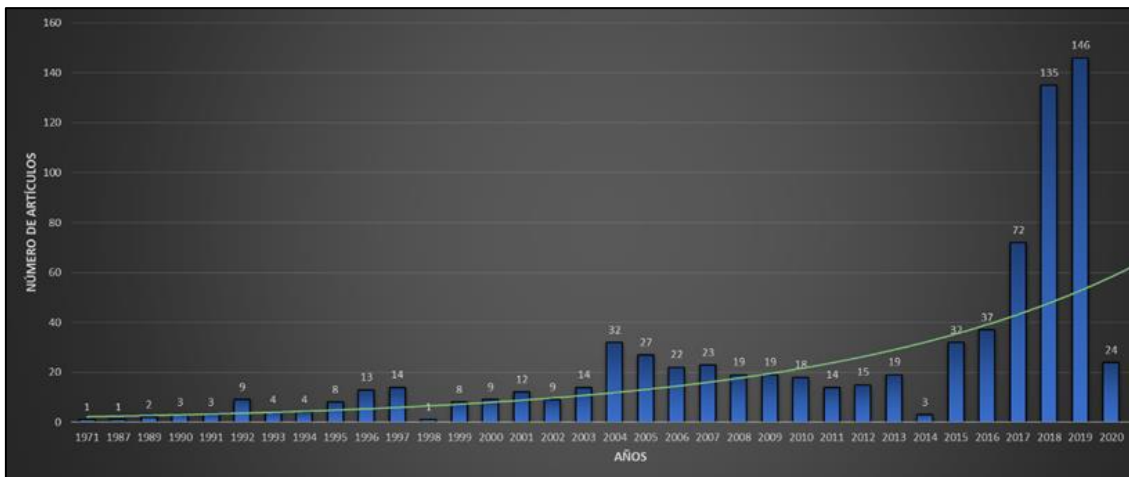


Figura 8. Artículos publicados relacionados con IA en oftalmología desde el año 1971.

Como se puede observar en la Figura 8, el número de publicaciones relacionadas con IA en oftalmología tiene un crecimiento exponencial sobre todo en el último lustro.

La irrupción de la IA en el ámbito de la oftalmología como especialidad médica fue sobre todo en los años 80 mediante la introducción de redes neuronales. Según un trabajo de revisión reciente publicado por WeiLu et al (64) las patologías oftalmológicas más estudiadas en relación con IA son retinopatía diabética, glaucoma, degeneración macular asociada a la edad (DMAE) y cataratas. Sin embargo otras enfermedades oculares como queratocono, ojo seco y retinopatía del prematuro tienen menos incidencia en la literatura publicada.

En relación al tema central de esta tesis, Clarke et al (8) fueron los primeros en publicar un trabajo en el que mediante una red neuronal (RN) entrenada únicamente con 200 casos, compararon la predicción de la potencia de la lente intraocular con una fórmula teórica robusta y fiable, vigente incluso hoy en día: Holladay I. Los autores concluyeron que los resultados obtenidos con la RN disminuían significativamente la varianza del error, es decir, la RN de Clarke mejoraba la precisión conseguida por la fórmula teórica.

Por otra parte, los autores no describen en la metodología de su trabajo ni el tipo de RN utilizada, ni cuantas capas ocultas tiene, ni los pesos o *weights* en las conexiones de la red, ni los *inputs* utilizados, ni el ratio de división de la muestra, ni cómo entrenó la muestra y tampoco la capacidad de computación empleada para esa época, datos estos fundamentales para entender el funcionamiento del método. De hecho, la RN de Clarke no se tiene constancia de su uso en clínica, ni se han publicado en revistas JCR comparativas con otras fórmulas o métodos.

3.3.3 Aplicación de Inteligencia Artificial en cálculos biométricos

En una revisión bibliográfica sistemática hecha con motivo de esta tesis doctoral, es evidente la escasez de publicaciones en revistas *peer review* relacionadas con IA en cálculos biométricos. Desde que Clarke et al en 1997 publicaran el primer trabajo hasta la actualidad, únicamente se han encontrado seis artículos (65-70). No obstante, cuatro de dichos trabajos no son métodos de cálculo propiamente dichos, son algoritmos de IA soportados con *Redes Neuronales (RN)*, *Support Vector Machines (SVM)* y *Self-Organizing Maps (SOM)*, implementados sobre fórmulas ya instauradas (Hoffer Q, SRK/T, Holladay I,

etc.) para seleccionar la mejor. Los otros dos trabajos, Sramka et al (69) y Fernández-Álvarez et al (65), sí son métodos de cálculo propiamente dichos, ya que predicen la potencia de una lente intraocular.

En el primero de ellos se proponen dos métodos de IA basados en regresión lineal con SVM y *Multilayer Neural Network Ensemble model* (MNNE) respectivamente, con un dataset de 2.194 ojos para calcular la potencia de la lente que se tiene que implantar. Quizás un punto importante en su trabajo hubiera sido comparar los resultados de ambos modelos con las fórmulas y métodos más utilizados en clínica actualmente. Los autores hacen una comparación únicamente con Barrett Universal II con resultados extraídos de otro trabajo. No publican su fórmula en una web propia de acceso libre para dar a conocer el potencial reflejado en su trabajo.

En el segundo trabajo, Fernández-Álvarez et al proponen una RN multicapa tipo *Perceptrón* con dos capas ocultas y entrenada con 15.728 ojos, introduciendo el 95% de los ojos dentro del intervalo ± 0.50 D, un resultado excepcional comparado con lo que está publicado hoy en día. También se echa de menos en este trabajo una comparación con las fórmulas teóricas y basadas en IA de última generación. Tal vez los autores deberían haber utilizado los descriptores estadísticos aceptados en las comparativas de fórmulas y/o métodos: *Mean Predictive Error*, *Mean Absolute Error* y *Median Absolute Error*, para de este modo que otros investigadores puedan hacer comparaciones con sus resultados. Los autores tampoco publican su fórmula en una web libre para probar su potencial.

Por otro lado, hasta la fecha se cuenta con varias fórmulas o métodos de acceso vía web para calcular la potencia de la lente intraocular que tienen relación con IA, pero

sin embargo no están publicados en los circuitos convencionales de divulgación científica. Estas son por orden cronológico: FullMonte IOL, Hill-RBF, Ladas Super Formula AI (LSF-AI), Kane formula, Pearl-DGS y Hoffer-QST. Excepto la primera, las demás fórmulas y métodos de cálculo han aparecido en los últimos tres años. El problema es que no se han publicado las mecánicas de funcionamiento, ni las metodologías y principios en las que se basan, solo se disponen de estudios en los que sus autores se remiten a los resultados de comparación con otros métodos.

3.4 Lenguaje R, nomenclatura básica y abordaje general de problemas de ML

Este capítulo describe el origen, funcionamiento y sintaxis básica del lenguaje de programación con el que se desarrollarán los modelos de IA, así como un análisis descriptivo general para acometer cualquier problema de índole regresiva en ML.

R es un lenguaje de programación que nació a raíz de otro entorno primigenio denominado S (*Statistical*). Este último fue desarrollado en AT&T Bell Laboratories por John M. Chambers en la década de los 70 (71).

El lenguaje S estaba orientado a la estadística para mejorar la interactividad con el usuario además de facilitar las interpretaciones gráficas de análisis estadísticos complejos. En el año 1980 los laboratorios Bell lo hicieron público incluso con sus códigos fuentes, lo que provocó una evolución continua del entorno. En 1993, Robert Gentleman y Ross Ihaka, del departamento de estadística de la Universidad de Auckland en Nueva Zelanda, desarrollaron R (72) basándose en S.

Su desarrollo actual es responsabilidad del equipo central de desarrolladores R (*R Development Core Team*). La gran ventaja de este entorno de programación es el

concepto abierto y gratuito, que apoyado por una gran comunidad de usuarios permite un desarrollo continuo. Los usuarios pueden definir sus propias funciones e incluso pueden programar e implementar algoritmos computacionalmente exigentes y complejos en C, C++ o Fortran, generando bibliotecas que se cargan dinámicamente en R. Actualmente existen 17687 librerías de R (<https://cran.r-project.org/web/packages/>), consulta hecha en junio de 2021, aunque se actualizan con bastante frecuencia.

R es un lenguaje interpretado, como Java™, y no compilado como Basic®. Es decir, los comandos escritos en el teclado son ejecutados directamente sin la necesidad de construir un archivo ejecutable, lo que facilita mucho el trabajo con el análisis de datos complejos.

Su principal característica es que R utiliza una consola de comandos en lugar de una interfaz gráfica, lo que conlleva que el dominio de la sintaxis sea exigente y requiera mucha práctica. Existen diversas interfaces o IDE (Integrated Development Environment) que facilitan el trabajo con R: R Commander, JGR, Sage, Cantor, RStudio®, etc. En la presente tesis se utilizó RStudio® por conocimiento previo de la herramienta (Figura 9).

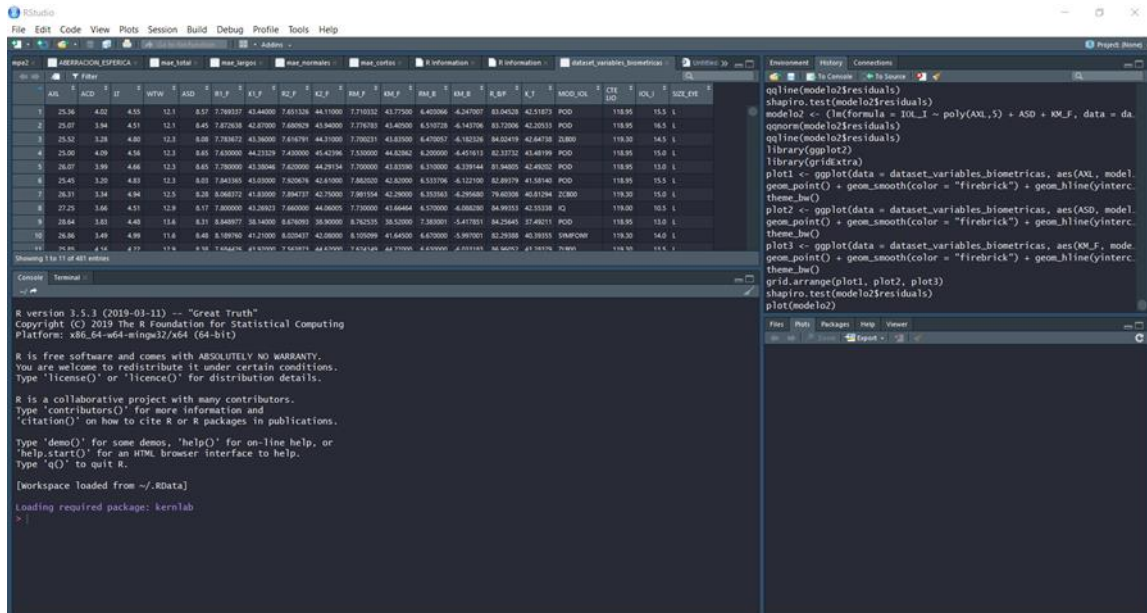


Figura 9. Interfaz de usuario de RStudio® versión 1.1.423 con sus cuatro espacios de trabajo por defecto. El panel superior izquierda es el editor. El panel superior derecha es el entorno de variables. El panel inferior izquierda es la consola de comandos. Panel inferior derecha es el panel de utilidades.

La nomenclatura y sintaxis del lenguaje no es más difícil que en otros entornos, pero tiene particularidades que deben comprenderse. En lo elemental, R está orientado a objetos y los más utilizados son vectores, matrices, listas, data frames y factores. Cada uno de ellos tiene unos atributos y las funciones genéricas ya definidas como plot, head o print por ejemplo, reconocen estos atributos y llevan a cabo distintos tipos de acciones en función del tipo de objeto.

Todos estos objetos anteriormente descritos son potencialmente contenedores de datos que pueden configurar grandes bases de datos. De hecho, la base de datos de la presente tesis está alojada en un *data frame* ya que contiene todo tipo de datos, variables numéricas y categóricas, aunque en alguno de los procesos matemáticos involucrados más adelante se convertirá en matriz al prescindir de las variables cualitativas.

Evidentemente no es objeto de esta tesis argumentar cada instrucción, librería, comando y función que posee R, pero al menos sí explicar brevemente cómo funciona, puesto que así es como se construyó la fórmula en esta tesis. Se ilustrará esto con un sencillo ejemplo con código generado e importado directamente de RStudio®, en el que se ejecutarán unos gráficos de histogramas de distribución normal (Figura 10):

```
## Genera una distribución de datos normal aleatoria contenidos en un vector llamado 'vector' gracias a la función "rnorm":
```

```
set.seed(123)
vector<-rnorm(10000,6,2)
```

```
## La función "Hist" genera un histograma de esos datos con unas características de color y textos:
```

```
hist(vector, breaks = 50, col = "red", freq = F, main = "Histograma", xlab = "Histograma con la función de densidad" )
```

```
## Con esto se dibuja la curva de densidad sobre el histograma previo:
```

```
dz<-seq(min(vector), max(vector), 0.001)
lines(dnorm(dz,6,2)~ dz, type="l", lwd=3)
```

```
## Ahora se utilizará una librería externa "ggplot2" que dibuja histogramas más llamativos:
```

```
library(ggplot2)
```

```
qplot (vector, geom = "histogram", binwidth = 0.8,
main = "Histograma",
xlab = "Histograma con la función de densidad",
fill=I("blue"),
col=I("red"),
alpha=I(.2))
```

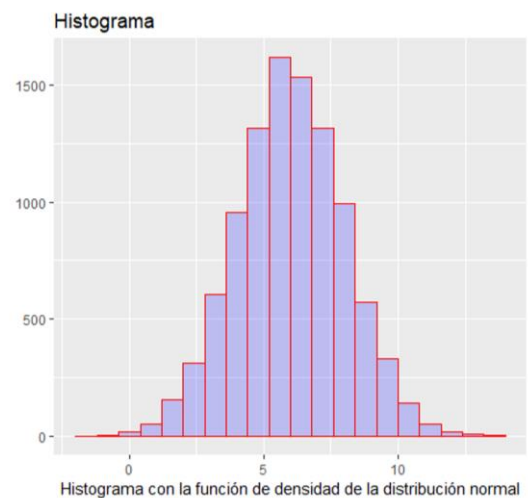
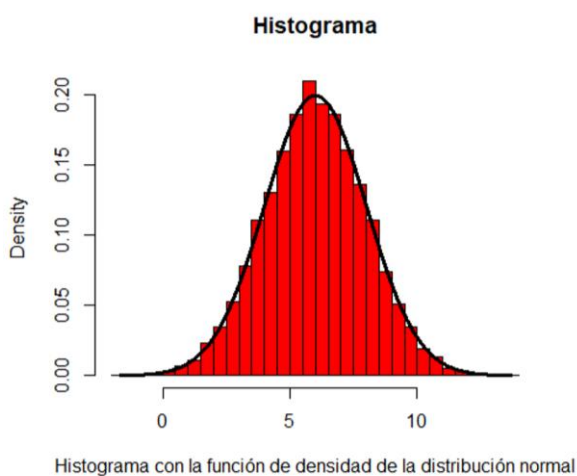


Figura 10. Histogramas de distribución normal generados con el código exportado de RStudio®.

3.4.1 Objetivos y selección del tipo de problema

Es fundamental clasificar las etapas por las que pasa cualquier problema de clasificación y/o regresión de ML, ya que esto permitirá comprender mejor el desarrollo posterior de los modelos propuestos en esta tesis.

En primer lugar se tendrá que definir el objetivo del problema mediante las siguientes preguntas ¿qué se pretende predecir? ¿de qué datos se dispone? y ¿qué calidad tienen?.

En segunda instancia, es necesario plantear el abordaje del problema que se quiere resolver, clasificatorio versus regresivo. En esencia, los problemas clasificatorios son aquellos en los que el resultado es una variable no numérica o clase, sin embargo en los problemas de regresión la variable resultado es un número. Un ejemplo típico de problema clasificatorio binario en ML es para saber si un tumor es maligno o no. Sin embargo un problema regresivo en ML sería por ejemplo estimar cuántos productos de un tipo concreto se van a vender en una determinada empresa.

Por tanto, para estimar la potencia de la lente intraocular que se quiere implantar en una cirugía de cataratas, el abordaje será puramente regresivo. Dentro de este ámbito regresivo, se plantearán posteriormente diferentes modelos matemáticos que se enfrenten al problema desde un punto de vista lineal y no lineal.

En ML la idea general no es encontrar un solo modelo, es construir muchos, ya que más de una técnica puede funcionar para el mismo tipo de situación. No obstante, en los últimos años está en boga combinar o ensamblar varios modelos para potenciar el rendimiento predictivo de los individuales (*ensemble methods*).

La clave para que la combinación entre modelos o *ensembling* consiga mejorar los resultados, reside en la diversidad de los modelos componentes. Si todos los modelos combinados fueran similares entre ellos, no podrán compensarse unos a otros. Por esta misma razón, se tienen que intentar combinar modelos que sean lo mejor posible a nivel individual y lo más diferentes entre ellos. Esta estrategia se basa en la asunción de que, distintos modelos entrenados independientemente, emplean distintos aspectos de los datos para realizar las predicciones, es decir, cada uno es capaz de identificar parte de la “verdad” pero no toda ella. Combinando la perspectiva de cada uno de ellos, se obtiene una descripción más detallada de la verdadera estructura subyacente en los datos.

3.4.2 Análisis exploratorio de los datos

El siguiente punto será hacer un análisis exploratorio de los datos. Esta etapa consiste en verificar los tipos de variables que componen el *dataset*, categóricas o factores y numéricas o cuantitativas. Es en esta fase en la que es básico conocer el número de observaciones disponibles y si todas ellas están completas, es decir, verificar si para cada observación se ha registrado el valor de cada una de las variables.

El análisis exploratorio de los datos suele caracterizarse por realizar en primera instancia una estadística descriptiva, al igual que un análisis gráfico para observar la distribución de los datos y la correlación entre variables si es que la hubiera.

Si dos variables numéricas están muy correlacionadas, añaden información redundante al modelo, por lo tanto, no conviene incorporar ambas. Si esto ocurre, se puede: excluir aquella que, acorde al criterio del analista, no está realmente asociada con

la variable respuesta; o combinarlas para recoger toda su información en una única nueva variable.

Si una variable tiene varianza igual o próxima a cero (su valor es el mismo o casi el mismo para todas las observaciones) añada al modelo más ruido que información, por lo que suele ser conveniente excluirla.

3.4.3 División de los datos

La siguiente etapa preceptiva en el modelaje de ML es la división de los datos para obtener dos conjuntos: el de entrenamiento (*training*) y el de validación (*test*). Dicha partición permite entrenar el modelo propuesto en la partición de entrenamiento, para posteriormente validarlo en los datos reservados como *test*.

Para evaluar la capacidad predictiva de un modelo se debe comprobar cómo de próximas son sus predicciones a los verdaderos valores de la variable respuesta. Para poder cuantificar de forma correcta este error, se necesita disponer de un conjunto de observaciones, de las que se conozca la variable respuesta, pero que el modelo no haya “visto” previamente, es decir, que no hayan participado en su ajuste, de lo contrario se incurrirá en el conocido *overfitting* o sobreajuste (73).

El tamaño adecuado de las particiones dependerá en gran medida de la cantidad de datos disponibles y la seguridad que se necesite en la estimación del error. No obstante un ratio 80-20 (80% de los datos para el *training* y el 20% par la validación) suele dar buenos resultados en la literatura revisada (74-76).

Es fundamental verificar que la distribución de la variable respuesta es similar en el conjunto de entrenamiento y en el de validación. Para ello, existen técnicas que garantizarán una distribución semejante. Esto asegura que las dos particiones sean similares en cuanto a la variable respuesta, sin embargo, no garantiza que ocurra lo mismo con los predictores. Por ejemplo, en un set de datos con 100 observaciones, un predictor binario que tenga 90 observaciones de un grupo y solo 10 de otro, tiene un alto riesgo de que, en alguna de las particiones, el grupo minoritario no tenga representantes. Si esto ocurre en el conjunto de entrenamiento, algunos modelos darán error al aplicarlos al conjunto de test, ya que no entenderán el valor que se le está pasando.

3.4.4 Preprocesado de los datos

Antes de plantear el primer modelo y entrenarlo, se hará un último paso previo denominado preproceso de datos en el que si fuera necesario se realizarán transformaciones sobre los datos para que los algoritmos de ML operen de forma homogénea. El preproceso bien podría incluir tres etapas básicas: exclusión de variables con varianza próxima a cero, reducción de dimensionalidad y estandarización de las variables numéricas.

No es conveniente incluir predictores que tengan una varianza próxima a cero, es decir, predictores que toman solo unos pocos valores, de los cuales, algunos aparecen con muy poca frecuencia. El problema con estos últimos es que pueden convertirse en predictores con varianza cero cuando se dividen las observaciones por validación cruzada o *bootstrap*.

Respecto a la reducción de dimensionalidad cuando se entrena un modelo, es importante incluir como predictores únicamente aquellas variables que están realmente relacionadas con la variable respuesta, ya que son estas las que contienen información útil para el modelo. Incluir un exceso de variables suele conllevar una reducción de la capacidad predictiva del modelo cuando se expone a nuevos datos. La selección de predictores puede suponer la diferencia entre un modelo normal y uno muy bueno.

Cuando los predictores son numéricos, la escala en la que se miden, así como la magnitud de su varianza, pueden influir en gran medida en el modelo. Algunos algoritmos de machine learning como SVM o RN son sensibles a esto, de forma que, si no se igualan de alguna forma los predictores, aquellos que se midan en una escala mayor o que tengan más varianza, dominarán el modelo aunque no sean los que más relación tienen con la variable respuesta. Existen principalmente dos estrategias para evitarlo: centrado y normalización.

3.4.5 Modelización en crudo

En esta etapa se puede modelizar en bruto, lo que significa que se construyen muchos modelos regresivos con los datos de la partición de entrenamiento (80% de la muestra) en crudo, sin hiperpametrizar. Dichos modelos se entrenan y se comparan en una fase inicial para poder evaluar su funcionalidad, descartándose los que rindan peor según las métricas seleccionadas previamente.

La finalidad última de un modelo es predecir la variable respuesta en observaciones futuras o en observaciones que el modelo no ha “visto” antes. Para conseguir una estimación más acertada, y antes de recurrir a la partición de datos llamada

test, se emplean estrategias de validación basadas en *resampling*. Habitualmente los más conocidos son: *cv*, *repeatedcv boot*, *boot632*, *optimism_boot*, *boot_all*, *LOOCV* y *LGOCV*. Cada uno funciona internamente de forma distinta, pero todos ellos se basan en la idea de ajustar y evaluar el modelo múltiples veces con distintos subconjuntos creados a partir de los datos de entrenamiento, obteniendo en cada repetición una estimación del error. El promedio de todas las estimaciones tiende a converger en el valor real del error de *test*.

A efectos prácticos, cuando se aplican métodos de *resampling* hay que tener en cuenta dos cosas: el coste computacional que implica ajustar múltiples veces un modelo, cada vez con un subconjunto de datos distinto, y la reproducibilidad en la creación de las particiones. En este sentido, paralelizar el proceso para que sea más rápido y establecer semillas para asegurar que cada partición pueda crearse de nuevo con exactamente las mismas observaciones, puede resultar de gran ayuda.

3.4.6 Optimización de hiperparámetros

Por último, una vez lanzados los modelos en bruto y habiendo descartado los menos potentes, se optimizan los hiperparámetros de los que mejor rendimiento predictivo han mostrado.

Cuando se habla de optimización de hiperparámetros o *parameter tuning*, se hace referencia a cómo podemos cambiar ciertos hiperparámetros del algoritmo matemático para optimizar o mejorar su rendimiento. Por tanto es necesario distinguir entre parámetros e hiperparámetros. Los parámetros son las variables que se estiman durante el proceso de entrenamiento con los conjuntos de datos, por lo que sus valores

no los indica manualmente el científico de datos, sino que son obtenidos. Sin embargo, los hiperparámetros de un modelo son los valores de las configuraciones utilizadas durante el proceso de entrenamiento. Son valores que generalmente no se obtienen de los datos, por lo que tienen que ser indicados por el científico de datos.

Los resultados de un modelo pueden depender en gran medida del valor que tomen sus hiperparámetros, sin embargo, no se puede conocer de antemano cuál es el adecuado. Aunque con la práctica los expertos en ML ganan intuición sobre qué valores pueden funcionar mejor en cada problema, no hay ninguna norma al respecto. La forma más común de encontrar los valores óptimos para los hiperparámetros es probando diferentes posibilidades basándose en ensayo y error.

Por ejemplo, el algoritmo SVM tiene un hiperparámetro llamado coste o *cost* en inglés. Este hiperparámetro controla la penalización que se aplica a las clasificaciones erróneas cuando se entrena el modelo. Si el valor de *C* es alto, el modelo resultante es más flexible y se ajusta mejor a las observaciones de entrenamiento, pero con el riesgo de sobreajuste.

Para llevar a cabo la hiperparametrización en ML existen técnicas que permiten obtener la combinación idónea de parámetros para cada modelo en concreto, téngase en cuenta que puede haber algoritmos que presenten muchos parámetros susceptibles de *tuning*. Habitualmente se utilizan métodos de validación cruzada mediante la implementación de *grids* de parámetros, pero suelen ser grabosos en tiempo. No obstante hay librerías programadas en R que permiten hacer esto sin el coste de tiempo computacional.

Se pueden distinguir dos tipos de procesamiento computacional, el secuencial y el paralelo. En el primero, la ejecución de un algoritmo de ML se llevará a cabo de una forma progresiva con una tarea después de otra, por orden, hasta la consecución final. Sin embargo en el procesamiento paralelo varios eventos ocurren y se desarrollan al mismo tiempo, con el beneficio temporal que eso otorga.

En el entrenamiento de muchos algoritmos de ML o en el proceso de optimización, la demanda computacional puede resultar exorbitante, lo que incurre en un gran problema no de precisión pero sí de eficiencia. Se podría disponer de un modelo predictivo con una altísima tasa de predictibilidad pero por el contrario unos tiempos computacionales durante su entrenamiento superiores a los 10 minutos, lo que favorece la inoperancia práctica del modelo. Así que se necesitarán modelos robustos eficientes y eficaces para una posible utilización futurible en los mercados clínicos reales.

3.4.7 Métricas de precisión en modelos de regresión

Para evaluar el potencial o rendimiento predictivo de un modelo ML se analizarán varias métricas que son las que se utilizarán en esta investigación.

Considerando dos series de datos de variables continuas, unos calculados y otros observados, relativos a un mismo fenómeno, existen una serie de métricas que sirven para cuantificar la precisión de una técnica de predicción comparando por ejemplo los valores predichos frente a los observados.

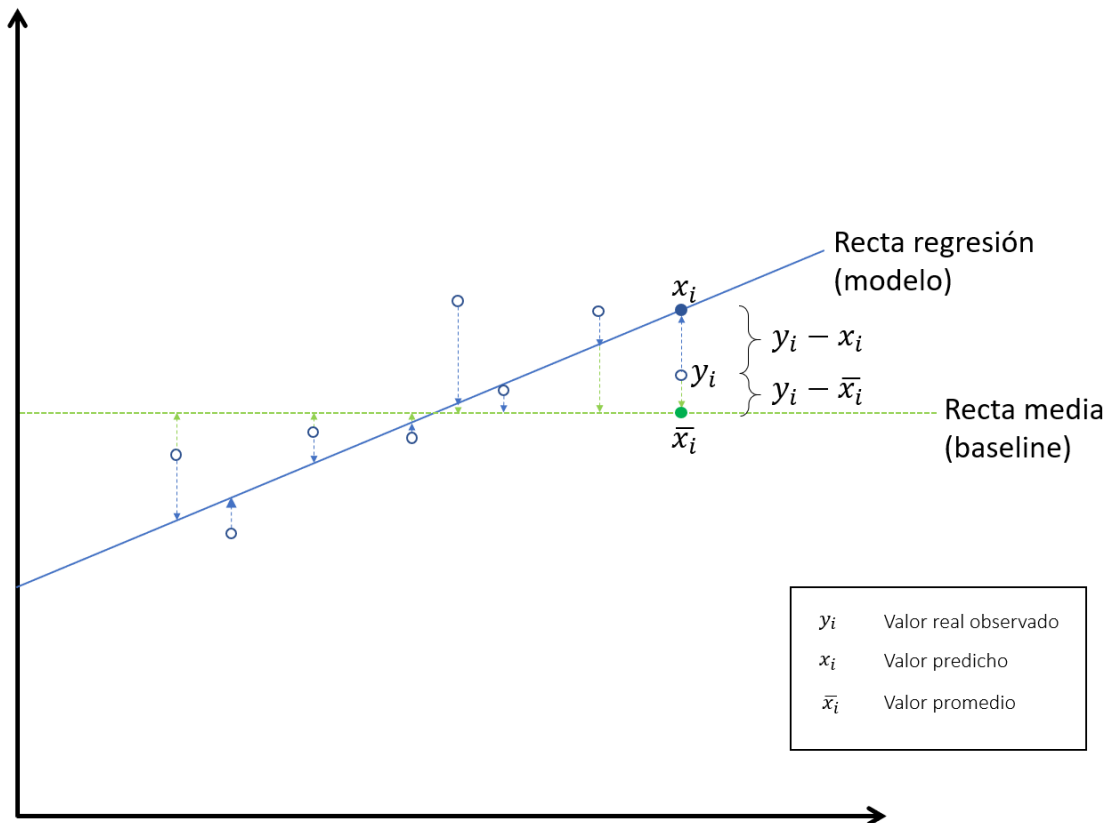


Figura 11. Datos observados y predichos en un modelo de regresión.

- **Mean Square Error (MSE)** o error cuadrático medio que se corresponde a la media del cuadrado de la diferencia entre el valor observado y el predicho por el modelo (Figura 11).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2 \quad (19)$$

- **Mean Absolute Error (MAE)** o error medio absoluto es la media en valor absoluto de la diferencia entre el valor real y el valor predicho.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - x_i| \quad (20)$$

Donde y_i es el valor real observado y x_i es el valor predicho por el modelo.

- **Root Mean Square Error (RMSE)** o la raíz del error cuadrático medio representa la raíz cuadrada del cuadrado de la diferencia entre el valor observado y el predicho.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2} \quad (21)$$

- Coeficiente de determinación R^2 y R^2 ajustado conocidos también como bondad del ajuste, son estadísticos que representan qué porcentaje de variación de la variable dependiente es explicado colectivamente por todas las variables independientes. Se aconseja utilizar mejor R^2 ajustado por no ser sensible al sobreajuste.

$$R^2 = 1 - \frac{MSE_{modelo}}{MSE_{baseline}} \quad (22)$$

Donde MSE_{modelo} es el error cuadrático medio calculado como se expresó al comienzo del apartado (Figura 11) y $MSE_{baseline}$ es el error cuadrático medio calculado en vez de con la diferencia entre el valor real y el predicho, con la diferencia entre el valor real y el promedio de los residuos (Figura 11).

$$R_{adj}^2 = 1 - \frac{(n-1)(1-R^2)}{(n-k-1)} \quad (23)$$

Donde n es el número de observaciones de la muestra, k es el número de variables independientes y R^2 es el coeficiente de determinación.

A lo largo de la presente tesis, como métricas de selección para los modelos ML más potentes y de mejor rendimiento predictivo, se utilizarán estas tres métricas: RMSE, MAE y R^2 .

MAE como métrica de rendimiento predictivo para modelos ML, es el mismo tipo de error absoluto utilizado en los estudios de comparación entre fórmulas o métodos de cálculo. No obstante, el MAE utilizado para evaluar modelos de ML que predicen la potencia de la LIO, es el que corresponde a la diferencia entre la lente emetropizante predicha y la lente emetropizante real implantada, por tanto el MAE es un valor de error que tiene sentido en el plano capsular.

Sin embargo, si la pretensión fuera comparar el rendimiento de un modelo ML con una fórmula de uso clínico habitual, el MAE deberá convertirse al plano corneal. Para ello, se deberá aplicar una fórmula de conversión distométrica teniendo en cuenta variables como la distancia entre lente intraocular y córnea, así como la potencia corneal, tal y como se verá más adelante.

De este modo un MAE de 0.35 D obtenido por un modelo regresivo en el ámbito del ML cuando el modelo predice la potencia de la lente intraocular, a efectos prácticos, es un muy buen rendimiento, porque equivaldría por distometría aproximadamente a un

MAE de 0.25 D en el plano corneal, lo que implica en el ámbito comparativo de fórmulas una valor óptimo del orden del cuarto de dioptría.

3.4.8 Interpretabilidad de modelos complejos

En una fórmula de cálculo biométrico teórica es relativamente sencillo la interpretación óptica-física de su funcionamiento interno, siempre y cuando la fórmula se conozca, ya que se pueden establecer relaciones causales para evaluar el comportamiento entre las variables predictoras y la variable respuesta. Por ejemplo, en la fórmula SRK/T es sabido que una combinación entre una queratometría por encima de 45.00 D y una longitud axial por encima de 28.00 mm produce una sobreestimación de la posición efectiva de la lente (77).

Sin embargo, en el modelaje de ML la situación se complica, ya que en muchas ocasiones aun obteniendo muy buenas predicciones con modelos de alta complejidad, estos son tratados como una *black box* al no aportar de forma directa una explicación de los mecanismos internos empleados o de cómo obtienen sus predicciones. Por tanto, son modelos que no pueden entenderse a partir de los parámetros que aportan, es fácil perderse en la lógica introspectiva de estos modelos, de ahí la terminología de “caja negra”, ya que preprocesan los datos, abarcan cantidades ingentes de combinaciones entre variables predictoras y emplean metodologías de validación cruzada con velocidades imponentes de computación.

Básicamente, la interpretación de modelos implementados con ML puede atacarse desde dos perspectivas: la intrínseca y el *post hoc*, la primera se refiere a la

utilización de modelos fácilmente interpretables que no estarían en la categoría de “caja negra”. Respecto al segundo abordaje, la interpretabilidad *post hoc* hace referencia a si se aplican métodos que analizan el modelo después de su entrenamiento.

Generalmente, el abordaje para dar interpretabilidad a los modelos ML se hace en base a cuatro métodos, que son los más usados en minería de datos. Estos se conocen como métodos agnósticos al modelo, porque pueden ser utilizados independientemente del mismo, son los siguientes:

- Importancia de Variables
- Gráfico de Dependencia Parcial
- Gráfico de Esperanza Condicional individual
- Interacción entre variables

Importancia de las variables

También conocida como VI (*Variables Importance*). La importancia de las variables corresponderá al aumento en el error de predicción del modelo después de permutar los valores de la variable, rompiendo así la asociación entre la variable y su verdadero resultado. El concepto es simple: ¿cómo funcionará el modelo si se mezclan aleatoriamente los valores de una variable? en otras palabras, si se inutiliza una variable para el modelo, cuál es el impacto en el rendimiento predictivo teniendo en cuenta que dicho modelo sí se ha entrenado con esa variable.

Gráfico de Dependencia Parcial

También conocidos con gráficos PDP (*Partial Dependence Plot*), muestran cómo varían las predicciones promedio a medida que se modifica el valor marginalmente de un

regresor de interés, manteniendo constantes el resto, mientras se toma en cuenta la no linealidad y se promedian los efectos de todas las demás variables de entrada.

Se puede representar esta misma variación del promedio de predicción teniendo en cuenta las variaciones marginales de dos regresores mediante mapas de calor (*Heatmaps*).

Gráfico de Esperanza Condicional individual

También conocidos como gráficos ICE (*Individual Conditional Expectation*). Un gráfico ICE visualiza la dependencia de la predicción en un predictor para cada observación por separado, lo que da como resultado una línea por observación, en comparación con una línea general en los gráficos de dependencia parcial. Se trata de un equivalente al PDP para datos individuales, ya que un PDP representa el promedio de las líneas de un diagrama ICE. Los valores para una línea (una observación) se pueden calcular manteniendo todos los otros predictores iguales, creando variantes de esta observación reemplazando el valor del predictor y obteniendo predicciones con el modelo para estas observaciones recién creadas.

Otro aspecto importante es que, a diferencia de los PDP, los gráficos ICE proporcionan más información respecto a posibles interacciones en cuanto a la evolución de la predicción.

Hay un gráfico particular ICE llamado d-ICE (*derivative-ICE*). Este tipo de gráficos se construyen mediante las derivadas parciales individuales de la función de predicción con respecto a un predictor de interés. Estas derivadas pueden indicar si ocurren cambios y en qué dirección. Los d-ICE hacen que sea visualmente más fácil detectar la heterogeneidad debido a la aparición de la desviación estándar de las derivadas parciales.

Si no existiera ninguna interacción entre las variables, todas las curvas generadas en un gráfico d-ICE son aproximadamente paralelas, sus derivadas aproximadamente iguales y, por lo tanto, el gráfico de derivadas mostraría una única recta. Si existen interacciones, entonces la representación de las derivadas parciales es heterogénea.

Interacción entre variables

Es posible obtener medidas de interacción de las variables disponibles en el modelo, representada en gráficos como valor de cuánta varianza de la función predictora es explicada por la interacción.

El rango de la medida de interacción se encuentra entre 0 y 1, siendo 0 la inexistencia de interacción y 1 el 100% de varianza debida a interacciones.

4 ESTADO DEL ARTE

4. ESTADO DEL ARTE

En este capítulo se hará un repaso de todas las fórmulas y métodos de cálculo de la potencia de la lente intraocular consolidados, accesibles y basados en IA, que actualmente son seis: FullMonte IOL, Hill-RBF, Ladas superformula-AI, Kane formula, PEARL-DGS y Hoffer-QST.

FullMonte IOL (www.fullmonteiol.com):

Es de acceso privado mediante una cuota mensual aunque a la finalización de esta tesis la web no está activa (Figura 12). Es un método híbrido basado en IA combinado con una parte teórica, ya que calcula la potencia de la lente emetropizante mediante la fórmula SRK/T incluídas las optimizaciones por axialidad para determina el error probabilístico que comete dicha fórmula en base a un tamaño poblacional.

Los autores de Fullmonte IOL son, el pionero que publicó en 1997 el primer trabajo con RN, Gerald P. Clarke y Adam Kapelner, matemático precursor del algoritmo en el que se basa la metodología.

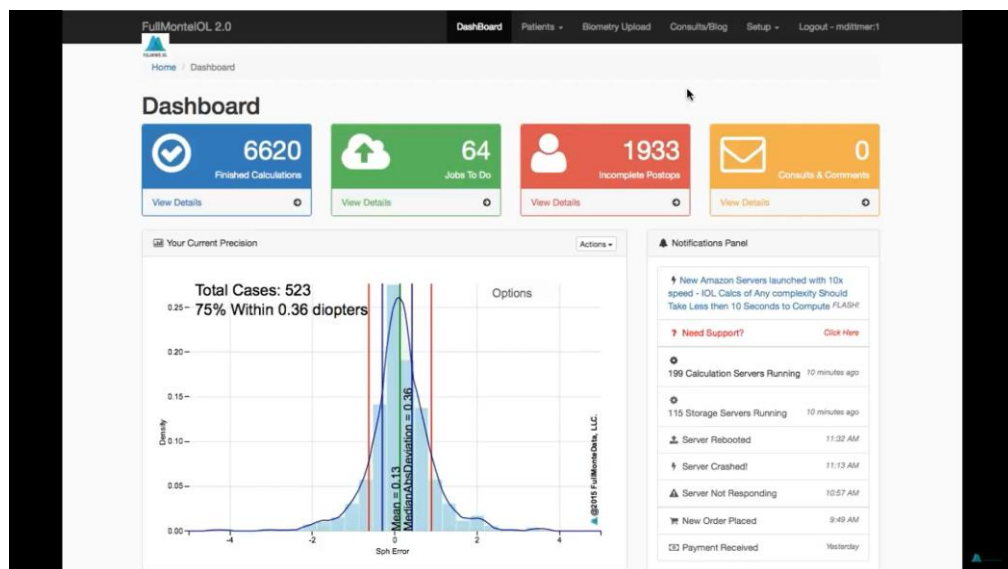


Figura 12. Pantalla principal de FullmonteIOL 2.0. (Imagen extraída de <https://www.youtube.com/watch?v=-p6f8XKJXBY&list=PLm85NWhdte2j1eSzDOLnoVpwW0A2kT3Rw&index=12>)

Su sistema se basa en un algoritmo de ML llamado Bayesian Additive Regression Trees (BART) (78), optimizado con técnicas de *Markov Chain Monte Carlo* (MCMC) (79). Esto significa que es un método regresivo que construye modelos que predicen la probabilidad de posibles resultados, en base al Teorema de Bayes, también conocido como teorema de la probabilidad condicionada. Los métodos MCMC son técnicas de simulación para generar muestras de distribuciones y estimar cantidades de interés a posteriori.

Para la predicción de la potencia de la LIO a implantar (Figura 13) la estrategia de FullMonte IOL es calcular la potencia del implante mediante SRK/T con el ajuste de Wang-Koch (80) para ojos axialmente por encima de 25.00 mm, para después, a través del modelo probabilístico corregir el error entre el cálculo teórico de la potencia de la lente con SRK/T y la verdadera potencia de LIO después de la operación. Obsérvese que no se está prediciendo la posición efectiva de la lente (81), como hacen todas las fórmulas teóricas, sino el error u *offset* de potencia de la lente respecto a SRK/T, operando de igual modo con el equivalente esférico.

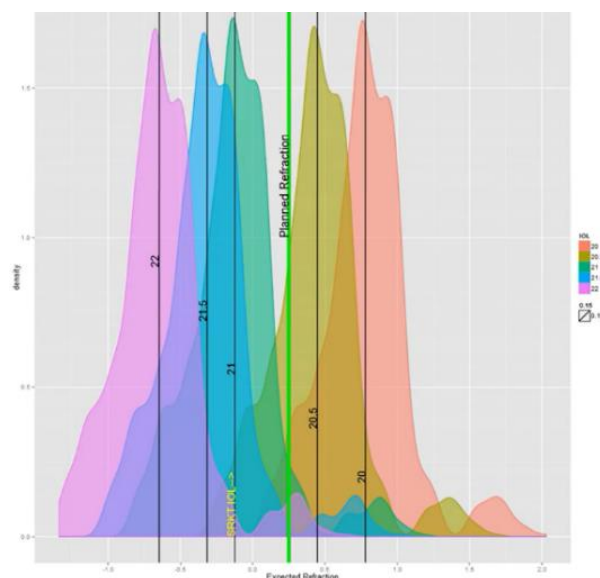


Figura 13. Sistema gráfico de resultados en FullMonte IOL mediante curvas de densidad de probabilidad. (Imagen extraída de <https://www.youtube.com/watch?v=T9UwecAw75o>).

Hill-RBF (www.rbfcalsculator.com):

Es de acceso libre a través de su web sin necesidad de registro (Figura 14). Es otro método basado íntegramente en IA y promulgado por el Dr. Warren E. Hill, que en el momento de finalización de esta tesis está actualizado a la versión 3.0.

Target Refr.[D] 0.00

Target Refr.[D] 0.00

1 ZEISS IOLMASTER 700

AL 23 mm K1 44 0 D[°]

CCT 555 μm K2 44 90.00 D[°]

ACD 3 mm n 1.3375

LT 4 mm WTW 12 mm

2 Biconvex 1:1

Manufacturer

Model

A-Constant 119

Hill - RBF

IOL[D]	REFR[D]
21.50	0.57
22.00	0.22
22.50	-0.13
23.00	-0.48
23.50	-0.84

Constants: A=119.00
IOL Power @ Emmetropia [D]: 22.32

Please select used measuring device...

AL mm K1 D[°]

CCT μm K2 D[°]

ACD mm n 1.3375

LT mm WTW mm

2 Biconvex 1:1

Manufacturer

Model

A-Constant

Hill - RBF

Figura 14. Calculador Hill-RBF v3.0.

La versión 3.0 se basa en un algoritmo no supervisado entrenado con más de 30.000 ojos de ML (Figura 15) llamado *Self-Organizing Map* (SOM) (82) que corresponde a un subtipo de RN, también conocido como *mapas de Kohonen* (83). A partir de un proceso iterativo de comparación con un conjunto de datos reales y cambios para aproximarse a los mismos, crea un modelo de esos mismos datos que puede servir para

agruparlos por criterios de similitud (84). Por tanto, se puede decir que SOM es un clasificador de observaciones. Si se busca la potencia de una lente intraocular para unas características biométricas concretas, SOM habrá aprendido previamente la lente que le corresponde a ese ojo y la clasificará en base a ese entrenamiento de reconocimiento de patrones.

Ongoing updates to Hill-RBF	
Upcoming version 3	
March 2018 (Version 2)	12,419 eyes measured by optical biometry. Improved accuracy for high axial hyperopes to +30.00 D. Addition of meniscus design IOLs down to -5.00 D. Expanded normal eye boundary model.
December 2020 (Version 3)	>30,000 eyes & further refinement of the AI database. Improved accuracy for high axial hyperopes to +34.00 D. Improved accuracy for odd anterior segment eyes: <ul style="list-style-type: none"> Central corneal thickness. White-to-white. Lens thickness. Gender. Additional entries Greatly expanded normal eye boundary model.

Figura 15. Diapositiva del Webinar impartido por W.Hill (07-01-2021) titulado "IOL Power Selection Driven by Artificial Intelligence". (Imagen extraída de <https://www.youtube.com/watch?v=EoZRnSHIZs4>)

El objetivo del aprendizaje en los SOM es provocar que diferentes partes de la red respondan de manera similar a ciertos patrones de entrada (Figura 16). Los SOM difieren de otras RN, ya que aplican el aprendizaje competitivo en oposición al aprendizaje de corrección de errores, además de que usan una función de vecindario para preservar las propiedades topológicas del espacio de entrada.

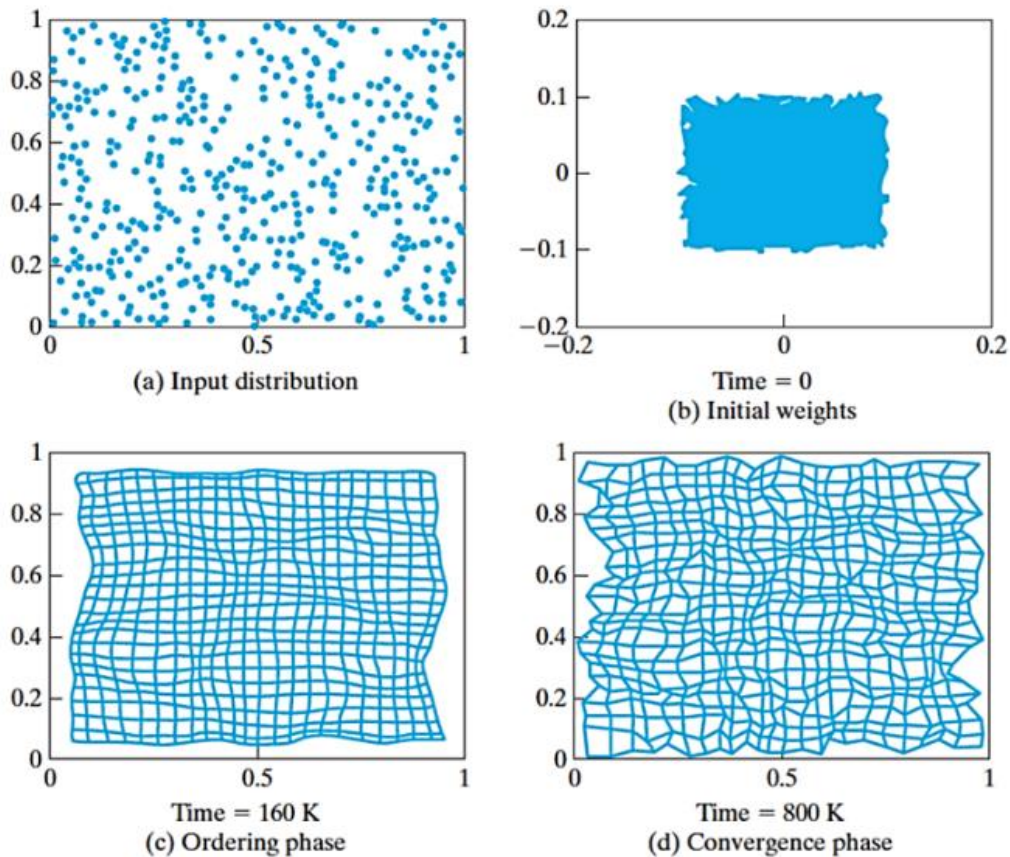


Figura 16. A) Distribución de datos de entrada. B) En el primer momento las funciones de vecindario no se han aplicado (Time indica el número de iteraciones). C) Tras el entrenamiento inicial de la red el patrón de va dibujando. D) Concluido el entrenamiento o fase de convergencia el mapa final habría reconocido el patrón de entrada. (Imagen extraída del trabajo de Haykin (84))

En la Figura 16 se puede observar como al aumentar los ciclos en el proceso de aprendizaje o entrenamiento, el mapa final se autoorganiza pareciéndose mucho a la distribución de datos iniciales, lo que implica que SOM reconoció el patrón en el que estaban distribuidos las variables de entrada.

Uno de los problemas a los que se enfrenta inicialmente cualquier modelo de ML independientemente del algoritmo o algoritmos que lo compongan, es la reducción de la dimensionalidad, o lo que es lo mismo, la selección de las variables predictoras adecuadas para maximizar la capacidad predictora del modelo sin sacrificar la información más importante del conjunto de datos y sin que por ello se resienta el

desempeño final del modelo. Existen muchas técnicas para la selección de características, entre otras: estrategias iterativas *Subset Selection*, *Random Forest*, *Ridge Regression*, *Lasso*, *PCA* y *PLS*, pero en el caso de Hill-RBF se utilizó un algoritmo genético inspirado en la teoría evolutiva de selección natural (82), aunque no hay referencias de cuál.

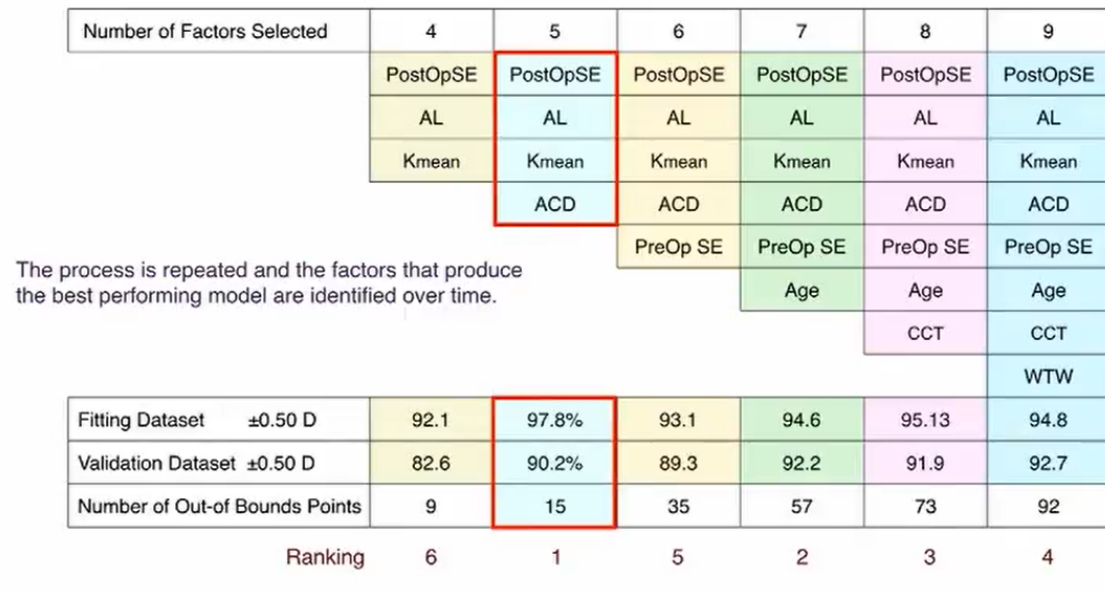


Figura 17. Algoritmo genético usado en el diseño de Hill-RBF para selección de características. (Imagen extraída de una conferencia de W. Hill en la ASCRS (82))

El modelo final de Hill-RBF (Figura 17) que mejor rendía es el que solo tenía cuatro inputs: AXL, ACD, Kmean y equivalente esférico postoperatorio y que además en la fase de validación introducía hasta un 90.20% de ojos en el intervalo refractivo postoperatorio de ± 0.50 D.

En cuanto a las siglas RBF hacen referencia a *Radial Basis Function* (85), que es una función matemática usada para construir aproximaciones de funciones. La función RBF (Figura 18) es empleada en el modelo como función de activación de la RN SOM, aunque no hay documentación de qué función es aplicada a este nivel ya que podrían ser

muchas las opciones (multicuadráticas, sigmoideas, etc). La función de activación en una RN es la que se encarga en el nodo o sinapsis neuronal de producir una señal de salida en función de la de entrada.

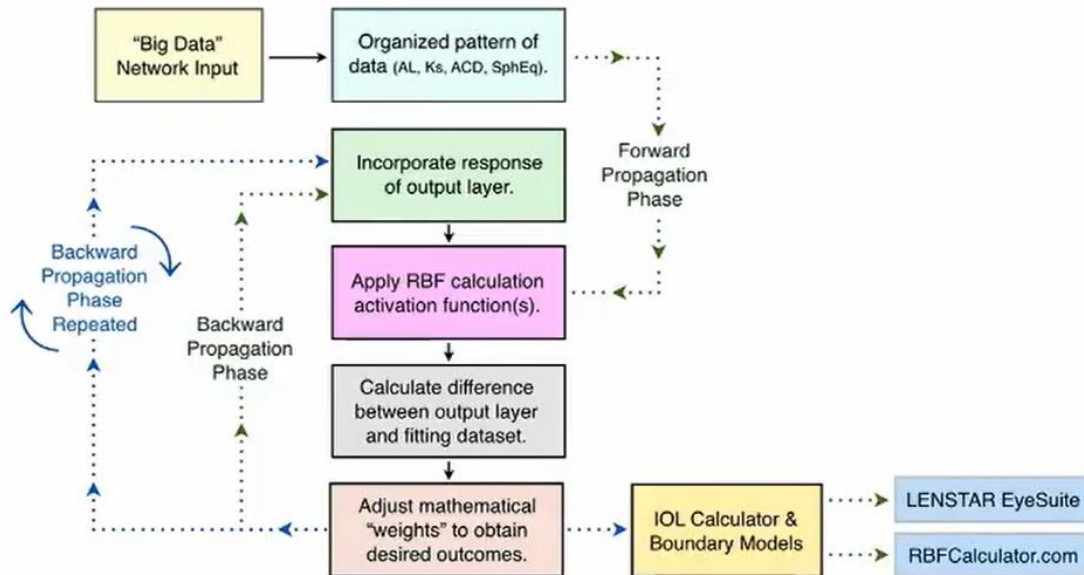


Figura 18. Esquema de funcionamiento de la red neuronal en la que está basada Hill-RBF. (Imagen extraída de una conferencia de W. Hill en la ASCRS (82))

LSF-AI (Ladas Super Formula Artificial-Intelligence) (www.iolcalc.com):

Es de acceso libre a través de su portal con registro previo (Figura 19). Es una fórmula que combina una parte teórica (*versión 1*) con IA (*versión 2*) fundamentada en Redes Neuronales artificiales multicapa (86). Los autores de esta fórmula son John G. Ladas, Aazim A. Siddiqui, Uday Devgan y Albert S. Jun.

Figura 19. Pantalla de cálculo de la fórmula de Ladas.

La parte teórica de la fórmula (87) se basa en el cálculo de la potencia de la lente intraocular con cinco fórmulas: Hoffer Q, Holladay I, Haigis, SRK/T y Holladay I con ajuste de Wang-Koch (88). Partiendo de la base que cada una de estas fórmulas funciona mejor en unas características anatómicas que otras, como por ejemplo Hoffer Q en ojos axialmente cortos (51) o Holladay I con ajuste Wang-Koch en ojos largos (89), los autores defienden así la cobertura total del espectro axial y queratométrico ocular usando estas cinco fórmulas. Las fórmulas tienen zonas de convergencia en base a las características biométricas del ojo estudiado y otras zonas de discrepancia. Precisamente por esto los autores incluyeron y excluyeron partes de cada fórmula en función de las características anatómicas y los resultados publicados en la literatura publicada revisada por pares.

Además, construyeron un modelo de análisis gráfico al que llaman *super-surface* incluido en la fórmula (Figura 20), que es la representación isométrica de un plano 3D dependiente de la longitud axial, queratometría media y potencia de la LIO, de las confluencias de las cinco fórmulas estudiadas. La super-superficie mostrada es una

composición de las cinco fórmulas, representadas en cinco capas distintas superpuestas para generar un plano único.

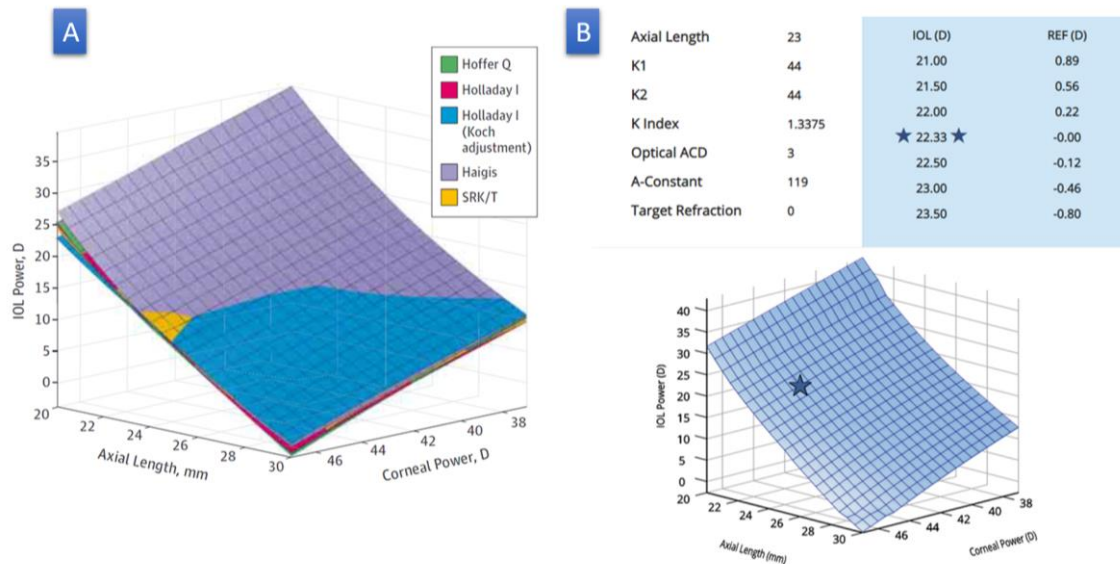


Figura 20. A) Superposición de “super-superficies” correspondientes a las distintas fórmulas que integran LSF para conformar un único plano. B) Cálculo realizado con Ladas Super Formula y la inclusión de la super-superficie que representa isométricamente la potencia de la lente, la longitud axial y la potencia corneal. La estrella marca la potencia de la lente que hay que implantar y la zona de influencia en el gráfico de la fórmula que subyace en el cálculo.

Respecto a la implementación de IA en la versión 2.0 de LSF-AI, el algoritmo de ML utilizado en este caso es una RN multicapa compuesta por cuatro capas: la de entrada o de inputs biométricos (AXL, ACD y K) , dos ocultas (*hidden layer*) de proceso y la capa de salida o de resultados (Figura 21).

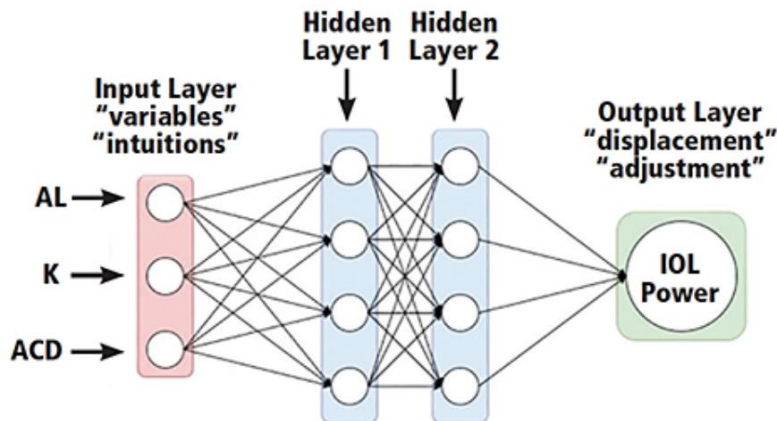


Figura 21. Esquema simple de la Red Neuronal en la que se basa parte de la fórmula de Ladas. (Imagen extraída de (86))

En este sentido, no existe documentación publicada de cómo se entrena la RN, ni de qué funciones de activación se emplean y tampoco de otras características relativas al desempeño de la red neuronal artificial. No obstante, el enfoque de los autores para la parte basada en IA es el mismo que planteó Clarke con su Fullmonte IOL, predecir la diferencia o el desplazamiento entre el resultado obtenido por la fórmula teórica ya existente (LSF) y la ideal para un ojo dado. Según los autores, estos ajustes no son aleatorios y generalmente ocurren en el orden de 0.20 D (86).

Kane Formula (www.iolformula.com):

Es de acceso libre y gratuito a través de la web sin registro previo (Figura 22). Desarrollada en 2017 por Jack X. Kane, se trata de una fórmula teórica combinada con algoritmos de IA sobre una muestra de 30.000 ojos. No existe documentación acerca del tipo de algoritmos de ML empleados para el refinamiento de la potencia obtenida con la parte teórica de la fórmula. Hay alguna publicación en la que se compara con otras fórmulas (90, 91).

Figura 22. Pantalla de cálculo de la fórmula de Kane.

PEARL-DGS (*Precision Enhancement using ARTificial intelligence and Linearization of the outputs*) (www.iolsover.com):

Es de acceso libre con registro previo en su página web (Figura 23). Combina una parte teórica basada en lente gruesa con IA. Sus autores son G. Debellemanière, D. Gatinel y A. Saad.

Al igual que la fórmula de Kane, tampoco se dispone de información técnica sobre qué o cuáles algoritmos de ML utiliza, ni en qué punto confluyen la parte teórica con la IA y tampoco cómo la formula afronta el cálculo de la potencia de la lente intraocular. Hasta la fecha no existen publicaciones *peer review* en las que se comparen otras fórmulas con esta, sin embargo sí existe un trabajo preliminar de los autores sobre 400 ojos presentado en la ASCRS de San Diego en 2019 (Figura 24), reportando como promedio de error absoluto 0.23 ± 0.20 D (0.17) con un error máximo de 1.08 D.

Pearl DGS Calculator® v2.0

This formula performs best for the PhysiOL Finevision IOL. This calculator is for research purposes only and should not be used to make clinical decisions. Meniscal IOL handling will be added soon.

If you want to evaluate the formula for research purposes, please do not hesitate to write us. We will perform the calculations for you in a blind manner. You will be asked to send the biometric measurements, the IOL model and power for each eye, and the mean postoperative outcome for the entire set (in order to adjust the formula).

The second eye enhancement function and post-refractive surgery function are currently being enhanced.

Right side

Vitreous type Vitreous	Vitreous Refractive Index 1.336	IOL A-constant 119	Keratometric Index 1.3375
K1 (flat) D	R1 Computed from K1 mm	K2 (steep) D	R2 Computed from K2 mm
Post. corneal radius 1 mm, optional		Post. corneal radius 2 mm, optional	

Figura 23. Pantalla principal de la fórmula PEARL-DGS v2.0.

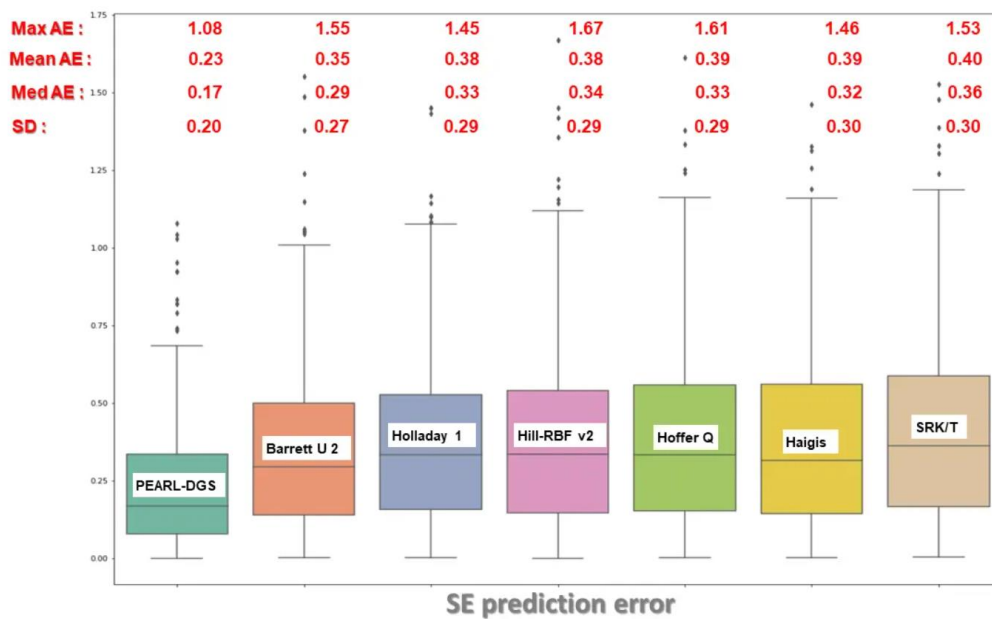


Figura 24. Boxplot de errores predictivos refractivos comparando PEARL-DGS con otras fórmulas. (Imagen extraída del blog del Dr. D. Gatinel: <https://www.gatinel.com/recherche-formation/biometrie-oculaire-calcul-dimplant/pearl-dgs-formula-for-iol-power-calculation/>).

Hoffer QST (<https://hofferqst.com/>):

La más reciente de todas, es de acceso libre sin registro previo en su web (Figura 25). Sus autores son K. Hoffer, G. Savini y L. Taroni. Es una fórmula híbrida que combina IA con la fórmula de vergencia clásica de Hoffer, la conocida Hoffer Q.

Los autores, siendo conscientes de las limitaciones de Hoffer Q, tiende a sobreestimar la potencia de la LIO en ojos con cámaras anteriores poco profundas y a subestimarla en ojos con cámaras anteriores profundas, han implementado ciertas mejoras basadas en el ajuste retrospectivo de la ELP sobre una muestra de 537 ojos implantados con la misma lente monofocal (92).

Utilizaron modelos basados en ML, pero sin concretar cómo y con qué algoritmos exactamente.

Figura 25. Pantalla principal del calculador Hoffer-QST.

El autor de esta tesis se puso en contacto mediante correo electrónico en varias ocasiones con el Dr. D. Gatinel, con el Dr. J. Kane y con el Dr. K. Hoffer para obtener más

información del funcionamiento de sus respectivas fórmulas, pero lamentablemente no se consiguió resultado.

5 DISEÑO DEL ESTUDIO

5.1 Diseño del estudio y flujo de trabajo

El diseño del estudio que se siguió en la presente tesis para la consecución de la fórmula basada en IA y que se explicará con detalle en el capítulo de metodología, se dividió en seis fases claramente diferenciadas (Figura 26). Las cuatro primeras son fases de construcción de la fórmula dentro de la metodología y las dos últimas son etapas de comprobación englobadas en los resultados de esta tesis:

- *Etapa I: Recogida, análisis y preparación de los datos.*
- *Etapa II: Equivalencias de potencia entre planos.*
- *Etapa III: Implementación, selección y optimización de modelos ML.*
- *Etapa IV: Conversión del equivalente esférico capsular a plano corneal.*
- *Etapa V: Validación del modelo resultante sobre la partición test.*
- *Etapa VI: Comparación del modelo con otras fórmulas y métodos de cálculo.*

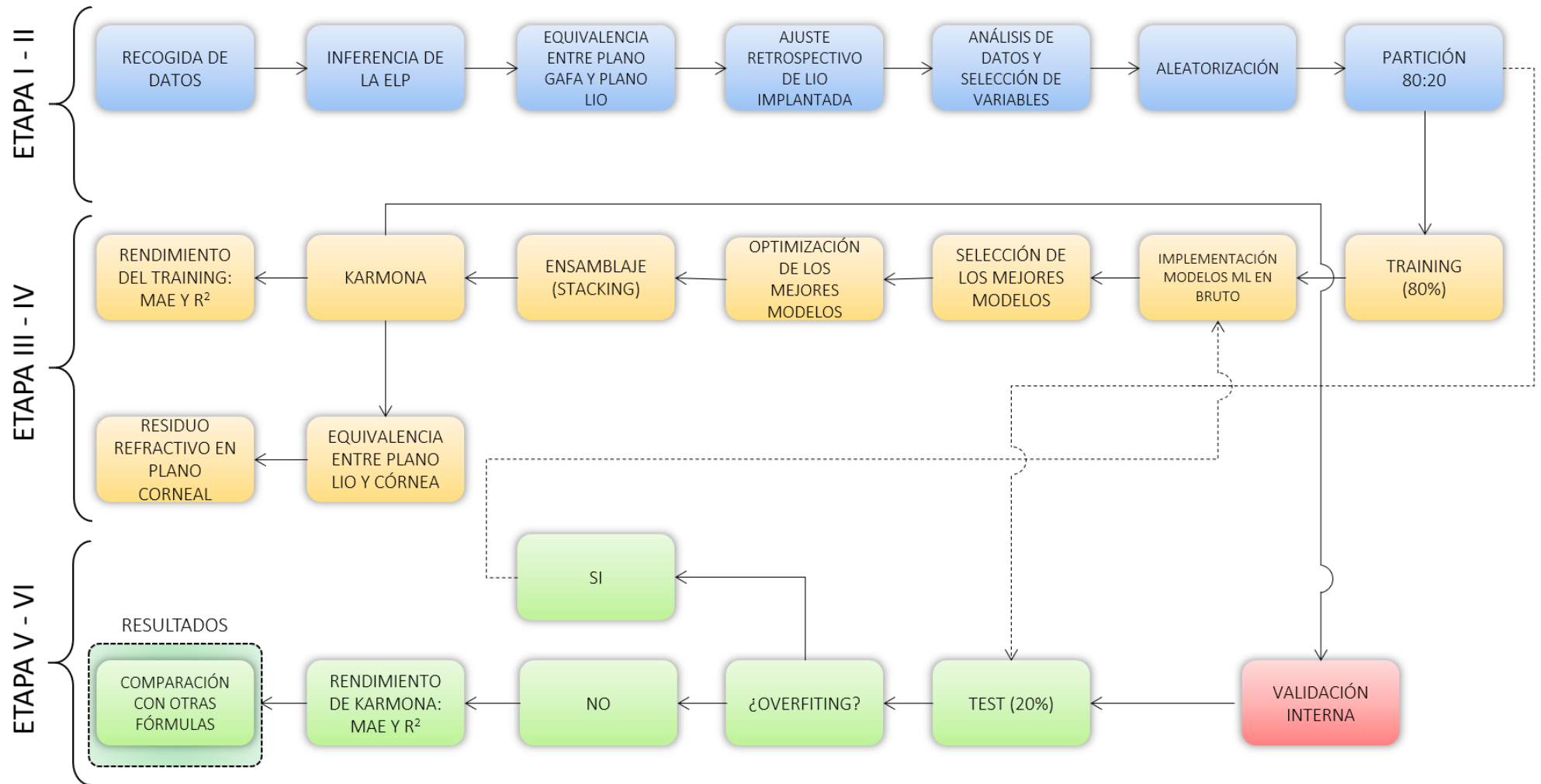


Figura 26. Diagrama de flujo para la tesis.

El flujo de trabajo que se establecerá es el siguiente:

- **Etapa I:** es una fase previa en la que se reclutará toda la información. Se buscarán correlaciones entre variables y se seleccionarán las características finales que conformarán la base de datos final.

- **Etapa II:** esta fase permitirá calcular las equivalencias distométricas entre los distintos planos ópticos de fuera hacia adentro: gafa-córnea-LIO. De esta forma se podrá determinar retrospectivamente la variable respuesta para el modelaje.

En esta fase se distinguen los siguientes puntos:

- En primera instancia se inferirá la ELP necesaria para determinar la variable respuesta descrita en el siguiente punto de esta fase óptica.
- En segundo lugar, habrá que calcular la variable respuesta de la base de datos, y para ello se dará respuesta a la siguiente pregunta ¿qué lente se tendría que haber puesto para dejar el ojo sin graduación postoperatoria? Sobre el *dataset* y desde la retrospectiva, se corregirá el error cometido por la lente intraocular implantada teniendo en cuenta el equivalente esférico postoperatorio refraccionado a 12.00 mm y la ELP calculada en el primer punto.
- Se transformarán los datos, y por último se dividirá aleatoriamente la muestra siguiendo la razón 80/20, para utilizar el 80% de los datos en el entrenamiento de los modelos propuestos y el 20% para validarlos.

- **Etapa III:** esta fase es matemática y engloba todos los procesos relacionados con la IA (modelaje y entrenamiento). Determina la potencia refractiva de la lente intraocular. El flujo de trabajo dentro de esta fase será el siguiente:
 - Se implementarán multitud de modelos ML regresivos sin ninguna optimización, en bruto, para comprobar su rendimiento. Este punto es importante, ya que se probarán hasta 100 estrategias distintas, con diferentes abordajes como por ejemplo el conocido algoritmo XGBOOST u otros como SVM, BART, RRNN, etc.
 - Se seleccionarán los mejores descartándose los demás, se hiperparametrizarán aumentando su rendimiento predictivo y se entrenarán definitivamente.
 - Se buscarán alianzas entre los mejores modelos si fuera posible.
 - Se obtendrá un modelo único final, al que se llamará Karmona[®], con buen rendimiento y competitivo que tendrá la capacidad de predecir la potencia de un lente intraocular basado en los *inputs* biométricos seleccionados en la primera etapa.

- **Etapa IV:** para determinar el equivalente esférico en el plano corneal asociado a la potencia predicha por Karmona[®] en la fase anterior será necesario esta etapa óptica.
 - Teniendo en cuenta la precisión de manufactura de 0.50 D de lentes intraoculares, la predicción hecha por Karmona[®] se redondeará con la misma precisión y la diferencia se convertirá a plano corneal.

- La distometría aplicada en este caso será de dentro hacia afuera: LIO-córnea. Con la ELP calculada en la segunda fase óptica se determinará dicha equivalencia, que corresponderá al equivalente esférico en el plano corneal.
- **Etapa V:** en esta fase se validará la nueva fórmula Karmona® mediante el 20% de los datos del *dataset* original reservados a tal efecto, comprobando sobre todo que no hay un sobreajuste del potencial predictivo.
- **Etapa VI:** en esta fase final se harán comparaciones para evaluar la competitividad de Karmona® con el resto de las fórmulas utilizadas globalmente en clínica.

5.2 Diseño del estudio retrospectivo de la Etapa I

Para la realización de la presente tesis, durante la *Etapa I: Recogida, análisis y preparación de los datos*, será necesaria la inclusión de un estudio retrospectivo a partir del cual se recogerán los datos a emplear para las siguientes fases.

Se recogerán datos de pacientes intervenidos de cataratas o cristalino transparente por cuatro experimentados cirujanos con implantes intraoculares no tóricos monofocales, multifocales y de rango extendido en el Hospital Universitario Quironsalud Madrid.

Este estudio sigue los preceptos éticos dictados en la Declaración de Helsinki de la Asociación Médica mundial. Por ello, previo a su inicio, se obtuvo la aprobación por

parte del Comité Ético de la Fundación Jiménez Díaz del grupo hospitalario Quironsalud (EO153-18_HUQM) (ver [11.5 ANEXO 5.- APROBACIÓN COMITÉ ÉTICO](#)).

Para la base de datos se seleccionarán deliberadamente ojos axialmente muy cortos (≤ 20.00 mm) y extralargos (≥ 30.00 mm), intentando cubrir todo el rango de fabricación de lentes intraoculares, desde meniscos negativos a lentes biconvexas de la máxima potencia. Así como ojos con características anatómicas singulares, como queratometrías excesivamente planas o curvas o profundidades anormales del segmento anterior.

Todos los ojos del *dataset* deberán haber sido medidos con el mismo modelo de biómetro, IOL Master® 700 (Carl Zeiss Meditec AG, Jena, Alemania) tomando las siguientes variables biométricas: AXL, ACD, LT, WTW y KM (1.3375). Mientras que las medidas de curvatura posterior corneal se deberán haber adquirido con Pentacam® HR (Oculus Optikgeräte GmbH, Wetzlar, Alemania).

La base de datos dispondrá de 5 tipos distintos de lentes intraoculares ya que son las que se utilizan en el departamento de oftalmología del Hospital Universitario Quironsalud Madrid para cubrir todos los rangos de axialidad: Tecnis 1 ZCB00 (Johnson & Johnson, New Jersey, EE. UU.), Tecnis ZXR00 (Johnson & Johnson, New Jersey, EE. UU.), Sensor AR40M (Johnson & Johnson, New Jersey, EE. UU.), Finevision PodF (Physiol, Lieja, Bélgica) y Aspira-Aay (Humanoptics, Erlangen, Alemania) cuyas constantes fueron extraídas de www.iolcon.org (Tabla 1).

	SRKT	A ₀	A ₁	A ₂	SF	pACD
ZCB00	119.39	-1.272	0.1874	0.2542	2.05	5.829
ZXR00	119.06	-0.1886	0.1716	0.1977	1.826	5.591
AR40M	118.70				1.63	5.41
PODF	118.95	1.36	0.4	0.1	1.83	5.59
ASPIRA	118.72	-0.8658	0.2131	0.2141	1.626	5.374

Tabla 1. Constantes de las lentes utilizadas en la base de datos extraídas de iolcon.org.

La técnica quirúrgica utilizada independientemente del cirujano deberá ser siempre la misma, facoemulsificación ultrasónica con cuchilletas calibrados a 2.40 mm y a 2.50 mm incidiendo por córnea clara en zona temporal a 180°. Los facoemulsificadores podrán ser Centurion® *visión System* (Alcon, Fort Worth, EE. UU.) y *Stellaris® PC next generation* (Bausch & Lomb, Rochester, EE. UU.).

Todos los casos deberán haber sido graduados con gafa de prueba (distancia al vértice de 12.00 mm) al mes de la cirugía como mínimo y al tercer mes como máximo, por el mismo optometrista experimentado (D.C.G.), mediante técnicas refractivas subjetivas previa medición objetiva con autorefractómetro, con la única directriz de conseguir la máxima cota de agudeza visual con la menor cantidad de refracción esferocilíndrica posible.

5.2.1 Criterios de inclusión

Los criterios de inclusión establecidos para los ojos seleccionados serán los siguientes:

- Pacientes sin ningún procedimiento quirúrgico a nivel corneal o implantacional previo a la cirugía de cataratas.
- Pacientes sin ambliopías, ni estrabismos.
- Pacientes sin patología corneal u ocular general que pudiera afectar a las medidas biométricas y/o queratométricas además de su visión. En esta categoría se incluyen como enfermedades oculares: glaucoma, retinopatía diabética y degeneración macular asociada a la edad diagnosticadas antes de la cirugía de cataratas.
- Pacientes con astigmatismo queratométrico inferior a 1.00 D, calculado a partir de la queratometría simulada con índice queratométrico estándar 1.3375.
- Pacientes con la lente intraocular correctamente implantada en el saco capsular y sin descentramientos ni *tilts* (medido con Pentacam® HR).
- La agudeza visual de lejos con corrección esperada tras la cirugía debía ser superior a 0.50 en escala decimal (0.30 logMAR).
- Pacientes con refracción postoperatoria subjetiva realizada por D.C.G. al mes postquirúrgico.
- Pacientes sin alteraciones maculares importantes que afectaran en exceso a la agudeza visual, sobre todo en el caso de grandes miopes.

5.2.2 Criterios de exclusión

Los criterios de exclusión a considerar se enumeran a continuación:

- Pacientes con algún tipo de opacidad capsular objetivada mediante biomicroscopía.

- Quedarán excluidos del estudio aquellos pacientes con síndrome de *Sjögren* o que tengan un diagnóstico previo a la cirugía de cataratas de ojo seco.
- Igualmente se excluirán del estudio a los pacientes en los que se hubiese producido algún evento adverso intraoperatorio o en las revisiones postoperatorias posteriores, así como a aquellos a los que hubiese habido que suturar la incisión.
- Pacientes con alteraciones de curvatura corneal anterior/posterior o asimetrías corneales analizadas según Pentacam® HR.

5.3 Desarrollo computacional

Para el desarrollo computacional de esta tesis doctoral se utilizará RStudio® versión 1.1.423 (RStudio® Inc, Boston, EE. UU.) debido a su acceso libre, facilidad de manejo, a su optimización computacional, a la cantidad de librerías disponibles y sobre todo al conocimiento previo profundo de la herramienta en la etapa postgrado. En el momento del desarrollo de modelos, entrenamiento y análisis se empleará la última versión disponible del lenguaje de programación R, la 4.0.0.

La mayor parte del modelado y entrenamiento se desarrollará con la librería CARET versión 6.0-78 (*Classification And REgression Training*) de Max Kuhn (93). Pero también se explorarán otras opciones como e1071 versión 1.7-3 de David Meyer et al (94), EARTH versión 5.1.2 de Stephen Milborrow (95), CUBIST de Ross Quinlan (96) y BRNN de Paulino Rodríguez (97) , para la hiperparametrización y optimización de alguno de los modelos ya que permitían más profundidad, flexibilidad y atomización que CARET.

Para la presentación de gráficos y otros análisis estadísticos se utilizarán otros paquetes como GGLOT2 y SKIMR respectivamente.

En cuanto al hardware que se empleará para el trabajo computacional, será un *laptop* con Windows 10 (Microsoft, Redmond, EE.UU.). El procesador utilizado para implementar y entrenar los múltiples modelos de la presente tesis doctoral, es un Intel® core i7-6700 HQ de sexta generación dotado de 4 núcleos y 8 subprocesos con una frecuencia de reloj basal de 2.60 GHz y máxima de 3.50 GHz. Esta configuración hardware permite 8 *cores* de proceso en paralelo como máximo, que es lo que se utilizó con alguno de los algoritmos más demandantes.

Inicialmente todos los modelos se entrenarán mediante procesado secuencial para ver sus tiempos computacionales. En caso de que los algoritmos fueran “perezosos” superando los 10-15 segundos, se optará por la paralelización mediante la librería *doParallel* (98) versión 1.0.15 y sus dependencias en R, dedicando los 8 subprocesos de la CPU al proceso en paralelo del *training*. Con ello se obtendrá una notable mejora en los rendimientos temporales de los algoritmos más complejos como las redes neuronales (Figura 27).

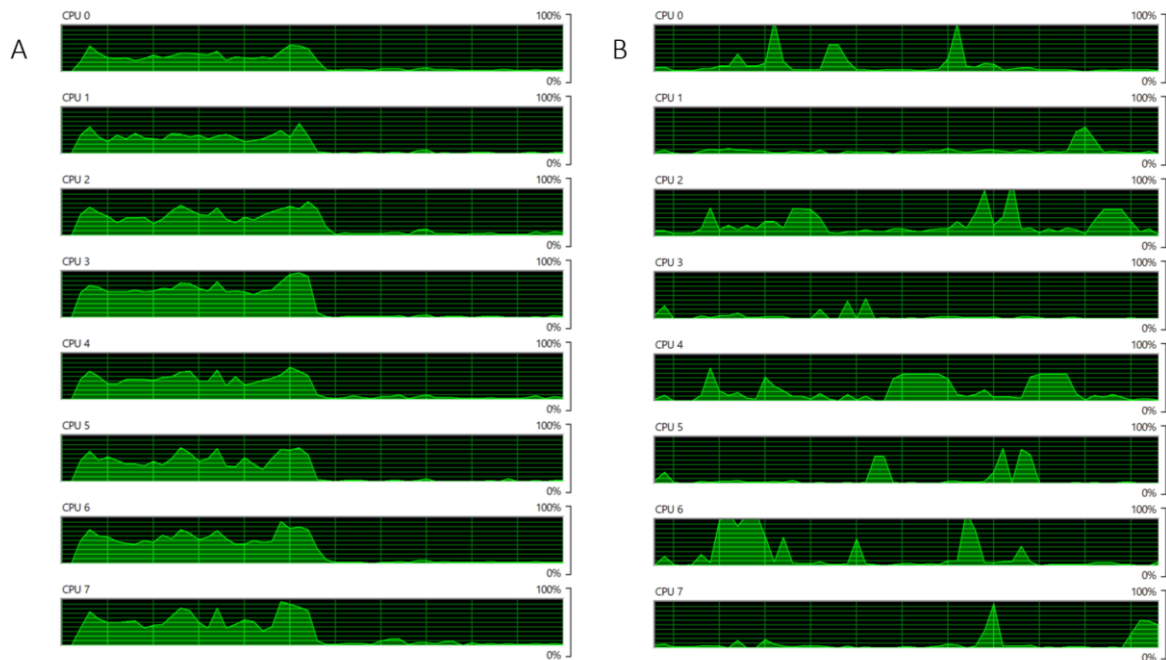


Figura 27. (A) Paralelización: rendimiento de los 8 cores durante el entrenamiento de una Red Neuronal multicapa, el tiempo de computación fue de 19 segundos. (B) Proceso secuencial: sin la activación del procesamiento en paralelo se ve como los núcleos van demandando de forma alternativa, el tiempo de computación en del entrenamiento de la misma Red Neuronal fue de 120 segundos (6 veces más).

5.4 Análisis estadístico

Para el análisis estadístico en los estudios de comparación entre fórmulas se utilizó RStudio versión 1.1.423.

El test de *Kolmogorov-Smirnov* se utilizó para estudiar la normalidad de las distribuciones de los datos. Las comparaciones de los errores absolutos entre las fórmulas se realizaron utilizando la prueba de *Friedman* con análisis de *Conover post hoc* y corrección de *Bonferroni* para las comparaciones múltiples.

Para evaluar si los errores de predicción refractivos dentro de los intervalos establecidos difieren estadísticamente entre fórmulas, se empleó *Cochran Q test* con análisis *post hoc* usando el test *McNemar* para comparaciones múltiples.

Se estableció un *p valor* menor de 0.05 como punto de corte para la significación estadística.

5.5 Cálculo del tamaño muestral

El cálculo del tamaño muestral se realizará mediante el software G*power versión 3.1.9.6 (99) (Universidad de Düsseldorf).

Como habrá una fragmentación de la muestra en dos partes (112), una para entrenar los modelos implementados (80%) y otra para validarlos (20%), los tamaños muestrales mínimos necesarios en cada submuestra tendrán justificaciones y orígenes diferentes. Para la muestra de entrenamiento, el objetivo es encontrar un mínimo muestral que permita un buen rendimiento predictivo. Sin embargo, para la muestra de validación se necesitará un tamaño mínimo de cara a las comparaciones posteriores con otras fórmulas en la fase final.

Respecto a la submuestra de entrenamiento, no será lo mismo predecir con modelos clasificatorios que con modelos regresivos, o que con patrones de reconocimiento sobre imágenes por ejemplo. En cada caso concreto el tamaño muestral requerido será distinto (100). Otro elemento clave será la dimensionalidad de la base de datos. En esta tesis se usarán modelos matemáticos regresivos de dimensionalidad

reducida, lo que redundaría en la posibilidad de una muestra poblacional más contenida, ahora bien, determinar el mínimo ideal es lo siguiente.

En base a estas ideas, el coeficiente de *Pearson* (r) evalúa la “fuerza” lineal entre dos variables aleatorias cuantitativas, siendo independiente de la escala de medida de las variables y oscila entre -1 y 1. Teniendo en cuenta esto, mediante la matriz de correlaciones entre las variables predictoras y la variable respuesta (Figura 28) se establece un mapa de coeficientes de *Pearson*. Fijando α en 0.05 y el poder estadístico ($1 - \beta$) en 0.95, en el mejor de los casos, un modelo lineal que incluyera la variable AXL y ACD por ejemplo, siendo estas las variables más correladas con la variable respuesta, ($r = -0.94$ y $r = -0.62$ respectivamente), el tamaño muestral necesario para establecer dicha relación no sería excesivo (Figura 29). Para obtener tamaños del efecto inferiores a 0.5 con una potencia estadística del 95% haría falta un n de 100 aproximadamente.

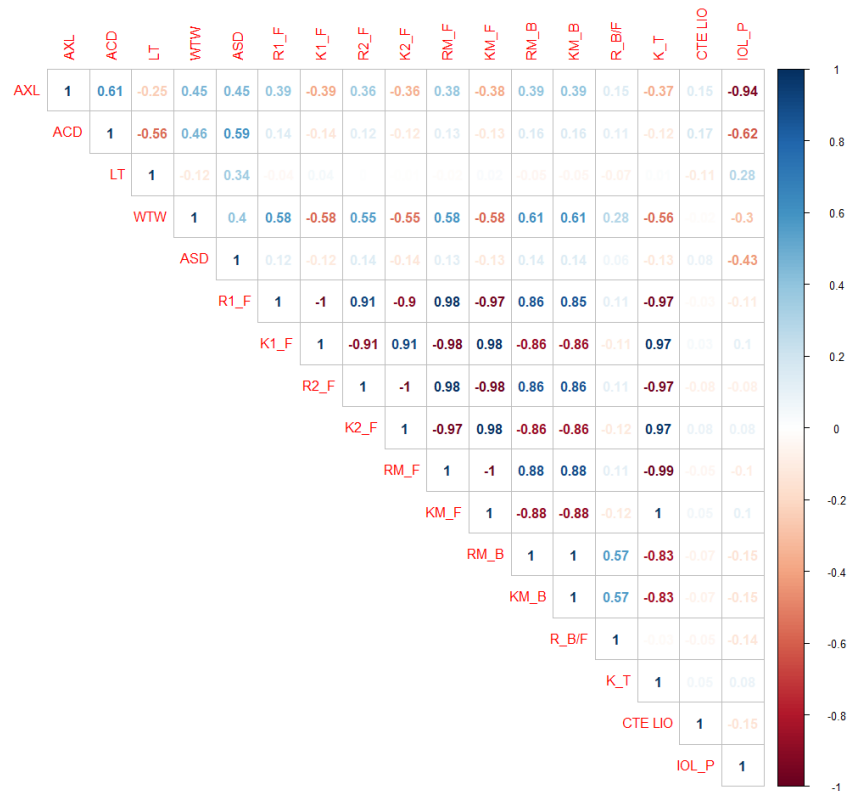


Figura 28. Diagrama de matriz de correlaciones entre variables predictoras y respuesta.

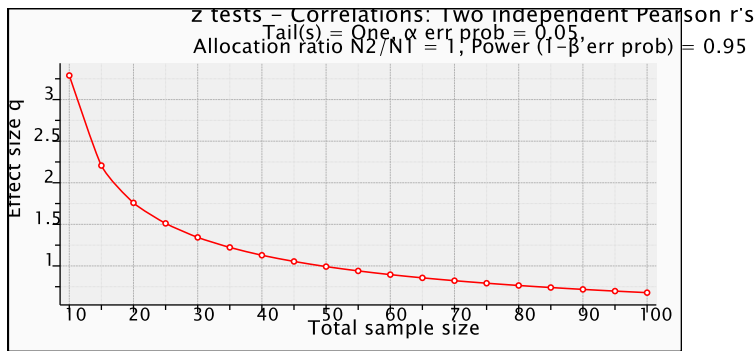


Figura 29. Tamaño muestral en función del tamaño del efecto cuando se comparan dos coeficientes de Pearson.

En el peor de los casos, sin incluir en la regresión ninguno de los predictores que más peso tienen, (Figura 30) se tendría un tamaño del efecto calculado de 0.22 y por tanto se tendría la necesidad de aumentar la muestra hasta aproximadamente 1000 ojos.

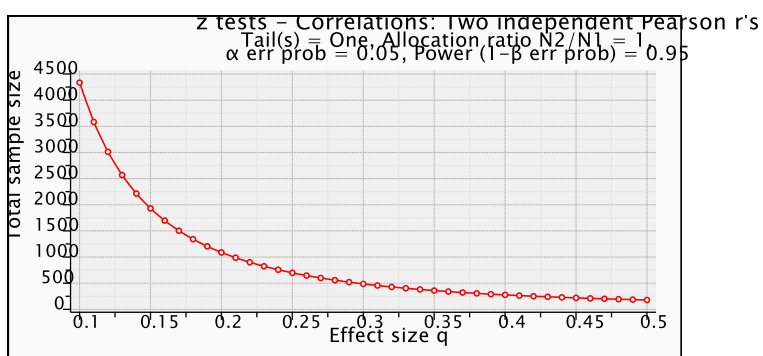


Figura 30. Tamaño muestral en función del tamaño del efecto en una comparación de coeficientes de correlación de Pearson de dos predictores con la variable respuesta.

Por tanto, desde un punto de vista regresivo, con tamaños muestrales para la partición de entrenamiento que oscilen entre 100 y 600 ojos se obtendrán a priori buenos resultados predictivos.

Por otro lado, en cuanto al tamaño mínimo necesario en la submuestra de validación (Figura 31), para un nivel de significación estadística $\alpha = 0,05$ y un poder estadístico $(1 - \beta)$ del 99%, la muestra debe ser de 95 datos.

