UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS

Departamento de Estadística e Investigación Operativa

TESIS DOCTORAL

Clasificacion de distribuciones y datos atipicos [Manuscritio]

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Paloma Main Yaque

DIRECTOR:

Miguel Ángel Gómez Villegas

Madrid, 2015

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Ciencias Matemáticas

Departamento de Estadística e Investigación Operativa

CLASIFICACION DE DISTRIBUCIONES Y DATOS ATIPICOS

Paloma Main Yaque

Madrid, 1988

Colección Tesis Doctorales. N.º 226/88

Edita e imprime la Editorial de la Universidad
Complutense de Madrid. Servicio de Reprografía
Noviciado, 3 - 28015 Madrid
Madrid, 1988
Ricoh 3700
Depósito Legal: M-7988-1988

CLASIFICACION DE DISTRIBUCIONES

Y DATOS ATIPICOS

Paloma Main Yaque

Memoria para optar al grado de
Doctor en Ciencias Matemáticas
realizada bajo la dirección del
Dr. D. Miguel Angel Gómez Villegas

Madrid, Mayo de 1986

INDICE

PROLOGO

PROLOGO

En esta memoria se analiza el problema de los datos atípicos desde un punto de vista diferente al usual en este tipo de situaciones.

Aunque en el capítulo I daremos cumplido espacio a la discusión del concepto y su tratamiento, se considera en general que dato atípico es aquél que, dentro de una muestra, nos sorprende de alguna manera. Con nuestro enfoque pretendemos educar esa sorpresa antes de extraer cualquier muestra, suponiendo un modelo para la población.

El método utilizado es la clasificación de dichos modelos subyacentes, según su facilidad para producir datos atípicos, en el sentido de maximos extremadamente grandes ó mínimos muy pequeños respecto a los demás valores presentes en la muestra.

Evidentemente si consideramos las funciones de distribución, todo dependerá de la forma en que convergen los extremos.

Son pieza fundamental en nuestra aportación los conceptos de función de variación regular, lenta y rápida introducidos por Karamata (1933), que aplicaremos a la variación de la cola derecha de las funciones de distribución para discriminar entre los diferentes comportamientos. Todos los resultados son trivialmente trasladables a la cola izquierda de la función de distribución.

Así se desarrollarán estas ideas en el capítulo II partiendo de una primera clasificación efectuada por Green (1976) que también hemos completado y relacionado con los dominios de atracción de los límites de máximos normalizados. Utilizaremos para nuestras definiciones el comportamiento asintótico de la diferencia y la razón entre los dos extremos de la muestra, resultando ésta última la más potente para reflejar el alejamiento del máximo , inversamente con la primera.

Con este planteamiento pasamos en el capítulo III a analizar las definiciones obtenidas, en los elementos que conforman un problema bayesiano. En un primer apartado mantenemos la muestra fija y estudiamos que situaciones pueden aparecer en el paso de la distribución a priori a la distribución a posteriori, llegando a conclusiones totalmente congruentes con la idea lógica de que con una muestra como mínimo mantenemos nuestro conocimiento acerca del parámetro.

A continuación consideramos el posible efecto de la forma de la verosimilitud en la diferencia entre distribución a priori y a posteriori, que es como se trata el problema en los trabajos de Dawid (1973) y O'Hagan (1979). No obstante en lugar de utilizar límites de esperanzas a posteriori para una clase de funciones ó conceptos de dominancia estocástica como ellos hacen, analizamos el comportamiento asintótico de la divergencia de Kullback-Leibler entre las distribuciones a priori y a posteriori como aproximación mas correcta desde nuestro punto de vista.

Los resultados obtenidos concuerdan con los trabajos citados y la idea primera debida a de Finetti (1961), de que las observaciones anómalas, en distribuciones propensas a producirlas, se desacreditan a si mismas al converger la distribución a posteriori a la distribución a priori cuando la observación crece hacia infinito. Nosotros consideramos esta aproximación en el sentido de que la divergencia de Kullback--Leibler entre las mismas tiende a anularse cuando la verosi militud es propensa.

También estudiamos la convergencia de esta medida para verosimilitudes de los demas tipos, consiguiendo reflejar la separación total entre ambas para resistentes, con una divergencia que crece indefinidamente.

Consideramos toda esta problemática con parámetro de localización, siguiendo la linea de los trabajos citados y además la extendemos al caso de parámetros de escala.

Por último en el capítulo IV hacemos una aplicación a funciones de distribución en el semieje positivo, propias de teoría de la fiabilidad y relacionamos los conceptos de razón de fallo creciente y decreciente con la resistencia y propensión respectivamente. Incluso generalizando las definiciones a las de NBUE y NWUE sigue existiendo la compatibilidad con los dos comportamientos señalados.

Quiero expresar mi más sincero agradecimiento al profesor Dr. D. Miguel Angel Gómez Villegas por su estímulo y ayuda incondicionales en todo momento. Hago extensivo el mismo a los compañeros del Departamento.

CAPITULO I

TRATAMIENTO DE DATOS ATIPICOS

CAPITULO I

TRATAMIENTO DE DATOS ATIPICOS

1. SUMARIO

Consideraremos en primer lugar el concepto de dato atípico en sus diferentes acepciones así como los posibles mecanismos generadores. A continuación y despues de una breve introducción histórica, daremos algunas nociones sobre la forma usual de tratar este tipo de observaciones: tests de discordancia y métodos de acomodación.

Los tests de discordancia han sido desarrollados principalmente para muestras univariantes y distribuciones conocidas, obteniéndose en cada caso los estadísticos más apropiados. Para comparar el comportamiento de los diferentes contrastes utilizaremos medidas específicas para este tipo de problemas.

Los métodos de acomodación se aplican para realizar inferencias válidas en la población aún con la presencia de datos atípicos. Los procedimientos de robustez, siendo éste un concepto mucho más amplio, han proporcionado la mayoría de las herramientas para estimación y contraste.

En el último apartado se trataran las nociones de resistencia y propensión. Según este enfoque podemos clasificar las distribuciones de manera que nuestra predisposición a considerar ciertas observaciones como atípicas, dependa del caracter de la distribución subyacente. Evidentemente es un tratamiento previo a la muestra y supone la existencia de un modelo para la población en la que se muestrea.

En esta linea se desarrollarán los demás capítulos, am pliando y modificando las definiciones así como aplicándolas a diferentes situaciones.

2. INTRODUCCION AL CONCEPTO DE DATO ATIPICO. ORIGENES HISTORICOS

Como paso inicial al desarrollo de este trabajo, será interesante detenernos en la significación y descripción de lo que entendemos por dato atípico. Quizá el aspecto más general puede ser el de la subjetividad del concepto (Collet y Lewis (1976)).

Ante un conjunto de datos, la persona que los observa puede indicar ciertos elementos que le parece no son consistentes con el resto de la masa global de los datos, ó que en principio le causan cierta sorpresa ó sospecha que son anómalos en algún sentido.

En ciertos casos puede ser que se tenga un modelo prefijado para estos datos y haya algunos que no se ajusten al mismo, pero en muchas otras ocasiones es el propio impacto que causa sobre el observador lo que determina totalmente un dato atípico.

La naturaleza del dato atípico puede ser aleatoria, siendo su causa, por ejemplo, la variabilidad propia del modelo, ó determinística, estando dentro de este tipo los producidos por errores de medidas o por influencias externas sin ningún interés.

La influencia de estos valores en el tratamiento posterior de los datos va desde la introducción de un grave error en los parámetros que se estimen, hasta la manifestación de un modelo diferente para la población, que englobe todos los datos, pasando por la detección de un fenómeno especialmente interesante para el estudio a realizar.