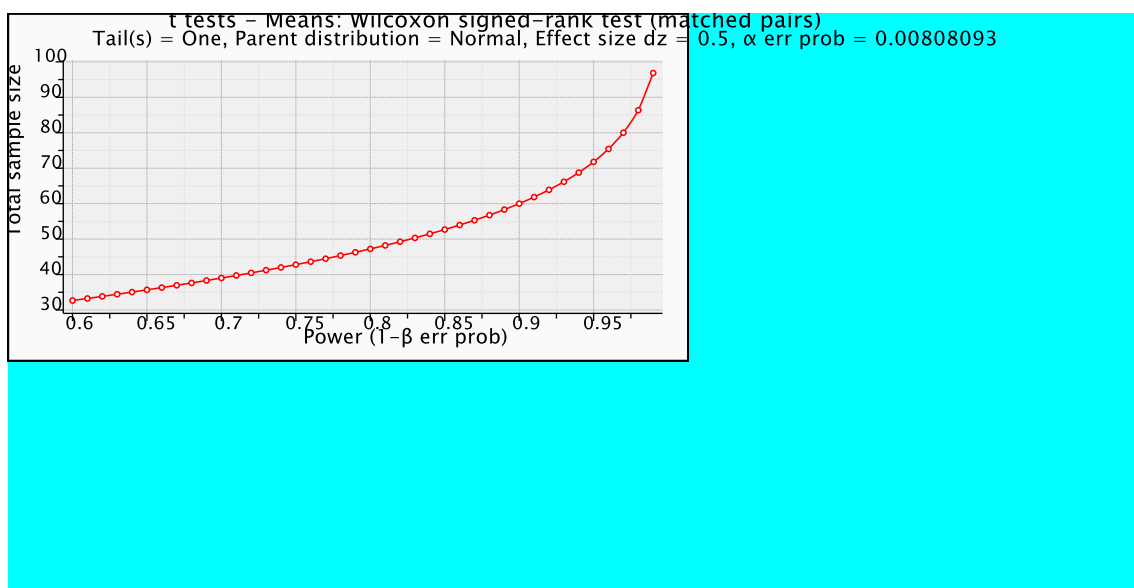


Figura 31. Poder estadístico en base el tamaño muestral para la partición de validación.

5.6 Diversidad muestral

La información de alta calidad es una herramienta muy poderosa en ML, sin embargo no siempre cuantos más datos se tengan más potente será el modelo (106). En ocasiones la adecuada selección de características, la laxitud en la complejidad del problema reduciendo la dimensionalidad además del uso y selección correcta de los

algoritmos de ML pueden compensar la falta de datos (107, 108), no obstante es fundamental poseer datos de calidad aunque no sean demasiados (109-111).

Esto significa que no es lo mismo un paciente operado de cataratas y revisado con un autorefractómetro una semana tras la cirugía y que se incluya en la base de datos, que ese mismo paciente graduado subjetivamente por un optometrista experimentado un mes después de la operación cumpliendo todos los protocolos. En el primer caso, dependiendo del tipo de lente intraocular implantada, el dato refractivo postoperatorio probablemente será susceptible de sesgo contaminando la muestra afectando a los valores de centralidad y dispersión, mientras que en el segundo la precisión y exactitud será mayor, propósito fundamental en esta investigación.

En este sentido, esta tesis cuenta a su favor con un precedente, que es la continuación de una investigación preliminar en la que se obtuvieron buenos resultados con una muestra pequeña (113). El modelo que se consiguió, rendía bien y su poder predictivo obtuvo mejores resultados en comparación con otras fórmulas de cálculo con las que se comparó. Por tanto, la pretensión en la presente investigación será extender la muestra en todas direcciones para sobre todo incrementar los datos en ojos extremos, y de esta forma se conseguirá un modelo predictivo más eficaz sobre todo en ojos marginales.

Otro de los factores que pueden influir positivamente en el potencial de predictibilidad de los modelos de ML es la variedad de la información. Para que el modelo sea lo más realista posible es necesario entrenarlo con datos de calidad procedentes de varias fuentes (114-116).

Como se expresó en el punto [5.2 Diseño del estudio retrospectivo de la Etapa I](#) se quiere potenciar la variabilidad quirúrgica con cuatro cirujanos, con tipos distintos de facoemulsificadores, con dos tipos de cuchilletas calibrados a 2.40 y 2.50 mm respectivamente, con cinco tipos distintos de lentes intraoculares con sus respectivas constantes optimizadas según www.iolcon.org, con la medición preoperatoria de dos dispositivos distintos y con la introducción de ojos únicos además de algún paciente con los dos en el total de la muestra (117). Consiguiéndose de este modo cierta variedad controlada pre-entrenamiento que concuerda con la realidad clínica con la que se encontrará el médico que afronte un cálculo biométrico.

Teniendo en cuenta que todas las inferencias hechas sobre cualquier modelo regresivo, sea cuál sea, y las fórmulas teóricas en alguna parte deductiva de la ELP lo son, están influenciadas por la realidad del *dataset* con el que están contruidos dichos modelos. Por tanto, se ven influenciados por cómo se graduó a esos pacientes, por cómo se operó, por cuánto astigmatismo se indujo quirúrgicamente, con qué dispositivos se midió preoperatoriamente y un largo etcétera de factores contribuyentes. Si en ML se construye un modelo con información de una sola lente y un solo cirujano, las predicciones para esa lente y ese cirujano serán probablemente excelentes, pero poco extrapolables a otras condiciones o circunstancias.

6 METODOLOGÍA

6.1 Etapa I: Recogida, análisis y preparación de los datos.

En los siguientes apartados se recogerá y estudiará la muestra para conocer cuáles son las variables que la componen con sus características y atributos, así como un análisis descriptivo por rangos de axialidad, determinación de correlaciones entre variables y analizar e investigar qué tratamiento previo al modelaje se dará a los datos.

6.1.1 Recogida de la base de datos

Se ha recogido una muestra total de 481 ojos. En la división *training* y *test*, la partición de entrenamiento constará de 386 ojos (80%) y la de validación de 95 ojos (20%), de acuerdo a los cálculos del tamaño muestral establecidos en el capítulo [5.5 Cálculo del tamaño muestral](#).

La base de datos consta de hasta 38 variables, aunque se pueden diferenciar dos partes principales: la de modelaje y la de comparación (Tabla 2). En la primera de ellas hay 14 variables que podrían tener potencialmente peso específico en la predicción, el resto de parámetros son los que conforman la otra parte y se utilizarán en la fase comparativa final del modelo obtenido con las fórmulas y métodos ya existentes.

BASE DE DATOS						
PARTE MODELAJE			PARTE COMPARACIÓN			
AXL	R_B/F	SEXO	T_HOL1	T_HILL	EP_HOL1	EP_HILL
ACD	IOL_P	EDAD	T_SRKT	T_KANE	EP_SRKT	EP_KANE
LT	SIMK_M		T_HOFQ	T_LADAS	EP_HOFQ	EP_LADAS
WTW	KM_B		T_HOL2	T_EVO	EP_HOL2	EP_EVO
ASD	CTELIO_OPT		T_HAIG	T_OLSEN	EP_HAIG	EP_OLSEN
K_T	EE_POSTOP		T_BARR	T_PEARL	EP_BARR	EP_PEARL

Tabla 2. Conjunto de variables incluidas en la base de datos de la tesis y su diferenciación en función del uso al que están destinadas.

En el siguiente apartado se explicará cada una de las variables mostradas en la Tabla 2 correspondientes a la parte del modelaje.

6.1.2 Datos para la predicción de modelos

Las variables del dataset de la parte de modelaje son:

1. **AXL** (*Axial Length*) o longitud axial. Es una variable numérica medida en milímetros y hace referencia a la longitud antero-posterior del globo ocular, desde el epitelio corneal hasta el epitelio pigmentario de la retina en la zona foveolar (Figura 32). Se determinó clínicamente con un biómetro IOL Master® 700 basado en tecnología Swept Source Optical Coherence Tomography (SS - OCT).
2. **ACD** (*Anterior Chamber Depth*) o profundidad de la cámara anterior. Es una variable numérica medida en milímetros y representa la distancia ortogonal entre el epitelio corneal y la superficie anterior del cristalino (Figura 32). Se midió con el mismo biómetro reseñado en el párrafo anterior.

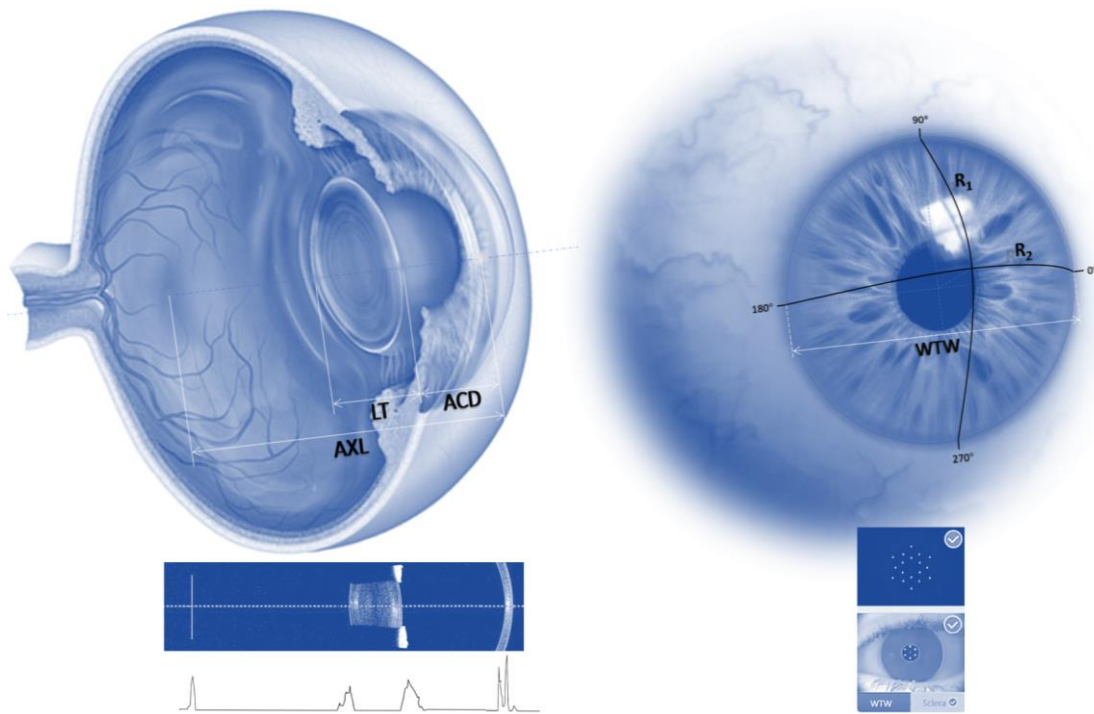


Figura 32. Anatomía intra y extraocular con sus variables biométricas.

3. **LT** (*Lens Thickness*) o espesor del cristalino. Es una variable numérica medida en milímetros que establece el grosor cristaliniano central, de cristaloides anterior a cristaloides posterior sobre el eje óptico (Figura 32). Se determinó también con IOL Master® 700.
4. **WTW** (*White To White*) o blanco a blanco. Es una variable numérica medida en milímetros y hace referencia al diámetro horizontal de la córnea de limbo a limbo esclerocorneal (Figura 32). Se utiliza como variable biométrica que caracteriza anatómicamente el ojo, porque según algunos autores puede tener cierta correlación con el saco capsular donde va alojada la lente intraocular ya que este no se puede medir directamente en clínica (118) lo que a su vez tiene relación directa con el tamaño del implante, no obstante existen otros estudios que apuntan en dirección contraria (119, 120). Se determinó con IOL Master® 700.

5. **ASD** (*Anterior Segment Depth*) o profundidad del segmento anterior. Es una variable numérica expresada en milímetros que representa la suma de la ACD y de LT (Figura 32). Se incluyó este parámetro por la posible relevancia anatómica ocular, ya que clásicamente los ojos se han dividido en tres categorías discriminando por longitud axial; cortos, normales y largos respectivamente. En consecuencia, podría haber ojos cortos con longitudes anteroposteriores inferiores a 21.00 mm y sin embargo poseer una ASD alta, o al contrario, presentarse un ojo axialmente largo con un segmento anterior pequeño (77).

6. **SIMK_M** (*Simulated keratometry Mean*) o queratometría simulada media. Se le ha designado como **KM_F** (*KeratometryMean_Front*) para una mejor identificación, aunque no se corresponda literalmente con la queratometría de la superficie anterior. Es una variable numérica expresada en dioptrías que representa la media de las queratometrías simuladas de los meridianos principales de la córnea. Se midió con IOL Master® 700.

El valor de queratometría simulada equivalente al radio de curvatura medido por el dispositivo, se obtiene a través de la siguiente fórmula gaussiana efectiva solo en la zona paraxial:

$$SIMK = \frac{(n' - n)}{R} \quad (24)$$

Donde K es la queratometría medida en dioptrías, R es el radio de curvatura corneal en milímetros y la diferencia de índices de refracción del numerador se corresponde al índice queratométrico 1.3375 y al índice del aire respectivamente, de manera que para un radio de curvatura medio 7.88 mm la $SIMK_M$ correspondiente sería:

$$SIMK_M = \frac{(1.3375 - 1)}{7.88 * 10^{-3}} = 42.83 D \quad (25)$$

El índice queratométrico 1.3375 no es un índice de refracción tisular real, es un artificio matemático que permite aproximar la potencia corneal calculada a la total, habiendo medido solamente la curvatura anterior de la córnea. Hasta hace tres lustros no se podía medir clínicamente de forma seriada la curvatura posterior de la córnea, de modo que a través de inferencias basadas en otras tecnologías como imágenes de Purkinje o fotografía con lámpara de hendidura (121, 122), se vió que el ratio entre curvaturas de ambas superficies corneales se mantenía más o menos constante en una mayoría poblacional.

Así que en una córnea normal y en la que el ratio de curvaturas anterior/posterior es aproximadamente 1.21, el índice queratométrico funciona relativamente bien y la estimación de la potencia corneal total no se aleja de la real (123, 124).

7. **KM_B** (*KeratometryMean_Back*) o queratometría media de la córnea posterior. Es una variable numérica expresada en dioptrías que mide la potencia queratométrica media de los meridianos principales de la superficie posterior corneal en los 3.00 mm centrales de la córnea. El dispositivo que se utilizó para determinar los radios de curvatura de la cara posterior fue Pentacam® HR.

Del mismo modo que se hizo anteriormente con la potencia de la superficie anterior, para determinar la equivalencia dióptrica del radio de curvatura y el poder dóptrico de la cara posterior, se utilizó la Ecuación (24) con índices de refracción reales del humor acuoso (1.336) y de la córnea (1.376). Precisamente la diferencia de índices del numerador es

negativa lo que implica que la potencia dióptrica de la superficie posterior corneal también lo sea. Para un radio de curvatura medio de la superficie posterior de la córnea de 6.33 mm su potencia equivalente sería a -6.32 D:

$$KM_B = \frac{(1.336 - 1.376)}{6.33 * 10^{-3}} = -6.32 D \quad (26)$$

8. **KM_T (Keratometry_Total)**: o queratometría total corneal. Es una variable numérica expresada en dioptrías y que representa la potencia total corneal central. Determinada mediante la fórmula gaussiana de acoplamiento de dos sistemas ópticos finos prescindiendo del efecto modulador que introduce en la ecuación el grosor corneal, ya que este no se recopiló en la muestra de la tesis.

$$K_T = K_{anterior} + K_{posterior} - \left[\frac{\text{paquimetría}}{n_{cornea}} * K_{anterior} * K_{posterior} \right] \quad (27)$$

Donde

$$K_{anterior} = \frac{(n_{córnea} - n_{aire})}{R_{anterior}} \quad (28)$$

$$K_{posterior} = \frac{(n_{humor\ acuoso} - n_{córnea})}{R_{posterior}} \quad (29)$$

El segundo término de la Ecuación (27) es muy pequeño en comparación con la suma algebraica de las potencias anterior y posterior de la córnea. El cociente entre el espesor corneal y su índice de refracción para una córnea media de 555 µm, será del orden de 0.00040 m, lo que implica que multiplicado respectivamente por las potencias

queratométricas anterior y posterior, se obtendrá un efecto modulador sobre la suma del primer término de la ecuación pequeño.

Para una córnea extremadamente fina (450 μm) y con curvaturas anteriores y posteriores medias normales se tendrá:

$$K_T = \frac{(1.376 - 1)}{7.67 * 10^{-3}} + \frac{(1.336 - 1.376)}{6.80 * 10^{-3}} - \left[\frac{450 * 10^{-6}}{1.376} * \frac{(1.376 - 1)}{7.67 * 10^{-3}} * \frac{(1.336 - 1.376)}{6.80 * 10^{-3}} \right] \quad (30)$$

El factor que introduce el grosor corneal en la determinación de la potencia total corneal central paraxial es únicamente 0.09 D, por tanto se tendrá:

$$K_T = 43.14 + 0.09 = 43.23 \text{ D} \quad (31)$$

Ahora bien, si con los mismo valores queratométricos se tiene un grosor corneal de 650 μm , el efecto modulador que induce el espesor corneal y todo su término sobre la potencia total del dioptrio corneal en el centro, únicamente sería de 0.14 D:

$$K_T = \frac{(1.376 - 1)}{7.67 * 10^{-3}} + \frac{(1.336 - 1.376)}{6.80 * 10^{-3}} + 0.14 \quad (32)$$

$$K_T = 43.14 + 0.14 = 43.28 \text{ D} \quad (33)$$

Por tanto, para diferencias de 200 μm en el espesor de la córnea, solo se producen variaciones nominales de 0.05 D en la potencia corneal total. De modo que se puede asumir la exclusión de la paquimetría corneal en el cálculo de K_T .

9. **R_B/F** (*Ratio_Back/Front*) o ratio posterior/anterior de la córnea. Esta variable es numérica adimensional. Se refiere al ratio entre el radio de curvatura central de la cara posterior de la córnea y anterior respectivamente multiplicado por 100. El radio de curvatura de la cara posterior fue medido con Pentacam® HR y el de la superficie anterior con IOL Master® 700.

Se incluye como caracterizador biométrico porque puede ser un factor diferencial y concluyente en ojos anatómicamente diferentes aportando la información necesaria de la cara posterior corneal. De hecho, ante la posibilidad de escoger el ratio extraído íntegramente de Pentacam® HR calculado con los radios de curvatura medios centrales corneales, se optó deliberadamente por mezclar la medida posterior obtenida con Pentacam® HR y la anterior de IOL Master® 700.

Según la literatura revisada en ojos normales que no han sufrido ningún procedimiento quirúrgico a nivel ocular, el promedio poblacional de este ratio es del 82.2% (125-127).

10. **SE_POST** (*Spherical Equivalent_post*) o Equivalente Esférico_postcirugía. Hace referencia al equivalente esférico residual postoperatorio al mes de la cirugía medido en el plano de la lente oftálmica, de hecho este fue medido con gafa de prueba en el total de la muestra a 12.00 mm del apex corneal. Es una variable numérica y expresada en dioptrías.

Matemáticamente el equivalente esférico es la combinación lineal de los dos componentes refractivos que conforman el defecto refractivo total de un individuo: la esfera y el cilindro, con independencia de la orientación que tenga la componente astigmática. La ecuación para el equivalente esférico es la suma algebraica de estos dos componentes:

$$SE_{POST} = esf + \frac{cil}{2} \quad (34)$$

De manera que para un defecto esferocilíndrico dado:

$$-0.50 \text{ esf} \approx -0.50 \text{ cil } 90^\circ$$

Se tendrá el siguiente equivalente esférico:

$$SE_{POST} = -0.50 + \frac{-0.50}{2} = -0.75 \text{ D} \quad (35)$$

Del mismo modo para este otro defecto esferocilíndrico:

$$+0.50 \text{ esf} \approx -2.50 \text{ cil } 180^\circ$$

su equivalente esférico será:

$$SE_{POST} = +0.50 + \frac{-2.50}{2} = -0.75 \text{ D} \quad (36)$$

Es decir, ante estos dos ejemplos tan dispares el equivalente esférico es idéntico, por lo que se puede aseverar que esta composición matemática esferocilíndrica omite toda la información relacionada con el astigmatismo y su orientación. No obstante, las fórmulas biométricas predicen un valor dióptrico para la lente intraocular y un valor asociado de equivalente esférico como residuo refractivo. Esto es así porque el cálculo que hace la fórmula o método empleado, se realiza para el círculo de mínima confusión (CMC), es decir, para el promedio de las dos potencias queratométricas de los dos meridianos principales de la córnea (Figura 33).

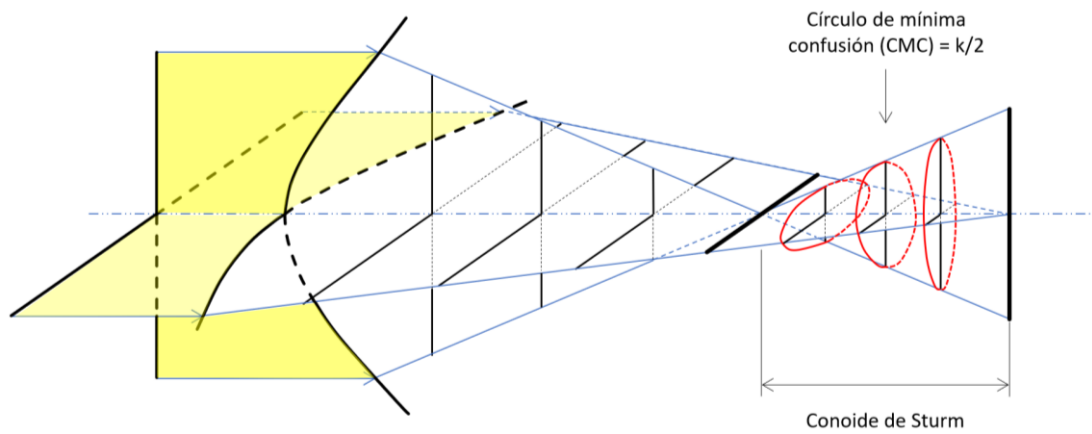


Figura 33. Conoide de Sturm y distribución de focales astigmáticas para un astigmatismo directo. Localización del círculo de mínima confusión en relación con la queratometría media.

Ópticamente hablando, el equivalente esférico convierte un sistema astigmático de focalidad doble en un sistema esférico de focalidad simple sobre el círculo de mínima confusión aportando únicamente información de su localización respecto a la retina tal y como se aprecia en la Figura 33. El CMC es el punto medio del conoide de Sturm, o lo que es lo mismo, el área en la que el foco se convierte en un círculo y no una elipse, ya que las dos focales astigmáticas tienen la misma longitud.

$$-0.50 \text{ esf} \approx -0.50 \text{ cil } 90^\circ$$

$$EE = \text{esf} + \frac{\text{cil}}{2} = -0.50 + \frac{-0.50}{2} = -0.75 \text{ D}$$

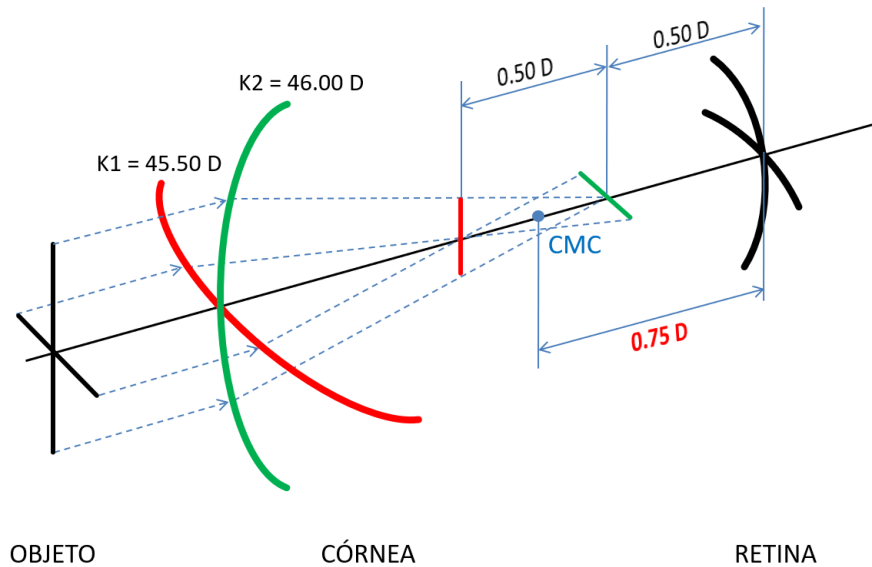


Figura 34. Posicionamiento respecto a la retina del círculo de mínima confusión (CMC) en función de la refracción esferocilíndrica.

En el *dataset* habrá equivalentes esféricos negativos, positivos y cero. Los negativos significan que la fórmula empleada sobreestimó la potencia de la lente a implantar colocando el CMC por delante de la retina (Figura 34). Por el contrario, los casos positivos implican infravaloración de la potencia refractiva del implante dejando el CMC por detrás de la retina.

11. **CTE_IOL** (*Constante_IntraOcular Lens*) o constante A óptica de la lente intraocular. Es una variable numérica, adimensional y hace referencia a la constante A de la lente que se ha implantado en el ojo. Esta variable se ha tenido en cuenta porque la operativa matemática con cualquier fórmula y o método de cálculo biométrico necesita discriminar por tipo de lente, para predecir un valor de potencia y de residuo refractivo concreto y específico para esa lente.

Cada modelo de lente intraocular tiene un diseño óptico, un espesor, radios de curvatura, un material, un tamaño, una geometría y todos estos parámetros físico-ópticos están aglutinados en una constante llamada constante del fabricante o constante A, única para cada modelo de lente.

Además de esto, la constante es el corrector universal para cualquier sesgo de la práctica clínica, permitiendo ajustarla en función de los resultados obtenidos para mejorar los futuros.

12. **IOL_P** (*IntraOcular_Lens_Perfect*) o lente intraocular perfecta. Se refiere a la potencia de la lente intraocular que se tendría que haber implantado en vez de la implantada realmente, es decir, la lente óptima que hubiera desplazado el CMC a retina teniendo en cuenta el equivalente esférico postoperatorio residual. Es una variable numérica y expresada en dioptrías.

Esta es la variable primordial y relevante, que permitirá predecir la potencia del implante teniendo en cuenta retrospectivamente el error cometido a nivel refractivo postoperatorio, ya que será la variable respuesta en los modelos. Es por ello que en la siguiente etapa metodológica se analiza cómo se obtiene.

13. **Sexo y edad**: son las últimas variables que se recogieron para ver si contribuían al rendimiento predictivo.

6.1.3 Datos para la comparación entre fórmulas

Por otro lado, se tiene la parte del *dataset* que se utilizará para la comparación de los resultados obtenidos con el modelo propuesto en la presente tesis. Respecto a la

selección de las 12 fórmulas que se utilizarán en la comparación del modelo obtenido en la tesis, se seleccionaron las fórmulas teóricas que históricamente y según la literatura publicada han funcionado mejor. Al mismo tiempo se quiso introducir las que están relacionadas con IA difundidas por sus autores en los últimos años (ver [4 ESTADO DEL ARTE](#)). Por tanto, todas las fórmulas utilizadas para la comparación con el modelo obtenido en la tesis fueron: Hoffer Q, Holladay I, SRK/T, Holladay II, Haigis, Barrett Universal II, Olsen PhacoOptics versión 1.10.100.2030-35, Ladas Superformula v1b, Kane, EVO v2.0, Hill-RBF v3.0 y PEARL-DGS v2.0.

Esta parte del *dataset* se compone de 24 variables, de las cuales 12 corresponden a la predicción de residuo refractivo en equivalente esférico que hace cada fórmula asociado a la lente intraocular que se implantó en cada caso y las otras 12 por tanto son los errores predictivos medidos al mes de la cirugía que han cometido dichas fórmulas y métodos respecto al equivalente esférico real obtenido postquirúrgicamente (Tabla 3).

FÓRMULA	IOL_I	TARGET REFRACTIVO_FÓRMULA	EE_POSTOP	ERROR PREDICTIVO
HOLLADAY II	+14.50 D	T_HOL2 = -0.13 D	-0.25 D	EP_HOL2 = -0.25 -(-0.13) = -0.12 D
SRK/T	+14.50 D	T_SRKT = -0.33 D	-0.25 D	EP_SRKT = -0.25 -(-0.33) = +0.08 D

Tabla 3. Obtención del error predictivo (EP) cometido por las fórmulas mediante del equivalente esférico postoperatorio.

En el ejemplo de la Tabla 3, para una potencia de LIO implantada de +14.50 D, la fórmula Holladay II asociado a esa potencia implantada predice un equivalente esférico miópico postoperatorio de -0.13 D, sin embargo la fórmula SRK/T predice -0.33 D para la

misma potencia. El residuo refractivo postquirúrgico al mes de la cirugía fue de -0.25 D, lo que implica que la fórmula Holladay II cometió un error de -0.12 D en términos absolutos, mientras que SRK/T solo fallo en +0.08 D. Lo que implica que Holladay II sobreestimó la potencia de la lente intraocular y SRK/T la infravaloró levemente. Obviamente el valor de error más cercano a 0.00 D será el mejor, en este ejemplo SRK/T.

Estas 12 fórmulas recopiladas para la comparación se pueden dividir en dos grandes subgrupos: GRUPO A: integrado por fórmulas teóricas y GRUPO B: conformado por fórmulas y métodos basadas en IA o alguna técnica de ML (Tabla 4).

GRUPO A		GRUPO B
FÓRMULAS TEÓRICAS		FÓRMULAS IA
HOLLADAY I	HAIGIS	HILL-RBF V3.0
HOFFER Q	HOLLADAY II	LADAS SPF
SRK/T	BARRETT U2	KANE
OLSEN	EVO V2.0	PEARL-DGS V2.0

Tabla 4. División en dos grupos de todas las fórmulas incluidas en la tesis para la comparación futura entre fórmulas teóricas y basadas en IA.

Se utilizará esta división para evaluar qué grupo es más preciso en el cálculo de la potencia de la lente intraocular, fórmulas teóricas versus IA.

6.1.4 Correlación de variables y selección de características

Para buscar relaciones y dependencias entre variables predictoras habitualmente se utilizan matrices de correlación. Frecuentemente dichas dependencias se estudian en base al coeficiente de *Pearson*, de *Spearman* o de *Kendall*, en función de la distribución de los datos y de otros condicionantes.

Para las variables que no se distribuyen normalmente, se utilizarán el coeficiente *rho de Spearman* y la *Tau de Kendall* como alternativa para ver sus relaciones con el resto de variables. Además del valor del coeficiente, es necesario calcular su significancia estadística. Solo si el *p-valor* es significativo se puede aceptar que existe correlación entre las variables estudiadas. Por muy cercano que sea el valor del coeficiente de correlación a +1 o -1, si no es significativo, se ha de interpretar que la correlación de ambas variables es 0 ya que el valor observado podría ser debido al azar.

La colinealidad y multicolinealidad son problemas insidiosos en los enfoques regresivos, y el de la presente tesis lo es. Si en un modelo de regresión lineal alguna variable independiente es combinación lineal de otra u otras, el modelo matemáticamente hablando es irresoluble. La multicolinealidad produce los siguientes problemas a nivel predictivo:

- Intervalos de confianza muy amplios.
- Errores estándar muy altos.
- Valores de T muy bajos.
- Estimadores poco robustos ya que pequeños cambios en la base de datos generan cambios importantes en los coeficientes estimadores.

El problema es esencialmente la falta de suficiente información para poder estimar coeficientes con precisión, no son eficientes ni consistentes aunque sigan siendo insesgados.

La solución a la multicolinealidad es reducir la dimensionalidad del modelo, es decir, seleccionar las adecuadas características o variables no correladas entre ellas que permitan un buen desempeño del modelo regresivo, a este proceso se le llama *Feature Selection*.

6.1.5 Transformación de datos

En muchas ocasiones, se hace necesario la transformación de los datos previa a la implementación de modelos con el fin de obtener los mejores y más eficientes modelos de regresión. En este tipo de modelaje el proceso de transformación de datos que se necesita es la estandarización y normalización para igualar la información en una misma escala porque al pretender obtener información de diversas escalas el peso que posee la variable se hace en base a su ponderación, lo que resalta las diferencias entre variables.

A priori, es difícil saber qué metodología de preprocesamiento o transformación (128) de datos aplicar o si es necesario aplicarla, aunque existen determinadas normas que suelen funcionar bien. Los modelos basados en distancias, como los *k-Nearest Neighbour* (kNN) por ejemplo, funcionan mejor si las variables de entrada tienen la misma escala. Generalmente, las transformaciones tienen más probabilidades de ser útiles para todos los algoritmos de regresión. Sin embargo, para métodos basados en árboles y reglas como por ejemplo *Random Forest* (RF) y *Tree Decision* (TD), las transformaciones suelen resultar menos útiles.

No obstante, incluso empleando dichas transformaciones de datos en los modelos predictores, puede que no se mejore la capacidad predictiva de dichos modelos, produciéndose mejores resultados con los datos en crudo. De modo que lo mejor y más prudente, es chequear la eficiencia del modelo implementado con y sin transformación.

En la base de datos de la tesis las unidades de medida y escalas son diferentes para algunas de las variables predictoras, todas las queratometrías están expresadas en dioptrías, el resto de variables biométricas se expresan en milímetros y dos variables adimensionales: la constante de la lente y el ratio de curvaturas corneales. Muchos algoritmos regresivos de ML son sensibles a esto, de forma que si no se igualan los predictores, aquellos que se midan en una escala mayor o que tengan más varianza, dominarán el modelo aunque no sean los que más relación tienen con la variable respuesta.

Por otro lado, en el *dataset* de esta tesis no fue necesario imputar variables ausentes, se ha reducido la dimensionalidad de la base de datos disminuyendo la multicolinealidad, eliminando variables como el sexo, la edad y los radios de curvatura anteriores y posteriores.

6.2 Etapa II: Equivalencias de potencia entre planos

En esta fase óptica del desarrollo se distinguen tres planos de trabajo: la gafa, la córnea y la lente intraocular con sus respectivas distancias. En esta investigación, la gafa y la córnea están separados 12.00 mm, mientras que la córnea y la lente intraocular

tienen una distancia variable llamada ELP que dependerá de ciertos parámetros como se verá más adelante.

La idea que prima en esta etapa metodológica es corregir retrospectivamente la potencia de la lente implantada en todo el *dataset* teniendo en cuenta el equivalente esférico residual postoperatorio, para que la predicción de los algoritmos de ML utilizados posteriormente sea para la lente que emetropiza y no la que produce errores refractivos postoperatorios. Para esto es necesario convertir el equivalente esférico en el plano de la gafa al plano capsular del implante pasando por la córnea.

Más tarde, cuando el modelo de ML determina la potencia emetropizante lo hace en el plano del implante y por tanto es necesario determinar la equivalencia distométrica sobre el plano corneal que será el resultado final que verá el usuario de la fórmula.

En los sucesivos apartados se disertará sobre estas equivalencias y conversiones entre estos planos, así como de la obtención de la ELP.

6.2.1 Residuo refractivo postoperatorio en plano corneal

Si en el abordaje teórico clásico del cálculo biométrico, lo que se infiere es la posición respecto a la córnea en la que va a quedar el implante intraocular, en la predicción mediante IA, la situación cambia sustancialmente, ya que como se explicó anteriormente la predicción apunta directamente a la potencia de la lente que se implantará desconociendo cuál será su posición postoperatoria (Figura 35).

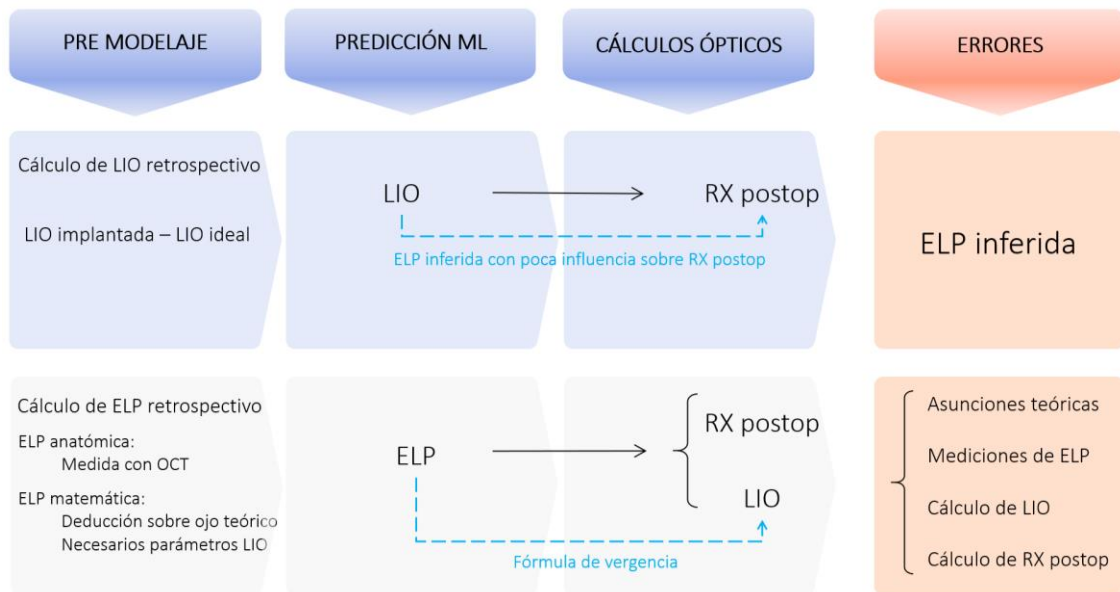


Figura 35. Esquema de posibles abordajes en la predicción con ML para una calculadora de potencia emetropizante.

La Figura 35 muestra los dos caminos posibles que existen para calcular la potencia dióptrica de la lente emetropizante en el modelaje con ML con todas las capas de abstracción. Como ya se ha expresado anteriormente, en esta tesis se decidió aplicar lo que se consideró técnicamente más sencillo y que menos error final produciría, la predicción de la potencia de la lente.

Si la predicción de potencia emetropizante ofrecida por el modelo de IA fuese por ejemplo 12.092 D, teniendo en cuenta la precisión manufacturera de 0.50 D (129), la lente intraocular que dejaría el menor residuo refractivo postoperatorio y que se tendría que implantar sería la más próxima, es decir 12.00 D. De modo que la diferencia entre la potencia predicha y la real implantada en el plano capsular sería el residuo refractivo postoperatorio, o lo que es lo mismo, 0.092 D.

Este dato carece de interés práctico ya que es efectivo solo en el plano capsular, por tanto, aplicando una fórmula distométrica dependiente de la potencia corneal y de

la distancia nominal entre planos principales, se podría convertir dicho residuo refractivo a plano corneal. Para una córnea de 44.00 D y una posición efectiva de 4.50 mm, 0.092 D en plano capsular es equivalente a 0.06 D teniendo en cuenta el índice de refracción del humor acuoso.

Por tanto, la dificultad técnica que se presenta en el planteamiento anteriormente expuesto es la consecución del posicionamiento del implante. Sin este dato, la equivalencia real entre ambos planos, córnea y lente intraocular no es posible.

En cualquier caso y antes de evaluar las diferentes estrategias seguidas para determinar la posición efectiva de la lente, es necesario destacar la repercusión a nivel de precisión que tiene dicha distancia en la conversión de potencias dióptricas.

6.2.1.1 Repercusión en la precisión debida a la distometría con ELP

Sabiendo que la predicción de potencia dióptrica del implante intraocular en el plano capsular será con precisión de centésima, pero al mismo tiempo la ofrecida generalmente por parte del fabricante será de media dioptría, la diferencia en el plano del saco cristalino entre la potencia emetropizante predicha y la lente seleccionada para el implante nunca superará el cuarto de dioptría. De lo que se deduce que la influencia de la distancia para la equivalencia entre plano capsular y plano corneal de cantidades tan pequeñas no es muy relevante, siendo cada vez mayor si hubiera una desviación cada vez mayor de la potencia emetropizante.

Supóngase que la potencia predicha por el modelo oscila entre 15.00 y 15.50 D respectivamente, pero en un primer caso se tiene concretamente 15.24 D como potencia

emotropizante y en un segundo caso 15.26 D, los residuos refractivos en el plano de la lente serán:

$$\text{Caso 1} \quad 15.24 - 15.00 = 0.24 \text{ D}$$

$$\text{Caso 2} \quad 15.26 - 15.50 = 0.24 \text{ D}$$

En el supuesto de que el fabricante ofreciera la posibilidad de seleccionar una lente con precisiones de cuarto de dioptría, la diferencia entre potencias en el plano capsular sería todavía menor:

$$\text{Caso 1} \quad 15.24 - 15.25 = 0.01 \text{ D}$$

$$\text{Caso 2} \quad 15.26 - 15.25 = 0.01 \text{ D}$$

Imagínese que tenemos una ELP de 4.83 mm y que la diferencia entre la potencia predicha y la seleccionada para el implante es de 0.24 D en plano capsular, la equivalencia en el plano corneal sería 0.17 D para una córnea de 44.00 D.

Ahora supóngase un error de ± 0.50 mm en la medición de la ELP anterior, y a continuación se estudiará la variación distométrica entre planos para un valor de 0.24 D como el del ejemplo anterior:

$$\text{Error máximo: } 4.83 + 0.50 = 5.33 \text{ mm, equivalencia} = 0.16 \text{ D en córnea}$$

$$\text{Error mínimo: } 4.83 - 0.50 = 4.33 \text{ mm, equivalencia} = 0.18 \text{ D en córnea}$$

Es decir, que en el hipotético caso de que la estimación de la ELP oscile en 1.00 mm, la repercusión sobre la distometría entre planos con residuos inferiores al cuarto de dioptría es muy pequeña, del orden de 0.02 D en los extremos del intervalo. Si el residuo dióptrico

en el plano capsular fuera menor, por ejemplo, la mitad de los extremos, el impacto es despreciable tal y como se aprecia en el siguiente ejemplo:

$$4.83 + 0.50 = 5.33 \text{ mm, equivalencia de } 0.12 \text{ D} = 0.08 \text{ D}$$

$$4.83 - 0.50 = 4.33 \text{ mm, equivalencia de } 0.12 \text{ D} = 0.09 \text{ D}$$

Y para residuos más pequeños todavía, la influencia sería absolutamente despreciable en las precisiones que se manejan:

$$4.83 + 0.50 = 5.33 \text{ mm, equivalencia de } 0.06 \text{ D} = 0.04 \text{ D}$$

$$4.83 - 0.50 = 4.33 \text{ mm, equivalencia de } 0.06 \text{ D} = 0.04 \text{ D}$$

Lo que demuestra la escasa influencia de variaciones importantes de la ELP sobre el residuo refractivo en plano corneal cuando en el plano capsular se implantan lentes con precisiones de media dioptría, permitiendo esto una tolerancia alta para la inferencia de la ELP.

Todo lo expuesto anteriormente tiene validez cuando se habla de la potencia emetropizante y su residuo refractivo asociado a nivel corneal, pero para un gradiente de potencias por encima o debajo de esta, como muestran los calculadores reales utilizados en clínica, y sus respectivos residuos asociados la situación podría cambiar.

La presentación de resultados en los diferentes calculadores y fórmulas siempre se hace con una tabla de potencias dióptricas incluida la emetropizante y sus correspondientes predicciones refractivas a nivel corneal (Figura 36).

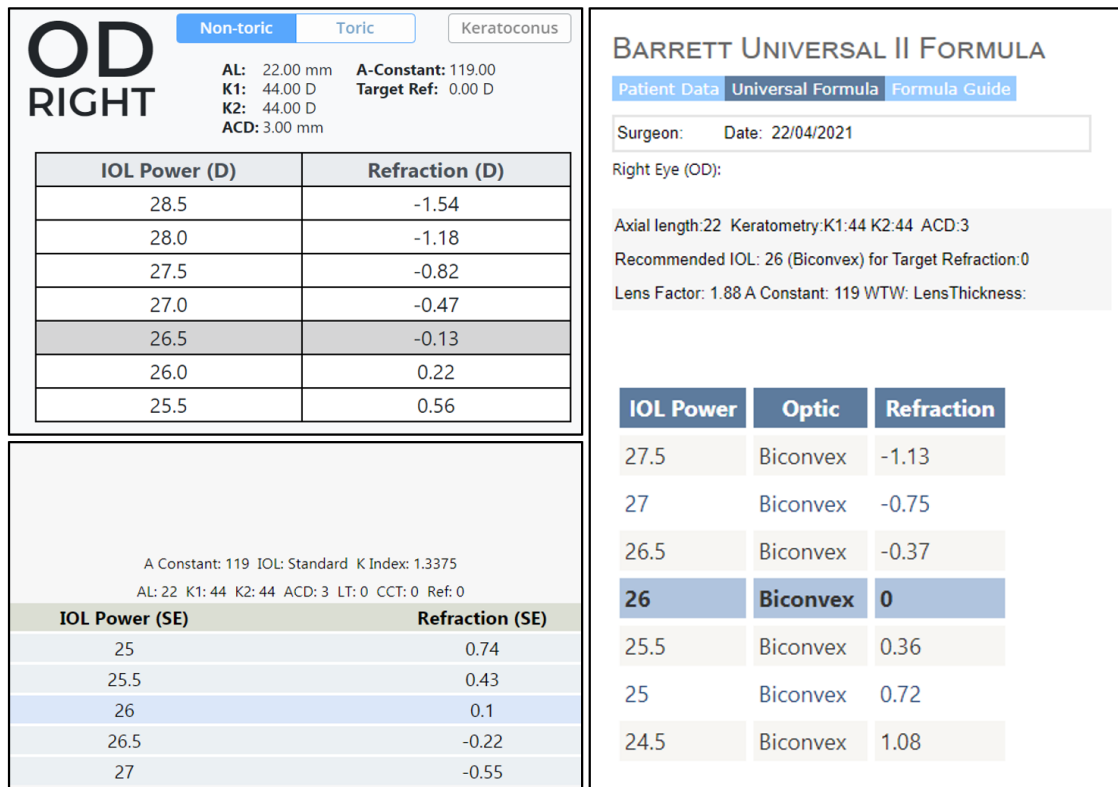


Figura 36. Presentación de resultados de diferentes fórmulas (Kane, Barrett Universal II y EVO v2.0) mostrando el gradiente de potencias con precisión de 0.50 D y sus residuos refractivos asociados en plano corneal.

Siguiendo el ejemplo de la Figura 36 si la potencia de la lente emetropizante según Barrett es 26.00 mm y se decide implantar por la razón que sea 27.50 D (se quiere dejar un residuo miópico postoperatorio), la diferencia entre ambas potencias en el plano capsular será $26.00 - 27.50 = -1.50$ D, que según la ELP estimada por Barrett y la potencia corneal, equivalen a -1.13 D en plano córnea tal y como se aprecia en el cálculo de la Figura 36.

Aplicando la fórmula de equivalencia entre planos, para una córnea de 44.00 D la ELP podría oscilar entre 4.53 y 4.67 mm sin que el residuo refractivo a nivel corneal se resienta. Si la horquilla de error en la estimación de ELP fuese ± 0.50 mm la variación en el resultado refractivo corneal sería del orden de 0.03 D en los extremos del intervalo.

Errores asumibles tratándose de potencias no emetropizantes y de asunciones solo probables en escenarios adversos.

6.2.1.2 Obtención de la ELP

Por tanto, a partir de aquí los esfuerzos se concentraron en obtener una posición efectiva del implante lo suficientemente precisa como para que el impacto sobre el residuo corneal fuese mínimo, sobre todo en las potencias alejadas de la potencia emetropizante central. Las estrategias que se plantearon fueron las siguientes:

1. ELP constante
2. ELP variable

Para estas estrategias se utilizó como referencia la fórmula de Barrett Universal II y los resultados refractivos corneales que esta consiguió para diferentes características anatómicas y biométricas.

Se registraron en una hoja de cálculo 46 ojos desde 20.00 mm hasta 30.00 mm de longitud axial, con diferentes profundidades de cámara, mismas queratometrías, y todo para una misma lente intraocular. Para todos ellos se calculó con Barrett Universal II la potencia emetropizante además de tres potencias por encima y por debajo de esta. Se comprobó la diferencia en plano capsular entre la potencia de la lente emetropizante y el resto de las potencias que pertenecían a un mismo rango de cálculo biométrico. Posteriormente, se validaron las diferentes estrategias planteadas para la posición efectiva (ELP constante y ELP variable) comprobando la diferencia existente entre el

residuo en plano corneal producido por Barrett y el producido teniendo en cuenta las distintas maneras de obtener las ELP.

6.2.2 Obtención del equivalente esférico en plano capsular

Considerando el equivalente esférico postoperatorio como el error en el cálculo de la lente intraocular implantada, sería una equivocación técnica sustraerlo sin más a la potencia del implante para obtener la lente intraocular idónea.

Se deberá tener en cuenta y especial precaución con las equivalencias de potencia refractiva entre los diferentes planos implicados en este proceso, ya que el equivalente esférico postoperatorio está medido a 12.00 mm del apex corneal en el plano de una lente oftálmica y la prótesis ocular quedará alojada aproximadamente a unos 5.50 mm del apex corneal en el saco capsular.

Por tanto, los planos útiles (Figura 37) y necesarios son el de la gafa, el corneal y el de la lente intraocular, separados uno de otro respectivamente por la distancia al vértice corneal o *Vertex Distance* y por la posición efectiva de la lente (ELP). Desde el punto de vista de la óptica geométrica esencial, estas distancias se miden desde los planos principales de cada uno de los elementos ópticos implicados.

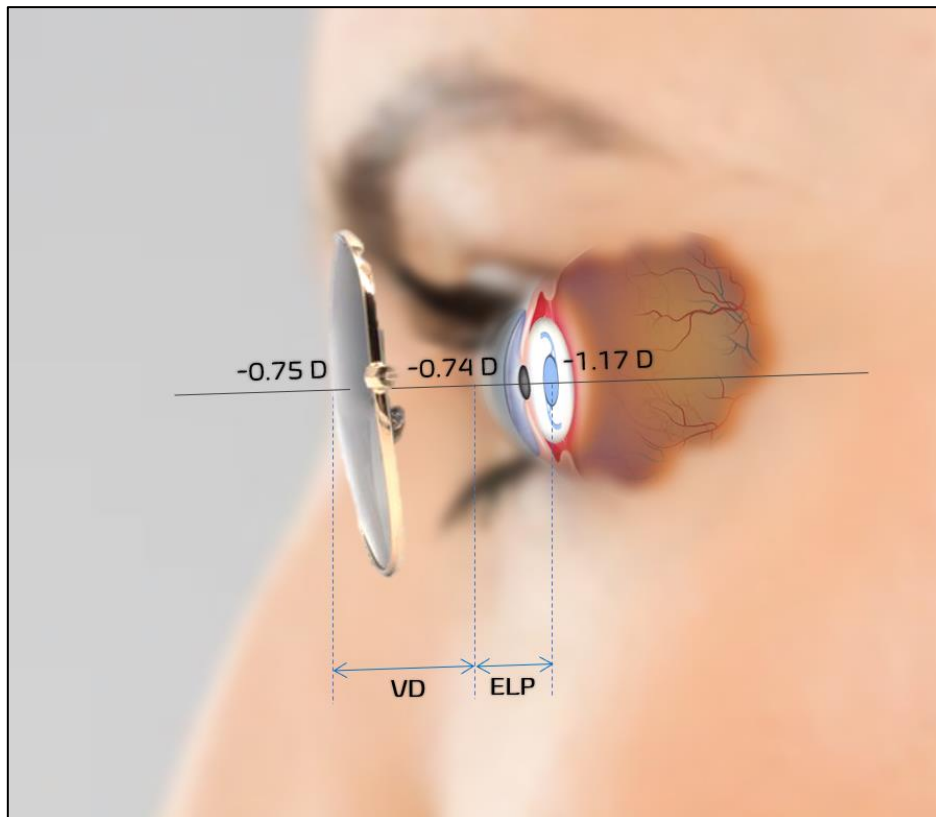


Figura 37. Planos de equivalencia dióptrica entre lente oftálmica, córnea y lente intraocular con sus respectivas distancias: distancia al vértice o Vertex Distance (VD) y posición efectiva de la lente (ELP).

En primera instancia, como el equivalente esférico obtenido es en el plano de la gafa se deberá aplicar la distometría entre lente oftálmica y córnea teniendo en cuenta que están separadas 12.00 mm por aire con un índice de refracción de 1.000. Posteriormente esta potencia se convertirá al plano del saco capsular teniendo en cuenta la posición efectiva de la lente y el índice de refracción del humor acuoso 1.336.

El cálculo de las distometrías o equivalencias dióptricas bien se podría abordar desde dos puntos de vista, considerando todos los sistemas ópticos involucrados como finos, o por el contrario como sistemas con espesor lo que complica la operativa matemática notablemente por los cálculos asociados de los planos principales imagen y objeto y sus respectivas distancias. Por tanto y por simplificación, se optó por asumir los sistemas ópticos como delgados despreciando su grosor basándonos en dos argumentos,

que los defectos refractivos manejados para calcular las equivalencias son muy bajos, inferiores en la mayoría de los casos a 1.00 D. Y por otro lado, que los radios de curvatura de las superficies implicadas son muy superiores a dichos espesores.

En primer lugar, para el cálculo distométrico entre gafa y córnea teniendo en cuenta las vergencias imagen de ambas superficies (Figura 38), separadas una distancia d por el índice de refracción de 1.000 del aire y teniendo en consideración la equivalencia de potencias en los distintos planos, se tiene:

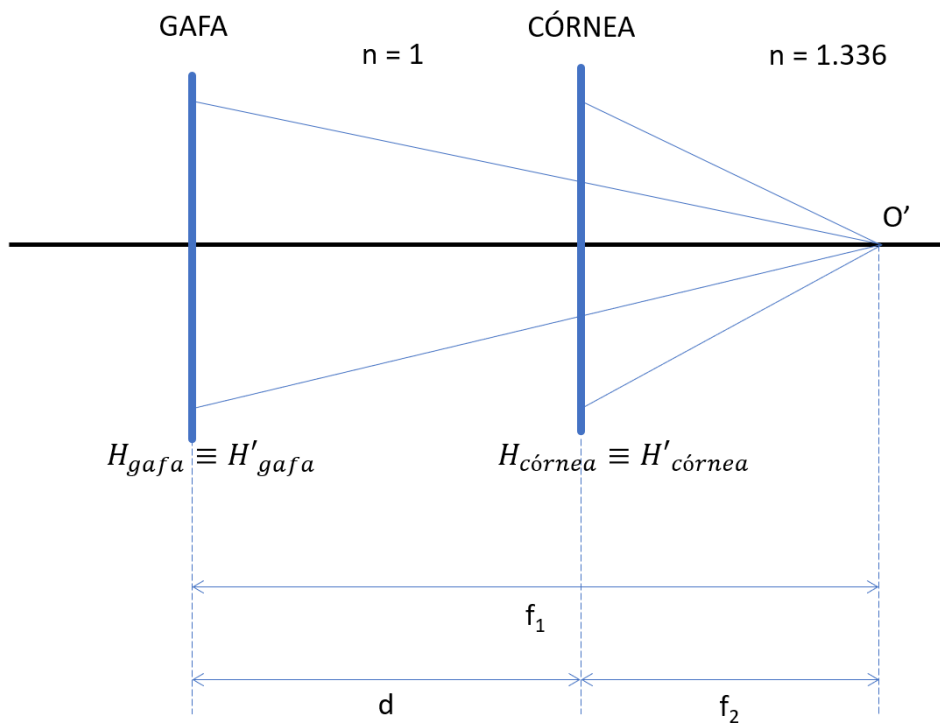


Figura 38. Esquema de equivalencia dióptrica entre lente oftálmica y ápex corneal considerando ambos sistemas ópticos como delgados.

$$P_{gafa} = \frac{n_{aire}}{f_1} \quad ; \quad f_1 = \frac{1}{P_{gafa}} \quad (37)$$

$$P_{córnea} = \frac{n_{aire}}{f_2} = \frac{1}{f_2} \quad (38)$$

Para que las potencias sean equivalentes, las vergencias imagen de ambas superficies deben coincidir en el mismo punto focal O' (Figura 38). De este modo la distancia focal de la córnea será:

$$f_1 = d + f_2 \quad ; \quad f_2 = f_1 - d \quad (39)$$

Como la equivalencia se quiere encontrar en el plano corneal:

$$P_{córnea} = \frac{1}{f_2} = \frac{1}{f_1 - d} = \frac{1}{\frac{1}{P_{gafa}} - d} \quad (40)$$

Simplificando la ecuación queda:

$$P_{córnea} = \frac{P_{gafa}}{1 - d * P_{gafa}} \quad (41)$$

Esta será la expresión que se utilizará para determinar la distometría entre gafa y córnea, donde d es la separación expresada en metros entre ambas superficies.

En este trabajo se graduó a toda la muestra con gafa de prueba separada del apex corneal 12.00 mm. Así para un equivalente esférico residual postoperatorio de -0.75 D determinado en gafa de prueba a 12.00 mm del ojo, tendremos su equivalencia dióptrica en plano corneal de:

$$Potencia_{córnea} = \frac{-0.75}{1 - (-0.75 * 0.012)} = -0.743 D \quad (42)$$

El siguiente paso será convertir esta potencia al plano capsular dónde está alojada la lente. En estas circunstancias se necesitará una variable nueva, la posición

efectiva de la lente que determinará dicha correspondencia entre planos y ya se dedujo en el apartado [6.2.1.2 Obtención de la ELP](#).

Y por último, la fórmula de equivalencia dióptrica entre córnea y lente intraocular se dedujo del siguiente modo:

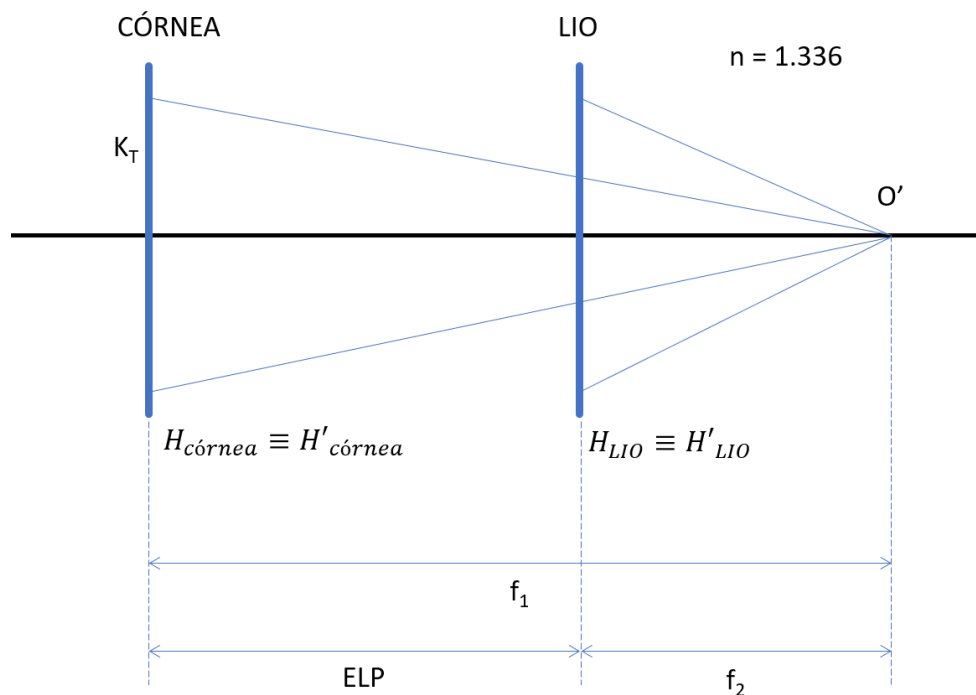


Figura 39. Esquema de equivalencia dióptrica entre córnea y lente intraocular considerando ambos sistemas ópticos como delgados y separados por la ELP y sus respectivos índices de refracción: humor acuoso (1.336) y humor vítreo (1.336).

En este caso, para satisfacer la condición de equivalencia de potencias entre los distintos planos separados por la posición efectiva de la lente se tiene que cumplir que la producción de la imagen O' sea en el mismo punto (Figura 39), pero los índices de refracción en este caso son el del humor acuoso y el del vítreo respectivamente, asumiendo los dos como iguales ya que la diferencia es únicamente de 0.001 se tiene:

$$P_{\text{córnea}} = \frac{n_{\text{h.acuoso}}}{f_1} = \frac{1.336}{f_1} \quad ; \quad f_1 = \frac{1.336}{P_{\text{córnea}}} \quad (43)$$

$$P_{L10} = \frac{n_{\text{vitreo}}}{f_2} = \frac{1.336}{f_2} \quad (44)$$

$$f_2 = f_1 - ELP \quad (45)$$

Como la equivalencia se quiere determinar en el plano del implante, se tiene que:

$$P_{L10} = \frac{1.336}{f_2} = \frac{1.336}{f_1 - ELP} = \frac{1.336}{\frac{1.336}{P_{\text{córnea}}} - ELP} \quad (46)$$

Simplificando se obtiene la expresión de equivalencia entre planos despreciando el grosor de los sistemas:

$$P_{L10} = \frac{1.336 * P_{\text{córnea}}}{1.336 - P_{\text{córnea}} * ELP} \quad (47)$$

Ya solo queda determinar la contribución de la potencia corneal total sobre el equivalente esférico en plano córnea, y para ello se procede del siguiente modo. Supóngase que la potencia total de la córnea (K_t) es 44.00 D y la ELP determinada con la Ecuación (38) de la fórmula de Haigis es de 6.23 mm. En primer lugar se calcula la equivalencia en el plano de la LIO con la suma algebraica de potencia corneal y la equivalencia entre gafa y córnea designada como (K_g), es decir, $44.00 - 0.74 = 43.26$ D:

$$P_{L10} = \frac{1.336 * 43.26}{1.336 - 43.26 * 0.00623} = 54.1921 \text{ D}$$

A continuación, se calcula la equivalencia entre córnea y plano-lente de la potencia corneal total sin la influencia del plano de la gafa:

$$P_{L10} = \frac{1.336 * 44.00}{1.336 - 44.00 * 0.00623} = 55.3584 D$$

Por último, con la diferencia entre ambos cálculos de equivalencia se obtendrá la distometría o conversión entre córnea y plano del implante del equivalente esférico con la contribución de la potencia corneal (K_t):

$$54.1921 - 55.3584 = -1.166 D \approx -1.17 D$$

En un ejercicio de simplificación se puede agrupar todo en una sola expresión matemática:

$$P_{L10} = P_{L10-Kg} - P_{L10-Kt} = \frac{1.336 * K_g}{1.336 - K_g * ELP} - \frac{1.336 * K_t}{1.336 - K_t * ELP} \quad (48)$$

$$P_{L10} = \frac{(1.336 - K_t * ELP) * 1.336 * K_g - (1.336 - K_g * ELP) * 1.336 * K_t}{(1.336 - K_g * ELP) * (1.336 - K_t * ELP)} \quad (49)$$

$$P_{L10} = \frac{1.336^2(K_g - K_t)}{(1.336 - K_g * ELP) * (1.336 - K_t * ELP)} \quad (50)$$

Teniendo en cuenta la expresión de equivalencia entre plano gafa y córnea deducida anteriormente Ecuación (41) y sabiendo que:

$$K_g = K_t + P_{córnea} \quad (51)$$

Se sustituye en la Ecuación (50) obteniéndose una expresión que relaciona la potencia del equivalente esférico en el plano de la gafa y el plano del implante intraocular, obteniéndose:

$$P_{LIO} = \frac{1.336^2 \left(\frac{P_{gafa}}{1 - d * P_{gafa}} \right)}{\left(1.336 - \left(K_t + \frac{P_{gafa}}{1 - d * P_{gafa}} \right) * ELP \right) * (1.336 - K_t * ELP)} \quad (52)$$

Donde

P_{LIO} es la potencia del equivalente esférico en el plano del implante.

P_{gafa} es la potencia del equivalente esférico en el plano de la gafa.

D es la distancia expresada en metros de la gafa al apex corneal.

K_t es la potencia media total de la córnea (superficie anterior + posterior).

ELP es la posición efectiva de la lente intraocular expresada en metros.

Por tanto y siguiendo el ejemplo que ha servido de argumento es este apartado, esto implica que un equivalente esférico postquirúrgico en plano gafa de prueba de -0.75 D equivale a -1.17 D en el plano del implante teniendo en cuenta unas características biométricas particulares de ACD, AXL y K (Figura 37).

Recuperando el objetivo primario de este apartado, que es calcular retrospectivamente la lente intraocular ideal que hubiera dejado el ojo intervenido con el CMC en retina, se necesitaba esta equivalencia para adicionarla a la potencia implantada, tal y como se expresa a continuación.

Si por ejemplo se hubiera implantado una lente de +15.00 D y se obtiene un equivalente esférico postoperatorio de -0.75 D significa que se ha hiper corregido al paciente, esto es, la potencia del implante ha sido excesiva produciendo un adelantamiento del CMC respecto a la retina. Aplicando todo lo expuesto anteriormente en este apartado y habiendo calculado la distometría entre planos, la lente perfecta (IOL_P) hubiera sido:

$$IOL_P = IOL_I + EE_POST_planoLIO$$

$$IOL_P = +15.00 + (-1.17) = +13.83 \text{ D}$$

Si se hubiera seleccionado +13.83 D como lente para implantar, el CMC hubiera estado situado en retina y el equivalente esférico postquirúrgico hubiera sido 0.00 D. Como la precisión de manufactura es de media dioptría, al menos en este estudio, habría que haber seleccionado o bien +14.00 D que es la que más se aproxima o bien +13.50 D y la diferencia respecto a la lente calculada sería el residuo de equivalente esférico postoperatorio en el plano del saco capsular. De modo que si se hubiera seleccionado +14.00 D como lente a implantar el equivalente esférico sería ligeramente negativo.

$$EE_POST_PlanoLIO = +13.83 - (+14.00) = -0.17 \text{ D}$$

Por el contrario, en caso de haber seleccionado para el implante +13.50 D, el equivalente esférico postquirúrgico sería positivo, lo que implica un CMC retrasado respecto a la retina o lo que es lo mismo, una infravaloración de la potencia de la lente intraocular, por tanto:

$$EE_POST_PlanoL10 = +13.83 - (+13.50) = +0.33 D$$

6.3 Etapa III: Implementación, selección y optimización de modelos ML.

Una vez filtrada y organizada la base de datos, analizada, seleccionadas las variables predictoras, aleatorizada, dividida y calculadas todas las equivalencias entre los distintos planos de trabajo, se está en disposición de implementar diferentes modelos ML regresivos.

Se planterán cuatro modelos distintos que contienen diferentes predictores (Tabla 5). Inicialmente la construcción de modelos se hace en crudo, sin ajustes, para seleccionar únicamente los que rinden mejor. En la siguiente fase se pasará a la hiperparametrización, para aumentar la potencia predictiva. Y en una última fase se intentarán ensamblar varios de los mejores modelos para obtener el definitivo que tendrá que ser validado posteriormente de dos maneras: con el 20% del total de la muestra, y en segunda instancia comparando con las 12 fórmulas de uso clínico común ([6.1.3 Datos para la comparación entre fórmulas](#)).

TIPOS DE MODELOS PREDICTIVOS			
MODELO 1	MODELO 2	MODELO 3	MODELO 4
AXL	AXL	AXL	AXL
ASD	ASD	ASD	ASD
WTW	WTW	WTW	WTW
KM_F	KM_T	KM_T	KM_F
R_B/F	R_B/F		KM_B

Tabla 5. Variables biométricas predictoras incluidas en los distintos modelos.

Los cuatro modelos (Tabla 5) en bruto se construirán desde cinco perspectivas distintas: las Redes Neuronales, la teoría probabilística bayesiana, las máquinas de vectores de soporte, las regresiones no lineales multivariante con ajustes polinómicos, los árboles de decisión y los Bosques Aleatorios.

Los cuatro modelos (Tabla 5) tienen en común las variables biométricas lineales: AXL, ASD y WTW, diferenciándose únicamente en las de curvatura.

De esta manera se probaron varios algoritmos basados en Bosques Aleatorios o *Random Forest* (131) con los cuatro modelos de la Tabla 5 y el tiempo de computación fue excesivo, en algún caso superior a 5 minutos, con resultados poco prometedores, al menos no tan buenos como con otras técnicas, de modo que se decidió no avanzar en esa dirección y se descartaron.

6.4 Etapa IV: Conversión del equivalente esférico capsular a plano corneal

La diferencia entre la potencia emetropizante calculada por el algoritmo y el redondeo con precisión de 0.50 D, corresponde al residuo refractivo postoperatorio en el plano del implante. Este habrá que convertirlo a su equivalencia dióptrica en el plano de la córnea que será el resultado final que verá el usuario de la fórmula.

Para ello, Holladay (130) basado en el sistema de vergencias gaussianas propone una fórmula de conversión que es la que se utilizará para hallar la equivalencia del equivalente esférico en plano capsular a plano corneal.

$$EE_{córnea} = \frac{\frac{1336}{\frac{1336}{K} - ELP} + EE_{LIO}}{\frac{1336}{1336} + ELP} - K \quad (53)$$

Donde K es la potencia queratométrica media, ELP es la posición efectiva de la lente y $EE_{córnea}$ y EE_{LIO} son el equivalente esférico en el plano de la córnea y del implante respectivamente. No se tiene en cuenta la superficie posterior ni el espesor de la córnea, por considerarse despreciable la contribución a la equivalencia con residuos refractivos tan pequeños, así que la potencia corneal está calculada con índice queratométrico 1.3375.

Respecto a la ELP en esta fórmula, los cálculos son en base a un sistema óptico formado por superficies refractivas (lente intraocular y córnea) de espesor despreciable y por tanto la distancia es de centro a centro entre superficies.

7 RESULTADOS

7.1 Resultados de la etapa I

7.1.1 Exploración descriptiva analítica y gráfica del dataset

El *dataset* con el que se trabajó consta de 481 ojos (Tabla 6) de 355 pacientes con un 45.92 % de mujeres y un 54.08 % de hombres, con una edad media de 65.33 ± 13.22 años [45,81].

En la estratificación por longitud axial hubo 114 ojos tienen una longitud axial por encima de 25.00 mm pertenecientes a un subgrupo que se llamó *Long*, 57 ojos con un eje anteroposterior por debajo de 22.00 mm correspondientes al subgrupo *Short* y 310 ojos con longitudes axiales comprendidas entre 22.01 mm y 24.99 mm pertenecientes al subgrupo *Normal*.

La Tabla 6 muestra la estadística descriptiva de todas las variables biométricas empleadas en el estudio de la muestra total y segmentadas por longitud axial.

MUESTRA TOTAL (n = 481)

	AXL	ACD	LT	WTW	ASD	KM_F(1.3375)	KM_B	R_B/F	K_T
MEAN	23,91	3,18	4,52	11,97	7,70	44,18	-6,27	82,71	42,49
SD	2,06	0,41	0,35	0,43	0,36	1,61	0,27	1,71	1,51
MAX	33,50	4,68	5,52	13,60	8,74	49,71	-5,40	89,62	47,54
MIN	19,33	1,96	3,18	10,90	6,74	38,90	-7,23	76,91	37,49

GRUPO SHORT (n = 57)

	AXL	ACD	LT	WTW	ASD	KM_F(1.3375)	KM_B	R_B/F	K_T
MEAN	21,23	2,81	4,50	11,58	7,31	45,16	-6,50	82,32	43,80
SD	0,65	0,28	0,30	0,42	0,31	1,92	0,34	1,75	1,83
MAX	21,98	3,50	5,38	12,40	8,01	49,10	-5,73	88,04	47,54
MIN	19,33	1,96	4,05	10,90	6,74	40,38	-7,23	77,88	39,25

GRUPO LONG (n = 114)

	AXL	ACD	LT	WTW	ASD	KM_F(1.3375)	KM_B	R_B/F	K_T
MEAN	26,81	3,55	4,37	12,22	7,92	43,25	-6,17	83,12	42,01
SD	1,84	0,34	0,30	0,43	0,32	1,62	0,27	1,86	1,58
MAX	33,50	4,68	4,99	13,60	8,74	47,20	-5,40	88,48	45,99
MIN	25,00	2,76	3,31	11,00	7,18	38,52	-6,93	76,91	37,49

GRUPO NORMAL (n = 310)

	AXL	ACD	LT	WTW	ASD	KM_F(1.3375)	KM_B	R_B/F	K_T
MEAN	23,32	3,11	4,57	11,95	7,68	43,72	-6,27	82,63	42,42
SD	0,74	0,35	0,36	0,37	0,32	1,32	0,23	1,63	1,28
MAX	24,98	4,09	5,52	13,20	8,44	48,10	-5,71	89,62	46,54
MIN	22,02	2,02	3,18	11,10	6,85	40,61	-7,04	78,49	39,38

Tabla 6. Estadística descriptiva de todas las variables biométricas empleadas en el estudio de la muestra total y estratificadas por longitud axial.

Todas las variables biométricas (Figura 40) muestran una correlación de proporcionalidad entre la longitud axial y la categoría a la que corresponde dicha variable, es decir, ojos axialmente cortos tienen distancias estructurales en promedio pequeñas como WTW o ACD y por el contrario ojos largos poseen en promedio estructuras más

grandes, exceptuando el espesor cristaliniano, dónde se puede comprobar que para los ojos que pertenecen al grupo *Long* la mediana y la media están por debajo de 4.50 mm, cabiendo esperar a priori tamaños promedio mayores que en ojos normales y cortos.

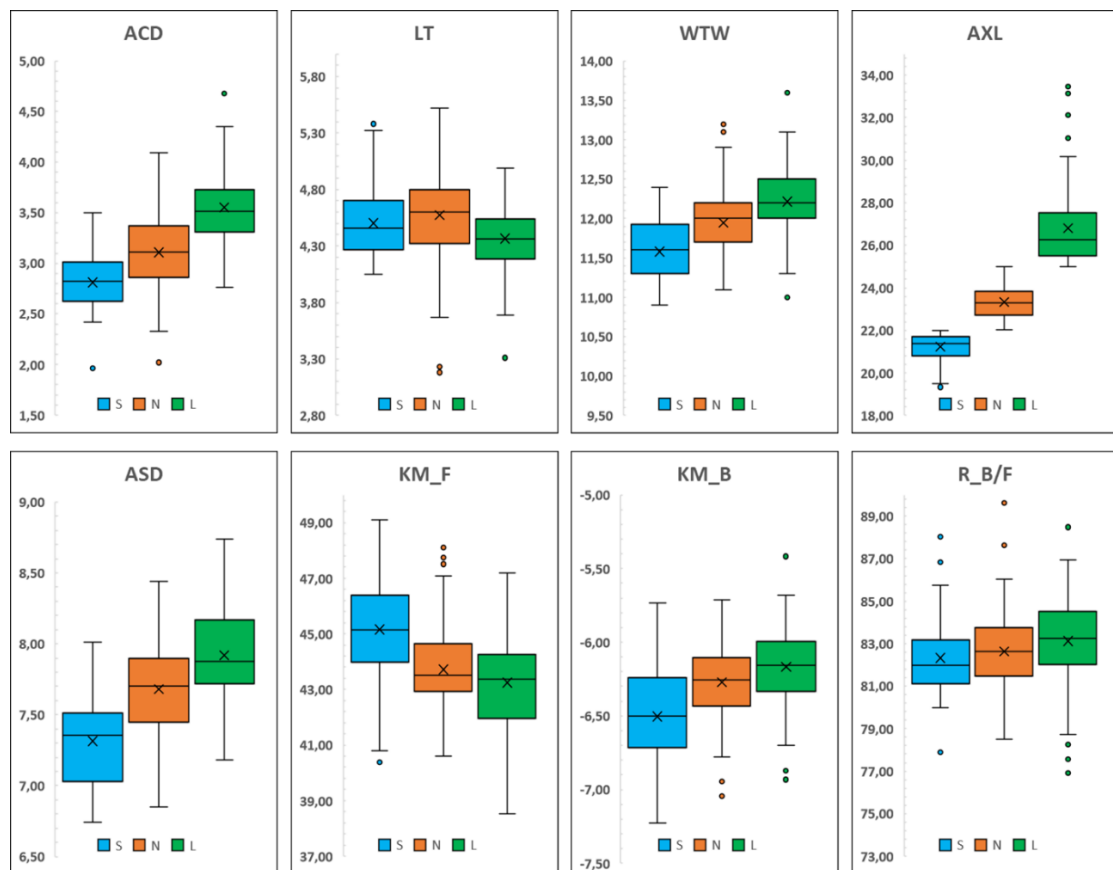


Figura 40. Diagrama de cajas y bigotes que muestra la distribución de todas las variables biométricas exploradas discriminando por tamaños oculares Short, Normal y Long respectivamente.

En cuanto a la distribución de todas las variables biométricas (Figura 41), todas ellas mostraron una distribución normal excepto en la variable AXL dónde se observó un sesgo a la izquierda y un *p* valor inferior a 0.001 con el test de *Kolmogorov-Smirnov* con corrección de *Lillefors*. Esta desviación de la normalidad, se conocía previamente debido a la acumulación artificial de ojos extremos con longitudes axiales largas por encima de 25.00 mm para darle más robustez al modelo en la predicción de esos valores.

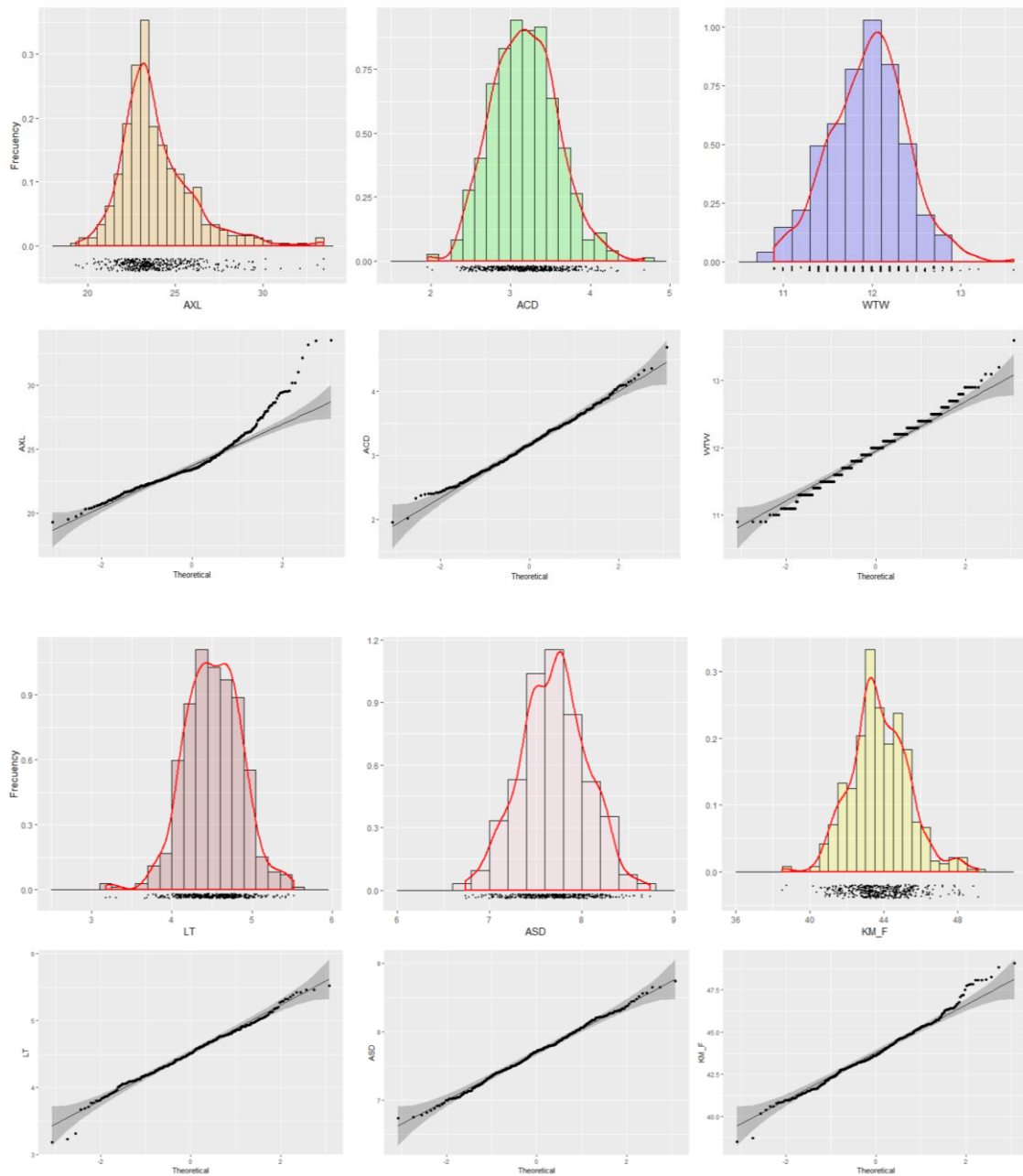


Figura 41. Histogramas y qqnorm de las variables relevantes de la base de datos.

7.1.2 Correlación entre variables y selección de características

Hay algunas variables implicadas en el modelaje que no se distribuyen normalmente según el análisis gráfico preliminar (Figura 41) mediante histogramas y qqplots. Efectivamente AXL, WTW, CTE_LIO e IOL_P son variables de distribución no

normal cuando se aplica un contraste de normalidad de *Kolmogorov-Smirnov* con corrección de *Lilliefors* ($p < 0.001$), detalle que se tendrá en cuenta en el análisis de correlaciones posteriormente.

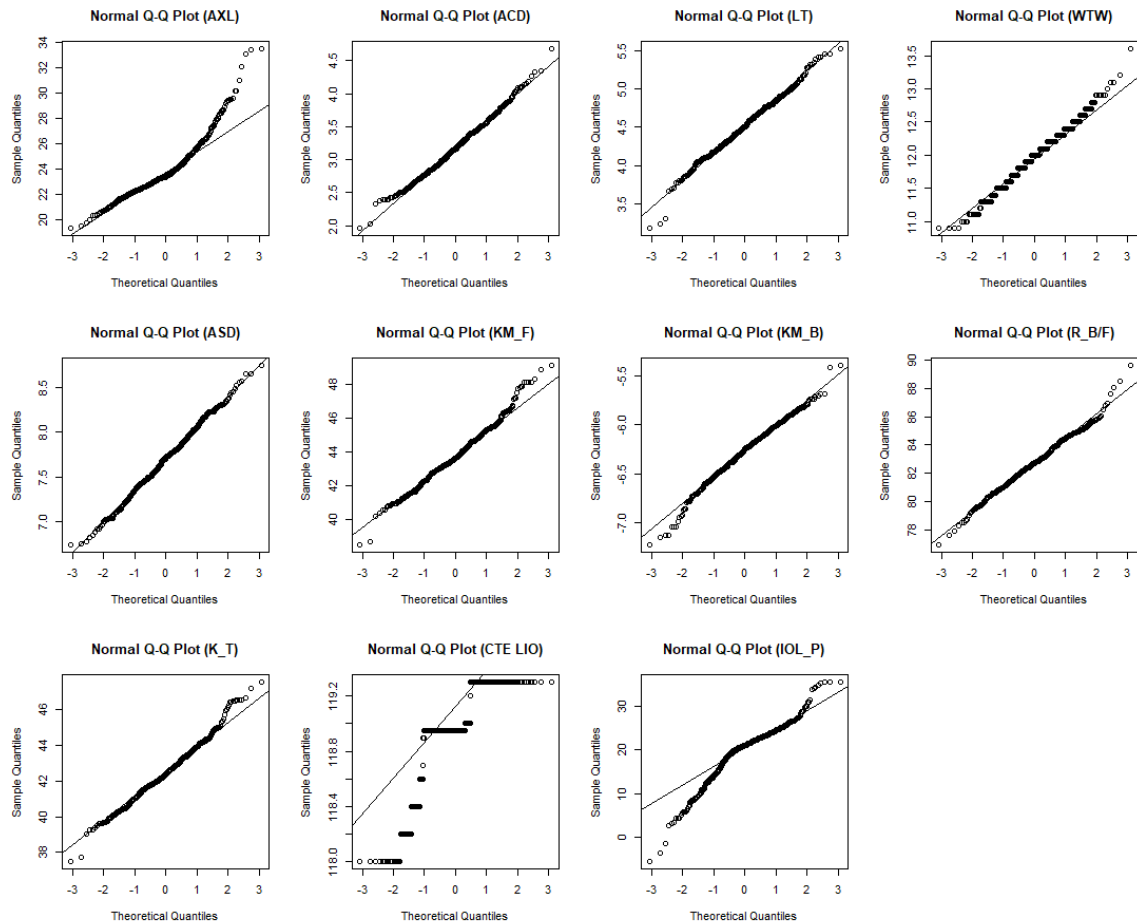


Figura 42. Gráficos de normalidad (Q-Q plot) de todas las variables de modelaje del dataset.

En un primer análisis bivariado de dependencias (Figura 43-A) se puede observar que hay algunas variables que presentan correlación entre ellas dos a dos: ACD, AXL, WTW, KT, ASD, a esto se le llama colinealidad y multicolinealidad cuando la combinación lineal es entre varios de estos predictores.

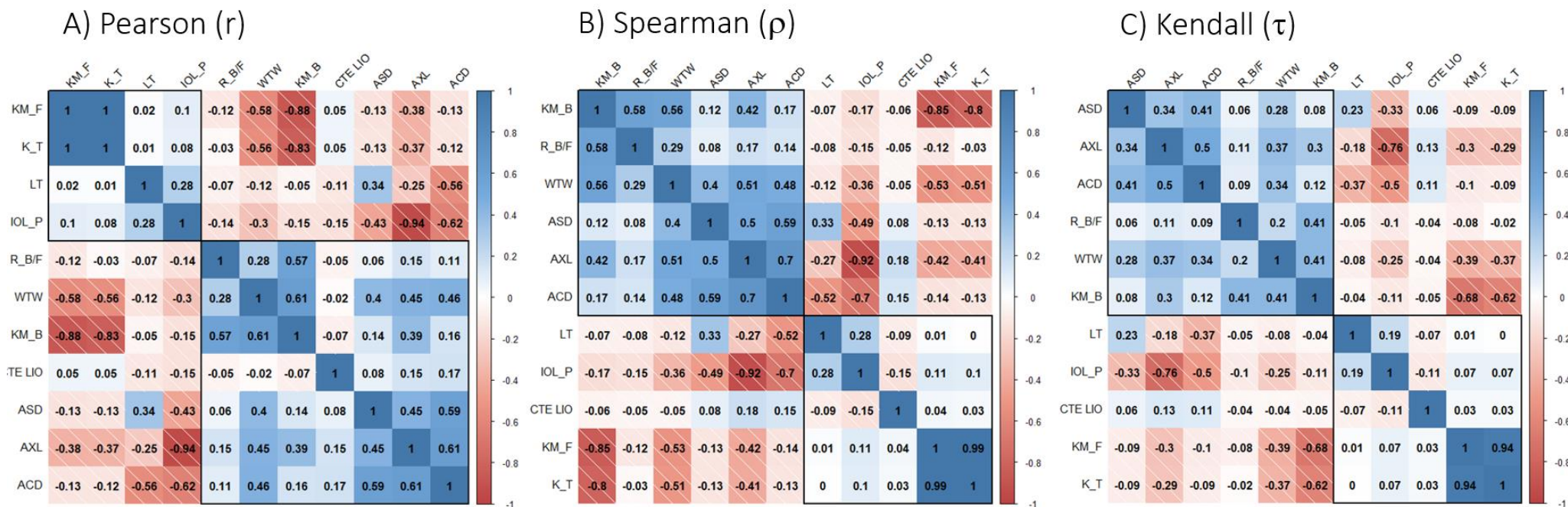


Figura 43. Heatmaps-cluster con las matrices de correlación según los diferentes coeficientes estudiados de todas las variables regresoras y de la variable respuesta (IOL_P).

Según *r de Pearson* (Figura 43-A) las variables con distribución normal más correladas son KM_B con R_B/F ($r = 0.57$, $p < 0.001$), con K_T ($r = -0.83$, $p < 0.001$) y con KM_F ($r = -0.88$, $p < 0.001$). En el caso de las variables que no siguen la normalidad, según *rho de Spearman* y *tau de Kendall* (Figura 43-B-C), WTW con KM_B ($\rho = 0.56$ y $\tau = 0.41$; $p < 0.001$), con AXL ($\rho = 0.51$ y $\tau = 0.37$; $p < 0.001$), con KM_F ($\rho = -0.53$ y $\tau = -0.39$; $p < 0.001$) y con K_T ($\rho = -0.51$ y $\tau = -0.37$; $p < 0.001$). en el caso de AXL con ASD ($\rho = 0.50$ y $\tau = 0.34$; $p < 0.001$) y con ACD ($\rho = 0.7$ y $\tau = 0.50$; $p < 0.001$).

Excepto para todas las medidas queratométricas, para el resto de variables las dependencias con otras variables no son demasiado fuertes.

De modo que si se quiere que el modelo predictor contenga información de curvatura o potencia de la cara posterior corneal, y esta era una premisa inquebrantable y diferencial respecto a lo ya existente, no se debe incluir como variables independientes las potencias o curvaturas de ambas superficies por la colinealidad que presentan y lo que se pudiera resentir el modelo. Para cumplir con este precepto, las opciones se reducen por tanto a cuatro:

- a) Construcción del modelo con potencia corneal total central (K_T) y R_B/F excluyendo a KM_F y KM_B que están altamente correladas con K_T, ($r = 1.00$, $p < 0.001$) y ($r = -0.83$, $p < 0.001$) respectivamente (Figura 44).
- b) Introducir el ratio de curvaturas corneales posterior/anterior (R_B/F) que presenta una colinealidad suave ($r = 0.57$, $p < 0.001$) con KM_B, excluyendo esta última y el resto de variables queratométricas excepto KM_F que solo tiene una despreciable relación inversa ($r = -0.12$, $p < 0.001$) (Figura 44).

- c) Otra opción sería introducir ambas variables como predictoras, K_T y R_B/F que no tienen relación ($r = -0.03$, $p < 0.001$) (Figura 44).
- d) Por último construir un modelo solo con K_T, excluyendo el resto de variables queratométricas.

Por tanto, se implementarán varios modelos en bruto que contemplen estas posibilidades remarcadas anteriormente para analizar cual rinde mejor, lo que se estudiará en la etapa III.

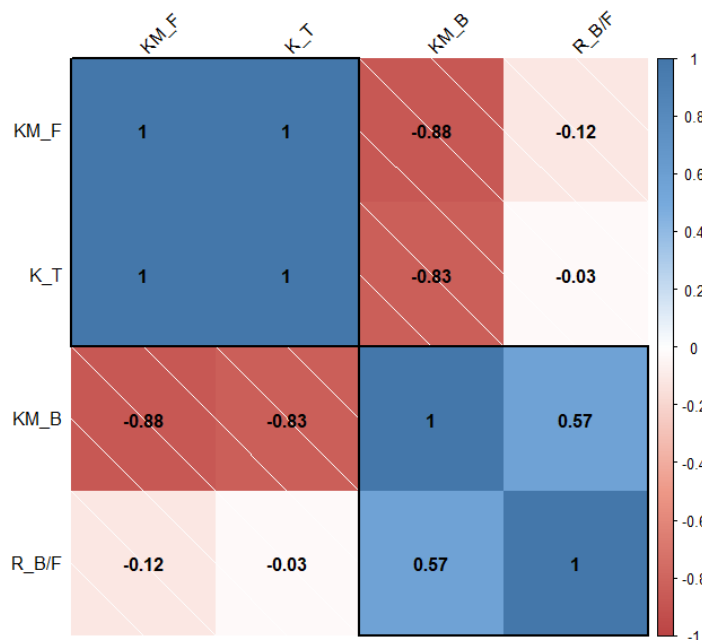


Figura 44. Heatmap-cluster con matriz de correlación en base al coeficiente r de Pearson de las variables queratométricas.

Respecto a las características anatómicas axiales del segmento anterior (Figura 45), ACD y AXL tienen una alta correlación de $\rho = 0.7$ y $\tau = 0.50$; $p < 0.001$, mientras que ACD y LT la tienen inversa con $r = -0.56$, $p < 0.001$, no presentando ninguna relación lineal AXL y LT. Evidentemente, a priori no se puede prescindir de la AXL como variable regresora ya que el peso que supone sobre la predicción del valor de lente intraocular es muy grande ($\rho = -0.92$ y $\tau = -0.76$; $p < 0.001$) (Figura 45). Del mismo modo, ACD es otra

variable fundamental en las fórmulas clásicas, pero tiene una relación inversa mucho más suave con la potencia de la lente intraocular implantada ($r = -0.62$, $p < 0.001$), por tanto es más factible prescindir de ACD mientras la información aportada al modelo de distancia al implante sea sustituida por otra variable con menos correlación como es ASD.

Precisamente por estas razones, se introdujo en el *dataset* una nueva variable predictora que no utiliza ninguna fórmula ni método que se sepa, la ASD que es una combinación lineal de otras dos variables: ACD y LT. La generación de dicho parámetro rompe la multicolinealidad existente entre AXL, ACD y la potencia perfecta de la lente implantada (IOL_P). ASD se relaciona con AXL ($r = 0.45$, $p < 0.001$) mucho menos que con ACD. ASD con WTW ($\rho = 0.40$ y $\tau = 0.28$; $p < 0.001$) se relaciona también menos que con ACD y mucho menos con la variable respuesta ($\rho = -0.49$ y $\tau = -0.33$; $p < 0.001$), condiciones ideales para implementar modelos regresivos sin perder información fundamental del segmento anterior.

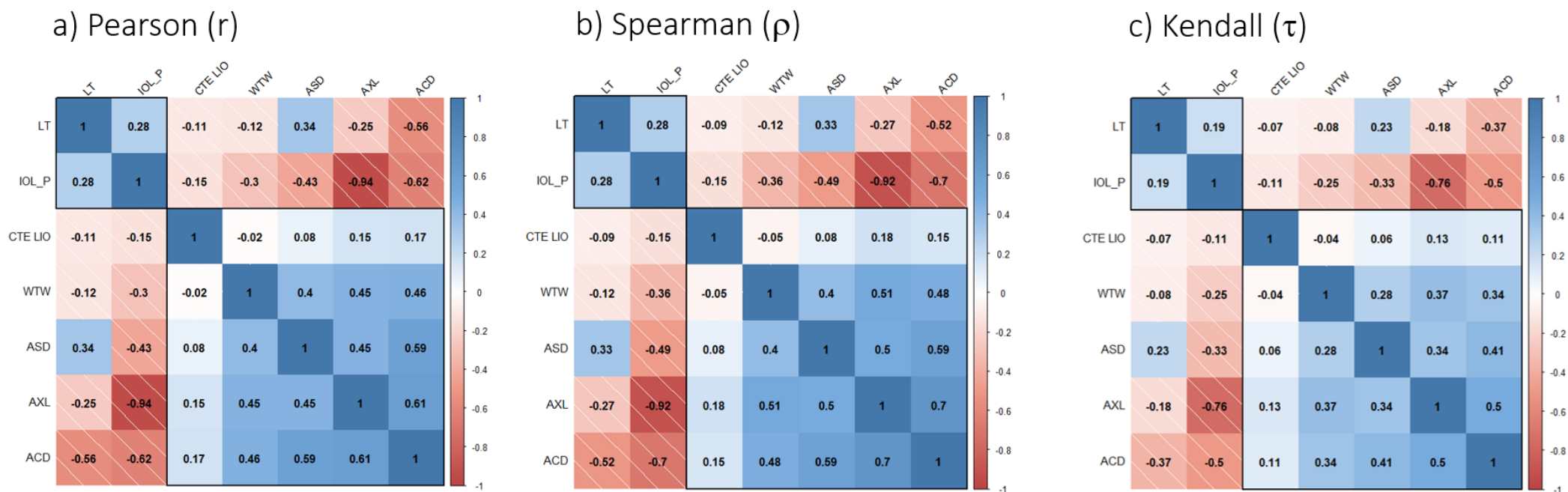


Figura 45. Heatmaps-cluster con matrices de correlación para coeficientes de Pearson, Spearman y Kendall en las variables no queratométricas.

Con todos estos cambios establecidos en las variables *inputs* del modelo relacionadas con la curvatura corneal y con la axialidad del segmento anterior, se consigue una reducción de la colinealidad sin haber perdido información esencial y necesaria para la predicción de la potencia de la lente intraocular.

En resumen, las características biométricas seleccionadas tras el análisis de los datos para el siguiente paso, la construcción de modelos, serán:

AXL, ASD, WTW, KM_T, KM_F, R_B/F, CTE_LIO y IOL_P.

7.1.3 Transformación de datos

Se aplicó una estandarización máximo-mínimo (Figura 46) que consiste en transformar los datos para que estén contenidos en el rango [0,1] mediante la siguiente fórmula:

$$X_{normalizado} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{54}$$

DATASET SIN TRANSFORMACIÓN								
	AXL	WTW	ASD	KM_B	R_B.F	K_T	CTE.LIO	IOL_P
1	25.36	12.1	8.57	-6.247007	83.04528	42.51873	118.95	14.27633
2	25.07	12.1	8.45	-6.143706	83.72006	42.20533	118.95	16.08955
3	25.52	12.3	8.08	-6.182326	84.02419	42.64738	119.30	15.34482
4	25.00	12.3	8.65	-6.451613	82.33732	43.48199	118.95	14.58649
5	26.07	12.3	8.65	-6.339144	81.94805	42.49202	118.95	12.58735
6	25.45	12.3	8.03	-6.122100	82.89379	41.58140	118.95	15.50000

DATASET CON TRANSFORMACIÓN								
	AXL	WTW	ASD	KM_B	R_B.F	K_T	CTE.LIO	IOL_P
1	0.2575482	0.1527483	0.1248490	0.007743279	0.7134620	0.3931617	0.9972338	0.1699488
2	0.2552562	0.1527483	0.1239006	0.008559713	0.7187951	0.3906848	0.9972338	0.1842795
3	0.2588128	0.1543290	0.1209763	0.008254486	0.7211988	0.3941785	1.0000000	0.1783936
4	0.2547030	0.1543290	0.1254813	0.006126184	0.7078666	0.4007748	0.9972338	0.1724001
5	0.2631597	0.1543290	0.1254813	0.007015076	0.7047901	0.3929506	0.9972338	0.1566000
6	0.2582596	0.1543290	0.1205811	0.008730477	0.7122647	0.3857536	0.9972338	0.1796200

Figura 46. Base de datos en bruto y transformada mediante estandarización.

7.2 Resultados de la etapa II

7.2.1 Obtención de la ELP

Para una ELP constante:

A priori, trabajar con una ELP constante independientemente de las características anatómicas y biométricas, resulta arriesgado e incluso impreciso.

La ELP constante se fijó en 5.23 mm porque es la que producía residuos refractivos en plano corneal idénticos en todos los rangos calculados respecto a los producidos por Barrett para la lente emetropizante. Sin embargo, a medida que la potencia en plano capsular crecía en ambos sentidos el residuo refractivo corneal se desviaba respecto al ofrecido por Barrett impuesto como referencia, incluso hasta valores de 0.16 D.

Por tanto, el sistema de fijar constante la posición efectiva se desechó como posibilidad de trabajo, únicamente útil en valores centrales, pero inexacto en posiciones extremas.

Para una ELP variable:

La otra alternativa al uso de ELP constante independientemente de las características anatómicas del ojo, es usar una posición efectiva variable en función de dichas características anatómicas. Por tanto, habrá que encontrar una ELP adecuada para cada caso anatómico sin que se resienta la distometría en términos de precisión comparando con Barrett Universal II.

Obviamente, para residuos inferiores al cuarto de dioptría en plano capsular, como ya se ha expuesto anteriormente, la ELP no supondrá mayor problema en la equivalencia en términos de precisión. Sin embargo, los valores dióptricos grandes por

encima de 1.00 D sí que se ven afectados con una ELP constante en la equivalencia distométrica.

Se trabajó sobre la hoja de cálculo explicada en el apartado metodológico (Figura 47). Se ajustó manualmente la ELP para obtener en plano corneal el mismo residuo que el que ofrecía Barrett independientemente de la tipología de ojo y del residuo refractivo en plano capsular.

AXL	ACD	Km	CTE LIO	LIO EMET.	LIO RND.	Plano LIO (D)	Refr. Postop Barrett	ELP variable	Plano corneal (D) ELP corr.	dif ELP
20,00	2,50	42,00	118,99	38,09	39,50	-1,41	-1,11	3,69	-1,11	0,00
20,00	2,50	42,00	118,99	38,09	39,00	-0,91	-0,71	3,69	-0,71	0,00
20,00	2,50	42,00	118,99	38,09	38,50	-0,41	-0,32	3,69	-0,32	0,00
20,00	2,50	42,00	118,99	38,09	37,50	0,59	0,45	3,99	0,45	0,00
20,00	2,50	42,00	118,99	38,09	38,00	0,09	0,07	3,99	0,07	0,00
20,00	2,50	42,00	118,99	38,09	37,00	1,09	0,82	4,10	0,82	0,00
20,00	2,50	42,00	118,99	38,09	36,50	1,59	1,19	4,20	1,19	0,00
22,00	3,00	42,00	118,99	28,39	30,00	-1,61	-1,24	4,00	-1,24	0,00
22,00	3,00	42,00	118,99	28,39	29,50	-1,11	-0,85	4,00	-0,85	0,00
22,00	3,00	42,00	118,99	28,39	29,00	-0,61	-0,47	4,00	-0,47	0,00
22,00	3,00	42,00	118,99	28,39	28,50	-0,11	-0,08	4,10	-0,08	0,00
22,00	3,00	42,00	118,99	28,39	28,00	0,39	0,29	4,15	0,29	0,00
22,00	3,00	42,00	118,99	28,39	27,50	0,89	0,66	4,28	0,66	0,00
22,00	3,00	42,00	118,99	28,39	27,00	1,39	1,03	4,35	1,03	0,00
24,00	3,00	42,00	118,99	21,20	22,00	-0,80	-0,60	4,36	-0,60	0,00
24,00	3,00	42,00	118,99	21,20	21,50	-0,30	-0,23	4,15	-0,23	0,00
24,00	3,00	42,00	118,99	21,20	22,50	-1,30	-0,97	4,45	-0,97	0,00
24,00	3,00	42,00	118,99	21,20	21,00	0,20	0,14	4,90	0,14	0,00
24,00	3,00	42,00	118,99	21,20	20,50	0,70	0,50	4,90	0,50	0,00
24,00	3,00	42,00	118,99	21,20	20,00	1,20	0,86	4,90	0,86	0,00
24,00	3,00	42,00	118,99	21,20	19,50	1,70	1,21	4,90	1,21	0,00
26,00	3,50	42,00	118,99	15,43	17,00	-1,57	-1,10	5,30	-1,10	0,00
26,00	3,50	42,00	118,99	15,43	16,50	-1,07	-0,74	5,40	-0,74	0,00
26,00	3,50	42,00	118,99	15,43	16,00	-0,57	-0,39	5,40	-0,39	0,00
26,00	3,50	42,00	118,99	15,43	15,50	-0,07	-0,05	5,40	-0,05	0,00
26,00	3,50	42,00	118,99	15,43	15,00	0,43	0,30	5,40	0,30	0,00
26,00	3,50	42,00	118,99	15,43	14,50	0,93	0,63	5,50	0,63	0,00

Figura 47. Hoja de cálculo Excel con los cálculos de ajuste de ELP para la obtención de los mismos resultados refractivos en plano corneal que Barrett Universal II.

En la columna ELP variable se observa como la posición efectiva de la lente varía en función del tipo de ojo para obtener un residuo corneal idéntico a Barrett ubicado en la columna contigua a la derecha. En la última columna se puede comprobar la diferencia entre el residuo corneal ofrecido por la fórmula de Barrett y el obtenido ajustando la posición efectiva en función de particularidades biométricas.

Posteriormente, con estos datos se construyó un modelo de regresión lineal multivariante para la obtención de la posición efectiva en función de todas estas variables: la longitud axial, la queratometría media, la profundidad de la cámara anterior, la potencia de lente emetropizante obtenida por el algoritmo de cálculo y el residuo refractivo en plano capsular. La ecuación regresiva obtenida con un R² ajustado de 0.985 (p < 0.001) fue:

$$ELP = -28,999 + (0,851 * AXL) + (K * 0,242) + (ACD * 0,010) + (0,161 * LIO) + (0,130 * Rx) \quad (55)$$

Finalmente se optó por utilizar esta ecuación de regresión para determinar la posición efectiva de la lente y así poder asociar residuos refractivos de predicción a todas las potencias mostradas en la tabla de resultados final incluida la lente emetropizante.



Figura 48. Pantalla de resultados de Karmona® (fórmula obtenida en esta tesis) con Barrett Universal II y con Hill-RBF v3.0.

En la Figura 48 se puede observar la comparación de resultados de fórmulas como Barrett Universal II y Hill RBF v3.0 con Karmona® que se corresponde al modelo de esta tesis. Tal y como se puede apreciar, para una misma lente emetropizante (22.42 D),

así como dos potencias por encima y dos por debajo, los residuos refractivos en plano corneal predichos asociados a dichas potencias de lentes intraoculares son prácticamente los mismos para todas las 3 fórmulas de la Figura.

7.3 Resultados de la etapa III

7.3.1 Modelos en bruto de Redes Neuronales

Se probaron los diferentes modelos construyendo una RN tipo SOM de Kohonen ya que se cuenta con los antecedentes de Hill-RBF. Los resultados no fueron nada satisfactorios (probablemente por tener una muestra 70 veces inferior a la de W. Hill), las métricas de precisión no fueron las mejores, así que se desechó la posibilidad de optimizarla. Cierto es, que los sistemas de reconocimiento de patrones como SOM necesitan ser alimentados para su aprendizaje por cantidades ingentes de datos, de manera que se desechó como vía de investigación.

Como se vió en el apartado [3.3.3 Aplicación de Inteligencia Artificial en cálculos biométricos](#), hay un antecedente en la literatura publicada de una RN con dos capas ocultas con muy buenos resultados, incluyendo hasta el 95% de los ojos en el rango de ± 0.50 D pero con una muestra sustancialmente mayor, 15.728 ojos. Fernández-Álvarez et al (65) construyeron su RN con 4 predictores en la capa de *inputs* , 5 neuronas en la primera capa oculta y 3 en la segunda, su capa de salida era bivariada (potencia de IOL y refracción postoperatoria). Lamentablemente no hay referencia de la función de activación utilizada.

Emulando a Fernández-Álvarez et al, para probar los cuatro modelos (Tabla 5) con la librería CARET se construyeron la clásica RN Perceptrón con una neurona en la capa oculta y otra RN con dos capas ocultas y variando el número de neuronas (como los autores), pero se vió que incluyendo más de una capa oculta el tiempo de computación se disparaba por encima de los 6 minutos con tasas de ocupación de la CPU y de la RAM muy altas (Figura 49), además el rendimiento no mejoraba sustancialmente al obtenido con una sola capa (Tabla 7).

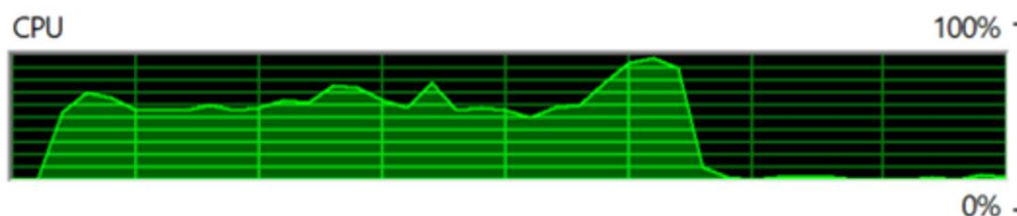


Figura 49. Rendimiento de la CPU en pleno proceso de aprendizaje de una RN, con tasas de ocupación de hasta el 100%.

Consumos de tiempos de ejecución tan prolongados para modelos tan demandantes y exigentes, penalizan la elección del algoritmo, en sí mismo obviamente no es un problema de precisión pero afecta claramente a su rendimiento. Se paralelizó el procesamiento con este algoritmo reduciéndose los tiempos computacionales drásticamente.

Se decidió utilizar una librería MxNet (<https://mxnet.apache.org/>) por ser una librería muy potente y versátil para poder manipular las redes e hiperparametrizarlas adecuadamente, aún teniendo dependencia con CARET mostró mucho mejor rendimiento en la construcción de RRNN.

Probando varias configuraciones para nuevas RRNN con MxNet, se consiguió mejorar el tiempo de computación ostensiblemente respecto a las RN construidas con

CARET, reduciéndolo hasta tiempos inferiores a los 5 segundos. Para el modelo 4, la mejor RN fue con tres neuronas en la única capa oculta, con funciones de activación SoftReLu y *back propagation* para el entrenamiento (Figura 50).

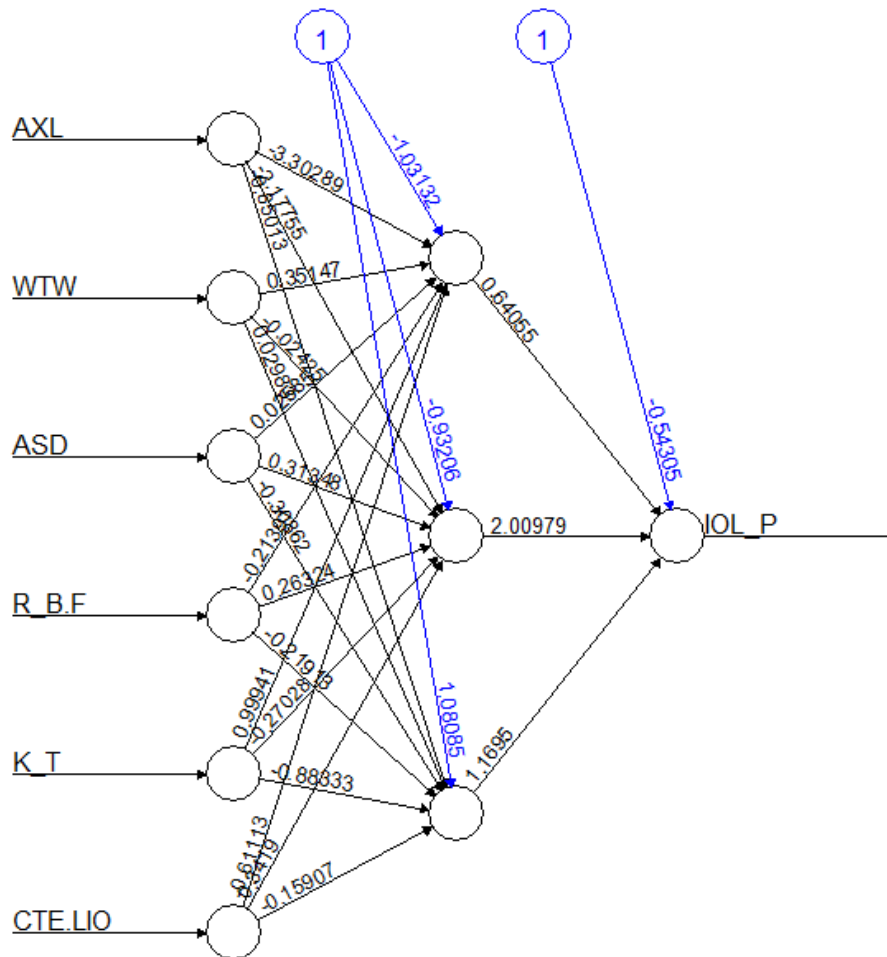


Figura 50. Esquema de red neuronal para modelo 4 con 3 neuronas en capa oculta y una salida para la potencia de la lente intraocular. Funciones de activación SoftReLu.

De todos los modelos propuestos, el que mejores métricas presentó fue el modelo 1 para la RN construida con MxNet y el modelo 3 para los perceptrones (Tabla 7). Teniendo en cuenta estos resultados, se trabajará sobre la RN construida con MxNet y la monocapa de CARET, desechando la RN con dos capas ocultas implementadas con CARET por sus excesivos tiempos de computación, a pesar de los buenos rendimientos.

Modelo	Hidden layers	RMSE	Rsquared	MAE	Time computation
RN-MxNet	1	0.8575414	0.9776920	0.6111457	4.72 seg
Perceptrón	2	0.5856949	0.9895204	0.4736063	> 6min
Perceptrón	1	0.5936281	0.9892858	0.4744041	1.26 min

Tabla 7. Métricas de rendimiento para el modelo 4 implementado en tres redes neuronales: CARET+MxNet (4.72 seg), CARET+Monmlp (Perceptrón) 2 capas (> 6min) y CARET+Monmlp (Perceptrón) 1 capa (1.26 min).

7.3.2 Modelos en bruto Bayesianos

Se implementaron varios modelos con enfoque bayesiano utilizando la librería CARET sobre los cuatro modelos base. En primer lugar se construyó el más sencillo de todos, un modelo lineal generalizado o *Generalized Linear Models* (GLM) (132) al que se le llamó BAY-GLM. Estos modelos permiten hacer análisis análogos a los de los modelos lineales clásicos, pero sin presuponer normalidad en los errores, basta que estos pertenezcan a la familia exponencial de distribuciones. De los cuatro modelos probados el mejor de todos fue el modelo 2 para BAY-GLM (Tabla 8).

Otro de los modelos implementados fue *Bayesian Regularized Artificial Neural Networks* (BRANNs) (133) que aunque es una red neuronal, se introduce en esta sección debido a cómo este método de aprendizaje bayesiano se basa en la optimización de la función de densidad de probabilidad de los hiperparámetros α y β , los cuales explican la varianza del error, que se expresa como la diferencia entre la salida esperada y la salida estimada de la red neuronal. Además de esto, BRANNs utiliza el algoritmo de Nguyen-Widrow (134) para mejorar el sistema de elección de los *weights* iniciales en las conexiones de la RN lo que acelera la convergencia en el proceso de entrenamiento. Y

también se usa el algoritmo de Gauss-Newton para resolver el problema no lineal de mínimos cuadrados. El modelo que mejor rinde bajo estas condiciones es el modelo 2 que tiene como predictores la queratometría total media y el ratio de curvaturas posterior y anterior, al que se llamó BRANN con 2 neuronas en una única capa oculta.

Se construyó otro modelo algo más complejo que los anteriores basado en un algoritmo llamado árboles aditivos de regresión bayesiana o *Bayesian Additive Regression Trees* (BART) utilizando la librería *bartMachine* (135). De hecho es el mismo algoritmo central en el que se basa la Fórmula FullMonte pero con una abordaje de concepción experimental diferente.

Los árboles de regresión y/o clasificación fueron propuestos por Breiman (136), y esencialmente se trata de ramificaciones dicotómicas anidadas en forma de árbol invertido con reglas de decisión sobre las covariables para determinar la variable respuesta. Dichos árboles conforman el esqueleto de los modelos BART, porque la suma de estos árboles de decisión generan un modelo aditivo con múltiples componentes multivariante que de forma natural puede incorporar las interacciones entre covariables. Ahora bien, la manera o el método de ensamblaje por el que varios árboles se suman generando una combinación lineal entre ellos, es el promedio del modelo bayesiano aplicado a posteriori que surge de un modelo bayesiano de un solo árbol. Tal promedio de modelos usa probabilidades a posteriori como pesos para promediar las predicciones de árboles individuales. Para ajustar el modelo de suma de árboles, BART utiliza una versión personalizada de cadenas de *Markov Monte Carlo*.

Se implementaron los cuatro modelos BART y el que mejor rindió fue el modelo 1 (Tabla 8) pero con tiempos de computación extremadamente largos, incluso paralelizando el proceso de ejecución del algoritmo.

Modelo	RMSE	Rsquared	MAE	Time computation
BAY-GLM	0.9458957	0.9762417	0.6796613	1.24 seg
BRANN	0.5862493	0.9895926	0.4771280	6.36 seg
BART	0.7962113	0.9817613	0.5788335	132 seg

Tabla 8. Resultados de rendimiento de modelos bayesianos.

7.3.3 Modelos en bruto con Máquinas de Vector de Soporte

Como en el *dataset* de la tesis, se tiene un espacio de entrada multidimensional para los cuatro modelos diseñados y tal como se vió en el apartado [7.1.1 Exploración descriptiva analítica y gráfica del dataset](#), los datos no son de manera estricta linealmente separables, aún así los kernel empleados para la implementación de modelos SVM fueron de tres tipos: lineal, radial y polinómico.

El mejor modelo fue el modelo 4 con queratometría simulada media y queratometría media posterior, cuyos resultados en bruto (Tabla 9) no fueron nada malos con tiempos de computación bastante contenidos y métricas muy interesantes, sobre todo con el kernel polinómico.

Modelo	RMSE	Rsquared	MAE	Time computation
SVM-LIN	0.966723	0.9771502	0.6760511	0.26 seg
SVM-RBF	1.449547	0.9379153	0.8207383	3.84 seg
SVM-POLY	0.5963513	0.9894152	0.4746620	26.26 seg

Tabla 9. Resultados de los modelos SVM con diferentes kernel (lineal, radial y polinómico).

7.3.4 Modelos en bruto con Splines adaptativos

Se construyeron cuatro modelos en crudo de regresión multivariante con *splines*: BAG-SPLINE, GCV-SPLINE, BCV-SPLINE y SPLINE para los cuatro modelos propuestos.

En el primero de ellos (BAG-SPLINE) se utilizó una técnica de reducción de la varianza llamada *Bagging* o *Bootstrap Aggregating* creada por Breinan (137). El *bagging* consiste en subdividir la partición de entrenamiento en pequeños dataset aleatorios con reemplazamiento llamados *bags*, para generar en este caso varios modelos de regresión multivariante con splines y ponderar la media resultante de la variable independiente de cada uno. El modelo 2 fue el que mejor rindió de los cuatro probados en esta categoría (Tabla 10).

Sin embargo, el modelo GCV-SPLINE utiliza otra técnica llamada *Generalized Cross Validation* (GCV). Utilizada como técnica de regularización, ya que como se refirió anteriormente se utilizará para determinar el número óptimo de puntos de corte o *knots*. El término *Cross-Validation* abarca distintas estrategias que permiten estimar el ratio del error sobre la partición *test*. Para ello se excluyen una serie de observaciones del *dataset*

de entrenamiento disponible, se ajusta el modelo y finalmente se evalúa con los datos excluidos (Figura 69). En este caso fue también el modelo 2 el que mejores métricas de rendimiento ofreció respecto al resto de los modelos.

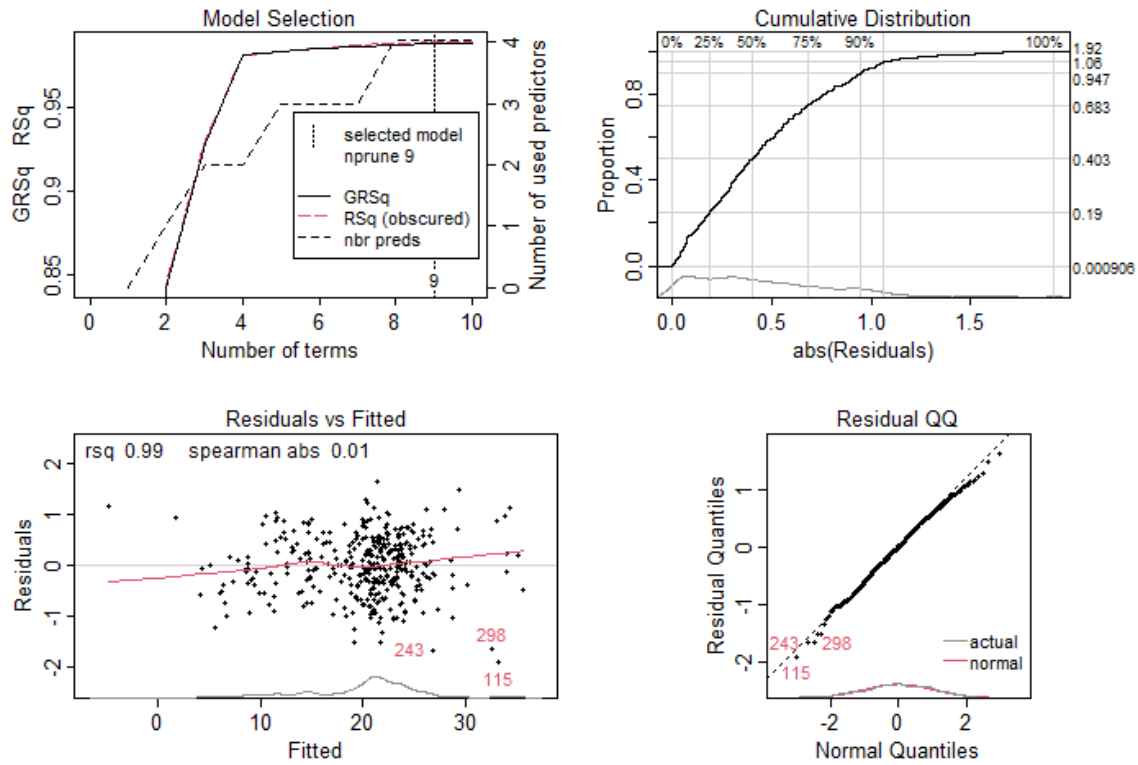


Figura 51. Gráficos resultantes del proceso de regularización mediante cross-validation en el modelo GCV-SPLINE implementado sobre el modelo 1 con predictores corneales KM_F y R_B/F.

En el modelo BCV-SPLINE se mezclan los dos conceptos de los modelos anteriores, el *bagging* como técnica de ensamblaje y mejoramiento, y la *cross-validation* para estimación del error y optimización de *knots*, siendo este sistema el que mejor rindió sobre el modelo 1 (Tabla 10).

Y por último, el modelo SPLINE es el más sencillo ya que no se aplicó ninguna técnica de mejoramiento, simplemente se implementó el modelo en bruto. Las métricas de rendimiento en general fueron muy buenas aunque los tiempos de computación con

las técnicas de *bagging* se disparaban un poco respecto al resto (Tabla 10). En este caso fue el modelo 1 el que mejor funcionó con este algoritmo de ML.

Modelo	RMSE	Rsquared	MAE	Time computation
BAG-SPLINE	0.6168589	0.9885993	0.4879891	48.84 seg
GCV_SPLINE	0.6287931	0.9879977	0.5028037	1.87 seg
BCV_SPLINE	0.6148159	0.9886721	0.4866504	16.76 seg
SPLINE	0.6295215	0.9879865	0.5038868	3.59 seg

Tabla 10. Resultados para los modelos de regresión adaptativa con splines.

7.3.5 Modelos basados en árboles de decisión

Por último tras haber construido multitud de modelos desde muchas perspectivas, en este último apartado se probarán los árboles de decisión como abordaje final. Los árboles de regresión y/o de clasificación son algoritmos no paramétricos propuestos por Leo Breiman.

Existen muchas metodologías para construir árboles de decisión, pero la más conocida es el algoritmo *Classification And Regression Trees* (CART). Un árbol de decisión básico divide los datos de entrenamiento en subgrupos homogéneos, es decir, grupos con valores respuesta similares y luego ajusta una constante simple en cada subgrupo (por ejemplo, la media de los valores de respuesta dentro del grupo para la regresión) (Figura 52). Los subgrupos, también llamados nodos, se forman de forma recursiva utilizando particiones binarias formadas haciendo preguntas simples de sí o no sobre cada característica (por ejemplo, ¿AXL < 22.33 mm?). Esto se hace varias veces hasta que se

cumple un criterio de detención adecuado. Después de que se haya realizado toda la partición, el modelo predice el resultado en función de:

- o Los valores de respuesta promedio para todas las observaciones que caen en ese subgrupo (problema de regresión).
- o La clase que tiene representación mayoritaria (problema de clasificación). Para la clasificación, las probabilidades predichas se pueden obtener utilizando la proporción de cada clase dentro del subgrupo.

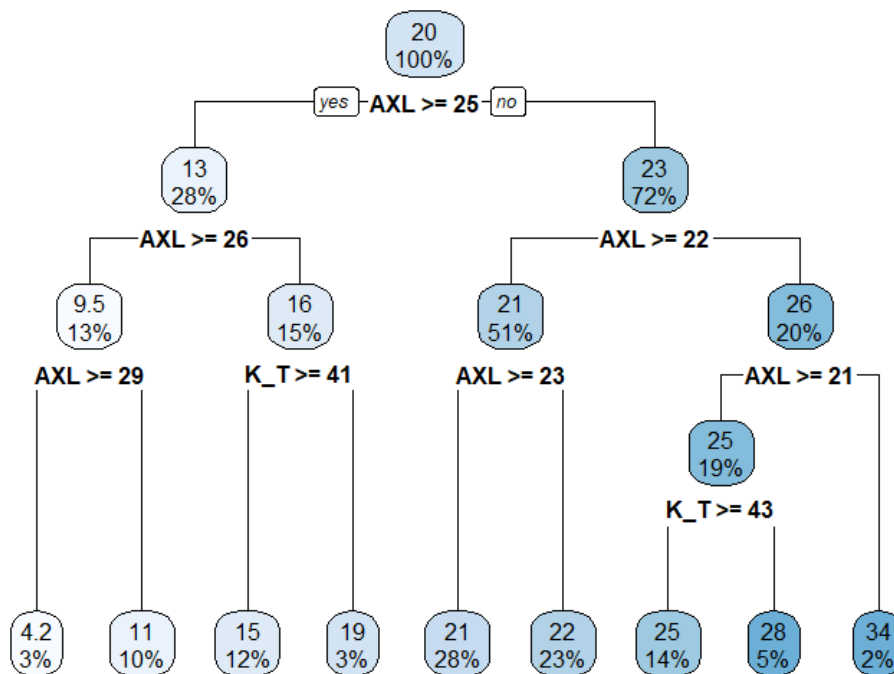


Figura 52. Árbol de decisión (CART) para el dataset original de la tesis.

Las predicciones se obtienen ajustando un modelo más simple (por ejemplo, una constante como valor de respuesta promedio) en cada región. Tales métodos se llaman “divide y vencerás” y pueden producir reglas simples que son fáciles de interpretar y visualizar con diagramas de árbol (Figura 52). Los árboles de decisión generalmente no

tienen tanto rendimiento predictivo en comparación con algoritmos más complejos como RRNN y MARS.

Se implementaron varios tipos de árboles de decisión: *bagged CART*, *boosted tree* y *conditional inference tree* entre otros, mostrando todos rendimientos muy pobres y bastante más bajos que los vistos anteriormente con otros enfoques, además de ser algoritmos muy perezosos.

Solo se encontró un algoritmo llamado *Cubist* (138) que ofreció buenas métricas y tiempo de ejecución (Tabla 11) ajustándose a lo que se exigía como mínimo para pasar la criba del modelaje en crudo y posterior hiperparametrización.

Para maximizar la interpretabilidad, los modelos *Cubist* generan los llamados *committees* que son agrupamientos de datos que cumplen ciertas condiciones, estos conformarán los submodelos. Cada *committee* se expresa como colecciones de reglas que pueden acumular varios condicionantes del tipo (si $AXL < 22.00$ y $KM \geq 44.00$ entonces), y cada regla tiene un modelo de regresión lineal multivariado asociado. Siempre que una situación coincide con las condiciones de una regla, el modelo asociado se utiliza para calcular el valor predicho. El algoritmo *Cubist* incluye un KNN para extraer la conclusión.

Por tanto, la importancia de variables se medirá teniendo en cuenta cuántas veces participa una variable predictora en las regresiones multivariantes establecidas al final de cada submodelo y de cada grupo de reglas.

Un problema que se encontró en la presentación de resultados para el modelo *Cubist*, es que al ser un modelo minorista y poco conocido no tiene implementadas librerías específicas para mostrar gráficos resumen de resultados, y por tanto se hace

difícil graficar tablas resumen del conjunto de submodelos, reglas y regresiones obtenidos.

El modelo que mejor desempeño mostró de los cuatro que se probaron sobre el algoritmo de decisión *Cubist* fue el modelo 1.

Modelo	RMSE	Rsquared	MAE	Time computation
CUB-TREE	0.6277091	0.9884701	0.4858134	17.28 seg

Tabla 11. Resultado del modelo de árbol de decisión *Cubist* (CUB-TREE).

7.3.6 Selección de los mejores modelos en bruto

Se hicieron muchas pruebas de construcción de modelos que no vieron la luz debido a su falta de rendimiento. Multitud de intentos se desecharon ante la inoperancia computacional, la poca precisión u otros factores, en total 50 tipos distintos de tecnologías y algoritmos fueron probados abarcando todas las estrategias posibles como se vió anteriormente. Pero solo 14 (Tabla 12) son los que pasarán al siguiente nivel, la optimización, gracias a su alto rendimiento.

En lo sucesivo se optimizarán los modelos cambiando y ajustando ciertos hiperparámetros (*knots*, capas ocultas, grados polinómicos, algoritmos de pesos, etc.) de los seleccionados para incrementar su rendimiento. Todos ellos tendrán que pasar una auditoría interna de validación con la partición del dataset (20%) dedicada a esto, desechando los modelos hiperparametrizados que produzcan *overfitting*.

Por último, una vez seleccionados los mejores modelos, se intentará mediante alguna de las técnicas disponibles, ensamblarlos para generar uno nuevo que potencie su predictibilidad aún más si fuera posible.

Modelo	RMSE	Rsquared	MAE	Time computation
RN-MxNET	0.8575414	0.9776920	0.6111457	4.72 seg
PERCEPTRÓN 1C	0.5936281	0.9892858	0.4744041	1.26 min
PERCEPTRÓN 2C	0.5856949	0.9895204	0.4736063	> 6min
BAG-SPLINE	0.6168589	0.9885993	0.4879891	48.84 seg
GCV_SPLINE	0.6287931	0.9879977	0.5028037	1.87 seg
BCV_SPLINE	0.6148159	0.9886721	0.4866504	16.76 seg
SPLINE	0.6295215	0.9879865	0.5038868	3.59 seg

Modelo	RMSE	Rsquared	MAE	Time computation
BAY-GLM	0.9458957	0.9762417	0.6796613	1.24 seg
BRANN	0.5862493	0.9895926	0.4771280	6.36 seg
BART	0.7962113	0.9817613	0.5788335	132 seg
SVM-LIN	0.966723	0.9771502	0.6760511	0.26 seg
SVM-RBF	1.449547	0.9379153	0.8207383	3.84 seg
SVM-POLY	0.5963513	0.9894152	0.4746620	26.26 seg
CUB-TREE	0.6277091	0.9884701	0.4858134	17.28 seg

Tabla 12. Resultados de todos los modelos en bruto implementados con diferentes técnicas mostrando todas las métricas y el tiempo de computación consumido. Los MAE obtenidos por los modelos en bruto son en el plano óptico de la lente intraocular.

Dos de los mejores modelos que mejor desempeño tienen (Tabla 12) son la RN Perceptrón de una y dos capas ocultas respectivamente (PERCEPTRÓN 1C y 2C), sin embargo el consumo de tiempo de computación superior al minuto y por encima de los seis respectivamente, se presupone excesivo para obtener la potencia de la lente emetropizante.

En este sentido y tal como se expresó en el apartado [3.4.8 Interpretabilidad de modelos complejos](#), se intentó paralelizar el proceso e introducir más cores de procesamiento consiguiéndose mejores resultados de tiempo, mejorando notablemente el tiempo de cálculo reduciéndose hasta en un 60%. No obstante, se intentó sin éxito hiperparametrizar para exprimir el máximo rendimiento de ellos, por tanto se desechó la posibilidad de utilizar RN perceptrón para modelizar un método definitivo.

De los tres modelos construidos con la tecnología SVM, los que tienen kernel radial y lineal (SVM-RBF y SVM-LIN) no se seleccionaron para pasar al siguiente nivel, ya que igualmente que con los Perceptrones, se intentaron optimizar sin una mejoría sustancial.

Del mismo modo, los modelos bayesianos BAY-GLM y BART con MAEs excesivamente altos para lo que se busca y un tiempo de procesamiento superior a los 2 minutos en el segundo caso, también se desecharon para la hiperparametrización.

Por el contrario, los modelos implementados con técnicas regresivas multivariantes con *splines* funcionan muy bien, con tiempos de computación aceptables. El modelo SVM-POLY, la RN bayesiana (BRANN) y el árbol de decisión (CUB-TREE) fueron los que mostraron mejores rendimientos y por tanto son los que se seleccionaron para el proceso de optimización y mejora.

Mención especial a los modelos extra gradient boost (XGBOOST) que se han puesto de moda entre los científicos de datos en los últimos años en las competiciones de plataformas como Netflix (www.netflixprize.com), Kaggle (www.kaggle.com), DrivenData (www.drivendata.org) o por ejemplo Crowdanalytix (www.crowdanalytix.com).

Se probó este tipo de algoritmo también en el modelaje en bruto, no consiguiéndose el rendimiento esperado. En los entrenamientos, las métricas fueron muy buenas, de hecho las mejores, consiguiéndose un espectacular MAE de 0.05 D en plano capsular, sinónimo de probable sobreajuste. Efectivamente, en la predicción posterior sobre la partición *test* de validación tanto el MAE como el RMSE y el R^2 empeoraron sustancialmente (MAE = 0.52 D), poniendo de manifiesto un gran *overfitting* en el entrenamiento (Figura 53). No obstante, se intentó corregir sin éxito disminuyendo el número de predictores y optimizando la parametrización del modelaje en el entrenamiento, por lo que se desechó la posibilidad de utilizar XGBOOST para predecir la potencia de la lente emetropizante.

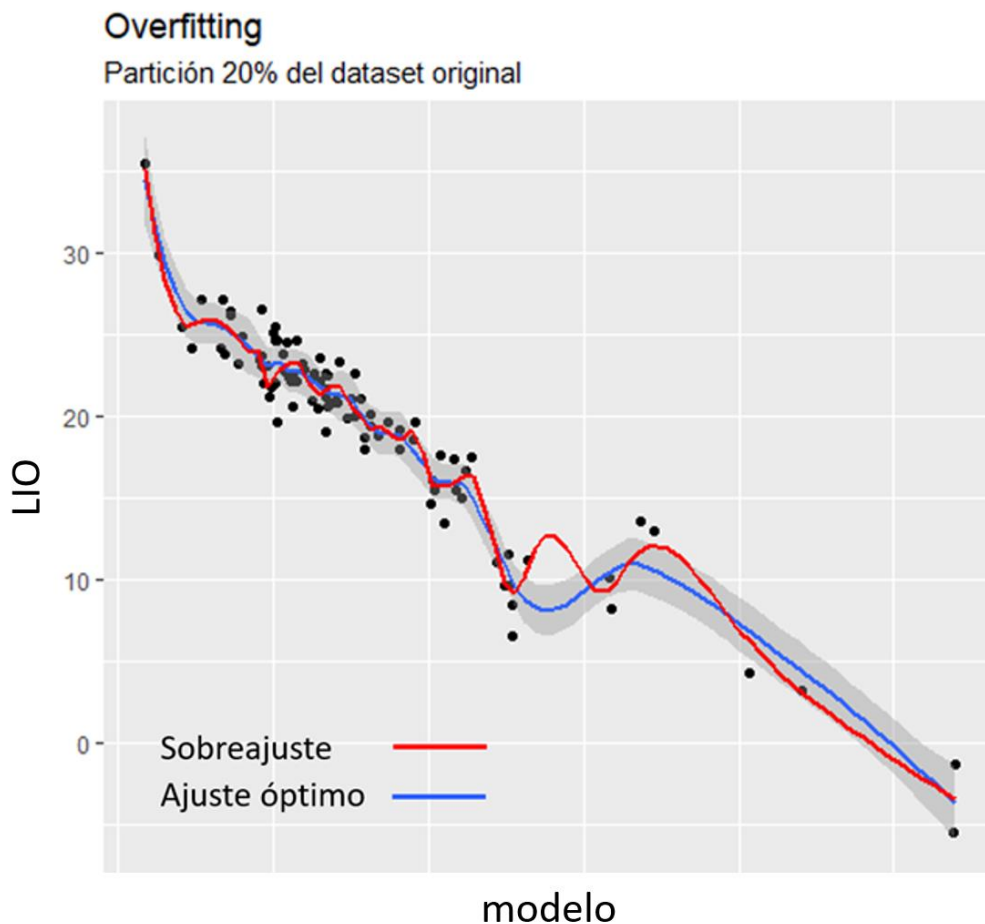


Figura 53. Predicción con XGBOOST sobre la muestra test (20%) poniendo de manifiesto un overfitting.

7.3.7 Hiperparametrización: optimización de los mejores modelos en bruto

El mejor de los modelos en crudo ofreció un MAE en plano capsular de 0.47 D, lo que implica por distometría aproximadamente 0.38 D en plano corneal. De manera que se estableció arbitrariamente este punto de corte MAE capsular para seleccionar los mejores modelos hiperparametrizados. Aquellos modelos que estén por encima de este valor se desecharán y no pasarán al siguiente nivel.

Para obtener los parámetros definitivos de optimización de cada modelo matemático habitualmente se establece un *grid* de búsqueda, en el que se definen todos los parámetros necesarios con unos rangos máximos y mínimos para que en el proceso

de aprendizaje se combinen todos entre sí y encontrar las mejores combinaciones, que ofrezcan las mejores métricas maximizando la función modelada. Este es un proceso laborioso, complejo y gravoso en tiempo.

Una estrategia alternativa, automática y mucho más rápida para encontrar los parámetros óptimos de modelaje, o al menos los de partida para poder manipularlos después, es utilizar la librería CARET que mediante validación cruzada permite hallar dichos parámetros de *tuning*.

Se optó por esta segunda fórmula, consiguiendo unos valores iniciales de dichos parámetros (Figura 54). No obstante, una vez obtenidos se fueron modificando y ajustando manualmente para potenciar el rendimiento predictivo y conseguir mejores resultados.

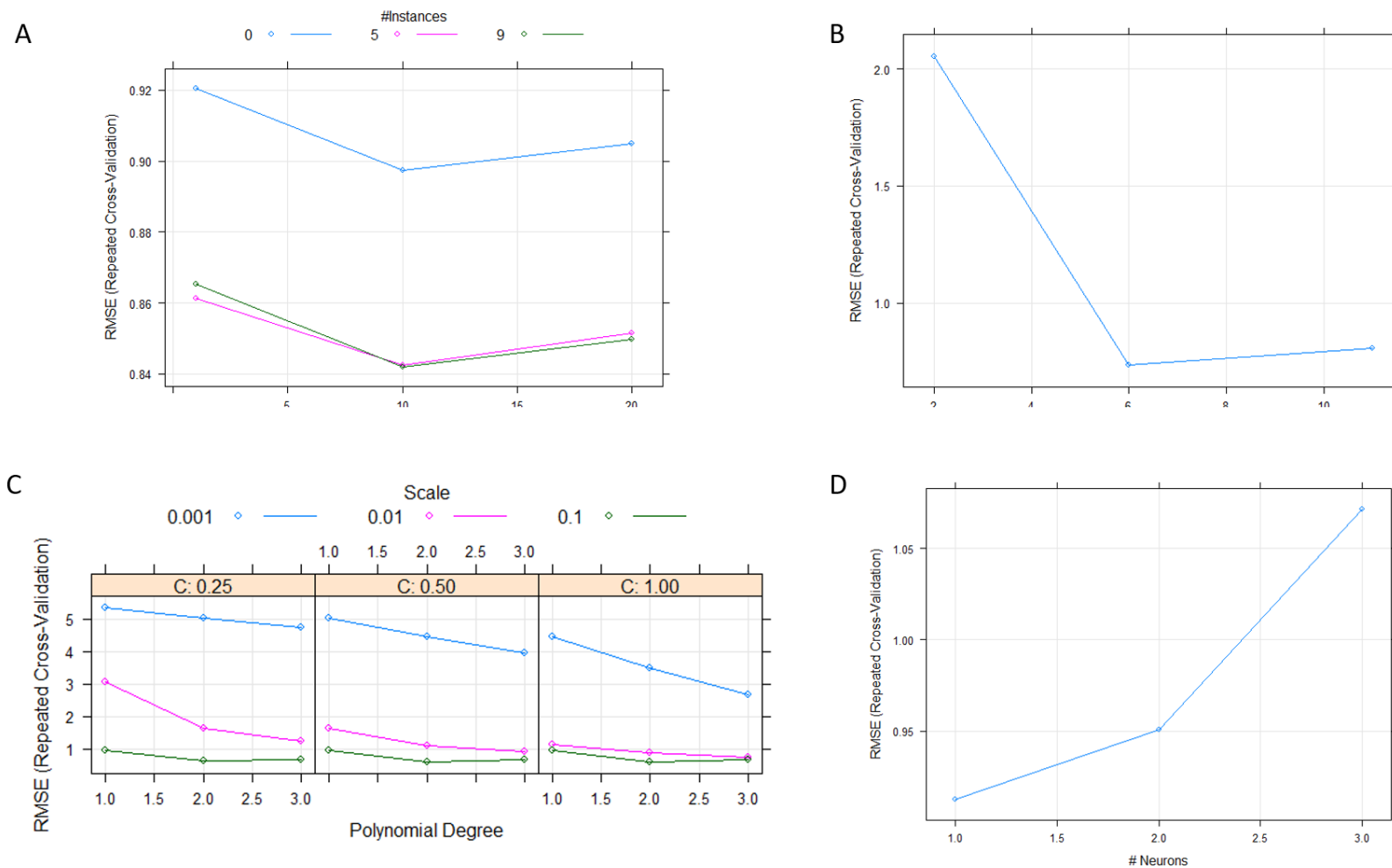


Figura 54. Optimización de parámetros para los cuatro modelos finales seleccionados. A) modelo CUB-TREE: committees = 10. B) modelo BAG-SPLINE: Prune = 6. C) modelo SVM-POLY: Scale = 0.1, degree = 2, C = 0.5. D) modelo BRNN: neurons = 2.

Los únicos modelos que rindieron un MAE capsular por debajo de 0.38 D (Tabla 13) tras la optimización fueron:

SVM-POLY : *Support Vector Machine* con Kernel polinómico de segundo grado sobre el modelo 4. Algoritmo construido y optimizado con Librería CARET.

CUB-TREE: árbol de decisión *Cubist* con 5 submodelos (*committees*) sobre el modelo 1 y un $k = 2$ (*neighbours*) para el KNN de predicción. Algoritmo construido y optimizado con Librería CUBIST.

BRANN: red neuronal bayesiana con dos neuronas en la capa oculta sobre el modelo 2. Algoritmo construido y optimizado con Librería BRNN.

BAG-SPLINE: *Multivariate Adaptive Regression Splines* con regularizado de varianza mediante *bagging* y 6 *splines* sobre el modelo 2. Algoritmo construido y optimizado con Librería CARET.

Así que los demás se descartaron. Véase que hay dos modelos que destacan por encima de los demás (Tabla 13), la red neuronal bayesiana (BRANN) y el árbol de decisión (CUB-TREE).

MÉTRICA	SVM-POLY	BAG-SPLINE	BRANN	CUB-TREE
MAE	0.3715	0.3844	0.3486	0.3572
R ²	0.9937	0.9936	0.9949	0.9946

Tabla 13. Métricas de los modelos optimizados que rinden por debajo de 0.38 D en plano capsular.

Por tanto, estos cuatro modelos son los candidatos perfectos para la siguiente fase y última, el ensamblaje.

7.3.8 Ensamblaje de los mejores modelos optimizados

El ensamblaje o combinación de modelos (*ensemble methods*) son una serie de técnicas (*voting, bagging, boosting y stacking*) utilizadas en ML para incrementar la potencia predictiva de modelos unitarios cuando se suman sus capacidades.

Se optó por la técnica denominada *stacking* (139) también conocida hoy en día como *Super Learner* (140), ya que los modelos que iban a conformar el conjunto de ensamblaje eran de distinto tipo, característica típica esta para desarrollar estas combinaciones. Es mejor ensamblar modelos procedentes de distintos algoritmos y con distintas variables regresoras.

El ensamblaje de modelos cuenta con tres fases. La primera es la selección de los *base learners* que se corresponden a los mejores modelos entrenados independientemente y con algoritmos diferentes, seleccionados en el apartado anterior (Tabla 13). En segundo lugar, se recopilan las predicciones mediante validación cruzada de cada uno de los modelos originales para entrenar el *ensemble* con un algoritmo específico que no tiene porque ser igual que en los modelos base. Y por último se generan las predicciones nuevas con el modelo ensamblado final.

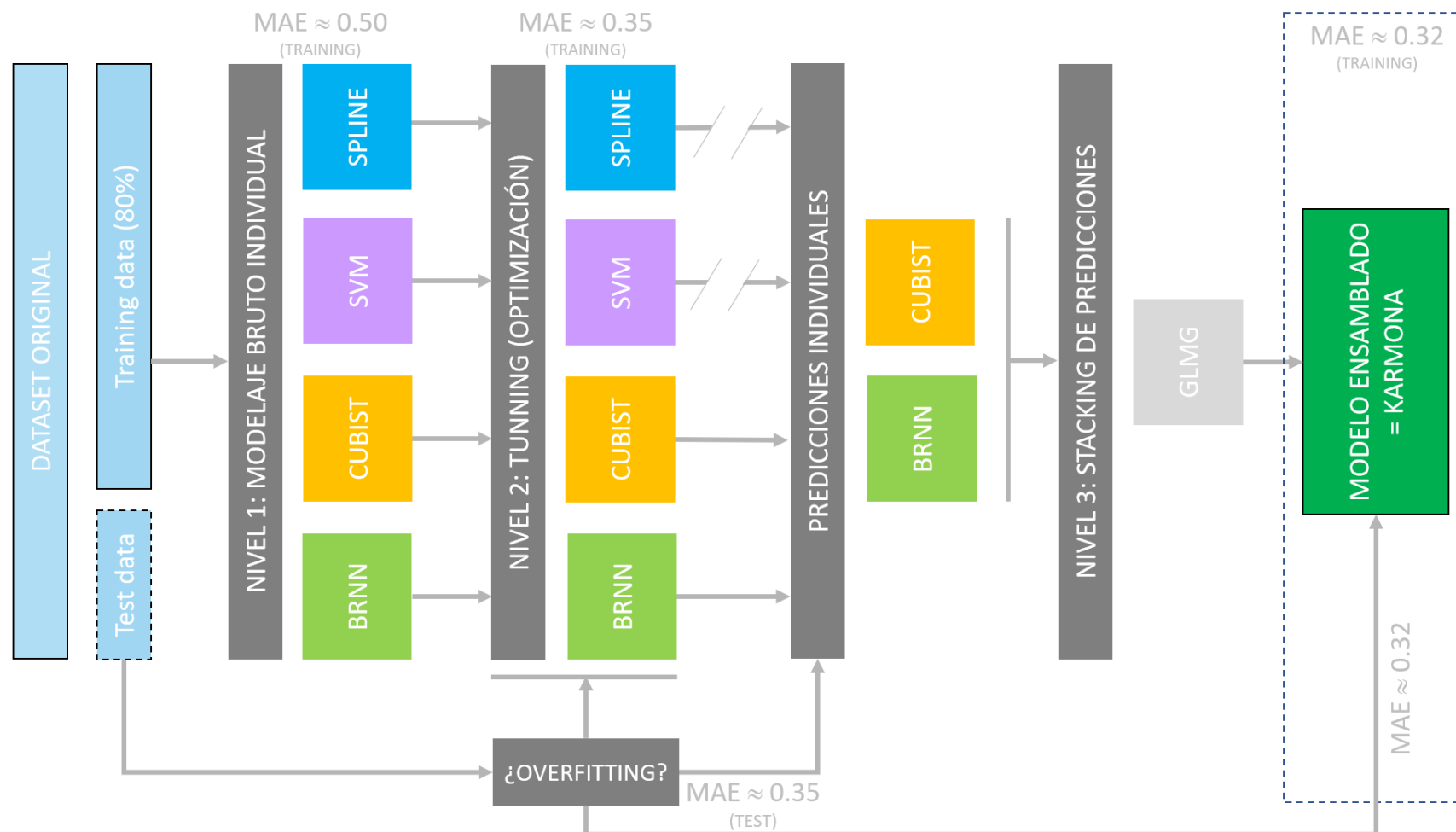


Figura 55. Esquema de modelaje en todos sus niveles de trabajo. El último nivel corresponde al modelo ensamblado que se llamó Karmona®, una combinación de BRNN y CUBIST, que no comete sobreajuste obteniendo el mismo MAE con la muestra de validación (20%).

Se probó a ensamblar todas las combinaciones posibles con los cuatro mejores modelos vistos en el apartado anterior, aunque la mejor unión y que más rendimiento ofreció fue BRNN + CUB-TREE (Figura 55). Aunque se probaron multitud de algoritmos de entrenamiento que se descartaron, el modelo ensamblado definitivo fue entrenado con el meta algoritmo *Generalized Linear Model Gaussian* (GLMG), y es el que da nombre al método final de cálculo de esta tesis, Karmona®.

Las métricas obtenidas mejoraban levemente las conseguidas por los modelos individuales, llegando hasta valores de MAE capsulares de 0.319 y R^2 de 0.9954, lo que implica una mejora del orden de 0.03 D sobre el mejor de los modelos individuales optimizados. La conversión de este MAE capsular a plano corneal resultará en unos valores por debajo de 0.24 D, haciendo muy competitiva a Karmona® con el resto de fórmulas actuales.

Hasta aquí, los valores de MAE, RMSE y R^2 se han obtenido sobre la partición de la muestra dedicada al *training* (80%). En adelante será necesario convalidar estos resultados con los datos reservados para la auditoría externa, el 20% del total de la muestra.

7.3.9 Interpretabilidad de modelos

En cuanto a la interpretabilidad del modelo Karmona® y atendiendo a la metodología expuesta en el capítulo [3.4.8 Interpretabilidad de modelos complejos](#), se tiene en primer lugar la importancia de las variables regresoras en cada uno de los modelos establecidos (Figura 56):

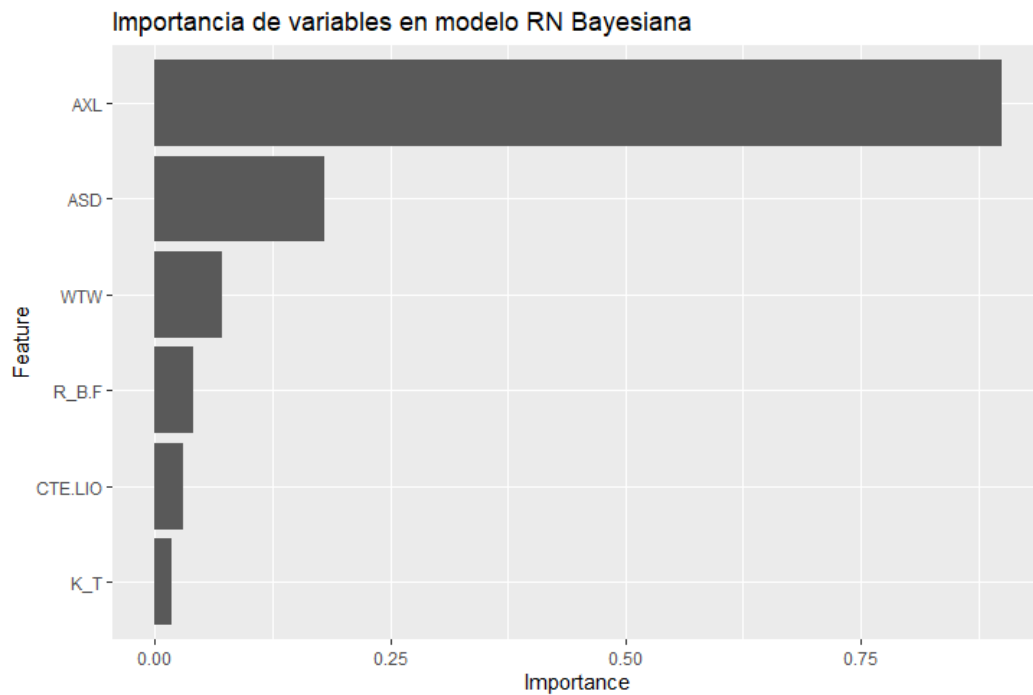


Figura 56. Importancia de las variables en el modelo de red neuronal bayesiana (BRANN).

En la red neuronal bayesiana construida con el modelo 2, la variable que más peso tiene en la predicción es la AXL seguida de ASD y de WTW, esto es, todas las variables lineales, dejando las variables que contienen información de curvatura en última posición, generando poca influencia en el resultado predictivo. De hecho, una variación notable de 2 DP en la queratometría media posterior afecta del orden de 0.40 D a la potencia emetropizante de la lente como variable respuesta.

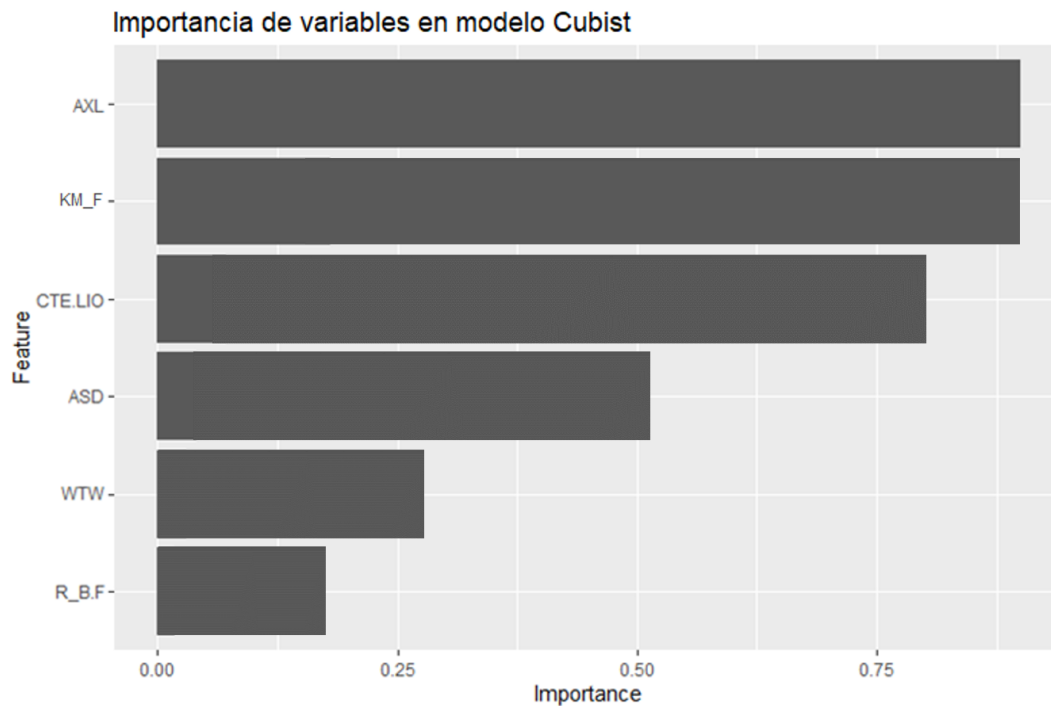


Figura 57. Importancia de variables en el modelo Cubist (CUB-TREE).

En el modelo *Cubist* el aporte predictivo de la queratometría calculada con 1.3375 es idéntico a la AXL (Figura 57), un 99%, lo que significa que prácticamente en todas las regresiones lineales obtenidas en cada submodelo (*committee*), ambas variables participan como regresoras.

```

Number of samples: 386
Number of predictors: 6

Number of committees: 5
Number of rules per committee: 6, 4, 2, 2, 3

```

Figura 58. Resumen de la optimización para Cubist. Se han empleado 386 ojos para el modelaje, 6 variables predictoras (AXL, ASD, WTW, CTE, SIMK, R_B/F), se han establecido 5 submodelos con 17 reglas por cada uno de ellos (en total 85 reglas).

```

Model 1:
Rule 1/1: [31 cases, mean 8.154109, range -3.52046 to 14, est err 0.568396]
  if
    AXL > 26.77
  then
    outcome = 142.258735 - 2.22 AXL - 1.31 KM_F - 0.18 R_B.F
Rule 1/2: [165 cases, mean 17.795965, range 7 to 24, est err 0.445833]
  if
    AXL > 23.41
    AXL <= 26.77
  then
    outcome = 21.09405 - 3.1 AXL - 1.33 KM_F + 1.1 CTE.LIO + 0.4 ASD
      - 0.07 R_B.F + 0.2 WTW
Rule 1/3: [42 cases, mean 20.189714, range 5.781973 to 27.30456, est err 0.561550]
  if
    KM_F > 45.495
  then
    outcome = 14.527673 - 3.41 AXL - 0.77 KM_F + 1 CTE.LIO
Rule 1/4: [28 cases, mean 21.658249, range 18 to 26.19012, est err 0.465061]
  if
    KM_F <= 45.495
    CTE.LIO <= 118.2
  then
    outcome = 513.967544 - 3.26 AXL - 1.03 KM_F - 3.2 CTE.LIO + 0.5 WTW
Rule 1/5: [132 cases, mean 23.750944, range 18.78692 to 33.84469, est err 0.484955]
  if
    AXL > 20.51
    AXL <= 23.41
    KM_F <= 45.495
    CTE.LIO > 118.2
  then
    outcome = 39.771599 - 3.78 AXL - 1.28 KM_F + 1 CTE.LIO + 0.9 ASD
Rule 1/6: [7 cases, mean 34.266098, range 30.95119 to 35.39616, est err 1.414781]
  if
    AXL <= 20.51
  then
    outcome = 60.667147 - 4.39 AXL + 5.2 WTW

```

Figura 59. Las 6 reglas establecidas para el submodelo1 con sus regresiones lineales multivariante correspondientes.

Como se puede observar en la Figura 59, las reglas se construyen en base a los casos que cumplen ciertas condiciones, por ejemplo, la regla 1 del submodelo 1 son 31 casos del *dataset* en los que $AXL > 26.77$ mm y su regresión lineal asociada solo tiene 3 predictores, no están ni ASD, ni CTE, ni SIMK.

Respecto a los gráficos PDP (Figura 60) muestran cómo varían en promedio las predicciones a medida que se modifica el valor de un predictor, manteniendo constantes

el resto. En la Figura 60 se muestra el efecto marginal promedio que AXL tiene en el resultado predicho por ambos modelos.

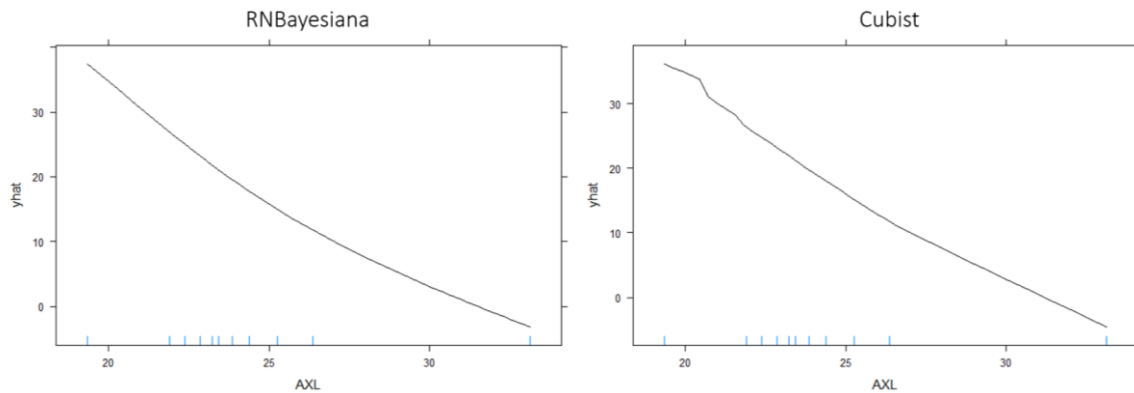


Figura 60. Gráficos PDP para la variable AXL en los dos modelos finales antes del ensamblaje.

No existe una relación lineal en ninguno de los modelos entre la variable respuesta y la longitud axial en ninguno de los tramos de axialidad explorados. Si se superponen ambos modelos sumando sus resultados se puede comprobar (Figura 61) que en un rango de 23.00 a 28.00 mm de AXL, ambos algoritmos predicen en promedio la misma potencia emetropizante cuando el resto de los *inputs* biométricos son constantes, pero fuera de ese intervalo, sobre todo en ojos cortos y extralargos ambos modelos complementan sus predicciones en el ensamblaje.

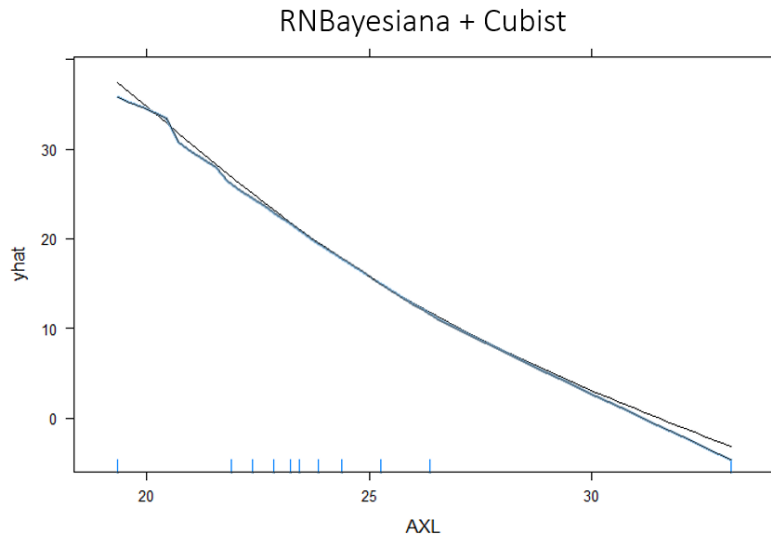


Figura 61. Gráfico PDP para el predictor AXL conjunto de los modelos seleccionados como finales antes del ensamblaje.

Sin embargo, en cuanto a la ASD, sí existe una relación lineal fuerte en ambos modelos con la predicción promedio, incluso para los extremos por ambas colas.

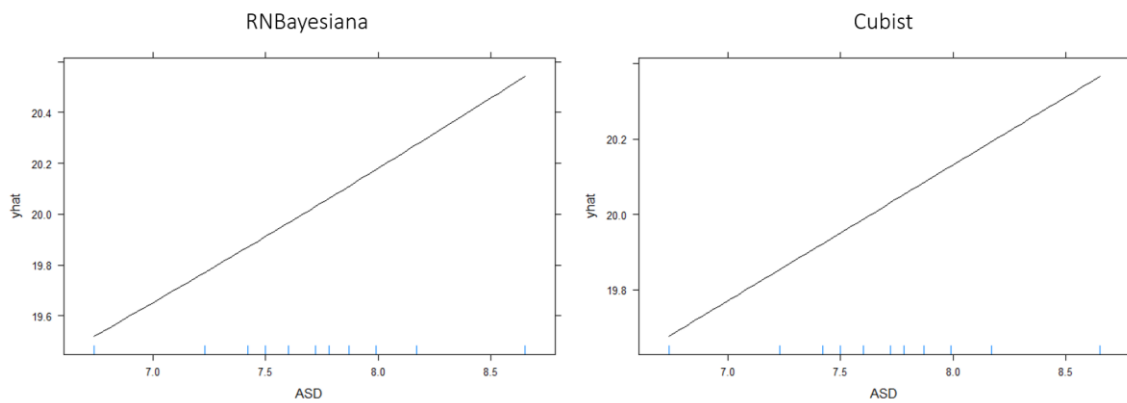


Figura 62. Gráficos PDP para la variable ASD en los dos modelos finales antes del ensamblaje.

Es interesante comprobar mediante un *Heatmap PDP* cómo se relacionan los dos predictores más importantes en el modelo *Cubist*, la longitud axial y la queratometría calculada con 1.3375 (Figura 63).

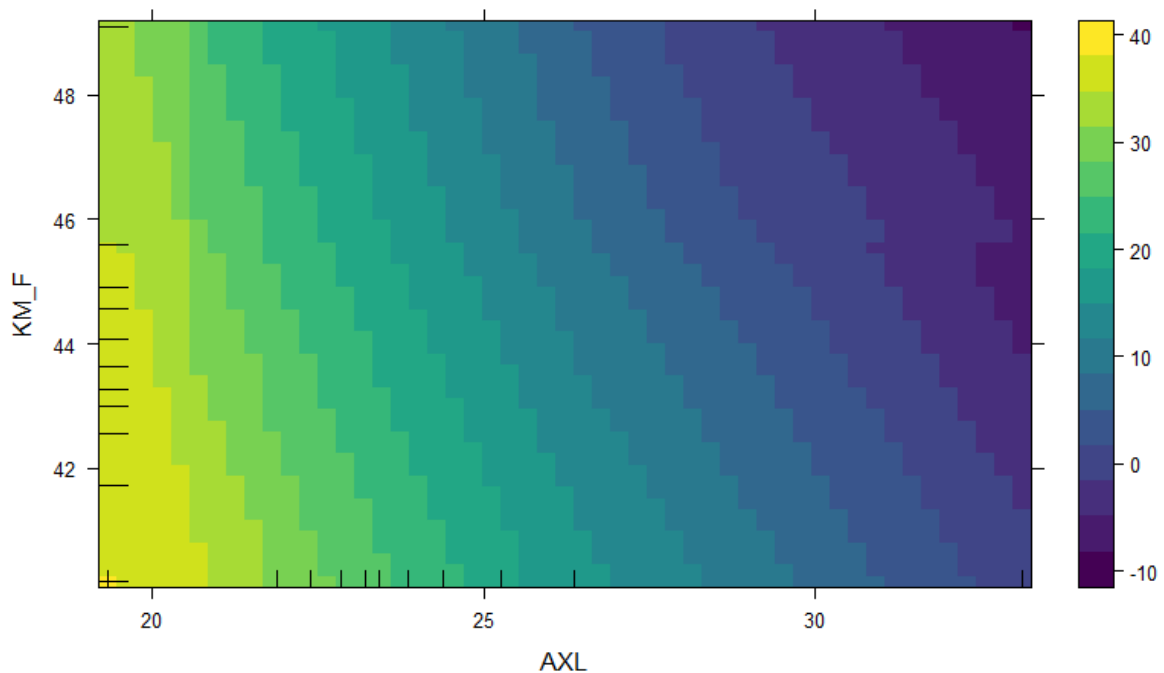


Figura 63. Heatmap PDP para AXL y KM_F del modelo Cubist.

A medida que la longitud axial es mayor, la queratometría tiene menos peso en la predicción media del modelo *Cubist* (Figura 63). Para AXL superiores a 30.00 mm las queratometrías medias inferiores a 42.00 D cuentan poco en la predicción de la lente emtopizante, pero las queratometrías medias superiores a 47.00 D no tienen peso. Sin embargo, para AXL por debajo de 23.00 mm cobran mucho protagonismo sobre todo las queratometrías medias por debajo de 44.00 D.

Por otro lado, en la relación entre AXL y el *ratio back/front* (R_B.F) que contiene información de la cara posterior corneal y es el regresor que menos cuenta en la predicción con *Cubist* (Figura 64), se puede observar que hasta AXL de 25.00 mm el ratio posterior/anterior influye de la misma manera en la predicción promedio, pero a medida que la longitud aumenta desde ahí, los ratios por debajo de 82% tienen más influencia sobre el cálculo de la potencia emtopizante, aunque de una forma muy suave.

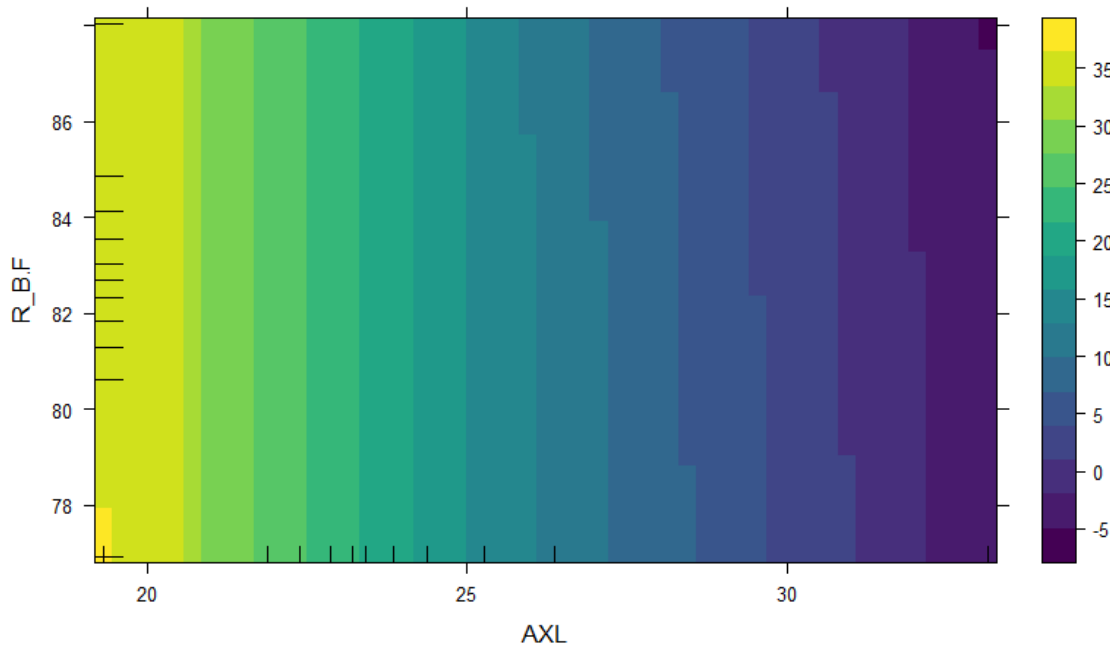


Figura 64. Heatmap PDP para AXL y R_B/F del modelo Cubist.

El Heatmap PDP para el modelo de red neuronal bayesiano (Figura 65) muestra la relación entre las dos variables más importantes e influyentes en el modelo, AXL y ASD. Donde se puede observar que tienen más influencia sobre la predicción de la lente intraocular, los segmentos anteriores profundos en ojos más largos que cortos cuando la AXL está por debajo de 24.00 mm, sin embargo a partir de aquí el modelo predice igual sin que ASD aporte demasiado.

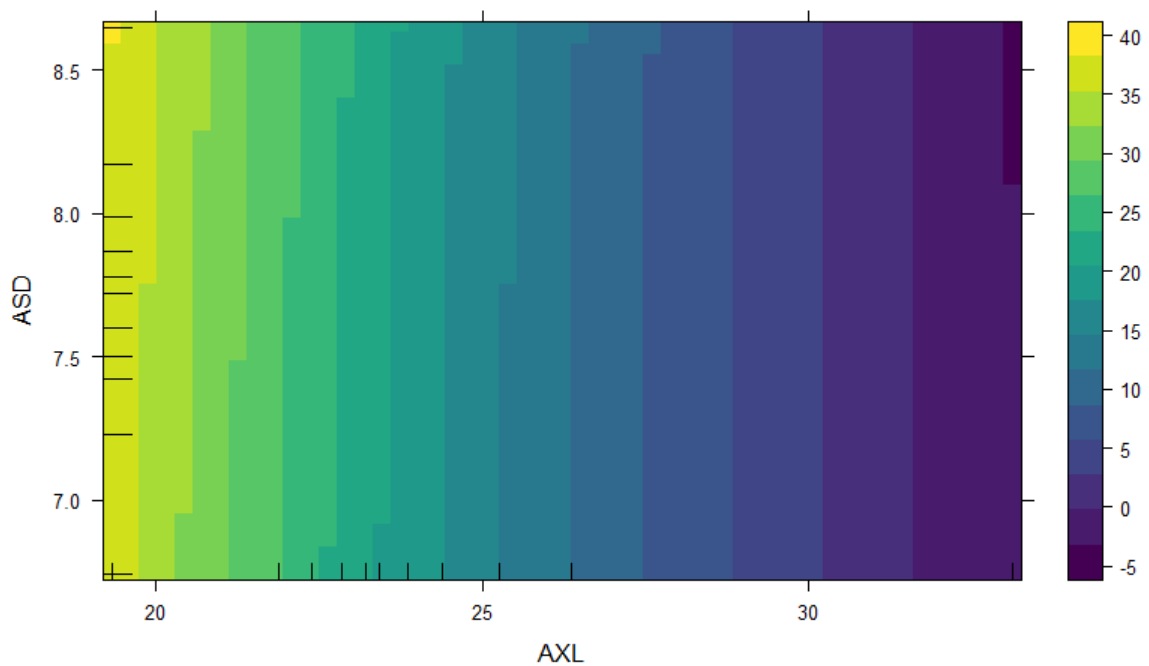


Figura 65. Heatmap PDP para AXL y ASD del modelo bayesiano.

En segundo lugar y atendiendo a los gráficos de esperanza condicional individual, en la (Figura 66-arriba) se muestra un gráfico ICE con selección de AXL como predictor de interés sobre el modelo de red bayesiana y también un gráfico d-ICE (Figura 66-abajo). En el gráfico ICE se puede observar como la función predicción promedio (línea con perfil amarillo) es monótona decreciente no lineal según aumenta la longitud axial, lo que implica menor potencia emetropizante a mayor eje anteroposterior. Cada predicción discurre continua y paralela a la PDP lo que denota inexistencia de interacciones con covariables en la evolución de la predicción. Efectivamente, en el gráfico d-ICE (Figura 66-abajo) se pone de manifiesto la no existencia de heterogeneidades ni interacciones salvo para una AXL entre 27.50 y 28.00 mm donde se puede observar un pico negativo de discontinuidad en la PDP pero que no afecta a la

desviación estándar a pie de gráfico, aunque englobada en un rango desde 27.00 mm a 29.00mm en el que se observa ciertas interacciones leves con otros predictores.

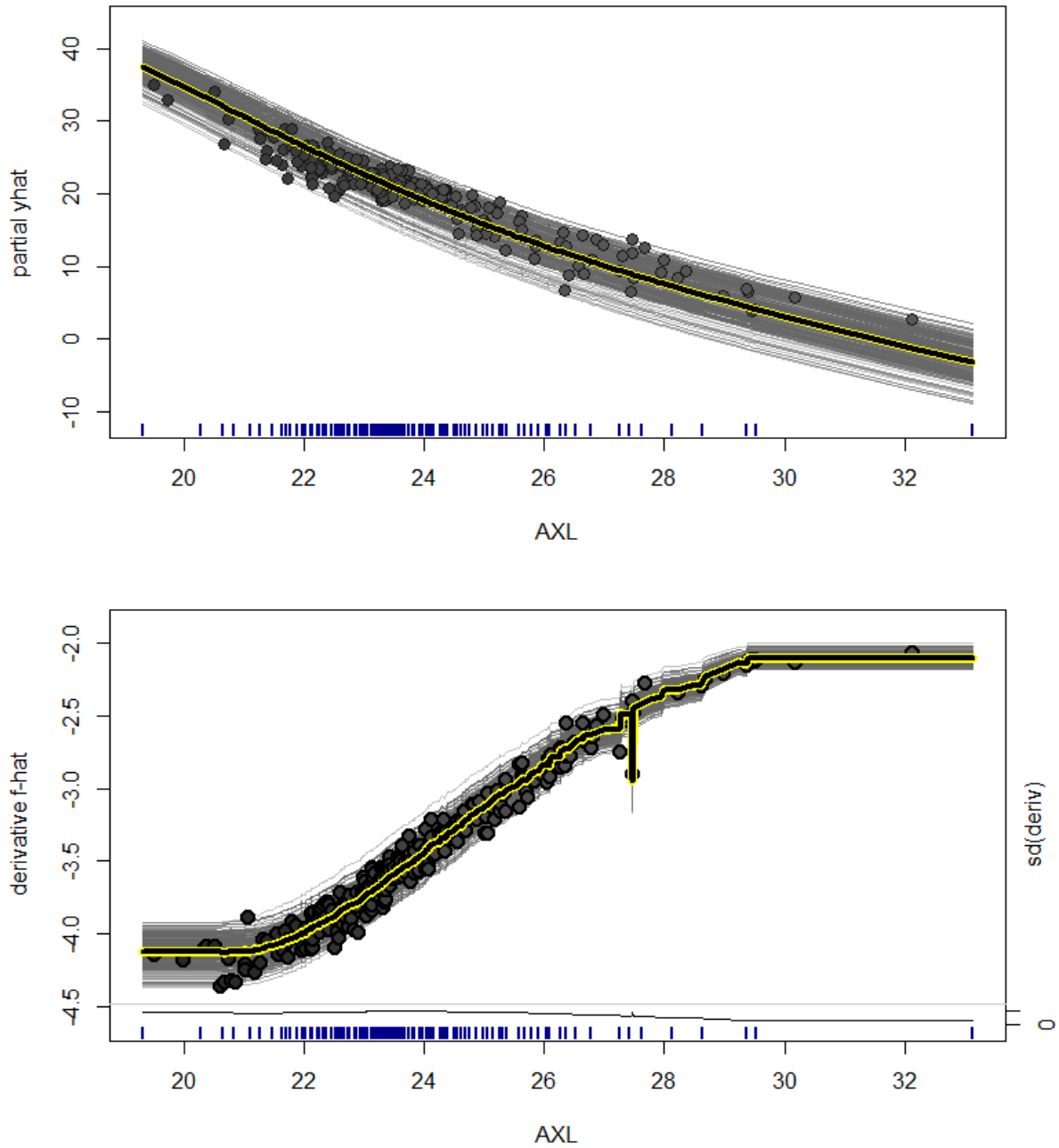


Figura 66. Arriba: gráfico ICE con predictor AXL en modelo bayesiano. Abajo: d-ICE para predictor AXL en el modelo bayesiano.

En cuanto a la interacción entre predictores (Figura 67 y Figura 68) representadas en ambos modelos entre AXL, que es la variable con más peso en ambos casos, y el resto de covariables, se tiene que en el modelo bayesiano la mayor interacción es entre la queratometría total media y la longitud axial, seguida de la interacción entre AXL y WTW por delante de la interacción con la profundidad del segmento anterior. Por el contrario, en el modelo *Cubist* la mayor varianza se explica por la interacción entre AXL y ASD.

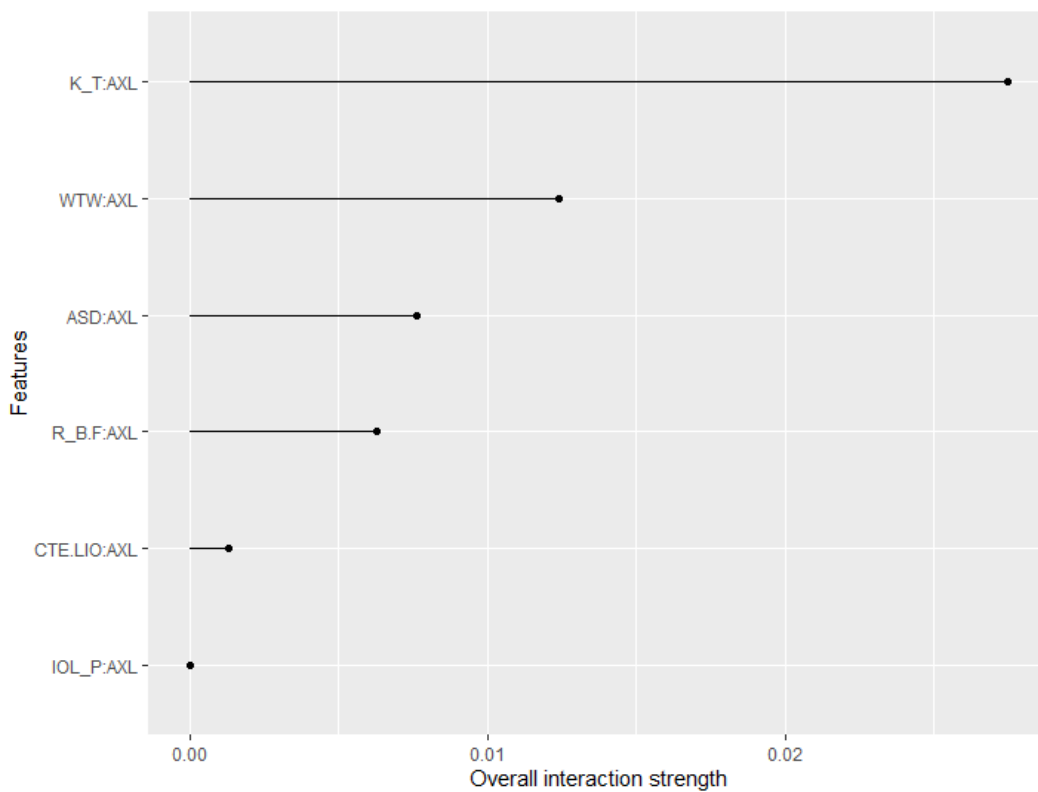


Figura 67. Interacciones entre AXL y el resto de predictores en el modelo bayesiano.

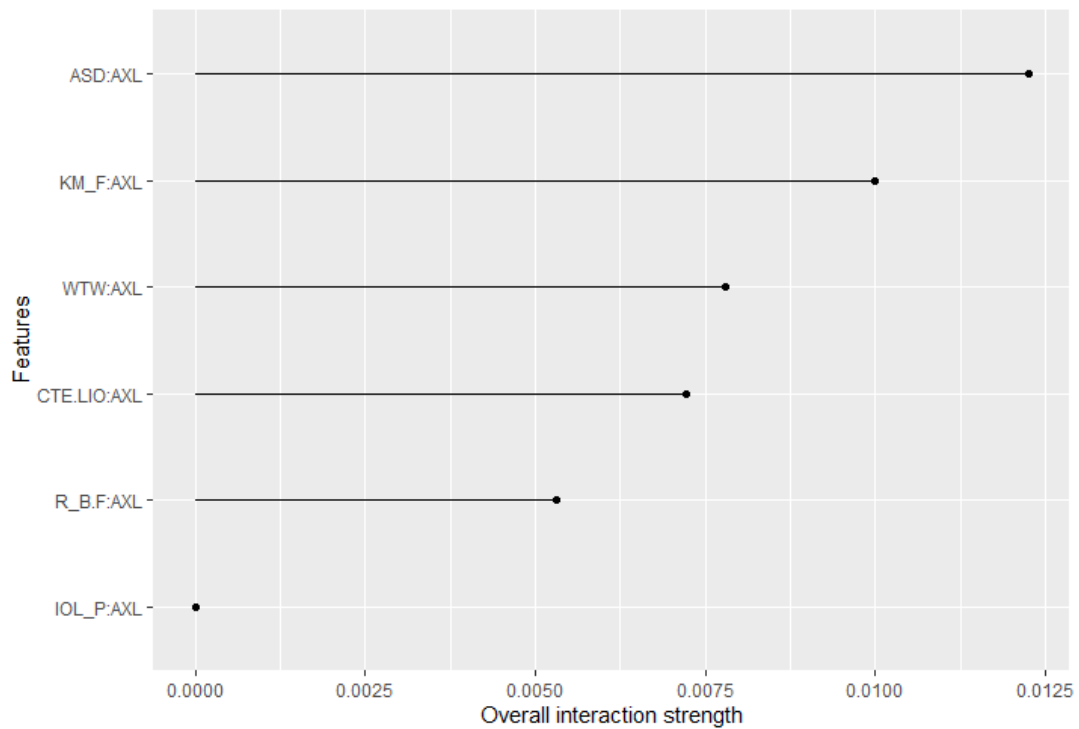


Figura 68. Interacciones entre AXL y el resto de predictores en el modelo Cubist.

7.4 Resultados de la etapa V

7.4.1 Validación del modelo resultante sobre la partición test

Recuérdese que la partición *test* o de validación del *dataset* original está aleatorizada, lo que significa que ese 20% de ojos ($n = 95$) consta de una mezcla de todas las tipologías anatómicas. Dicha diversidad producirá veracidad en los resultados medios auditados.

La Tabla 14 muestra los resultados obtenidos con Karmona® para los 95 ojos de la partición *test*, dónde se puede observar las siguientes métricas de rendimiento del modelo, un MAE de 0.33 D y un R^2 de 0.995. Lo que confirma que Karmona® no produce sobreajuste en la fase predictiva, puesto que obtiene el mismo resultado en todas las métricas que en la fase de entrenamiento.

	ENTRENAMIENTO	TEST-VALIDACIÓN (LIO)	TEST-VALIDACIÓN (CÓRNEA)
MAE (D)	0.33	0.33	0.23
R ²	0.992	0.995	0.995

Tabla 14. Métricas de rendimiento del modelo Karmona® en la fase de entrenamiento con el 80% del dataset original (columna izquierda) y en la fase de validación con el 20% de los datos (columna central). En la columna de la derecha el MAE se convierte por distometría a plano corneal.

Por otro lado, atendiendo a las métricas clásicas para evaluar la precisión y exactitud de cualquier fórmula de cálculo, y teniendo presente que la aleatorización de la partición de validación conforma una muestra con diferentes tipos de lentes intraoculares, por tanto de constantes que provienen de optimizaciones extraídas de <http://www.IOLcon.org>, se deberá proceder a centrar la distribución de los datos de error predictivo (*zeroed out*) para que la comparación posterior con otras fórmulas o métodos sea en las mismas condiciones de partida ([11.2 ANEXO 2.- ANÁLISIS CRÍTICO: MÉTRICAS PARA EVALUAR LA PRECISIÓN Y EXACTITUD DE FÓRMULAS](#)).

Tal y como se muestra en la Tabla 15 el error predictivo medio sin ajustar es -0.02 D lo que implica un equilibrio hacia ninguna tendencia ametrópica del residuo refractivo postoperatorio, no obstante la tendencia es levemente miópica. El MAE de Karmona® es 0.23 ± 0.20 D y el MedAE es de 0.19 D, con un error máximo en valor absoluto de 0.87 D. El porcentaje de ojos incluidos en el intervalo refractivo postquirúrgico de ± 0.25 D es de 66.32 % así como un 91.58% en el rango de ± 0.50 D.

Las métricas cuando la distribución se ajusta centrándola no son muy distintas, mejorando levemente en MedAE y en el porcentaje de ojos incluidos en el intervalo refractivo ± 0.50 D (Tabla 15).

	n	MPE (D)	MAE (D)	SD (D)	MedAE (D)	MAX ERROR (D)	± 0.25 (%)	± 0.50 (%)
SIN AJUSTE	95	-0.02	0.23	0.20	0.19	0.87	66.32	91.58
CON AJUSTE	95	0.00	0.23	0.20	0.18	0.88	65.26	92.63

Tabla 15. Métricas clásicas de evaluación de precisión y exactitud de una fórmula o método de cálculo biométrico. La fila SIN AJUSTE muestra las métricas sin la corrección de centrado sobre el MPE. La fila CON AJUSTE corresponde a las mismas métricas con la corrección sobre MPE.