En general se observa una tendencia a considerar las cuestiones relativas a la integridad ó no de los datos, a si se ajustan ó no a modelos prefijados ó a si su efecto es ó no importante en las conclusiones finales. Sin embargo, a veces estos datos, en principio anómalos, pueden ser un aviso ante una situación errónea en el planteamiento inicial. Por lo tanto no todo se reduce a una discusión sobre la línea que delimita el rechazo ó aceptación de ciertas observaciones, sino a una visión general de unos datos intentando extraer de ellos la mayor información posible a cerca del fenómeno en cuestión. Solo en los casos de conocimiento exacto del error de tipo determinístico se llevaría a cabo el rechazo ó la repetición de la observación. En el resto de las ocasiones se realizará un estudio de discordancia ó acomodación directa de las observaciones.

Para concluir este apartado daremos una primera evaluación histórica de los trabajos sobre datos atípicos.

El comienzo del tratamiento del problema se puede resumir en un conjunto de opiniones a favor y en contra del rechazo de observaciones en principio dudosas o anómalas, con diferentes grados de justificación.

Dentro de este período se encuentran diferentes publicaciones, sobre todo de Astronomía como la de Boscovich (1755) que para determinar la elipticidad de la Tierra con un promedio de diez medidas del exceso del grado polar sobre el ecuatorial decide desechar los dos valores extremos por exceso, y calcular la media con las ocho observaciones restantes. Y así, a pesar de que Bernouilli (1777), en la primera discusión sobre el tema, no estuviese a favor del rechazo drástico de observaciones raras, éste se siguió practicando casi 75 años después sin ningún tipo de justificación.

El primer trabajo de caracter objetivo sobre el rechazo de observaciones atípicas, se debe al astrónomo americano Peirce (1852). Su criterio fue plasmado, tres años después, en unas tablas por Gould y así , aceptado por unos y criticado por otros llegamos dentro de esta línea al test de Chauvenet (1863). Este, considerando que el número esperado de observaciones en una muestra de tamaño n, de una $N(0,\sigma)$ que superen $c\sigma$ en valor absoluto es $n\,\Phi(-c)$, rechaza los mayores que $c\sigma$, en valor absoluto, donde c verifica: $n\,\Phi(-c) = 0{,}5$.

Al poco tiempo Stone (1868) introduce un test de rechazo, basado en el concepto de "módulo de descuido": m, si en media un observador comete un error en cada m observaciones que toma. Con un razonamiento análogo al de Chauvenet exige que c verifique: $m\,\Phi(-c) = 0{,}5$.

Los siguientes trabajos relativos a la detección de datos atípicos consideran diferentes pesos para las observaciones. Uno de los primeros fué Glaisher (1873), y supuso una cierta ruptura en la forma de tratar observaciones atípicas ya que empieza a utilizar métodos robustos, en algún sentido, y por primera vez señala que las observaciones podrían venir de poblaciones diferentes.

Estos dos autores anteriores, mantuvieron una polémica sobre sus planteamientos, recogidas en los diferentes comentarios que se dedican en sus publicaciones.

Otro método de ponderación fue propuesto por Newcomb (1886), y en el aparece un interesante principio del uso de una cierta función de pérdida. También es notable , dentro de este período de acomodación de los datos mediante ponder ciones el trabajo de Mendelev (1895) en el que se encuentra lo que ahora denominaríamos "truncamiento".

La aproximación, debida a Irwin (1925), proporcionó la base para un test de varios datos atípicos, desarrollándose a partir de él todos los trabajos sobre tests de huecos entre observaciones.

Un test basado en un criterio de "studentización" $\frac{X_i - \bar{X}}{s}$ fue publicado por Thompson (1935). Y por fin el tratamiento razonado de los datos atípicos empieza a tomar forma con el clásico de Pearson y Chandra Sekar (1936) donde además de criterios para tests de rechazo analizan el efecto de "enmascaramiento" producido al tener varios datos atípicos y aplicar contrastes para una única observación anómala.

A partir de este momento se aplicaron los tratamientos que veremos a continuación.

3. TESTS DE DISCORDANCIA

La forma más general de tratar los datos atípicos ha sido la aplicación de contrastes para detectarlos. En muestras pequeñas, la aplicación de los tests sólo se hacía en el caso de observaciones sospechosas, pero desde que se pueden manejar grandes masas de datos es casi imprescindible una aplicación sistemática de los mismos.

Según David (1981) parece deseable que los tests de este tipo sirvan para al menos uno de los siguientes objetivos:

i) Visualizar los datos de una forma rutinaria previa al análisis.

ii) Detectar la presencia de datos atípicos, indicando la necesidad de un estudio más profundo del mecanismo generador.

iii) Señalar las observaciones, que precisamente por ser extremas, son de especial interés.

3.1. Elementos de un test de discordancia

- Hipótesis Nula: Considera que los datos han sido generados por un modelo probabilístico básico, sin considerar los datos atípicos.

- Hipótesis Alternativa: Introduce unas ciertas modificaciones en el modelo, que explican la aparición de esos datos anómalos.

Hasta ahora las consideraciones que han prevalecido al plantear las pruebas de discordancia han sido fundamentalmente intuitivas y precisamente la hipótesis alternativa es uno de los puntos claves de la polémica que aún hoy se mantiene.

Diferentes formas en las que puede expresarse son las siguientes:

i) ALTERNATIVA DETERMINISTICA: Se engloban aquí todos los errores de medida o ejecución que pueden identificarse. Por tanto no hace falta contrastar la discordancia, y la decisión a tomar es la del rechazo o corrección de las observaciones detectadas como erróneas.

ii) ALTERNATIVA INHERENTE: Con ella se enfrenta la idea de que todos los datos provienen de una distribución F, con que todos provienen de una distribución G, bajo la cual el grado de sorpresa ante ciertas observaciones desaparezca prácticamente.

iii) ALTERNATIVA DE MIXTURA: Se consideran los datos atípicos como miembros de una población contaminante que aparece junto al modelo básico. Esta contaminación puede resultar en algunos casos, el descubrimiento de un fenómeno importante aisladamente estudiado, y en otros casos la introducción de un posible error en las estimaciones.

Se podría plantear este tipo de hipótesis alternativa con una expresión: $\bar{H}$: $(1-\lambda)F+\lambda G$, que refleja una probabilidad λ de que los datos atípicos surjan de una distribución G, diferente del modelo inicial F. Desde luego λ debe presuponerse pequeño, con lo que hay poca evidencia, si sólo aparecen uno o dos datos atípicos, para garantizar el uso del modelo alternativo. Además hay un cierta contradicción con la idea de dato atípico como valor extremo alejado, ya que bajo esta hipótesis alternativa ciertas observaciones son generadas por un modelo G, y no tienen por que ser justamente los extremos. Es decir bajo este enfoque, el dato atípico es un valor contaminante, evidentemente difícil de visualizar dentro de una muestra.

La aproximación bayesiana descrita por Box y Tiao (1968) puede considerarse como un ejemplo de la alternativa de mixtura, aunque es un método mas de acomodación que de discordancia.

iv) ALTERNATIVA DE DESLIZAMIENTO: Tal vez sea el tipo de alternativa más estudiada, como modelo apropiado para la generación de datos atípicos.

A grandes rasgos se considera que el modelo inicial F, tiene parámetros de localización μ y escala σ , y que en la muestra un cierto número de observaciones provienen de un modelo del tipo inicial con los parámetros transformados. Se suele estudiar por separado dos tipos de modelos alternativos: uno en que la trasformación implica una traslación del parámetro de origen y otro en el cual se considera la alteración del parámetro de escala.

Como ejemplo citaremos el planteamiento de Ferguson (1961), que describe dos posibles modelos.