Por otra parte, la Tabla 16 muestra los resultados de Karmona® con las métricas propuestas en esta tesis en [11.2 ANEXO 2.- ANÁLISIS CRÍTICO: MÉTRICAS PARA EVALUAR LA PRECISIÓN Y EXACTITUD DE FÓRMULAS](#). Se pueden observar unos índices IOF muy altos, 0.839 y 0.882 para el rango ± 0.25 D y ± 0.50 D respectivamente, lo que denota un altísimo rendimiento predictivo en ambos rangos. El promedio de errores absolutos tras la anti-transformación es de 0.194 D lo que se traduce en una alta exactitud por debajo del cuarto de dioptría, con una media recortada absoluta al 10% prácticamente idéntica de 0.189 D, lo que significa que en 95 ojos de la partición de validación no hay *outliers*.

MÉTRICAS	KARMONA
IOF25	0,839
IOF50	0,882
ANTI-MAE (D)	0,194
ANTI-TMAE10 (D)	0,189
ANTI-MedAE (D)	0,190
ANTI-MAD (D)	0,025

Tabla 16. Métricas propuestas en esta tesis en relación con Karmona®. IOF (Index Of Formula), ANTI-MAE es el promedio del error absoluto de la distribución transformada. ANTI-TMAE10 es el promedio del error absoluto de la distribución transformada, pero con un acotamiento bilateral del 10%. ANTI-MedAE es la mediana del error absoluto de la distribución transformada y ANTI-MAD es la distancia absoluta entre la mediana y el error de predicción absoluto de la distribución transformada.

La mediana tras la transformación de los datos es de 0.19 D coincidente con MAE reforzando el argumento de la normalización. Por último la métrica que representa la dispersión respecto a un valor de centralidad, ANTI-MAD, posee un valor muy bajo de 0.025 D indicando por ende una alta precisión.

7.5 Resultados de la etapa VI

7.5.1 Comparación del modelo con otras fórmulas y métodos de cálculo

Una vez comprobada la precisión y exactitud de Karmona® sobre la muestra de validación, para evaluar la competitividad será necesario establecer una comparación con el resto de fórmulas y métodos de cálculo existentes empleados actualmente.

Tal y como se refirió en [6.1.3 Datos para la comparación entre fórmulas](#), las 12 fórmulas seleccionadas para la comparación con Karmona® fueron Holladay I y II, SRK/T, Haigis, Hoffer Q, Barrett Universal II, Olsen PhacoOptics versión 1.10.100.2030-35, Ladas Superformula v1b, Kane, EVO v2.0, Hill-RBF v3.0 y PEARL-DGS v2.0.

La comparación del modelo Karmona® con el resto de fórmulas y métodos se hizo con la partición de 95 ojos correspondientes al dataset de validación. La Tabla 17 muestra las características demográficas de la muestra en cuestión.

	AXL (mm)	ACD (mm)	LT (mm)	WTW (mm)	KM ANT (D)	KM POST (D)	POTENCIA LIO (D)
MEAN	24,05	3,17	4,53	11,92	43,76	-6,26	20,02
SD	2,39	0,41	0,33	0,47	1,68	0,29	6,24
MEDIAN	23,36	3,13	4,51	11,90	43,61	-6,25	22,00
MAX	33,50	4,33	5,38	13,60	48,10	-5,40	35,00
MIN	20,43	2,38	3,77	10,90	38,52	-7,04	-4,00

Tabla 17. Variables biométricas del total de la muestra de validación (n = 95).

En la Tabla 18 se muestran los resultados de las 13 fórmulas, incluida Karmona[®], en base a las métricas tradicionales para toda la muestra (n = 95) sin discriminar por longitud axial y sin ajustar el MPE.

Karmona[®], PEARL-DGS y Kane se desmarcan respecto a las demás fórmulas, presentando resultados semejantes en el entorno del cuarto de dioptría, siendo Karmona[®] la mejor: 0.23 ± 0.20 D , 0.27 ± 0.25 D y 0.27 ± 0.31 D respectivamente. Además de esto Karmona[®] tiene las cotas de error máximo y mínimo más bajas introduciendo más cantidad de ojos en el rango ± 0.50 D, un 91.58%.

	HOLL I	SRK/T	HOFFERQ	HOLL II	HAIGIS	BARRETT	HILL RBF	KANE	LADAS	EVO	OLSEN	PEARL	KARMONA
MAE (D)	0,38	0,36	0,40	0,30	0,32	0,30	0,32	0,27	0,30	0,30	0,30	0,27	0,23
SD (D)	0,41	0,45	0,41	0,37	0,33	0,31	0,31	0,31	0,37	0,25	0,32	0,25	0,20
MedAE (D)	0,28	0,24	0,31	0,18	0,21	0,19	0,22	0,17	0,17	0,27	0,20	0,22	0,19
MAX ERROR (D)	1,24	0,61	1,46	0,77	0,60	0,58	1,00	0,63	0,70	1,63	0,86	1,49	0,87
MIN ERROR (D)	-2,57	-2,28	-2,39	-2,27	-1,89	-1,79	-1,59	-1,53	-1,90	-0,85	-1,58	-0,85	-0,86
± 0.25 D (%)	32,63	51,58	35,79	46,32	61,05	54,74	44,21	50,53	54,74	48,42	45,26	61,05	66,32
± 0.50 D (%)	56,84	84,21	58,95	75,79	82,11	91,58	80,00	81,05	86,32	84,21	76,84	88,42	91,58

*Sin ajuste sobre Mean Predictive Error.

**Muestra para un n = 95 ojos.

Tabla 18. Resultados con métricas clásicas para las 13 fórmulas sin ajustar MPE en toda la muestra (n = 95 ojos).

La Tabla 19 muestra los resultados empleando las métricas tradicionales haciendo este proceso de *zeroed out*. Con el ajuste sobre MPE se produce una mejoría notable en muchas fórmulas en términos de MAE ± SD, siendo de nuevo Karmona® y Kane las mejores con 0.23 ± 0.20 D y 0.25 ± 0.27 D respectivamente. En cuanto a MedAE, Kane presenta un mejor resultado que Karmona® con el ajuste: 0.16 D contra 0.18 D.

	HOLL I	SRK/T	HOFFERQ	HOLL II	HAIGIS	BARRETT	HILL RBF	KANE	LADAS	EVO	OLSEN	PEARL	KARMONA
MAE (D)	0,39	0,32	0,41	0,30	0,27	0,25	0,32	0,25	0,28	0,28	0,28	0,27	0,23
SD (D)	0,40	0,38	0,40	0,35	0,31	0,29	0,30	0,27	0,32	0,26	0,28	0,26	0,20
MedAE (D)	0,30	0,21	0,30	0,21	0,18	0,19	0,24	0,16	0,18	0,21	0,21	0,21	0,18
MAX ERROR (D)	2,53	1,98	2,31	2,16	1,68	1,59	1,48	1,36	1,68	1,74	1,40	1,52	0,88
MIN ERROR (D)	0,01	0,00	0,03	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
± 0.25 D (%)	38,95	55,79	41,05	63,16	58,95	58,95	51,58	65,26	60,00	69,47	58,95	58,95	65,26
± 0.50 D (%)	78,95	85,26	76,84	82,11	86,32	88,42	82,11	90,53	87,37	91,58	84,21	87,37	92,63

*Con ajuste sobre Mean Predictive Error.

**Muestra para un n = 95 ojos.

Tabla 19. Resultados con métricas clásicas para las 13 fórmulas con ajuste sobre MPE en toda la muestra (n = 95 ojos).

El test de *Friedman* confirmó que hubo diferencias estadísticamente significativas ($p < 0.01$) entre fórmulas tanto en la muestra sin ajustar MPE como en la ajustada a 0.00 D. El análisis *post hoc* usando *Wilcoxon signed-rank* con ajuste de *Bonferroni* para comparaciones múltiples mostró diferencias significativas en la muestra sin ajuste sobre MPE entre Holladay II y Hoffer Q ($p < 0.01$), Kane versus Holladay I y Hoffer Q ($p < 0.01$), así como entre Ladas Superformula y Hoffer Q ($p < 0.01$), no encontrándose otras diferencias entre fórmulas para toda la muestra.

Por otro lado, en la muestra ajustada se encontraron diferencias estadísticamente significativas ($p < 0.01$) entre Holladay II, Haigis, Kane, Ladas Superformula y Karmona® con Hoffer Q, además de entre Holladay II, Haigis, Kane y Ladas Superformula con Holladay I.

Figura 70 se pueden observar los *boxplot* de las 13 fórmulas tanto sin ajuste sobre MPE como con ajuste respectivamente, comprobando que Karmona® es la que menos outliers tiene, concretamente 3 y por debajo de 1.00 D. Por el contrario EVO y Kane tienen 6 y 7 respectivamente en el entorno de 1.50 D.

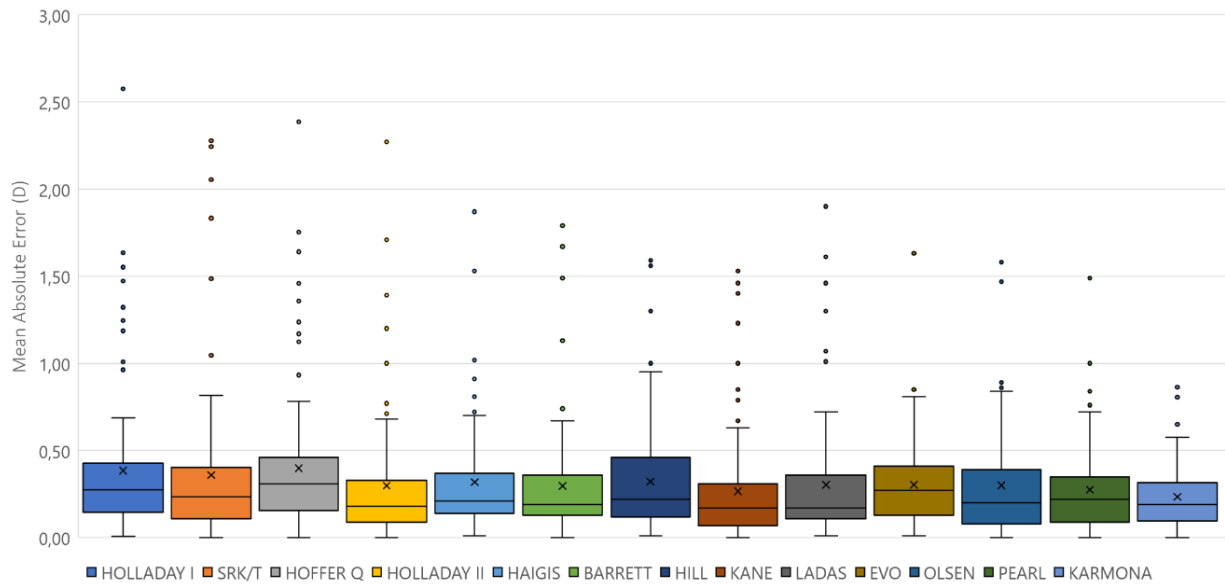


Figura 69. Boxplot de todas las fórmulas sin ajuste sobre MPE para el total de la muestra (n = 95).

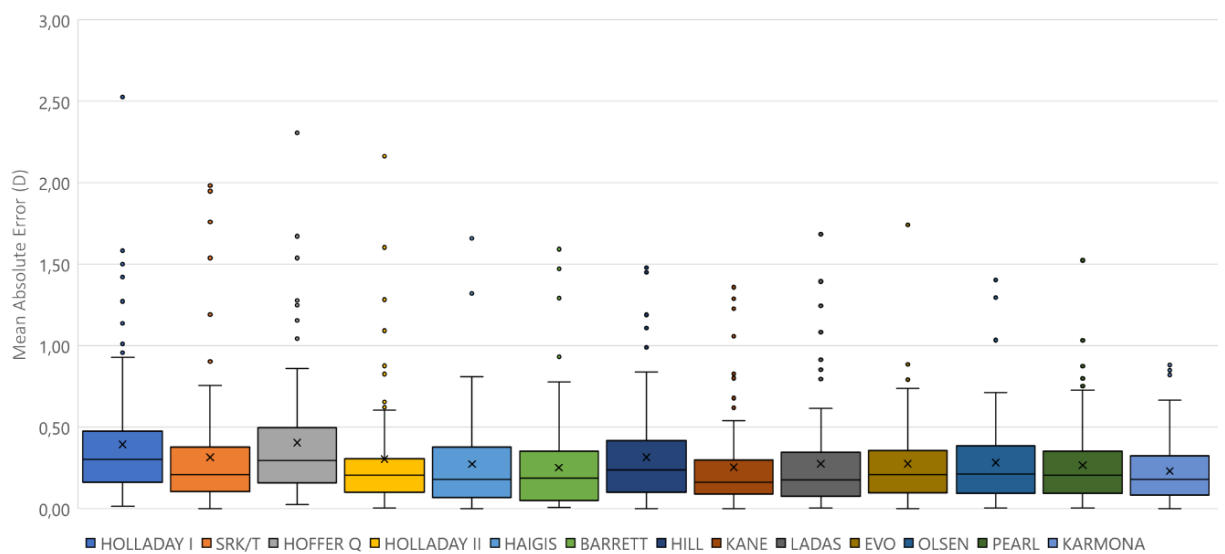


Figura 70. Boxplot de todas las fórmulas con ajuste 0.00 D sobre MPE para el total de la muestra (n = 95).

La Tabla 20 muestra el rendimiento de las fórmulas medido con las métricas propuestas en esta tesis para toda la muestra (n = 95).

	HOLL I	SRK/T	HOFFERQ	HOLL II	HAIGIS	BARRETT	HILL RBF	KANE	LADAS	EVO	OLSEN	PEARL	KARMONA
IOF25	0,704	0,756	0,698	0,796	0,774	0,781	0,763	0,822	0,794	0,766	0,786	0,805	0,848
IOF50	0,742	0,801	0,731	0,846	0,805	0,840	0,824	0,874	0,844	0,822	0,802	0,850	0,882
MEAN IOF	0,723	0,779	0,714	0,821	0,789	0,811	0,794	0,848	0,819	0,794	0,794	0,828	0,865
ANTI-MAE (D)	0,309	0,267	0,318	0,227	0,259	0,242	0,261	0,204	0,234	0,256	0,234	0,226	0,194
ANTI-TMAE10 (D)	0,540	0,495	0,551	0,457	0,493	0,477	0,498	0,435	0,465	0,252	0,473	0,469	0,434
ANTI-MedAE(D)	0,292	0,245	0,303	0,209	0,243	0,227	0,248	0,189	0,216	0,270	0,224	0,220	0,189
ANTI-MAD (D)	0,038	0,048	0,044	0,035	0,032	0,028	0,038	0,035	0,035	0,030	0,039	0,029	0,025

*Muestra para un n = 95 ojos.

Tabla 20. Rendimiento de las 13 fórmulas empleando las métricas descritas en esta tesis. IOF25 Index Of Formula en el rango ± 0.25 D. IOF50 en el rango de ± 0.50 D. MEAN IOF es el promedio de IOF25 e IOF50. ANTI-MAE es el error medio absoluto de la distribución normalizada. ANTI-TMAE10 es el error medio absoluto de la distribución normalizada pero acotada al 10% por ambas colas. ANTI-MedAE es la mediana de los errores absolutos de la distribución normalizada. ANTI-MAD es la distancia promedio entre el error absoluto y la mediana de la distribución normalizada.

Karmona® y Kane son las fórmulas que tuvieron un mejor comportamiento para la muestra total, pero destaca Karmona® por presentar mejores índices IOF25 e IOF50 así como mejores métricas de exactitud ANTI-MAE, ANTI-TMAE10 y ANTI-MedAE y precisión ya que Karmona® obtuvo un ANTI-MAD de 0.025 D. En base a estas métricas, el ordenamiento de las fórmulas sería Karmona® con un índice medio de 0.865, Kane con un índice medio de 0.848, PEARL-DGS con un índice de 0.828, Holladay II con un índice de 0.821 y en quinto lugar Ladas Superformula con 0.819. Obsérvese que cuatro de las cinco primeras fórmulas tienen relación con la IA.

Tras la transformación de los datos y el empleo de las métricas propuestas, el análisis de varianza no mostró diferencias estadísticamente significativas ($p < 0.01$) entre ninguna de las fórmulas en la categoría de ANTI-MAE, confirmándose este hecho ulteriormente en las comparaciones múltiples *post hoc* con ajuste de *Bonferroni*.

Como el *dataset* de validación de 95 ojos está previamente aleatorizado tal y como se expresó en el apartado, consta de diferentes longitudes axiales, se segmentó la muestra en función de estas para ver el rendimiento de las distintas fórmulas en los

diferentes tramos anatómicos. Debido al escaso tamaño muestral ($n = 95$), se optó por seleccionar los puntos de corte clásicos en la longitud axial. Por tanto, los ojos cortos por debajo de 22.00 mm fueron 12, ojos normales entre 22.01 mm y 24.99 mm fueron 58, y ojos largos por encima de 25.00 mm hubo 25.

La Tabla 21 muestra los resultados de todas las fórmulas para ojos cortos ($n = 12$) por debajo de 22.00 mm de AXL sin ajuste sobre los errores predictivos y con ajuste.

	HOLL I	SRK/T	HOFFERQ	HOLL II	HAIGIS	BARRETT	HILL RBF	KANE	LADAS	EVO	OLSEN	PEARL	KARMONA
SIN AJUSTE SOBRE MPE													
MAE (D)	0,45	0,58	0,38	0,35	0,48	0,42	0,43	0,45	0,46	0,42	0,48	0,42	0,35
SD (D)	0,52	0,73	0,46	0,45	0,42	0,48	0,40	0,44	0,50	0,47	0,45	0,44	0,30
MedAE (D)	0,29	0,38	0,22	0,17	0,43	0,22	0,36	0,27	0,24	0,30	0,41	0,26	0,21
MAX ERROR (D)	0,46	0,02	0,49	0,23	0,04	0,16	0,16	0,22	0,28	1,63	0,12	1,49	0,65
MIN ERROR (D)	-1,55	-2,24	-1,36	-1,39	-1,53	-1,79	-1,56	-1,40	-1,63	-0,79	-1,58	-0,85	-0,86
CON AJUSTE SOBRE MPE													
MAE (D)	0,49	0,52	0,42	0,39	0,28	0,37	0,33	0,38	0,46	0,41	0,41	0,42	0,31
SD (D)	0,44	0,49	0,39	0,36	0,31	0,36	0,32	0,32	0,36	0,47	0,28	0,44	0,30
MedAE (D)	0,42	0,46	0,27	0,33	0,17	0,22	0,19	0,28	0,36	0,26	0,38	0,26	0,20
MAX ERROR (D)	0,61	0,60	0,61	0,41	0,52	0,53	0,53	0,60	0,63	1,69	0,55	1,48	0,80
MIN ERROR (D)	-1,40	-1,66	-1,23	-1,21	-1,05	-1,42	-1,20	-1,02	-1,28	-0,73	-1,15	-0,86	-0,72

*Muestra para un $n = 12$ ojos cortos ($AXL \leq 22$ mm).

Tabla 21. Rendimiento de las 13 fórmulas en ojos cortos ($AXL \leq 22.00$ mm) sin y con ajuste de centrado sobre MPE ($n = 12$ ojos) de la partición test.

En ojos cortos sin corrección sobre el MPE, Karmona® y Holladay II se posicionan como las mejores con MAE \pm SD de 0.35 ± 0.30 D y 0.35 ± 0.45 D respectivamente. Con el centrado de distribución las mejores son Haigis y Karmona® con 0.28 ± 0.31 D y 0.31 ± 0.30 D respectivamente.

El análisis *post hoc* usando *Wilcoxon signed-rank* con ajuste de *Bonferroni* en las comparaciones múltiples de la muestra de ojos cortos ($n = 12$), no se mostraron

diferencias significativas entre los errores absolutos en la serie ajustada y sin ajustar sobre MPE entre ninguna de las fórmulas estudiadas.

La Tabla 22 muestra el resultado para los ojos largos (n = 25). Karmona® vuelve a ser la mejor fórmula con un MAE ± SD de 0.21 ± 0.19 D tanto con ajuste del MPE como sin él, seguida de PEARL-DGS con 0.26 ± 0.21 D sin ajuste sobre MPE. Kane pasa a ser la segunda mejor cuando se centra la distribución con un MAE ± SD de 0.22 ± 0.18 D.

	HOLL I	SRK/T	HOFFERQ	HOLL II	HAIGIS	BARRETT	HILL RBF	KANE	LADAS	EVO	OLSEN	PEARL	KARMONA
SIN AJUSTE SOBRE MPE													
MAE (D)	0,54	0,53	0,57	0,48	0,41	0,33	0,34	0,29	0,41	0,30	0,33	0,26	0,21
SD (D)	0,62	0,60	0,58	0,57	0,50	0,40	0,38	0,38	0,52	0,16	0,40	0,21	0,19
MedAE (D)	0,32	0,34	0,37	0,22	0,25	0,18	0,21	0,17	0,17	0,29	0,20	0,24	0,20
MAX ERROR (D)	0,62	0,35	0,44	0,72	0,12	0,36	0,67	0,37	0,36	0,68	0,47	1,00	0,87
MIN ERROR (D)	-2,57	-2,28	-2,39	-2,27	-1,89	-1,67	-1,59	-1,53	-1,90	-0,55	-1,58	-0,42	-0,52
CON AJUSTE SOBRE MPE													
MAE (D)	0,57	0,49	0,51	0,43	0,35	0,32	0,32	0,28	0,42	0,22	0,33	0,26	0,21
SD (D)	0,47	0,45	0,47	0,44	0,38	0,35	0,36	0,34	0,40	0,18	0,38	0,21	0,19
MedAE (D)	0,46	0,35	0,37	0,30	0,22	0,26	0,20	0,17	0,28	0,19	0,22	0,22	0,20
MAX ERROR (D)	0,96	0,78	0,86	1,13	0,50	0,58	0,84	0,56	0,68	0,86	0,60	1,02	0,86
MIN ERROR (D)	-2,23	-1,84	-1,96	-1,86	-1,51	-1,45	-1,42	-1,34	-1,58	-0,37	-1,45	-0,40	-0,53

*Muestra para un n = 25 ojos largos (AXL ≥ 25 mm).

Tabla 22. Rendimiento de las 13 fórmulas en ojos cortos (AXL ≥ 25.00 mm) sin y con ajuste de centrado sobre MPE (n = 25 ojos) de la partición test.

Tras el análisis *post hoc* usando *Wilcoxon signed-rank* con ajuste de *Bonferroni* en las comparaciones múltiples de la muestra de ojos largos (n = 25), Karmona® y Barrett Universal II mostraron diferencias significativas (p < 0.01) únicamente respecto a Hoffer Q tanto en la serie ajustada como sin ajustar sobre MPE, no encontrándose otras diferencias significativas.

La Tabla 23 muestra los resultados en ojos axialmente normales (n = 58), y se ve cómo se produce mucha diversificación entre las fórmulas porque todas tienen un

comportamiento semejante. Con ajuste sobre MPE, Ladas Superformula y SRK/T son las mejores con MAE \pm SD de 0.19 \pm 0.18 D y 0.19 \pm 0.18 D respectivamente. En la ordenación por MAE y ajuste sobre MPE, Holladay II, Kane y Karmona[®] son las siguientes con 0.20 \pm 0.18 D, 0.21 \pm 0.19 D y 0.22 \pm 0.17 D respectivamente.

	HOLL I	SRK/T	HOFFERQ	HOLL II	HAIGIS	BARRETT	HILL RBF	KANE	LADAS	EVO	OLSEN	PEARL	KARMONA
SIN AJUSTE SOBRE MPE													
MAE (D)	0,31	0,24	0,33	0,21	0,24	0,25	0,29	0,22	0,22	0,28	0,25	0,25	0,22
SD (D)	0,22	0,21	0,27	0,18	0,18	0,21	0,26	0,22	0,20	0,22	0,22	0,21	0,17
MedAE (D)	0,25	0,20	0,27	0,15	0,19	0,20	0,21	0,13	0,17	0,22	0,19	0,20	0,16
MAX ERROR (D)	1,24	0,61	1,46	0,77	0,60	0,58	1,00	0,63	0,70	0,79	0,86	0,84	0,81
MIN ERROR (D)	-0,69	-1,05	-0,94	-0,68	-0,81	-1,13	-0,95	-1,00	-1,01	-0,85	-0,89	-0,76	-0,58
CON AJUSTE SOBRE MPE													
MAE (D)	0,28	0,19	0,32	0,20	0,23	0,20	0,29	0,21	0,19	0,26	0,23	0,24	0,22
SD (D)	0,23	0,19	0,27	0,18	0,18	0,21	0,26	0,19	0,18	0,23	0,19	0,21	0,17
MedAE (D)	0,25	0,14	0,29	0,14	0,17	0,14	0,23	0,15	0,14	0,21	0,20	0,19	0,16
MAX ERROR (D)	1,14	0,79	1,38	0,73	0,68	0,73	1,03	0,75	0,84	0,88	1,00	0,89	0,80
MIN ERROR (D)	-0,79	-0,87	-1,02	-0,72	-0,73	-0,98	-0,92	-0,88	-0,87	-0,76	-0,75	-0,71	-0,58

*Muestra para un n = 58 ojos largos (AXL > 22 y AXL < 25 mm).

Tabla 23. Rendimiento de las 13 fórmulas en ojos cortos (AXL > 22.00 y AXL < 25.00 mm) sin y con ajuste de centrado sobre MPE (n = 25 ojos) en la partición test.

El análisis *post hoc* usando *Wilcoxon signed-rank* con ajuste de *Bonferroni* en las comparaciones múltiples de la muestra de ojos axialmente normales (n = 58), no mostró diferencias significativas entre los errores absolutos en la serie ajustada y sin ajustar sobre MPE entre ninguna de las fórmulas.

La Figura 71 muestra el diagrama de barras de los porcentajes de ojos introducidos en los diferentes intervalos refractivos postoperatorios: ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D. Karmona[®] es la que más ojos incluye en los dos primeros intervalos: un 66.32% en ± 0.25 D y un 91.58% en ± 0.50 D respectivamente. Además es la única fórmula que introduce el 100% de los resultados refractivos en el rango ± 1.00 D. En segunda instancia, las fórmulas que incluye más ojos en los dos primeros intervalos son

PEARL-DGS y Haigis que introducen un 61.05% en el rango de ± 0.25 D y Barrett Universal II que introduce un 91.58% en el rango de ± 0.50 D igualando a Karmona®.

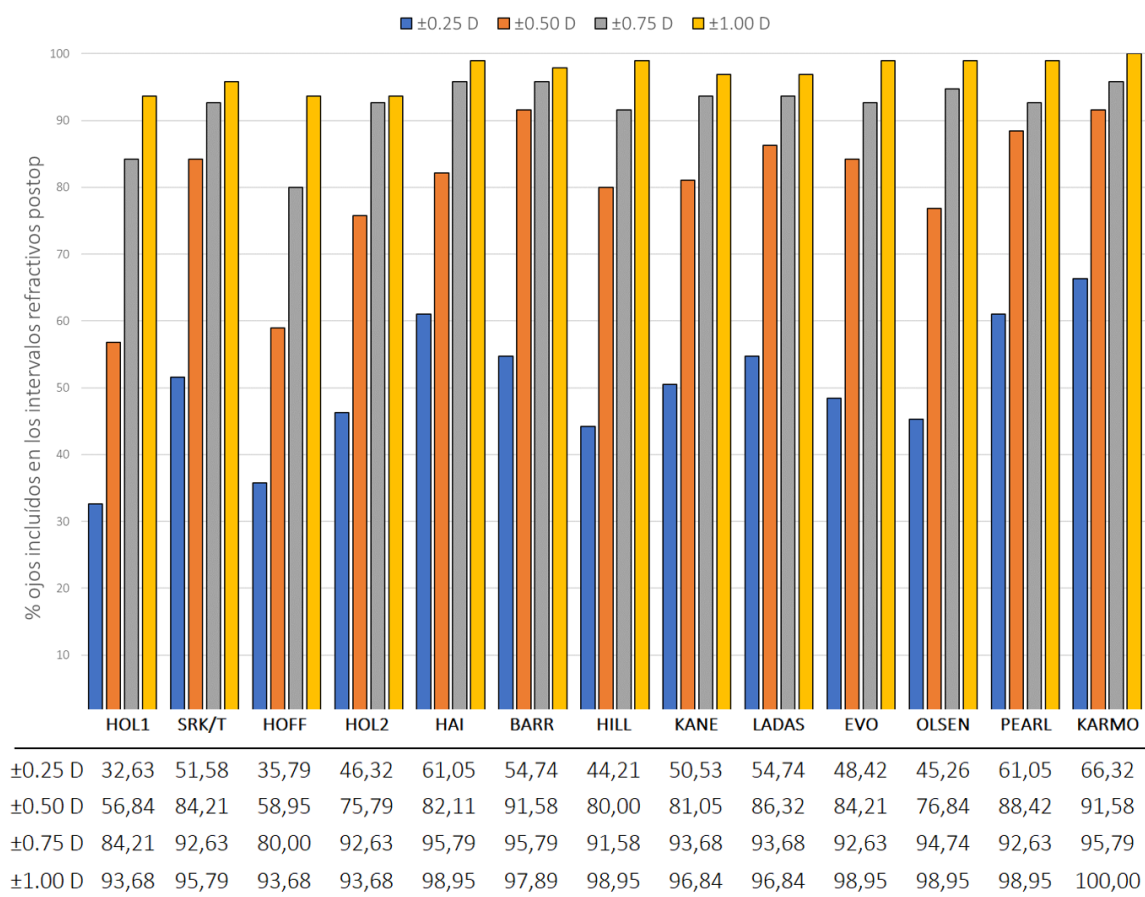


Figura 71. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra ($n = 95$) sin ajustar sobre MPE.

El test de *Cochrane Q* confirmó que no hubo diferencias estadísticamente significativas ($p < 0.01$) entre fórmulas en la muestra sin ajustar MPE en ninguno de los rangos refractivos postoperatorios estudiados, confirmándose este particular mediante el análisis *post hoc* usando *McNemar* para comparaciones múltiples.

La Figura 72 muestra los porcentajes de ojos incluidos pero con los errores predictivos corregidos. En esta ocasión Karmona® vuelve a ser la fórmula que más ojos introduce en los dos primeros intervalos junto con Kane, un 65.26% y un 90.53% en ± 0.25 D y ± 0.50 D respectivamente. En el intervalo ± 0.75 D Karmona® y Olsen introducen un 92.63% de ojos y por último Karmona® es la que más introduce en el rango ± 1.00 D, un 94.74% de los ojos del total de la muestra.

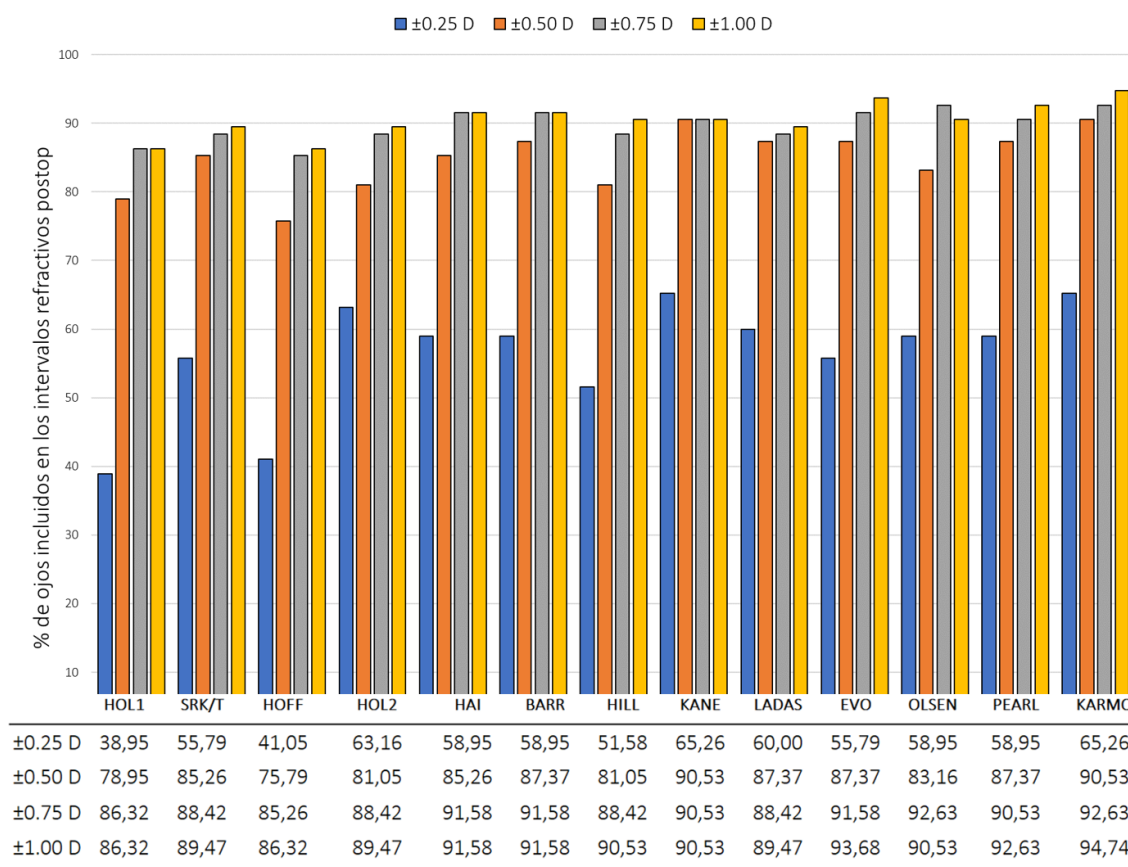


Figura 72. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra (n = 95) con ajuste sobre MPE.

El test de *Cochrane Q* demostró diferencias estadísticamente significativas ($p < 0.01$) entre fórmulas en la muestra con ajuste sobre el MPE únicamente en el rango refractivo postoperatorio de ± 0.25 D. El análisis *post hoc* usando el test *McNemar* para comparaciones múltiples confirmó que Karmona® tenía diferencias estadísticamente

significativas ($p < 0.01$) con Holladay I, no encontrándose otras diferencias para el resto de fórmulas en las comparaciones dos a dos.

7.6 Validación del modelo resultante sobre ojos cortos

La estratificación de la partición de validación por rangos de longitud axial, tal y como se ha visto en el apartado anterior, empobrece la comparación en ojos cortos o en largos dónde solo hay 12 y 25 ojos respectivamente.

Por otro lado, sabiendo que ninguna de las fórmulas empleadas hoy en día tienen un comportamiento estable y confiable en todos los rangos biométricos y con toda suerte de características biométricas, se decidió poner a prueba a Karmona® con otras muestras accesorias que contuvieran estas propiedades (Figura 73).

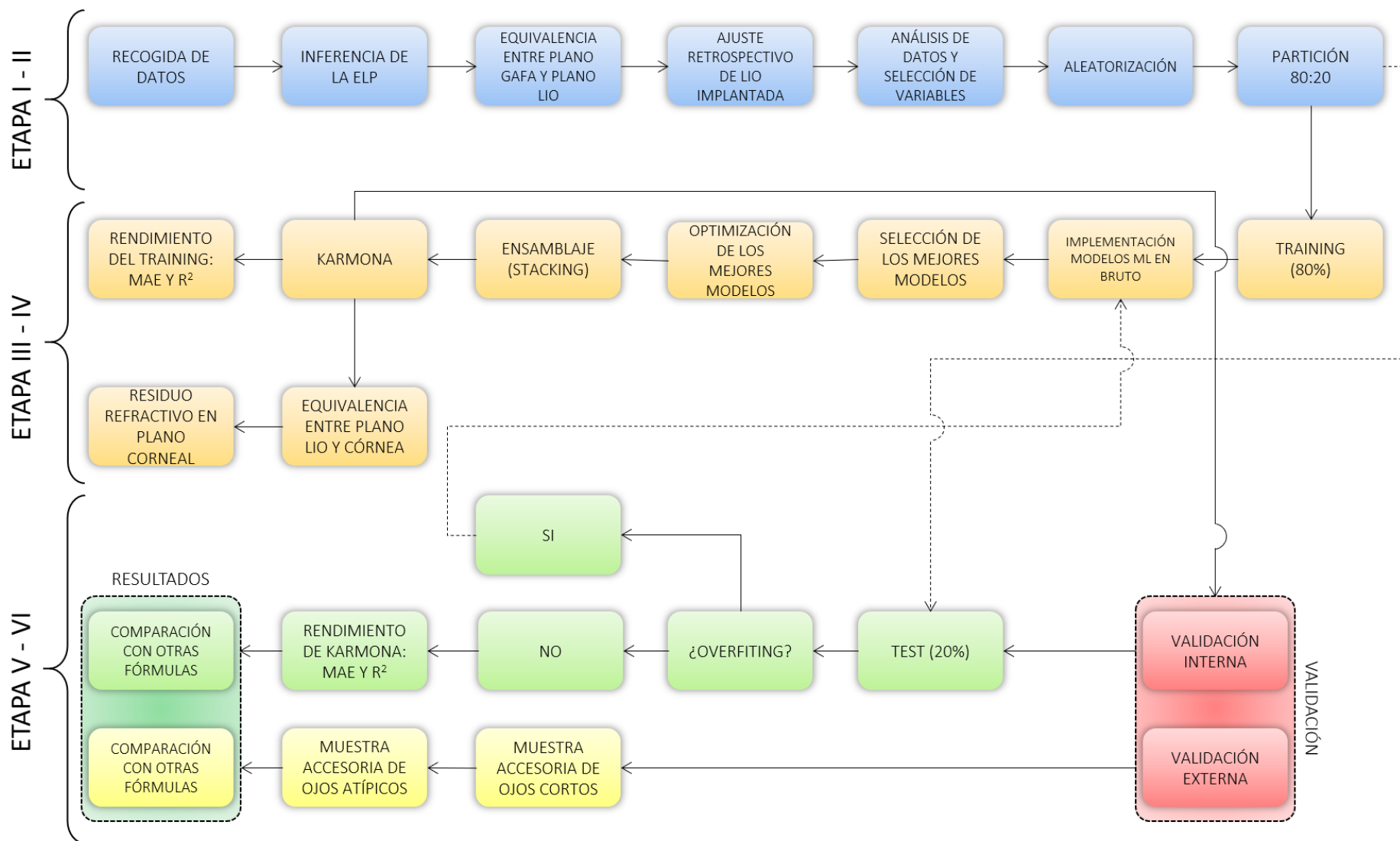


Figura 73. Esquema de flujo de trabajo con la inclusión de las nuevas muestras accesorias (amarillo) de ojos axialmente cortos y anatómicamente atípicos.

Se consiguió una muestra accesoria de 33 ojos axialmente cortos que cumplían los mismos criterios de exclusión e inclusión que en el estudio retrospectivo, con la salvedad que la longitud axial debía ser inferior o igual a 22.00 mm. Las fórmulas que se utilizaron para comparar con Karmona® fueron dos teóricas de las más utilizadas hoy en día: Haigis, Barrett Universal II y tres basadas en IA de las más actuales: EVO v2.0, Hill-RBF v3.0 y PEARL-DGS v2.0.

La única lente intraocular que se utilizó en esta muestra fue monofocal TECNIS® 1 ZCB00 con su constante A optimizada para Barrett Universal II cuyo valor fue de 119.3. Fueron medidos preoperatoriamente con los mismo dispositivos, graduados al mes de la cirugía por el mismo optometrista (D.C.G.) y operados por el mismo cirujano siguiendo el mismo protocolo para todos.

Habiendo auditado sobre la partición de validación con 95 ojos de tipología variada y confirmando su rendimiento Karmona® sin sobreajuste, además de comparado con 12 fórmulas de uso habitual en la práctica diaria, se observó que Karmona® es un buen modelo matemático que rinde muy bien y capaz de competir con las mejores fórmulas y métodos de cálculo actuales.

Ahora bien, si en el apartado anterior se vió que para ojos axialmente normales el comportamiento de todas las fórmulas es muy parecido, con pequeñas diferencias no significativas estadísticamente, en ojos cortos no es tanto así.

A continuación se muestran todos los resultados comparativos entre Karmona® y otras fórmulas en la muestra accesoria de 33 ojos cortos. La Tabla 24 muestra la distribución demográfica de la muestra de ojos cortos (n = 33).

	AXL (mm)	ACD (mm)	LT (mm)	WTW (mm)	KM ANT (D)	KM POST (D)
MEAN	21,46	2,60	4,91	11,80	44,23	-6,38
SD	0,47	0,41	0,40	0,36	17,35	0,19
MEDIAN	21,58	2,57	4,88	11,80	44,28	-6,40
MAX	21,97	3,51	5,56	12,59	45,95	-6,00
MIN	20,29	1,98	4,18	11,10	41,83	-6,80

Tabla 24. Variables biométricas de la muestra de ojos cortos ($n = 33$)

La Tabla 25 muestra los resultados de comparación de rendimiento entre las 6 fórmulas con las métricas clásicas. Karmona[®] fue la mejor con un MAE \pm SD de 0.38 ± 0.28 D seguida de EVO y Hill-RBF v3.0 con 0.40 ± 0.29 D y 0.40 ± 0.32 D respectivamente. Los valores de MedAE fueron 0.31 D, 0.33 D, 0.34 D, 0.36 D y 0.46 D para Hill-RBF v3.0, Karmona[®], PEARL.DGS, EVO v2.0, Haigis y Barrett Universal II respectivamente. Karmona[®] es la fórmula que menor error comete sobre la muestra tanto en valores negativos como positivos, -0.70 D y 0.94 D respectivamente.

	KARMONA	BARRETT	HAIGIS	PEARL	HILL	EVO
MAE (D)	0,38	0,47	0,45	0,41	0,40	0,40
SD (D)	0,28	0,36	0,36	0,30	0,32	0,29
MedAE (D)	0,33	0,46	0,36	0,34	0,31	0,34
MAX ERROR (D)	0,94	1,82	1,45	1,07	1,04	1,07
MIN ERROR (D)	-0,70	-1,82	-1,09	-0,88	-1,01	-0,86
± 0.25 D (%)	39,39	33,33	33,33	33,33	42,42	36,36
± 0.50 D (%)	60,61	57,58	69,70	69,70	60,61	66,67

*Sin ajuste sobre Mean Predictive Error.

**Muestra para un $n = 33$ ojos cortos ($AXL \leq 22.00$ mm).

Tabla 25. Resultados con métricas clásicas para 6 fórmulas sin ajuste sobre MPE en toda la muestra ($n = 33$ ojos).

Aplicando el ajuste sobre MPE las métricas de rendimiento se pueden observar en la Tabla 26. Karmona® vuelve a ser la mejor con 0.38 ± 0.27 D seguida de PEARL-DGS con un MAE \pm DS de 0.40 ± 0.27 D. Karmona® introduce en el rango refractivo postoperatorio ± 0.25 D un 42.42 % de los ojos, así como es la que menos intervalo de error produce [0.86 , -0.70] D nuevamente.

	KARMONA	BARRETT	HAIGIS	PEARL	HILL	EVO
MAE (D)	0,38	0,47	0,44	0,40	0,41	0,41
SD (D)	0,27	0,36	0,36	0,27	0,30	0,28
MedAE (D)	0,35	0,45	0,33	0,35	0,31	0,37
MAX ERROR (D)	0,86	1,81	1,48	0,86	1,09	0,99
MIN ERROR (D)	-0,70	-1,82	-1,09	-0,86	-1,09	-0,94
± 0.25 D (%)	42,42	33,33	30,30	33,33	39,39	33,33
± 0.50 D (%)	63,64	57,58	69,70	69,70	57,58	63,64

*Con ajuste sobre Mean Predictive Error.

**Muestra para un $n = 33$ ojos cortos ($AXL \leq 22.00$ mm).

Tabla 26. Resultados con métricas clásicas para 6 fórmulas con ajuste sobre MPE en toda la muestra ($n = 33$ ojos).

El análisis *post hoc* usando *Wilcoxon signed-rank* con ajuste de *Bonferroni* en las comparaciones múltiples de la muestra de ojos cortos ($n = 33$), no reveló ninguna diferencia estadísticamente significativa entre ninguna de las fórmulas estudiadas, ni en la muestra ajustada sobre MPE ni en la muestra sin ajuste.

Es remarcable que de las cuatro primeras fórmulas, tres de ellas se basan en IA, relegando a las últimas posiciones de la clasificación las teóricas: Barrett Universal II y Haigis.

En los diagramas de cajas y bigotes (Figura 74) de los errores absolutos se observa como Karmona® es la fórmula que tiene la media de MAE más baja, sin la presencia de *outliers* y con una distribución de los datos muy simétrica.

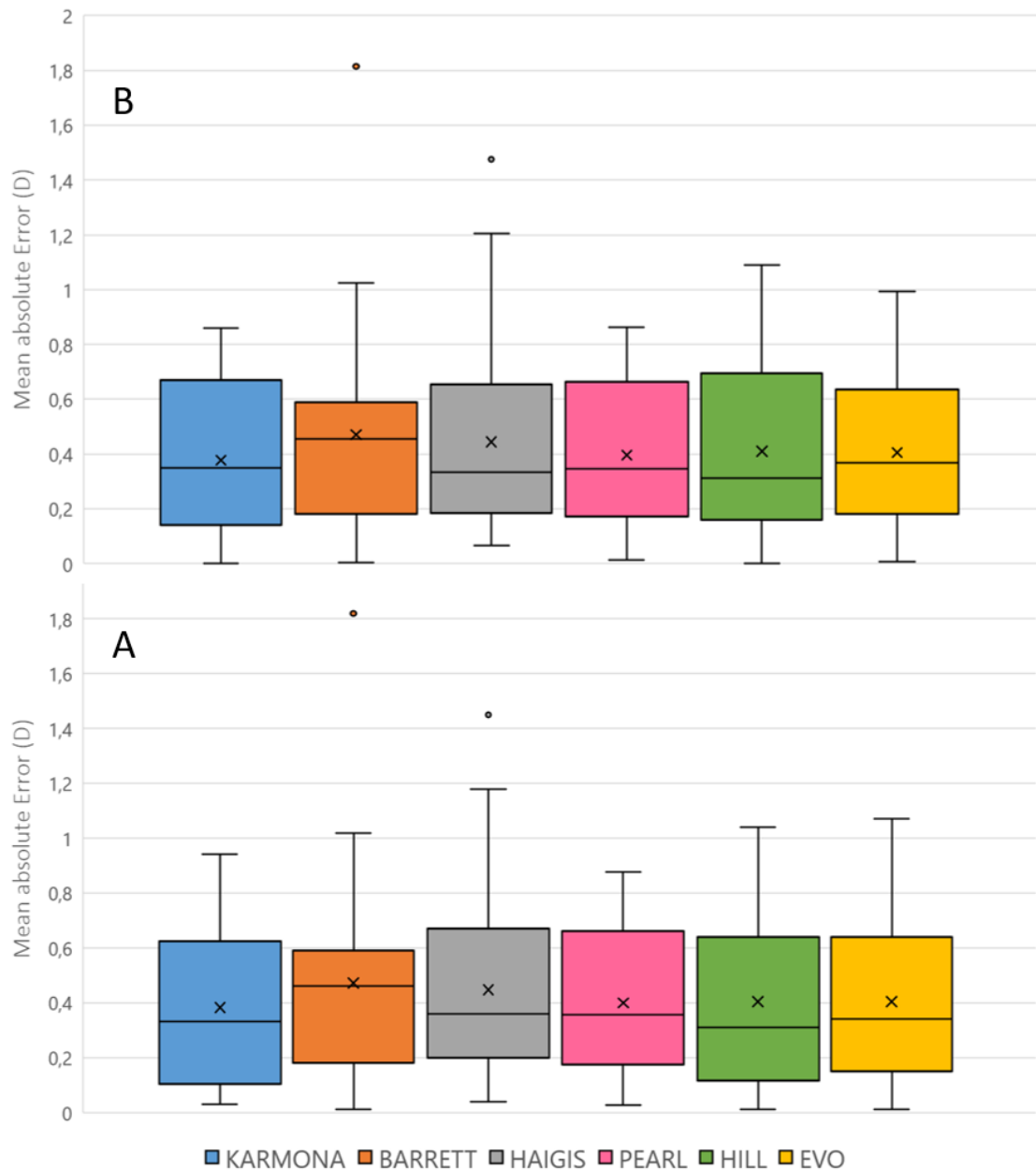


Figura 74. A) Boxplot de las 6 fórmulas para la muestra completa (n = 33) sin ajuste sobre MPE. B) Boxplot de las 6 fórmulas para la muestra completa (n = 33) con ajuste sobre MPE.

La Tabla 27 muestra el rendimiento de las fórmulas medido con las métricas propuestas en esta tesis en el capítulo (11.2 ANEXO 2.- ANÁLISIS CRÍTICO: MÉTRICAS PARA EVALUAR LA PRECISIÓN Y EXACTITUD DE FÓRMULAS) para toda la muestra (n = 33).

	KARMONA	BARRETT	HAIGIS	PEARL	HILL	EVO
IOF25	0,689	0,607	0,630	0,665	0,685	0,672
IOF50	0,714	0,633	0,697	0,703	0,710	0,680
MEAN IOF	0,702	0,620	0,664	0,684	0,698	0,676
ANTI-MAE (D)	0,325	0,408	0,384	0,351	0,328	0,345
ANTI-TMAE10 (D)	0,325	0,401	0,378	0,348	0,329	0,347
ANTI-MedAE(D)	0,330	0,460	0,360	0,340	0,310	0,340
ANTI-MAD (D)	0,044	0,038	0,040	0,043	0,057	0,040

*Muestra para un n = 33 ojos cortos (AXL ≤ 22.00 mm).

Tabla 27. Rendimiento de las 6 fórmulas empleando las métricas descritas en esta tesis sin ajuste sobre MPE.

Al igual que con el análisis clásico, Karmona® presenta las mejores métricas tras la transformación de los datos. Un IOF_{MEAN} de 0.702 frente al 0.698 de Hill-RBF v3.0 que es la segunda mejor fórmula. El ANTI-MAE de Karmona® es 0.325 D frente a 0.328 de Hill-RBF v3.0. El ANTI-TMAE10 en Karmona® es 0.325 D frente a 0.328 D de Hill-RBF v3.0, valores idénticos a ANTI-MAE en ambos casos, lo que implica la no aparición de *outliers* que afecten a los promedios.

Aplicando el ajuste *zeroed out* sobre MPE, las métricas propuestas en esta tesis se muestran en la Tabla 28.

	KARMONA	BARRETT	HAIGIS	PEARL	HILL	EVO
IOF25	0,705	0,607	0,631	0,669	0,671	0,661
IOF50	0,736	0,635	0,699	0,717	0,691	0,681
MEAN IOF	0,720	0,621	0,665	0,693	0,681	0,671
ANTI-MAE (D)	0,312	0,406	0,382	0,344	0,345	0,348
ANTI-TMAE10 (D)	0,318	0,401	0,372	0,349	0,349	0,351
ANTI-MedAE(D)	0,349	0,454	0,334	0,348	0,312	0,368
ANTI-MAD (D)	0,051	0,038	0,038	0,036	0,048	0,041

*Muestra para un $n = 33$ ojos cortos ($AXL \leq 22.00$ mm).

Tabla 28. Rendimiento de las 6 fórmulas empleando las métricas descritas en esta tesis con ajuste sobre MPE.

Karmona® es la que mejor rendimiento ofrece con un IOF_{MEAN} de 0.720, un ANTI-MAE de 0.312 D y un ANTI-TMAE10 de 0.318 D. PEARL-DGS es la que ocupa el segundo lugar con un IOF_{MEAN} de 0.693, un ANTI-MAE de 0.344 D y un ANTI-TMAE10 de 0.349 D.

Tras la transformación de los datos y el empleo de las métricas propuestas, el análisis de varianza no mostró diferencias estadísticamente significativas ($p < 0.01$) entre ninguna de las fórmulas en la categoría de ANTI-MAE ni en la muestra ajustada sobre MPE ni sin ajustar, confirmándose este hecho en las comparaciones múltiples *post hoc* con ajuste de *Bonferroni*.

Hill-RBF v3.0 es la fórmula que más valores introduce en el rango refractivo postoperatorio de ± 0.25 D, un 42.42 % de ojos, seguida de Karmona® con un 39.39 %. En el rango ± 0.50 D Haigis y Pearl-DGS son las que más ojos incluyen en ese intervalo, un 69.70 % en ambos casos (Figura 75).

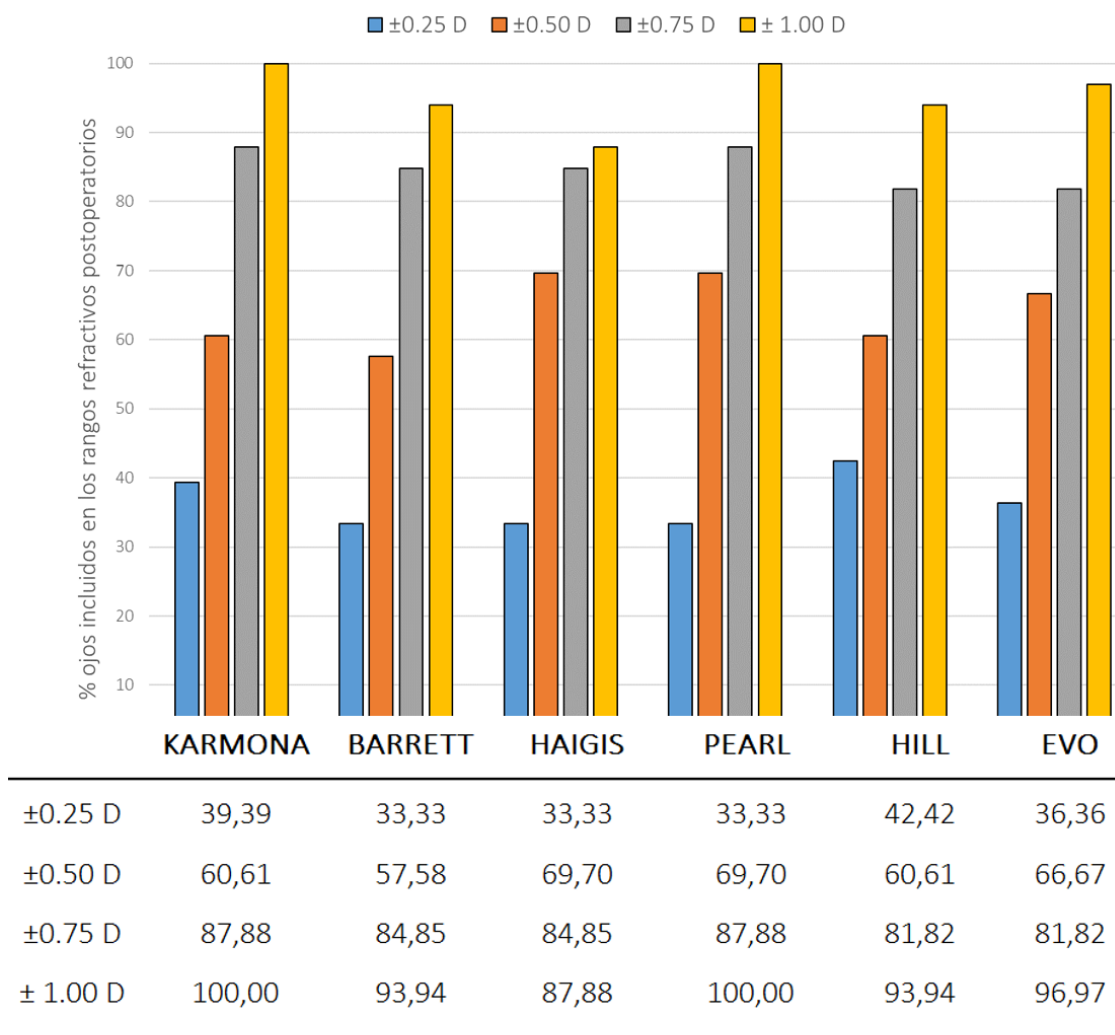


Figura 75. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra ($n = 33$) sin aplicar ajuste sobre MPE.

La Figura 76 muestra el porcentaje de ojos incluidos en los distintos intervalos refractivos postquirúrgicos con el ajuste de media a 0.00 D sobre MPE. Habiendo hecho el *zeroed out* sobre MPE, Karmona® es la fórmula que más ojos incluye en el primer rango de ± 0.25 D, un 42.42%.

El test de *Cochrane Q* no demostró diferencias estadísticamente significativas ($p < 0.01$) en la muestra sin y con ajuste sobre el MPE en ninguno de los rangos refractivos postoperatorios estudiados. Efectivamente, el análisis *post hoc* usando el test *McNemar*

para comparaciones múltiples dos a dos confirmó la no existencia de diferencias estadísticamente significativas ($p < 0.01$) entre ninguna de las fórmulas estudiadas.

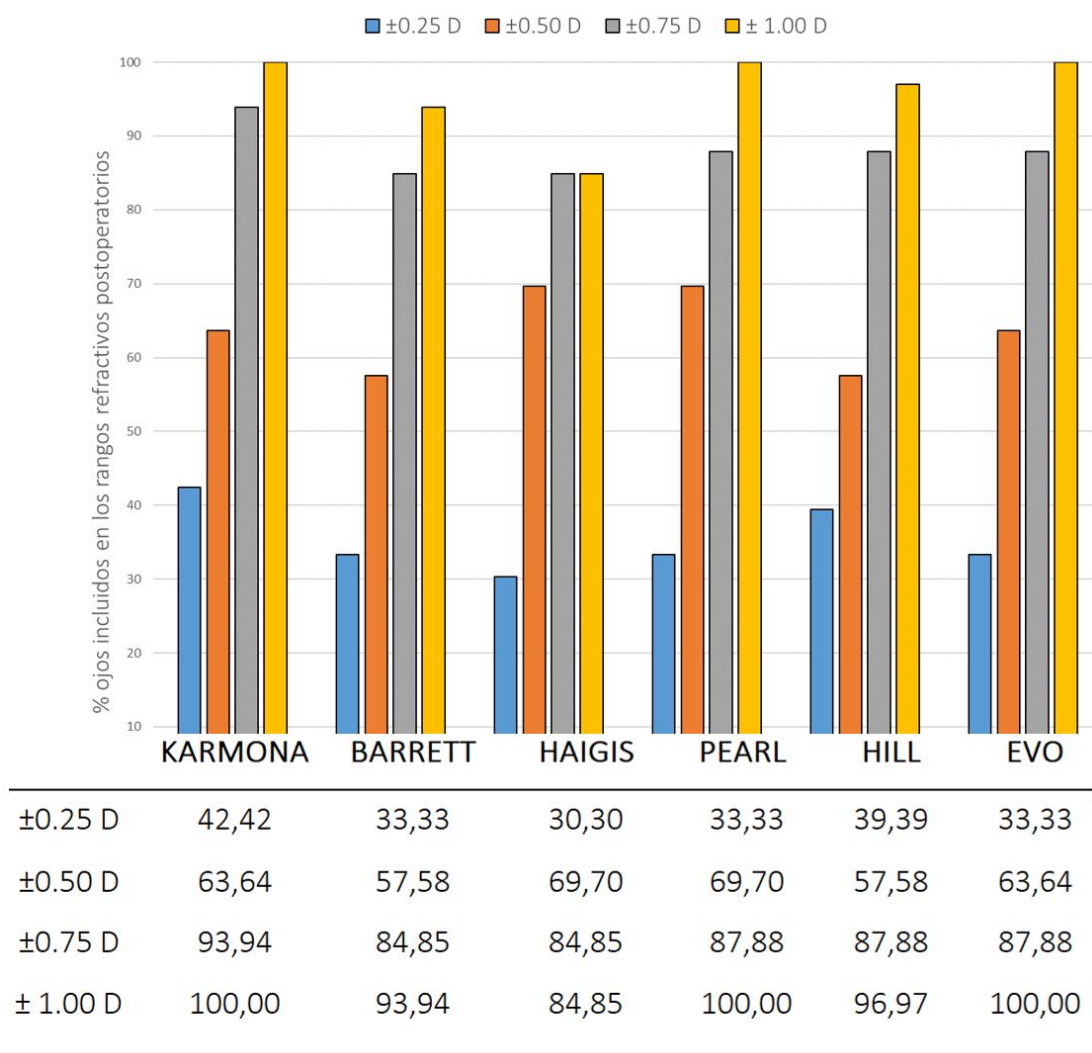


Figura 76. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra ($n = 33$) con ajuste sobre MPE.

7.7 Validación del modelo resultante sobre ojos atípicos

Se consiguió otra muestra accesoria de 41 ojos atípicos que cumplieran los mismos criterios de exclusión e inclusión que en el estudio retrospectivo, pero que computaron

como atípicos e infrecuentes según el índice de singularidad propuesto en esta tesis en el anexo [11.3 ANEXO 3.-NORMALIZACIÓN: ANÁLISIS DE POBLACIÓN BIOMÉTRICA](#). Las fórmulas que se utilizaron para comparar con Karmona® fueron dos teóricas de las más utilizadas hoy en día: Haigis, Barrett Universal II y tres basadas en IA de las más actuales: EVO v2.0, Hill-RBF v3.0 y PEARL-DGS v2.0.

La Tabla 29 muestra la distribución de las variables biométricas en la muestra de ojos atípicos, donde se pueden observar ojos con ACD de 1.85 mm, LT de 6.00 mm, KM de 33.81 D, WTW de 13.11 mm y KM posterior de -5.70 D por ejemplo.

	AXL (mm)	ACD (mm)	LT (mm)	WTW (mm)	KM ANT (D)	KM POST (D)
MEAN	23,33	3,00	4,81	12,07	43,01	-6,08
SD	1,89	0,48	0,88	0,37	2,16	0,21
MEDIAN	22,78	3,08	4,67	11,99	43,41	-6,10
MAX	33,50	3,82	6,00	13,11	46,29	-5,70
MIN	22,04	1,85	3,88	11,30	33,81	-6,60

Tabla 29. Descriptiva de las variables biométricas de la muestra de ojos atípicos (n = 41).

La Tabla 30 muestra los resultados de comparación de rendimiento entre las 6 fórmulas con las métricas clásicas. Karmona® fue la mejor con un MAE \pm SD de 0.35 ± 0.47 D seguida de EVO y Hill-RBF v3.0 con 0.42 ± 0.53 D y 0.42 ± 0.73 D respectivamente.

	KARMONA	BARRETT	HAIGIS	PEARL	HILL	EVO
MAE (D)	0,35	0,43	0,52	0,43	0,42	0,42
SD (D)	0,47	0,70	0,73	0,57	0,73	0,52
MedAE (D)	0,20	0,22	0,27	0,23	0,19	0,22
MAX ERROR (D)	1,84	1,55	1,80	1,61	1,77	1,58
MIN ERROR (D)	-2,13	-2,37	-2,04	-2,66	-2,40	-2,51
± 0.25 D (%)	61,90	57,14	40,48	54,76	54,76	52,38
± 0.50 D (%)	76,19	83,33	71,43	76,19	83,33	78,57

*Muestra para $n = 41$ ojos atípicos.

Tabla 30. Resultados con métricas clásicas para 6 fórmulas ($n = 41$ ojos).

Karmona® fue la fórmula que más ojos incluyó en el rango refractivo postoperatorio de ± 0.25 D, un 62%, seguida de Barrett Universal II con un 57% y de PEARL-DGS v2.0 y de Hill-RBF v3.0, ambas con un 55%.

El análisis *post hoc* usando *Wilcoxon signed-rank* con ajuste de *Bonferroni* en las comparaciones múltiples de la muestra de ojos atípicos ($n = 41$), demostró que Karmona® y Haigis tenían diferencias estadísticamente significativas ($p < 0.01$) sin que se observaran otras diferencias para el resto de comparaciones.

En los diagramas de cajas y bigotes (Figura 77) de los errores absolutos se observa como Karmona® es la fórmula que tiene la media de MAE más baja, con la menor presencia de *outliers*.

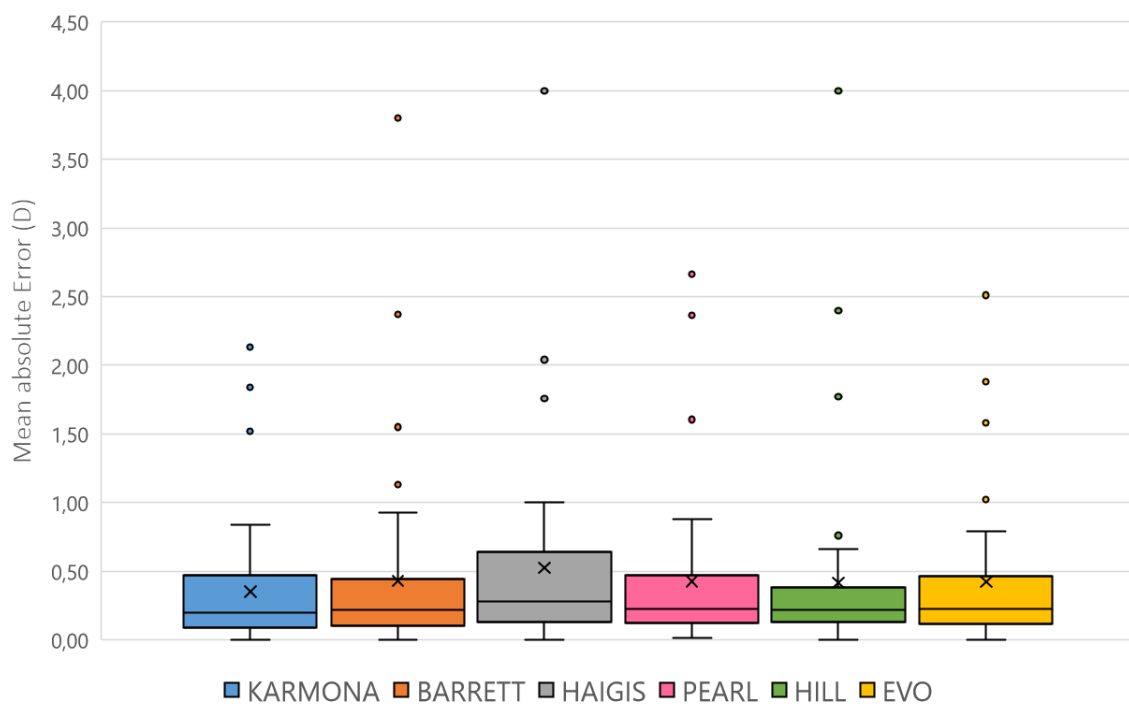


Figura 77. Boxplot para el Mean Absolute Error (D) de todas las fórmulas (n = 41).

La Tabla 31 muestra el rendimiento de las fórmulas medido con las métricas propuestas en esta tesis en el capítulo ([11.2 ANEXO 2.- ANÁLISIS CRÍTICO: MÉTRICAS PARA EVALUAR LA PRECISIÓN Y EXACTITUD DE FÓRMULAS](#)) para toda la muestra (n = 41).

	KARMONA	BARRETT	HAIGIS	PEARL	HILL	EVO
IOF25	0,773	0,730	0,632	0,689	0,741	0,708
IOF50	0,791	0,769	0,669	0,719	0,777	0,749
MEAN IOF	0,782	0,749	0,651	0,704	0,759	0,728
ANTI-MAE (D)	0,257	0,296	0,384	0,319	0,287	0,315
ANTI-TMAE10 (D)	0,234	0,256	0,350	0,286	0,244	0,290
ANTI-MedAE(D)	0,200	0,220	0,280	0,223	0,214	0,225
ANTI-MAD (D)	0,046	0,057	0,066	0,051	0,050	0,059

*Muestra para n = 41 ojos atípicos.

Tabla 31. Rendimiento de las 6 fórmulas empleando las métricas descritas en esta tesis.

Al igual que con el análisis clásico, Karmona® presenta las mejores métricas tras la transformación de los datos. Un IOF_{MEAN} de 0.782 frente al 0.759 de Hill-RBF v3.0 que es la segunda mejor fórmula. El ANTI-MAE de Karmona® es 0.257 D frente a 0.287 de Hill-RBF v3.0. El ANTI-TMAE10 en Karmona® es 0.234 D frente a 0.244 D de Hill-RBF v3.0.

Karmona® es la fórmula que más valores introduce en el rango refractivo postoperatorio de ± 0.25 D, un 61.90 % de ojos, seguida de Barrett con un 57.14 %. En el rango ± 0.50 D Barrett Universal II y Hill-RBF v3.0 son las que más ojos incluyen en ese intervalo, un 83.33 % en ambos casos (Figura 78). Karmona® es la que más ojos incluye en los siguientes rangos estudiados ± 0.75 D y ± 1.00 D.

El test de *Cochrane Q* demostró diferencias estadísticamente significativas ($p < 0.01$) en la muestra en el rango de ± 0.25 D y en ninguno más. Efectivamente, el análisis *post hoc* usando el test *McNemar* para comparaciones múltiples dos a dos confirmó la existencia de diferencias estadísticamente significativas ($p < 0.01$) entre Karmona® con Haigis y EVO v2.0.

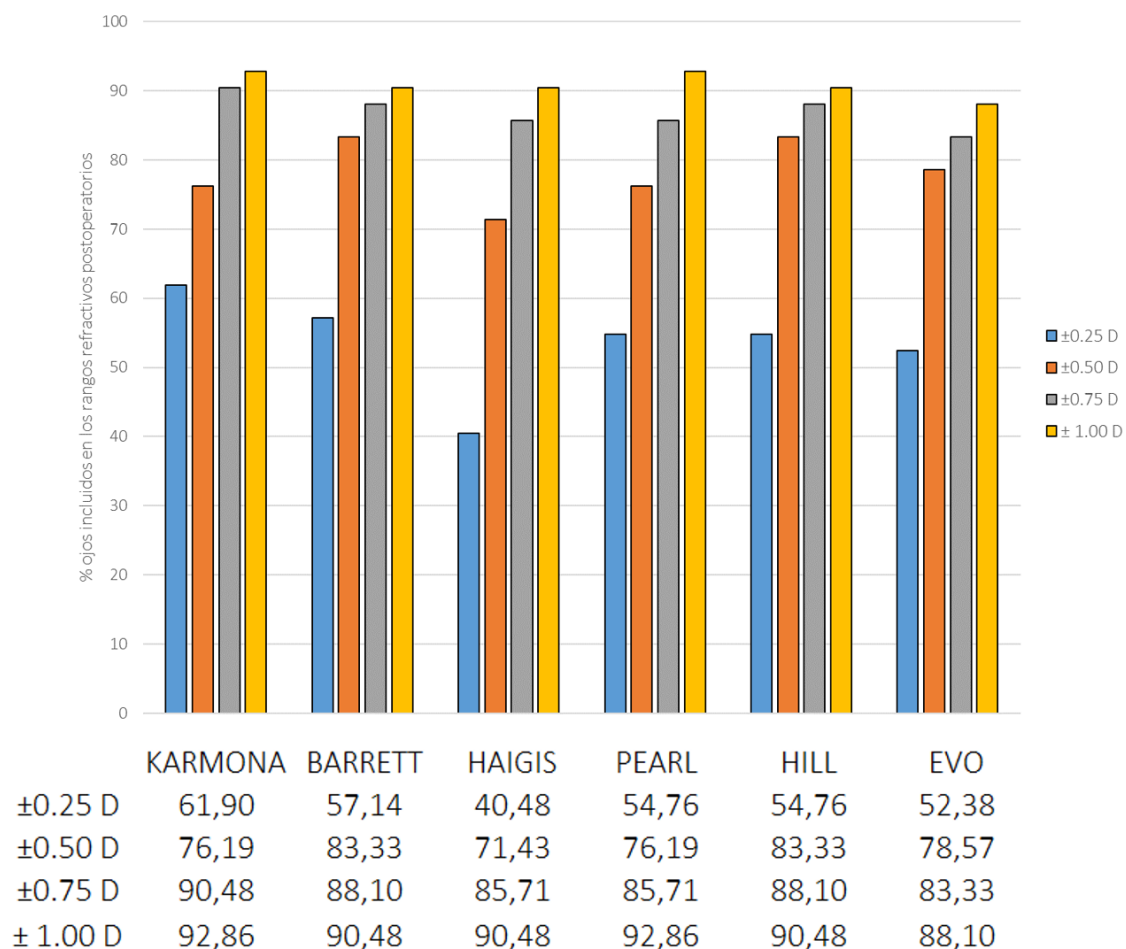


Figura 78. Porcentaje de ojos incluidos en los intervalos refractivos postoperatorios de ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D en toda la muestra (n = 41).

7.8 Comparación entre fórmulas teóricas y basadas en IA

En la muestra de validación de 95 ojos para el total de la muestra con ajuste sobre MPE y utilizando las métricas propuestas en la tesis, las fórmulas que mejor rendimiento tienen son las que funcionan con IA, siguiendo este orden: Karmona®, Kane y PEARL-DGS con IOF_{MEAN} 0.865, 0.848 y 0.828 respectivamente. Por debajo está la primera fórmula teórica, Holladay II con un IOF_{MEAN} de 0.821 seguida de otra con IA, Ladas SPF presentando un IOF_{MEAN} de 0.819. Barrett Universal II ocupa el sexto lugar con un IOF_{MEAN} de 0.811.

En la muestra accesoria de ojos cortos ($n = 33$) y atendiendo de nuevo al índice de fórmula propuesto en esta tesis, las fórmulas de IA vuelven a ser mejores que las teóricas con el siguiente ranking: Karmona[®], Hill-RBF v3.0 y PEARL-DGS con IOF_{MEAN} de 0.702, 0.698 y 0.684 respectivamente, seguidas de las teóricas por este orden: EVO v2.0, Haigis y Barrett Universal II con IOF_{MEAN} de 0.676, 0.664 y 0.620 respectivamente.

Por último, en la muestra de ojos atípicos ($n = 41$) y basándose en IOF_{MEAN} las fórmulas que funcionan con IA de nuevo vuelven a estar por encima de las teóricas. En este caso la ordenación será Karmona[®] y Hill-RBF v3.0 con IOF_{MEAN} de 0.782 y 0.759 respectivamente, seguidas de las teóricas Barrett Universal II y EVO v2.0 con IOF_{MEAN} de 0.749 y 0.728 respectivamente. Completan la clasificación PEARL-DGS otra fórmula basada en IA y en último lugar la teórica Haigis.


En ninguno de los rangos evaluados hubo diferencias estadísticamente significativas entre ninguna de las fórmulas basadas en IA.

7.9 Presentación de resultados mediante página web

Se ha diseñado y programado una página web profesional de libre acceso, con un dominio propio (www.karmona-iol.com) que incluye el algoritmo Karmona[®] IOL Power Calculation, y permite al usuario calcular la potencia de la lente intraocular que quiere implantar en función de los inputs biométricos (Figura 79).

La página y su algoritmo están en proceso registral en el Registro de la Propiedad Intelectual de Madrid con número de expediente 09/832047.9/21 .

Karmona
Singularity
Researcher
About
Contact



Patient name

ID

Surgeon

Gender

Unknown ▾

Print

Biometric data

AXL (mm)
 (15,00 / 40,00)

ACD (mm)
 (0,00 / 7,00)

LT (mm)
 (1,00 / 10,00)

WTW (mm)
 (6,00 / 18,00)

IOL A-CST.

Calculate

Keratometric data

KM ANT (D)
 (n = 1.3375)

KM POST (D)
 (-2,00 / -10,00)

Ktotal ~
 Dioplers

Ratio B/F
 %

Target (D)
 (-30,00 / +30,00)

Clean All

Select Anterior K Source


IOL Master® 700 ▾

Select Posterior K Source

None ▾

Select Eye for Surgery

Right Eye ▾



Results

IOL power (D)	Refraction postop (D)

Powered by David Carmona Gonzalez

Figura 79. Pantalla principal de la web www.karmona-iol.com.

El *website* está programado con Shiny de RStudio (<https://shiny.rstudio.com/>) en lenguaje R (<https://www.r-project.org>). Los *scripts*, bases de datos, imágenes y todas las dependencias de R, están *dockerizados* y alojados en un servidor Linux con certificado SSL y nodo localizado en Francia, que permite un acceso global sin latencias.

La pantalla principal (Figura 79) consta de 5 pestañas superiores que dirigen al usuario a diferentes escenarios dentro del mismo portal web: Karmona®, Singularity, Researcher, About y Contact.

Karmona®, la primera pestaña, aloja el propio algoritmo de cálculo y permite obtener al usuario un resultado imprimible una vez rellenos los datos del formulario principal. Como cualquier calculador moderno hay que completar los campos de identificación del caso en cuestión: nombre y apellidos del paciente con un número de identificación asociado para la trazabilidad postquirúrgica, así como el nombre del cirujano y el género del individuo. En cuanto al género, únicamente tiene relevancia para el análisis de la singularidad pero no para el cálculo de la potencia de la lente intraocular.

A continuación, tras cumplimentar todos los campos de datos biométricos necesarios, se mostrará en la parte superior de la columna derecha el resultado de la potencia de la lente emetropizante que se necesita para el *target* refractivo seleccionado previamente, además del despliegue ordenado de mayor a menor de cinco potencias de lentes que permiten saber al usuario qué lente es mejor para implantar según sus necesidades (Figura 80).


En el formulario de datos biométricos el usuario podrá utilizar tanto comas como puntos en la introducción de datos decimales, a diferencia de otros calculadores, así como no será necesario introducir los datos queratométricos de ambos meridianos

principales de la córnea con sus respectivas orientaciones, con introducir el valor medio queratométrico será suficiente. Una vez completado el formulario con todos los datos necesarios y habiendo calculado, se mostrarán la potencia total corneal y el ratio entre la cara posterior y anterior como información complementaria para el usuario, ya que estas dos variables son dos parámetros predictores para el modelo Karmona[®], tal y como ya se explicó en [7.1.2 Correlación entre variables y selección de características](#).

En la zona media de la pantalla principal (Figura 80), el usuario podrá seleccionar el dispositivo queratométrico con el que se midió al paciente la potencia corneal anterior y posterior. Aunque la web ofrece varias posibilidades para la queratometría anterior: “IOL Master[®] 700”, “Pentacam[®]” y “Sirius[®]”, solo está activa la primera ya que la optimización para las demás no está preparada actualmente. De igual modo, las únicas opciones activas hoy por hoy para la queratometría posterior son “Pentacam[®]” y “None”, porque Karmona[®] también permite calcular sin los datos de la superficie posterior corneal. En un futuro todas estas opciones estarán implementadas para ofrecer mejores resultados.

El algoritmo Karmona[®] funciona con cualquier dispositivo queratométrico tanto de cara anterior como posterior de la córnea. No obstante, el algoritmo está optimizado para IOL Master[®] 700 en medidas de la córnea anterior y para Pentacam[®] HR en cara posterior.

Karmona Singularity Researcher About Contact



Patient name

ID

Surgeon

Gender

[Print](#)

Biometric data

AXL (mm)
 (15,00 / 40,00)

ACD (mm)
 (0,00 / 7,00)

LT (mm)
 (1,00 / 10,00)

WTW (mm)
 (6,00 / 18,00)

IOL A-CST.

[Calculate](#)

Keratometric data

KM ANT (D)
 (n = 1.3375)

KM POST (D)
 (-2,00 / -10,00)

Ktotal ~
 Diopters

Ratio B/F
 %


Target (D)
 (-30,00 / +30,00)

[Clean All](#)

Select Anterior K Source

Select Posterior K Source

Select Eye for Surgery



Results

Predicted IOL: 19.124 for target refraction 0.00 D.

IOL power (D)	Refraction postop (D)
20.00	-0.6
19.50	-0.26
19.00	0.09
18.50	0.43
18.00	0.77

Powered by David Carmona Gonzalez

Figura 80. Pantalla de resultados de Karmona® en la que se muestra la potencia de la lente emetropizante y 5 potencias redondeadas con precisión de 0.50 D, así como las predicciones refractivas asociados a dichas potencias dióptricas.

Además la web es absolutamente operativa permitiendo seleccionar el ojo a operar así como imprimir los resultados para incorporarlos al plan quirúrgico. No se guardan, ni se exportan, ni almacenan datos empleados en el cálculo, ni ningún dato relacionado con la filiación del paciente, acorde con la Ley Orgánica 3/2018 de 5 de diciembre de protección de datos personales y garantía de los derechos digitales.

Por otro lado, nada más iniciar la web se despliega una ventana *popup* (Figura 81) con un pliego de descargo (*Disclaimer*) en el que se exime al autor de responsabilidad sobre el uso que se le de al calculador al igual que la información exhibida no es garantía de nada, recordando además al usuario de la única y última responsabilidad del cirujano sobre la elección de la potencia de la lente intraocular a implantar, del mismo modo que aparece en otros calculadores.

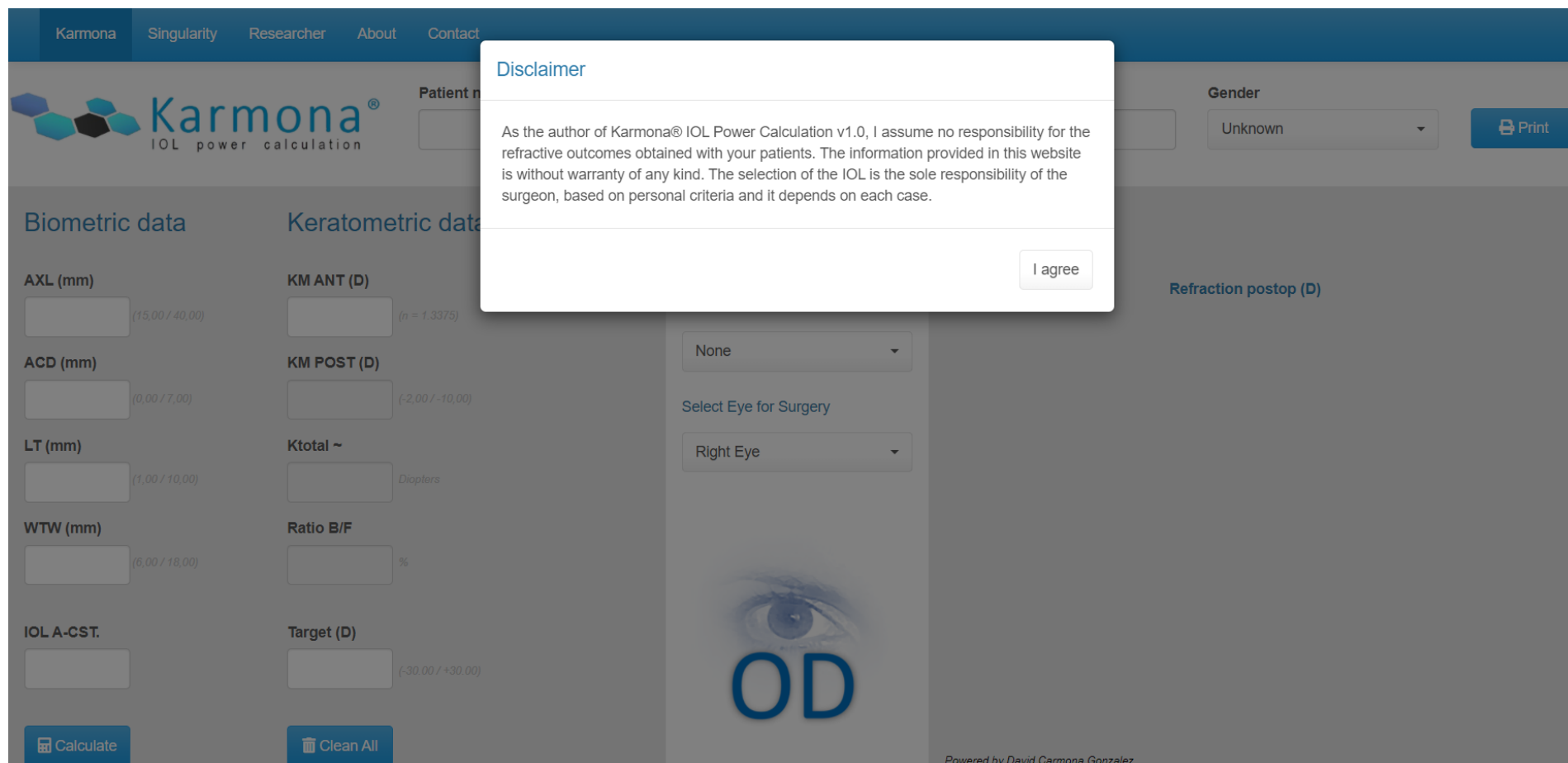


Figura 81. Mensaje de descargo de responsabilidad en el uso de Karmona® que aparece nada más acceder a la web.

La segunda pestaña de la zona superior, *Singularity* dirige al usuario a la pantalla de análisis de singularidad en la que se mostrarán tal y como se expresó [11.3 ANEXO 3.- NORMALIZACIÓN: ANÁLISIS DE POBLACIÓN BIOMÉTRICA](#) los resultados de normalización poblacional (Figura 82 y Figura 83) en función de los *inputs* biométricos que haya introducido el usuario en la pantalla principal.

El usuario puede estudiar mediante un triple diagrama, cuánto se desvían sus datos biométricos utilizados para el cálculo de la población normal, gracias a la comparación con una muestra de 10.295 ojos ([11.3 ANEXO 3.-NORMALIZACIÓN: ANÁLISIS DE POBLACIÓN BIOMÉTRICA](#)). Además se muestra en la parte inferior de la derecha (Figura 83) la cuantificación porcentual del índice de singularidad (SI).

En la esquina superior derecha de esta pantalla sale un signo de interrogación que enseña al usuario un texto de ayuda explicándole brevemente el origen del índice SI, con la siguiente leyenda: *“Singularity Index (SI) is a parameter to quantify the anatomical rarity of an eye. It ranges from 0 to 100, being 0 normal and 100 very atypical. It is composed of the biometric rarity of 5 variables: AXL, ACD, LT, WTW and KM. Each of these variables will compute at most 20% over IS in the case of maximum rarity, or 0% when the variable in question is normal. This is based on the robust statistical analysis of a database measured with IOL Master® 700 on 10.295 caucasian eyes, without surgeries and without pathologies.”*

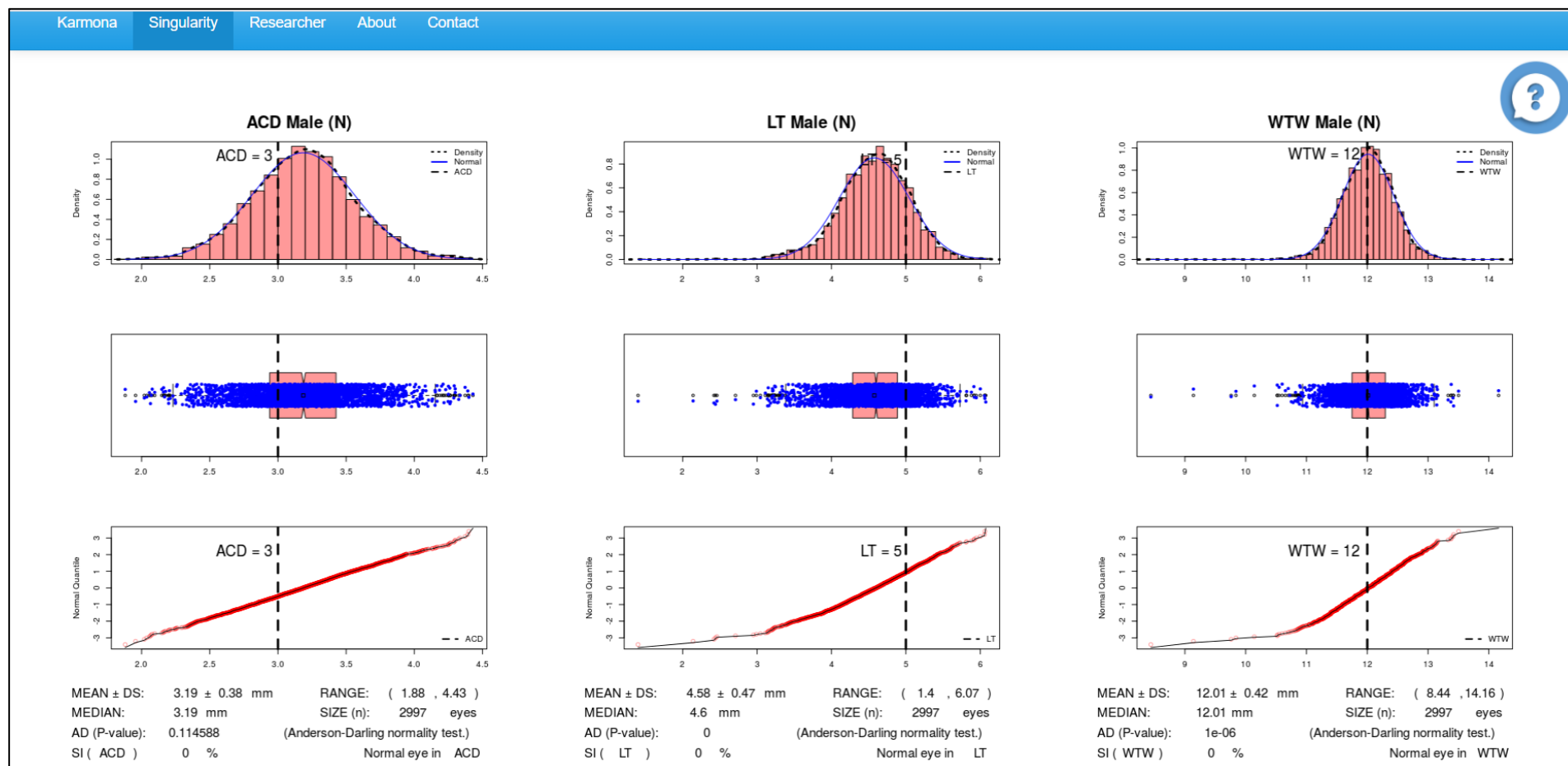


Figura 82. Resultados de Singularity en ACD, LT y WTW (primera parte).

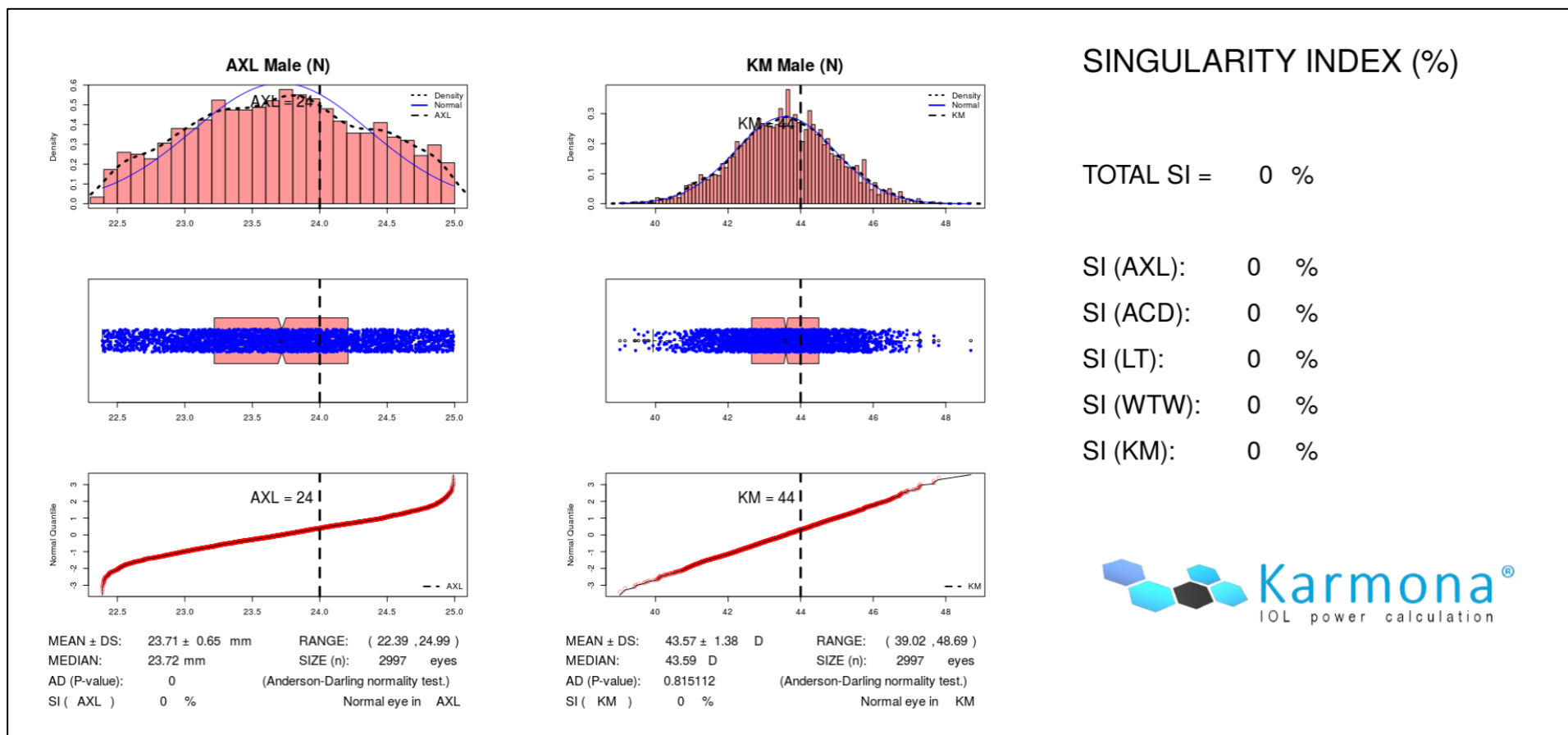


Figura 83. Resultados de Singularity en AXL y KM con el cálculo de SI (segunda parte).