Modelo A : $X_1,\ldots,X_n$ son variables aleatorias independientes, $N(\mu_i, \sigma)$ respectivamente. Existen constantes conocidas $a_1,\ldots,a_n$, muchas de las cuales serán cero, un parámetro desconocido Δ y una permutación de $(1,2,\ldots,n)$ desconocida $(\nu_1,\ldots,\nu_n)$, tales que:

$$\mu_i = \mu + \sigma \Delta a_{\nu_i} \qquad (i = 1,\ldots,n)$$

entonces las hipótesis alternativas serán del tipo:

$$\bar{H} : \Delta \neq 0 \quad ; \quad \bar{H}' : \Delta > 0$$

Modelo B: $X_1,\ldots,X_n$ variables aleatorias independientes $N(\mu, \sigma_i)$, que en las condiciones anteriores verifican:

$$\sigma_i^2 = \sigma^2 \exp(\Delta a_{\nu_i}) \qquad (i = 1,\ldots,n)$$

en este caso la hipótesis alternativa será $\bar{H} : \Delta > 0$ ya que $\bar{H}' : \Delta < 0$ será irrelevante para el estudio de datos atípicos.

De nuevo se plantea la anomalía de ser hipótesis alternativas que no especifican la observación que corresponde a la distribución trasformada en origen y escala.

v) ALTERNATIVA INTERCAMBIABLE: Se suele aplicar este modelo cuando lo que interesa es una estimación robusta o se busca la acomodación de los datos atípicos.

Para el caso de una única observación atípica la verosimilitud para una muestra de tamaño n, nos queda:

$$L(x_1,\dots,x_n) = \frac{1}{n} \sum_{i=1}^{n} \left\{ g(x_i) \prod_{j\neq i} f(x_j) \right\}$$

donde se supone que f es la densidad del modelo básico y g la de la población contaminante, siendo ambas f y g conocidas.

En general suponiendo k posibles datos atípicos

$$L_k(x_1,\dots,x_n) = \frac{1}{\binom{n}{k}} \sum \left\{ \prod_{i\in J} g(x_i) \prod_{j\notin J} f(x_j) \right\}$$

donde $J = (i_j,\dots,i_k)$ es una selección de k enteros de (1,2,...,n) y el sumatorio se extiende sobre todas las posibles formas de construir J.

3.2. Tipos básicos de estadísticos

Con vistas a la aplicación y comparación de los diversos estadísticos es conveniente:

- En primer lugar conocer el criterio de rechazo que se utiza, para lo que debemos saber la distribución del estadístico bajo la hipótesis nula.

- En segundo lugar que los tests sean comparables por su función de potencia, lo que nos lleva a conocer la distribución bajo la hipótesis alternativa, o al menos la posibilidad de manejarla.

- Y por último que posean ciertas propiedades de optimalidad; buscar tests uniformemente mas potentes o al menos óptimos localmente, insesgados, invariantes y de no ser así que su construcción se haga con un método como el de razón de verosimilitudes que puede trasmitir ciertas características útiles al problema de datos atípicos.

Este planteamiento reduce mucho los estadísticos a tener en cuenta, ya que es relativamente fácil construirlos bajo aspectos intuitivos, pero verificar las condiciones de manejabilidad, optimalidad y capacidad de comparación, hace inútiles muchas ideas totalmente razonables.

Otro problema que puede surgir con tests para un único dato atípico, es la existencia de otras observaciones extremas, no consideradas en principio como atípicas y que pueden producir un efecto llamado de "enmascaramiento", e impedir la detección de datos atípicos aunque de hecho existan varios. El problema se ha intentado solucionar proponiendo tests para varios datos atípicos o examinando los mismos sucesivamente, utilizando un test de tipo jerárquico.

Pasamos a ennumerar las clases más generales de estadísticos propuestos para la detección de datos atípicos siguiendo a Barnett y Lewis (1984).

i) Estadísticos de exceso y dispersión

Ejemplos: DIXON (1951):

$$\frac{X_{n:n} - X_{n-1:n}}{X_{n:n} - X_{2:n}}$$

($X_{k:n}$: estadístico de orden k en una muestra de tamaño n) para detectar una desviación del máximo, sin considerar el mínimo y así eliminar el posible efecto de distorsión si és te fuese demasiado pequeño.

IRWIN (1925): $$\frac{X_{n:n} - X_{n-1:n}}{\sigma}$$

suponiendo conocida la desviación típica de la distribución básica.

ii) <u>Estadísticos de dispersión y rango</u>

Ejemplo: DAVID, HARTLEY and PEARSON (1954); PEARSON and STEPHENS (1964):

$$\frac{X_{n:n} - X_{1:n}}{S} ,$$

donde S puede ser cualquier valor muestral o poblacional que refleje dispersión. Sirve para detectar que hay valores alejados del resto, pero no los señala claramente.

iii) <u>Estadísticos de dispersión y desviación</u>

Ejemplos: GRUBBS (1950):

$$\frac{\bar{X} - X_{1:n}}{S}$$

HALPERIN et al. (1955):

$$\max \frac{| X_i - \bar{X} |}{S}$$

donde $\bar{X}$ y S pueden ser cualquier valor muestral o poblacional que refleje localización y dispersión respectivamente.

iv) Estadísticos de sumas de cuadrados

Las sumas consideradas en este tipo de estadísticos, son la total y la obtenida excluyendo algunas observaciones. Sirva como ejemplo el propuesto por GRUBBS (1950) para contrastar dos datos atípicos superiores:

$$\frac{\sum_{i=1}^{n-2} (X_{i:n} - \bar{X}_{n,n-1})^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

donde

$$\bar{X}_{n,n-1} = \sum_{i=1}^{n-2} \frac{X_{i:n}}{n-2}$$

v) Estadísticos de momentos de orden elevado

Ejemplo: FERGUSON (1961):

$$\frac{\sqrt{n} \sum_{i=1}^{n} (X_i - \bar{X})^3}{\left(\sum_{i=1}^{n} (X_i - \bar{X})^2 \right)^{3/2}} \quad \text{y} \quad \frac{n \sum_{i=1}^{n} (X_i - \bar{X})^4}{\left(\sum_{i=1}^{n} (X_i - \bar{X})^2 \right)^{2}}$$

medidas de asimetría y curtosis respectivamente.

vi) Estadísticos de localización y extremos

Un representante de este tipo de estadísticos es $\frac{X_{n:n}}{\bar{X}}$

principalmente utilizado para el estudio de datos atípicos en las distribuciones GAMMA. Hay varios autores que tratan con estos estadísticos, citaremos a un especialista en el problema: KIMBER (1982).

Como tipo general de estadístico, eligiendo los parámetros i,j,k y m según donde se sospeche la contaminación, están los de DIXON (1950):

$$\frac{X_{n-i+1:n} - X_{j:n}}{X_{n-k+1:n} - X_{m:n}}$$

También pueden considerarse los estadísticos que surjan de aplicar en cada caso el test de razón de verosimilitudes o aproximaciones intuitivas de los obtenidos con la aplicación de dicho método.

Otro principio estadístico para la construcción de tests puede considerarse la optimalidad local. Así, Ferguson (1961) obtiene los tests invariantes localmente mejores con hipótesis alternativas de deslizamiento, tanto en la media, su modelo A, como en la varianza, modelo B, basados en el coeficiente muestral de asimetría

$$b_1 = \frac{\sqrt{n} \sum_{i=1}^{n} (X_i - \bar{X})^3}{\left(\sum_{i=1}^{n} (X_i - \bar{X})^2 \right)^{3/2}}$$

para alternativas unilaterales, y en el coeficiente muestral de curtosis:

$$b_2 = \frac{n \sum_{i=1}^{n} (X_i - \bar{X})^4}{\left(\sum_{i=1}^{n} (X_i - X)^2 \right)^2}$$

para alternativas bilaterales.

LOs estadísticos para contrastar desviaciones del modelo básico F_0, como por ejemplo el de Anderson-Darling (1954)

$$A_n^2 = n \int_{-\infty}^{+\infty} \frac{(F_n(x) - F_0(x))^2}{F_0(x)\,(1 - F_0(x))} \, d\,F_0(x)$$

donde $F_n(x)$ es la función de distribución empírica, así como la familia general de estadísticos que propone Tiku (1975), parecen en principio muy interesantes como tests para datos atípicos, aunque todavía su comportamiento no ha sido suficientemente investigado.

3.3 Datos atípicos multiples

Como ya hemos apuntado anteriormente, la existencia de varias observaciones atípicas provoca ciertos problemas como el del "enmascaramiento", ya definido y detectado por Pearson y Chandra Sekar (1936) y el de "superación" descrito por Fieller (1976) que nos hace identificar como atípica alguna observación que no lo es, al ser arrastrada por otra demasiado extrema. Esto nos lleva a utilizar planteamientos propios de este tipo de situación.

Las dos lineas más generales en las que se han desarrollado los tests para datos atípicos múltiples son:

- procedimientos consecutivos
- procedimientos de bloque

Según indican sus denominaciones, los procedimientos consecutivos consisten en analizar sucesivamente las observaciones extremas y eliminarlas hasta obtener valores no significativos, por lo que son mas propensos al efecto de "enmascaramiento". Por otra parte los de bloque analizan los correspondientes datos extremos sospechosos y contrastan su discordancia en conjunto ; aquí el posible error se puede producir por el efecto de "superación".

Y para concluir, como ejemplo de otros métodos para construir tests de discordancia, citaremos a Kitagawa (1979) que utiliza el criterio de información de Akaike (AIC), que aproxima menos dos veces la Entropía Esperada con:

AIC = - 2 log (máxima verosimilitud) + 2 (número de parámetros independientes ajustados)

En su trabajo propone la comparación de los diferentes modelos que se consiguen suponiendo las r_1 observaciones pequeñas, de una población y las r_2 mayores de otra, con el criterio de Akaike, seleccionando el del mínimo AIC. Este método aparece duramente criticado en Barnett y Lewis (1984).

A pesar del inmenso desarrollo del estudio sobre datos atípicos, aun queda mucho por hacer para conseguir tests con buenas propiedades y asequibles desde un punto de vista práctico, sobre todo en el terreno de los datos atípicos múltiples.

3.4. Medidas de comportamiento

Para poder comparar los diferentes tests de discordancia, David y Paulson (1965) y David (1981) proponen una serie de medidas para el caso de un único dato atípico:

P_1 : Función de potencia: probabilidad de que el tests concluya correctamente que existe un dato atípico.

P_2 : Probabilidad de que el valor del estadístico aplicado al contaminante sea significativo.

P_3 : Probabilidad de que el test concluya que existe un dato atípico y lo identifique correctamente.

P_4 : Probabilidad de que sólo sea significativo el estadístico en el valor contaminante.

P_5 : Potencia condicional: probabilidad de que el test sea significativo en el contaminante suponiendo que es extremo.

Consideremos el caso de un único dato atípico por la derecha, por ejemplo X_{i^*} , con hipótesis alternativa de deslizamiento. Aunque las medidas expuestas pueden aplicarse a situaciones más generales, supongamos que el estadístico es del tipo:

$$T = \max_i T_i$$

como por ejemplo : $T = \max\limits_i \dfrac{|X_i - \bar{X}|}{S}$

Sea t_α tal que $\Pr (T > t_\alpha / H_0) = \alpha$, entonces:

$$P_1 = \Pr (T > t_\alpha / \bar{H})$$

$$P_2 = \Pr (T_{i^*} > t_\alpha / \bar{H})$$

$$P_3 = \Pr (T_{i^*} > t_\alpha \wedge X_{i^*} > X_j \quad \forall j \neq i^* / \bar{H})$$

$$P_4 = \Pr (T_{i^*} > t_\alpha \wedge T_j < t_\alpha \quad \forall j \neq i^* / \bar{H})$$

$$P_5 = \Pr (T_{i^*} > t_\alpha / X_{i^*} > X_j \quad \forall j \neq i^* ; \bar{H})$$

Por su propia definición, $P_1 \geqslant P_2 \geqslant P_3 \geqslant P_4$

$$y\ P_5 = \frac{P_3}{Pr\ (X_{i^*} > X_j \quad \forall\ j \neq i^* \ /\ \bar{H})}$$

La diferencia entre P_1 y P_3 radica en que el valor contaminante puede aparecer en el centro o en un extremo de la muestra.

La utilización de cada una de estas medidas en particular, dependerá del tipo de hipótesis alternativa considerada, y de hecho no existen siempre todas las medidas.

Dado que el calculo, en algunos casos, se complica podremos utilizar ciertas acotaciones. Así, para la función de potencia haremos uso de las desigualdades de Bonferroni para n sucesos $(A_1,\ldots,A_n)$

$$\sum_{i=1}^{n} Pr(A_i) - \sum_{i<j}\sum Pr(A_i \cap A_j) \leqslant Pr(\bigcup_{i=1}^{n} A_i) \leqslant \sum_{i=1}^{n} Pr(A_i)$$

ya que identificando A_i con el suceso $(T_i > t_\alpha)$ nos quedaría:

$$P_1 = Pr(\bigcup_{i=1}^{n} A_i)$$

Para datos atípicos múltiples se pueden utilizar las mismas medidas, por supuesto con un significado diferente que originará interesantes medidas de tipo parcial.

4. METODOS DE ACOMODACION

Con este tratamiento se pretende eliminar el efecto nocivo de los datos atípicos en las conclusiones sin tener en cuenta su identificación.

Como concepto más amplio, en esta linea, esta la robustez, entendiendo como Huber (1981) que: "robustez significa insensibilidad a pequeñas desviaciones de las suposiciones". De toda la colección de problemas que se tratan en esta parcela de la Estadística, nos quedaremos con aquellos métodos específicos para situaciones en las que aparecen o pueden aparecer datos atípicos.

Podremos utilizar métodos robustos de estimación, que incluyen los L- estimadores , M-estimadores y R-estimadores. Como ilustración sirve el trabajo de Kale y Sinha (1971) que utilizan L-estimadores ponderando las observaciones y dando menos peso a los extremos.

En esta linea podríamos incluir una de las primeras publicaciones con enfoque bayesiano sobre el tema debida a de Finetti (1961).

Dentro de un tratamiento más específico de la acomodación de datos atípicos tenemos el trabajo de Anscombe (1960) en el cual se introducen unas ciertas medidas de eficiencia que denomina: premio y protección comparando con la contratación de una póliza de seguros. Si por ejemplo H_0 es la hipotesis nula de no contaminación y $\bar{H}$ es la alternativa que incluye algún tipo de contaminación, podremos construir un estadístico T óptimo bajo H_0 y T_R , un estadístico robusto para el modelo descrito en $\bar{H}$, entonces:

$$\text{premio} = \frac{\text{Var}(T_R / H_0) - \text{Var}(T / H_0)}{\text{Var}(T / H_0)}$$

$$\text{protección} = \frac{\text{Var}(T / \bar{H}) - \text{Var}(T_R / \bar{H})}{\text{Var}(T / \bar{H})}$$

También se podrán utilizar tests de hipótesis e intervalos de confianza basados en estadísticos del tipo $T - \theta / S_T$ donde T es un estimador robusto para θ, y que poseen propiedades de robustez.

5. PROPENSION Y RESISTENCIA A DATOS ATIPICOS

Según este planteamiento, ya desde el conocimiento del modelo que sigue la población, podemos conseguir distintos grados de expectativa sobre los datos atípicos que van a aparecer en la muestra.