“Researchers” es la tercera pestaña y conducirá al usuario a una página en la que podrá calcular de forma rápida solo en tres pasos y simultáneamente los casos que quiera con Karmona® (Figura 84). Muy útil cuando se quiere estudiar el comportamiento de rendimiento de una fórmula y compararlo con otras, pero también para la optimización de constantes con esta fórmula de una manera eficiente y rápida.

El procedimiento será muy sencillo, el usuario se descargará un archivo Excel (Microsoft Corporation, Washington, EE.UU.) estándar desde la web karmona-iol, llamado *Karmona.xlsx*, con todos los campos vacíos. A continuación se rellenarán las celdas correspondientes a cada parámetro biométrico de este excel tipo con la casuística propia del usuario y posteriormente se sube a la web de nuevo mediante el *Browse* que aparece en el panel de la izquierda de la pantalla (Figura 84).

Cuando el usuario presione “Calculate” ocurrirán dos eventos; uno la descarga automática y directa desde el servidor al ordenador del usuario del archivo excel que él mismo subió con su casuística personal, con una nueva columna añadida a la derecha titulada KARMONA que contiene los residuos refractivos predichos postoperatorios para cada caso concreto. El otro evento es la aparición de información útil para el investigador en las subpestañas ubicadas en la zona superior de la derecha (Figura 84) llamadas “Data” y “Summary”

La pestaña “Data” permite ver el archivo en tiempo real con los cálculos hechos de cada caso para comprobar posibles errores y anomalías. La pestaña “Summary” contiene un gráfico que aglutina histogramas, gráficos de dispersión y valores de correlación r entre variables.

Karmona
Singularity
Researchers
About
Contact

1. Download the standard excel file (karmona.xlsx):

Download

2. Complete excel file with your data (blind calculation: refractive outcomes is unknown).

3. Upload the excel file with your data to the server clicking on BROWSE:

Choose XLSX file:

No file selected

4. Check the excel file in the 'Data' tab. If any data is missing, please insert it and upload again before calculating.

5. Click on CALCULATE for karmona predictions:

Calculate

6. The same excel file with the calculations made will automatically be downloaded to your computer.

Help
About file
Data
Summary

For Researchers :

Massive calculations your cases with Karmona® IOL power calculation using this application. You can quickly and easily obtain results to publish your cases with three simple steps, no captchas, no limits and free:

A	B	C	D	E	F	G	H	I	J
n	AXL	ACD	LT	WTW	KM_ANT (diopters)	KM_POST (diopters)	CTE_IOL	IOL_implanted	KARMONA

1. DOWNLOAD THE EXCEL FILE.

A	B	C	D	E	F	G	H	I	J
n	AXL	ACD	LT	WTW	KM_ANT (diopters)	KM_POST (diopters)	CTE_IOL	IOL_implanted	KARMONA
1	20,36	2,30	4,69	11,50	42,78	-6,10	119,30	34,00	
2	22,55	3,39	4,38	11,90	44,46	-6,30	119,30	24,00	
3	22,50	3,46	4,39	12,10	44,19	-6,30	119,30	24,50	
4	21,77	2,04	5,09	11,30	44,05	-6,40	119,30	27,00	
5	21,64	2,07	5,15	11,50	44,31	-6,40	119,30	27,00	

2. COMPLETE THE EXCEL FILE AND UPLOAD.

A	B	C	D	E	F	G	H	I	J
n	AXL	ACD	LT	WTW	KM_ANT (diopters)	KM_POST (diopters)	CTE_IOL	IOL_implanted	KARMONA
1	20,36	2,30	4,69	11,50	42,78	-6,10	119,30	34,00	-0,10
2	22,55	3,39	4,38	11,90	44,46	-6,30	119,30	24,00	-0,17
3	22,50	3,46	4,39	12,10	44,19	-6,30	119,30	24,50	-0,07
4	21,77	2,04	5,09	11,30	44,05	-6,40	119,30	27,00	-0,22
5	21,64	2,07	5,15	11,50	44,31	-6,40	119,30	27,00	0,00

3. KARMONA CALCULATIONS.

Figura 84. Pantalla de "Researchers" en la web karmona-iol.com

La cuarta pestaña del menú de navegación es “About”, en la que se explica de forma somera el funcionamiento del algoritmo Karmona® (Figura 85), así como pequeñas apreciaciones sobre las variables biométricas con las que opera. Asimismo hay un *link* directo al *abstract* del *paper* original en el que la versión primigenia de Karmona® vió por primera vez la luz.

También se muestra el mismo pliego de descargo de responsabilidad (*Disclaimer*) que se expone nada más acceder a la web, y los términos de uso.

Karmona
Singularity
Researcher
About
Contact

About...

What is Karmona® v1.0?

Karmona® IOL power calculation is a new calculation method, based on artificial intelligence, to calculate the power of the intraocular lens implanted in lens surgery. It is based on two non-linear regression mathematical models that have been ensembled in a final model considering stacking method, cross validation and generalized linear model training techniques, to enhance the prediction. Obtaining the following performance metrics with the test sample (n = 100 eyes):

$R^2 = 0.995$, $MAE = 0.24$ D.

About Karmona®

Karmona® IOL power calculation is designed and programmed in Shiny-RStudio (Version 1.1.423) by David Carmona Gonzalez (Spain-Madrid, 2018-2021). It is optimized for anterior keratometry with IOL Master® 700 (Carl Zeiss Meditec, Jena, Germany) and posterior keratometry with Pentacam® (Oculus Optikgeräte GmbH, Wetzlar, Germany), however it can be used with measurements from any other devices. The required biometric variables that *Karmona®* works with are AXL, ACD, WTW, ANTERIOR KM (1.3375) and the optional ones are LT and POSTERIOR KM. It uses the A constant as the intraocular lens constant, recognizable for everyone and easy to obtain. Unlike its competitors, *Karmona®* has been developed only with a database of 386 eyes, but as any other 'Artificial Intelligence based model', it is constantly learning, so its predictive capacity is improving continuously. It is remarkable that *Karmona®* has been designed to increase the accuracy in extreme and atypical eyes. In this respect, there is a paper published to see corresponding clinical results (<https://doi.org/10.1038/s41433-020-0883-3>).

Disclaimer

As the author of *Karmona®* IOL Power Calculation v1.0, I assume no responsibility for the refractive outcomes obtained with your patients. The information provided in this website is without warranty of any kind. The selection of the IOL is the sole responsibility of the surgeon, based on personal criteria and it depends on each case.

Terms of use

Karmona® is registered and under license of intellectual property. Any attempt to copy, modify or decompile will be prosecuted by law.

2021 Karmona® IOL Power Calculation. All rights reserved.
 Madrid - Spain

Figura 85. Pantalla "About" de la web karmona-iol.com

Por último, la pestaña “Contact” en la que se muestra información personal del autor (D.C.G.): nombre y apellidos, empleo, fotografía, lugar de trabajo con dirección completa y dirección de correo electrónico para cualquier consulta acerca del funcionamiento o de otras cuestiones relacionadas con el contenido de la web (Figura 86).

En la parte inferior hay dos iconos de Tweeter y LinkedIn con acceso directo a dichas redes sociales en relación al uso de la fórmula Karmona[®], para mantenerse informado de posibles actualizaciones y futuras versiones, aclaraciones sobre su uso, publicaciones científicas u otras cuestiones.

Karmona Singularity Researcher About Contact


Contact



If you have any suggestions, questions or comments do not hesitate to contact me.

My name is David Carmona Gonzalez and I am optometrist. My email address and place of work:

davcargon@hotmail.com

Hospital Universitario QuironSalud Madrid
Calle Diego de Velazquez, 1.
28223 - Pozuelo de Alarcon
Madrid (Spain)



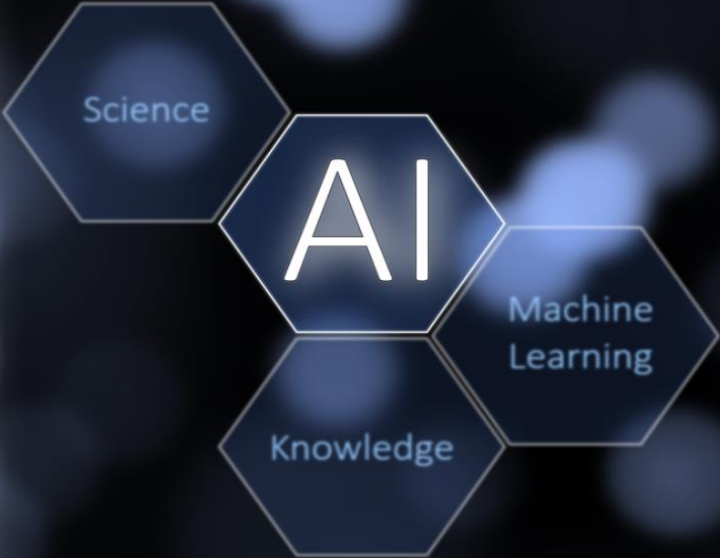




Figura 86. Pantalla "Contact" de la web karmona-iol.com

8 DISCUSIÓN

8. DISCUSIÓN

Actualmente no existe una fórmula o un método de cálculo biométrico con un comportamiento predictivo estable, exacto y preciso e independiente de la tipología de ojo, a pesar de las tecnologías de medición, de fabricación y del conocimiento adquirido durante más de cinco décadas.

Con estas pretensiones, en esta tesis se propone una nueva fórmula basada en técnicas actuales de inteligencia artificial, para predecir la potencia de la lente intraocular que se implantará en la cirugía de cristalino.

El algoritmo Karmona® ha sido diseñado teniendo en cuenta la curvatura de la superficie posterior corneal además de la queratometría simulada y hasta la fecha no existe ninguna fórmula para cálculos no tóricos que haga esto. Lo que tiene dos consecuencias directas importantes, en primer lugar, la adquisición de información aportada por otra variable queratométrica diferente, otorgando mayor precisión en el cálculo respecto al resto de fórmulas utilizadas actualmente. Por otro lado, surge una posible línea de investigación futura, el cálculo del equivalente esférico de lentes intraoculares en queratoconos, córneas ectásicas con afectación de la curvatura posterior y queratoplastias.

Otro aspecto fundamental del diseño de Karmona® es el uso de nuevo de un factor diferencial respecto al resto de las fórmulas conocidas, como es la utilización de la profundidad del segmento anterior (ASD) como variable predictora. Aunque el usuario introduce las variables que caracterizan esta longitud de forma separada, ACD y LT como se haría en cualquiera de los métodos usados habitualmente, el uso interno matemático que hace Karmona® es mediante la suma de ambas.

La idea que subyace tras esto es la diversidad anatómica y su potencial influencia en el cálculo. En este sentido, no es lo mismo un ojo con un ASD profundo, por ejemplo 8.45 mm, que otro con un segmento anterior corto de 6.00 mm. La experiencia dice que la diferencia de posicionamiento intraocular de la lente podrá ser superior a 0.50 mm en estos casos, lo que se traducirá en errores refractivos del orden de 0.75 D en ojos axialmente cortos (77, 141) no siendo tan relevante en ojos largos.

Desde una perspectiva puramente matemática el encaje de ASD no es complicado, ya que es una combinación lineal de ACD y LT que rompe la multicolinealidad existente entre AXL, ACD y la potencia implantada de la lente intraocular, condiciones todas estas idóneas para implementar modelos regresivos sin perder la información fundamental del segmento anterior.

Sabiendo que para casos no tóricos la potencia de la lente intraocular que calcula cualquier fórmula independientemente de la tecnología en la que se base, se realiza con la queratometría media para obtener en realidad el equivalente esférico. Desde el punto de vista del usuario no tiene ningún sentido introducir las potencias queratométricas de los dos meridianos principales de la córnea así como sus respectivas orientaciones. Por tanto, la interfaz de Karmona® se diseñó teniendo en cuenta este sencillo aspecto práctico, permitiendo introducir como único valor la media queratométrica sin orientaciones, confiriendo de esta forma mayor rapidez y mejor experiencia de usuario.

Un factor más diferencial respecto a otras fórmulas modernas como Hill-RBF o Kane, y a nuestro juicio relevante, es la limitación que ponen los autores a sus parámetros biométricos en los formularios de relleno, impidiendo al usuario calcular ciertos casos complejos o atípicos. Por ejemplo W. Hill impone unos límites a sus *inputs*

queratométricos entre 37.00 D y 52.00 D, no pudiendo calcular fuera de ese rango. La fórmula de Kane solo permite calcular longitudes axiales entre 18.00 mm y 35.00 mm, igualmente Hill-RBF v3.0 acorta este rango entre 19.00 mm y 35.00 mm.

Sin embargo, Karmona® aunque anuncia al usuario un límite en ciertos parámetros biométricos avisando de potenciales resultados inconsistentes, permite realizar el cálculo introduciendo cualquier valor que caracterice el ojo en cuestión, cosa que ninguna otra fórmula deja hacer.

Esto deja una puerta abierta para cálculos muy complejos en nanofthalmos, megalocórneas, atalamias, esferofaquias, queratoplastias y ojos extremadamente largos. De hecho, en esta tesis se seleccionaron dos casos muy raros en la muestra de atípicos, el primero con una córnea virgen de 33.00 D de queratometría que no se pudo calcular con Hill-RBF v3.0 y el segundo con una AXL de 35.34 mm que tampoco pudo calcularse ni con Kane, ni con Hill-RBF v3.0.

La constante más usada y más reconocida por los clínicos, es la constante A óptica de la lente, y aunque muchas fórmulas de uso actual como Barrett Universal II, Olsen, Holladay I y II o Hoffer-QST utilicen sus propias constantes y conversiones, se hace más complicado para usuarios inexpertos tratar con ellas. De manera que parece más práctico el uso masivo de la constante A como estándar en los cálculos biométricos. Karmona® se diseñó teniendo en cuenta este precepto y es la constante que utiliza para los cálculos.

El total de la muestra con la que se trabajó en la presente tesis fue de 481 ojos, de los cuales el 80% se emplearon para implementar y entrenar los algoritmos de Machine Learning y el 20% restante se utilizaron como validación posterior. Lo que

significa que para la consecución de Karmona® solo se modelizó con 385 ojos, una nimiedad si se compara con el tamaño de la base de datos de Hill-RBF v3.0 en 2021, más de 30.000 ojos, por compararlo con otro algoritmo basado solo en Inteligencia Artificial. Actualmente la versión 3.0 de Hill-RBF publicada recientemente en diciembre de 2020 contiene más ojos, aunque no se dispone del dato exacto en su página web.

Una debilidad destacable de Karmona®, así como de cualquier diseño basado en Machine Learning, podría ser el escaso tamaño muestral. Teniendo en cuenta que el dataset utilizado en Karmona® es aproximadamente 75 veces mas pequeño que el utilizado en Hill-RBF v3.0, sin embargo los resultados son competitivos con las fórmulas comparadas, el planteamiento podría cambiar y hace pensar que el conocimiento profundo de la información, la selección adecuada de esta, la calidad de los datos y el uso conveniente de los algoritmos, podrían estar por encima del tamaño muestral.

En un futuro y lejos de ser un sistema estático, Karmona® podrá incrementar su potencial predictivo y su rendimiento general todavía más en sucesivas versiones con el aumento de la muestra que permitirá cada vez un mayor aprendizaje. A este respecto si se le enseña al algoritmo con ojos axialmente muy cortos, con ojos extremadamente largos o con ojos de características biométricas singulares, será información privilegiada con la que contará Karmona® para aumentar su poder predictivo.

De los 386 ojos con los que aprendió Karmona®, las potencias de las lentes intraoculares que se utilizaron oscilaban entre +36.00 D y -10.00 D, por tanto se puede aseverar que en esos rangos dióptricos el algoritmo funcionará gracias a la experiencia previa del sistema. Sin embargo, Hill-RBF en su versión 3.0 con su última actualización

(aproximadamente 30.000 ojos) es capaz de realizar cálculos dentro de los límites de +34.00 D a +6.00 D para lentes biconvexas y de +5.00 D a -5.00 D para meniscos.

El algoritmo Karmona® está basado íntegramente en IA igual que Hill-RBF, de hecho son los únicos métodos actualmente que lo hacen, ya que Fullmonte, Ladas Superformula, Kane, PEARL-DGS y Hoffer-QST son fórmulas híbridas que combinan un cuerpo teórico vergencial con alguna mejora basada en IA. Este hecho implica que Karmona® predice directamente la potencia de la lente emetropizante y no la posición efectiva de la lente como haría cualquier fórmula teórica pura o híbrida.

Habitualmente la estimación del posicionamiento de la lente intraocular se basa en regresiones lineales multivariantes que conllevan un error derivado de las medidas de las propias variables predictoras e incluso del diseño regresivo *per se*, siendo la fuente de error que más contribuye (142). A su vez este error se traslada posteriormente a la fórmula vergencial para el cálculo final de la potencia de la lente emetropizante, que independientemente de si se basa en lente fina o gruesa, tiene fundamentos gaussianos con asunciones de aproximación. Si además de todo esto, se tiene en cuenta que cada fórmula está concebida con pequeñas variaciones teóricas según sus autores, existe una propagación acumulativa de errores sobre el resultado final.

Por el contrario, en el diseño de Karmona®, desde el momento en el que retrospectivamente se ajusta la potencia de la lente implantada teniendo en cuenta el resultado refractivo postoperatorio, se elimina el sesgo producido por la estimación de la ELP así como el de la fórmula empleada para la selección de la potencia. Se minimiza de esta forma los errores y quedando expuesta únicamente al error inherente al método de ML regresivo no lineal seleccionado.

Mientras que las fórmulas teóricas deben adaptar sus fundamentos y principios al tipo de lente que se implanta, obviamente no tiene las mismas características físicas una lente biconvexa de +35.00 D por ejemplo que un menisco de -5.00 D, y por tanto la ELP se calculará de manera diferente para cada caso. Karmona® no necesita discriminar entre esto, simplemente necesita haber entrenado previamente con una muestra suficientemente representativa de este tipo de casos extremos para *entenderlos*.

Por otro lado, en esta tesis además del algoritmo Karmona® hay otra importante aportación que podría contribuir en un futuro al consenso general en la comparación y evaluación del rendimiento entre fórmulas, ya que en la actualidad hay mucha controversia. Se trata de un nuevo enfoque del tratamiento de los datos refractivos postoperatorios, acompañado de unas nuevas métricas que representan mejor el ordenamiento de las fórmulas en los estudios comparativos, identificando la mejor con más garantías que con las métricas utilizadas hasta ahora.

En el análisis de validación del algoritmo Karmona® se estudiaron 95 ojos de todos los tipos, siendo Karmona® la que mejor rendimiento ofreció con 0.23 ± 0.20 D (MedAE 0.18 D) en comparación con las otras 12 fórmulas en el total de la muestra sin encontrarse diferencias estadísticamente significativas, introduciendo además un 92% de los ojos operados en el intervalo ± 0.50 D. Por otro lado y teniendo en cuenta las métricas propuestas en esta tesis, Karmona® vuelve a ser la mejor en toda la muestra presentando un IOF_{MEAN} de 0.865 sin diferencias estadísticamente significativas con las demás fórmulas, siendo el segundo lugar ocupado por Kane con un IOF_{MEAN} de 0.848.

Hay multitud de estudios publicados sobre la comparación entre distintas fórmulas (5, 90, 143-146), pero hasta donde hemos podido saber, este es el primer

estudio en el que se consiguen MAEs por debajo del cuarto de dioptría y cotas superiores al 90% en el intervalo refractivo postoperatorio de ± 0.50 D. Sin embargo, hasta la fecha no hay estudios publicados con las últimas versiones de Hill-RBF y PEARL-DGS, la versión 3.0 y 2.0 respectivamente.

Sabiendo que no es una muestra amplia, solo 95 ojos, los resultados son muy prometedores en comparación con el resto de fórmulas de este trabajo y con los resultados de la literatura publicada. No obstante, en un futuro será necesario comparar Karmona® en otros escenarios, por otros investigadores independientes y en condiciones diferentes a las planteadas en esta tesis.

De esta misma muestra de validación ($n = 95$), en el rango de ojos largos por encima de 25.00 mm de longitud axial ($n = 25$) Karmona® obtuvo también el mejor resultado con 0.21 ± 0.19 D (MedAE 0.20 D) respecto a las demás fórmulas comparadas sin diferencias estadísticamente significativas, siendo EVO 2.0 la segunda mejor con un 0.22 ± 0.18 D (MedAE 0.19 D).

Wang et al (1), en un meta-análisis que involucraba a 4047 ojos largos, demostraron la superioridad de Barrett Universal II sobre el resto de fórmulas estudiadas, pero no estaban incluidas las de última generación: Ladas Superformula, Hill-RBF, Kane, PEARL-DGS y EVO. En la presente tesis, Karmona®, Kane, PEARL-DGS v2.0 y EVO 2.0 son mejores fórmulas que Barrett Universal II en ojos largos sin que haya diferencias estadísticamente significativas entre ellas. No se han encontrado estudios en la literatura publicada donde se comparen estas fórmulas y sus últimas versiones, del mismo modo que no se han encontrado trabajos con métricas de rendimiento tan buenas como las obtenidas por Karmona® en ojos largos.

De la misma muestra de validación ($n = 95$) pero en el rango de ojos cortos por debajo de 22.00 mm de longitud axial ($n = 12$) Karmona® obtuvo también el mejor resultado con 0.35 ± 0.30 D (MedAE 0.21 D) respecto al resto de fórmulas.

Con una representación tan escasa de ojos cortos extraída de la muestra de validación y sabiendo que el punto débil de todas las fórmulas es precisamente este tipo de ojos, se decidió analizar una muestra externa de 33 ojos axialmente por debajo de 22.00 mm diferente a los 95 ojos de validación del algoritmo Karmona®. Las fórmulas más recientes como PEARL-DGS v2.0, Hill-RBF v3.0 y EVO v2.0 son mejores que Barrett Universal II o Haigis para esta tipología de ojo sin diferencias estadísticamente significativas.

No obstante, Karmona® volvió a obtener los mejores resultados de todas las fórmulas estudiadas con las métricas clásicas: 0.38 ± 0.28 D (MedAE 0.33 D) introduciendo un 64% de los ojos en el intervalo ± 0.50 D, sin obtener diferencias estadísticamente significativas, pero también con las métricas propuestas en la tesis: $\text{IOF}_{\text{MEAN}} = 0.702$, la mejor de todas las fórmulas testadas.

Tradicionalmente, Hoffer Q ha sido la fórmula de primera elección de los clínicos para ojos cortos (54, 147, 148), pero con el advenimiento de fórmulas nuevas y estudios publicados más recientes no corroboran este hecho. Wang et al (149) publicaron un meta-análisis sobre 1161 ojos axialmente cortos reportando que Haigis es mejor fórmula que Hoffer Q en ojos por debajo de 22.00 mm.

Roberts et al (150) sobre una muestra de 21 ojos con longitudes axiales por debajo de 22.00 mm comparan Barrett Universal II, Hill-RBF v1.0, Holladay 2, SRK/T y Hoffer Q, obteniendo 0.43 ± 0.36 D, 0.37 ± 0.33 D, 0.52 ± 0.52 D, 0.44 ± 0.41 D y $0.45 \pm$

0.41 D respectivamente. Resultados comparables a los obtenidos en esta tesis donde Barrett Universal II no es la mejor fórmula en ojos con estas características anatómicas, viéndose superada por Hill-RBF v1.0 con valores similares a Karmona[®] aunque con peor desviación estándar.

Carmona et al (113) compararon 11 fórmulas sobre una submuestra de 57 ojos cortos, siendo Haigis y EVO v2.0 las mejores con un 0.39 ± 0.35 D y 0.39 ± 0.39 D respectivamente, valores ligeramente superiores a los obtenidos por Karmona[®] en la presente tesis.

Connell y Kane (90) sobre 46 ojos cortos compararon Kane, Olsen, Hill-RBF v2.0, Barrett Universal II, Haigis, Hoffer Q, Holladay 2 y SRK/T. Ninguna de las fórmulas estudiadas bajaba de 0.441 D en la categoría de MAE, valores muy superiores a los conseguidos por Karmona[®] en este estudio.

No hay un estudio publicado en el que específicamente se comparen fórmulas teóricas versus basadas en IA y que además ese sea el objetivo principal del trabajo, sin embargo sí hay trabajos en los que se comparan muchas fórmulas, permitiendo comprobar qué tipo de fórmula es mejor. En este sentido, el estudio de Savini et al (145) no coincide con nuestros resultados ya que una fórmula teórica como Barrett Universal II, sin la inclusión de ACD, es la mejor de todas las fórmulas comparadas, seguida de una basada en IA como Kane. Hill-RBF v2.0 y PEARL-DGS v1.0 que son las otras fórmulas basadas en IA están por detrás de otra teórica como EVO v2.0, probablemente debido a que en esta tesis se utilizaron las últimas actualizaciones de Hill-RBF y PEARL-DGS.

En esta investigación las fórmulas teóricas fueron superadas por las que estaban basadas en IA, incluida Karmona[®], sin diferencias estadísticamente significativas, siendo

mejores en un total de 169 ojos: 95 correspondientes a la muestra de validación, 33 a los ojos axialmente cortos y 41 a los ojos biométricamente atípicos.

9 CONCLUSIONES

9. CONCLUSIONES

Las conclusiones obtenidas en esta investigación son las que se enumeran a continuación:

1. Se ha conseguido predecir la potencia refractiva de la lente intraocular que se implantará en una cirugía de cristalino mediante la creación de una nueva fórmula basada en inteligencia artificial, así como el residuo refractivo en plano corneal asociado a dicho implante.
2. Karmona[®], el nuevo método de cálculo de la potencia de lentes intraoculares para la cirugía del cristalino propuesto en esta investigación, es competitivo respecto a cualquiera de las fórmulas utilizadas actualmente en la clínica, tanto las que están fundamentadas en teoría óptica, como las más modernas basadas en IA. Karmona[®] obtuvo la validación sin sobreajuste sobre el 20% de la muestra, con unos rendimientos predictivos comparables a los de cualquier metodología actual de cálculo.
3. Karmona[®] obtiene un error absoluto medio para la población en general por debajo del 0.25 D con una dispersión muy contenida en comparación con el resto de fórmulas y métodos actuales. Karmona[®] introduce en el rango refractivo postoperatorio de ± 0.50 D más del 90% de los casos para la población en general. Tiene un comportamiento firme y estable independientemente de la tipología axial de ojo y del resto de características anatómicas biométricas.

4. Las últimas versiones publicadas de las fórmulas basadas en IA, incluida Karmona[®], son mejores en términos predictivos que las teóricas puras basadas en óptica geométrica.

5. Se diseñó, desarrolló y programó una página web de acceso libre conteniendo el algoritmo final obtenido en la tesis.

10 BIBLIOGRAFÍA

10. BIBLIOGRAFÍA

1. Wang Q, Jiang W, Lin T, Zhu Y, Chen C, Lin H, et al. Accuracy of intraocular lens power calculation formulas in long eyes: a systematic review and meta-analysis. *Clin Exp Ophthalmol*. 2018;46(7):738-49.
2. Faramarzi A AA, Ghiasian L. Accuracy of Various Intraocular Lens Power Calculation Formulas in Steep Corneas. *J Ophthalmic Vis Res*. 2017.
3. Jeong J SH, Lee JK, Chuck RS, Kwon JW. The effect of ocular biometric factors on the accuracy of various IOL power calculation formulas. *BMC Ophthalmol*. 2017.
4. Abulafia A, Barrett GD, Koch DD, Wang L, Assia EI. Protocols for Studies of Intraocular Lens Formula Accuracy. *Am J Ophthalmol*. 2016;164:149-50.
5. Cooke DL, Cooke TL. Comparison of 9 intraocular lens power calculation formulas. *J Cataract Refract Surg*. 2016;42(8):1157-64.
6. Gokce SE, Zeiter JH, Weikert MP, Koch DD, Hill W, Wang L. Intraocular lens power calculations in short eyes using 7 formulas. *J Cataract Refract Surg*. 2017;43(7):892-7.
7. Reitblat O, Levy A, Kleinmann G, Lerman TT, Assia EI. Intraocular lens power calculation for eyes with high and low average keratometry readings: Comparison between various formulas. *J Cataract Refract Surg*. 2017;43(9):1149-56.
8. Clarke GP, Burmeister J. Comparison of intraocular lens computations using a neural network versus the Holladay formula. *J Cataract Refract Surg*. 1997;23(10):1585-9.
9. Maloney WF, Kratz RP, Mazzocco TR, Davidson B. Posterior chamber intraocular lens power calculation in 441 cases. *J Am Intraocul Implant Soc*. 1979;5(4):349.
10. Raufi N, James C, Kuo A, Vann R. Intraoperative aberrometry vs modern preoperative formulas in predicting intraocular lens power. *J Cataract Refract Surg*. 2020.

11. Apple DJ, Sims J. Harold Ridley and the invention of the intraocular lens. *Surv Ophthalmol.* 1996;40(4):279-92.
12. Sanders DR, Kraff MC. Improvement of intraocular lens power calculation using empirical data. *Journal-American Intra-Ocular Implant Society.* 1980;6(3):263-7.
13. Mundt GH, Jr., Hughes WF, Jr. Ultrasonics in ocular diagnosis. *Am J Ophthalmol.* 1956;41(3):488-98.
14. Kanki K, Yoshimoto M, Uesugi T, Kimura T. [Measurement of the axial length of the eye by the application of ultra-sonic wave]. *Nippon Ganka Gakkai Zasshi.* 1961;65:1877-83.
15. Otsuka J, Tokoro T, Araki M. [Comparative studies of phacometry with ultrasonic method on refractive elements of the human eyes]. *Nippon Ganka Gakkai Zasshi.* 1961;65:1777-92.
16. Araki M. [Studies on reflective elements of the human eye by ultrasonic waves. I. Accuracy of the measurement of ocular axial length by ultrasonic echography]. *Rinsho Ganka.* 1961;15:111-9.
17. Leary GA, Sorsby A, Richards MJ, Chaston J. Ultrasonographic measurement of the components of ocular refraction in life: I. Technical considerations. *Vision Research.* 1963;3(11):487-IN5.
18. Koch DD, Hill W, Abulafia A, Wang L. Pursuing perfection in intraocular lens calculations: I. Logical approach for classifying IOL calculation formulas. *J Cataract Refract Surg.* 2017;43(6):717-8.
19. Mendoza G. El sistema de vergencias ópticas como derivación de la óptica geométrica. *Ciencia & Tecnología para la Salud Visual y Ocular.* 2007:87.
20. Hoffer KJ. Preoperative cataract evaluation: intraocular lens power calculation. *Int Ophthalmol Clin.* 1982;22(2):37-75.

21. Binkhorst CD. Power of the prepupillary pseudophakos. *Br J Ophthalmol.* 1972;56(4):332-7.
22. Fedorov SN KA, Kolinko AI. Estimation of optical power of the intraocular lens. *Vestn Oftalmol* 1967;80:27–31.
23. A. Gernet HO, and H. Werner, . Ein neues Haftschalen-Nomogramm far Aphakie, Die praoperative Berechnung intraocularer Binkhorst-Linsen. *Sitzungsberichte der.* 1970.
24. LEARY GA. Ultrasonographic Assessment of the Implant Lens Required to Produce Emmetropia After Implantation. *Investigative Ophthalmology & Visual Science.* 1971;10(10):745-9.
25. Colenbrander MC. Calculation of the power of an iris clip lens for distant vision. *Br J Ophthalmol.* 1973;57(10):735-40.
26. Heijde GL. A nomogram for calculating the power of the prepupillary lens in the aphakic eye. *Bibl Ophthalmol.* 1975(83):273-5.
27. Oguchi Y, Van Balen AT. Determination of the expected power of the implant lens by ultrasound. *Ophthalmologica.* 1975;171(4-5):281-3.
28. Thijssen JM. The emmetropic and the iseikonic implant lens: computer calculation of the refractive power and its accuracy. *Ophthalmologica.* 1975;171(6):467-86.
29. Binkhorst RD. Pitfalls in the determination of intraocular lens power without ultrasound. *Ophthalmic Surg.* 1976;7(3):69-82.
30. Epstein LI, Geeraets WJ. On the determination of the power of a lens implant. *American Intra-Ocular Implant Society Journal.* 1977;3(2):87-93.
31. Kollarits FJ, Kollarits CR, Torchia RT. A Fortran IV program for intraocular lens power calculation. *American Intra-Ocular Implant Society Journal.* 1979;5(4):330-6.

32. Drews RC. Programs for the HP-25/C calculators for lens implant power. American Intra-Ocular Implant Society Journal. 1977;3(1):48-9.
33. Drews RC. A practical approach to lens implant power. American Intra-Ocular Implant Society Journal. 1977;3(3):170-6.
34. Drews RC. Calculation of intraocular powera program for the hewlitt packard 97 calculator. American Intra-Ocular Implant Society Journal. 1977;3(3):209-12.
35. Apple DJ, Mamalis N, Loftfield K, Googe JM, Novak LC, Kavka-Van Norman D, et al. Complications of intraocular lenses. A historical and histopathological review. Surv Ophthalmol. 1984;29(1):1-54.
36. Margo C. Intraocular Lenses: Evolution, Designs, Complications, and Pathology. Archives of Ophthalmology. 1990;108:650.
37. Binkhorst R. Intraocular lens power calculation manual: A guide to the authors TI 58/59 IOL power module. New York 1981. Binkhorst RD.
38. Hoffer K. The Effect of Axial Length on Posterior Chamber Lens and Posterior Capsule Position. Current Concepts in Ophthalmic Surgery. 1984;1:20-2.
39. Retzlaff J. A new intraocular lens calculation formula. American Intra-Ocular Implant Society Journal. 1980;6(2):148-52.
40. Gills JP. Minimizing postoperative refractive error. Eye & Contact Lens. 1980;6(1):56-9.
41. Hillis A. Use of regression formulas for IOL power calculation. American Intra-Ocular Implant Society Journal. 1981;7(1):62.
42. Sanders D, Retzlaff J, Kraff M, Kratz R, Gills J, Levine R, et al. Comparison of the accuracy of the Binkhorst, Colenbrander, and SRK™ implant power prediction formulas. American Intra-Ocular Implant Society Journal. 1981;7(4):337-40.
43. Thompson JT, Maumenee AE, Baker C. A new posterior chamber intraocular lens formula for axial myopes. Ophthalmology. 1984;91(5):484-8.

44. Olsen T. Prediction of intraocular lens position after cataract extraction. *Journal of Cataract & Refractive Surgery*. 1986;12(4):376-9.
45. Olsen T. Theoretical approach to intraocular lens calculation using Gaussian optics. *Journal of Cataract & Refractive Surgery*. 1987;13(2):141-5.
46. Barrett GD. Intraocular lens calculation formulas for new intraocular lens implants. *Journal of Cataract & Refractive Surgery*. 1987;13(4):389-96.
47. Holladay JT, Musgrove KH, Prager TC, Lewis JW, Chandler TY, Ruiz RS. A three-part system for refining intraocular lens power calculations. *Journal of Cataract & Refractive Surgery*. 1988;14(1):17-24.
48. Retzlaff J, Sanders D, Kraff M. *Lens Implant Power Calculation: A Manual for Ophthalmologists & Biometrists*. Thorofare, NJ: Slack. Inc; 1990.
49. Retzlaff JA, Sanders DR, Kraff MC. Development of the SRK/T intraocular lens implant power calculation formula. *Journal of Cataract & Refractive Surgery*. 1990;16(3):333-40.
50. Olsen T, Olesen H, Thim K, Corydon L. Prediction of postoperative intraocular lens chamber depth. *Journal of Cataract & Refractive Surgery*. 1990;16(5):587-90.
51. Hoffer KJ. The Hoffer Q formula: a comparison of theoretic and regression formulas. *Journal of Cataract & Refractive Surgery*. 1993;19(6):700-12.
52. Barrett GD. An improved universal theoretical formula for intraocular lens power prediction. *Journal of Cataract & Refractive Surgery*. 1993;19(6):713-20.
53. Olsen T, Corydon L, Gimbel H. Intraocular lens power calculation with an improved anterior chamber depth prediction algorithm. *Journal of Cataract & Refractive Surgery*. 1995;21(3):313-9.
54. Hoffer KJ. Clinical results using the Holladay 2 intraocular lens power formula. *Journal of Cataract & Refractive Surgery*. 2000;26(8):1233-7.

55. Haigis W, editor Einfluß der Optikform auf die individuelle Anpassung von Linsenkonstanten zur IOL-Berechnung. 9 Kongreß der Deutschsprachigen Gesellschaft für Intraokularlinsen Implantation; 1995: Springer.
56. Haigis W, editor Biometrie bei komplizierten Ausgangssituationen. 9 Kongreß der Deutschsprachigen Gesellschaft für Intraokularlinsen Implantation; 1995: Springer.
57. Ishikawa T, Hirano A, Murai K, Kumagai M, Nakayasu K, Kanai A. Intraocular lens calculation for cataract treated with photorefractive keratectomy using ray tracing method. Japanese journal of ophthalmology. 2000;44(5):575.
58. Ishikawa T, Hirano A, Inoue J, Nakayasu K, Kanai A, Takeuchi K, et al. Trial for new intraocular lens power calculation following phototherapeutic keratectomy. Japanese journal of ophthalmology. 2000;44(4):400-6.
59. Preussner PR, Wahl J, Lahdo H, Dick B, Findl O. Ray tracing for intraocular lens calculation. J Cataract Refract Surg. 2002;28(8):1412-9.
60. Norrby S. Using the lens haptic plane concept and thick-lens ray tracing to calculate intraocular lens power. Journal of Cataract & Refractive Surgery. 2004;30(5):1000-5.
61. Ghoreyshi M, Khalilian A, Peyman M, Mohammadinia M, Peyman A. Comparison of OKULIX ray-tracing software with SRK-T and Hoffer-Q formula in intraocular lens power calculation. Journal of current ophthalmology. 2018;30(1):63-7.
62. Nabil KM. Accuracy of minus power intraocular lens calculation using OKULIX ray tracing software. International ophthalmology. 2019;39(8):1803-8.
63. McCarthy J. Measures of the value of information. Proceedings of the National Academy of Sciences of the United States of America. 1956;42(9):654.
64. Lu W, Tong Y, Yu Y, Xing Y, Chen C, Shen Y. Applications of artificial intelligence in ophthalmology: general overview. Journal of ophthalmology. 2018;2018.

65. Fernández-Álvarez JC, Hernández-López I, Cruz-Cobas PP, Cárdenas-Díaz T, Batista-Leyva AJ. Using a multilayer perceptron in intraocular lens power calculation. *Journal of Cataract & Refractive Surgery*. 2019;45(12):1753-61.
66. Findl O, Struhal W, Dorffner G, Drexler W. Analysis of nonlinear systems to estimate intraocular lens position after cataract surgery. *Journal of Cataract & Refractive Surgery*. 2004;30(4):863-6.
67. Fukuda T, Kamiura N, Saitoh A, Isokawa T, Matsui N, Tabuchi H, editors. Formula selection for intraocular power calculation using support vector machines and self-organizing maps. 2011 IEEE International Conference on Systems, Man, and Cybernetics; 2011: IEEE.
68. Kamiura N, Takehara N, Saitoh A, Isokawa T, Matsui N, Tabuchi H, editors. On selection of intraocular power formula based on data classification using self-organizing maps. 2010 IEEE International Conference on Systems, Man and Cybernetics; 2010: IEEE.
69. Sramka M, Slovak M, Tuckova J, Stodulka P. Improving clinical refractive results of cataract surgery by machine learning. *PeerJ*. 2019;7:e7202.
70. Yarmahmoodi M, Arabalibeik H, Mokhtaran M, Shojaei A. Intraocular lens power formula selection using support vector machines. *Frontiers in Biomedical Technologies*. 2015;2(1):36-44.
71. Chambers JM. *Computational methods for data analysis*. 1977.
72. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *Journal of computational and graphical statistics*. 1996;5(3):299-314.
73. Dietterich T. Overfitting and undercomputing in machine learning. *ACM computing surveys (CSUR)*. 1995;27(3):326-7.
74. Larsen J, Goutte C, editors. On optimal data split for generalization estimation and model selection. *Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (Cat No 98TH8468)*; 1999: IEEE.

75. Reitermanova Z, editor Data splitting. WDS; 2010.
76. Korjus K, Hebart MN, Vicente R. An efficient data partitioning to improve classification performance while keeping parameters interpretable. PloS one. 2016;11(8).
77. Aramberri J. Cálculo de la lente intraocular. Lorente R, Mendicute J Cirugía del Cristalino LXXXIV Ponencia Oficial de la Sociedad Española de Oftalmología. 2008:234-75.
78. Kapelner A, Bleich J. Prediction with missing data via Bayesian Additive Regression Trees. Canadian Journal of Statistics. 2015;43(2):224-39.
79. Clarke GP, Kapelner A. Data-driven intraocular lens power computation: a comparison of outcomes from two new methods, and a third generation intraocular lens formula. Unpublished work https://www.academia.edu/32075532/Data-driven_intraocular_lens_power_computation_a_comparison_of_outcomes_from_two_new_methods_and_a_third_generation_intraocular_lens_formula_OptiVision_EyeCare_Corresponding_Author. 2016.
80. Wang L, Koch DD. Modified axial length adjustment formulas in long eyes. 2018.
81. Clarke GP, Kapelner A. The Bayesian Additive Regression Trees Formula for Safe Machine Learning-Based Intraocular Lens Predictions. Frontiers in big data. 2020;3:572134-.
82. Hill W. Hill RBF version 2.0 update pattern recognition IOL power selection. ASCRS https://www.youtube.com/watch?v=wDz_FuLMPW4. 2018.
83. Kohonen T. Self-organized formation of topologically correct feature maps. Biological cybernetics. 1982;43(1):59-69.
84. Haykin SS. Neural networks and learning machines/Simon Haykin. New York: Prentice Hall; 2009.
85. Bors A. Introduction of the Radial Basis Function (RBF) Networks. 12001. p. 1-7.
86. Ladas JG, Siddiqui AA. Using AI in IOL calculations. Ophthalmology Managements. 2018;22(April):36-8, 40, 62.

87. Ladas JG, Siddiqui AA, Devgan U, Jun AS. A 3-D “super surface” combining modern intraocular lens formulas to generate a “super formula” and maximize accuracy. *JAMA ophthalmology*. 2015;133(12):1431-6.
88. Wang L, Koch D. Modified axial length adjustment formulas in long eyes. *Journal of cataract and refractive surgery*. 2018;44(11):1396.
89. Cheng H, Liu L, Sun A, Wu M. Accuracy of Modified Axial Length Adjustment for Intraocular Lens Power Calculation in Chinese Axial Myopic Eyes. *Curr Eye Res*. 2019:1-7.
90. Connell BJ, Kane JX. Comparison of the Kane formula with existing formulas for intraocular lens power selection. *BMJ Open Ophthalmol*. 2019;4(1):e000251.
91. Savini G, Hoffer KJ, Balducci N, Barboni P, Schiano-Lomoriello D. Comparison of formula accuracy for intraocular lens power calculation based on measurements by a swept-source optical coherence tomography optical biometer. *Journal of Cataract & Refractive Surgery*. 2020;46(1):27-33.
92. Olivia Li J-P, Mehta JS, Shu-Wei Ting D, Agarwal A, Ambrósio RJ, Vinciguerra R, et al. Emerging AI and Big Data. *Cataract and Refractive Surgery Today*. 2021;March.
93. Kuhn M. Building predictive models in R using the caret package. *Journal of statistical software*. 2008;28(5):1-26.
94. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2017. R package version. 2018;1(8).
95. Milborrow S, Hastie T, Tibshirani R. Earth: multivariate adaptive regression spline models. R package version. 2014;3:2-7.
96. Quinlan JR, editor *Combining Instance-Based and Model-Based Learning*. ICML; 1993.

97. Pérez-Rodríguez P, Gianola D, Weigel K, Rosa G, Crossa J. Technical Note: An R package for fitting Bayesian regularized neural networks with applications in animal breeding. *Journal of animal science*. 2013.
98. Analytics R, Weston S. doParallel: Foreach parallel adaptor for the parallel package. *R package version*. 2014;1(8).
99. Faul F, Erdfelder E, Lang A, Buchner A. G* Power 3: un programa flexible de análisis de poder estadístico para las ciencias sociales, conductuales y biomédicas. *Métodos de investigación del comportamiento*. 2007;39(2):175-91.
100. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. *PloS one*. 2019;14(11).
101. Maxwell SE, Kelley K, Rausch JR. Sample size planning for statistical power and accuracy in parameter estimation. *Annu Rev Psychol*. 2008;59:537-63.
102. Cohen J. *Statistical power analysis for the behavioral sciences*: Academic press; 2013.
103. Ellis PD. *The Essential Guide to Effect Sizes: An Introduction to Statistical Power, Meta-Analysis and the Interpretation of Research Results*. 2010. United Kingdom: Cambridge University Press.
104. Darcy K, Gunn D, Tavassoli S, Sparrow J, Kane JX. Assessment of the accuracy of new and updated intraocular lens power calculation formulas in 10 930 eyes from the UK National Health Service. *Journal of Cataract & Refractive Surgery*. 2020;46(1):2-7.
105. Melles RB, Holladay JT, Chang WJ. Accuracy of intraocular lens calculation formulas. *Ophthalmology*. 2018;125(2):169-78.
106. Liu B, Wei Y, Zhang Y, Yang Q, editors. *Deep Neural Networks for High Dimension, Low Sample Size Data*. *IJCAI*; 2017.
107. Fuchs M, Jiang X, Anne-Laure B. The computationally optimal test set size in simulation studies on supervised learning. 2016.

108. van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC Medical Research Methodology*. 2014;14(1):137.
109. Al-Anazi A, Gates I. Support vector regression to predict porosity and permeability: effect of sample size. *Computers & geosciences*. 2012;39:64-76.
110. Olden JD, Jackson DA. Torturing data for the sake of generality: how valid are our regression models? *Ecoscience*. 2000;7(4):501-10.
111. Cui Z, Gong G. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *Neuroimage*. 2018;178:622-37.
112. Steyerberg E, Harrell Jr F. Chapter 8: Statistical models for prognostication. *Symptom Research: Methods and Opportunities* Bethesda, MD: National Institutes of Health Interactive textbook retrieved from http://painconsortium.nih.gov/symptomresearch/chapter_8/sec7/cess7pg6.htm. 2012.
113. Carmona González D, Palomino Bautista C. Accuracy of a new intraocular lens power calculation method based on artificial intelligence. *Eye*. 2020.
114. Crammer K, Kearns M, Wortman J. Learning from multiple sources. *Journal of Machine Learning Research*. 2008;9(Aug):1757-74.
115. Cesa-Bianchi N, Hardoon DR, Leen G. Guest Editorial: Learning from multiple sources. *Machine Learning*. 2010;79(1-2):1.
116. Wu P, Dietterich TG, editors. Improving SVM accuracy by training on auxiliary data sources. *Proceedings of the twenty-first international conference on Machine learning*; 2004.
117. Simpson MJ, Charman WN. The effect of testing distance on intraocular lens power calculation. *Journal of Refractive Surgery*. 2014;30(11):726-.

118. Dong EY, Joo C-K. Predictability for proper capsular tension ring size and intraocular lens size. *Korean Journal of Ophthalmology*. 2001;15(1):22-6.
119. Khng C, Osher RH. Evaluation of the relationship between corneal diameter and lens diameter. *Journal of Cataract & Refractive Surgery*. 2008;34(3):475-9.
120. Kim K-H, Shin H-H, Kim H-M, Song J-S. Correlation between ciliary sulcus diameter measured by 35 MHz ultrasound biomicroscopy and other ocular measurements. *Journal of Cataract & Refractive Surgery*. 2008;34(4):632-7.
121. Dunne MC, Royston JM, Barnes DA. Normal variations of the posterior corneal surface. *Acta ophthalmologica*. 1992;70(2):255-61.
122. Royston J, Dunne M, Barnes D. Measurement of the posterior corneal radius using slit lamp and Purkinje image techniques. *Ophthalmic and Physiological Optics*. 1990;10(4):385-8.
123. Tonn B, Klaproth OK, Kohnen T. Anterior surface-based keratometry compared with Scheimpflug tomography-based total corneal astigmatism. *Investigative ophthalmology & visual science*. 2015;56(1):291-8.
124. Næser K, Savini G, Bregnhøj JF. Corneal powers measured with a rotating Scheimpflug camera. *British Journal of Ophthalmology*. 2016;100(9):1196-200.
125. Fam H-B, Lim K-L. Validity of the keratometric index: large population-based study. *Journal of Cataract & Refractive Surgery*. 2007;33(4):686-91.
126. Ho J-D, Tsai C-Y, Tsai RJ-F, Kuo L-L, Tsai I-L, Liou S-W. Validity of the keratometric index: evaluation by the Pentacam rotating Scheimpflug camera. *Journal of Cataract & Refractive Surgery*. 2008;34(1):137-45.
127. Barboni P, Savini G, Carbonelli M, Hoffer KJ. Agreement between Pentacam and videokeratography in corneal power assessment. *Journal of Refractive Surgery*. 2009;25(6):534-8.

128. Kotsiantis S, Kanellopoulos D, Pintelas P. Data preprocessing for supervised learning. *International Journal of Computer Science*. 2006;1(2):111-7.
129. ENGLISH D. Ophthalmic implants-Intraocular lenses-Part 2: Optical properties and test methods (ISO 11979-2: 1999); English version of DIN EN ISO 11979-2. 2000.
130. Holladay JT. Refractive power calculations for intraocular lenses in the phakic eye. *Am J Ophthalmol*. 1993;116(1):63-6.
131. Breiman L. Random forests. *Machine learning*. 2001;45(1):5-32.
132. Nelder JA, Wedderburn RW. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*. 1972;135(3):370-84.
133. Burden F, Winkler D. Bayesian regularization of neural networks. *Artificial neural networks: Springer*; 2008. p. 23-42.
134. Nguyen D, Widrow B, editors. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. 1990 IJCNN International Joint Conference on Neural Networks; 1990: IEEE.
135. Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. *The Annals of Applied Statistics*. 2010;4(1):266-98.
136. Breiman L, Friedman J, Stone CJ, Olshen RA. *Classification and regression trees*: CRC press; 1984.
137. Breiman L. Bagging predictors. *Machine learning*. 1996;24(2):123-40.
138. Quinlan JR, editor *Learning with continuous classes*. 5th Australian joint conference on artificial intelligence; 1992: World Scientific.
139. Breiman L. Stacked regressions. *Machine Learning*. 1996;24(1):49-64.
140. Laan MJvd, Polley EC, Hubbard AE. Super Learner. *Statistical Applications in Genetics and Molecular Biology*. 2007;6(1).

141. Vega Y, Gershoni A, Achiron A, Tuuminen R, Weinberger Y, Livny E, et al. High Agreement between Barrett Universal II Calculations with and without Utilization of Optional Biometry Parameters. *J Clin Med*. 2021;10(3).
142. Norrby S. Sources of error in intraocular lens power calculation. *J Cataract Refract Surg*. 2008;34(3):368-76.
143. Hipólito-Fernandes D, Elisa Luís M, Gil P, Maduro V, Feijão J, Yeo TK, et al. VRF-G, a New Intraocular Lens Power Calculation Formula: A 13-Formulas Comparison Study. *Clin Ophthalmol*. 2020;14:4395-402.
144. Rocha-de-Lossada C, Colmenero-Reina E, Flikier D, Castro-Alonso FJ, Rodriguez-Raton A, García-Madróna JL, et al. Intraocular lens power calculation formula accuracy: Comparison of 12 formulas for a trifocal hydrophilic intraocular lens. *Eur J Ophthalmol*. 2020:1120672120980690.
145. Savini G, Di Maita M, Hoffer KJ, Næser K, Schiano-Lomoriello D, Vagge A, et al. Comparison of 13 formulas for IOL power calculation with measurements from partial coherence interferometry. *Br J Ophthalmol*. 2021;105(4):484-9.
146. Carmona-González D, Castillo-Gómez A, Palomino-Bautista C, Romero-Domínguez M, Gutiérrez-Moreno M. Comparison of the accuracy of 11 intraocular lens power calculation formulas. *Eur J Ophthalmol*. 2020:1120672120962030.
147. Day AC, Foster PJ, Stevens JD. Accuracy of intraocular lens power calculations in eyes with axial length <22.00 mm. *Clin Exp Ophthalmol*. 2012;40(9):855-62.
148. Narváez J, Zimmerman G, Stulting RD, Chang DH. Accuracy of intraocular lens power prediction using the Hoffer Q, Holladay 1, Holladay 2, and SRK/T formulas. *J Cataract Refract Surg*. 2006;32(12):2050-3.
149. Wang Q, Jiang W, Lin T, Wu X, Lin H, Chen W. Meta-analysis of accuracy of intraocular lens power calculation formulas in short eyes. *Clin Exp Ophthalmol*. 2018;46(4):356-63.

150. Roberts TV, Hodge C, Sutton G, Lawless M. Comparison of Hill-radial basis function, Barrett Universal and current third generation formulas for the calculation of intraocular lens power during cataract surgery. *Clin Exp Ophthalmol*. 2018;46(3):240-6.
151. Livni R, Shalev-Shwartz S, Shamir O, editors. On the computational efficiency of training neural networks. *Advances in neural information processing systems*; 2014.
152. Ramachandran P, Zoph B, Le QV. Searching for activation functions. *arXiv preprint arXiv:171005941*. 2017.
153. Lang B, editor *Monotonic multi-layer perceptron networks as universal approximators*. International conference on artificial neural networks; 2005: Springer.
154. Minin A, Velikova M, Lang B, Daniels H. Comparison of universal approximators incorporating partial monotonicity by structure. *Neural Networks*. 2010;23(4):471-5.
155. Rosenblatt F. Perceptron simulation experiments. *Proceedings of the IRE*. 1960;48(3):301-9.
156. Vapnik V. *Estimation of dependences based on empirical data*: Springer Science & Business Media; 2006.
157. Friedman JH. Multivariate adaptive regression splines. *The annals of statistics*. 1991:1-67.
158. Boehmke B, Greenwell BM. *Hands-On Machine Learning with R*: CRC Press; 2019.
159. Lever J, Krzywinski M, Altman N. Model selection and overfitting. *Nature Methods*. 2016;13(9):703-4.
160. Zhang Y, Liang XY, Liu S, Lee JW, Bhaskar S, Lam DS. Accuracy of Intraocular Lens Power Calculation Formulas for Highly Myopic Eyes. *J Ophthalmol*. 2016;2016:1917268.
161. Terzi E, Wang L, Kohnen T. Accuracy of modern intraocular lens power calculation formulas in refractive lens exchange for high myopia and high hyperopia. *J Cataract Refract Surg*. 2009;35(7):1181-9.

162. Roh YR, Lee SM, Han YK, Kim MK, Wee WR, Lee JH. Intraocular lens power calculation using IOLMaster and various formulas in short eyes. *Korean J Ophthalmol.* 2011;25(3):151-5.
163. Rong X, He W, Zhu Q, Qian D, Lu Y, Zhu X. Intraocular lens power calculation in eyes with extreme myopia: Comparison of Barrett Universal II, Haigis, and Olsen formulas. *J Cataract Refract Surg.* 2019;45(6):732-7.
164. Savini G, Hoffer KJ, Shammas HJ, Aramberri J, Huang J, Barboni P. Accuracy of a New Swept-Source Optical Coherence Tomography Biometer for IOL Power Calculation and Comparison to IOLMaster. *J Refract Surg.* 2017;33(10):690-5.
165. Abulafia A, Barrett GD, Rotenberg M, Kleinmann G, Levy A, Reitblat O, et al. Intraocular lens power calculation for eyes with an axial length greater than 26.0 mm: comparison of formulas and methods. *J Cataract Refract Surg.* 2015;41(3):548-56.
166. Voytsekhivskyy OV. Development and Clinical Accuracy of a New Intraocular Lens Power Formula (VRF) Compared to Other Formulas. *Am J Ophthalmol.* 2018;185:56-67.
167. Bang S, Edell E, Yu Q, Pratzner K, Stark W. Accuracy of intraocular lens calculations using the IOLMaster in eyes with long axial length and a comparison of various formulas. *Ophthalmology.* 2011;118(3):503-6.
168. Shrivastava AK, Behera P, Kumar B, Nanda S. Precision of intraocular lens power prediction in eyes shorter than 22 mm: An analysis of 6 formulas. *J Cataract Refract Surg.* 2018;44(11):1317-20.
169. Hoffer KJ, Savini G. Update on Intraocular Lens Power Calculation Study Protocols: The Better Way to Design and Report Clinical Trials. *Ophthalmology.* 2020.
170. Gavin EA, Hammond CJ. Intraocular lens power calculation in short eyes. *Eye (Lond).* 2008;22(7):935-8.
171. Hoffer KJ, Aramberri J, Haigis W, Olsen T, Savini G, Shammas HJ, et al. Protocols for studies of intraocular lens formula accuracy. *Am J Ophthalmol.* 2015;160(3):403-5.e1.

172. Wang L, Koch DD, Hill W, Abulafia A. Pursuing perfection in intraocular lens calculations: III. Criteria for analyzing outcomes. *J Cataract Refract Surg.* 2017;43(8):999-1002.
173. Kane JX, Van Heerden A, Atik A, Petsoglou C. Intraocular lens power formula accuracy: Comparison of 7 formulas. *J Cataract Refract Surg.* 2016;42(10):1490-500.
174. Wilcox RR, Keselman HJ. Modern robust data analysis methods: measures of central tendency. *Psychol Methods.* 2003;8(3):254-74.
175. Aristodemou P, Knox Cartwright NE, Sparrow JM, Johnston RL. Statistical Analysis for Studies of Intraocular Lens Formula Accuracy. *Am J Ophthalmol.* 2015;160(5):1085-6.
176. Huber P. *Robust Statistics.* New York: John Wiley and Sons. HuberRobust statistics1981. 1981.
177. Basulto J. Transformacion raiz cuadrada versus logaritmica de una distribucion Gamma. *Trabajos de Estadistica Y de Investigacion Operativa.* 1978;29(3):83-8.
178. Wilcox RR. *Introduction to robust estimation and hypothesis testing:* Academic press; 2011.
179. Pitman EJG. Tests of Hypotheses Concerning Location and Scale Parameters. *Biometrika.* 1939;31(1/2):200-15.
180. Hampel FR. The Influence Curve and Its Role in Robust Estimation. *Journal of the American Statistical Association.* 1974;69(346):383-93.
181. Rousseeuw PJ, Croux C. Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association.* 1993;88(424):1273-83.
182. Pedersen AR, editor *How to Obtain Those Nasty Standard Errors from Transformed Data | and Why They Should Not Be Used*2007.

183. Kim M, Eom Y, Song JS, Kim HM. Comparative evaluation of refractive outcomes after implantation of two types of intraocular lenses with different diopter intervals (0.25 diopter versus 0.50 diopter). *BMC Ophthalmol.* 2018;18(1):176.
184. Hoffer KJ. Biometry of 7,500 cataractous eyes. *Am J Ophthalmol.* 1980;90(3):360-8.
185. Fotedar R, Wang JJ, Burlutsky G, Morgan IG, Rose K, Wong TY, et al. Distribution of axial length and ocular biometry measured using partial coherence laser interferometry (IOL Master) in an older white population. *Ophthalmology.* 2010;117(3):417-23.
186. Fernández J, Rodríguez-Vallejo M, Poyales F, Burguera N, Garzón N. New method to assess the accuracy of intraocular lens power calculation formulas according to ocular biometric parameters. *J Cataract Refract Surg.* 2020;46(6):849-56.
187. Shamma HJ, Jabre JF. Validating e-norms methodology in ophthalmic biometry. *BMJ Open Ophthalmol.* 2020;5(1):e000500.
188. Jabre JF. Do you define the limits of normalcy from looking at the patient or the healthy subject? - An e-norms reply. *Clin Neurophysiol.* 2018;129(8):1517-8.

11 ANEXOS

11.1 ANEXO 1.- BASES TEÓRICAS DE LOS ALGORITMOS DE MACHINE LEARNING EMPLEADOS EN LA TESIS

11.1.1 Funcionamiento de redes neuronales

Se le llama RN por su arquitectura y funcionamiento similar e inspirado en el sistema nervioso humano (Figura 87). Unos *inputs* ($x_1, x_2, x_3 \dots x_i$) que conforman la primera capa son modulados o ponderados por unos pesos o *weights* ($w_1, w_2, w_3 \dots w_i$) que inicialmente se fijan aleatoriamente y que se irán autoajustando con el proceso iterativo de aprendizaje, gracias al error cometido por la salida respecto al dato real, es lo que se conoce como *back propagation*.

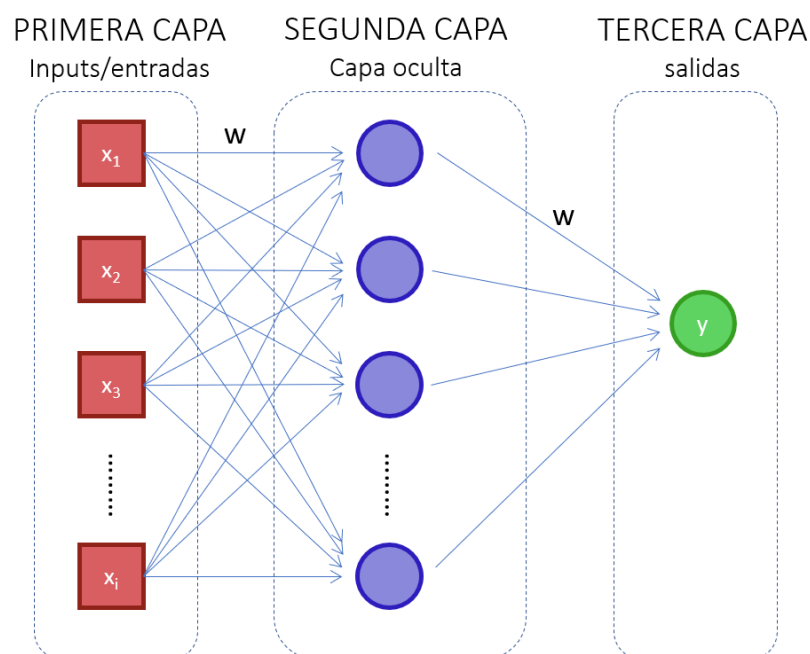


Figura 87. Esquema simple de una red neuronal con tres capas.

Desde un punto de vista matemático sencillo y teniendo en cuenta lo anterior, a la primera neurona de la segunda capa o *hidden layer* (Figura 87) le llega la siguiente señal:

$$w_1x_1 + w_2x_2 + w_3x_3 + \dots + x_iw_i + U = 0 \quad (56)$$

Donde x_i es el predictor, w_i es el coeficiente que lo acompaña (peso) y U es el término independiente. Por tanto esta ecuación se corresponde a la de un plano, o mejor hiperplano ya que no le afecta la dimensionalidad, formado por el producto de dos vectores (x, w) y el sumatorio de todos ellos:

$$f(x) = U + \sum_i w_i x_i \quad (57)$$

Una vez que a la neurona le llega la señal, existe una función de activación que permitirá, al igual que hace una neurona normal, dejar pasar la información al siguiente nivel o no, es decir, inhibirla o excitarla. En este sentido, la función sigmoide (σ) ha sido históricamente la más utilizada y la primera, aunque hay muchas otras (ReLU, tangente hiperbólica, softmax, etc). La función sigmoide tiene esta expresión:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (58)$$

Donde e es la constante 2.71828. Si se establece que z equivale a la suma ponderada de entrada Ecuación (57) y después se pasa a través de la función sigmoide o activadora, se tiene:

$$z = U + \sum_i w_i x_i \quad (59)$$

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (60)$$

Y la gráfica resultante tiene esta forma:

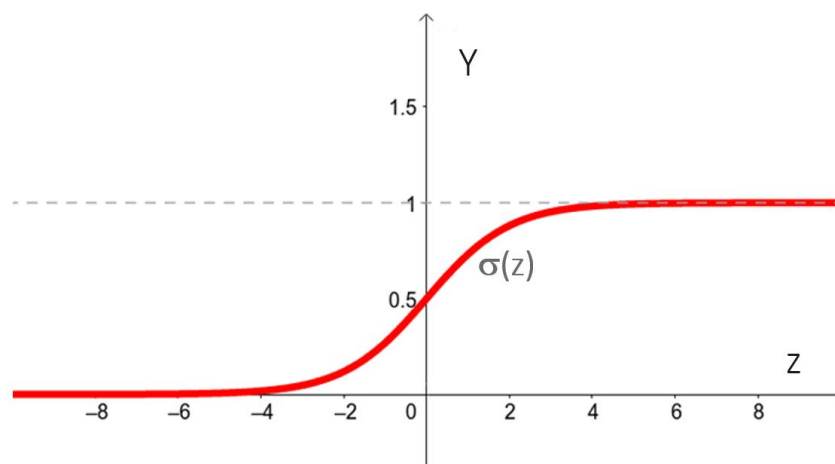


Figura 88. Función sigmoide de activación en una red neuronal.

Como se puede comprobar la función sigmoide (σ) comprime el rango de salida a un valor entre 0 y 1 (Figura 88. Función sigmoide de activación en una red neuronal.). Valores de z positivos harán que la función crezca exponencialmente siendo asintótica a 1, por tanto se obtendrán valores de salida $\sigma(z)$ de 1, por el contrario valores negativos de entrada z generan un valor 0 de salida en la función. Por último cuando z es 0 la función $\sigma(z) = 0.5$ (Figura 88).

Pero como se dijo anteriormente, existen muchas otras funciones activadoras, no obstante junto con las funciones sigmoides, la ReLu del inglés *rectified linear unit* y todas sus variantes, son las más empleadas (151, 152). La función ReLu tiene la siguiente expresión:

$$R(z) = \max(0, z) \quad (61)$$

Y su representación gráfica es:

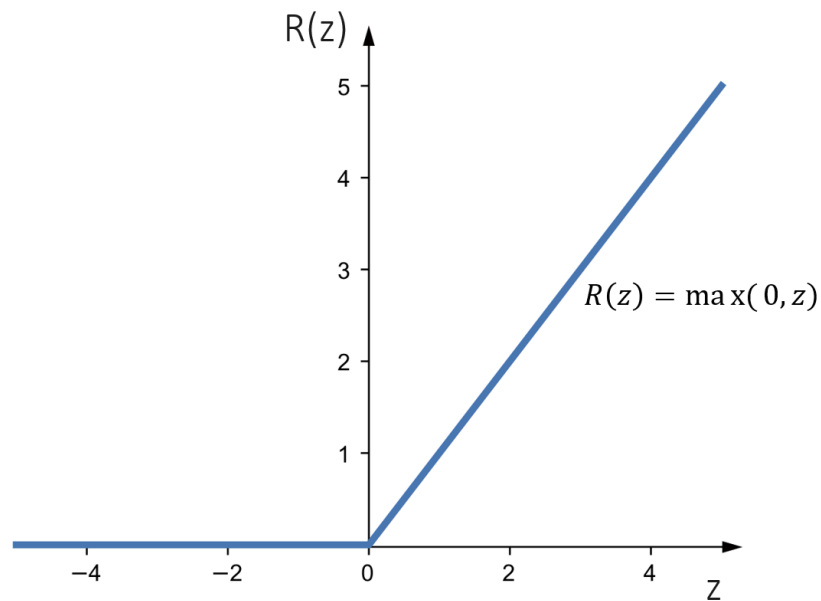


Figura 89. Función ReLu de activación en una red neuronal.

Las funciones de activación ReLU permiten el paso de todos los valores positivos sin cambiarlos, pero asigna todos los valores negativos de la función a 0 (Figura 89). Esta particularidad para RRNN de convolución en reconocimiento de imágenes es muy interesante, ya que con funciones sigmoideas la información negativa no tiene sentido y la abstracción de particularidades de datos en capas ocultas es capturada con mayor precisión. A priori, en RRNN para predecir la potencia dióptrica de una lente intraocular la información negativa no tiene mucho peso específico ya que pocas lentes son meniscos negativos, de manera que teóricamente funciones sigmoideas como activadoras funcionarían mejor. Sin embargo, se probaron varias funciones matemáticas activadoras: sigmoide, ReLu, tangente hiperbólica y SoftReLu, siendo la que mejor funcionó SoftReLu o SoftPlus. Esta es una variante de la función ReLu cuya función matemática es:

$$R(z) = \ln(1 + e^z) \quad (62)$$

Lo que permite suavizar la curva sin tanta restricción en la salida como ReLu. Su representación gráfica es:

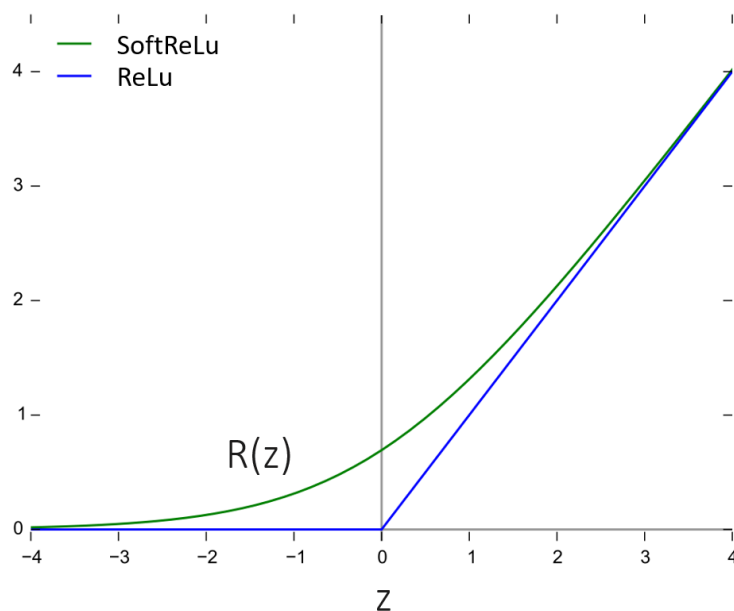


Figura 90. Función SoftReLu de activación en una red neuronal.

El hecho de no ser tan restrictiva con los valores negativos como ReLu, permite que las RRNN para predicción de meniscos negativos amplíen el rango mejorando resultados notablemente.

Pero, cuál es el propósito de una función de activación, por qué no es suficiente usar la suma ponderada directamente Ecuación (57). La razón es que la suma ponderada, z , es una combinación lineal con respecto a sus entradas (x). Sin embargo, las funciones de activación no-lineales como la sigmoide ayudan a modelar funciones curvas no triviales. La gran diferencia que existe en el uso de las RN con respecto a los modelos estadísticos clásicos, es su aplicación sin la necesidad de considerar el cumplimiento de

supuestos teóricos ni requisitos necesarios para el modelamiento estadístico, es por ello que las RN son definidas como técnicas no paramétricas.

Téngase en cuenta además, que el proceso matemático que se ha descrito para una neurona, se realiza de forma paralela para varias, con todas las que cuente la capa o capas ocultas de proceso (Figura 87).

¿Qué cometido tienen las capas ocultas? ¿Cuántas más capas ocultas, más rendimiento tiene la RN? Si la RN no tuviera capas ocultas y por tanto se tratara la información de forma directa entre las entradas y la salida, la contribución de cada entrada hacia el valor de salida sería independiente de las otras entradas. En la mayoría de los problemas del mundo real, las variables de entrada tienden a ser altamente interdependientes y afectan la salida de forma combinatoria y compleja, de manera que las neuronas de las capas ocultas permiten capturar interacciones sutiles entre los inputs.

Las capas ocultas representan características o atributos a nivel superior de los datos. Cada una de las neuronas de una capa oculta sopesa sus entradas de forma diferente, y de esta manera “aprende” características diferentes de los datos, así la neurona de salida logra capturar dichas propiedades “especiales”. Al incluir más de una capa oculta, se permite que la red neuronal pueda aprender sobre varios niveles de abstracción de los datos. Además de todo esto, las funciones de activación como se ha visto anteriormente permiten capturar relaciones no lineales entre entradas y salidas. Si se encadenan múltiples transformaciones no lineales a través de las capas, se incrementa la flexibilidad y capacidad de expresión de la red neuronal. Se puede demostrar que cualquier RN de dos capas con una función de activación no lineal y con suficientes

neuronas ocultas, es un aproximador de función universal (*function approximator*) (153, 154), o sea, capaz de interpretar cualquier mapeo arbitrario de entrada.

La RN más simple y la primera fue el Perceptrón creado por Frank Rosenblatt (155) en los años 50 (Figura 91). El Perceptrón simple o monocapa consta de una capa de entrada (inputs), una capa oculta (proceso) y una de salida (resultado), es un modelo muy extendido que resulta eficiente para modelar muchas situaciones del mundo real.

El algoritmo Perceptrón es equivalente a un hiperplano capaz de separar correctamente un conjunto de datos que sean linealmente separables si no lo son el algoritmo no funcionará bien. Siendo el hiperplano un plano aplicable a cualquier espacio geométrico independientemente del número de dimensiones.

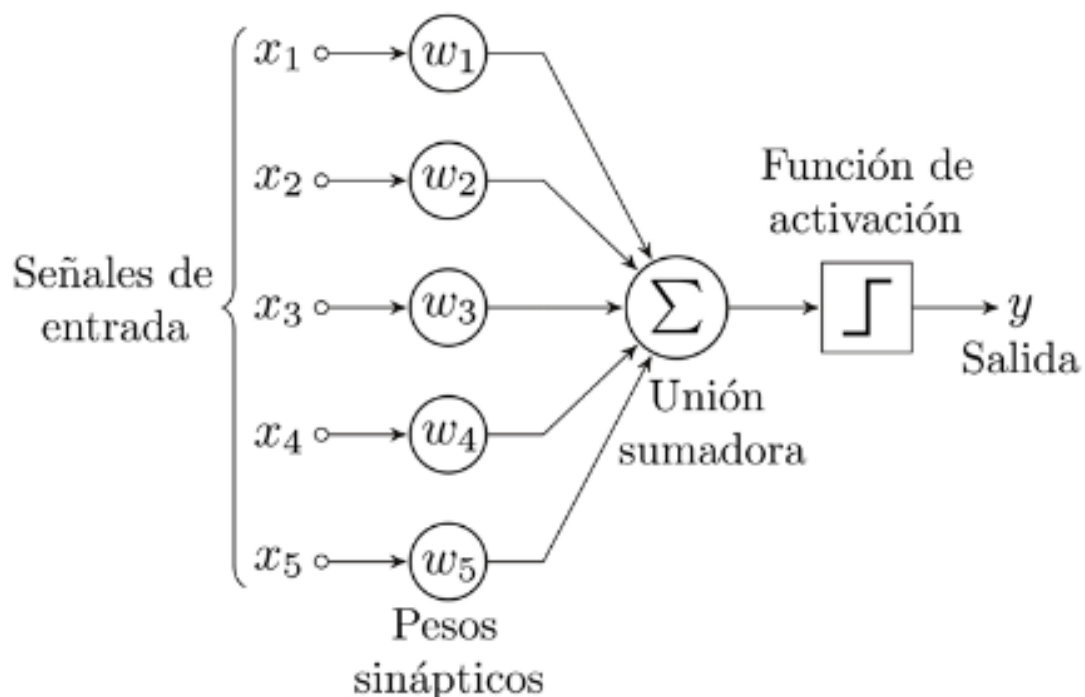


Figura 91. Esquema del algoritmo Perceptrón. (Imagen extraída de <https://es.wikipedia.org/wiki/Perceptr%C3%B3n>)

11.1.2 Funcionamiento de redes bayesianas

Afrontar el modelaje regresivo desde una perspectiva bayesiana implica introducir funciones de densidad de probabilidad como nuevo concepto para su resolución y por tanto se hablará de modelos probabilísticos o bayesianos. Tal y como se vió en el apartado [4 ESTADO DEL ARTE](#) existe un antecedente en la aplicación de estas técnicas estadísticas aplicadas al cálculo de la potencia de una lente intraocular, la fórmula FullMonte de Clarke y Kalpener.

El objetivo de todo análisis regresivo es estimar los coeficientes o parámetros que acompañan a las variables regresoras. En el análisis de regresión clásico, tras muchos supuestos teóricos; homocedasticidad, incorrelación y normalidad en los residuos, la estimación de coeficientes por el método de estimación de mínimos cuadrados ordinarios es lo convencional, dónde se intenta minimizar la suma de cuadrados de los errores o residuos entre los puntos generados por la función elegida y los correspondientes valores en los datos.

Sin embargo en un enfoque bayesiano, los coeficientes en la regresión se estimarán mediante la función de verosimilitud que indica la probabilidad de que una muestra observada dependa de los posibles valores de los parámetros. El método empleado para ello es la estimación de máxima verosimilitud que consiste en maximizar la función de verosimilitud para determinar los parámetros que tienen mayor probabilidad de producir los datos observados.

En la estimación Bayesiana se asigna una función de probabilidad a priori a los coeficientes, que se actualiza a través de la función de verosimilitud para producir una

estimación de los parámetros que tenga en cuenta la incertidumbre de los mismos. La función de probabilidad actualizada para los parámetros se obtiene empleando el teorema de Bayes, que hará las veces de algoritmo de aprendizaje (Figura 92).

La distribución a priori cumple un papel importante en el análisis bayesiano ya que mide el grado de conocimiento inicial con cierta incertidumbre que se tiene de los parámetros en estudio. Si bien su influencia disminuye a medida que más información muestral es disponible.

La inferencia bayesiana asienta directamente sobre el teorema de Bayes que formalmente se expresa diciendo que la probabilidad condicional de un evento aleatorio A dado B es igual a la distribución de probabilidad condicional del evento B dado A por la distribución de probabilidad marginal de solo A partido por la probabilidad marginal de solo B. Su definición matemática es:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (63)$$

dónde $P(A)$ es la probabilidad del suceso inicial también llamado a priori, $P(B|A)$ es la probabilidad de la nueva hipótesis basada en datos objetivos recientes que se llama verosimilitud (*likelihood*) y por último $P(A|B)$ es la probabilidad de un nuevo suceso condicionado por el primero al que se le llamará a posteriori (Figura 92).

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$


$$\underbrace{PROB(A|DATOS)}_{\text{A POSTERIORI}} = \left[\frac{1}{PROB(DATOS)} \right] \underbrace{PROB(DATOS|A)}_{\text{VEROSIMILITUD}} \underbrace{PROB(A)}_{\text{A PRIORI}}$$


Figura 92. Esquema del teorema de Bayes para modelaje probabilístico.

Sin embargo, la mayoría de las situaciones no son tan sencillas con una distribución a priori discreta y muy informativa, y por tanto resolver Bayes como a posteriori es proporcional a verosimilitud por a priori no resultará tan fácil.

Para obtener las distribuciones de densidad a posteriori de los parámetros del modelo es indispensable el uso de los métodos de aproximación, porque analíticamente no se puede especificar las densidades marginales a priori. Uno de estos sistemas para muestrear las funciones de probabilidad son los métodos MCMC (*Markov Chain Monte Carlo*) que surgen de la necesidad de simular el comportamiento de variables aleatorias y de estimar parámetros de las funciones de probabilidad de las mismas.

11.1.3 Funcionamiento de máquinas de vector de soporte

Las máquinas de vector soporte o *Support Vector Machine* (SVM) son un conjunto de algoritmos de aprendizaje supervisado desarrollados por Vladimir Vapnik (1963) y su equipo en los laboratorios AT&T.

Son algoritmos relativamente recientes ya que fueron desarrollados en la década de los 90 y son aplicables tanto a problemas clasificatorios como de regresión, como muchos de los algoritmos utilizados en ML, aunque en primera instancia se diseñaron para clasificar.

Para comprender el funcionamiento que subyace en estos sistemas, es necesario entender el concepto geométrico de hiperplano del que ya se habló previamente en [11.1.1 Funcionamiento de redes neuronales](#) en el subapartado de redes neuronales. No obstante, sin un abordaje profundo en el álgebra lineal que describe esta entidad matemática, a continuación se hará una descripción práctica y elemental para facilitar su entendimiento.

En geometría euclidiana el plano es un contenedor de infinitos puntos e infinitas rectas, pero para su descripción algebraica solo se necesitan como mínimo 3 puntos que lo compartan y que no estén alineados dentro de un espacio geométrico mínimo bidimensional.

Ahora bien, en un espacio euclídeo con n dimensiones \mathbf{R}^n , un hiperplano es un subespacio plano de dimensión $n-1$ que no tiene por qué pasar por el origen, que divide el espacio en dos mitades. De este modo, en un espacio \mathbf{R}^2 el hiperplano será una recta y en un espacio \mathbf{R}^3 será un plano. Para espacios de dimensión superior no es intuitivo

visualizar un hiperplano pero el concepto matemático de subespacio con $n-1$ dimensiones permanece.

Por simplicidad geométrica y matemática, en un espacio \mathbb{R}^2 y tal como se expresó anteriormente el hiperplano será una recta cuya ecuación es:

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 0 \quad (64)$$

Donde todos los pares de puntos x_1, x_2 que cumplen la igualdad son puntos contenidos en la recta, o lo que es lo mismo, en el hiperplano. La misma ecuación se puede generalizar para n dimensiones:

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n = 0 \quad (65)$$

Del mismo modo todos los puntos definidos por el vector $(x = x_1, x_2, \dots, x_n)$ que cumplen con la igualdad pertenecen al hiperplano. En caso de no satisfacer dicha ecuación, significa que x puede pertenecer al semiespacio de un lado u otro del hiperplano (Figura 93).

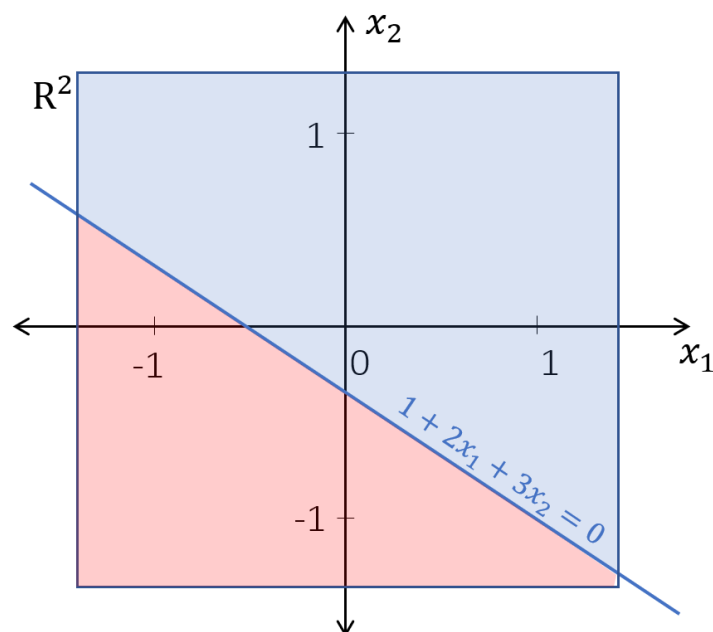


Figura 93. Hiperplano de separación en un espacio bidimensional.

Todos los puntos que pertenecen al subespacio geométrico azul (Figura 93) por encima del hiperplano que en un espacio bidimensional es una recta, no cumplen con la ecuación de la recta sino que saldrán valores positivos. En la región roja ocurrirá lo contrario, serán negativos.

Dados una serie de datos linealmente separables, existirían infinitos hiperplanos que clasificarán dichos datos (Figura 94), de manera que se dispone de un método llamado *maximal margin hyperplane* que selecciona un hiperplano como óptimo para generar la división de subespacios. Para obtenerlo, se tiene que calcular la distancia perpendicular de cada observación a un determinado hiperplano. La menor de estas distancias, conocida como *margin*, determina como de alejado está el hiperplano de las observaciones. Efectivamente, esto mismo habría que aplicarlo a las infinitas posibles soluciones lo que conlleva la imposibilidad de llevarlo a cabo, de manera que existen métodos de optimización, multiplicadores de *Lagrange*, que por su complejidad

matemática no se explicarán ya que se escapan del objetivo de esta tesis, pero que conseguirán obtener el hiperplano que cumpla el criterio marginal.

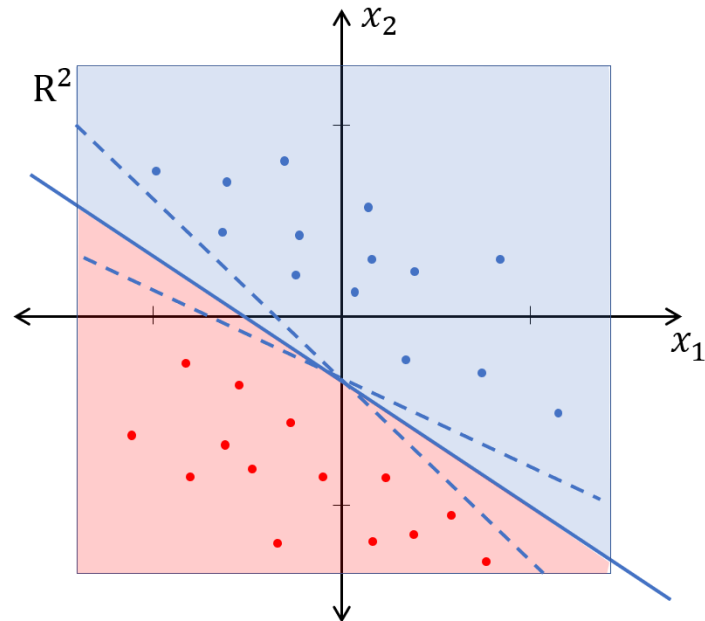


Figura 94. Infinitos hiperplanos de separación = infinitas soluciones al problema clasificadorio. El hiperplano óptimo está representado por la recta continua y las rectas discontinua representan los hiperplanos posibles.

Tras encontrar el hiperplano óptimo y sus márgenes (Figura 95), los vectores que dan nombre al método o algoritmo, los vectores soporte, son aquellos que están contenidos en los márgenes y equidistan del hiperplano.

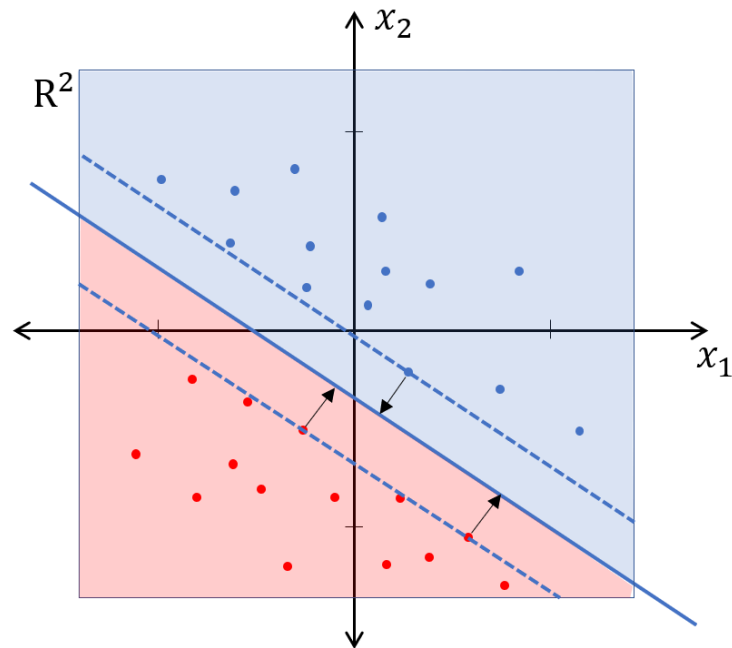


Figura 95. Hiperplano óptimo con sus márgenes y vectores soporte (3 datos equidistantes).

Hasta aquí, todo esto es válido cuando los datos son linealmente separables, pero qué ocurre cuando se tienen conjuntos de datos que no cumplen dicho criterio, algo habitual por otro lado. En este punto, se introduce un nuevo concepto llamado kernel, que hace referencia etimológicamente a un núcleo, es decir, a un algoritmo específico y fundamental del modelaje que permite encontrar el hiperplano óptimo en caso de tener conjuntos de datos no linealmente separables con espacios multidimensionales. Esto significa que un modelo SVM puede tener diferentes tipos de kernel que se corresponden a diferentes maneras de optimizar un espacio multidimensional.

El hecho de que los grupos no sean linealmente separables en el espacio original no significa que no lo sean en un espacio de mayores dimensiones (Figura 96). Una posible estrategia para enfrentarse a escenarios en los que la separación de los grupos es de tipo no lineal consiste en expandir las dimensiones del espacio original. Pero cómo saber que dimensión es la adecuada para analizar el problema, ahí es dónde los kernel

tienen un papel fundamental. Por tanto la selección del kernel es capital para el rendimiento del modelo, ya que de este dependerá la dimensionalidad, y por ende el hiperplano y los márgenes.

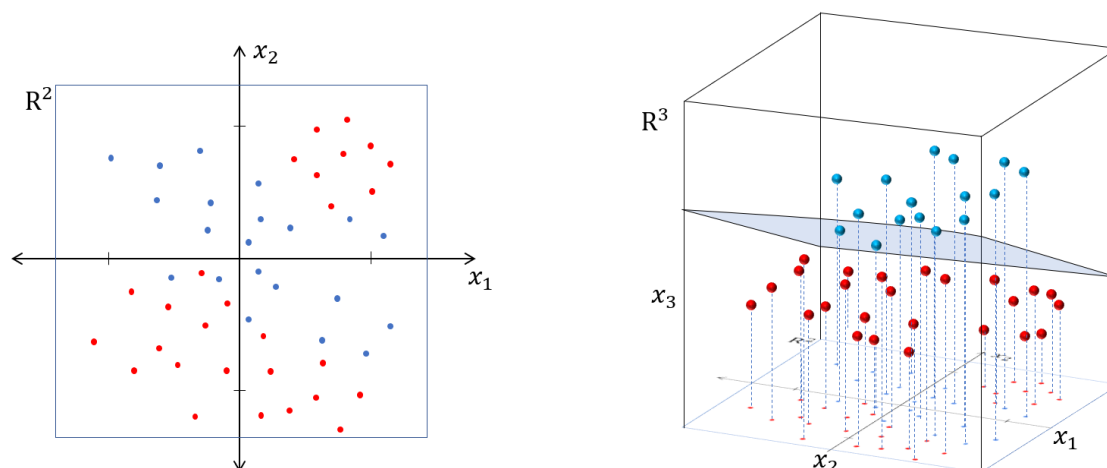


Figura 96. Mismo conjunto de datos analizado en diferentes dimensiones (R^2 y R^3). Cuando las observaciones pertenecen a un espacio bidimensional no son linealmente separables (diagrama de la izquierda). Cuando los mismos datos pertenecen a un espacio tridimensional con un hiperplano sí se pueden separar (diagrama de la derecha).

11.1.4 Funcionamiento de regresión multivariante adaptativa con splines

Multivariate Adaptive Regression Splines (MARS) hace referencia a una forma de análisis regresivo y clasificatorio introducido por Friedman (1985). El término *spline* quiere decir curva suavizada, sin zonas abruptas, desde un punto de vista matemático un *spline* es una curva diferenciable definida en porciones o segmentos mediante polinomios.

Muchos de los modelos intrínsecamente lineales se pueden adaptar a patrones no lineales de los datos agregando manualmente términos de modelos no lineales por ejemplo: términos cuadrados, efectos de interacción y otras transformaciones de las características originales. Sin embargo, para hacerlo, el analista debe conocer a priori la naturaleza específica de las no linealidades e interacciones.

Alternativamente, existen numerosos algoritmos que son inherentemente no lineales, como ya se ha visto: RN, modelos probabilísticos o SVM. Cuando se usan estos modelos, no es necesario conocer explícitamente o especificar la forma exacta de la no linealidad antes de la implementación del modelo. Más bien, estos algoritmos buscarán y descubrirán no linealidades e interacciones en los datos que ayudan a maximizar la precisión predictiva.

Existen otros algoritmos intrínsecamente no lineales como la regresión polinómica o las funciones escalonadas o *step functions* (158). En el primer caso, para conferir flexibilidad a un modelo lineal se introducen nuevos predictores conseguidos cuando se eleva a distintas potencias el predictor primigenio (Figura 97(B) y (C)), de esta forma se genera una única función global que describe el comportamiento de la variable dependiente en todo el rango del predictor x_i .

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_n x_i^n + \epsilon_i \quad (66)$$

La debilidad de este sistema es que para aumentar la flexibilidad del modelo con la intención de captar mejor la no linealidad se tiene que ampliar órdenes, pero polinomios de grados superiores a 3 o 4 incrementa la multicolinealidad y produce *overfitting* (159).

La alternativa a la regresión polinómica es la función escalonada (Figura 97 (D)). Es aquella función definida a trozos que en cualquier intervalo finito en que esté definida tiene un número finito de discontinuidades $c_1 < c_2 < \dots < c_n$ o puntos de corte también llamados *knots*. Se divide el rango del predictor x_i en subintervalos y se ajusta una constante C distinta para cada uno.

La principal desventaja de este método es que la mayoría de predictores no tienen *knots* establecidos, por lo que al imponerlos en determinadas posiciones se puede estar perdiendo la naturaleza de la relación.

Aunque tienen su utilidad, la implementación de la regresión polinómica y las funciones escalonadas requieren que el usuario identifique e incorpore explícitamente qué variables deben tener qué grado específico de interacción o en qué puntos de una variable x_i deben hacerse puntos de corte para las funciones escalonadas. Considerando que muchos conjuntos de datos actuales pueden contener fácilmente 50, 100 o más características, esto requeriría un compromiso de tiempo enorme e innecesario por parte de un analista para determinar estas configuraciones explícitas no lineales.