En su origen el concepto surge aplicandolo a una familia de distribuciones, si es que poseen un determinado tipo de observaciones extremas, Neyman y Scott (1971). Las definiciones fueron las siguientes:

DEFINICION 5.1.- Sea $(X_1,\ldots,X_n)$ una muestra aleatoria de una distribución F, y sea $(X_{1:n},\ldots,X_{n:n})$ el estadístico de orden. Diremos que $X_{n:n}$ es un (k,n) dato atípico para un determinado $k > 0$ si:

$$(X_{n:n} - X_{n-1:n}) > k\,(X_{n-1:n} - X_{1:n})$$

DEFINICION 5.2.- Si llamamos $P(k,n,F) = Pr\,(X_{n:n} - X_{n-1:n} > k(X_{n-1:n} - X_{1:n}) \,/\, F)$ y $\Pi(k,n,\mathcal{F}) = \sup_{F \in \mathcal{F}} P(k,n,F)$

diremos que la familia de distribuciones $\mathcal{F}$ es (k,n) PROPENSA (a datos atípicos) si $\Pi(k,n,\mathcal{F}) = 1$ y será (k,n) RESISTENTE (a datos atípicos) si $\Pi(k,n,\mathcal{F}) < 1$.

DEFINICION 5.3.- Si la familia de distribuciones $\mathcal{F}$ es PROPENSA para todo $k > 0$ y $n \geqslant 3$, entonces diremos que $\mathcal{F}$ es COMPLETAMENTE PROPENSA.

Posteriormente Green (1974), demuestra el siguiente resultado:

TEOREMA 5.1 : Una familia de distribuciones es propensa completamente, si y solo si es (k,n) propensa para algún $k > 0$ y $n > 2$.

Basándose en este tipo de definiciones Kale (1977) estudia el comportamiento de las mixturas $\{(1-p)\,F_1 + p\,F_2\}$ donde $F_1 \in \mathcal{F}_1$, $F_2 \in \mathcal{F}_2$ y $0 < p < 1$; variando estas distribuciones a lo largo de toda la familia se obtiene una nueva familia de distribuciones formada por mixturas de las anteriores. Los resultados podrían resumirse así:

i) Si $\mathcal{F}_1$ y $\mathcal{F}_2$ son resistentes, entonces la mixtura es resistente, sea cual sea el valor de p.

ii) Con p conocido ó al menos $0 < \epsilon_1 \leqslant p \leqslant \epsilon_2 < 1$, basta con que haya una familia de la mixtura que sea resistente, para que la mixtura lo sea.

iii) Sólo demuestra en casos particulares que una mixtura de dos propensas completamente, también lo es.

iv) Si en vez de mixturas se consideran modelos para acomodar datos atípicos en los que de las n observaciones que idealmente provienen de una población con función de distribución F, se sospecha que una de las observaciones es de una población contaminante con distribución G, en ese caso demuestra que son propensos completamente.

En el mismo artículo se señala la diferencia en principrio sorprendente de estos dos últimos resultados, justificado por el efecto de enmascaramiento ya que la definición de propensas corresponde al caso de un único dato atípico análogo al modelo iv) pero diferente de las mixturas, modelo iii) que pueden permitir más de un dato atípico.

Continua generalizando la definición al caso de m datos atípicos ($X_{n-m+1:n}, \ldots, X_{n:n}$) mediante:

$$P_m(k,n,F) = Pr(X_{n-m+1:n} - X_{n-m:n} > k\,(X_{n-m:n} - X_{1:n}) \;/\; F\,)$$

$$\pi_m(k,n,F) = \sup_{F\in\mathcal{F}} P_m(k,n,F)$$

y a partir de aquí repetir las definiciones.

Finalmente recomienda los modelos de tipo intercambiable, o sea verosimilitudes:

a) $$L(x_1,\ldots,x_n / \theta, \xi) = \frac{1}{n} \sum_{i=1}^{n} f(x_i, \xi) \prod_{j\neq i} f(x_j, \theta)$$ para

un dato atípico.

b) $$L(x_1,\ldots,x_n / \theta, \xi) = \frac{1}{\binom{n}{m}} \sum_{C_{n,m}} \prod_{i\in I_m} f(x_i, \xi) \prod_{j\in I_m} f(x_j, \theta)$$

para m datos atípicos, donde I_m: conjunto de m enteros contenido en (1,2,...,m) y $C_{n,m}$: combinaciones de n elementos tomados de m en m, para acomodar observaciones extremas.

No obstante se plantea la necesidad de aplicar estas definiciones a distribuciones en particular en vez de a familias de distribuciones como se había hecho anteriormente. Entonces aparece un artículo de Green (1976) que demuestra el siguiente resultado:

TEOREMA 5.2 : Ninguna familia de distribuciones, con un único miembro, puede ser propensa completamente.

Había que dar nuevas definiciones y así continúa Green su trabajo suponiendo que la función de distribución F(x) verifica:

1) $F(\infty) = 1$

2) $F(x) < 1$ para todo x finito

DEFINICION 5.4.- Una función de distribución F es ABSOLUTAMENTE RESISTENTE (a datos atípicos) si $\forall \varepsilon > 0$ tenemos que $Pr(X_{n:n} - X_{n-1:n} > \varepsilon) \longmapsto 0$ cuando $n \to \infty$.

DEFINICION 5.5.- Una función de distribución F es RELATIVAENTE RESISTENTE (a datos atípicos) si para todo $k > 1$, se tiene que $Pr\left(\frac{X_{n:n}}{X_{n-1:n}} > k\right) \longmapsto 0$ cuando $n \to \infty$.

DEFINICION 5.6.- Una función de distribución F es ABSOLUTAMENTE PROPENSA (a datos atípicos) si existe un $\varepsilon > 0$, $\delta > 0$ y un n_0 tales que $Pr(X_{n:n} - X_{n-1:n} > \varepsilon) \geqslant \delta$ $\forall n \geqslant n_0$.

DEFINICION 5.7: Una función de distribución F es RELATIVAMENTE PROPENSA (a datos atípicos) si existe $k > 1$, $\delta > 0$

y un n_0 tales que: $\Pr\left(\frac{X_{n:n}}{X_{n-1:n}} > k\right) \geq \delta \quad \forall n \geq n_0$

Y este es el punto de partida para el trabajo desarrollado en los capítulos posteriores.

Los demas resultados que aparecen en la publicación de Green (1976) y que relacionan las definiciones con el comportamiento asintótico del máximo, serán expuestos posteriormente dentro de un contexto más general, ampliando las definiciones anteriores.

Evidentemente y teniendo en cuenta que:

$$\min(X_1,\ldots,X_n) = -\max(-X_1,\ldots,-X_n)$$

se pueden generalizar los resultados al caso de dato atípico por excesivamente pequeño es decir un valor de $X_{1:n}$ demasiado alejado del resto de la muestra.

CAPITULO II

CLASIFICACION DE DISTRIBUCIONES SEGUN SU CAPACIDAD DE PRODUCIR EXTREMOS ALEJADOS

CAPITULO II

CLASIFICACION DE DISTRIBUCIONES SEGUN SU CAPACIDAD DE PRODUCIR EXTREMOS ALEJADOS

1. SUMARIO

Después de la relación de algunos resultados importantes, utilizados en el capítulo, pasamos a generalizar las definiciones de Green (1976); en un primer paso suprimiendo los supuestos sobre la función de distribución, y a continuación utilizando conceptos de convergencia en probabilidad tanto de la diferencia $S_n = X_{n:n} - X_{n-1:n}$ como de la razón $R_n = \frac{X_{n:n}}{X_{n-1:n}}$. De las definiciones correspondientes surgirán los resultados esenciales del presente trabajo.

La primera generalización de las definiciones de Green (1976) nos llevará a establecer ciertas implicaciones entre los diferentes comportamientos y la pertenencia a los distintos dominios de atracción de los límites de máximos normalizados. Con la siguiente ampliación, tendremos sencillas caracterizaciones según el comportamiento de la cola derecha de la función de distribución y de la función de densidad cuando exista.

Al igual que en el capítulo anterior, efectuamos el desarrollo considerando extremo al máximo, ya que de forma obvia puede pasarse de un mínimo a un máximo.

2. RESULTADOS PREVIOS

2.1. Distribuciones asintóticas de extremos

Consideremos $X_1,\ldots,X_n,\ldots$ una sucesión de variables a leatorias independientes e identicamente distribuidas según la función de distribución $F(x)$. Si construimos la sucesión de los correspondientes máximos: $X_{1:1}, X_{2:2}, \ldots, X_{n:n}, \ldots$, las funciones de distribución serán:

$$F_n(x) = P(X_{n:n} \leq x) = F^n(x), \; n=1,2,\ldots$$

El estudio asintótico de esta sucesión, fue el objetivo fundamental en la teoría clásica de extremos, con publicaciones como las de Dodd (1923), Frechet (1927), Fisher y Tippet (1928) y von Misses (1939), cuyos resultados fueron recopilados y completados por Gnedenko (1943). Más tarde de Haan (1970, 1976) aplica nuevas técnicas y demuestra más sencillamente los teoremas básicos.

2.1.1. Leyes de los grandes números

Siguiendo el ya citado trabajo de Gnedenko, tenemos las siguientes definiciones:

DEFINICION 2.1.1. La sucesión de máximos $(X_{n:n})$ cumple la LEY DE LOS GRANDES NUMEROS, si existe una colección de constantes (A_n) tales que $(X_{n:n} - A_n)$ converge en probabilidad hacia cero, cuando $n \to \infty$.

DEFINICION 2.1.2. Diremos que la sucesión de los máximos $(X_{n:n})$ es RELATIVAMENTE ESTABLE, si existen unas constantes positivas $\{B_n\}$ tales que $\left(\frac{X_{n:n}}{B_n}\right)$ converge en probabilidad hacia uno cuando $n \to \infty$.

En los dos teoremas siguientes aparecen condiciones sobre la función de distribución $F(x)$ que posteriormente servirán para caracterizar las distribuciones propensas y resistentes de Green (1976).

TEOREMA 2.1.1.: Suponiendo que $F(x) < 1$ para todo x finito, la sucesión de los máximos ($X_{n:n}$), cumplirá la ley de los grandes números si y solo si, para todo $\epsilon > 0$

$$\lim_{x \to \infty} \frac{1 - F(x+\epsilon)}{1 - F(x)} = 0$$

TEOREMA 2.1.2.: Suponiendo que $F(x) < 1$ para todo x finito, la sucesión de los máximos será relativamente estable si y solo si para todo $k > 1$ se verifica

$$\lim_{x \to \infty} \frac{1 - F(kx)}{1 - F(x)} = 0$$

Geffroy (1958, 1959) demuestra además, que si la sucesión de los máximos cumple la ley de los grandes números, también lo hace la sucesión de las (k+1)-mayores observaciones y con las mismas constantes; analogamente para la estabilidad relativa. De esta forma se puede establecer la equivalencia de:

i) La definición anterior de resistencia absoluta (Definición 5.4. del Capítulo I) con la ley de los grandes números.

ii) La definición anterior de resistencia relativa (Definición 5.5. del Capítulo I) con el concepto de estabilidad relativa, pudiendose caracterizar las funciones de distribución resistentes de Green (1976) con las condiciones de los Teoremas 2.1.1. y 2.1.2. ya enunciados.

Basándose en las condiciones de Gnedenko (1943), caracteriza Green (1976) las distribuciones propensas con los siguientes teoremas:

TEOREMA 2.1.3: Una función de distribución $F(x)$, bajo las condiciones i) $F(\infty) = 1$, ii) $F(x) < 1$ para todo x finito, será absolutamente propensa a datos atípicos si y solo si exis ten constantes $A > 0$ y $B > 0$ tales que: $\frac{1 - F(x+B)}{1 - F(x)} \geq A$ para todo x finito.

TEOREMA 2.1.4.: Bajo los supuestos i) y ii) del Teorema 2,1.3. una función de distribución F(x) será relativamente propensa a datos atípicos si y solo si existen constantes $k > 1$, $M > 0$ tales que:

$$\frac{1 - F(kx)}{1 - F(x)} \geq M \quad \text{para todo x finito}$$

Si suponemos la existencia de la función de densidad f(x), se podrán enunciar los correspondientes resultados que damos referidos a la propensión y resistencia de las distribuciones correspondientes.

TEOREMA 2.1.5.: Considerando que la función de distribución F(x) verifica los supuestos i) y ii) y que posee función de densidad f(x), las condiciones

a) $\lim\limits_{x \to \infty} \frac{f(x+\varepsilon)}{f(x)} = 0$ para todo $\varepsilon > 0$

b) $\lim\limits_{x \to \infty} \frac{f(kx)}{f(x)} = 0$ para todo $k > 1$

son suficientes para la resistencia absoluta y relativa respectivamente de la distribución F. También serán necesarios si la densidad posee una cola derecha monótona.

TEOREMA 2.1.6.: Si la función de distribución F(x) verifica las hipótesis i) y ii) y posee función de densidad f(x), son condiciones suficientes para la propensión absoluta y relativa respectivamente las siguientes:

a) existen constantes $A > 0$, $B > 0$ y x_0 tales que

$$\frac{f(x+B)}{f(x)} \geq A \quad \text{para todo } x \geq x_0$$

b) existen constantes $k > 1$, $M > 0$ y x_0 tales que

$$\frac{f(kx)}{f(x)} \geq M \quad \text{para todo } x \geq x_0$$

Estas condiciones no son necesarias ni siquiera añadiendo la condición de monotonía para la cola derecha de la densidad.

2.1.2. Tipos de distribución límite

Manteniendo la analogía con los resultados de sumas de variables aleatorias independientes e identicamente distribuidas, donde en vez de potencia n-ésima de F(x) se obtiene la convolución n-ésima, nos planteamos las condiciones necesarias para que existan colecciones de constantes (a_n) y (b_n) donde $a_n > 0$, que hagan converger debilmente la sucesión:

$$F^n(a_n x + b_n) = \Pr(a_n^{-1}(X_{n:n} - b_n) \leq x)$$

a una distribución G no degenerada, cuando $n \to \infty$. Como resultados importantes a utilizar daremos el teorema de Khintchine y la definición de distribución Max-estable.

<u>TEOREMA 1.7.</u>: Sea (F_n) una sucesión de funciones de distribución y G una función de distribución no degenerada. Sean $a_n > 0$ y b_n constantes tales que

$$F^n(a_n x + b_n) \xrightarrow{w} G(x) \quad \text{(convergencia débil)}$$

Entonces, que existan una función de distribuciones G_* y constantes $c_n > 0$ y d_n tales que:

$$F_n(c_n x + d_n) \xrightarrow{w} G_*(x)$$

es equivalente a que:

$$\frac{c_n}{a_n} \longmapsto a \qquad y \qquad a_n^{-1}(d_n - b_n) \longmapsto b$$

para algún $a > 0$ y b, entonces

$$G_*(x) = G(ax + b)$$

es decir G y G_* son distribuciones del mismo tipo. Por tanto más que una distribución límite tendremos un tipo límite.