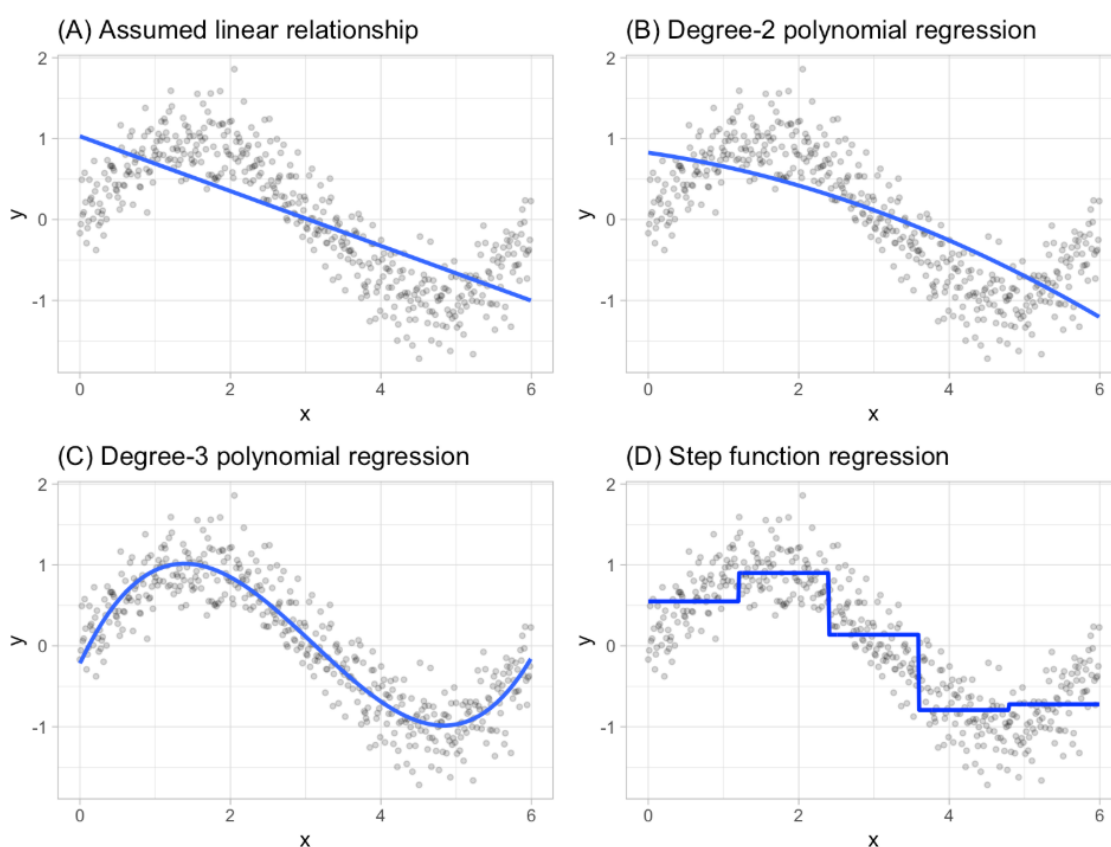


Figura 97. (A) Regresión lineal normal. (B) Regresión polinómica de segundo grado en la que se capta algo de no linealidad en el dataset. (C) Regresión polinómica de tercer orden en el que se capturó todas las relaciones. (D) Step Function aplicada al mismo dataset con 5 subintervalos. (Imagen extraída de <https://bradleyboehmke.github.io/HOML/mars.html>)

MARS es una extensión de la regresión polinómica y de las *step functions* consiguiendo una mayor flexibilidad ya que auna las fortalezas de ambos métodos. Consiste en dividir el rango del predictor x_i en K subintervalos. Para cada una de las nuevas regiones se ajusta una función polinómica, introduciendo una serie de restricciones que hacen que los extremos de cada función se aproximen a los de las funciones de las regiones colindantes.

Una de las restricciones consiste en forzar a cada polinomio a pasar por los puntos de corte que lo delimitan, de esta forma, el polinomio de la región i termina en el mismo punto donde empieza el polinomio de la región $i+1$ (Figura 98). A pesar de que esta restricción permite obtener una curva continua, el cambio de una región a otra es excesivamente abrupto y poco natural, así que para suavizar la curva en las transiciones se imponen otras condiciones, que la primera y segunda derivada sean continuas y que pasen por los puntos de corte del intervalo.

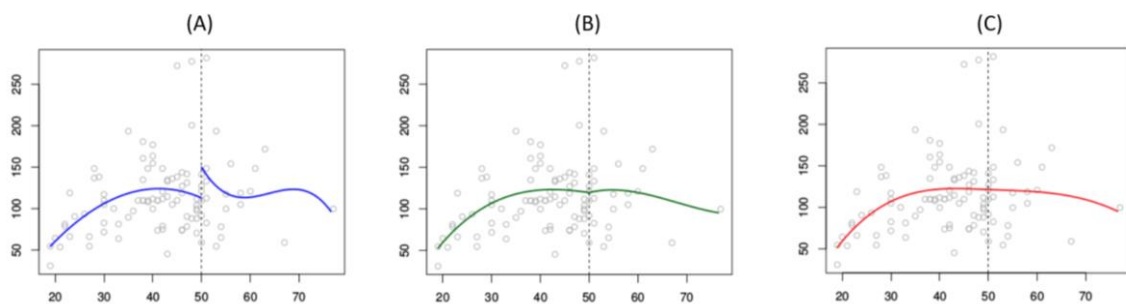


Figura 98. Tres fases de optimización que utiliza MARS para suavizar el spline. (A) Primera fase de *step function* donde se imponen los puntos de corte (*knots*) de la función escalonada. (B) Segunda fase en la que se obliga al segmento polinómico a que los puntos de corte coincidan. (C) Tercera fase o *smoothing* en la que se suaviza el *knot* mediante las primeras y segundas derivadas de las funciones contiguas.

El resultado de MARS dependerá en gran medida de la cantidad de *knots* que se introduzcan, así como de sus posiciones. Cuantos más *knots*, más flexible es la función y

mejor se adaptará a los datos reales captando mejor la no linealidad. Pero ya se advirtió anteriormente que esto podría aumentar la multicolinealidad, por tanto para obtener el número óptimo de puntos de corte en la práctica se utiliza como técnica la validación cruzada o *cross validation*.

El término "MARS" está registrado y con licencia exclusiva de Salford Systems (<https://www.salford-systems.com>). Se puede usar MARS como abreviatura, pero para implementar los modelos en R se utilizó una librería llamada *Enhanced Adaptive Regression Through Hinges* (EARTH) combinada con CARET.

11.2 ANEXO 2.- ANÁLISIS CRÍTICO: MÉTRICAS PARA EVALUAR LA PRECISIÓN Y EXACTITUD DE FÓRMULAS

11.2.1 Actualidad y discrepancias entre trabajos

Determinar objetivamente qué método o fórmula de cálculo es mejor que otra y establecer un ranking entre ellas, resulta muy complicado debido a que actualmente no hay un protocolo consensuado y universal sobre qué descriptor o métrica se debe usar para tomar una decisión.

Actualmente en la literatura oftalmológica publicada y revisada por pares, los descriptores aceptados y empleados para seleccionar una u otra fórmula como mejor son los siguientes:

Mean Predictive Error (MPE), *Mean Absolute Error* (MAE), *Median Absolute Error* (MedAE), *Standard Deviation* del error absoluto (SD-MAE), el rango de errores absolutos máximo-mínimo y el porcentaje de ojos dentro de un rango de *Predictive Error* (PE), habitualmente ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D.

El MPE o error predictivo medio se calcula promediando todos los errores predictivos o *Predictive Errors* (PE) o lo que es lo mismo, el sumatorio de la diferencia entre el equivalente esférico postoperatorio (EE_{post}) y la predicción refractiva ($T_{fórmula}$) que hace la fórmula en cuestión para cada individuo, dividido por el número de observaciones:

$$MPE = \frac{\sum_{i=1}^n (EE_{post} - T_{fórmula})}{n} \quad (67)$$

El MAE es el valor absoluto de error de la fórmula para cada individuo promediando toda la muestra:

$$MAE = \frac{\sum_{i=1}^n |EE_{post} - T_{fórmula}|}{n} \quad (68)$$

Por último MedAE es el valor absoluto de error de la fórmula de cada individuo, por tanto igual que MAE, pero obteniendo la mediana en vez de la media del total de la muestra.

En muchos trabajos recientes se seleccionan las mejores fórmulas según la desviación estándar (SD) (5, 90, 105) del error de predicción, en otros basándose en el Mean Absolute Error (MAE) (160-162) y en otros sobre el Median Absolute Error (MedAE) (163, 164). Incluso los nombres de las métricas utilizadas cambian de un trabajo a otro, a veces se le llama Mean Numerical Error (165) y en otras ocasiones Mean Arithmetic Refractive Error (166), cuando en realidad se refieren al error de predicción medio o Mean Prediction Error (MPE). Hay estudios en los que no se hace referencia a MedAE o MPE (167), otros únicamente en los que se habla de MAE y otros en los que no se describe ni la desviación estándar de los errores absolutos medios (164).

Respecto al tratamiento estadístico en este tipo de trabajos también ha habido controversia, ya que en algunas ocasiones no se analizan los resultados de forma adecuada. Hay estudios (167) en los que se aplica un análisis de varianza ANOVA con poblaciones de datos no normales como son los MAEs. Otros en los que se utiliza la prueba F de Fisher para establecer qué fórmula es mejor en base a sus MAEs (5). Se han publicado artículos (168) en los que no se estudia la distribución de los datos y por tanto

la aplicación posterior de los test paramétricos de contraste de hipótesis puede que no sean los correctos.

En resumen, deberían establecerse unos criterios de selección estandarizados, uniformes y coincidentes, probablemente renovados y revisados, sobre los que trabajar para decidir de una manera universal y homogénea, qué fórmula es mejor cuando se comparen varias. Este hecho facilitaría ulteriores meta-análisis así como exámenes comparativos entre artículos.

Además en múltiples ocasiones, cuando se trabaja con muchas fórmulas de cálculo biométrico que rinden de manera muy similar, con las métricas de las que se dispone hoy en día no es tarea fácil concluir una ordenación de dichas fórmulas como se verá a lo largo del presente capítulo. Asimismo, es frecuente encontrar discrepancias en los estudios comparativos de fórmulas dependiendo de la métrica seleccionada, lo que implica que una fórmula puede ser mejor que otra según se elija MAE o MedAE incurriendo en una incongruencia.

No obstante se ha intentado en el último lustro homogeneizar y universalizar en varias ocasiones el sistema de selección de fórmulas en estudios comparativos, siendo el mayor propulsor de esto KJ Hoffer (169) pero sin aportar ninguna novedad respecto a lo ya establecido y usado en los últimos 40 años.

En primer lugar conviene discriminar entre dos conceptos diferentes: precisión y exactitud de una fórmula de cálculo biométrico (Figura 99). La precisión hace referencia a la dispersión del conjunto de valores obtenidos de resultados refractivos en torno a 0.00 D como valor óptimo. Se podría hablar de precisión de MPE y precisión de MAE. Por

otro lado, exactitud se refiere a cuán cerca del valor óptimo (0.00 D) se encuentra el valor medio medido. Del mismo modo se podría distinguir exactitud de MPE y de MAE.

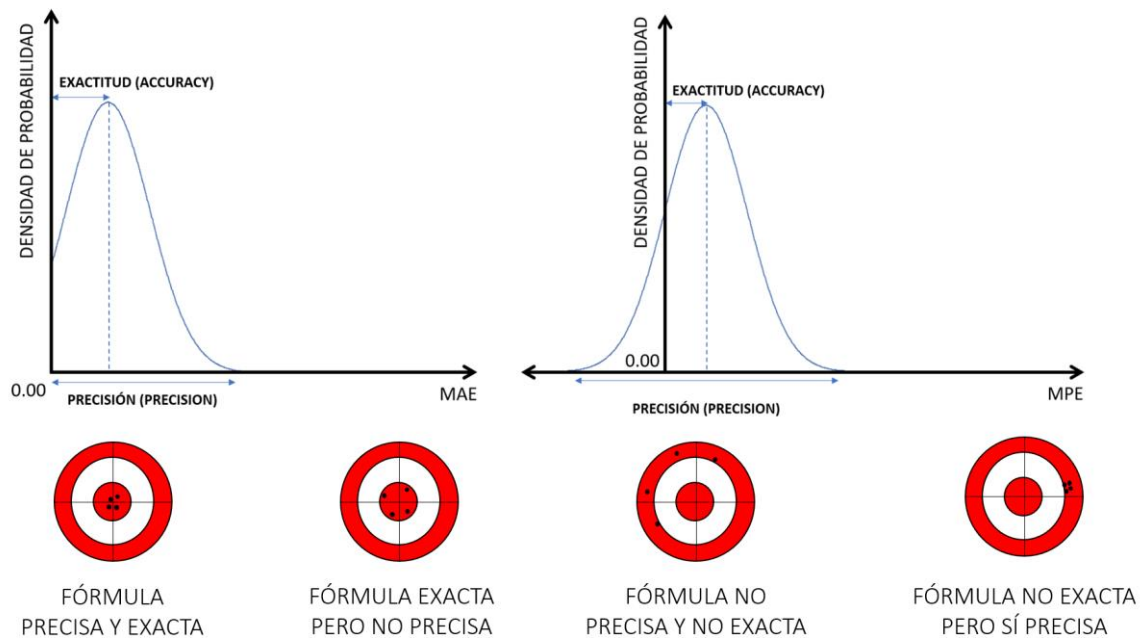


Figura 99. Diferencias gráficas entre precisión y exactitud tanto para el promedio de errores absolutos como para el promedio de errores predictivos en las fórmulas y métodos de cálculo.

Una fórmula puede ser muy precisa porque todos sus valores de MPE y/o MAE tienen poca dispersión respecto a un valor, pero sin embargo poco exacta porque dichos valores no se acercan a 0.00 D. O, al contrario, muy exacta porque muchos de sus resultados refractivos están muy cerca de 0.00 D, pero muy imprecisa porque presenta gran dispersión en esas magnitudes. De modo que aquella fórmula que tenga un MAE cercano a cero con una desviación estándar muy contenida tendente a cero, será una fórmula muy precisa y exacta al mismo tiempo, como se muestra en la primera diana empezando por la izquierda en la Figura 99.

11.2.2 Error predictivo de refracción

Se da por hecho que los PE se distribuyen normalmente, pero esto no es siempre así. Si bien el target de PE es 0.00 D de equivalente esférico postoperatorio y los residuos refractivos postquirúrgicos tienden a distribuirse a los dos lados de manera más o menos uniforme, hay fórmulas que tienen una tendencia manifiesta hacia la ametropía negativa o positiva en función de la tipología de ojo en la que esté calculando dicha fórmula. Por ejemplo, SRK/T en ojos con longitudes axiales por debajo de 22.00 mm tiende a subestimar la potencia (170) de la IOL haciendo que los PE en una muestra poblacional sean más positivos que negativos, de manera que la distribución de datos pudiera tener un sesgo o *skewness* hacia la derecha, alejándose sutil o abruptamente de la campana gaussiana. Por tanto, y dado que esto puede ocurrir con cualquier fórmula a priori, es preceptivo explorar la normalidad de la distribución de datos PE de forma analítica y gráfica. Para la parte matemática se propone una triple metodología mediante contrastes de hipótesis para los test de *Kolmogorov-Smirnov* con corrección de *Lilliefors*, *Shapiro-Wilk* y *Anderson-Darling*. Para la parte gráfica un *QQ-plot* acompañado de un histograma (Figura 100) será suficiente.

En la Figura 100 se muestra un ejemplo real de la no normalidad de PE cuando se compararon dos fórmulas para ojos con longitudes axiales por encima de 25.00 mm. Si se establece el nivel de significación en 0.05, la fórmula 1 no distribuye normalmente sus datos de error predictivo, como demuestran los tres test aplicados ($p < 0.05$), observándose un claro sesgo hacia la izquierda, que implica tendencia miópica en los resultados refractivos postoperatorios. Por el contrario, la fórmula 2 muestra una

distribución normal ($p > 0.05$) salvo para el test de *Shapiro-Wilk*, no obstante téngase en cuenta que el tamaño muestral es 114 observaciones.

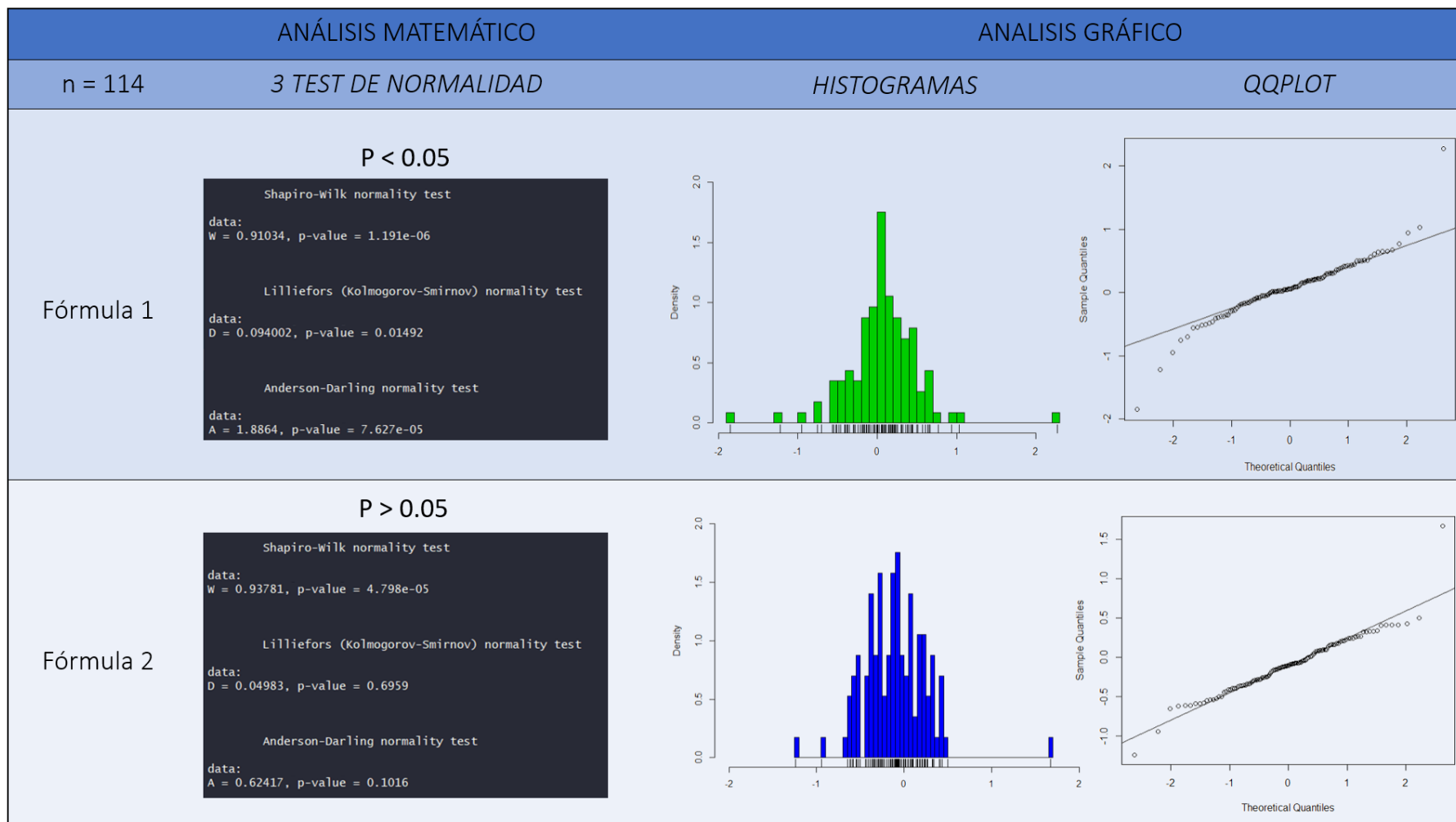


Figura 100. Análisis matemático y gráfico sobre una muestra de 114 observaciones en el que se está contrastando la normalidad o no de la distribución de errores predictivos para dos fórmulas de cálculo.

El hecho de utilizar tres tipos de contrastes de normalidad es porque *Shapiro-Wilk* está indicado para muestras inferiores a 50 observaciones, mientras que *Kolmogorov-Smirnov* para muestras superiores. *Anderson-Darling* se utiliza como tercera vía en caso de incertidumbre, siendo un test muy potente porque se adapta muy bien a cualquier tipo de distribución mostrando menos sensibilidad al tamaño muestral.

En cuanto al test estadístico a aplicar, si la distribución fuera normal un *one sample t test* sería perfecto, mientras que en distribuciones no normales un *Wilcoxon Signed Rank Test* sería el indicado.

Por otro lado, si en cualquier trabajo comparativo de fórmulas se utilizan distintos tipos de lentes intraoculares con diferentes constantes, todas las fórmulas deberían competir en las mismas condiciones para poder compararse. Para conseguir este escenario ideal, el MPE debería ser cero cambiando la constante de la lente individualmente para cada fórmula. De esta forma, la posible desviación que pueda surgir en los cálculos derivados del uso de diferentes constantes queda eliminada.

El procedimiento consistirá en una transformación por centrado de los datos de PE. Esto se consigue operando algebraicamente cada valor individual de PE con su media y promediando posteriormente toda la muestra para obtener el MPE ajustado (Figura 101). Cuando se obtienen valores positivos de MPE sin centrar, el operador algebraico que se tendrá que emplear en el centrado con el valor individual de PE será negativo y viceversa. En la Figura 101 se puede observar como para la primera fórmula el MPE antes de ajustar se sitúa en valores negativos, concretamente -0.29 D, lo que implica en términos generales que esta fórmula tiende a sobre corregir la potencia de IOL dejando a los pacientes miopes.

Una vez hecho el ajuste se conseguirá un MPE de 0.00 habiendo desplazado hacia la derecha la distribución. En el caso de la otra fórmula se obtiene un MPE de 0.20 D, lo que significa una tendencia hipermetrópica en la fórmula hipocorrigiendo la estimación de la potencia de IOL. Tras el ajuste, habiendo restado la media a cada valor individual de PE, se desplaza la distribución hacia la izquierda centrándola y obteniendo un *mean* de 0.00 D.

“CENTERING” DE UNA DISTRIBUCIÓN

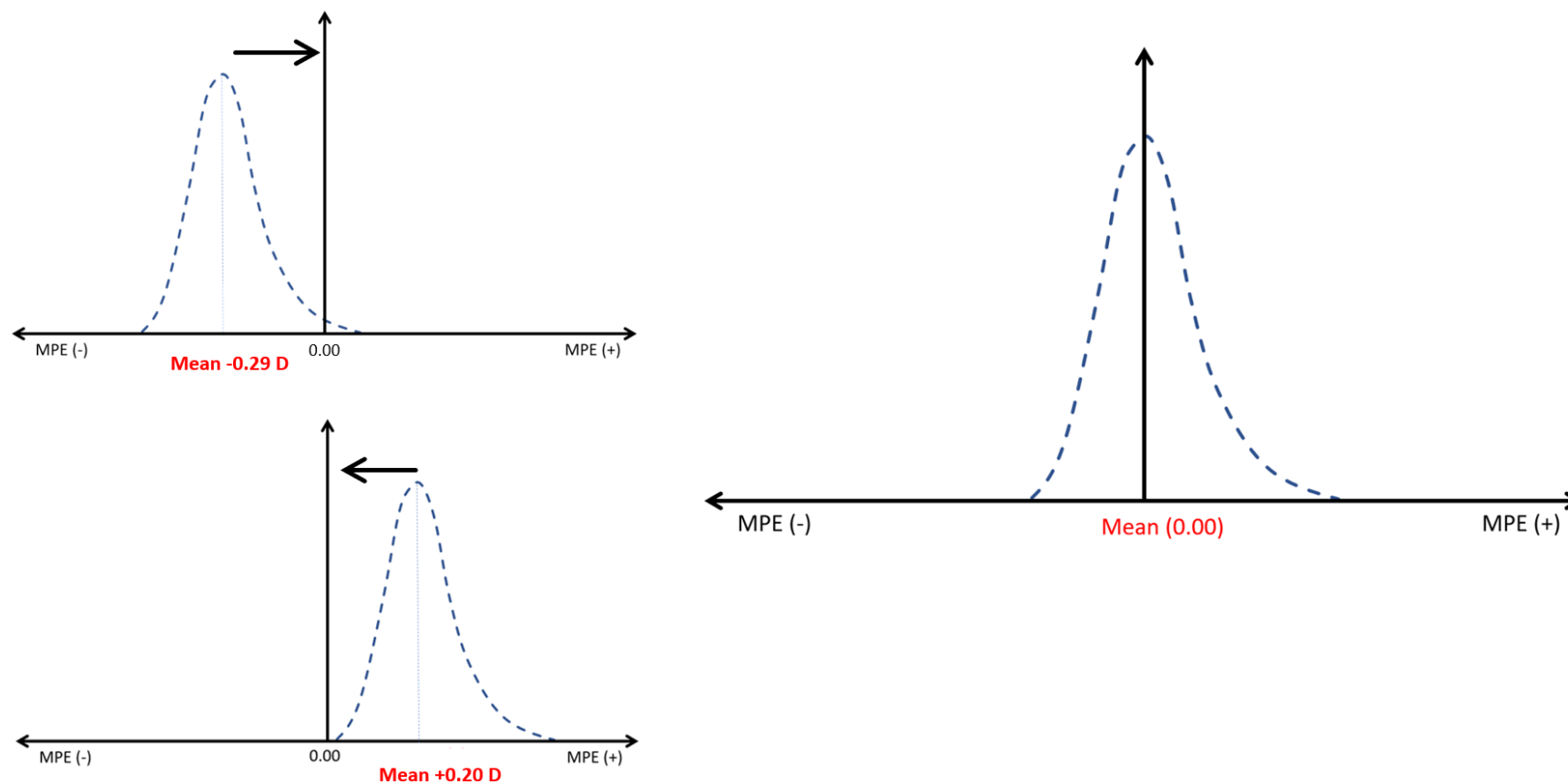


Figura 101. Centrado de una distribución para que la media de errores predictivos sea 0.00 D.

11.2.3 Error absoluto medio

Tras haber completado esta primera fase de análisis con los errores predictivos, se calculan los MAE y MedAE de cada fórmula aplicando el valor absoluto a cada PE ajustado previamente, determinando la media y la mediana respectivamente, tal y como ya se explicó anteriormente.

La distribución de los datos MAE deja de ser normal ya que solo hay valores positivos y por tanto se distribuyen únicamente a la derecha de 0.00 D. Se puede comprobar que los tres contrastes de normalidad arrojan un p-valor inferior a 0.05, lo que implica que la distribución no es normal, así como en el *QQ-plot* se muestra una convexidad alejada de la *QQ-line* o línea de normalidad. El tipo de distribución de datos a los que se ajusta la población de MAEs, usualmente es exponencial, del tipo *Gamma* o *Weibull* como se demuestra en la parte inferior derecha de la Figura 102.

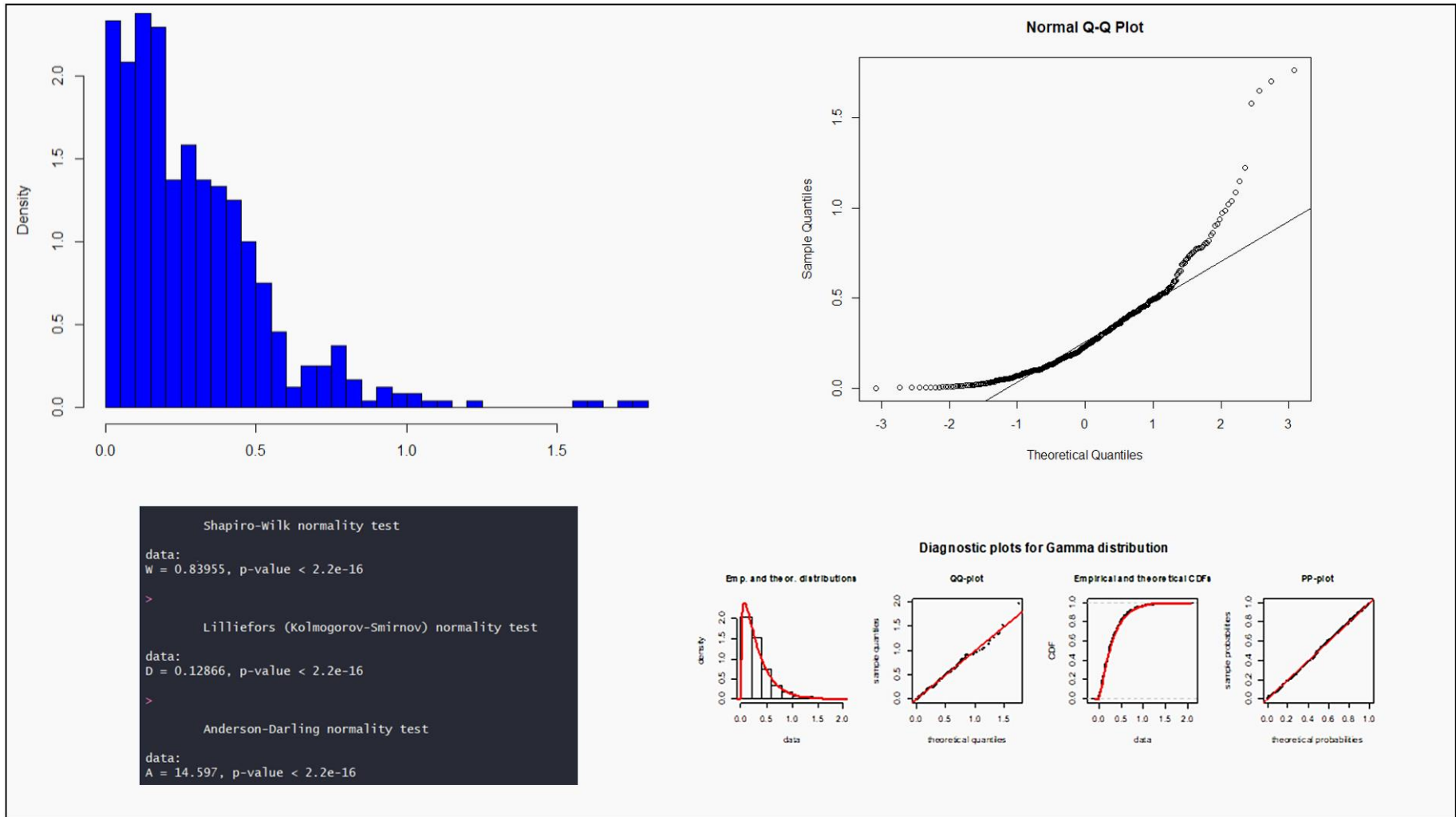


Figura 102. Demostración gráfica y analítica de que los errores absolutos se distribuyen según una distribución exponencial gamma alejada de la normal gaussiana.

En este punto, tienen mucha relevancia los valores atípicos o *outliers*. En la población no normal de MAEs los valores atípicos tienen mucha influencia sobre los valores de centralización y dispersión: media y desviación estándar respectivamente. Siguiendo esta línea, Hoffer et al (171) en primera instancia recomendaron utilizar como descriptor de comparación entre fórmulas MedAE que es menos sensible a los outliers, por el contrario, en un trabajo más reciente (169) aconsejan emplear las dos métricas: MAE y MedAE, así como Wang et al (172) que promulgan lo mismo. Pero sin embargo, Kane et al (173) abogan por utilizar MAE y no MedAE, ya que este último podría conducirnos a selecciones equivocadas de fórmulas en casos particulares.

Un punto importante, es analizar todas las métricas con precisiones superiores a las que hasta ahora se han empleado, lo habitual era hacerlo con dos decimales. Debido a la irrupción de nuevas tecnologías, fórmulas y métodos de cálculo en los últimos años cada vez más precisos, en ocasiones la utilización de precisiones centesimales impide cualificar el resultado, por lo que se propone en esta tesis la precisión milesimal.

En cualquier caso, es muy aconsejable estudiar y analizar los *outliers* gráficamente mediante un diagrama de cajas y bigotes (Figura 103). En esta figura de comparación real entre 11 fórmulas se pueden observar todos los valores atípicos por encima del último bigote de cada boxplot. En un diagrama de caja, el rango intercuartílico o *InterQuartile Range* (IQR) es la distancia entre el primer cuartil (Q1) y el tercero (Q3) y encierra el 50% de los datos de la muestra. IQR es un estadístico robusto ya que no está influido por los valores anómalos extremos. Se considera un valor atípico leve el que se encuentra por encima de 1,5 veces la distancia del rango intercuartílico más allá del último cuartil, y outlier extremo si esa distancia es 3 veces IQR.

En este caso todas las fórmulas presentan outliers en sus MAEs (Figura 103). La que tiene el valor atípico más extremo es la fórmula 2 que está en 2.50 D, y el outlier más suave es el de la fórmula 6 que se corresponde con un valor de 0.80 D.

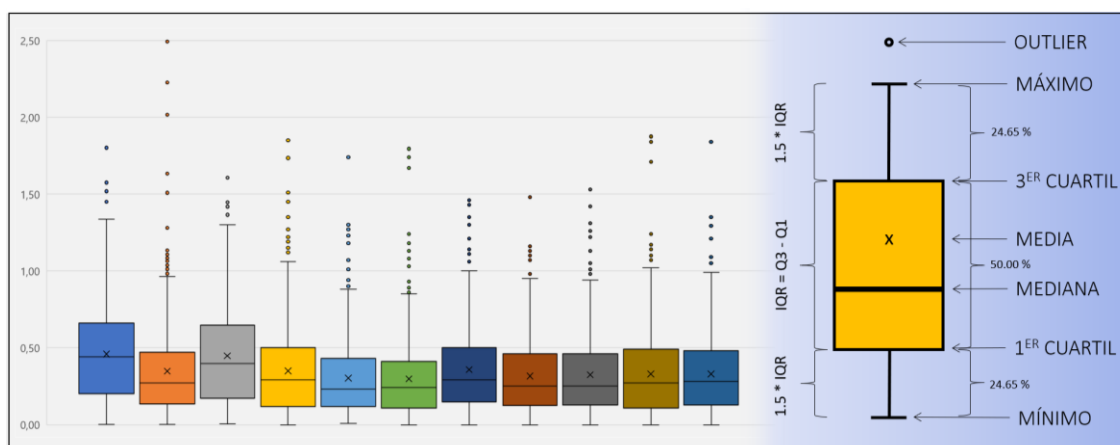


Figura 103. Boxplots para los MAE de 11 fórmulas y la explicación detallada en el margen derecho de interpretación de un gráfico de cajas y bigotes

Si se atiende al tercer decimal del descriptor MAE para la ordenación de fórmulas, muy susceptible a los outliers (Figura 104), se puede comprobar que la fórmula 6 tiene el MAE más bajo: 0.292 D seguida de las fórmulas 5, 8 y 10, con valores de MAE de 0.293 D, 0.305 D y 0.311 D respectivamente. La peor de todas en base a MAE es la fórmula 3 con un valor de 0.418 D. Ahora bien, si se clasifica por MedAE la ordenación cambia, el valor más bajo corresponde a las fórmulas 6 con 0.233 D, 9 con 0.238 D y 8 con 0.241 D respectivamente. El peor valor de MedAE vuelve a ser para la fórmula 3 con un valor de 0.355 D. Pero si se utiliza la SD-MAE como métrica de selección, el ranking de precisión ahora sería fórmula 8, 5 y 6 con valores 0.251 D, 0.254 D y 0.255 D respectivamente. Ahora la fórmula 1 tiene el valor de dispersión más alto, 0.328 D.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
MAE	0,404	0,352	0,418	0,327	0,293	0,292	0,328	0,305	0,321	0,311	0,312
SD	0,328	0,318	0,322	0,284	0,254	0,255	0,272	0,251	0,284	0,258	0,257
MedAE	0,335	0,277	0,355	0,263	0,243	0,233	0,265	0,241	0,238	0,254	0,253
RANKING POR					1	2	3				
MAE					F6	F5	F8				
SD					F8	F5	F6				
MedAE					F6	F9	F8				

Figura 104. Métricas clásicas (MAE, SD y MedAE) para clasificar precisión y exactitud de las fórmulas y métodos de cálculo biométrico.

En resumen, la fórmula 6 es la mejor si clasificamos por MAE y por MedAE y la fórmula 8 sería la más precisa al ordenar por SD-MAE. En este caso hay coincidencia en dos de las métricas para la primera fórmula lo que a priori podría fortalecer el argumento de selección. Pero qué pasa cuando hay discrepancia entre las tres métricas ¿Cómo se actúa? ¿A qué métrica o métricas se recurre para la selección?

Para ahondar en el funcionamiento de la sistemática de estas métricas clásicas se verá por simplificación otro ejemplo con solo tres fórmulas (Figura 105) y únicamente 10 valores en valor absoluto de PE.

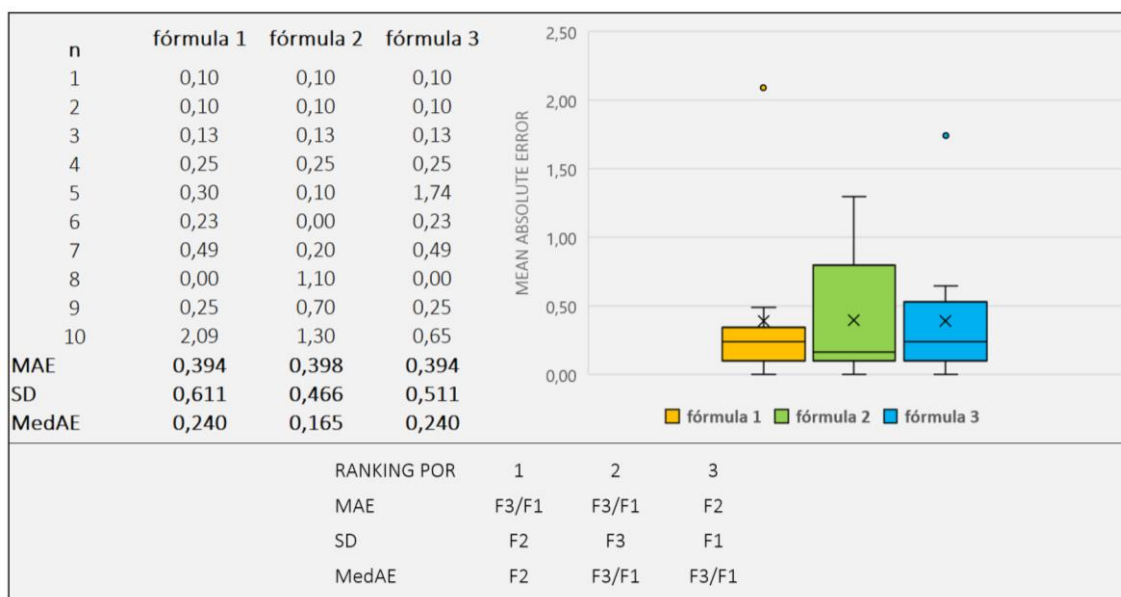


Figura 105. Ejemplo con 3 fórmulas y solo 10 observaciones en valor absoluto para calcular los MAE, SD y MedAE. Ordenación de las fórmulas en base a estas métricas.

En este caso, si se pone atención en los valores en valor absoluto de PE y únicamente evaluamos las métricas mostradas, MAE, MedAE y SD, surge un grave conflicto de selección ya que según SD y MedAE la fórmula 2 es la mejor, pero según MAE la fórmula 1 y 3 igualan sus valores de media a 0.394 D. Ahora bien, si se analizan los valores absolutos de PE de las tres fórmulas mostrados en la columna de la izquierda, la fórmula 1 y 3 repiten valores salvo en la posición 5 y 10, donde nos encontramos 0.30 D contra 1.74 D y 2.09 D contra 0.65 D respectivamente. Las diferencias de error entre fórmulas en ambas posiciones son iguales, 1.44 D en ambos casos, de ahí que el promedio sea el mismo para las dos fórmulas, pero cuando la fórmula 1 falla en esos dos valores, salvo uno (el outlier), el otro lo introduce por debajo de 0.50 D, sin embargo la fórmula 3 tiene los dos fallos por encima de 0.50 D. Por otro lado, la fórmula 3 tiene una precisión mayor que la fórmula 1 porque dispersa menos los resultados respecto a la media presentando una SD de ± 0.511 D respecto a ± 0.611 D. De manera que definitivamente

se puede asegurar que la fórmula 3 es mejor en términos de precisión que la fórmula 1, pero no en cuanto a exactitud, siendo la fórmula 1 ligeramente mejor.

Si se compara ahora la fórmula 3 con la fórmula 2 que según SD y MedAE es la mejor, en los cuatro primeros valores absolutos de PE coinciden, en los tres siguientes la fórmula 2 es notablemente mejor que ninguna de las otras dos, y en los tres últimos es donde la fórmula 2 falla sustancialmente respecto a las otras. La diferencia de errores entre la fórmula 2 y 3 es a favor de la fórmula 3, concretamente 2.16 D respecto a 2.20 D, de manera que la fórmula 3 falla menos que la 2 en términos generales. De lo que se deduce que la fórmula 1 es la mejor de las tres a pesar de tener una SD y un MedAE peores que la fórmula 2 e idénticos a la fórmula 3, seguida de la fórmula 3 y por último la 2.

Estas inferencias no son nada intuitivas si no se presta atención sobre los outliers que presente cada fórmula, ni sobre los datos de error absoluto de cada observación, no se sabrá nunca cuánto y cómo afectan a los valores de centralidad y dispersión de las fórmulas en cuestión. De ahí que las desviaciones típicas y los promedios de las fórmulas 1 y 3 con presencia de outliers se vean más afectados que en la fórmula 2. Pero si la mediana es un estadístico robusto que por definición no se ve influenciado por los valores atípicos, y en este caso la fórmula 2 presenta el mejor valor de MedAE (0.165 D) ¿por qué no es la mejor fórmula? El próximo apartado dará respuesta a esta pregunta.

11.2.4 Mediana de errores absolutos

La mediana divide la muestra en dos partes iguales, 50% a cada lado del estadístico con la distribución ordenada de menor a mayor. Si n es impar la mediana corresponde al dato central, pero si n es par la mediana equivale a la media aritmética de los dos datos centrales (Figura 106). Este es uno de los grandes inconvenientes de usar la mediana, la omisión de información de todas las observaciones excepto la central o centrales, sin embargo como contrapartida es muy poco sensible a los *outliers*.

Los valores absolutos centrales de PE (en negrita) ordenados de menor a mayor en las tres fórmulas son más bajos en la fórmula 2 que en las otras, lo que no asegura que ni por encima ni por debajo de este valor la distribución de errores absolutos sea mejor una que otra. De hecho, se observa que el valor cuarto y quinto de la fórmula 2 por debajo de la mediana son más bajos que en las otras. Sin embargo, el octavo y noveno son mayores que los demás. Es más, si los valores de la posición octava, novena y décima fueran por ejemplo 3.00 D, 3.22 D y 3.17 D respectivamente, la mediana de 0.165 D no cambiaría. De lo que se deduce que la mediana como métrica de selección no es suficientemente precisa en función de qué casos y de qué observaciones se disponga.

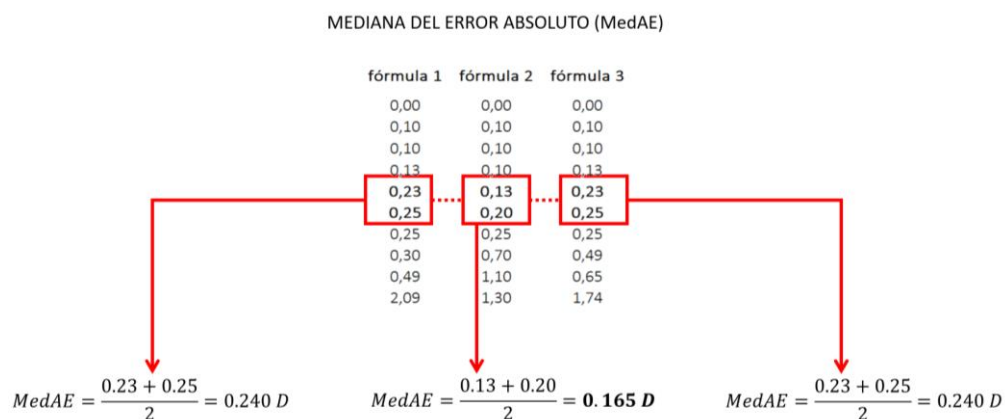


Figura 106. Cálculo de la mediana de los errores absolutos como métrica robusta en la clasificación de fórmulas de cálculo biométrico.

Para este ejemplo concreto, como ya se expresó anteriormente se puede asegurar que la fórmula 3 es la mejor, seguida de la 1 y de la 2 según todos los análisis antes expuestos. Sin embargo atendiendo a los estadísticos clásicos aceptados en la literatura oftalmológica publicada, la ordenación produciría confusión.

Imáginese que de una muestra de 99 observaciones, la mediana es 0.49 D, los 49 datos ordenados por encima llegan hasta 0.48 D desde 0.00 D y los 49 errores absolutos por debajo de la mediana llegan hasta 1.33 D. Si se atiende en la clasificación a la métrica MedAE ¿sería esta una buena fórmula en términos de exactitud? Aparentemente sí, pero si se compara con otra fórmula cuya mediana es 0.45 D cuyos errores absolutos llegan por debajo de la mediana hasta 1.37 D, la fórmula con mediana 0.45 D puede que sea mejor, pero sin la certeza absoluta de que la distribución de los datos tanto por encima como por debajo del punto medio sea mejor una que otra.

Si se extiende el análisis un poco más y se observa cómo se distribuyen los datos antes y después de la mediana, calculando por ejemplo el promedio de ambas muestras pre y post-mediana, la fórmula que tiene una mediana más baja (0.45 D), a priori la mejor fórmula, presenta las siguientes medias pre y post-mediana 0.32 D y 0.64 D respectivamente. Por otro lado, la fórmula que tenía una mediana de 0.49 D, a priori la peor de las dos, presenta las siguientes medias pre y post-mediana 0.29 D y 0.63 D respectivamente.

Ahora bien, ¿cuál de las dos fórmulas es mejor ahora? obviamente la que tiene medias por encima y debajo del punto medio más bajas porque esto implica que esos datos están más cerca de 0.00 el valor óptimo, es decir, la que peor mediana tiene, la de 0.49 D, por culpa de la ceguera de la mediana en los datos pre y post centro.

En resumen, la mediana como estadístico para evaluar la exactitud de una fórmula se descartará debido a la incertidumbre que plantea sobre cómo se distribuyen los datos a un lado u otro del punto medio. Es inadmisibles que una métrica que solo da información de la observación central de la distribución, se utilice para cualificar una fórmula o método de cálculo de potencia de lente intraocular, es más, que dicha clasificación se cimiente únicamente en esta métrica según algunos estudios.

Tras todo este argumentario, es fundamental homogeneizar los criterios de selección, consensuar las nomenclaturas y aplicar adecuadamente los test estadísticos. Es necesario introducir métricas diferentes a las utilizadas hasta ahora para reforzar el sistema de ordenación de fórmulas y métodos de cálculo, para no sufrir estas situaciones paradójicas en las que dependiendo del descriptor utilizado la clasificación y selección de las fórmulas cambie. El paradigma es necesario modificarlo para producir unos resultados repetibles, comparables y concordantes entre investigadores.

En la siguiente parte de este capítulo se explicará como abordar este cambio de paradigma para homogeneizar los resultados entre trabajos y sobre todo facilitar la selección de precisión y exactitud de una fórmula de cálculo de la potencia de IOL.

11.2.5 Tratamiento matemático de los errores absolutos

En general, no tener valores anómalos o *outliers* en una muestra de datos es una entelequia, y desde luego no poseerlos en una distribución relacionada con cálculos biométricos y resultados refractivos, hoy por hoy es una utopía, probablemente por la

variabilidad biológica no mensurable o por ciertos parámetros quirúrgicos sin control e incluso por el diseño teórico de algunas fórmulas, a saber.

Aunque el tratamiento de los *outliers* no es tarea fácil y lo más rápido sería eliminarlos puesto que son los valores que “contaminan” la muestra deformándola y modificando sus estimadores de centralidad y dispersión, esto conllevaría un coste sobre el proceso inferencial no paramétrico a posteriori. En todo caso y como se verá más adelante, una supresión racional de ciertos valores atípicos será necesaria para “depurar” la distribución y exprimir el máximo de información.

Por otro lado, un análisis exploratorio pormenorizado previo de los datos de que se disponen es preceptivo, porque en ocasiones alguno de los valores anómalos se corresponden simplemente a fallos de transcripción, lo que obligaría sin más a su eliminación.

Por otro lado, expresar un valor de centralidad en una distribución no normal, es arriesgado y probablemente no óptimo (174), lo que obliga a asumir ciertas condiciones teóricas que no se dan en las distribuciones de errores absolutos.

Por estas razones, se expondrán metodologías que actúan sobre la distribución de la población de errores absolutos para minimizar el impacto de los valores atípicos y sobre los estadísticos paramétricos de centralidad y dispersión, mejorando el rendimiento y precisión estadística.

En primer lugar se propondrá la transformación de datos como ya propuso Aristodemou et al (175) y posteriormente se introducirá algún estimador basado en estadística robusta (176). Aristodemou en su trabajo desechó la posibilidad de transformación de variables por posibles sesgos en la significación, alegando que una

transformación mediante raíz cúbica de los MAEs da el mismo peso para un cambio de PE de 0.00 D a 0.10 D que de 0.10 D a 0.60 D.

11.2.6 Transformación de la distribución de errores absolutos

Como ya se comentó en el apartado anterior la distribución de MAEs no se ajusta a una campana de Gauss, es una distribución exponencial con cola hacia la derecha, acumulándose la mayoría de valores positivos a la izquierda de la media. Mediante las técnicas de transformación de variable se puede “normalizar” la distribución, reducir la influencia de los valores atípicos así como mejorar la heterogeneidad de varianza. Existen multitud de tipos de transformación, entre otras, las más conocidas son las logarítmicas, la raíz cuadrada o cúbica y las inversas. Habiendo probado varios tipos, la más potente para el tipo de distribución de MAEs con valores muy pequeños y cercanos a cero, es la transformación raíz cuadrada (177) (Figura 107), donde cada valor individual de MAE corresponderá a su raíz cuadrada en la transformada. Tras la transformación se comprueba la normalidad de la distribución transformada aplicando los siguientes estadísticos: *Kolmogorov-Smirnov* con corrección de *Lilliefors*, *Shapiro-Wilk* y *Anderson-Darling*, como ya se relató anteriormente.

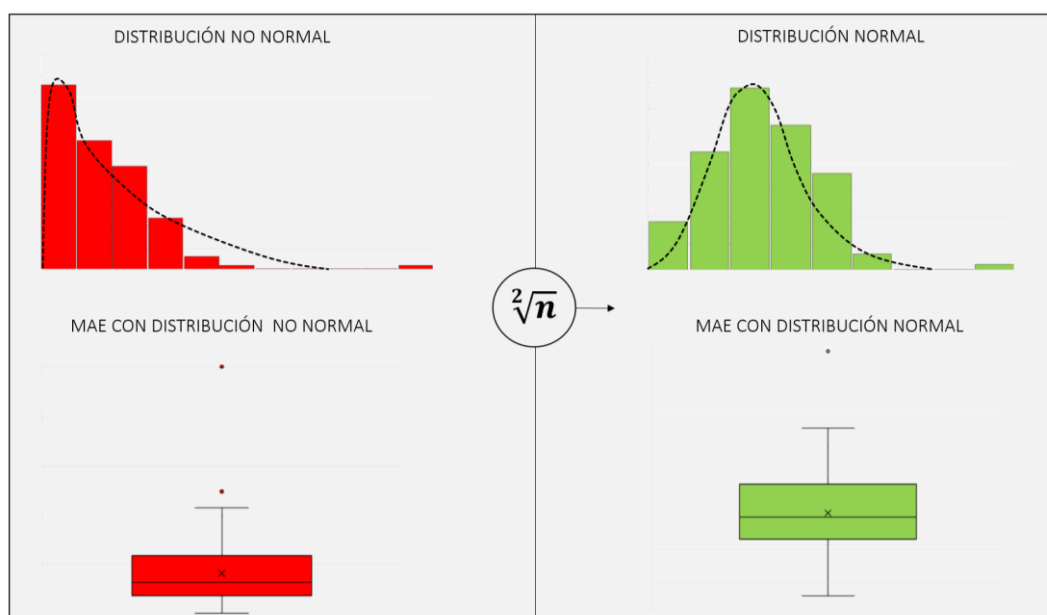


Figura 107. Normalización mediante raíz cuadrada de la distribución de errores absolutos.

Esta transformación permite “estirar” los valores cercanos a 0.00 D sin desplazar demasiado los próximos a 1.00 D haciendo que la curva cambie su *skewness* a positivo. Igualmente los valores altos de error absoluto que pudieran considerarse atípicos serán comprimidos mediante la transformación reduciendo su distancia a 0.00 D, lo que significa mayor control sobre las observaciones *outliers* (Figura 107). A su vez esto significa que se pueden emplear contrastes de hipótesis paramétricos para obtener el nivel de significación. Los test no paramétricos son más conservadores para alcanzar el nivel de significación estadística necesario para poder rechazar la hipótesis nula, de ahí que la potencia estadística de una prueba paramétrica sea mayor.

Posteriormente es necesario deshacer la transformación de la variable mediante la función inversa para la correcta interpretación de los resultados tras haber aplicado el test paramétrico correspondiente. Por tanto, al haber transformado con la raíz cuadrada a cada valor y haber obtenido el promedio, se debe aplicar a este su cuadrado. La anti-

transformación permite recuperar medias, medianas e intervalos de confianza sin problemas, no siendo así con la desviación estándar. Esto sucede porque la definición de desviación típica es la raíz cuadrada de la varianza:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}}$$

La transformación de cada error absoluto se hizo previamente con raíz cuadrada, de manera que al calcular la desviación de esa transformada, el numerador “deshace” dicha transformación, obteniéndose valores desvirtuados con la anti-transformación final. Posteriormente se explicará una vía alternativa a esto para determinar el error estándar.

11.2.7 Estadística robusta y centralidad de errores absolutos

Tras haber transformado la distribución se necesitan unas métricas robustas que “aguanten” todavía el envite de los posibles outliers remanentes tras la transformación.

La estadística robusta arroja metodologías que no se resienten por la influencia de valores atípicos u otras discrepancias en las asunciones del modelo, así que la mejor opción es quitarle peso a esas observaciones atípicas de los MAEs mediante estas técnicas robustas.

Existen los llamados M-estimadores o medidas robustas de localización (*Robust Location Measures*). Un M-estimador se define como *Maximun Likelihood Estimator* o estimador de máxima verosimilitud. Su objetivo es buscar un índice de localización a partir del conjunto de observaciones, ponderando a éstas en función de lo cerca o lejos

que se encuentren del centro de datos independientemente del tipo de distribución que sigan estos. El más conocido es el de Huber, aunque hay muchos más: onestep, Pitman, bponderado de Tukey, Hampel, Andrews, etc.

Para demostrar que el proceso de normalización por raíz cuadrada combinado con la inclusión de estadísticos robustos es suficientemente potente como para clasificar fórmulas, los M-estimadores se calcularán sobre la distribución no normal de MAEs para compararlos con los valores de medias, medias truncadas y medianas anti-transformados de la distribución normalizada de esos mismos MAEs. No obstante, por su complejidad matemática y dificultad de ejecución, no se propondrán como métricas finales de ordenación para las fórmulas de cálculo biométrico, solo como herramienta para la aprobación del proceso de transformación.

De manera que como métrica de centralidad (= exactitud) se utilizará la media recortada, truncada o acotada (*Trimmed Mean*) desechando un porcentaje concreto de valores extremos por ambas colas de la distribución de MAEs, su nomenclatura será TMAE (*Trimmed Mean Absolute Error*). La media recortada queda definida por la proporción de observaciones (α) que son excluidas desde cada extremo de la muestra ordenada. Una vez que se han eliminado los valores indicados de cada extremo, se calcula el promedio de los valores restantes. Si α es un múltiplo de $1/n$, se eliminan de cada extremo un número entero de valores (αn) y la media recortada es el promedio de los valores restantes.

El porcentaje de datos recortados suele oscilar entre un 5% y un 20% (178). Por ejemplo, para muestras extensas, recortar un 5% de los extremos de la muestra supone

poca pérdida de poder inferencial. Por ejemplo, para una muestra de 2000 ojos un 5% significa eliminar 100 datos, 50 por cada cola.

El problema surge cuando hay que tomar la decisión de qué porcentaje de muestra se recorta sin que por ello se afecte el poder inferencial, o cuál es el valor óptimo de recorte para determinar un valor de centralidad. En este sentido algunos autores abogan por el acercamiento empírico, es decir, recortar las colas progresivamente de 5% a 20% (178), para observar cuánto varía la media truncada hasta posicionarse en un valor relativamente estable.

Como la distribución de datos MAEs se ha normalizado provocando un mejor control sobre los valores anómalos, consecuentemente el porcentaje de recorte en el truncamiento de la media no deberá ser grande. La diferencia en estos casos entre recortar un 5 % y un 20% es escasa, esta actitud conservadora mantendrá poder inferencial. Por otro lado se comprobó que el promedio de los M-estimadores máximo y mínimo de los cuatro calculados sobre la distribución sin transformación, correspondía aproximadamente al valor de segundo nivel de la media truncada antitransformada, o lo que es lo mismo, al acotamiento del 5%, cuya sintaxis será TMAE5 (Figura 108).

	fórmula 1	fórmula 2	fórmula 3	fórmula 4	fórmula 5	fórmula 6	fórmula 7	fórmula 8	fórmula 9	fórmula 10	fórmula 11
MAE	0,630	0,539	0,620	0,539	0,505	0,497	0,553	0,516	0,520	0,522	0,526
ANTI-MAE	0,397	0,290	0,384	0,290	0,255	0,247	0,305	0,266	0,270	0,272	0,277
MedAE	0,664	0,522	0,632	0,539	0,480	0,490	0,539	0,500	0,500	0,520	0,529
ANTI-MedAE	0,441	0,273	0,399	0,290	0,230	0,240	0,290	0,250	0,250	0,270	0,280
TMAE5	0,630	0,533	0,619	0,534	0,500	0,492	0,550	0,514	0,516	0,519	0,524
ANTI-TMAE5	0,397	0,284	0,383	0,285	0,250	0,243	0,302	0,264	0,266	0,270	0,274
HUBER	0,439	0,304	0,426	0,319	0,265	0,267	0,326	0,290	0,284	0,303	0,305
ONESTEP	0,444	0,286	0,428	0,312	0,248	0,256	0,315	0,284	0,268	0,297	0,299
PITMAN	0,442	0,273	0,399	0,284	0,230	0,247	0,290	0,252	0,250	0,269	0,280
TUKEY BIPON.	0,413	0,256	0,378	0,269	0,212	0,223	0,269	0,233	0,225	0,256	0,261

Figura 108. Métricas robustas sobre la población normalizada de errores absolutos en el cálculo real de 11 fórmulas. Métricas paramétricas clásicas MAE y MedAE con sus respectivas anti-transformaciones. Métrica robusta TMAE5 correspondiente a la media acotada al 5% y su anti-transformación. En rojo están calculados cuatro M-estimadores sobre la distribución sin transformar para comparar con los datos normalizados (Huber, Onestep, Pitman y Tukey bponderado).

También se observó que en una distribución cuasi normal como la transformada de MAEs con raíz cuadrada, independientemente de la longitud axial y de la fórmula, el valor anti-transformado de mediana (ANTI-MedAE) es prácticamente idéntico al M-estimador de Pitman (179) (Figura 108) calculado sobre la muestra sin transformar, pero habiendo también muy pocas diferencias respecto a los otros estimadores de posición testados (Huber, Onestep y Tukey bponderado).

De esta manera, queda demostrado que la transformación de la distribución gamma a normal de la población de MAEs mediante raíz cuadrada, es correcta y suficientemente potente para la estimación de la precisión y exactitud de las fórmulas de cálculo biométrico.

11.2.8 Estadística robusta y dispersión de errores absolutos

Cómo se ha visto anteriormente, la desviación estándar es un estadístico muy sensible a los valores anómalos. Al transformar la distribución de datos, se ha conseguido ejercer un mayor control sobre estos datos atípicos, pero no obstante todavía pueden tener influencia en los estimadores de posición y dispersión, pudiendo perder eficiencia.

En cualquier caso, la desviación estándar no es anti-transformable como ya se explicó anteriormente, de manera que es preceptivo sustituir esta por un estimador de dispersión que caracterice la muestra de datos normalizados o en todo caso buscar una alternativa eficiente a la desviación típica.

Usualmente, en distribuciones sesgadas no normales se pueden emplear con garantías dos estadísticos robustos como estimadores de dispersión:

- Median Absolute Distance (MAD) atribuido a Gauss (180) que consiste en calcular la mediana de la diferencia en valor absoluto entre la observación y la mediana. Este estadístico tiene un *breakdown point* de 50%.
- Rango intercuartílico con un *breakdown point* del 25% representa la diferencia entre el tercer cuartil y el primero.

Estos métodos robustos se utilizan particularmente como estimadores de un parámetro de escala, y tienen las ventajas tanto de su robustez como de su eficiencia superior en series de datos contaminadas, a costa de una eficiencia inferior en series de datos limpias, especialmente de distribuciones como la distribución normal. Por lo que, en la normalización de la distribución de MAE, el uso de MAD podría verse afectado.

Para datos extraídos de una distribución normal, como es el caso que ocupa, MAD es solo un 37% eficiente comparado con la desviación estándar de la muestra (181). Además, MAD calcula una estadística simétrica sobre una estimación de ubicación (mediana), por lo que no contempla la asimetría estadística.

Precisamente por estas razones, Rousseeuw y Croux propusieron algunas alternativas a MAD, dos estimadores llamados S_n y Q_n , ambos más eficientes que MAD bajo una distribución gaussiana, 58% y 82% respectivamente. Ninguno de ellos requiere una estimación de localización como MAD, ya que se basan únicamente en diferencias entre valores. Sin embargo, se desecharon para esta tesis por coste computacional y complejidad de cálculo, ya que requieren cálculos iterativos y recurrentes.

Con todos estos argumentos, para encontrar un buen estimador de dispersión sobre una muestra transformada sabiendo que la desviación estándar sería correcta,

pero con el óbice de la anti transformación, la estrategia será encontrar una desviación típica suficientemente eficiente sin necesidad de recurrir a la función inversa.

Esta vía se basa en el trabajo de Jorgensen y Pedersen (182) en el que demuestran cómo transformar los errores estándar a la escala original tras haber normalizado una distribución mediante una sencilla fórmula. Para ello y sin reparar en demasiados formalismos matemáticos, en transformaciones que no sean lineales como la de raíz cuadrada, se puede aproximar cualquier función por una función lineal en la vecindad de un punto mediante la expansión de Taylor de primer orden. De esta manera se puede calcular la varianza y el valor esperado necesitándose únicamente la primera derivada de la función utilizada en la transformación.

Teniendo en cuenta que x_1, x_2, \dots, x_i son observaciones de una distribución de datos no normal, e y_1, y_2, \dots, y_i son datos normalizados y transformados mediante raíz cuadrada, la fórmula que derivan los autores para el cálculo de la desviación estándar es:

$$\tilde{\sigma}_x = 2\sqrt{\tilde{X}}\hat{\sigma}_y \quad (69)$$

Dónde $\tilde{\sigma}_x$ es una aproximación a la desviación estándar en la escala original sin transformación, \tilde{X} es la media en la escala original que se corresponde a Anti-MAE y $\hat{\sigma}_y$ es el error estándar en la escala transformada, al que no se le puede aplicar la función inversa para la anti-transformación como ya se explicó anteriormente. Por tanto, si por ejemplo se tiene que en la escala transformada la media de errores absolutos para una fórmula es 0.216 D (Anti-MAE) y su desviación estándar transformada es 0.217 D, aplicando la Ecuación (69) se tiene:

$$\tilde{\sigma}_x = 2\sqrt{0.216} * 0.217 = 0.202 D$$

Así que los datos de este caso concreto para esta fórmula se podrían expresar como Anti-MAE \pm SD: $0.216 \pm 0.202 D$.

La función raíz cuadrada es continua para todos los números no negativos y derivable para todos los números positivos menos para cero ya que la pendiente de la tangente ahí es infinita. Por esta razón, y como el método propuesto por Jorgensen y Pedersen se basa en el desarrollo de Taylor de primer orden, todos los valores de MAE que sean 0.00 D en la distribución sin normalizar no computarán para el cálculo del error estándar en la distribución transformada. De manera que se aconseja observar primero si existe algún dato que sea 0.00 D, en cuyo caso la transformación debería ser:

$$Y_i = \sqrt{X_i + k} \quad (70)$$

Donde k sería un escalar pequeño ($k > 0$) que $X_i + k \neq 0$. Por ejemplo, se podría probar 0.001.

11.2.9 Aprobación de las métricas robustas

Para ratificar la inclusión de los estadísticos robustos anteriormente descritos se seguirá utilizando el ejemplo simple de comparación de tres fórmulas con 10 observaciones (Figura 105). El *modus operandi* ante una población de errores absolutos en primer lugar sería aplicar la transformación para normalizar la distribución (Figura 109. Errores absolutos de tres fórmulas sin transformar con sus métricas paramétricas clásicas (MAE, SD y MedAE). A la derecha los mismos errores absolutos transformados mediante

raíz cuadrada con las nuevas métricas (Anti-MAE, SD, Anti-MedAE y Anti-TMAE20). Figura 109). Aquí se pueden observar los datos en crudo y transformados, los boxplot antes y después de la transformación, así como las métricas correspondientes al enfoque clásico: MAE, SD y MedAE y las métricas robustas: Anti-MAE, SD, Anti-MedAE y Anti-TMAE20. Obsérvese que la población de datos ($n = 10$) tiende a la normalidad tras la transformación además de haber eliminado un outlier de la fórmula 3 y haberlo suavizado en la fórmula 1.

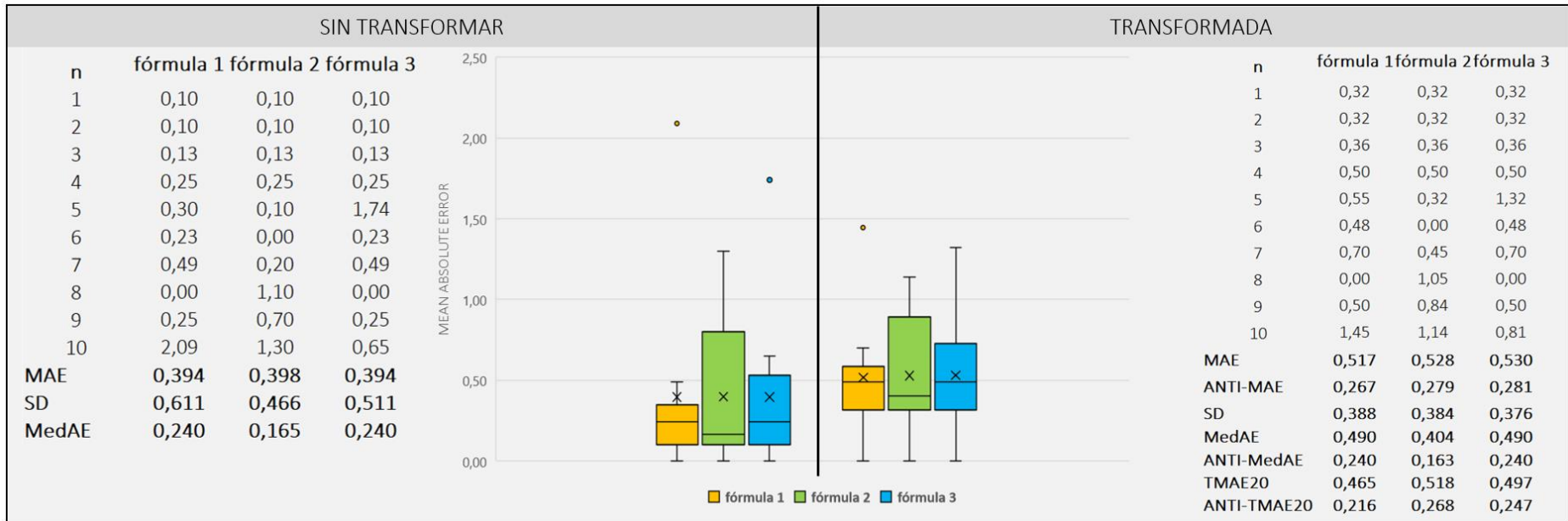


Figura 109. Errores absolutos de tres fórmulas sin transformar con sus métricas paramétricas clásicas (MAE, SD y MedAE). A la derecha los mismos errores absolutos transformados mediante raíz cuadrada con las nuevas métricas (Anti-MAE, SD, Anti-MedAE y Anti-TMAE20).

Dado el tamaño muestral del ejemplo, $n = 10$, el recorte de la media fue del 20% dado que un 5% no hubiera eliminado por las colas ningún valor.

Al aplicar la anti-transformación el promedio de errores absolutos se perfila a favor de la fórmula 1 (0.267 D) frente a 0.281 D y 0.279 D de la fórmula 3 y 2 respectivamente. Sin embargo antes de la transformación de variable, con una distribución gamma, había coincidencia en los MAEs de la fórmula 1 y 3 (0.394 D).

Se aplica sobre este ejemplo la Trimmed Mean Absolute Error (TMAE) que recortará en ambos extremos de la distribución normalizada un 20% de los datos, en este caso un valor por cola. La fórmula 1 vuelve a ser la mejor ya que al eliminar el único outlier en la cola derecha (1.45 D) y el mismo valor absoluto por la izquierda (0.00 D), la media de 0.216 D es más baja que para la fórmula 3 (0.247 D), ya que en este caso la eliminación será de 1.32 D por la derecha y de 0.00 D por la izquierda.

Este ejemplo es muy representativo de lo importante que es la media recortada, ya que si se pretende clasificar únicamente por Anti-MAE y SD el ranking final quedaría establecido como fórmula 1, 2 y 3 con 0.267 ± 0.388 D, 0.279 ± 0.384 D y 0.281 ± 0.376 D respectivamente. Esta clasificación es errónea por culpa de los valores atípicos a pesar de la transformación, de ahí que sea necesaria la media acotada.

Como se puede observar en la Figura 110, las columnas de la derecha representan el contraste de errores absolutos dos a dos, comprobando que habiendo suprimido un valor por ambas colas (valores en rojo), la fórmula 2 a pesar de tener un mejor MAE que la fórmula 3, tiene mayor diferencia de errores respecto a la fórmula 1 y a la 3, 0.64 D y 0.38 D respectivamente. Dando como resultado final clasificadorio un mejor comportamiento para la fórmula 1, luego la 3 y por último la 2.

			fórmula 1 vs fórmula 2		fórmula 1 vs fórmula 3		fórmula 2 vs fórmula 3		
fórmula 1	fórmula 2	fórmula 3	error fórmula 1	error fórmula 2	error fórmula 1	error fórmula 3	error fórmula 2	error fórmula 3	
0,00	0,00	0,00							
0,32	0,32	0,32	0,00	0,00	0,00	0,00	0,00	0,00	
0,32	0,32	0,32	0,00	0,00	0,00	0,00	0,00	0,00	
0,36	0,32	0,36	0,04	0,00	0,00	0,00	0,00	0,04	
0,48	0,36	0,48	0,12	0,00	0,00	0,00	0,00	0,12	
0,50	0,45	0,50	0,05	0,00	0,00	0,00	0,00	0,05	
0,50	0,50	0,50	0,00	0,00	0,00	0,00	0,00	0,00	
0,55	0,84	0,70	0,00	0,29	0,00	0,15	0,14	0,00	
0,70	1,05	0,81	0,00	0,35	0,00	0,11	0,24	0,00	
1,45	1,14	1,32							
ANTI-TMAE20									
ANTI-TMAE20	0,216	0,268	0,247	0,21	0,64	0,00	0,26	0,38	0,22

Figura 110. Contraste de errores absolutos dos a dos.

En síntesis, tras haber normalizado la distribución mediante transformación de variable con raíz cuadrada y haber aplicado estadísticos robustos, si se aplica la siguiente nomenclatura Anti-MAE \pm SD (Anti-MedAE), la fórmula 1 es mejor que las demás con 0.267 ± 0.388 D (0.240), seguida de la 3 con 0.281 ± 0.376 D (0.240), y de la 2 con 0.279 ± 0.384 D (0.163). Se podría perfectamente expresar la desviación estándar, calculada según el método de Jorgensen, respecto a la media truncada y por tanto la nomenclatura sería Anti-TMAE20 \pm SD20 (Anti-MedAE20): fórmula 1 con 0.216 ± 0.122 D (0.240) , fórmula 3 con 0.247 ± 0.177 D (0.240) y fórmula 2 con 0.268 ± 0.286 D (0.160).

En este ejemplo de solo 10 datos es fácil analizar los datos uno a uno, pero cuando la muestra es amplia la situación se torna inmanejable, no se tiene la capacidad de escrutar una a una las observaciones, y la facultad de inferenciar sobre muestras amplias no es nada intuitiva, de ahí que la utilización de estos nuevos estimadores ayude en la decisión de ordenación final de las fórmulas.

11.2.10 Aplicación de las nuevas métricas sobre un ejemplo real

A continuación se propone un ejemplo real (Figura 111) en el que se comparan 12 fórmulas y métodos de cálculo de la potencia de la lente intraocular sobre una muestra de 481 ojos con longitudes axiales entre 19.33 mm y 33.00 mm. Si atendemos a los descriptores clásicos (MAE, MedAE y SD-MAE) se genera controversia y confusión en la ordenación, ya que dependiendo de cuál se escoja una fórmula es mejor que otra.

Sin embargo, al transformar la población de MAEs y normalizarla, se favorece el control sobre observaciones anómalas, y al utilizar estadísticos robustos la perspectiva clasificatoria difiere.

En primer lugar, atendiendo a la forma clásica de ordenar, si se selecciona MAE como métrica primordial la fórmula 6 es la mejor con 0.299 ± 0.257 D siempre y cuando se utilice precisión milesimal, de lo contrario las fórmulas 5, 6 y 12 coincidirían en promedios de errores absolutos. Ahora bien, si la métrica fuese únicamente SD la mejor fórmula sería la fórmula 12 con ± 0.240 D, ya que es la que menos dispersa los datos respecto a la media. Pero por otra parte, si se escoge la mediana (clásicamente se sigue haciendo) se tendría un conflicto entre la fórmula 5 y la 12 por poseer valores idénticos con precisión a la milésima, ± 0.230 D.

Con la propuesta y tras la normalización de los datos, el Anti-MAE de la fórmula 6 es el mejor con 0.247 ± 0.226 (0.240) D. Para la otra métrica de centrado, la media truncada al 5%, la mejor fórmula con 0.243 ± 0.034 (0.240) D sigue siendo la fórmula 6. En segundo lugar la fórmula 12 con 0.249 ± 0.196 (0.230) D, seguida de la fórmula 5 con 0.250 ± 0.194 (0.230) D.

MÉTRICAS CLÁSICAS SIN TRANSFORMACIÓN												
	fórmula 1	fórmula 2	fórmula 3	fórmula 4	fórmula 5	fórmula 6	fórmula 7	fórmula 8	fórmula 9	fórmula 10	fórmula 11	fórmula 12
MAE	0,459	0,348	0,447	0,350	0,303	0,299	0,358	0,316	0,325	0,330	0,330	0,300
SD	0,315	0,308	0,319	0,295	0,249	0,257	0,272	0,248	0,276	0,274	0,261	0,240
MedAE	0,441	0,273	0,399	0,290	0,230	0,240	0,290	0,250	0,250	0,270	0,280	0,230
MÉTRICAS NUEVAS CON TRANSFORMACIÓN												
	fórmula 1	fórmula 2	fórmula 3	fórmula 4	fórmula 5	fórmula 6	fórmula 7	fórmula 8	fórmula 9	fórmula 10	fórmula 11	fórmula 12
MAE	0,630	0,539	0,620	0,539	0,505	0,497	0,553	0,516	0,520	0,522	0,526	0,501
ANTI-MAE	0,397	0,290	0,384	0,290	0,255	0,247	0,305	0,266	0,270	0,272	0,277	0,251
SD	0,314	0,259	0,311	0,262	0,220	0,226	0,253	0,231	0,245	0,249	0,243	0,222
MedAE	0,664	0,522	0,632	0,539	0,480	0,490	0,539	0,500	0,500	0,520	0,529	0,480
ANTI-MedAE	0,441	0,273	0,399	0,290	0,230	0,240	0,290	0,250	0,250	0,270	0,280	0,230
TMAE5	0,630	0,533	0,619	0,534	0,500	0,492	0,550	0,514	0,516	0,519	0,524	0,499
ANTI-TMAE5	0,397	0,284	0,383	0,285	0,250	0,243	0,302	0,264	0,266	0,270	0,274	0,249
SD5	0,284	0,220	0,284	0,231	0,194	0,195	0,225	0,206	0,215	0,217	0,215	0,196

Figura 111. Estadística clásica contra estadística propuesta en la comparativa real de 12 fórmulas de cálculo biométrico para 482 ojos con longitudes axiales entre 19.33 mm y 33.00 mm.

En cuanto a la aplicación de test estadísticos para contrastar la significación estadística, se optará obviamente por contrastes paramétricos ya que se ha normalizado la muestra transformándola. Si se comparan dos fórmulas una *t* de *Student* es perfecta, y para comparaciones múltiples dos a dos, se aconseja realizar un ANOVA de una vía con pruebas *post hoc* mediante ajuste de *Holm-Bonferroni* si hay menos de 6 fórmulas y ajuste de *Tukey HSD* para más de 6 comparaciones.

11.2.11 Porcentaje de errores predictivos pertenecientes a un intervalo refractivo

En la literatura oftalmológica publicada está plenamente aceptado reportar en los informes postoperatorios los porcentajes de errores refractivos predictivos incluidos en ciertos intervalos, habitualmente: ± 0.25 D, ± 0.50 D, ± 0.75 D y ± 1.00 D. Puesto que la precisión con la que se manufacturan la mayoría de las lentes intraoculares es de media dioptría (129), se asume que ± 0.50 D es un buen descriptor de exactitud para una fórmula (18), obviamente a priori cuanto más porcentaje de ojos haya incluidos en este rango dióptrico, mejor será la fórmula en cuestión, más exacta. Pero en cualquier caso, la mejor fórmula será la que introduzca más EP en el rango de ± 0.25 D, de hecho, hay algún estudio (183) en el que se pone de manifiesto esto: si la precisión de manufacturación es mejor, la precisión postoperatoria también mejora.

En ocasiones hay fórmulas que en un rango determinado de longitud axial colocan un porcentaje de ojos en el rango ± 0.25 D más alto que otras, pero sin embargo tienen un MAE peor. ¿Esto significa que es mejor fórmula la que introduce mayor porcentaje de ojos en el intervalo ± 0.25 D o la que tiene mejor MAE?

En el mismo ejemplo comparativo de 12 fórmulas (Figura 111), hilo conductor de este argumentario, ocurre esto mismo. La fórmula 5 introduce un 57.89% de ojos dentro del rango ± 0.25 D contra un 53.51% de la fórmula 6. Sin embargo la fórmula 5 tiene un anti-MAE de 0.255 D contra un 0.247 D de la fórmula 6. Este hecho resulta engañoso pero en realidad es perfectamente compatible.

Imagínese que de una muestra de 100 ojos una fórmula introduce el 50% dentro del rango ± 0.25 D y otra el 60%. No obstante los 40 valores de la segunda fórmula que no están dentro del rango estudiado son peores, más alejados de cero que los 50 valores de la primera fórmula que están fuera del rango. Esto implica que el promedio (anti-MAE) de todos los valores será mejor en la primera fórmula a pesar de que introduce menos valores en ± 0.25 D.

Esta circunstancia obliga a relacionar ambos parámetros de alguna manera para obtener otra métrica definitiva que permita distinguir estas situaciones correctamente. A continuación se explicará la inclusión de PE en intervalos modificados respecto a lo ya conocido como introducción a un nuevo índice que se propondrá más tarde para relacionar anti-MAE y porcentaje de PE.

En la literatura oftalmológica publicada hay una ausencia manifiesta de información de cómo se tienen que gestionar estas situaciones, cuando se afronta un análisis comparativo con varios modelos de lentes intraoculares, con distintas constantes de fabricación. Es importante incidir en esta cuestión, en los trabajos comparativos en los que se incluyan diferentes lentes y por tanto diferentes constantes A, los porcentajes de PE que se analizarán cuáles son, los tipificados para centrar todas las distribuciones de

datos con media 0.00, o por el contrario, se utilizarán los PE obtenidos en bruto sin corrección.

Tal y como se expresó anteriormente, la tipificación de una variable permite hacer comparaciones entre diferentes observaciones con distintas escalas. Teniendo en cuenta que el uso de diferentes constantes influiría en el resultado refractivo postoperatorio, la normalización contribuye a que la comparación se haga en las mismas condiciones cuando hay diferentes constantes A. De esto se desprende que de ahora en adelante se usarán los datos de PE tipificados o estandarizados.

11.2.12 Nuevo sistema de clasificación por porcentajes de rangos

¿Es mejor un resultado refractivo postoperatorio negativo o positivo? En líneas generales cuando se trata de cirugía implantacional con lentes monofocales, el cálculo de la potencia emetropizante y su resultado refractivo postoperatorio, es preferible que se encuentre en cero o en todo caso en valores ligeramente negativos (miópicos) a valores hipermetrópicos. De esta manera, la posibilidad de no empeorar la visión de cerca y mejorar la intermedia es mayor además de no afectarse excesivamente la visión de lejos. Valores ligeramente negativos, son los que se encuentran entre 0.00 D y -0.50 D de equivalente esférico, ya que la manufactura habitual de IOL es con precisión de media dioptría.

Sin embargo, con el advenimiento de la variedad de perfiles difractivos de los que se disponen hoy en día en los diseños de lentes multifocales intraoculares y de foco

extendido, no siempre es mejor un PE negativo, a veces leves residuos positivos capacitan visualmente al paciente más y producen mejor rendimiento óptico de la lente.

En cualquier caso, si se quiere hacer independiente el rango de PE estudiado del tipo y diseño de la lente o lentes implantadas, se deben escindir los rangos por tipo ametrópico, es decir, analizar el porcentaje de ojos incluidos en el rango de -0.25 D a 0.00 D y en el rango de 0.00 a +0.25 D como intervalos diferentes, dónde el 0.00 esté incluido por ejemplo en el intervalo negativo.

De esto se deduce que contabilizar resultados refractivos de PE positivos y negativos indistintamente dentro de los rangos, podría afectar al potencial que se busca en una fórmula dependiendo del tipo de lentes que se incluyan en el estudio.

Imagínese que se tienen 100 ojos de muestra con dos fórmulas sometidas a estudio que introducen en el rango de ± 0.25 D el 50% de los ojos. Ahora bien, de esos 50 ojos introducidos en el rango estudiado, 45 son valores de PE positivos para la primera fórmula, es decir, la fórmula hipermetropizó hipocorrigiendo la potencia de la lente emetropizante. Los otros 5 valores son negativos. Sin embargo, la segunda fórmula introduce 50 valores negativos en el intervalo estudiado. ¿Qué fórmula es mejor? La que introduce 5 valores negativos o la que introduce 50 PE miópicos en el rango ± 0.25 D. En términos absolutos, no importa el signo si solo se valora la exactitud. Pero en términos cualitativos, dependiendo del tipo de lente o lentes sometidas a estudio si afectará seleccionar un rango u otro.

Se propone una nueva aproximación basada en el porcentaje de ojos incluidos en los distintos rangos estudiados reportando el promedio de PE \pm SD en esos rangos concretos. Se aconseja por practicidad denominar a los MPE dentro de cada rango como

MPE-0.25 para el rango 0.00 a -0.25 D y MPE+0.25 para el rango 0.01 a +0.25 D. La misma nomenclatura para las siguientes cotas de estudio: MPE-0.50 y MPE+0.50, ect.

Con el método habitual clásico (rango positivo-negativo), se representa únicamente la proporción de precisión de la fórmula en el rango estudiado pero no la exactitud, es decir cuántos ojos están incluidos en el intervalo pero sin conocer valores de centralidad y dispersión, lo que puede restar potencia a la métrica. Sería conveniente saber dentro del rango estudiado cómo se colocan y distribuyen los valores respecto a 0.00 D como valor ideal.

Por ejemplo, en la competencia de dos fórmulas que ambas introducen un 80% de los ojos en el rango ± 0.50 D, la primera distribuye prácticamente todos los valores en el entorno de 0.00 D, sin embargo la segunda desplaza todos sus resultados dentro del rango estudiado hacia el +0.25 D. Aunque aparentemente ambas tienen la misma precisión colocando el mismo número de valores en el mismo rango, no tienen la misma exactitud, la primera es más exacta y por tanto mejor fórmula.

Abordando esto de la manera clásica, únicamente se reportaría el 80% de PE incluidos en ± 0.50 D. Sin embargo con el método propuesto (rango 0.00-negativo y 0.01-positivo) se exploran ambas características, exactitud y precisión, puesto que se introducen los conceptos de media y desviación típica en un rango concreto de estudio además del porcentaje de ojos introducidos en dichos intervalos.

En la Figura 112 se puede observar la comparación de 4 fórmulas en el rango de ± 0.25 D. Según el análisis clásico, se observa que la fórmula 2 es la que más ojos incluye en ese intervalo, un 59.65% seguida de la fórmula 1 con 57.89% y de la fórmula 3 con un

54.39%. Por tanto, en un primer análisis y atendiendo al abordaje clásico, se podría aseverar que la fórmula 2 es la más precisa de las cuatro en la comparativa.

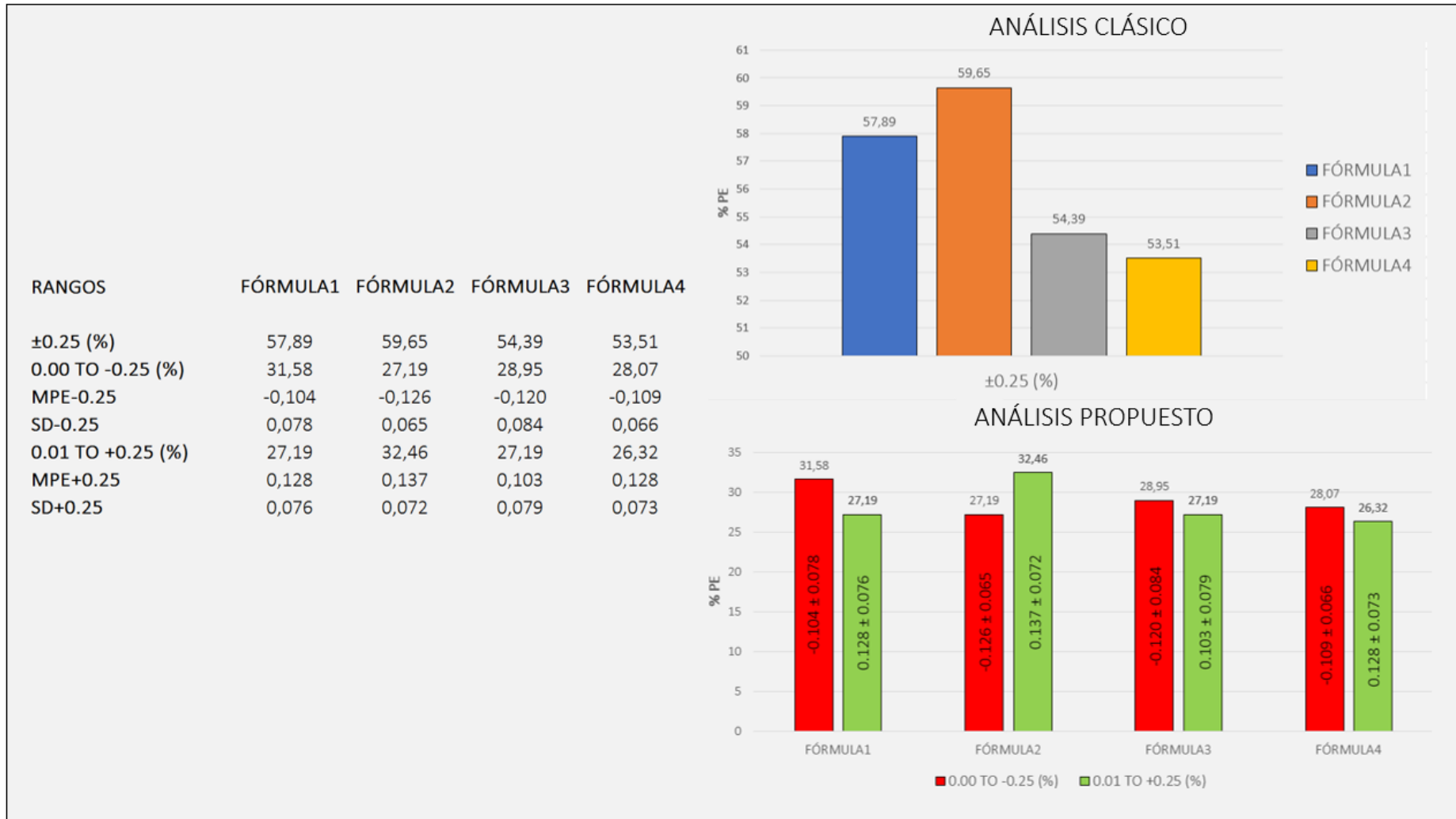


Figura 112. Representación gráfica y analítica de rangos dióptricos -0.25/0.00 y 0.01/+0.25 estudiados en la comparación de diferentes fórmulas de cálculo biométrico.

¿Pero es la más exacta? Dependerá del intervalo que se revise, disociando el rango $\pm 0.25 D$ en dos con 0.00 incluido en el rango negativo, se puede observar que la fórmula 1 introduce más valores negativos (31.58%) en el rango 0.00 a $-0.25 D$ que la fórmula 2 (27.19%) y con una media más baja: $-0.104 \pm 0.078 D$ versus $-0.126 \pm 0.065 D$. Esto implica que en este rango concreto la fórmula 1 es más exacta y más precisa que la fórmula 2. Sin embargo la fórmula 1 en valores positivos es más exacta pero menos precisa: $0.128 \pm 0.076 D$ versus $0.137 \pm 0.072 D$.

Esta manera de expresar los resultados permite profundizar más sobre la precisión y exactitud de las fórmulas, además de poder relacionar un intervalo con un tipo de lente concreto. Por supuesto, no son análisis excluyentes, son perfectamente compatibles y pueden cohabitar en los reportes científicos de estudios comparativos de fórmulas. Se puede decir que una fórmula incluye en el intervalo $\pm 0.25 D$ un 67% y al mismo tiempo expresar en los rangos disociados de este intervalo los porcentajes incluidos con sus medias y desviaciones estándar, haciendo el análisis más extensivo y más preciso si cabe.

Para el análisis estadístico de contraste de medias de PE de los rangos disociados, teniendo en cuenta que no se distribuyen normalmente puesto que se trata solo de valores negativos o positivos, se aplicará un test no paramétrico como el de *Wilcoxon*. Sin embargo para el porcentaje de PE incluidos en el intervalo estudiado se aplicará un test de *Cochran Q* y análisis *post hoc* con el test de *McNemar*.

En la siguiente sección se explicará la manera de relacionar el MAE y el porcentaje de PE introducidos en ciertos intervalos refractivos mediante un descriptor inédito al que llamaremos índice de fórmula o *Index Of Formula* (IOF) en inglés. IOF

expresará que fórmula o método se comporta mejor en ciertos rangos dióptricos. IOF servirá para determinar cuál es la fórmula que más valores negativos o positivos mete en los rangos dióptricos disociados, además de ser la más exacta en todos los valores absolutos de PE.

11.2.13 Nuevo índice IOF: Index Of Formula

En este tesis se propone un nuevo índice denominado *Index Of Formula* (IOF). Es un valor positivo, adimensional y oscila entre 0 y 1, siendo 1 mejor que 0. Es conveniente presentarlo con precisión de tres decimales para la distinción entre fórmulas con rendimientos similares.

IOF corresponde a la media entre los ratios de las dos métricas implicadas, ANTI-MAE y MPE en los rangos refractivos disociados. La expresión matemática general es:

$$IOF(n) = \frac{1}{2} \left[\left(\frac{1 - Anti_MAE}{1 - \left| \frac{P_1}{100} * MPE_1 \right|} \right) + \left(\frac{1 - Anti_MAE}{1 - \left(\frac{P_2}{100} * MPE_2 \right)} \right) \right] \quad (71)$$

Donde:

IOF(n) es el índice de fórmula dónde *n* hace referencia a si está calculado para un intervalo refractivo u otro, pudiendo *n* adquirir valores múltiplos de 0.25 en función del rango de cálculo. Esto es, IOF25 corresponderá al índice de fórmula calculado para el rango de 0.25 D, IOF50 para el de 0.50 D y así sucesivamente. Probablemente tenga interés solo el cálculo en esos dos rangos, puesto que son los más cercanos a 0.00.

Anti_MAE es el promedio de errores absolutos anti-transformados.

P₁ es el porcentaje de PE negativos introducidos en los intervalos refractivos.