DEFINICION 2.1.3.: Diremos que una función de distribución no degenerada $G(x)$ es MAX-ESTABLE si para todo $n = 2,3,...$ existen constantes $a_n > 0$ y b_n tales que:

$$G^n(a_n x + b_n) = G(x)$$

Así se pasa a resolver el posible comportamiento asintótico de los máximos normalizados de una forma completa mediante los siguientes teoremas:

TEOREMA 2.1.8.: Toda distribución max-estable es del tipo de valores extremos es decir igual a $G(ax + b)$ para algún $a > 0$ y algún b, donde:

TIPO I : (ó $\Lambda(x)$ en la notación clásica)

$$G(x) = \exp(-e^{-x}) \qquad -\infty < x < \infty$$

TIPO II: (ó $\Phi_\alpha(x)$)

$$G(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}) \text{ (para algún } \alpha > 0) & x > 0 \end{cases}$$

TIPO III: (ó $\Psi_\alpha(x)$)

$$G(x) = \begin{cases} \exp\left(-(-x)^{\alpha}\right) \text{ (para algún } \alpha > 0) & x \leq 0 \\ 1 & x > 0 \end{cases}$$

Inversamente, cada función de distribuciones del tipo da valores extremos es max-estable.

Y como punto final a esta parte de la Teoría de Extremos tenemos el Teorema de los Tipos Extremos, al que así llaman Leadbetter, Lindgren y Rootzen (1983).

TEOREMA 2.1.9.: Si la sucesión de los máximos ($X_{n:n}$) es tal que para algunas constantes $a_n > 0$ y b_n, verifica

$$\Pr(a_n^{-1}(X_{n:n} - b_n) \leq x) \xrightarrow{w} G(x)$$

donde G(x) es alguna función de distribución no degenerada, entonces G es de alguno de los tipos de valores extremos que aparecen en el Teorema 2.1.8.

Inversamente cada función de distribución G(x) del tipo de valores extremos puede aparecer como límite de máximos normalizados; de hecho es el límite si se considera que la sucesión de variables aleatorias tiene como función de distribución común G(x).

2.1.3. Caracterización de los dominios de atracción

Diremos que la función de distribución F pertenece al dominio de atracción (para máximos) de la función de distribución G(x) que será del tipo de valores extremos si existen constantes $a_n > 0$ y b_n tales que: $F^n(a_n x + b_n) \xrightarrow{w} G(x)$. Utilizaremos la notación $F \in D(G)$.

Desde luego el siguiente paso en la teoría clásica de extremos fue caracterizar estos dominios de atracción por el comportamiento de las funciones de distribución que los componían, más concretamente se vió que dependía de la variación de la cola derecha en la función de distribución.

Los siguientes teoremas se deben a Gnedenko (1943).

TEOREMA 2.1.10.: Para que una función de distribución F(x) pertenezca al dominio de atracción de una distribución del tipo II de valores extremos i.e. $\Phi_\alpha(x)$ es necesario y suficiente que para todo valor de $k > 0$

$$\lim_{x \to \infty} \frac{1 - F(kx)}{1 - F(x)} = k^{-\alpha} \qquad \text{con } \alpha > 0$$

y el punto final $x_F = \sup(x \,/\, F(x) < 1)$ sea infinito.

TEOREMA 2.1.11.: Para que una función de distribución F(x) pertenezca al dominio de atracción de alguna distribución del tipo III de valores extremos i.e. $\Psi_\alpha(x)$ es necesario y suficiente que:

i) exista un x_F tal que $\begin{cases} F(x_F) = 1 \\ F(x_F - \epsilon) < 1 \text{ para todo } \epsilon > 0 \end{cases}$

es decir que el punto final de F sea finito

ii) $$\lim_{h \to 0} \frac{(1 - F(x_F - kh)}{1 - F(x_F - h)} = k^{\alpha} \quad \text{para todo } k > 0$$

Respecto al dominio de atracción del tipo I, se dan dos teoremas aunque según el propio autor no tienen una forma tan concluyente y cómoda de aplicar como los casos anteriores.

TEOREMA 2.1.12: La función de distribución $F(x)$ pertenece al dominio de atracción de alguna distribución del tipo I si y solo si existen constantes (a_n) y (b_n) donde $a_n > 0$ tales que:

$$\lim_{n \to \infty} n(1 - F(a_n x + b_n)) = e^{-x}$$

TEOREMA 2.1.13: La función de distribución $F(x)$ pertenece al dominio de atracción de alguna distribución del tipo I si y solo si es posible encontrar una función $A(z)$ que verifique

$$z\, A(z) > 0 \quad \text{para } z \neq 0$$

$$\lim_{z \uparrow x_F} A(z) = 0 \quad \text{donde } x_F = \sup \{ x \; F(x) < 1 \}$$

y tal que para cada x:

$$\lim_{z \uparrow x_F} \frac{1 - F(z + z\,A(z)\, x)}{1 - F(z)} = e^{-x}$$

siendo $x_F = \sup (x \,/\, F(x) < 1)$ punto final de la distribución.

Observación: El enunciado del teorema 2.1.13, corresponde a L. de Haan (1971), ya que en el de Gnedenko (1943) se exige la continuidad de $A(x)$, que después no se utiliza en la demostración.

Para concluir este apartado resumimos el trabajo de von Misses (1936) en una serie de condiciones suficientes de pertenencia a los diferentes dominios de atracción en el caso de existir la función de densidad $f(x)$.

TEOREMA 2.1.14: Dada la función de distribución $F(x)$ que posee función de densidad $f(x)$, las condiciones suficientes para que pertenezca a cada uno de los tres dominios de atracción son las siguientes:

TIPO I:

- $f(x)$ tenga derivada negativa para todo x en un intervalo (x_0, x_F) donde $x_F \leqslant \infty$
- $f(x) = 0$ para $x \geqslant x_F$
- $\lim_{x \uparrow x_F} \dfrac{f'(x)\,(1 - F(x))}{f^2(x)} = -1$

TIPO II:

- $f(x) > 0$ para todo $x \geqslant x_0$ donde x_0 es finito.
- $\lim_{x \to \infty} \dfrac{x\, f(x)}{1 - F(x)} = \alpha > 0$

TIPO III:

- $f(x) > 0$ para todo x en un intervalo finito (x_0, x_F)
- $f(x) = 0$ para $x > x_F$
- $\lim_{x \uparrow x_F} \dfrac{(x_F - x)\, f(x)}{1 - f(x)} = \alpha > 0$

siendo $x_F = \sup \{x \,/\, F(x) < 1\}$, punto final de la distribución.

En la demostración de los teoremas anteriores se obtienen también las constantes normalizadas a_n y b_n. No obstante en de Haan (1970) se hace un estudio más exhaustivo de la obtención de dichas constantes en términos de los cuantiles de F y para situaciones más generales.

EJEMPLOS:

A continuación citaremos algunas distribuciones conocidas de los distintos dominios de atracción:

Al de tipo I pertenecen: Normal, Exponencial y la propia distribución $\Lambda(x)$.

En el de tipo II está la distribución de Pareto, y la de Cauchy así como toda la colección $(\Phi_\alpha(x))_{\alpha>0}$.

Por último dentro del dominio de atracción del tipo III estarán distribuciones con punto final finito como por ejemplo la Uniforme en (0,1), la Exponencial truncada en un valor x_F y las propias distribuciones de la familia $(\Psi_\alpha)_{\alpha>0}$.

También hay distribuciones que no pertenecen a ningún dominio de atracción, como es el caso de la Poisson, y la Geométrica, para las que el único posible límite del máximo normalizado es degenerado.

2.2. Funciones de variación regular. Aplicación a la teoría de extremos

El concepto de variación regular para una función positiva, lo introdujo Karamata en 1930. Desde entonces ha tenido múltiples aplicaciones, en particular dentro de la teoría de la probabilidad.

La justificación de esta gran aplicabilidad dentro del campo de las probabilidades está, según Feller (1966), en que las funciones de distribución son monótonas y se tiene el siguiente resultado:

LEMA 2.2.1: Sea U una función monótona positiva en $(0,\infty)$, tal que:

$$\frac{U(tx)}{U(t)} \longmapsto R(x) \leqslant \infty \quad \text{cuando } t \to \infty$$

en un conjunto denso A. Entonces:

$$R(x) = x^{\rho} \quad \text{donde} \quad -\infty \leqslant \rho \leqslant \infty$$

Este comportamiento es el de las funciones de variación regular, si ρ es finito, aunque se extiende a funciones no necesariamente monótonas.

<u>DEFINICION 2.2.1.</u>: Una función positiva U, definida en $(0, \infty)$ es de VARIACION REGULAR en infinito si:

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^{\rho} \quad \text{para todo } x > 0 .$$

El número real ρ recibe el nombre de EXPONENTE DE REGULARIDAD. En el caso particular de $\rho = 0$, se dice que la función es de VARIACION LENTA.

La variación regular de una función, es una propiedad de tipo local, unilateral y asintótico, que surge al extender la clase de funciones cuyo comportamiento asintótico cerca de un punto es el de una función potencial multiplicada por un factor que varía más lentamente que una función potencial.

Así puede considerarse esta nueva definición:

<u>DEFINICION 2.2.2</u>: Una función positiva (definida en $(0, \infty)$) varia regularmente, con exponente ρ (diremos que es ρ-VARIANTE) si y solo si es de la forma:

$$U(x) = x^{\rho} L(x)$$

donde L(x) es una función de variación lenta.

Diremos que una función $U(x)$ es de variación regular en el cero si $U(1/x)$ es de variación regular en infinito. Por tanto podemos definir la variación regular en cualquier punto finito simplemente trasladando el origen de la función.

Estas definiciones pueden extenderse al caso $\rho = \pm\infty$ si consideramos:

$$x^{\infty} = \begin{cases} 0 & \text{para } x < 1 \\ 1 & \text{para } x = 1 \\ \infty & \text{para } x > 1 \end{cases} \qquad x^{-\infty} = \begin{cases} \infty & \text{para } x < 1 \\ 1 & \text{para } x = 1 \\ 0 & \text{para } x > 1 \end{cases}$$

<u>DEFINICION 2.2.3</u>: Una función positiva U definida sobre $(0, \infty)$ es de VARIACION RAPIDA en infinito si:

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^{\rho} \qquad \text{para todo } x > 0$$

con $\rho = \pm\infty$.

<u>Ejemplos:</u>

- Cualquier función eventualmente positiva y medible que posea un límite finito cuando $x \to \infty$ es de variación lenta.

- El ejemplo más sencillo de función de variación lenta en infinito es $L(x) = \log x$ y cualquier iteración de las mismos.

- Entonces $U(x) = x^{\rho} \log(1+x)$ es ρ-variante.

- También se tiene que $L(x) = \operatorname{arctg} x$ es de variación lenta y $U(x) = e^{x}$ es de variación rápida en infinito.

- Por último funciones de tipo oscilatorio no amortiguadas no varían regularmente como en el caso de :

$$f(x) = 2 + \operatorname{sen} x$$

A continuación expondremos algunos resultados debidos a Karamata, posteriormente recogidos por de Haan (1970) de gran aplicación en la Teoría de Extremos.

TEOREMA 2.2.1: Sea U una función positiva definida en $(0, \infty)$ e integrable sobre intervalos finitos, entonces son equivalentes:

a) U es de variación regular con exponente $\rho > -1$

b) $$\lim_{x\to\infty} \frac{x\,U(x)}{\int_0^x U(t)\,dt} = \rho + 1 > 0$$

c) (Tª de Representación). Existen funciones reales c(x) y a(x) tales que

$$\lim_{x\to\infty} c(x) = c \qquad (0 < c < \infty)$$

$$\lim_{x\to\infty} a(x) = \rho > -1$$

y se verifica: $U(x) = c(x) \exp\left(\int_1^x \frac{a(t)}{t}\,dt \right)$

TEOREMA 2.2.2: El mismo enunciado anterior para $\rho < -1$ reemplazando el apartado b) por:

b)' $$\lim_{x\to\infty} \frac{x\,U(x)}{\int_x^\infty U(t)\,dt} = -\rho - 1 > 0$$

COROLARIO 2.2.1: a) Si U es de variación regular con exponen te $\rho > -1$ se tiene que: $U_1(x) = \int_0^x U(t)\,dt$ es de variación regular con exponente $\rho + 1$.

b) si U es de variación regular con exponente $\rho < -1$ se tiene que: $U_2(x) = \int_x^\infty U(t)\,dt$ es de variación regular con exponente $\rho + 1$.

Para funciones de variación rápida se puede dar una versión más debil de los resultados anteriores.

TEOREMA 2.2.3: a) Una función no-decreciente y positiva U, definida en $(0, \infty)$ es de variación rápida con $\rho = \infty$ si y solo si:

$$\lim_{x \to \infty} \frac{x\, U(x)}{\int_0^x U(t)\,dt} = \infty$$

b) Una función no creciente y positiva U definida en $(0, \infty)$ es de variación rápida con $\rho = -\infty$ si y solo si:

$$\lim_{x \to \infty} \frac{x^{-1}\, U(x)}{\int_x^\infty U(t)\, \frac{dt}{t^2}} = \infty$$

Dado que la variación regular es una propiedad asintótica es suficiente que las condiciones de los teoremas sobre la función sean válidas en intervalos $[A, \infty)$ para algún $A > 0$.

PROPIEDADES:

1 .- Si U es ρ-variante en infinito $(-\infty < \rho < \infty)$. entonces:

$$\lim_{x \to \infty} \frac{\log U(x)}{\log x} = \rho \qquad \text{y por tanto:}$$

$$\lim_{x \to \infty} U(x) = \begin{cases} 0 & \text{si } \rho < 0 \\ \infty & \text{si } \rho > 0 \end{cases}$$

2 .- Si U es ρ-variante en infinito $(-\infty < \rho < \infty)$ se verifica que para todas las sucesiones (a_n) y (b_n) de números positivos tales que:

$$\lim_{n \to \infty} a_n = \lim_{n \to \infty} b_n = \infty \quad \text{y} \quad \lim_{n \to \infty} \frac{a_n}{b_n} = c$$

donde $0 < c < \infty$ tenemos:

$$\lim_{n \to \infty} \frac{U(a_n)}{U(b_n)} = c^{\rho}$$

Si $\rho \neq 0$ $(-\infty < \rho < +\infty)$ el resultado es válido también para $c = 0$ y $c = \infty$.

Si $\rho = \pm\infty$ la conclusión es cierta para funciones monótonas U y $c \neq 1$ $(0 \leqslant c \leqslant \infty)$.

3 .- Si U es ρ-variante en infinito $(-\infty < \rho < \infty)$ la relación:

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^{\rho}$$

se da uniformemente sobre intervalos finitos de la forma: (x_1, x_2) con $0 < x_1 < x_2 < \infty$

Si $\rho < 0$ se puede eliminar la condición $x_2 < \infty$

Si $\rho > 0$ y U está acotada sobre intervalos acotados podemos tomar $x_1 = 0$

4 .- Toda función de variación regular con exponente $\rho \neq 0$ es asintótica a una función estrictamente monótona y de variación regular con el mismo exponente.

5 .- Sea U una función ρ-variante $(-\infty < \rho < \infty)$ y supongamos que existe una función monótona u tal que para todo x positivo

$$U(x) = \int_0^x u(t)\, dt$$

entonces

$$\lim_{x \to \infty} \frac{x\, u(x)}{U(x)} = \rho$$

Si se considera que $u(x) = U'(x)$ es continua para $x \geqslant B$ donde $B > 0$ y que