P₂ es el porcentaje de PE positivos introducidos en los intervalos refractivos.

MPE₁ es el promedio de errores predictivos negativos correspondientes en dichos intervalos.

MPE₂ es el promedio de errores predictivos positivos correspondientes en dichos intervalos. Nótese que el operador no está incluido en valor absoluto como **MPE₁** por tratarse ya de valores positivos.

Siguiendo el ejemplo mostrado en la Figura 15, se tiene para la fórmula 1 sabiendo que su Anti_MAE es 0.242 D:

$$IOF25 = \frac{1}{2} \left[\left(\frac{1 - 0,242}{1 - \left| \frac{31,58}{100} * (-0,104) \right|} \right) + \left(\frac{1 - 0,242}{1 - \left(\frac{27,19}{100} * 0,128 \right)} \right) \right]$$

$$IOF25 = 0,785$$

Y para la fórmula 2 sabiendo que su Anti_MAE es 0.216 D:

$$IOF25 = \frac{1}{2} \left[\left(\frac{1 - 0,216}{1 - \left| \frac{27,19}{100} * (-0,126) \right|} \right) + \left(\frac{1 - 0,216}{1 - \left(\frac{32,46}{100} * 0,137 \right)} \right) \right]$$

$$IOF25 = 0,816$$

La fórmula 2 tiene IOF25 0.816 versus 0.785 de la fórmula 1, lo que implica que la fórmula 2 es mejor. Para el rango ± 0.50 D se procedería del mismo modo pero

incluyendo en MPE_1 y MPE_2 los promedios de los rangos refractivos correspondientes, al igual que sus ponderadores P_1 y P_2 respectivamente.

La media de los dos índices (IOFm) es un buen descriptor de exactitud y precisión de una fórmula de cálculo

Cuanto más alto sea IOF mejor será la fórmula en cuestión, debido a que por un lado Anti-MAE tiende a cero y el numerador estará más cerca de 1 y por otro MPE con su ponderación también baja estando el denominador más cerca de 1, lo que implica que la fórmula es más robusta y de mejor desempeño en la predicción de potencia de IOL.

El producto en valor absoluto del denominador representa la ponderación del peso específico que arrastra MPE en esta relación. El denominador siempre será superior al numerador, ya que en el hipotético caso de que el 100% de los datos estuvieran entre 0.00 D y -0.25 D, lo que implicaría una muy buena fórmula, el producto de MPE_1 por P_1 sería MPE_1 , y $Anti_MAE$ y MPE_1 tendrían el mismo valor, lo que resultaría en un IOF de 1.0, la fórmula perfecta.

En cuanto al análisis estadístico aplicable a IOF, teniendo en cuenta que el ratio se puede tratar como la proporción de éxito de la fórmula, se aplicará el test de McNemar al tratarse de muestras pareadas.

Para concluir este capítulo se propone un estudio comparativo real en el que se analizan 11 fórmulas habiendo implantado 3 lentes intraoculares distintas en ojos con longitudes axiales por encima de 25.00 mm (Figura 113). Si se atiende a la clasificación clásica y ordenando las fórmulas por MAE, la primera fórmula es la 6 con 0.263 D, seguida de la 8 con 0.268 D y de la fórmula 5 con 0.271 D. Si se clasificara por SD la ordenación cambia y por MedAE también es diferente.

	FÓRMULA 1	FÓRMULA 2	FÓRMULA 3	FÓRMULA 4	FÓRMULA 5	FÓRMULA 6	FÓRMULA 7	FÓRMULA 8	FÓRMULA 9	FÓRMULA 10	FÓRMULA 11
MAE	0,414	0,334	0,355	0,309	0,271	0,263	0,290	0,268	0,295	0,283	0,272
SD	0,337	0,302	0,297	0,336	0,244	0,236	0,240	0,217	0,275	0,233	0,218
MedAE	0,323	0,271	0,285	0,230	0,210	0,240	0,219	0,225	0,216	0,242	0,220

Figura 113. Métricas clásicas aplicadas a 11 fórmulas en una muestra con AXL < 25.00 mm.

Sin embargo, analizando la comparativa desde el punto de vista propuesto en esta tesis, se podría aseverar sin paliativos que la primera fórmula es la 6 con un Anti-MAE de 0.216 D, un Anti-TMAE5 de 0.211 D y una SD de 0.202 D. Por otro lado si se pone atención sobre los índices de fórmulas, tanto IOF25 como IOF50, la fórmula 6 es la mejor con índices de 0.809 y 0.865 respectivamente. De manera que los resultados obtenidos son congruentes y sin disparidades.

Para discernir el segundo puesto en el ranking, las fórmulas 5, 8 y 11 tienen un rendimiento muy similar, sin embargo IOF50 decanta la balanza a favor de la fórmula 8 con un índice de 0.848. Tiene sentido, ya que aunque la fórmula 11 presente un Anti-MAE 0.01 D mejor, incluye aproximadamente un 1% menos de ojos en el rango ± 0.50 D. Además la fórmula 8 dispersa menos los datos que la 11 con una SD de ± 0.190 D y una SD de ± 0.154 D.

Por último, el tercer puesto de la clasificación en este estudio sería para la fórmula 11, de rendimiento muy parecido como se ha visto anteriormente a la fórmula 8.

CATEGORIAS	FÓRMULA 1	FÓRMULA 2	FÓRMULA 3	FÓRMULA 4	FÓRMULA 5	FÓRMULA 6	FÓRMULA 7	FÓRMULA 8	FÓRMULA 9	FÓRMULA 10	FÓRMULA 11
ANTI-MAE	0,353	0,277	0,297	0,242	0,229	0,216	0,242	0,229	0,244	0,238	0,228
ANTI-MedAE	0,323	0,271	0,285	0,230	0,210	0,240	0,219	0,225	0,216	0,242	0,220
ANTI-TMAE5	0,347	0,271	0,292	0,233	0,224	0,211	0,239	0,226	0,238	0,235	0,226
SD	0,295	0,251	0,263	0,256	0,197	0,202	0,216	0,190	0,223	0,207	0,202
SD5	0,243	0,198	0,215	0,198	0,156	0,162	0,184	0,154	0,174	0,170	0,171
±0.25 (%)	37,72	47,37	42,11	57,89	59,65	54,39	53,51	56,14	52,63	55,26	56,14
0.00 TO -0.25 (%)	19,30	20,18	21,93	31,58	27,19	28,95	28,07	32,46	24,56	28,95	31,58
MPE(-0.25)	-0,155	-0,119	-0,133	-0,104	-0,126	-0,120	-0,109	-0,117	-0,107	-0,119	-0,119
SD (-0.25)	0,066	0,068	0,087	0,078	0,065	0,084	0,066	0,071	0,064	0,075	0,079
0.00 TO +0.25 (%)	18,42	28,07	21,93	27,19	32,46	27,19	26,32	25,44	29,82	27,19	26,32
MPE (+0.25)	0,116	0,119	0,108	0,128	0,137	0,103	0,128	0,130	0,120	0,130	0,122
SD (+0.25)	0,078	0,074	0,078	0,076	0,072	0,079	0,073	0,068	0,071	0,077	0,075
±0.50 (%)	64,91	80,70	71,93	82,46	85,96	89,47	81,58	86,84	83,33	86,84	85,96
0.00 TO -0.50 (%)	32,46	33,33	36,84	42,11	41,23	45,61	41,23	47,37	40,35	44,74	50,00
MPE (-0.50)	-0,223	-0,213	-0,213	-0,174	-0,198	-0,210	-0,188	-0,198	-0,210	-0,217	-0,210
SD (-0.50)	0,107	0,136	0,128	0,143	0,120	0,145	0,136	0,139	0,145	0,153	0,142
0.00 TO +0.50 (%)	32,46	48,25	36,84	41,23	44,74	45,61	41,23	41,23	44,74	42,98	37,72
MPE (+0.50)	0,229	0,227	0,212	0,213	0,200	0,203	0,204	0,212	0,189	0,212	0,188
SD (+0.50)	0,149	0,147	0,149	0,140	0,125	0,143	0,124	0,121	0,117	0,134	0,123
IOF25	0,664	0,744	0,722	0,784	0,802	0,809	0,783	0,799	0,780	0,789	0,800
IOF50	0,698	0,795	0,763	0,824	0,843	0,865	0,824	0,848	0,825	0,841	0,847

Figura 114. Comparativa de 11 fórmulas mediante las métricas propuestas en esta tesis. Se muestran todos los pasos y medidas necesarias para obtener IOF.

11.2.14 Esquema de actuación para estudios comparativos entre fórmulas

El objetivo de este capítulo es encontrar unos criterios diferentes a los que ya existen en la comparación de estudios sobre precisión de fórmulas de cálculo de la potencia de la lente intraocular. Hasta la fecha tal y como se ha demostrado en el presente trabajo, hay controversia en la aplicación de test estadísticos, en la utilización de unas métricas u otras para establecer una ordenación en las fórmulas sometidas a estudio e incluso en cómo se denominan dichos descriptores.

En este trabajo se aborda una nueva manera de analizar los resultados refractivos obtenidos en la cirugía de cristalino cuando se calcula la potencia de la lente intraocular que se va a implantar y estos se someten a una comparativa con diferentes métodos y/o fórmulas.

Hasta ahora, no se habían establecido unos criterios estándar con los que hacer este tipo de análisis. Ha habido intentos mediante publicaciones en editoriales como ya se ha reseñado anteriormente, pero no un trabajo dedicado exclusivamente a esto.

La pretensión del presente trabajo es encontrar un camino mediante el cual se organiza el trabajo paso a paso, con notaciones claras y concisas, con métricas distintas más robustas y con una nueva metodología de trabajo que facilite el trabajo comparativo entre fórmulas.

Otro factor importante a tener en cuenta es la precisión de comparación, mientras que hasta ahora se han utilizado centésimas. Con la aparición de nuevas fórmulas y su mejora en la exactitud, en muchas ocasiones hay coincidencias en las métricas empleadas y no se puede esclarecer la ordenación. Por tanto, si utilizamos

precisión de milésimas será más fácil establecer el ranking cuando comparemos los distintos descriptores.

Las distribuciones de PE no son siempre normales, dependerá de la fórmula y de la tipología de ojo explorada, por tanto el contraste estadístico a aplicar será distinto. Se aconseja utilizar tres test de normalidad: *Kolmogorov-Smirnov* con corrección de *Lilliefors*, *Shapiro-Wilk* y *Anderson-Darling*. Si se confirma la normalidad de la distribución de los datos, se aplicará una *t* de Student, de lo contrario un test de Wilcoxon.

Si se introducen en el estudio muchos tipos de lentes intraoculares con diferentes constantes, será preceptivo centrar la distribución de MPE en 0.00 D tal y como se ha explicado anteriormente, lo que se llama tipificación. De esta manera evitamos los posibles errores derivados de la utilización de diferentes constantes.

Las distribuciones de MAEs no son normales, son exponenciales inversas como se demostró anteriormente, ya que tienen un sesgo manifiesto en la cola izquierda o *skewness* positivo. Se propone en este trabajo utilizar la técnica de transformación de variable para normalizar y controlar el número de *outliers* de la muestra. Una vez transformada la variable es aconsejable comprobar la normalidad mediante los contrastes estadísticos comentados anteriormente. Se aplica la anti-transformación o función inversa para obtener los Anti-MAEs de la nueva distribución. Se introduce la nueva métrica robusta TMAE y se calcula la SD mediante la técnica de Jorgensen para conferir robustez al trabajo comparativo.

Respecto al porcentaje de ojos incluidos en los intervalos dióptricos estudiados, se aconseja introducir una nueva variante recontando los PE miópicos e hipermetrópicos en dos rangos diferentes calculando su promedio y desviación típica en dichos rangos.

Por último, se propone un nuevo índice de éxito denominado IOF con el que se puede decidir qué fórmula es mejor que otras en base a dos métricas explicadas anteriormente: Anti_MAE y MEP.

11.3 ANEXO 3.-NORMALIZACIÓN: ANÁLISIS DE POBLACIÓN BIOMÉTRICA

11.3.1 Introducción

Para poder clasificar un ojo como atípico por sus características anatómicas, es necesario conocer cómo es un ojo normal y comprobar cómo se distribuye en la población (184-187).

Hay mucha variabilidad según autores en los criterios para clasificar axialmente a la población. Hoffer en 1980 publicó un trabajo con 7500 ojos medidos con biómetro ultrasónico en el que establece que los ojos cortos son aquellos que están por debajo de 22.00 mm de longitud axial, los ojos normales oscilan entre 22.00 y 24.50 mm, los medio-largos entre 24.50 y 26.00 mm y los muy largos por encima de 26.00 mm. Barrett (52) establece el punto de corte de ojos cortos en 22.50 mm igual que Melles (105), sin embargo para ojos largos Barrett lo fija en 24.50 mm y Melles en 25.50 mm.

En un estudio reciente de Shammas y Jabre mediante una metodología denominada *e-norms* (188) se concluye que los ojos cortos están por debajo de 22.00 mm, medios-cortos entre 22.01 y 22.49 mm, normales entre 22.50 y 24.50 mm, medio-largos entre 24.51 y 24.99 mm y largos por encima de 25.00 mm.

En cualquier caso, se podría decir que existe un consenso más o menos generalizado para fijar el punto de transición entre ojos normales y largos en 24.50 mm de longitud axial, sin embargo hay más controversia en el establecimiento del punto de corte para ojos cortos. No obstante, la clasificación más extendida entre investigadores y clínicos es la propuesta por Hoffer hace ya más de 4 décadas.

En los pocos estudios publicados hasta la fecha en relación a la estratificación axial ocular, los autores se limitan a explicar la distribución general de algunas variables biométricas, no todas, estableciendo puntos de corte para el agrupamiento sin esclarecer su procedencia, pero en ningún caso se explica cómo se distribuyen el resto de parámetros biométricos: ACD, LT, WTW y KM dentro de esos grupos.

Sería interesante, por ejemplo, conocer en ojos cortos la distribución del diámetro horizontal corneal así como del grosor cristalino. A priori esto solo tendrá importancia clasificatoria demográfica sin valor práctico clínico, aunque no es desdeñable la potencial correlación con el cálculo biométrico, no siendo el objeto principal de esta disertación.

En este anexo de la tesis, se dará respaldo y argumento estadístico a la siguiente pregunta ¿qué ojo es más atípico o más singular? un ojo de 20.87 mm en una mujer con una ACD de 2.20 mm, LT de 4.00 mm, WTW de 11.00 mm y KM de 42.00 D o bien un ojo de un hombre con una AXL de 21.23 mm, ACD de 3.00 mm, LT de 3.88 mm, WTW de 11.00 mm y KM de 42.00 D.

Para ello, se dispone de una gran base de datos, 10385 ojos, obtenida con IOL Master® 700 sobre pacientes caucásicos, no intervenidos con ningún procedimiento quirúrgico oftalmológico y sobre los que se han descartado patologías corneales.

El objetivo final es encontrar unos puntos de corte basados en percentiles de selección axial para clasificar anatómicamente, así como analizar el resto de variables biométricas dentro de cada rango axial. De esta manera, se propondrá un sistema novedoso clasificatorio basado en un algoritmo de decisión que detectará el grado de

rareza anatómica de un ojo por lo singular e infrecuente de sus características anatómicas.

11.3.2 Análisis del problema clasificatorio

Es razonable pensar que la longitud axial es el factor anatómico determinante para clasificar un ojo, puesto que es la mayor de todas las longitudes biométricas oculares. Carecería de sentido clasificar por profundidad de cámara anterior ya que aunque obviamente se pueda estratificar, es más fácil encontrar AXL de todo tipo en un rango pequeño de ACD que al contrario como demuestra la Figura 115.

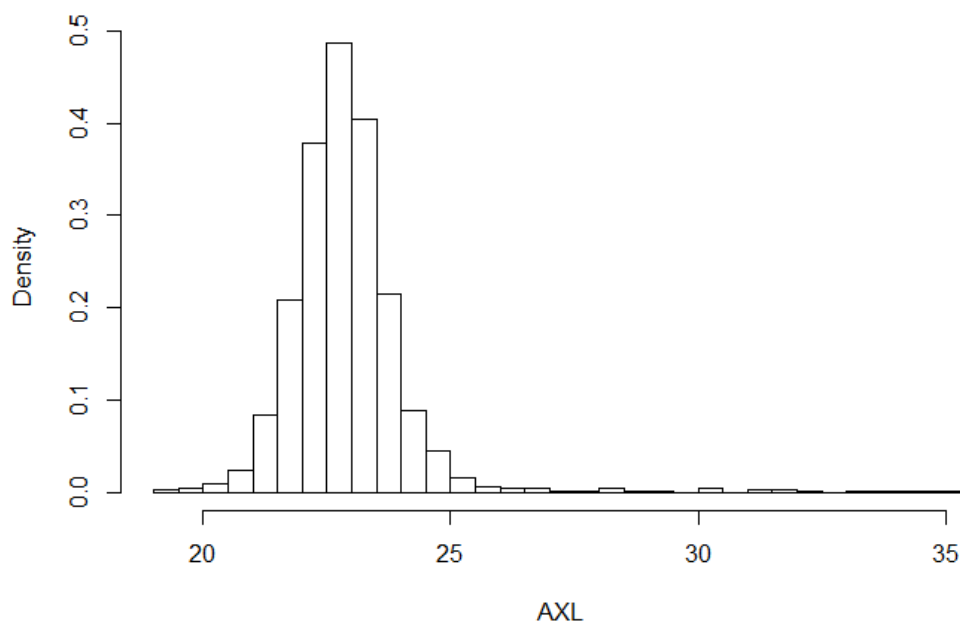


Figura 115. Histograma de distribución de la AXL dentro de un rango de ACD determinado (2.91 a 3.16 mm).

En un rango de ACD de 2.91 a 3.16 mm dónde se encuentra aproximadamente el 23% de la población, se puede comprobar que la distribución de AXL es muy amplia y

dispersa, desde valores inferiores a 20.00 mm hasta 35.00 mm o más, lo que hace inmanejable dicha clasificación.

Por el contrario, si se invierten las categorías exploradas y se estudia la ACD en un rango concreto de AXL, la segmentación de ACD es mucho más contenida con cotas máxima y mínima más manejables, lo que hace absolutamente improbable encontrar una ACD de 5.00 mm en un ojo de 22.00 mm.

Argumentos suficientes para proceder a categorizar anatómicamente por longitud axial y dentro de cada rango encontrado de esta, se estudiarán cómo se distribuyen el resto de las variables biométricas: ACD, LT, WTW y KM.

Como ya se expresó anteriormente, esta manera de clasificar a priori no tiene correlación con el cálculo de la potencia de la lente intraocular emetropizante, y por ende con la fórmula o método empleado, además se insiste en que no es el objeto del presente anexo.

Por otro lado, un BIPLLOT (Figura 116) aproxima la distribución de una muestra multivariante en un espacio de dos dimensiones, y superpone sobre la misma representaciones de las variables sobre las que se mide la muestra. Dichas representaciones son vectores y coinciden con las direcciones en las que mejor se muestra el cambio individual de cada variable, siendo los cosenos de los ángulos formados entre ellos el grado de relación existente.

Por tanto, se confirma la gran correlación positiva entre AXL, ACD y WTW y negativa entre LT, AXL, ACD y WTW. KM tiene una débil correlación con AXL y algo más fuerte con WTW. No obstante, las variables biométricas que más se relacionan son AXL y ACD.

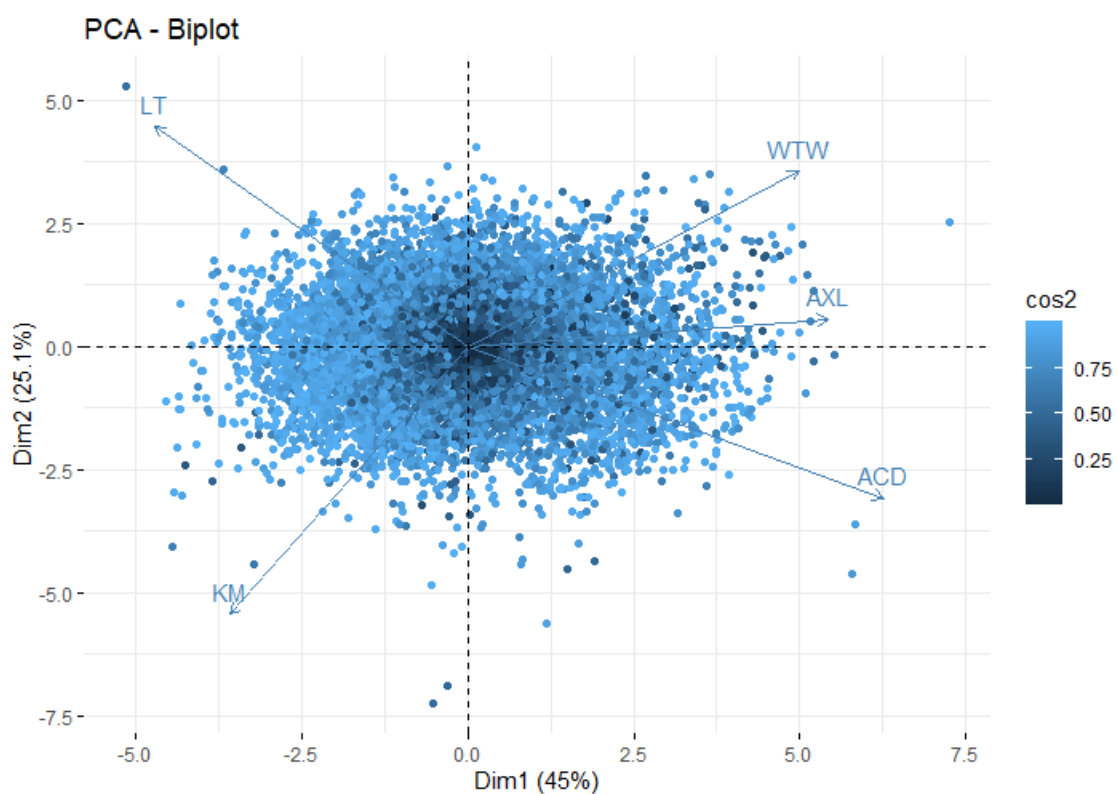


Figura 116. Biplot representando en dos ejes la correlación entre todas las variables biométricas.

Para complementar el análisis gráfico hecho con BIPLLOT se realiza un estudio analítico mediante matrices de correlación (Figura 117). En efecto, AXL y ACD tienen una correlación positiva de 0.50 y LT se correlaciona negativamente con ACD por un coeficiente de 0.62 pero con un escaso -0.23 con AXL.

De lo que se desprende que el argumento expuesto en el bloque anterior de segmentar la población por AXL para ver cómo se distribuyen el resto de variables biométricas en esos rangos, seguida de ACD siendo la variable que más se relaciona con el eje anteroposterior, es no solo factible sino recomendable.

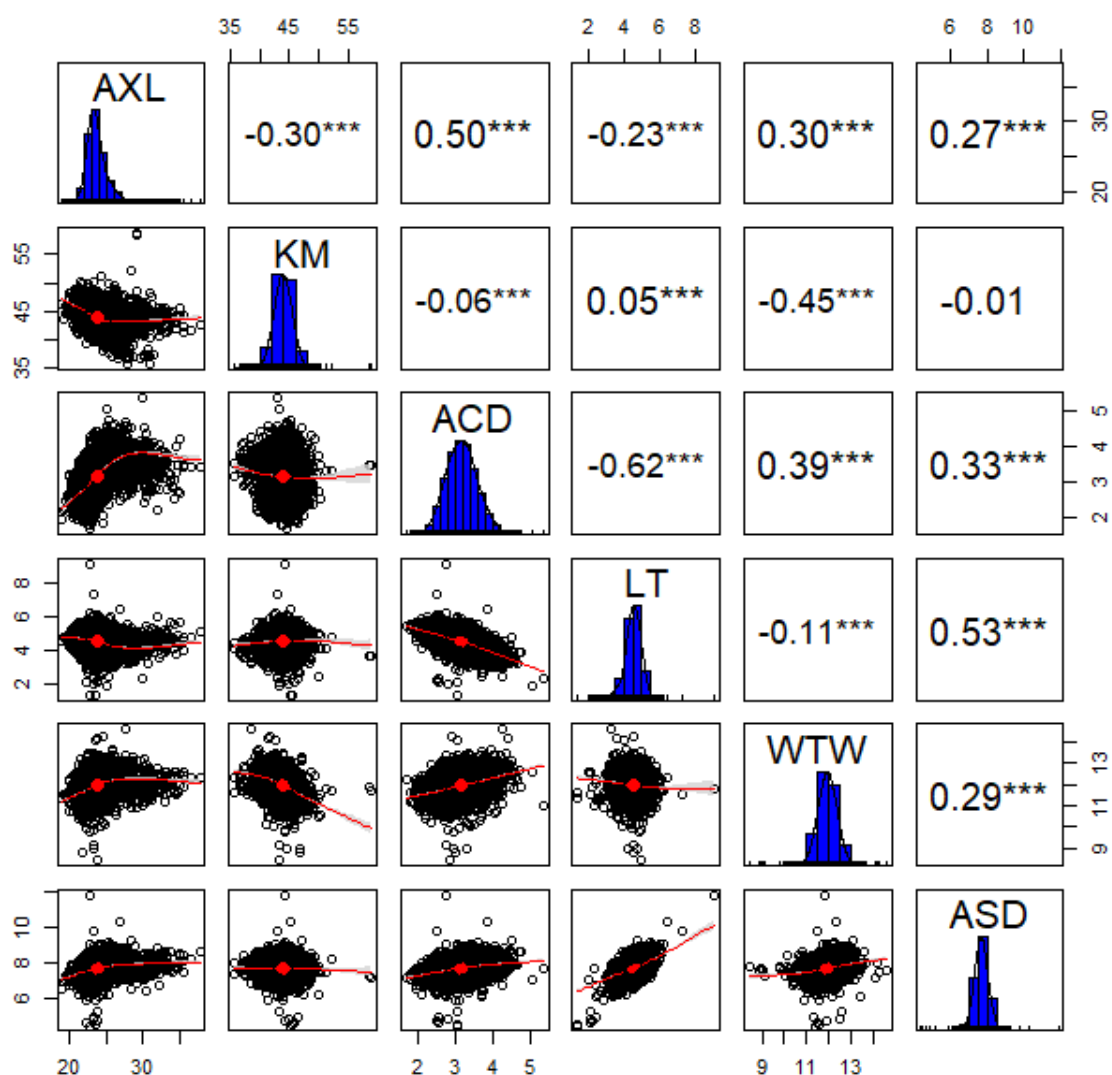


Figura 117. Matriz de correlaciones de todas las variables biométricas estudiadas. El diagrama contiene los scatter-plot de relación lineal, los histogramas y los factores de correlación de Pearson.

11.3.3 Tamaño muestral

¿Qué tamaño poblacional será necesario para representar fidedignamente a todo un universo a nivel biométrico? Previamente será necesario concretar ciertos parámetros: como el número de ojos que componen la población a estudiar, el margen de error, el nivel de confianza y la diversidad del universo a estudiar. Aplicando la siguiente fórmula:

$$n = \frac{Z^2 N p q}{e^2 (N - 1) + Z_{\alpha}^2 p q} \quad (72)$$

Donde:

N : es el tamaño del universo.

Z_{α} : es una constante que depende del nivel de confianza que se asigne.

e : es el error muestral deseado, en tanto por ciento.

p : proporción de individuos que poseen en la población la característica de estudio. Este dato es generalmente desconocido y se suele suponer que es 0.5 que es la opción más segura.

q : proporción de individuos que no poseen esa característica, es decir, es 1-p.

De esta manera se obtendrá una aproximación buena de la cantidad de ojos necesarios para estudiar y representar el comportamiento biométrico-anatómico de la población en general. Para un universo aleatorio de 50.000.000 de ojos caucásicos sin intervenciones quirúrgicas, con un nivel de confianza del 95.7% y un margen de error del 1%, será necesaria una muestra de 10.237 ojos.

Se extrajo una muestra final de 10.295 ojos de IOL Master® 700 cumpliendo las premisas citadas anteriormente. Estos fueron el total de ojos que se emplearon para las clasificaciones propuestas en los apartados anteriores.

11.3.4 Metodología de análisis

La sistemática de análisis que se siguió a lo largo de todas las clasificaciones que se propondrán en los sucesivos apartados de este anexo, fue la siguiente.

En primer lugar se estudió la normalidad en todas las variables biométricas mediante una triple combinación gráfica de histograma con curva de densidad, *boxplot* y *QQplot* (Figura 118), así como analíticamente mediante diferentes contrastes de hipótesis: *Kolmogorov-Smirnov* con corrección de *Lilliefors*, *Anderson-Darling* y *Jarque-Bera* como tercera vía ante una posible discrepancia entre los anteriores. De cada variable estudiada se calculó la media, desviación estandar, rangos máximo y mínimo, así como la mediana.

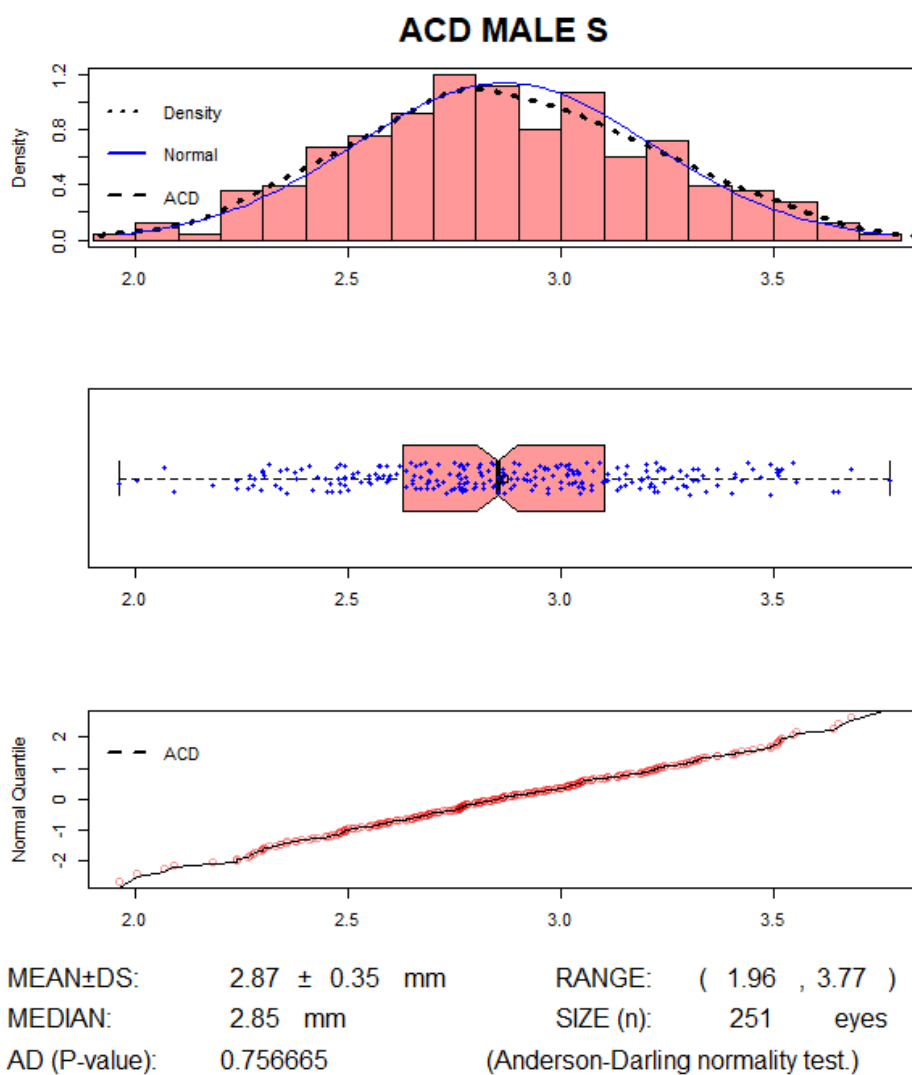


Figura 118. El gráfico superior representa el histograma de frecuencias con las curvas de densidad poblacional y la de normalidad. El gráfico del medio es un boxplot con la media, la mediana y su intervalo de confianza, así como la nube de puntos. El último gráfico es un Q-Qplot que representa la distribución de los datos comparándolos con una distribución normal. (La figura analiza la distribución de ACD en hombres con ojos cortos).

Por otro lado en todas las variables biométricas estudiadas, se computaron las tablas de distribución de frecuencias relativas, absolutas y acumuladas, así como sus diagramas asociados mediante representación polinómica para facilitar visualmente los puntos de inflexión y cambios de pendiente en las frecuencias (Figura 119).

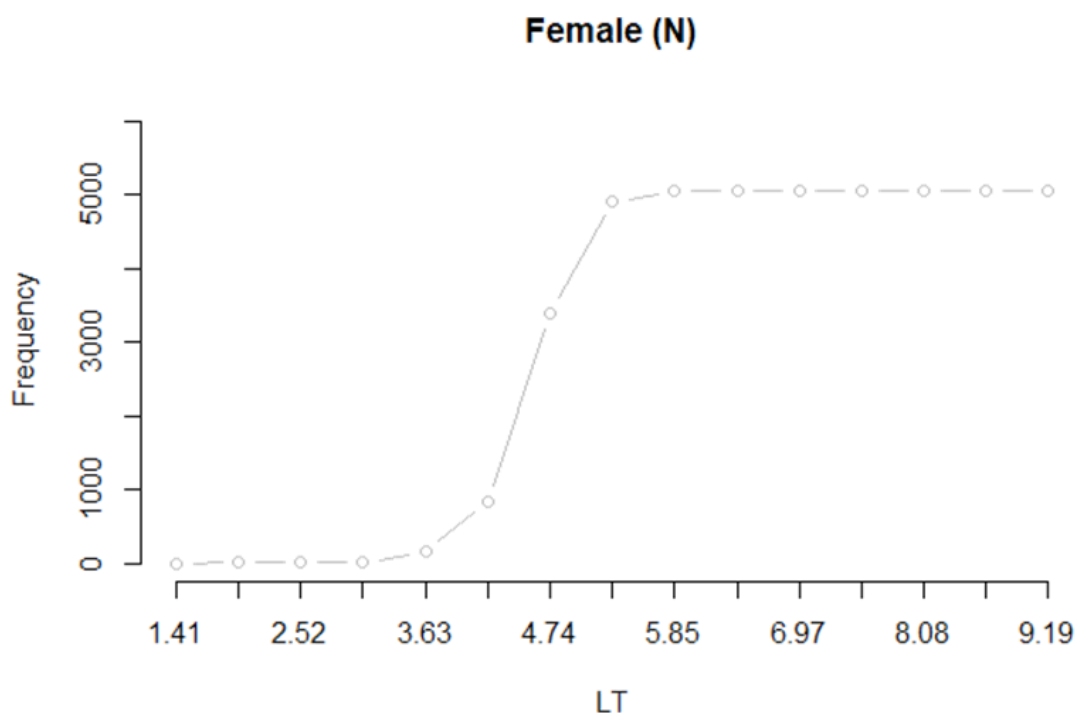


Figura 119. Diagrama polinómico para representar las frecuencias de una muestra. En este gráfico se ve la distribución de frecuencias del grosor cristaliniano en ojos considerados axialmente normales en mujeres.

Class limits	f	rf	rf(%)	cf	cf(%)
[1.406,1.962)	1	0.00	0.02	1	0.02
[1.962,2.518)	3	0.00	0.06	4	0.08
[2.518,3.074)	4	0.00	0.08	8	0.16
[3.074,3.63)	140	0.03	2.77	148	2.93
[3.63,4.186)	692	0.14	13.71	840	16.65
[4.186,4.742)	2551	0.51	50.55	3391	67.20
[4.742,5.298)	1511	0.30	29.94	4902	97.15
[5.298,5.854)	139	0.03	2.75	5041	99.90
[5.854,6.411)	2	0.00	0.04	5043	99.94
[6.411,6.967)	0	0.00	0.00	5043	99.94
[6.967,7.523)	1	0.00	0.02	5044	99.96
[7.523,8.079)	0	0.00	0.00	5044	99.96
[8.079,8.635)	0	0.00	0.00	5044	99.96

Figura 120. Tabla de distribución de frecuencias que acompaña a la representación polinómica de la figura anterior (Figura 119). La primera columna (Class limits) se corresponde a los intervalos generados, la segunda (f) es la frecuencia absoluta o número de casos en cada intervalo de clase. La tercera (rf) es la frecuencia relativa. La cuarta columna (rf%) es la frecuencia relativa porcentual. La quinta (cf) es la frecuencia acumulada y la sexta (cf%) es la frecuencia acumulada en porcentaje. Esta tabla de frecuencias representa el espesor de cristalino (LT) en ojos de mujeres axialmente normales.

Finalmente, se propondrá un índice que permita cuantificar la anormalidad o singularidad para tipificar y clasificar cada ojo a nivel biométrico en función de sus características anatómicas. Para ello se cuantificó la rareza en cada variable biométrica por separado para posteriormente totalizar obteniendo un valor único de singularidad.

11.3.5 Clasificación de la población biométrica

11.3.5.1 Clasificación por AXL

La **¡Error! No se encuentra el origen de la referencia.** muestra la distribución de la longitud axial de toda la muestra ($n = 10.295$ ojos) en función del género. A grandes rasgos, las distribuciones de longitud axial entre mujeres y hombres son semejantes, lo que hace pensar la no necesidad de discriminar por sexo en las clasificaciones pudiendo utilizar un sistema único.

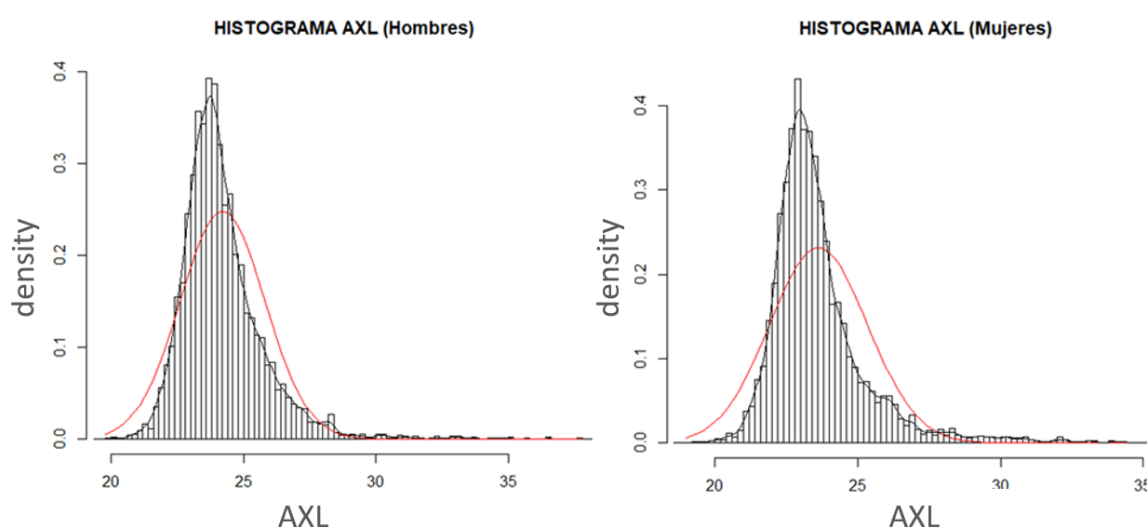


Figura 121. Histogramas de densidad para AXL distinguiendo por género de la muestra ($n = 10.295$).

No obstante, las tablas de frecuencia relativa y absoluta ponen de manifiesto diferencias significativas entre ambos sexos (Figura 122). En los hombres cerca del 78% de la población ($n = 4197$) se concentra entre 19.77 mm y 25.00 mm de longitud axial, sin embargo en mujeres ($n = 6046$) es hasta el 86%, casi un 10% más.

Si se exploran los percentiles con detalle se pone de manifiesto que en mujeres hay más cantidad de ojos cortos por debajo de 22.00 mm que en hombres, 9% versus 3% respectivamente. De 22.00 mm a 24.00 mm en hombres se acumula el 50% de la población, pero en mujeres el 64%.

Con todo esto, a efectos de categorización y clasificación axial es necesario la discriminación por sexo, de manera que se propondrá una organización bivariada en función del género, lo que implica que en todas las categorías biométricas analizadas se distinguirá por sexo.

Se puede observar un sesgo hacia longitudes axiales superiores a 25.00 mm. La curva de densidad normal se aleja bastante de la distribución de datos real. Efectivamente, aplicando un test de *Kolmogorov-Smirnov* con corrección de *Lilliefors* ($p < 0.0001$) se pone de manifiesto la distribución no normal de los datos. Se corrobora este hecho con otros contrastes como *Anderson-Darling* ($p < 0.0001$) y *Jarque-Bera* ($p < 0.0001$).

HOMBRES

Class limits	f	rf	rf(%)	cf	cf(%)
[19.77,21.078)	13	0.00	0.31	13	0.31
[21.078,22.387)	251	0.06	5.98	264	6.29
[22.387,23.695)	1465	0.35	34.91	1729	41.20
[23.695,25.003)	1540	0.37	36.69	3269	77.89
[25.003,26.311)	573	0.14	13.65	3842	91.54
[26.311,27.619)	218	0.05	5.19	4060	96.74
[27.619,28.927)	77	0.02	1.83	4137	98.57
[28.927,30.235)	20	0.00	0.48	4157	99.05
[30.235,31.543)	14	0.00	0.33	4171	99.38
[31.543,32.851)	7	0.00	0.17	4178	99.55
[32.851,34.159)	10	0.00	0.24	4188	99.79
[34.159,35.468)	5	0.00	0.12	4193	99.90
[35.468,36.776)	3	0.00	0.07	4196	99.98
[36.776,38.084)	1	0.00	0.02	4197	100.00

MUJERES

Class limits	f	rf	rf(%)	cf	cf(%)
[18.899,20.12)	11	0.00	0.18	11	0.18
[20.12,21.34)	127	0.02	2.10	138	2.28
[21.34,22.561)	1253	0.21	20.72	1391	23.01
[22.561,23.781)	2656	0.44	43.93	4047	66.94
[23.781,25.002)	1140	0.19	18.86	5187	85.79
[25.002,26.222)	441	0.07	7.29	5628	93.09
[26.222,27.443)	187	0.03	3.09	5815	96.18
[27.443,28.663)	96	0.02	1.59	5911	97.77
[28.663,29.884)	50	0.01	0.83	5961	98.59
[29.884,31.104)	45	0.01	0.74	6006	99.34
[31.104,32.325)	20	0.00	0.33	6026	99.67
[32.325,33.545)	11	0.00	0.18	6037	99.85
[33.545,34.766)	7	0.00	0.12	6044	99.97
[34.766,35.986)	2	0.00	0.03	6046	100.00

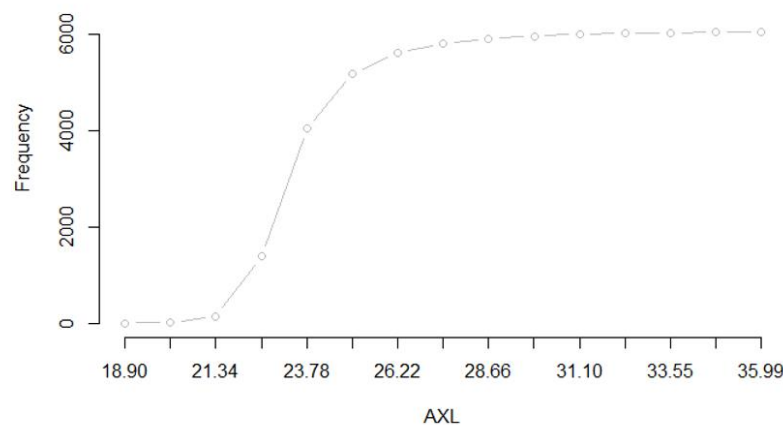
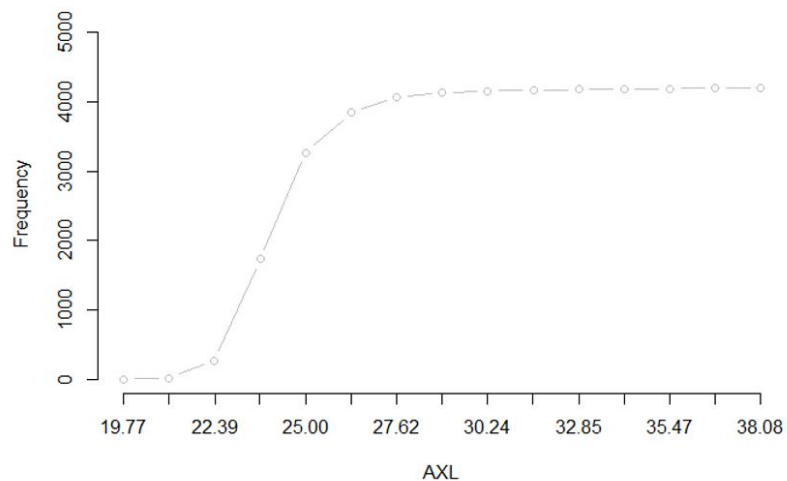


Figura 122. Distribución de frecuencias en hombres y mujeres para AXL.

El plot *QQnorm* con *QQline* de AXL (Figura 123) muestra claramente una distribución gaussiana solo en ojos considerados “normales” axialmente, de 20.00 a 25.00 mm aproximadamente, siendo no normal para el resto de longitudes axiales.

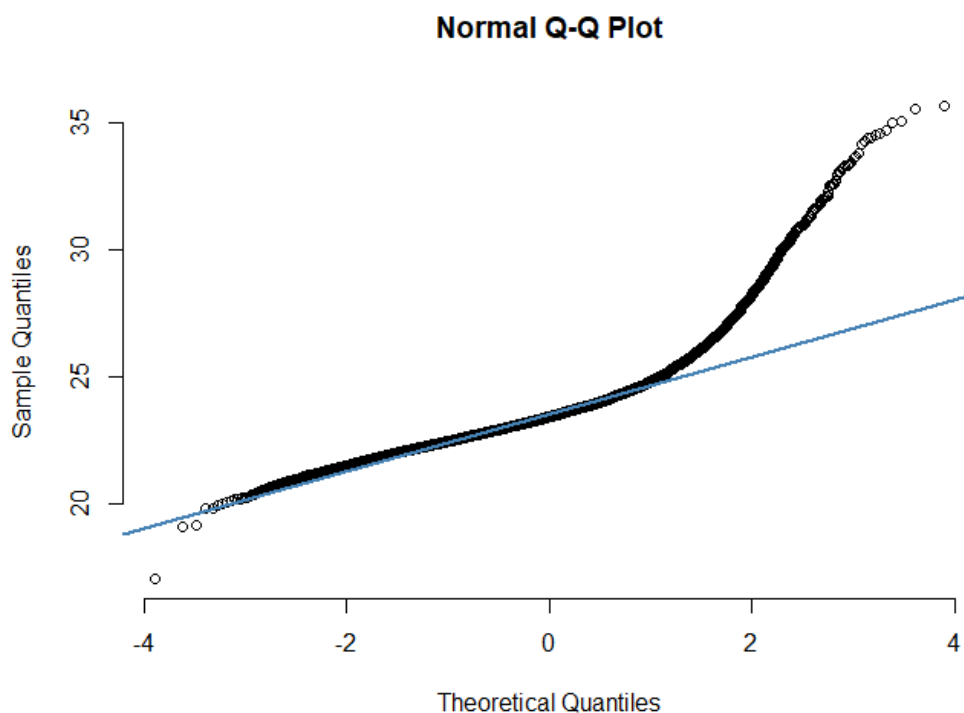


Figura 123. QQ-plot para longitud axial dónde los cuantiles de la distribución de los datos comparada con una normal se alejan de esta en los extremos, sobre todo en longitudes largas por encima de 25.00 mm.

La no normalidad en la distribución de longitudes axiales implica cierta inconsistencia estadística al seleccionar la media como valor de centralidad, cuyo resultado es para la muestra total de 23.87 mm, para la muestra de hombres 24.21 mm y para la de mujeres 23.63 mm. De manera que es más razonable tomar la mediana como valor de centralidad, en este caso 23.54 mm para la muestra total sin discriminar por género, 23.92 mm para hombres y 23.26 mm para mujeres.

Volviendo a las tablas de frecuencias relativas y absolutas, así como a su distribución polinómica de las longitudes axiales discriminando por género y aplicando la

regla de Sturges, son 4.5 las clases que se deben considerar para acumular población, lo que implica que se pueden fijar preliminarmente cinco puntos de corte que generan cinco escenarios axiales diferentes en función del género:

TIPO (ABR.)	GÉNERO	RANGO	% FRECUENCIA
Extracortos (XS)	HOMBRES	< 21.10 mm	0.31 %
	MUJERES	< 20.12 mm	0.18 %
Cortos (S)	HOMBRES	≥ 21.10 mm y < 22.39 mm	5.98 %
	MUJERES	≥ 20.12 mm y < 21.34 mm	2.10 %
Normales (N)	HOMBRES	≥ 22.39 mm y < 25.00 mm	72%
	MUJERES	≥ 21.34 mm y < 25.00 mm	84%
Largos (L)	HOMBRES	≥ 25.00 mm y < 27.62 mm	19 %
	MUJERES	≥ 25.00 mm y < 27.44 mm	10 %
Extralargos (XL)	HOMBRES	≥ 27.62 mm	3.26 %
	MUJERES	≥ 27.44 mm	3.72 %

Tabla 32. Clasificación poblacional por longitud axial en hombres y mujeres. Se proponen 5 categorías de ojos en función de la longitud del eje anteroposterior: extra cortos o extrashort (XS), cortos o shorts (S), normales o normal (N), largos o long (L) y extralargos o extra-Long (XL).

Estos puntos de corte seleccionados no son arbitrarios y obedecen a los cambios abruptos de pendiente de la curva que dibuja la frecuencia acumulada (Tabla 32).

No obstante, por qué el valor en mujeres que delimita la transición entre ojos cortos y normales es 21.34 mm y no otro. Evidentemente en algún punto hay que establecer el límite anatómico de clasificación, y por paradójico que parezca un ojo con 21.34 mm se considerará un ojo corto y otro con 21.35 mm será normal.

Así por ejemplo un ojo de 20.53 mm de un hombre se considerará extracorto (XS) al estar por debajo de 21.10 mm, sin embargo ese mismo ojo en una mujer solo será corto (S) al estar comprendido en el intervalo [20.12, 21.34). Esto conlleva la etiqueta de ojo anómalo o atípico independientemente del género, ya que en el caso de los hombres solo un 0.31 % de la población tiene esa característica anatómica y un 2.10% en el caso de las mujeres.

A continuación tal como se expresó en apartados anteriores, una vez clasificado el ojo axialmente, se estudiarán las otras variables biométricas y cómo se distribuyen dentro de cada rango axial para tener cubierto todo el espectro y sus combinaciones anatómicas.

11.3.5.2 Clasificación de ACD

En primer lugar, la ACD en ojos categorizados como extracortos (XS) en hombres presenta una distribución normal ($p\text{-value} = 0.8341$) con un promedio de 2.872 ± 0.319 mm (2.300, 3.341). En mujeres la distribución de ACD en ese rango es normal ($p\text{-value} = 0.1371$) con un promedio de 2.444 ± 0.319 mm (1.950, 2.689)

En ojos cortos (S), la distribución de ACD para hombres es normal ($p\text{-value} = 0.7039$) con un promedio de 2.865 ± 0.350 mm (1.963, 3.772). En mujeres la distribución de ACD también es normal ($p\text{-value} = 0.8132$) con un promedio de 2.662 ± 0.313 mm (1.968, 3.568). Las tablas respectivas de frecuencias de ACD, sus *plots* polinómicos correspondientes y los histogramas se representan en la *Histogramas, plots polinómicos de frecuencia y tablas de frecuencia de ACD en ojos cortos para hombres y mujeres*. Figura 124.

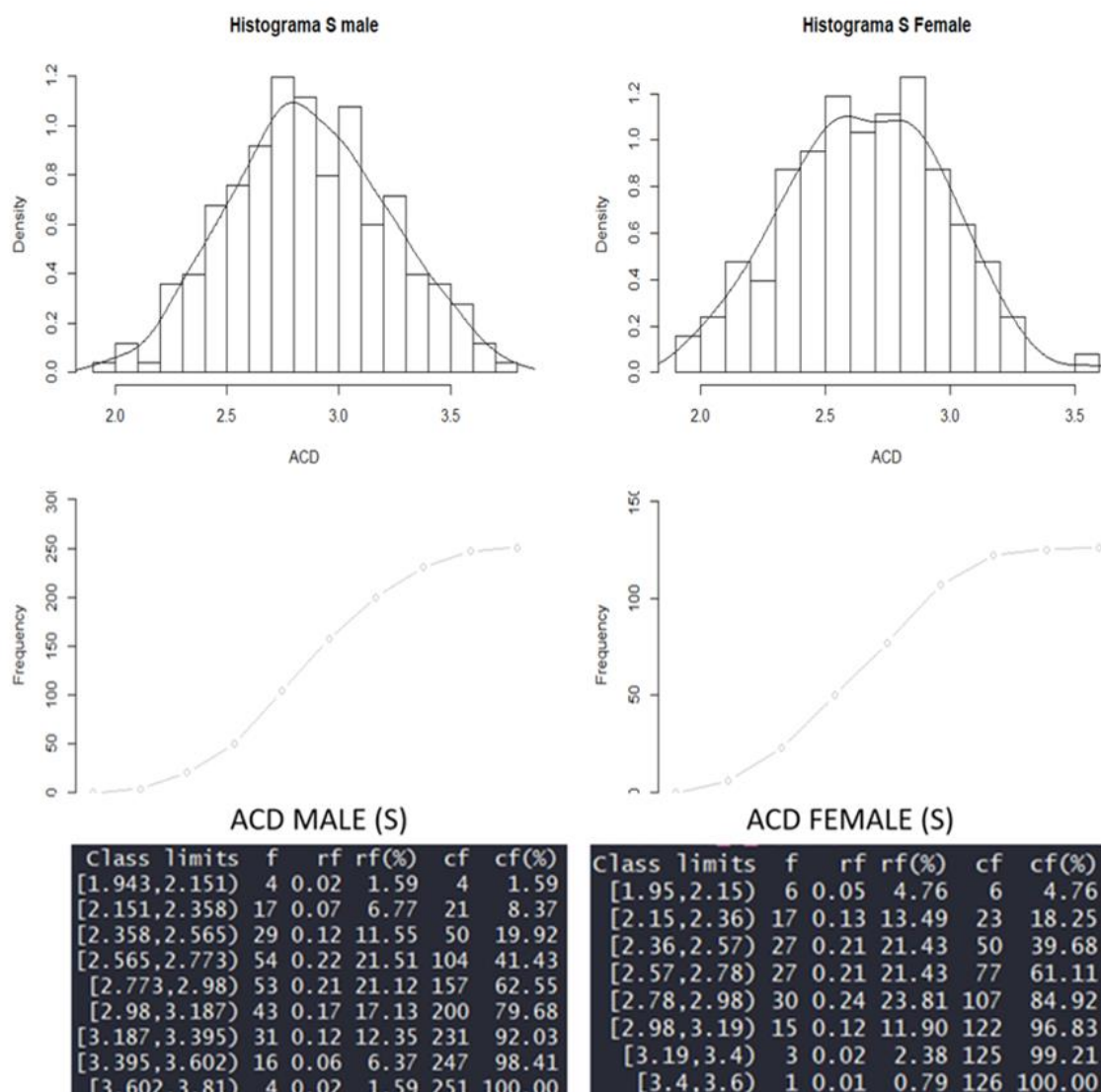


Figura 124. Histogramas, plots polinómicos de frecuencia y tablas de frecuencia de ACD en ojos cortos para hombres y mujeres.

Atendiendo a las frecuencias relativas de ACDs en ojos cortos se puede concluir que en hombres el 72% de los valores se encuentran en el intervalo [2.358, 3.187) y todos los valores que excedan dicho rango se considerarán atípicos, lo que implica por ejemplo que un ojo de 21.53 mm de longitud axial con una ACD de 3.42 mm será mucho más infrecuente que el mismo ojo con una ACD de 2.86 mm. Este ojo en particular sería atípico por su longitud axial y también por su profundidad de cámara anterior, infrecuente en ese rango de axialidad.

En mujeres el 79% de los valores de ACD se encuentran en el rango [2.150, 2.980), de manera que para un hombre una ACD de 3.18 mm en un ojo corto se consideraría normal puesto que está en el límite superior pero incluida, en una mujer con un ojo corto ese valor de profundidad de cámara es considerado atípico por superar la cota superior.

Como las ACDs se distribuyen normalmente en ambos sexos, habiendo calculado el valor de centralidad y de dispersión, se podrían establecer cuatro rangos de singularidad, los valores que estén dentro de una $\pm DS$ (68.2% del total), los que estén dentro de dos $\pm DS$ (27.2%), los que estén dentro de tres $\pm DS$ (4.2%) y por último los que estén fuera de tres DS (0.4%). De esta manera, una mujer con un ojo axialmente corto y una ACD de 3.30 mm (Figura 125) será mucho más atípico que el mismo ojo con una ACD de 3.00 mm y este a su vez más infrecuente que el mismo ojo con una ACD de 2.70 mm.

En ojos axialmente normales, la distribución de ACD (Figura 125) es normal tanto en hombres como mujeres (p -value = 0.1085 y 0.257 respectivamente). En hombres el promedio de ACD es 3.186 ± 0.375 mm (1.880, 4.430) y en mujeres 3.048 ± 0.368 mm (1.686, 4.396).

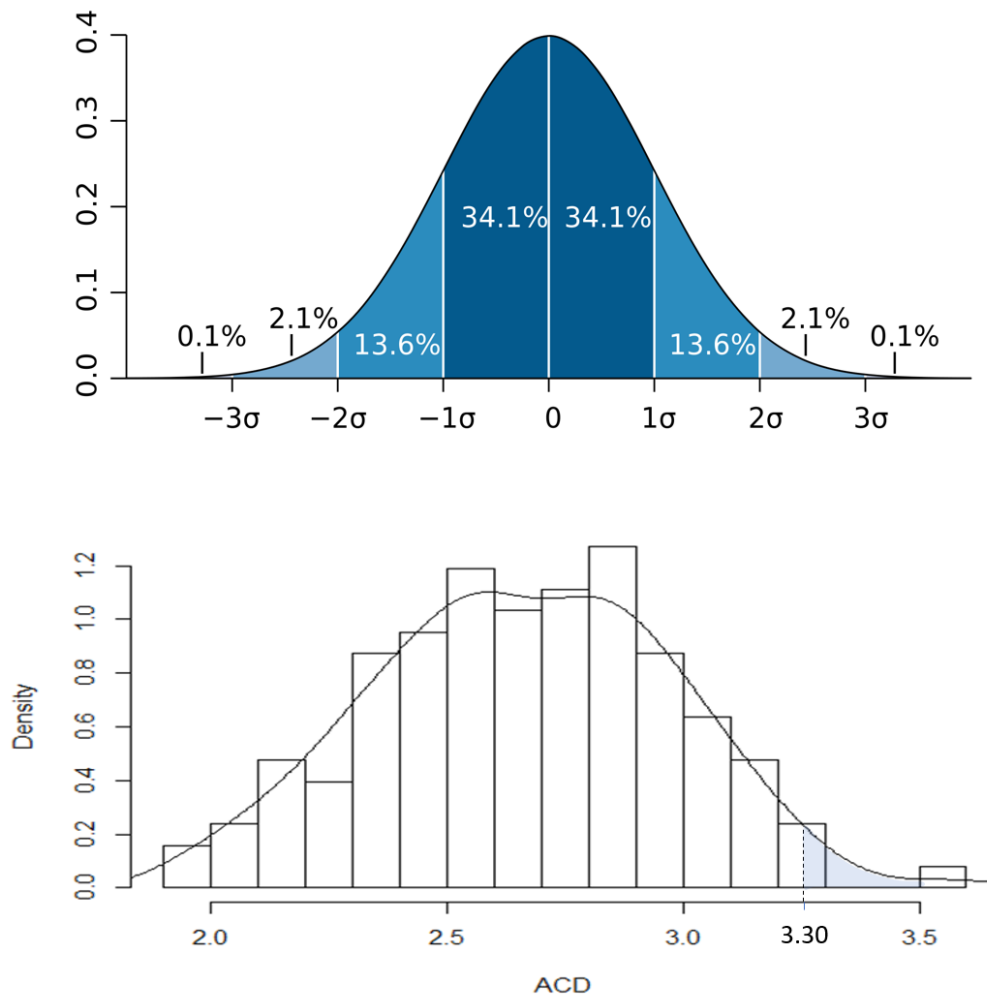


Figura 125. Histograma de ACD en ojos cortos para mujeres (inferior) en comparación con una distribución teórica normal (superior). Se puede observar un ejemplo de ACD de 3.00 mm recaería en una zona correspondiente al 2.1% de la población, es decir, muy infrecuente.

En ojos axialmente normales en hombres el 82% de los valores de ACD se encuentran en el rango [2.464, 3.670) que es dónde se produce más cambio de pendiente en el gráfico polinómico de frecuencias (Figura 126).

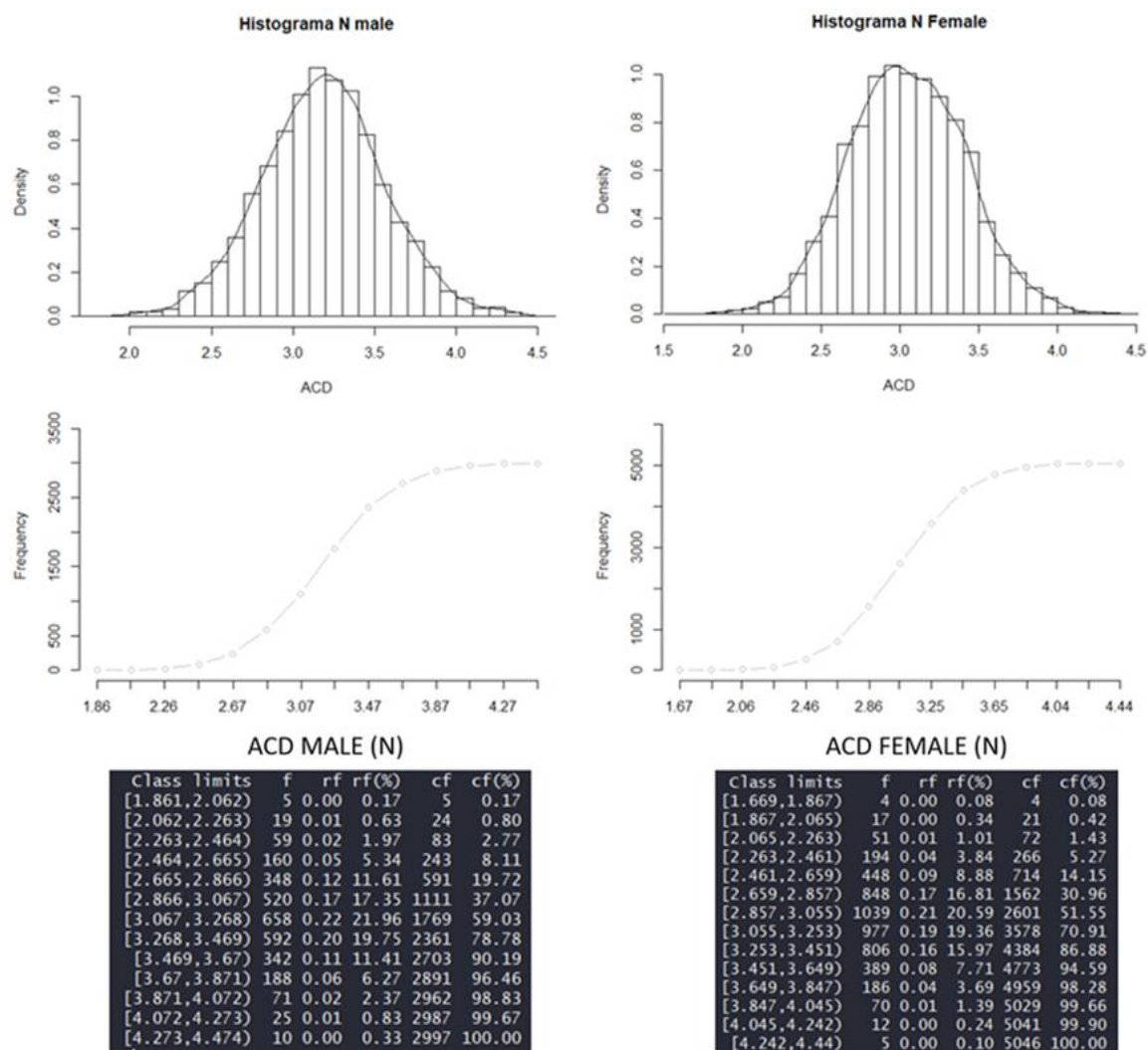
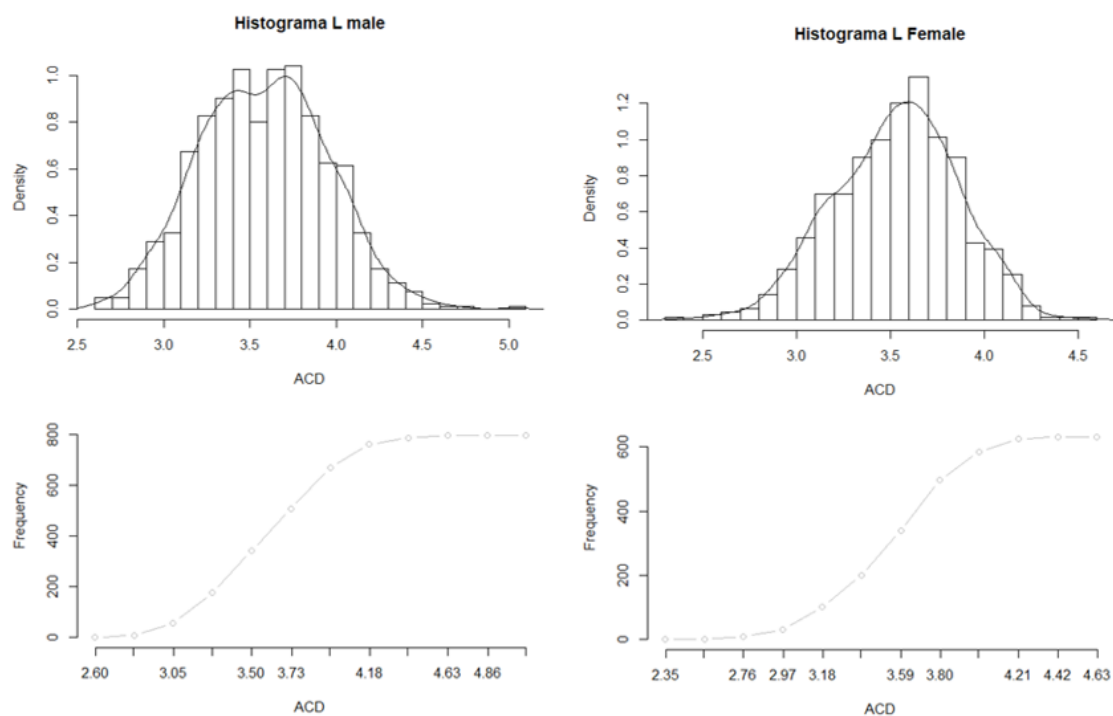


Figura 126. Histogramas, plots polinómicos y tablas de frecuencia de ACD en ojos normales para hombres y mujeres.

Para ojos largos (L) la distribución de ACDs tanto en hombres como en mujeres sigue la normalidad (p -value = 0.0566 y 0.07165 respectivamente) con un promedio de 3.581 ± 0.368 mm (2.622, 5.031) para hombres y un 3.531 ± 0.330 mm (2.370, 4.580) para mujeres. Una mujer con una longitud axial ubicada en el rango de normalidad y una ACD de 4.00 mm, es tan atípico como en un hombre con las mismas características anatómicas. Del mismo modo, un individuo con AXL normal pero una ACD pequeña de 2.20 mm, es muy infrecuente como se puede observar en las tablas de distribución de frecuencia.



ACD MALE (L)

Class limits	f	rf	rf(%)	cf	cf(%)
[2.596,2.822)	9	0.01	1.13	9	1.13
[2.822,3.048)	47	0.06	5.88	56	7.01
[3.048,3.274)	121	0.15	15.14	177	22.15
[3.274,3.5)	165	0.21	20.65	342	42.80
[3.5,3.726)	168	0.21	21.03	510	63.83
[3.726,3.952)	160	0.20	20.03	670	83.85
[3.952,4.178)	92	0.12	11.51	762	95.37
[4.178,4.404)	26	0.03	3.25	788	98.62
[4.404,4.63)	9	0.01	1.13	797	99.75
[4.63,4.856)	1	0.00	0.13	798	99.87
[4.856,5.082)	1	0.00	0.13	799	100.00

ACD FEMALE (L)

Class limits	f	rf	rf(%)	cf	cf(%)
[2.346,2.554)	2	0.00	0.32	2	0.32
[2.554,2.761)	7	0.01	1.11	9	1.42
[2.761,2.968)	22	0.03	3.48	31	4.91
[2.968,3.175)	70	0.11	11.08	101	15.98
[3.175,3.382)	98	0.16	15.51	199	31.49
[3.382,3.59)	140	0.22	22.15	339	53.64
[3.59,3.797)	158	0.25	25.00	497	78.64
[3.797,4.004)	89	0.14	14.08	586	92.72
[4.004,4.211)	40	0.06	6.33	626	99.05
[4.211,4.419)	5	0.01	0.79	631	99.84
[4.419,4.626)	1	0.00	0.16	632	100.00

Figura 127. Histogramas, plots polinómicos y tablas de frecuencia de ACD en ojos largos para hombres y mujeres.

Un hombre con un ojo axialmente largo, por ejemplo 26.12 mm, y una ACD de 4.41 mm se clasifica como ojo atípico por la AXL, pero también tiene una cámara muy infrecuente, ya que para 2 DS la ACD sería 4.317 mm y por tanto estaría en el rango de 3 DS, es decir, solo aproximadamente el 2.1% de la población tendría un valor así. Sin embargo, el mismo ejemplo de ojo con una ACD de 3.65 mm, sería tipificado como atípico únicamente por la AXL pero no por la profundidad de cámara anterior.

Por último, para los ojos extralargos (XL) la distribución de ACDs no es normal en hombres ($p\text{-value} < 0.001$), pero sí en mujeres ($p\text{-value} = 0.1905$) con un promedio de 3.451 ± 0.388 mm (2.232, 5.332), tal y como se puede comprobar en la Figura 128. Se aplicó un contraste de *Kolmogorov-Smirnov* además de un *Anderson-Darling* para confirmar este particular.

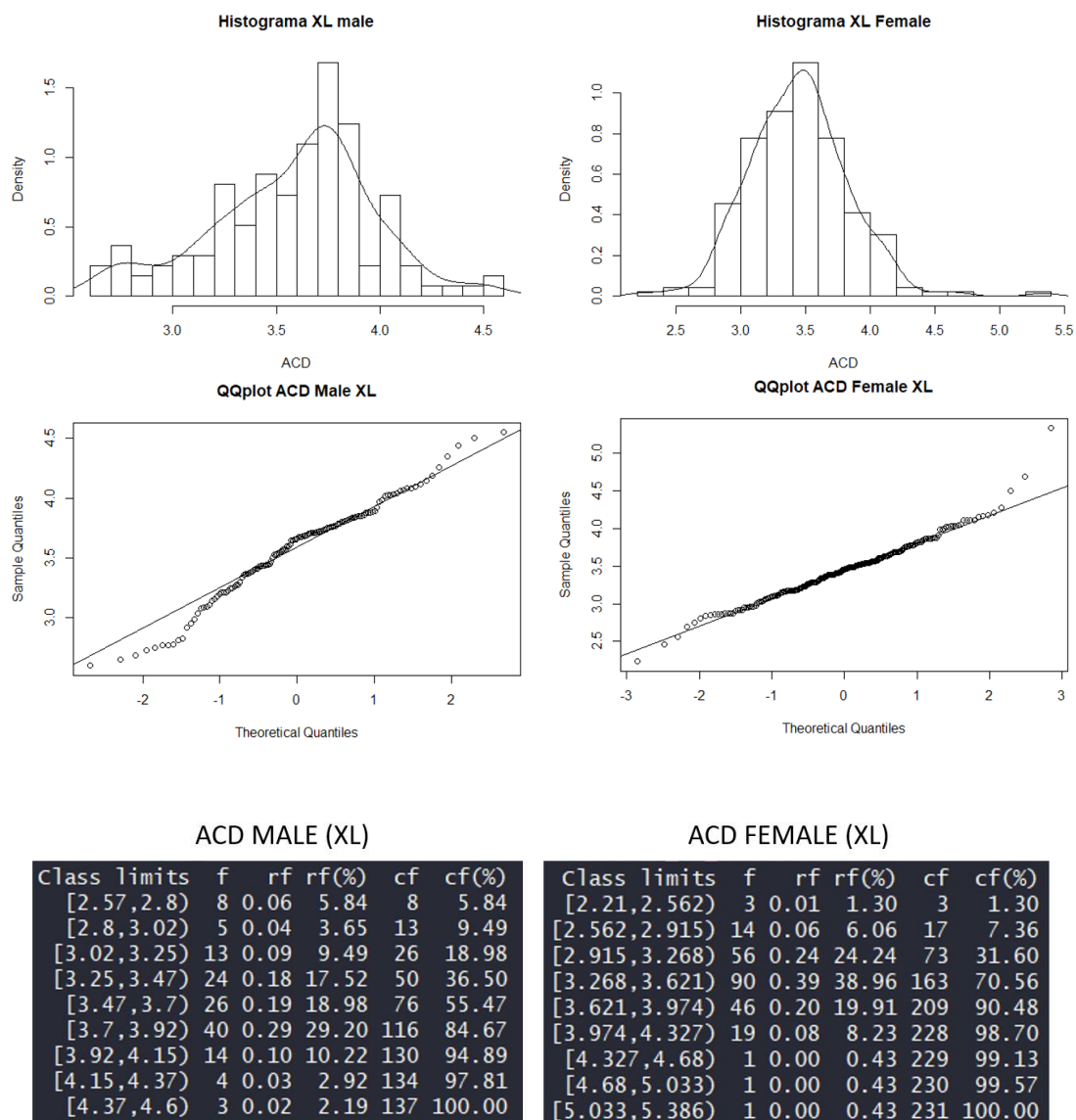


Figura 128. Histogramas, plots polinómicos y tablas de frecuencia de ACD en ojos extralargos para hombres y mujeres. En hombres hay un sesgo claro en ACD por debajo de 3.50 mm.

Por tanto, en el caso de afrontar un ojo XL en hombres se utilizará la tabla de frecuencias relativas para estratificar la singularidad y asemejar el escenario al de una distribución normal. El 66% de los casos estará en el intervalo [3.02, 3.92), habrá un 19% y un 15% respectivamente que se distribuirán a ambos lados por las dos colas de forma ligeramente asimétrica hacia la izquierda (Figura 128).

11.3.5.3 Clasificación de LT

A continuación se estudiarán las distribuciones del espesor de cristalino por género en todos los rangos axiales, tal como se ha hecho con la ACD.

Para ojos extracortos (XS) la distribución de LT en hombres sigue una distribución cuasi normal ($p\text{-value} = 0.03137$) con un promedio de 4.586 ± 0.361 mm (4.160, 5.299) y en mujeres sí es normal ($p\text{-value} = 0.4731$) con un promedio de 4.571 ± 0.199 mm (4.213, 4.809).

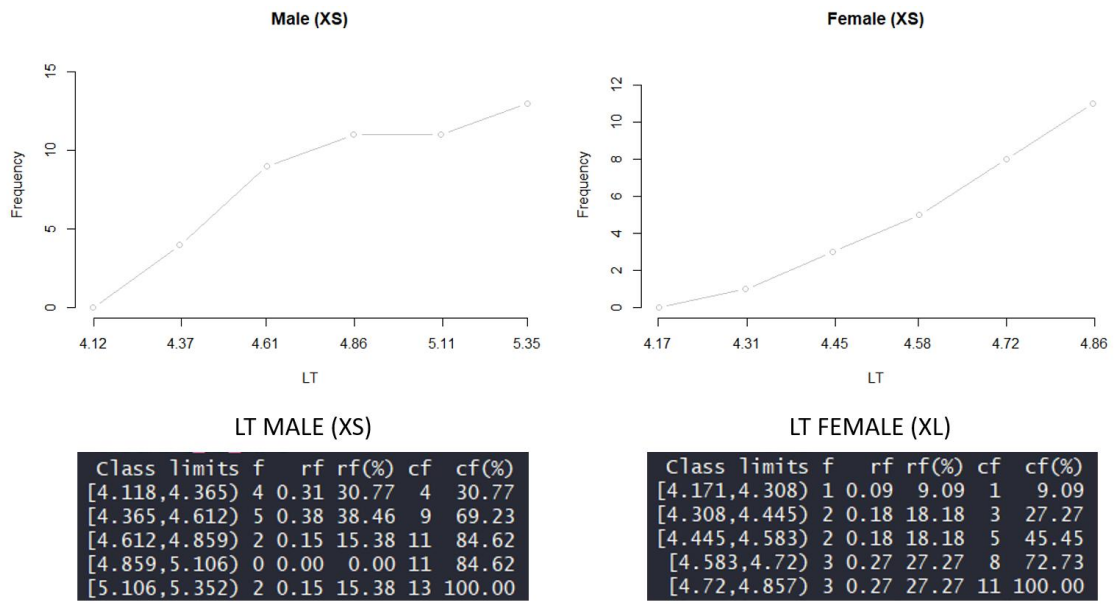


Figura 129. Representación polinómica de frecuencias absolutas y tabla de distribución de frecuencias del espesor de cristalino en ojos extra cortos para ambos sexos.

En el caso de ojos cortos (S) la distribución del LT es normal tanto en hombres como en mujeres (p -value = 0.117 y 0.5153 respectivamente) con un promedio de 4.680 \pm 0.422 mm (3.170, 5.693) para hombres y 4.625 \pm 0.388 mm (3.704, 5.697) para mujeres.

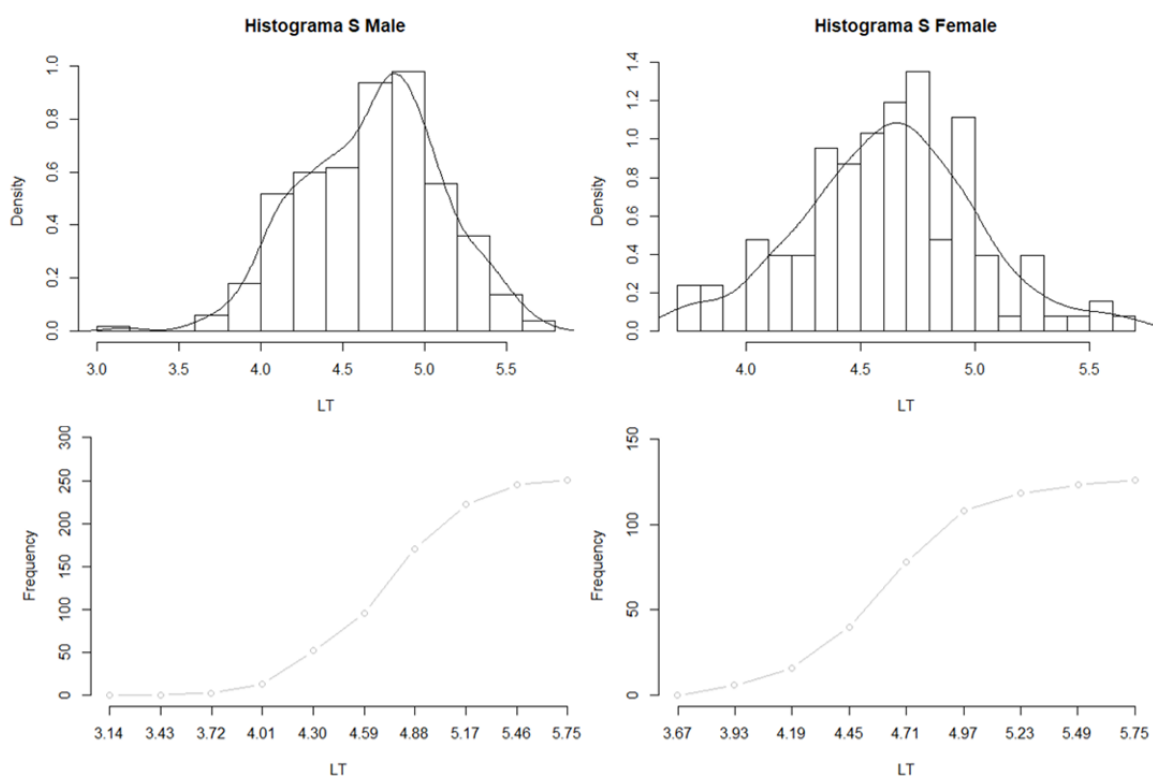


Figura 130. Histogramas de densidad y diagramas polinómicos de frecuencias del espesor cristalino en ojos axialmente cortos para ambos sexos.

LT MALE (S)						LT FEMALE (S)					
Class limits	f	rf	rf(%)	cf	cf(%)	Class limits	f	rf	rf(%)	cf	cf(%)
[3.138,3.428)	1	0.00	0.40	1	0.40	[3.667,3.928)	6	0.05	4.76	6	4.76
[3.428,3.719)	2	0.01	0.80	3	1.20	[3.928,4.189)	10	0.08	7.94	16	12.70
[3.719,4.009)	10	0.04	3.98	13	5.18	[4.189,4.449)	24	0.19	19.05	40	31.75
[4.009,4.299)	39	0.16	15.54	52	20.72	[4.449,4.71)	38	0.30	30.16	78	61.90
[4.299,4.589)	44	0.18	17.53	96	38.25	[4.71,4.971)	30	0.24	23.81	108	85.71
[4.589,4.879)	75	0.30	29.88	171	68.13	[4.971,5.232)	10	0.08	7.94	118	93.65
[4.879,5.17)	51	0.20	20.32	222	88.45	[5.232,5.493)	5	0.04	3.97	123	97.62
[5.17,5.46)	23	0.09	9.16	245	97.61	[5.493,5.754)	3	0.02	2.38	126	100.00
[5.46,5.75)	6	0.02	2.39	251	100.00						

Figura 131. Tabla de frecuencias de espesor de cristalino en ojos cortos para hombres y mujeres.

Por ejemplo, una mujer con un ojo axialmente corto de 21.00 mm, una ACD de 2.70 mm y un LT de 3.71 mm, se considerará atípico solo por dos categorías, la longitud axial y el grosor del cristalino, siendo la profundidad de cámara normal para ese rango de axialidad. El LT está en el rango de -3 DS, efectivamente, la tabla de frecuencias desvela que solo el 4.76% de la población (Figura 131) tiene ese LT en ojos cortos de mujeres.

En ojos axialmente normales (N), LT no sigue una distribución normal ($p\text{-value} < 2.2e-16$) ni en hombres ni en mujeres tal y como se puede comprobar en el *QQplot* y en el histograma con un sesgo en la cola izquierda.

No obstante, en este rango de ojos normales los outliers de ambos sexos que desplazan la distribución a la cola izquierda probablemente se correspondan con medidas erróneas, mal tomadas o mal estimadas por el dispositivo biométrico. No es normal que en ojos cortos el rango mínimo de espesor sea 3.138 mm para hombres y en ojos axialmente normales este grosor sea inferior a 2.5 mm. La supresión de estos valores atípicos normalizaría la distribución con un $p\text{-value} < 0.05$ en ambos casos.

La mayoría de los grosores cristalinianos en hombres (96%), en ojos axialmente normales, se concentran en el rango [3.577, 5.404) mm tal y como se puede comprobar en la tabla de frecuencias. En mujeres, el 99% se aglutina en el rango [3.074, 5.854) mm.

Para ojos largos (L) la distribución de LT tampoco es normal ($p\text{-value} = 0.001061$) ni en hombres ni en mujeres como se puede observar en la disposición del histograma y del gráfico *QQnorm*. (Figura 133).

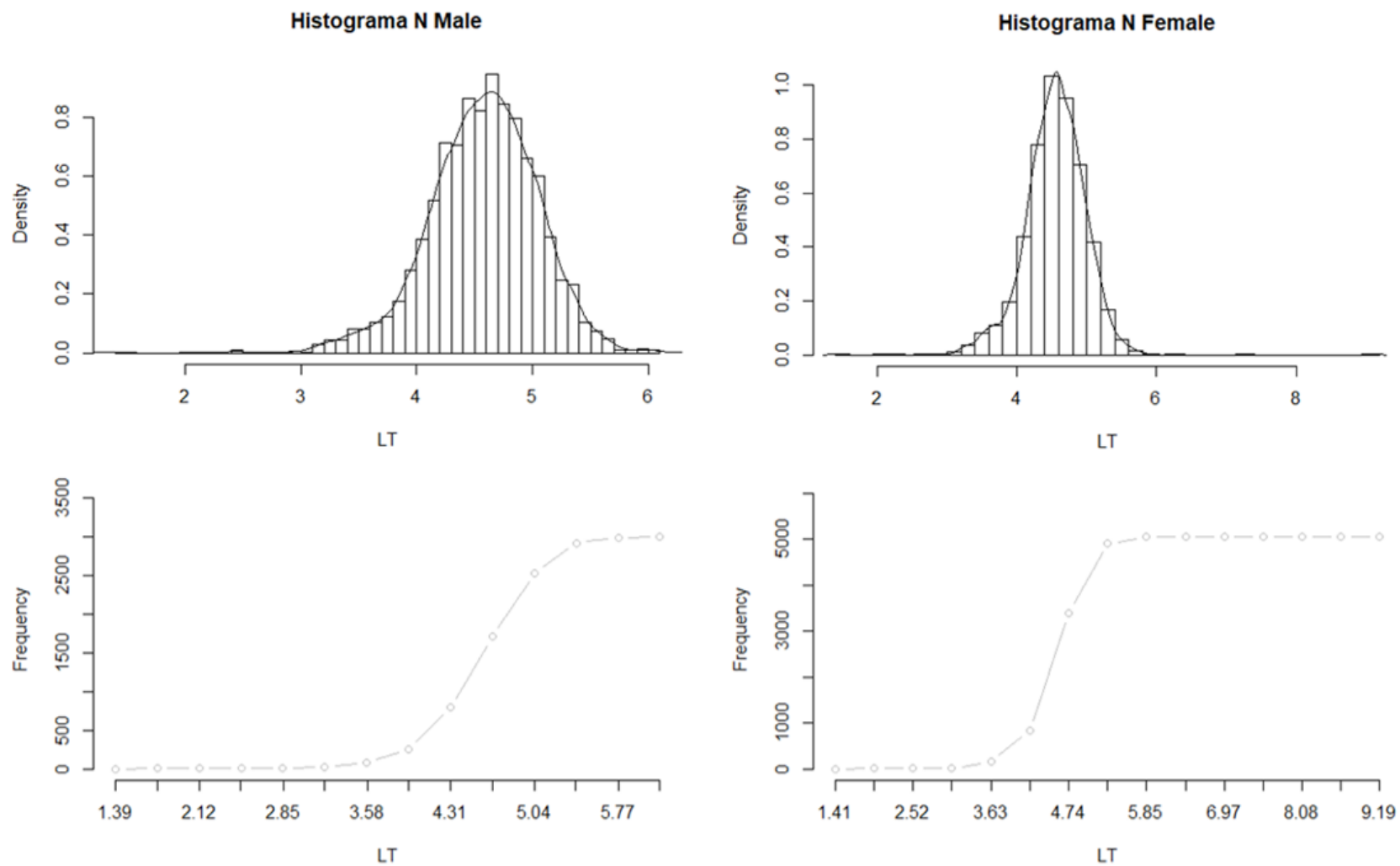
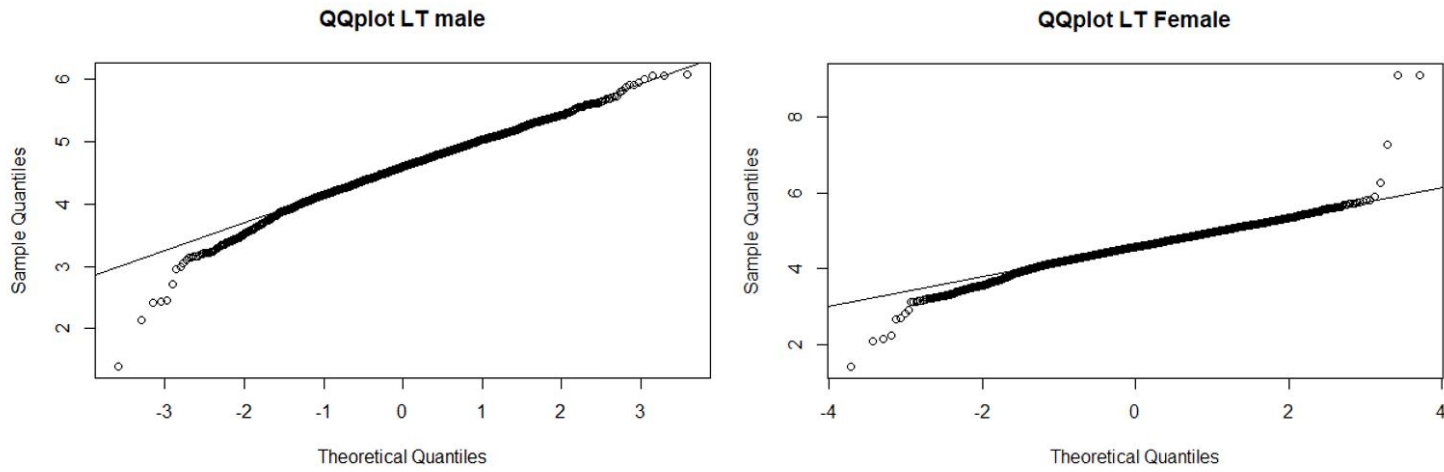


Figura 132. Histogramas y curvas polinómicas de frecuencias para grosor de cristalino en ojos normales para ambos sexos.



LT MALE (N)

Class limits	f	rf	rf(%)	cf	cf(%)
[1.386,1.751)	1	0.00	0.03	1	0.03
[1.751,2.116)	0	0.00	0.00	1	0.03
[2.116,2.482)	4	0.00	0.13	5	0.17
[2.482,2.847)	1	0.00	0.03	6	0.20
[2.847,3.212)	14	0.00	0.47	20	0.67
[3.212,3.577)	63	0.02	2.10	83	2.77
[3.577,3.943)	170	0.06	5.67	253	8.44
[3.943,4.308)	536	0.18	17.88	789	26.33
[4.308,4.673)	915	0.31	30.53	1704	56.86
[4.673,5.038)	828	0.28	27.63	2532	84.48
[5.038,5.404)	386	0.13	12.88	2918	97.36
[5.404,5.769)	69	0.02	2.30	2987	99.67
[5.769,6.134)	10	0.00	0.33	2997	100.00

LT FEMALE (N)

Class limits	f	rf	rf(%)	cf	cf(%)
[1.406,1.962)	1	0.00	0.02	1	0.02
[1.962,2.518)	3	0.00	0.06	4	0.08
[2.518,3.074)	4	0.00	0.08	8	0.16
[3.074,3.63)	140	0.03	2.77	148	2.93
[3.63,4.186)	692	0.14	13.71	840	16.65
[4.186,4.742)	2551	0.51	50.55	3391	67.20
[4.742,5.298)	1511	0.30	29.94	4902	97.15
[5.298,5.854)	139	0.03	2.75	5041	99.90
[5.854,6.411)	2	0.00	0.04	5043	99.94
[6.411,6.967)	0	0.00	0.00	5043	99.94
[6.967,7.523)	1	0.00	0.02	5044	99.96
[7.523,8.079)	0	0.00	0.00	5044	99.96
[8.079,8.635)	0	0.00	0.00	5044	99.96

Figura 133. QQplot y tablas de distribución de frecuencias de espesores cristalinos en ojos normales para ambos sexos.

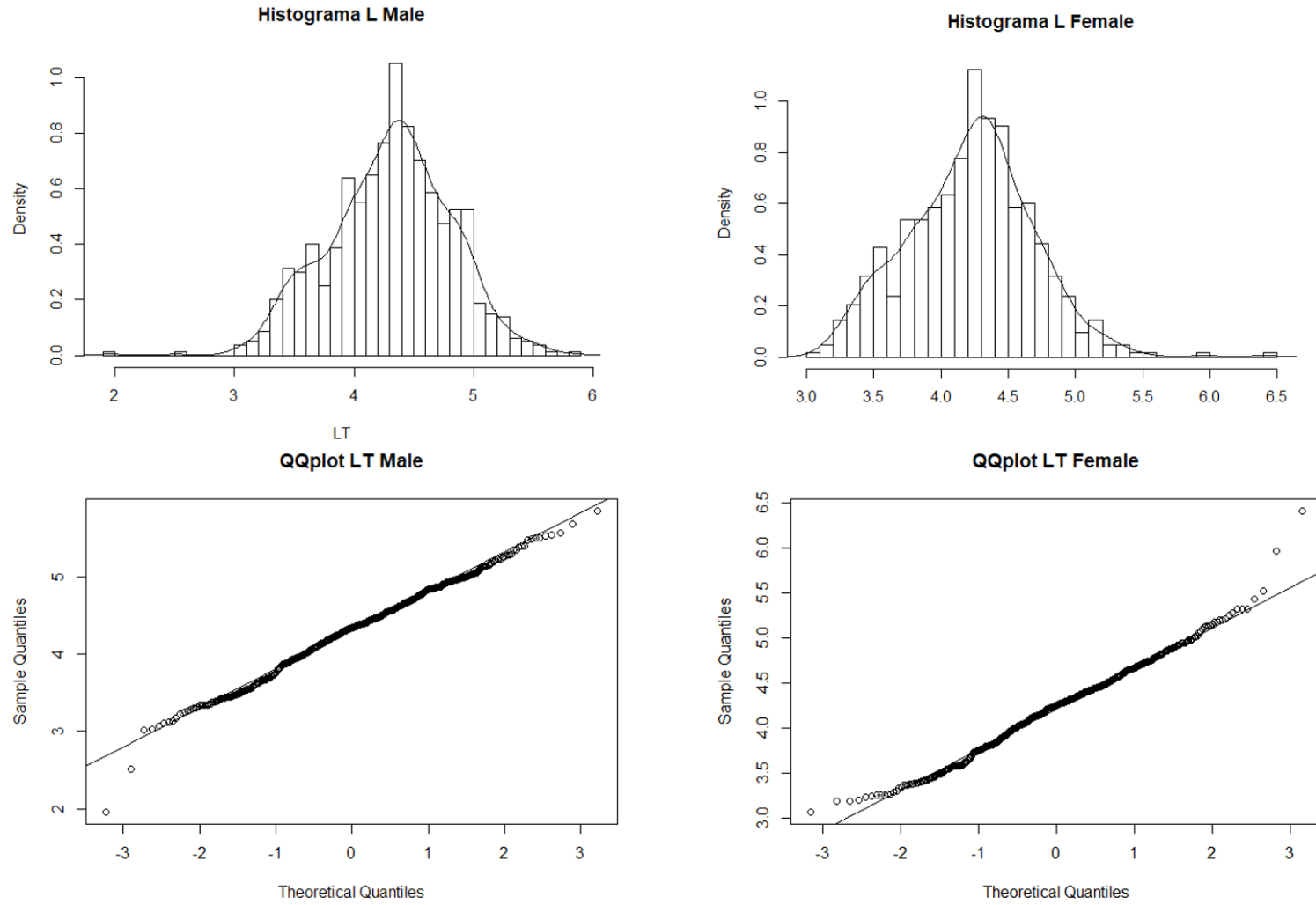


Figura 134. Histogramas y QQplot de LT en ojos axialmente largos para los dos sexos.

La no normalidad en LT (Figura 134) en ambos sexos en ojos largos se da por culpa de un sesgo marcado en la cola izquierda en hombres de 3 casos de cristalinos poco gruesos por debajo de 3.02 mm, lo que supone un 0.38% de la población. En el caso de las mujeres los *outliers* están en la cola derecha hay un 1.42 % de la población por encima de 5.22 mm de espesor cristaliniario.

Y por último en ojos extralargos (XL) la distribución de LT es normal en hombres ($p\text{-value} = 0.2793$) y en mujeres ($p\text{-value} = 0.5335$), con un valor medio de 4.383 ± 0.481 mm (3.241, 5.690) en hombres y un promedio de 4.347 ± 0.458 mm (2.343, 5.973) en mujeres.

A pesar de la extremada longitud axial del ojo independientemente del sexo, se ve que el grosor de cristalino se mueve en valores muy similares a los anteriores tramos de axialidad, es más, en ojos normales donde se presupondría un espesor cristaliniario menor en proporción a la longitud, hay un porcentaje alto tanto en hombres como en mujeres que tienen un LT superior a los límites por arriba en ojos XL.

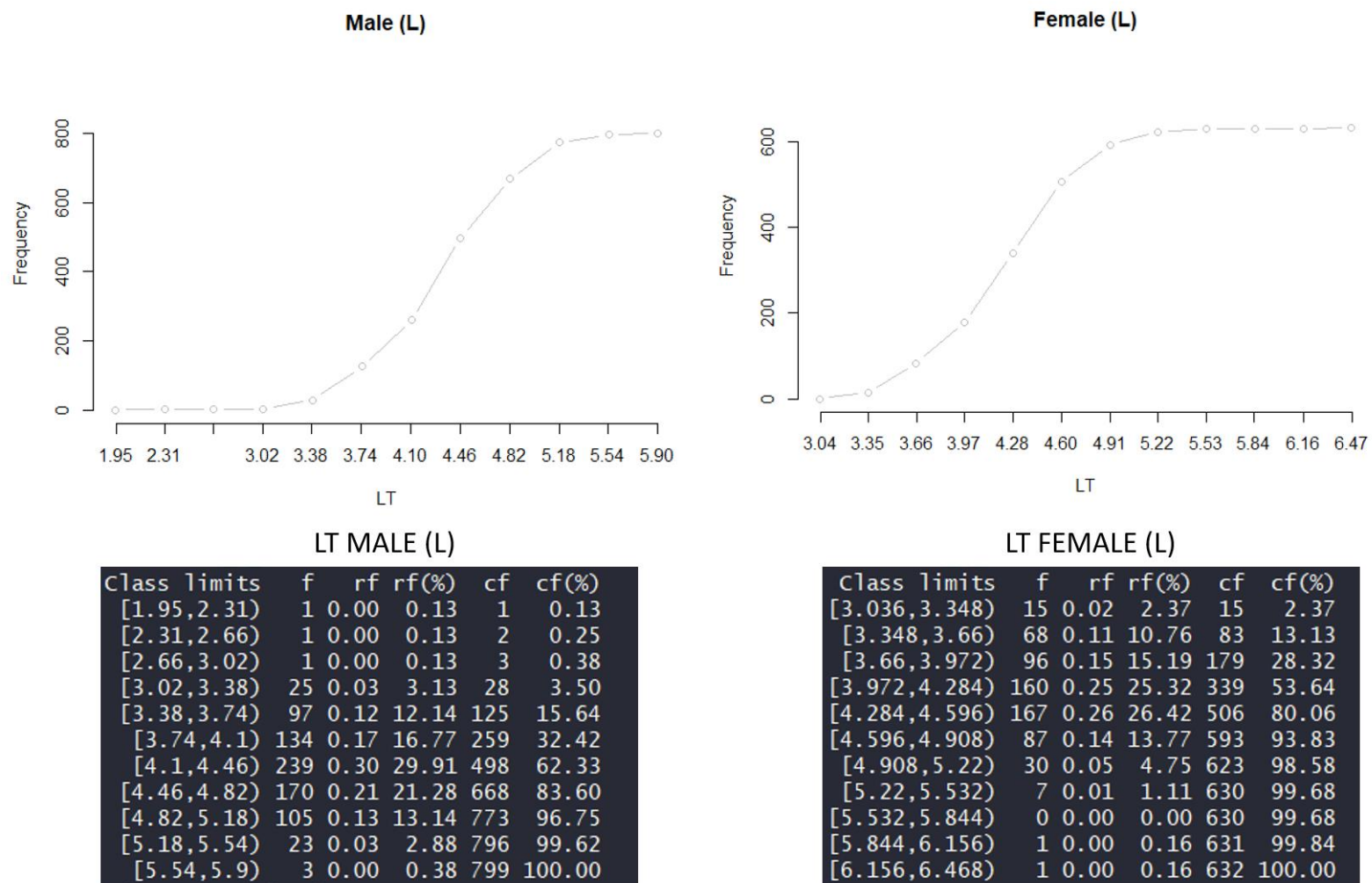


Figura 135. Plots polinómicos y tablas de frecuencia de LT en ojos largos para los dos sexos.

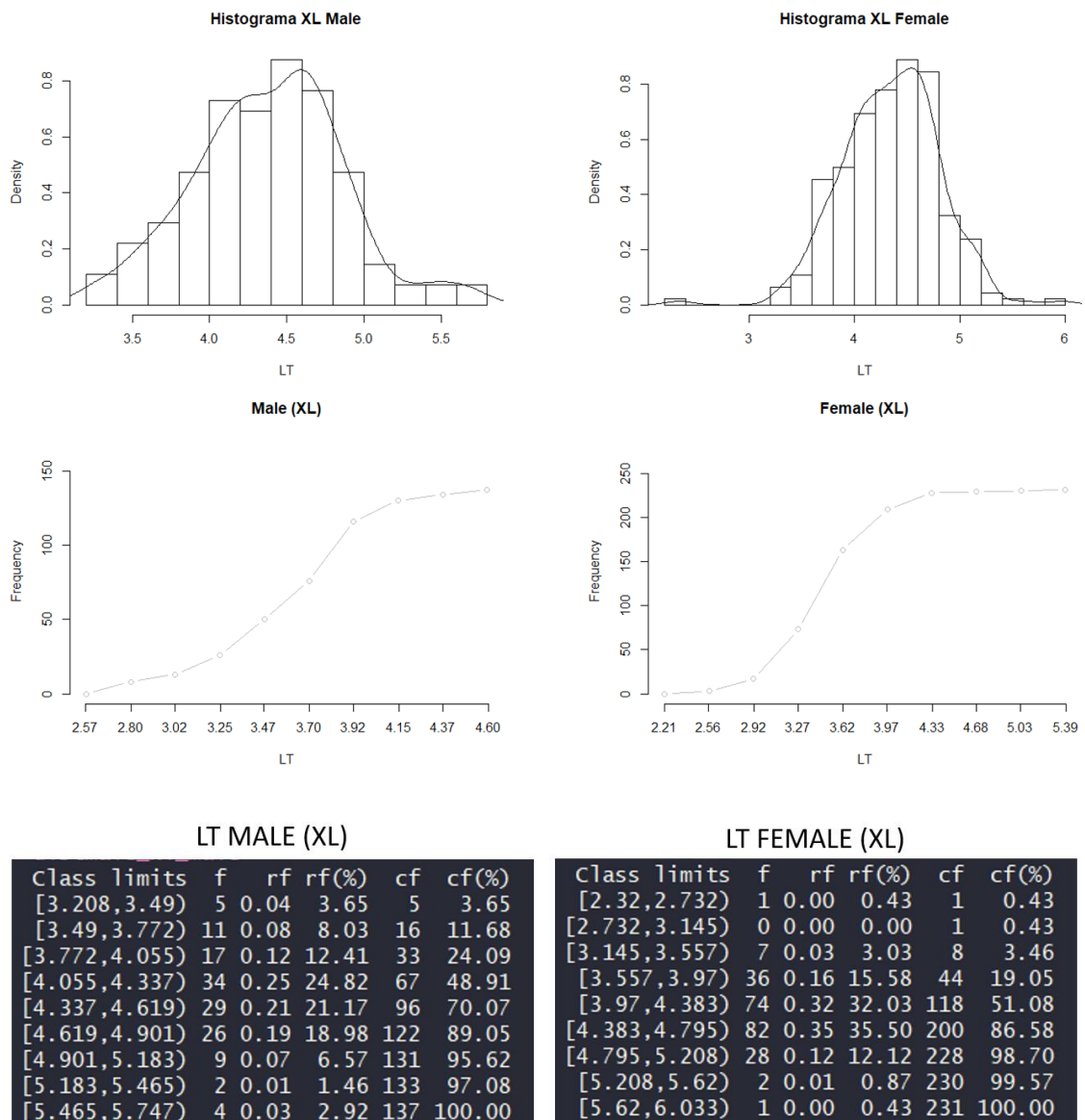


Figura 136. Histogramas, diagramas polinómicos y tablas de frecuencia para ojos extralargos en ambos sexos.

11.3.5.4 Clasificación por WTW

Ahora se analizará el diámetro horizontal corneal por tramos de longitud axial como se ha hecho con las otras variables biométricas.

Para ojos extracortos (XS) la distribución de WTW es normal tanto en hombres como en mujeres ($p\text{-value} = 0.4633$ y 0.1634 respectivamente). El promedio para hombres es de 11.920 ± 0.511 mm (11.070, 12.620) y en mujeres 11.470 ± 0.178 mm (11.100, 11.690).

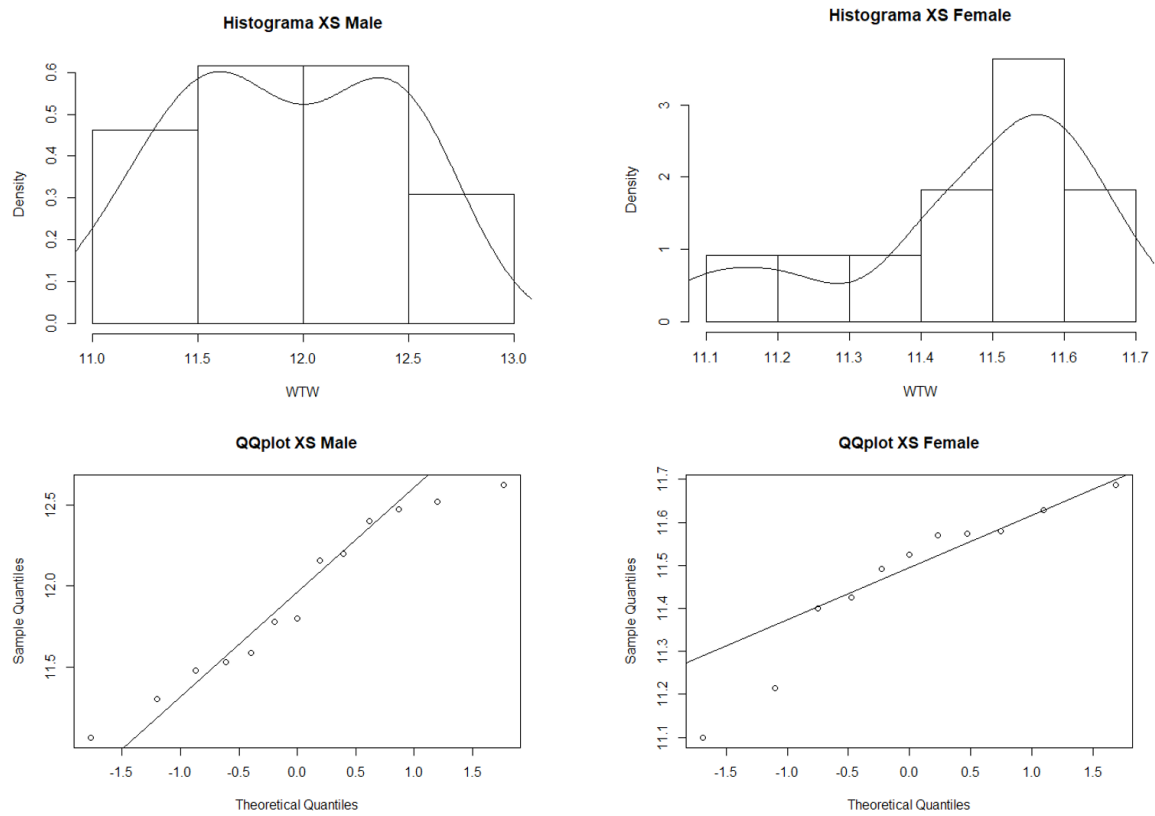


Figura 137. Histogramas y QQplots de normalidad para WTW en ojos extra cortos en ambos sexos.

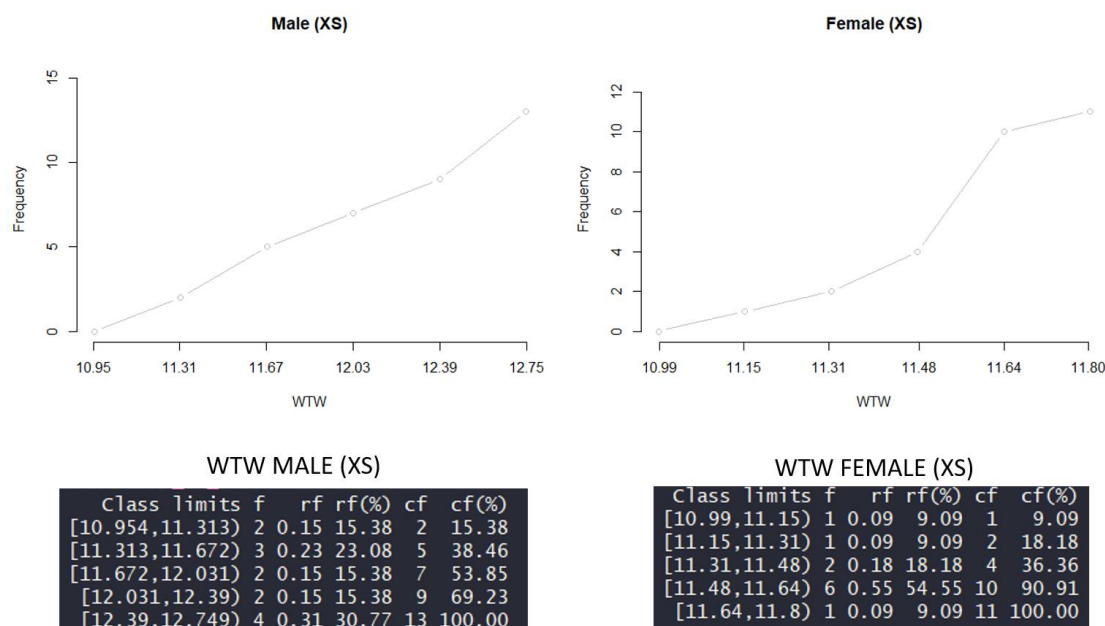


Figura 138. Diagramas de frecuencia polinómico y las tablas de distribución de frecuencias para WTW en ojos extra cortos en ambos sexos.

En hombres el diámetro horizontal corneal oscila entre 10.954 mm y 12.749 mm, sin embargo en mujeres los límites son mucho más estrechos, entre 11.000 mm y 11.800 mm. Parece lógico pensar que en un ojo muy corto axialmente dónde las estructuras anatómicas a priori tendrían que ser más pequeñas, el 46% de los hombres tienen $WTW > 12.031$ mm.

En ojos cortos (S), WTW en hombres se distribuye normalmente (p -value = 0.05204) con un promedio de 11.780 ± 0.438 mm (8.850, 12.900) y en mujeres también la distribución es normal (p -value = 0.665) con un promedio de 11.550 ± 0.417 mm (10.380, 12.930). El valor mínimo de 8.85 mm en hombres es un outlier imputable probablemente a una medida errónea, ya que es el único valor en ese entorno, el siguiente es casi 11.00 mm.

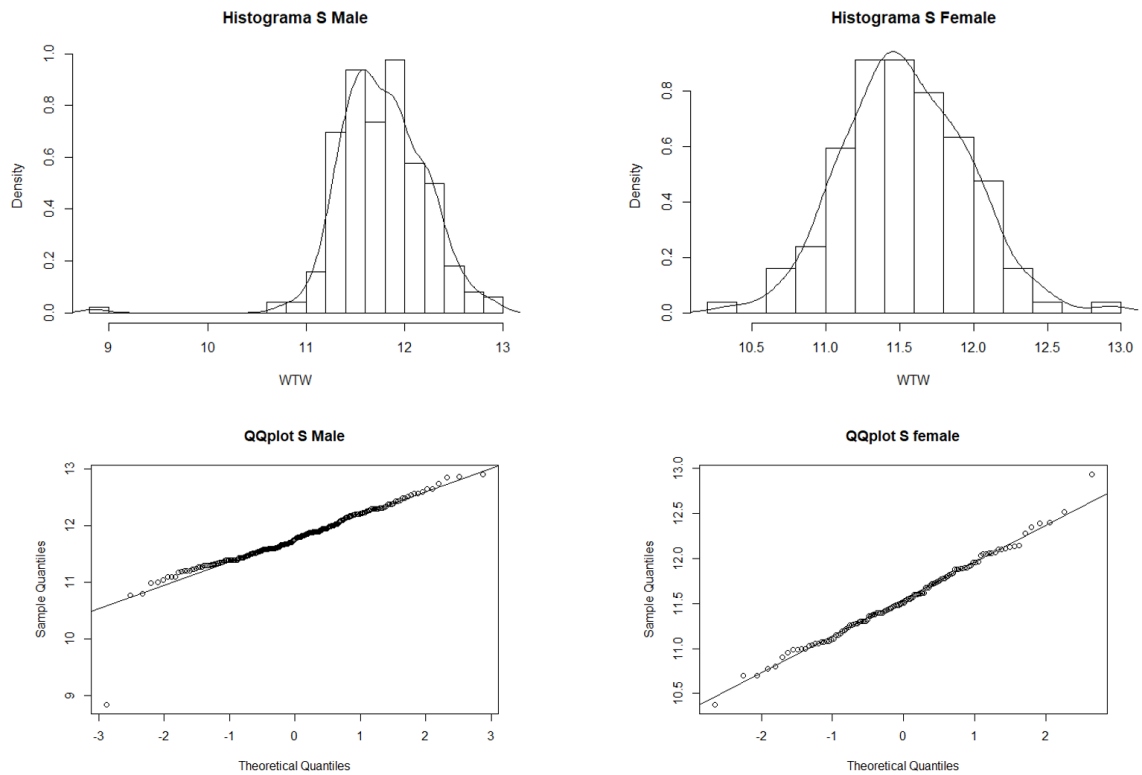


Figura 139. Histogramas y QQplots para ojos cortos en ambos sexos.

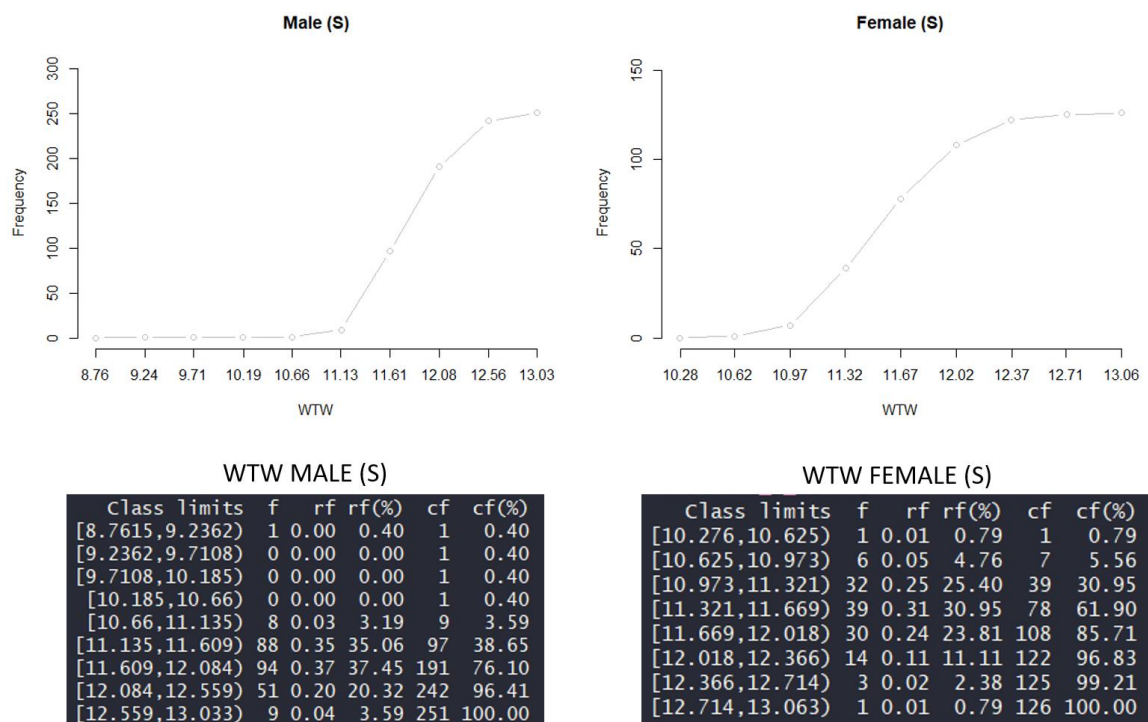


Figura 140. Diagramas de frecuencia polinómico y las tablas de distribución de frecuencias para WTW en ojos cortos en ambos sexos.

En ojos normales (N), el WTW no se distribuye normalmente ni en hombres ni en mujeres ($p\text{-value} = 7.988\text{e-}08$) por la aparición de outliers (Figura 141).

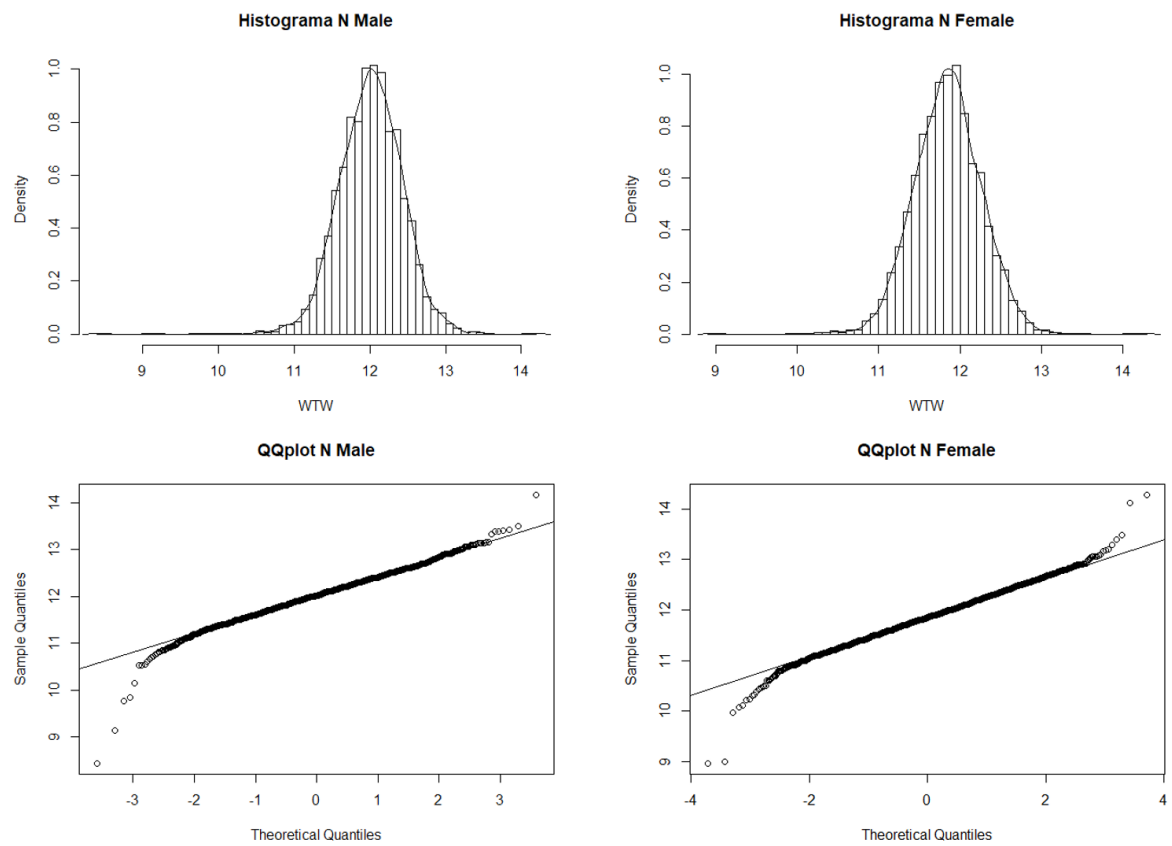
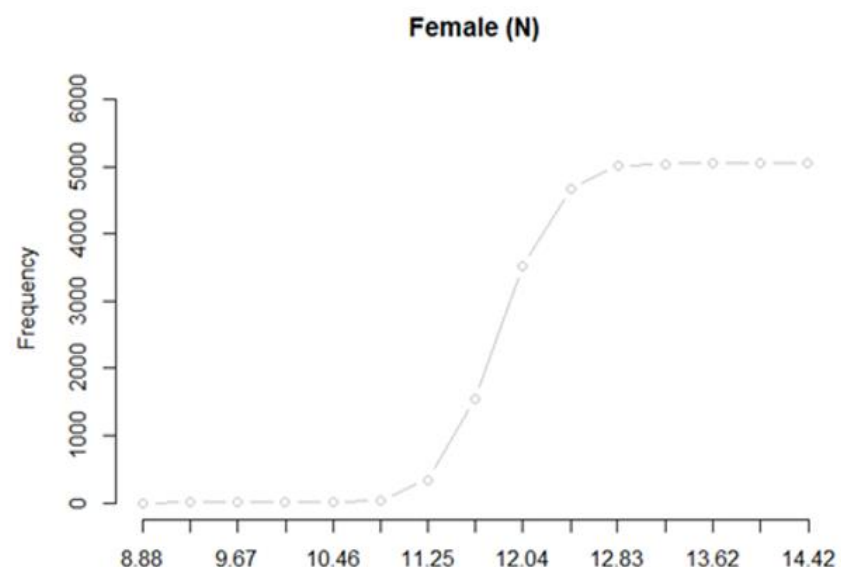
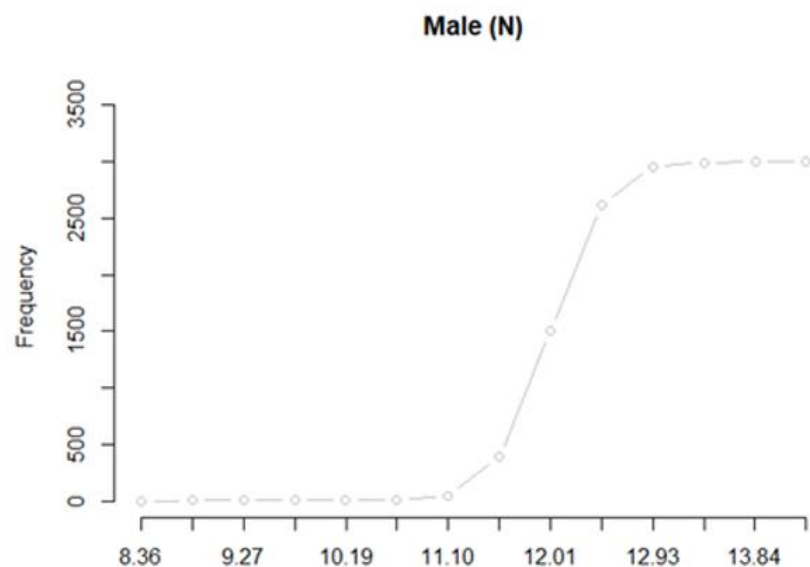


Figura 141. Histogramas de densidad y QQplots de normalidad para WTW en ojos normales en ambos sexos.



WTW MALE (N)

Class limits	f	rf	rf(%)	cf	cf(%)
[8.356,8.813)	1	0.00	0.03	1	0.03
[8.813,9.27)	1	0.00	0.03	2	0.07
[9.27,9.728)	0	0.00	0.00	2	0.07
[9.728,10.19)	3	0.00	0.10	5	0.17
[10.19,10.64)	4	0.00	0.13	9	0.30
[10.64,11.1)	37	0.01	1.23	46	1.53
[11.1,11.56)	339	0.11	11.31	385	12.85
[11.56,12.01)	1114	0.37	37.17	1499	50.02
[12.01,12.47)	1118	0.37	37.30	2617	87.32
[12.47,12.93)	340	0.11	11.34	2957	98.67
[12.93,13.39)	36	0.01	1.20	2993	99.87
[13.39,13.84)	3	0.00	0.10	2996	99.97
[13.84,14.3)	1	0.00	0.03	2997	100.00

WTW FEMALE (N)

Class limits	f	rf	rf(%)	cf	cf(%)
[8.8803,9.2756)	2	0.00	0.04	2	0.04
[9.2756,9.671)	0	0.00	0.00	2	0.04
[9.671,10.066)	1	0.00	0.02	3	0.06
[10.066,10.462)	10	0.00	0.20	13	0.26
[10.462,10.857)	28	0.01	0.55	41	0.81
[10.857,11.252)	303	0.06	6.00	344	6.82
[11.252,11.648)	1197	0.24	23.72	1541	30.54
[11.648,12.043)	1972	0.39	39.08	3513	69.62
[12.043,12.438)	1160	0.23	22.99	4673	92.61
[12.438,12.834)	335	0.07	6.64	5008	99.25
[12.834,13.229)	33	0.01	0.65	5041	99.90
[13.229,13.624)	3	0.00	0.06	5044	99.96
[13.624,14.02)	0	0.00	0.00	5044	99.96
[14.02,14.415)	2	0.00	0.04	5046	100.00

Figura 142. Diagramas polinómicos y tablas de distribución de frecuencias de WTW en ojos normales en ambos sexos.

Para ojos largos (L), WTW se distribuye normalmente tanto en mujeres como en hombres ($p\text{-value} = 0.1016$ y $p\text{-value} = 0.01032$ respectivamente). En hombres con un promedio de 12.240 ± 0.418 mm (10.750, 14.600) y en mujeres de 12.080 ± 0.403 mm (10.580, 13.400).

Independientemente del sexo, el WTW se mueve en valores muy parecidos ya que los rangos de normalidad mínimo y máximo son aproximadamente entre 11.00 mm y 13.00 mm respectivamente, aunque la distribución y estratificación del diámetro horizontal corneal sea diferente discriminando por sexo (Figura 143).

Por último, para ojos extralargos (XL), el diámetro horizontal corneal se distribuye normalmente en hombres ($p\text{-value} = 0.7735$) con un promedio de 12.220 ± 0.413 mm (11.130, 13.280) y en mujeres ($p\text{-value} = 0.1497$) con un promedio de 12.010 ± 0.416 mm (10.900, 13.110).

Se puede observar que el WTW en ojos XL oscila en los mismos rangos aproximadamente que para ojos L (Figura 144).

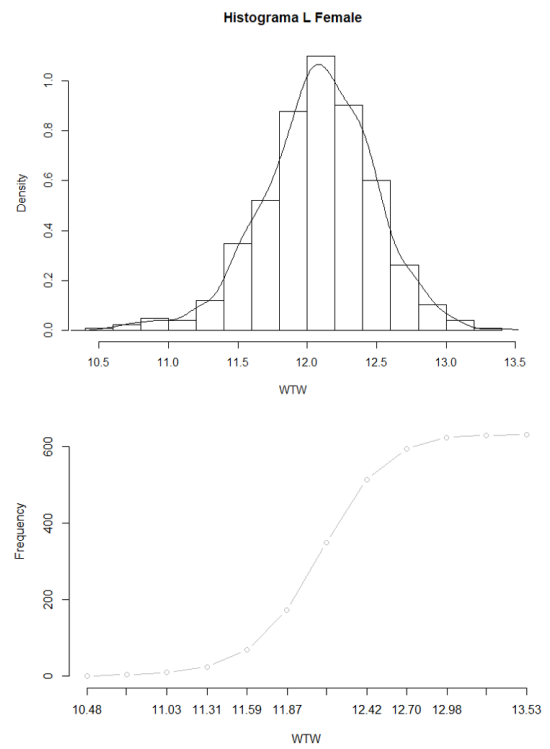
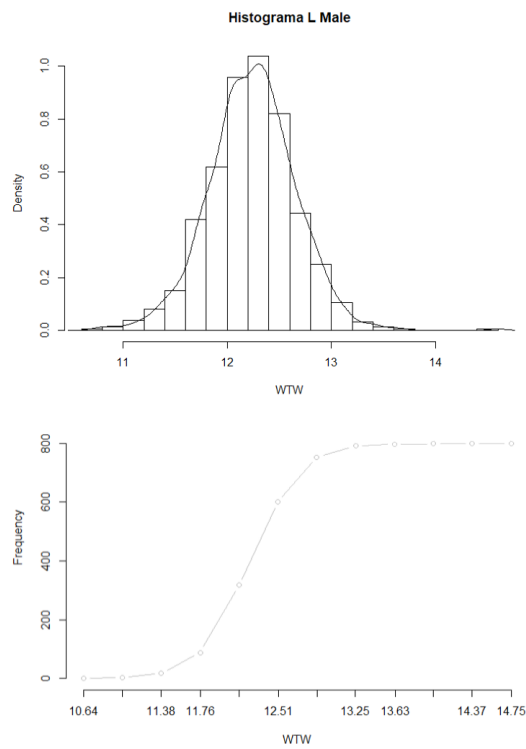
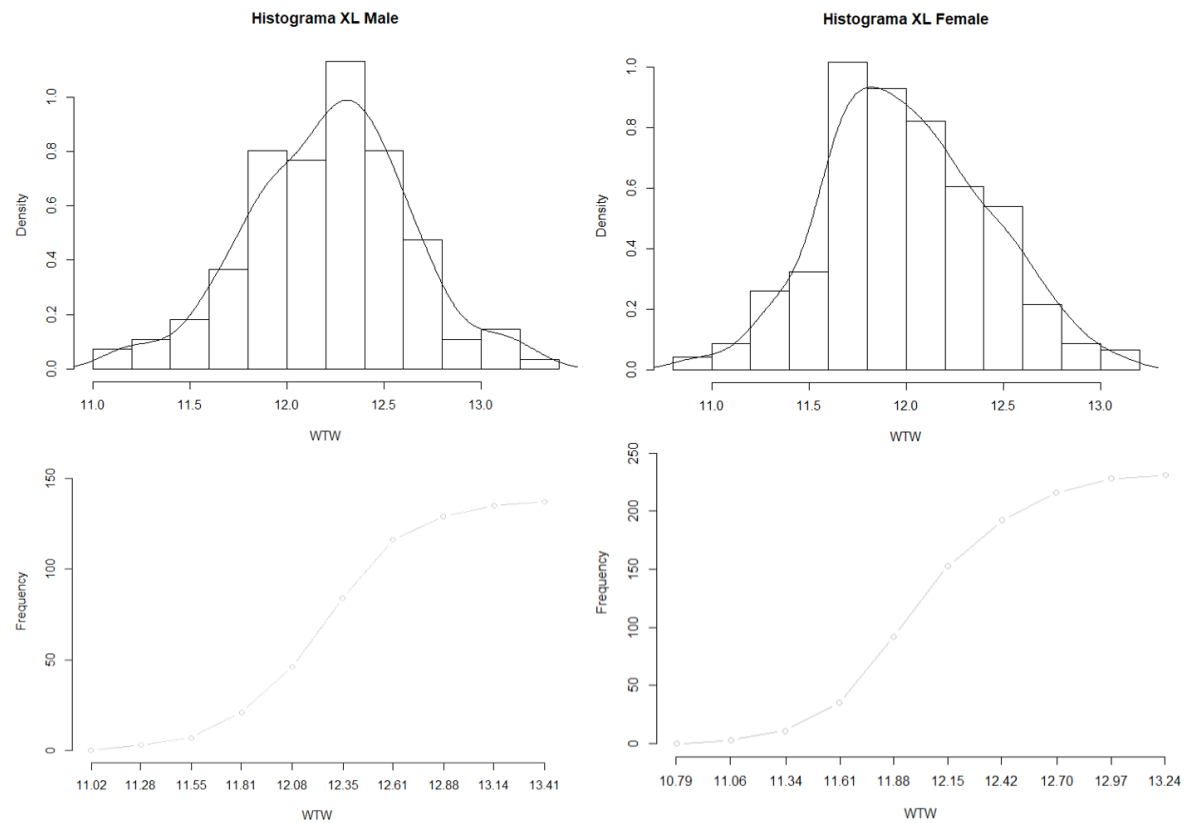


Figura 143. Histogramas, curvas polinómicas y tablas de frecuencia de WTW en ojos largos para hombres y mujeres.



WTW MALE (XL)

Class limits	f	rf	rf(%)	cf	cf(%)
[11.016,11.282)	3	0.02	2.19	3	2.19
[11.282,11.548)	4	0.03	2.92	7	5.11
[11.548,11.814)	14	0.10	10.22	21	15.33
[11.814,12.08)	25	0.18	18.25	46	33.58
[12.08,12.346)	38	0.28	27.74	84	61.31
[12.346,12.612)	32	0.23	23.36	116	84.67
[12.612,12.878)	13	0.09	9.49	129	94.16
[12.878,13.144)	6	0.04	4.38	135	98.54
[13.144,13.41)	2	0.01	1.46	137	100.00

WTW FEMALE (XL)

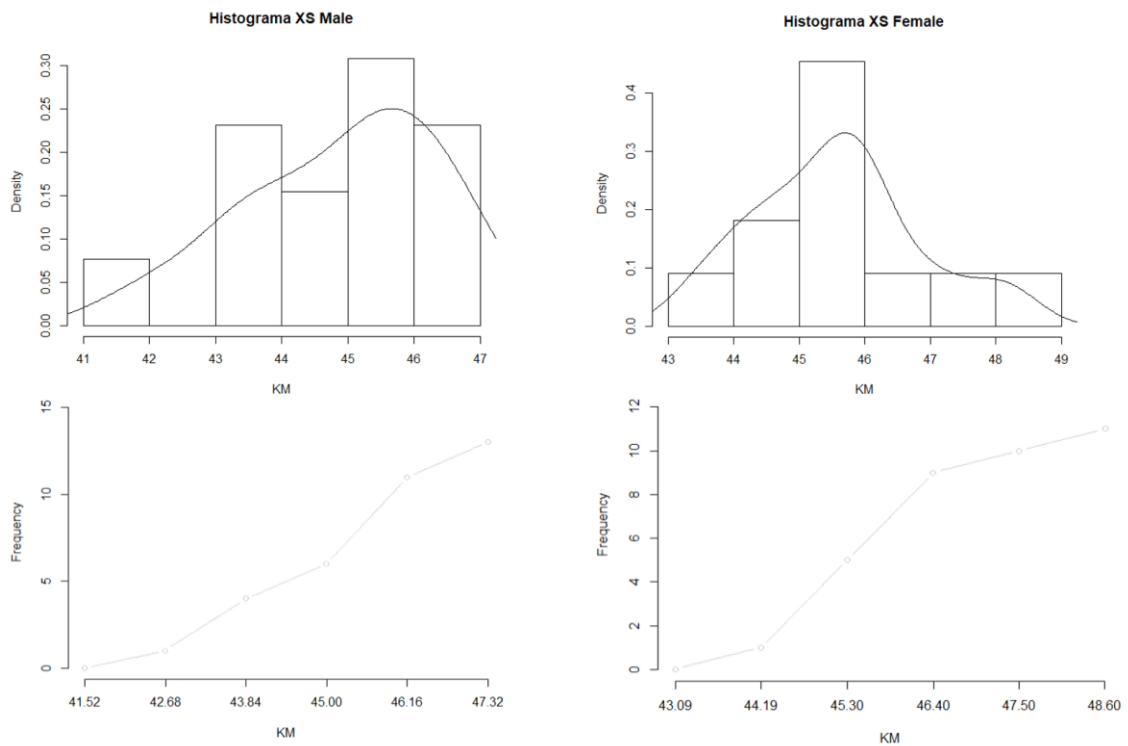
Class limits	f	rf	rf(%)	cf	cf(%)
[10.791,11.063)	3	0.01	1.30	3	1.30
[11.063,11.335)	8	0.03	3.46	11	4.76
[11.335,11.608)	24	0.10	10.39	35	15.15
[11.608,11.88)	57	0.25	24.68	92	39.83
[11.88,12.152)	61	0.26	26.41	153	66.23
[12.152,12.424)	39	0.17	16.88	192	83.12
[12.424,12.697)	24	0.10	10.39	216	93.51
[12.697,12.969)	12	0.05	5.19	228	98.70
[12.969,13.241)	3	0.01	1.30	231	100.00

Figura 144. Histogramas, curvas polinómicas y tablas de frecuencia de WTW en ojos extralargos para hombres y mujeres.

11.3.5.5 Clasificación de KM

En función de la longitud axial, para ojos extracortos (XS) la KM en hombres se distribuye normalmente ($p\text{-value} = 0.6197$) y en mujeres también ($p\text{-value} = 0.9442$). El

promedio de KM en hombres es de 44.910 ± 1.467 D (41.940, 46.850) y en mujeres de 45.580 ± 1.298 D (43.530, 48.120).



KM MALE (XS)

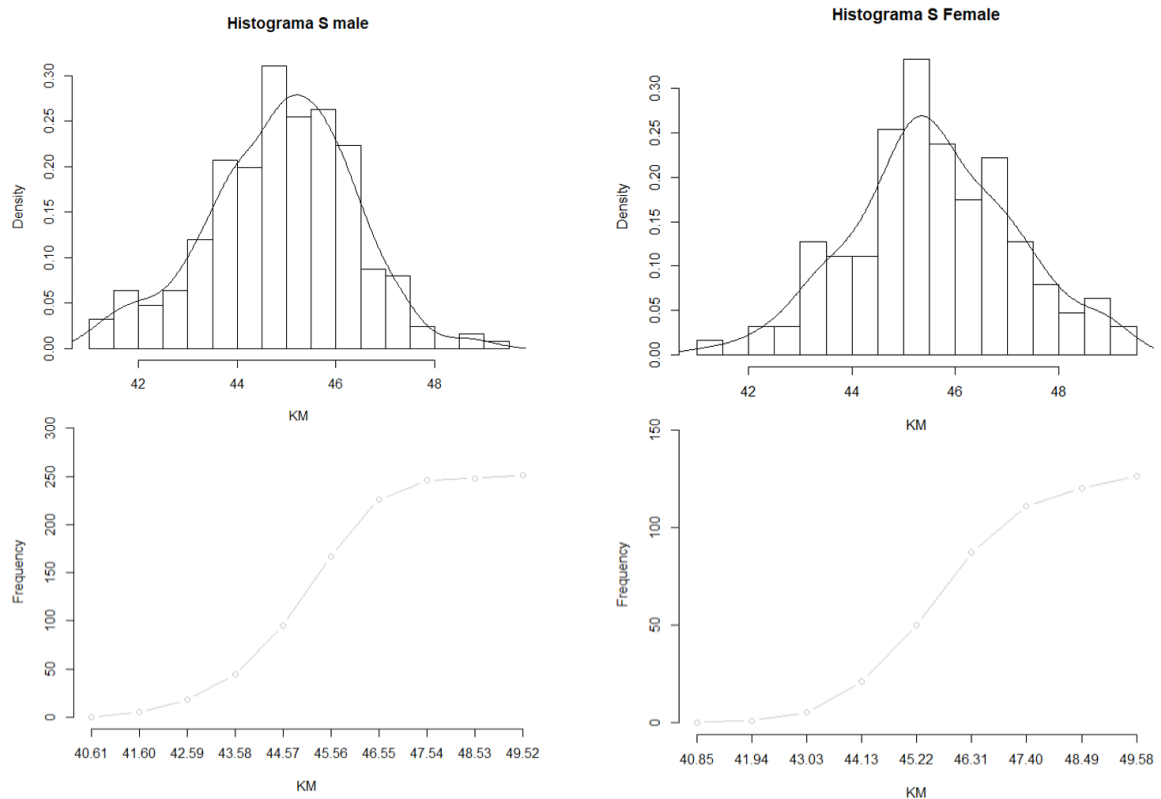
Class limits	f	rf	rf(%)	cf	cf(%)
[41.516,42.676)	1	0.08	7.69	1	7.69
[42.676,43.837)	3	0.23	23.08	4	30.77
[43.837,44.997)	2	0.15	15.38	6	46.15
[44.997,46.158)	5	0.38	38.46	11	84.62
[46.158,47.319)	2	0.15	15.38	13	100.00

KM FEMALE (XS)

Class limits	f	rf	rf(%)	cf	cf(%)
[43.09,44.19)	1	0.09	9.09	1	9.09
[44.19,45.3)	4	0.36	36.36	5	45.45
[45.3,46.4)	4	0.36	36.36	9	81.82
[46.4,47.5)	1	0.09	9.09	10	90.91
[47.5,48.6)	1	0.09	9.09	11	100.00

Figura 145. Histogramas, diagrama polinómico y tablas de frecuencias de KM en ojos extra cortos para ambos sexos.

En ojos cortos (S) la distribución de KM es normal tanto en hombres como en mujeres (p -value = 0.08252 y 0.472 respectivamente). El promedio en hombres es de 44.870 ± 1.464 D (41.020, 49.030) y en mujeres de 45.590 ± 1.572 D (41.270, 49.090).



KM MALE (S)

Class limits	f	rf	rf(%)	cf	cf(%)
[40.61,41.6)	5	0.02	1.99	5	1.99
[41.6,42.59)	13	0.05	5.18	18	7.17
[42.59,43.579)	26	0.10	10.36	44	17.53
[43.579,44.569)	51	0.20	20.32	95	37.85
[44.569,45.559)	72	0.29	28.69	167	66.53
[45.559,46.549)	59	0.24	23.51	226	90.04
[46.549,47.539)	20	0.08	7.97	246	98.01
[47.539,48.529)	2	0.01	0.80	248	98.80
[48.529,49.519)	3	0.01	1.20	251	100.00

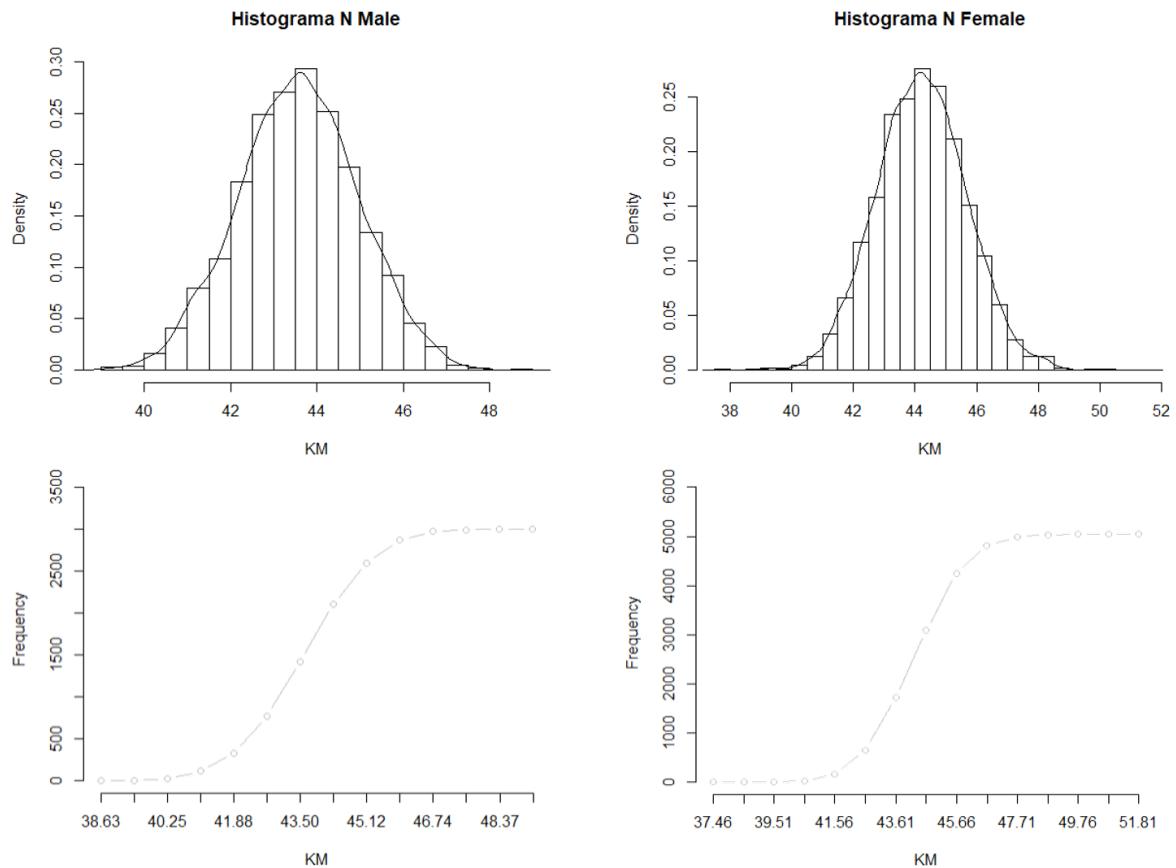
KM FEMALE (S)

Class limits	f	rf	rf(%)	cf	cf(%)
[40.853,41.944)	1	0.01	0.79	1	0.79
[41.944,43.035)	4	0.03	3.17	5	3.97
[43.035,44.126)	16	0.13	12.70	21	16.67
[44.126,45.217)	29	0.23	23.02	50	39.68
[45.217,46.308)	37	0.29	29.37	87	69.05
[46.308,47.399)	24	0.19	19.05	111	88.10
[47.399,48.49)	9	0.07	7.14	120	95.24
[48.49,49.581)	6	0.05	4.76	126	100.00

Figura 146. Histogramas, diagrama polinómico y tablas de frecuencias de KM en ojos cortos para ambos sexos.

El grosor de KM en ojos cortos independientemente del género, se encuentra entre 42.59 D y 47.54 D pero con especial prevalencia de potencias corneales altas entre 44.50 D y 46.55 D, es muy raro encontrar un ojo corto con córneas planas.

En ojos normales (N) la KM tanto en hombres como mujeres se distribuyen normalmente (*p-value* = 0.9499 y 0.3856 respectivamente). El promedio en hombres es de 43.570 ± 1.376 D (39.020, 48.690) y en mujeres de 44.220 ± 1.463 D (37.840, 51.300).



KM MALE (N)

Class limits	f	rf	rf(%)	cf	cf(%)
[38.63,39.441)	3	0.00	0.10	3	0.10
[39.441,40.252)	20	0.01	0.67	23	0.77
[40.252,41.064)	91	0.03	3.04	114	3.80
[41.064,41.875)	219	0.07	7.31	333	11.11
[41.875,42.686)	441	0.15	14.71	774	25.83
[42.686,43.498)	648	0.22	21.62	1422	47.45
[43.498,44.309)	681	0.23	22.72	2103	70.17
[44.309,45.12)	492	0.16	16.42	2595	86.59
[45.12,45.932)	276	0.09	9.21	2871	95.80
[45.932,46.743)	102	0.03	3.40	2973	99.20
[46.743,47.554)	20	0.01	0.67	2993	99.87
[47.554,48.366)	3	0.00	0.10	2996	99.97
[48.366,49.177)	1	0.00	0.03	2997	100.00

KM FEMALE (N)

Class limits	f	rf	rf(%)	cf	cf(%)
[37.463,38.488)	2	0.00	0.04	2	0.04
[38.488,39.512)	6	0.00	0.12	8	0.16
[39.512,40.537)	17	0.00	0.34	25	0.50
[40.537,41.562)	139	0.03	2.75	164	3.25
[41.562,42.587)	489	0.10	9.69	653	12.94
[42.587,43.612)	1080	0.21	21.40	1733	34.34
[43.612,44.637)	1357	0.27	26.89	3090	61.24
[44.637,45.661)	1163	0.23	23.05	4253	84.28
[45.661,46.686)	565	0.11	11.20	4818	95.48
[46.686,47.711)	174	0.03	3.45	4992	98.93
[47.711,48.736)	46	0.01	0.91	5038	99.84
[48.736,49.761)	4	0.00	0.08	5042	99.92
[49.761,50.785)	3	0.00	0.06	5045	99.98
[50.785,51.81)	1	0.00	0.02	5046	100.00

Figura 147. Histogramas, diagrama polinómico y tablas de frecuencias de KM en ojos normales para ambos sexos.

En ojos axialmente largos (L), KM muestra tanto en hombres como mujeres una distribución no normal ($p\text{-value} = 0.00812$ para hombres y $p\text{-value} = 0.0001945$ para mujeres) tal y como se puede observar en la Figura 147.

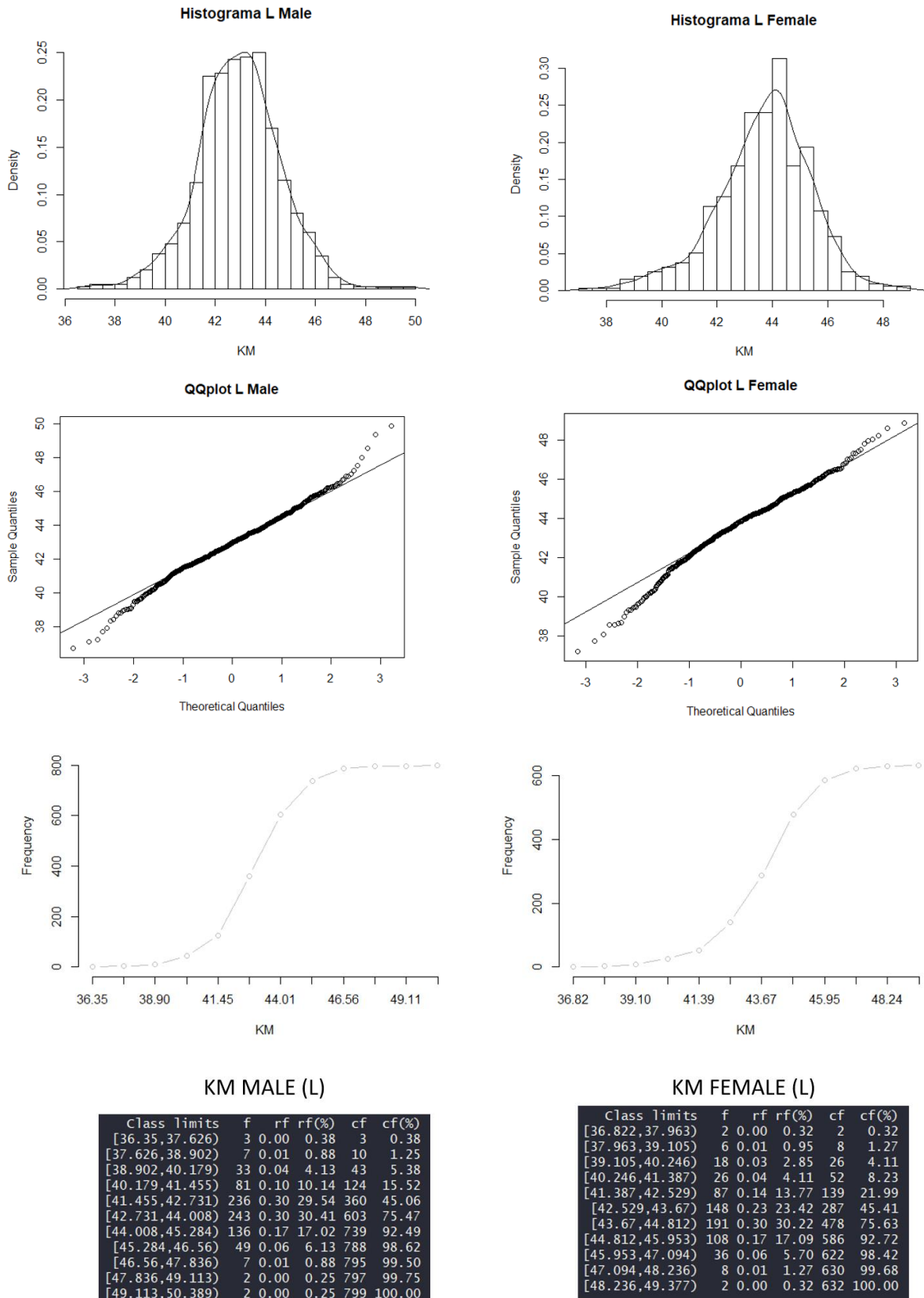


Figura 148. Histogramas, diagramas polinómicos, QQplots y tablas de frecuencias de KM en ojos largos para ambos sexos.

Por último, en los ojos extralargos (XL) la KM para hombres la distribución es cuasi normal ($p\text{-value} = 0.06963$) debido a unos outliers de potencias corneales muy bajas por debajo de 39.89 D (Figura 149). En mujeres pasa algo similar, la distribución aunque sale normal ($p\text{-value} = 0.8712$), el QQplot muestra outliers en potencias bajas.

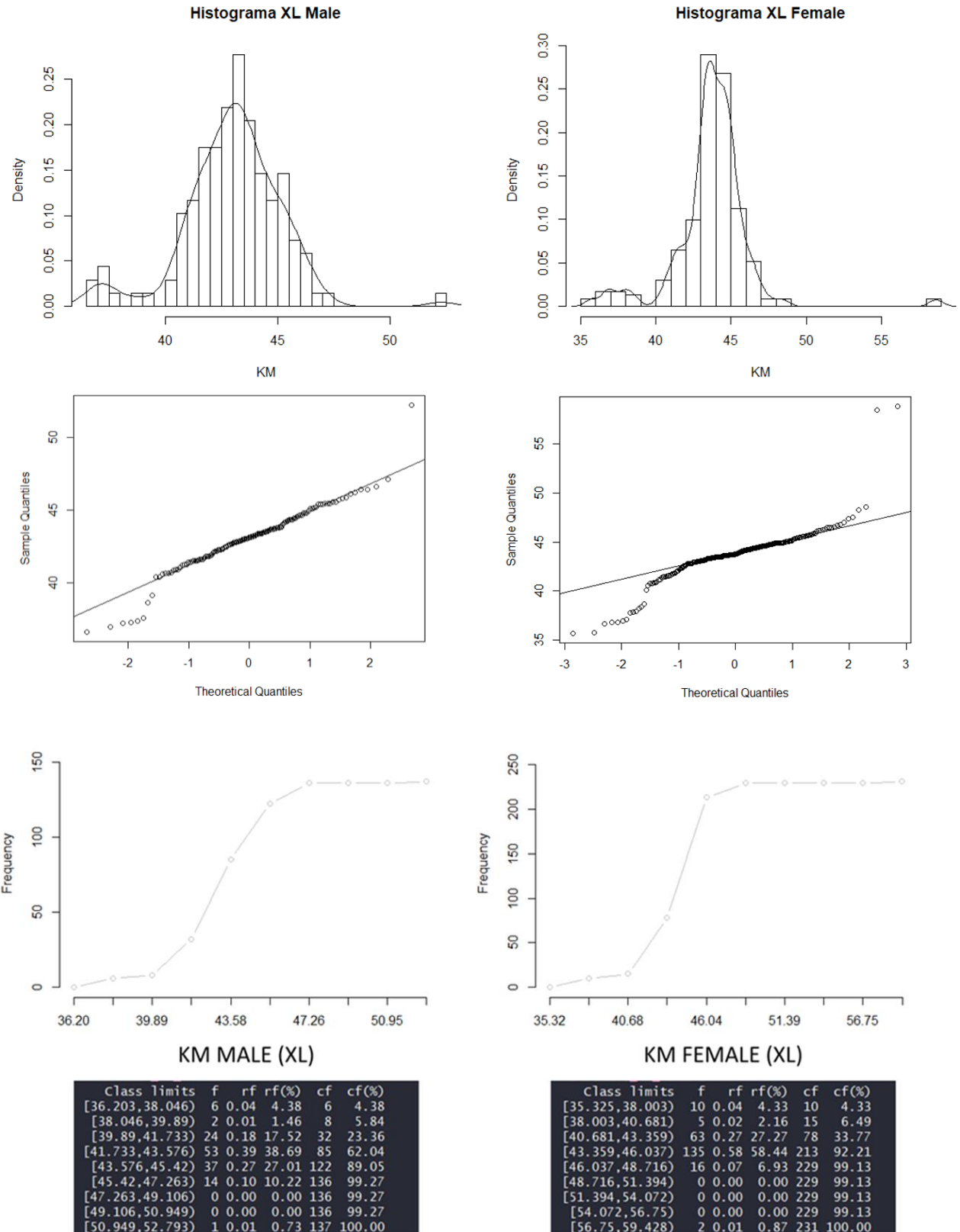


Figura 149. Histogramas, diagramas polinómicos, QQplots y tablas de frecuencias de KM en ojos extra cortos para ambos sexos.

11.3.6 Instauración de un índice de singularidad

Una vez organizadas todas las variables biométricas, se pretende instaurar un índice teórico que permita cuantificar la rareza o infrecuencia anatómica, englobando la información de todas las variables biométricas medidas para un ojo en cuestión.

Se le llamará SI (*Singularity Index*). Cuanto más alto sea SI más atípico será el ojo, pudiendo oscilar entre 0 y 100.

Por simplicidad cada uno de los cinco parámetros oculares medidos (AXL, ACD, LT, WTW y KM) se corresponderá con un 20% de rareza, de modo que computando al máximo en dos categorías, IS será del 40%, si fueran raras tres de las variables biométricas IS sería 60% y así sucesivamente. Esto implica que un SI de 100% representará un ojo extramadamente atípico ya que las cinco variables biométricas serán infrecuentes.

Cada una de las cinco variables biométricas presentará a su vez un grado de infrecuencia, es decir, no será lo mismo un hombre con un ojo corto que con un ojo extracorto, en ambos casos la longitud axial es atípica pero mucho más en el segundo caso. Esto obliga a escalar de manera graduada la cuota del 20% que le corresponde a cada categoría de SI. Por ejemplo, si se trata la AXL, el 20% es para un ojo XS, el 15% para un ojo S, el 10% para un ojo XL, el 5% para un ojo L y el 0% para un ojo N. Para el resto de categorías biométricas se asimilará una distribución normal de manera que el 68.4% de los datos están dentro de una $\pm SD$, el 27.2% en dos $\pm SD$, el 4.2% en tres $\pm SD$ y el 0.2% en más de tres $\pm SD$.

Para representar y explicar lo anterior un ejemplo será suficiente. Imagínese una mujer que presenta un ojo con los siguientes parámetros biométricos: AXL = 20.35 mm,

ACD = 2.15 mm, LT = 3.66 mm, WTW = 11.40 mm y KM = 47.41 D. Históricamente este ojo se calificaría como corto sin más independientemente del género y del resto de parámetros biométricos, pero ¿qué hay detrás del resto de variables biométricas y cómo se relacionan con la longitud axial? ¿Cuál será el Índice de Singularidad? ¿Es un ojo realmente atípico por su longitud axial o por algo más?

En primer lugar, si se atiende a la clasificación segmentada propuesta en este anexo de la tesis, una longitud axial de 20.35 mm en una mujer corresponde a un ojo corto (≥ 20.12 mm y < 21.34 mm) con una incidencia únicamente del 2.10 %. De modo que en términos de axialidad este ojo es atípico, infrecuente y singular, así que computa como tal con un 15% de SI (Figura 150) como se explicó al comienzo del apartado.

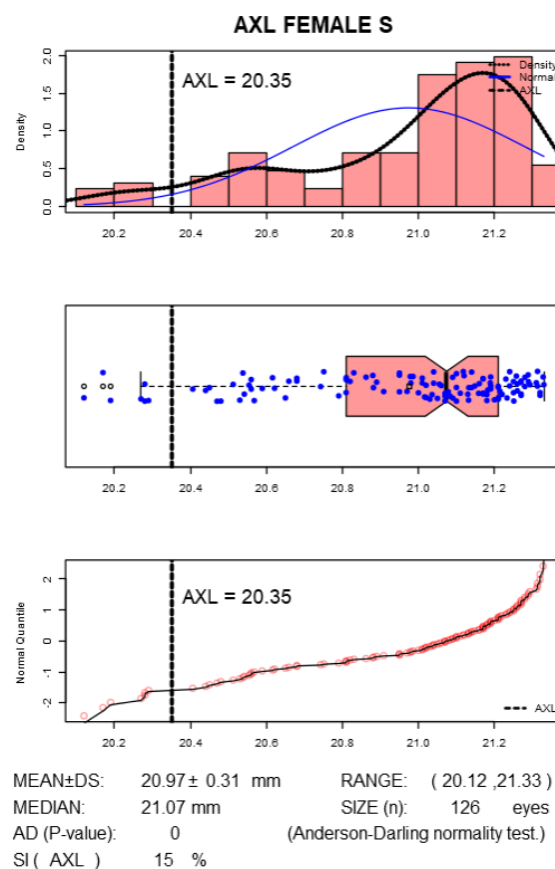


Figura 150. Representación triple (histograma + boxplot + qqnorm) de la ubicación de una AXL de 20.35 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de axialidad.

Continuando con la ACD, en mujeres con ojos cortos (S), 2.15 mm está a más de una DS del promedio poblacional, efectivamente solo hay una incidencia del 4.76% según la tabla de frecuencias. En este caso ACD también computará como singular, pero antes hay que cuantificar el grado de infrecuencia. En una distribución normal 2.15 mm está en la cola izquierda dentro del 13.6%, es decir, sin superar las 2 SD, lo que implica que computará dentro de SI como un 6.66% teniendo en cuenta que habrá cuatro posibles zonas en ambas colas (Figura 151), totalizando un SI de 21.66% de momento entre AXL y ACD.

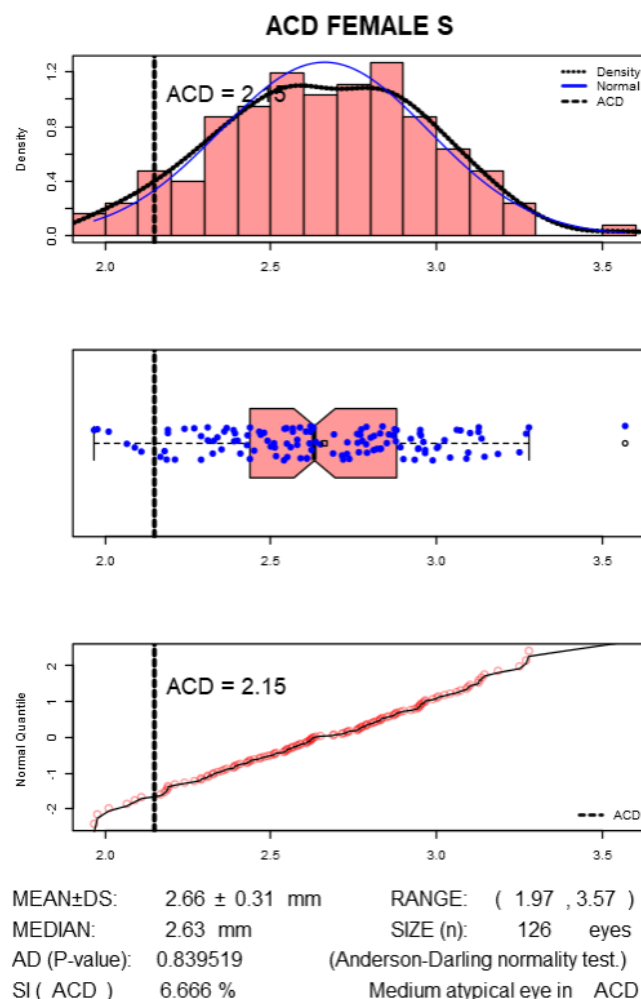


Figura 151. Representación triple (histograma + boxplot + qqnorm) de la ubicación de una ACD de 2.15 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de profundidad de cámara anterior.

En cuanto al grosor de cristalino, un LT de 3.66 mm es rarísimo para el rango axial al que pertenece, teniendo en cuenta que en mujeres con ojos cortos el promedio de LT es 4.625 ± 0.388 mm (3.704, 5.697). De manera que al ser distribución normal y estar alejado mas de 2 SD de la media, solo el 2.1% de la población tendrá este valor (Figura 152). Computará dentro de SI con un 13.333% sumándose a 21.66%, obteniéndose un 35.332% de SI.

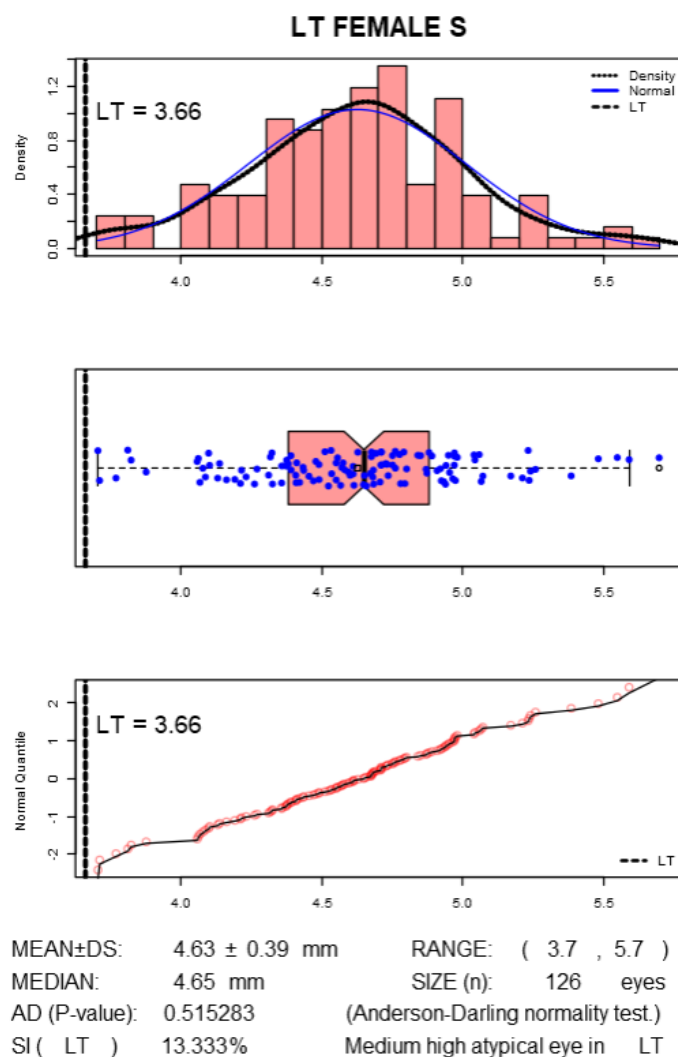


Figura 152. Representación triple (histograma + boxplot + qqnorm) de la ubicación de un LT de 3.66 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría espesor de cristalino.

Solo quedan dos variables biométricas del ejemplo propuesto, WTW y KM. Un WTW de 11.40 mm implica que es normal en esa tipología de ojo (Figura 153), por tanto no agrega porcentaje de rareza a IS que se mantendrá de momento en 41.66%.

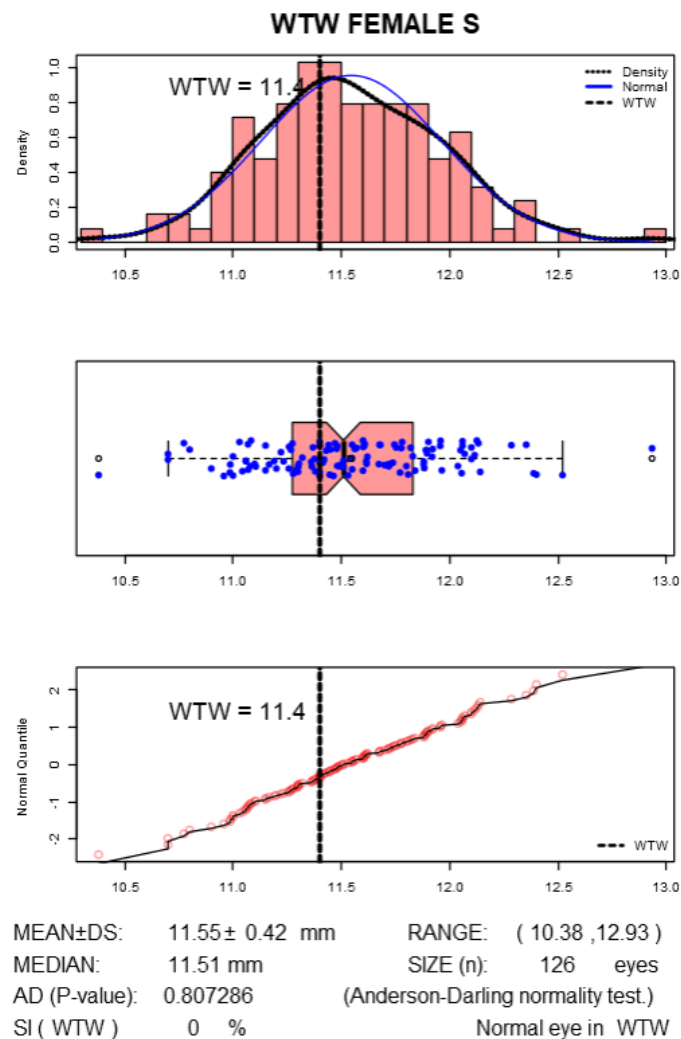


Figura 153. Representación triple (histograma + boxplot + qqnorm) de la ubicación de un WTW de 11.40 mm sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de diámetro horizontal corneal.

Por último una KM de 47.41 D teniendo en cuenta que en mujeres con ojos cortos el promedio de KM es 45.590 ± 1.572 D (41.270, 49.090) y que la distribución de datos es normal, 47.41 D está entre 1DS y 2DS (Figura 154). De manera que este valor se

localizará dentro del 13.6% de la población en distribuciones normales hacia la cola derecha y por tanto el porcentaje de rareza correspondiente a KM será de 6.66%.

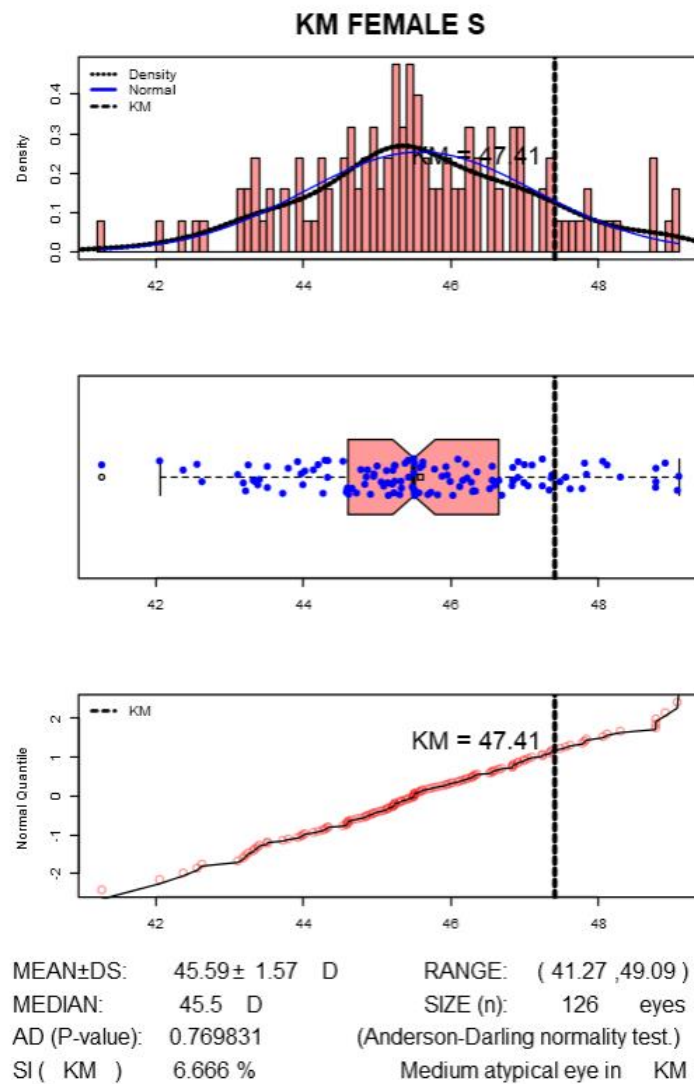


Figura 154. Representación triple (histograma + boxplot + qqnorm) de la ubicación de un KM de 47.41 D sobre la distribución real de los datos en ojos cortos de mujeres. Se muestran todas las métricas de centralidad y dispersión, así como el índice de singularidad para la categoría de queratometría media con índice queratométrico 1.3375.

El computo general de IS tras sumar el valor de KM será de 42% que determinará el grado de singularidad o de rareza del ojo en cuestión (Figura 155).

SINGULARITY INDEX (%)	
TOTAL SI =	42 %
SI (AXL):	15 %
SI (ACD):	6.67 %
SI (LT):	13.33 %
SI (WTW):	0 %
SI (KM):	6.67 %

Figura 155. Desglose del índice de singularidad (SI) por categorías anatómicas y el sumatorio de todas representando el grado de rareza anatómica del ojo estudiado.

En conclusión, manejando SI como parámetro que pondera el grado de rareza se puede instaurar lo atípico que es un ojo anatómicamente hablando, así como comparar ojos entre sí, establecer ordenaciones de infrecuencia y sentar posibles correlaciones con las diferentes fórmulas de cálculo biométrico.

11.3.7 Aplicación web para clasificar un ojo por infrecuencia

Se ha desarrollado una aplicación web susceptible de incorporación a Karmona® iol power calculation que permite clasificar el grado de rareza biométrica calculando el índice de singularidad de un ojo en función de los parámetros biométricos introducidos por el usuario para el cálculo de la potencia de la lente intraocular.

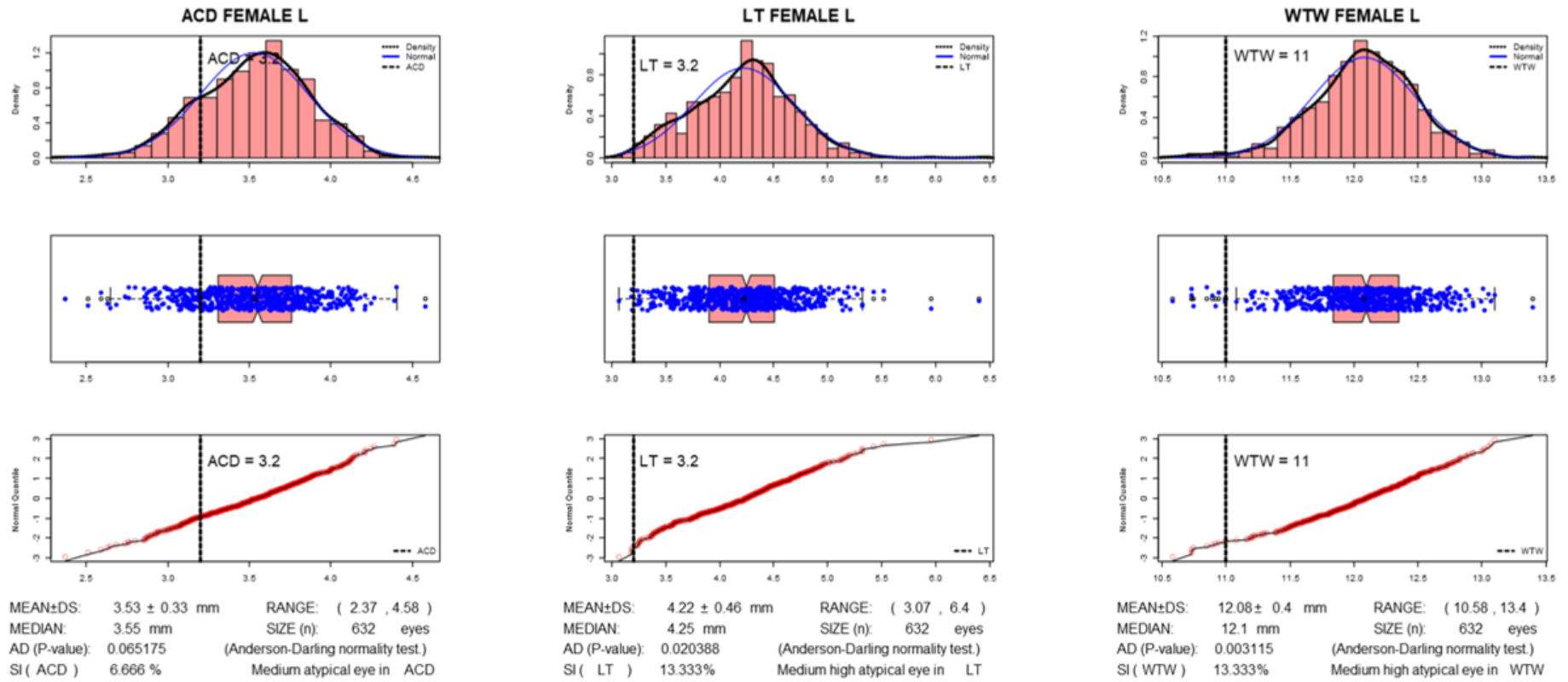


Figura 156. Captura de pantalla de la aplicación web diseñada para evaluar el índice de singularidad de un ojo (1ª parte).

SINGULARITY INDEX (%)

TOTAL SI = 45 %

SI (AXL): 5 %

SI (ACD): 6.67 %

SI (LT): 13.33 %

SI (WTW): 13.33 %

SI (KM): 6.67 %

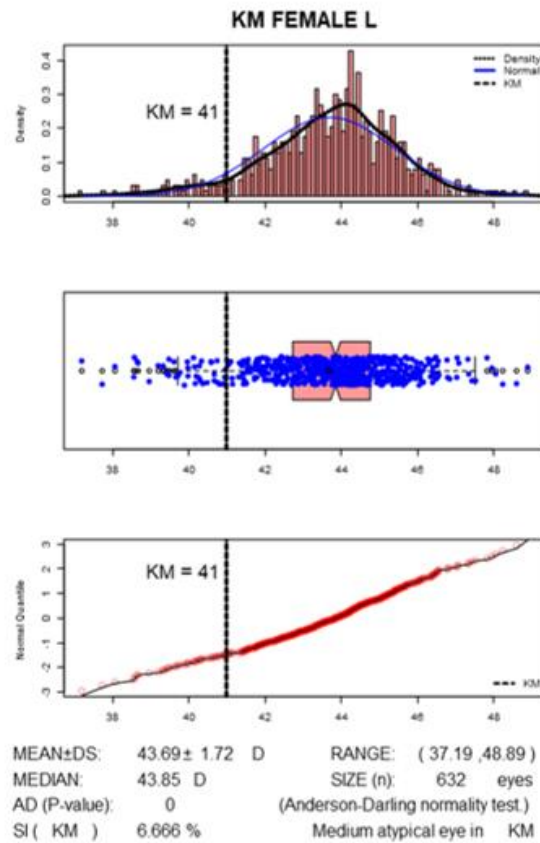
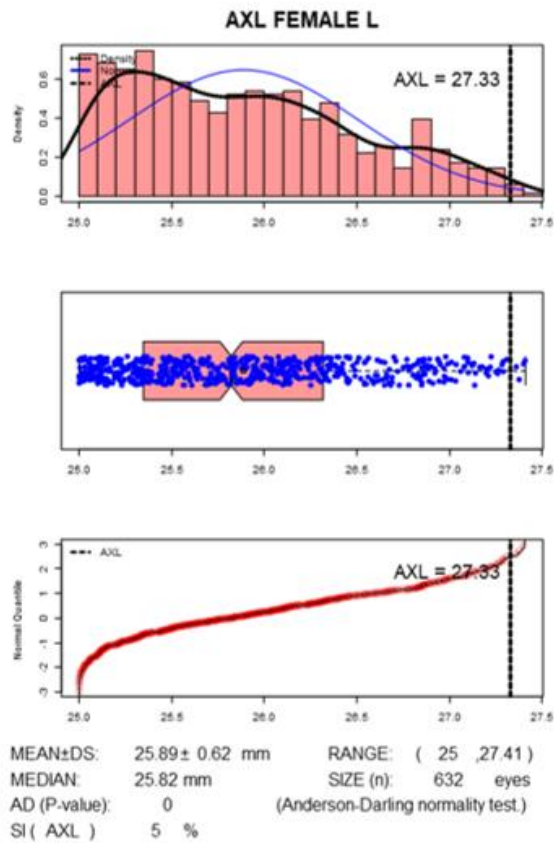


Figura 157. Captura de pantalla de la aplicación web diseñada para evaluar el índice de singularidad de un ojo (2ª parte).

En este ejemplo se tiene un SI del 45% con el desglose de SI para cada variable biométrica en la columna de la derecha, concluyendo que es un ojo singular en todas las categorías, dando lugar a un ojo bastante atípico desde el punto de vista anatómico.

11.4 ANEXO 4.- APORTACIONES CIENTÍFICAS DERIVADAS DE LA TESIS DOCTORAL

- **Publicaciones en revistas indexadas**

1. Carmona González, D., Palomino Bautista, C. Accuracy of a new intraocular lens power calculation method based on artificial intelligence. *Eye* 35, 517–522 (2021).
2. Carmona-González D, Castillo-Gómez A, Palomino-Bautista C, Romero-Domínguez M, Gutiérrez-Moreno MÁ. Comparison of the accuracy of 11 intraocular lens power calculation formulas. *Eur J Ophthalmol.* 2020 Oct 15:1120672120962030.

- **Capítulos en libros**

David Carmona González. “Fórmulas de cálculo biométrico basadas en Inteligencia Artificial”. Monografía SECOIR 2021. (*In press*).

- **Cursos impartidos**

1. David Carmona. “Biometría. En busca de la predictibilidad”. SIFI. Mayo 2020.

2. David Carmona. "Pruebas preoperatorias para cirugía de cataratas". SIFI. Mayo 2020.
3. David Carmona. "Fórmulas de cálculo de lente intraocular". Sesión Online Hospital clínico San Carlos. Marzo 2021. Madrid.

- **Congresos internacionales**

1. Palomino C, Carmona D. "Aproximación al desarrollo de un nuevo método de cálculo biométrico basado en Inteligencia Artificial". Asociación Latinoamericana de Cirujanos de Catarata, Segmento Anterior y Refractiva. Marzo 2019. Bogotá. Colombia.

- **Software generado**

Karmona® IOL Power Calculation: página web de acceso libre que contiene el algoritmo Karmona® para calcular la potencia emetropizante de la lente intraocular que se se implantará en la cirugía de cristalino. (www.Karmona®-iol.com).

11.5 ANEXO 5.- APROBACIÓN COMITÉ ÉTICO

INFORME DEL COMITE DE ÉTICA DE LA INVESTIGACION

Dra. Lucía Llanos Jiménez, Secretaria Técnica del COMITE DE ÉTICA DE LA INVESTIGACION DE LA FUNDACION JIMENEZ DIAZ

CERTIFICA:

Que en la reunión del CEIm-FJD que tuvo lugar el 09 de Octubre de 2018 (acta nº 18/18) se evaluó el estudio referido y, ha decidido, una vez aceptadas las respuestas a las aclaraciones menores:

A P R O B A R

La propuesta para que se realice el estudio titulado: **“Cálculo de una lente intraocular mediante técnicas de Machine Learning”**. Version de Protocolo: noviembre 2018

Además, hace constar que:

1. En dicha reunión se cumplieron los requisitos establecidos en la legislación vigente –Real decreto 1090/2015 y Decreto 39/94 de la CAM– para que la decisión del citado CEIm sea válida.
2. El Estudio reúne las normas éticas estándar de nuestra Institución para la realización de este tipo de estudios.
3. Se cumplen los preceptos éticos formulados en la Orden SAS 3470/2009 y la Declaración de Helsinki de la Asociación Médica mundial sobre principios éticos para las investigaciones médicas en seres humanos y en sus posteriores revisiones, así como aquellos exigidos por la normativa aplicable en función de las características del estudio.
4. El CEImFJD, tanto en su composición como en sus procedimientos, cumple con las normas de BPC (CPMP/ICH/135/95) y con la legislación vigente que regula su funcionamiento, y que la composición del CEIm FJD es la indicada en el anexo I, teniendo en cuenta que en el caso de que algún miembro participe en el estudio o declare algún conflicto de interés no habrá participado en la evaluación ni en el dictamen.
5. Asimismo, hacemos constar que no existe contraprestación económica para el centro ni para los investigadores.
6. Además, este comité recuerda la obligación de realizar el seguimiento del estudio de acuerdo a la legislación vigente.

Lo que firmo en Madrid a 8 de noviembre de 2018



Dra. Lucía Llanos Jiménez
 Secretaria Técnica CEImFJD

Anexo I COMPOSICIÓN DEL CEIm

Presidente	Dr. Javier Bécares Martínez. Farmacéutico de FJD
Vicepresidenta	Dra. Macarena Bonilla Porras. Farmacéutica de FJD
Secretaria Técnica	Dra. Lucía Llanos Jiménez. Farmacóloga de la FJD
Vocales	Dr. Nicolás Alejandro Alba. Médico Asistencial. Oftalmología de FJD
	Dra. Miriam Blanco Rodríguez. Médico Asistencial. Pediatría de FJD
	Dr. Alfonso Cabello Ubeda. Médico Asistencial. Medicina Interna FJD
	Dr. Raúl Cordoba Mascuñano. Médico Asistencial. Hematología FJD
	Dra. Ana Díez Alcántara. Farmacéutico de Atención Primaria. Dirección Asistencial Noroeste
	Dra. Isabel Egocheaga Cabello. Médico de Atención Primaria. CS Isla de Oza.
	Dña. Ana García Díaz, abogada
	Dr. Bernard Gastón Doger de Speville. Médico Asistencial. Oncología FJD
	Dr. Sergio Hoyos Simón. Médico Asistencial. Oncología HRJC
	Dña. María del Mar Jiménez del Castillo. DUE de la FJD
	D. Carlos Macía de la Asunción. Lego no vinculado a la Institución
	Dr. Ricardo Pardo García. Médico Asistencial. Cirugía Digestivo de FJD
	Dr. Gonzalo Pizarro Sánchez. Médico Asistencial. Cardiología HUQM
	Dr. Francisco Javier Ruiz Hornillos. Miembro Comité de Investigación y Miembro del Comité de Ética Asistencial. Médico Asistencial. Alergología del HIE
	Dra. Olga Sánchez Pernaute. Médico Asistencial. Reumatología de la FJD
	Dra. Sandra Zazo Hernández. Bióloga. Anatomía Patológica de la FJD

Anexo II

CENTROS E INVESTIGADORES PRINCIPALES Y COLABORADORES

Título “Cálculo de una lente intraocular mediante técnicas de Machine Learning” Version de Protocolo: noviembre 2018

Fecha de actualización del anexo II: 8 de noviembre de 2018

INVESTIGADOR PRINCIPAL	INVESTIGADORES COLABORADORES	CENTROS DE REALIZACIÓN
David Carmona González		Hospital U. Quirón Madrid
		Servicio Oftalmología

EO153-18_HUQM

11.6 ANEXO 6.- INFORME DE ORIGINALIDAD

A continuación se presenta el informe final de originalidad reportado por Turnitin™:

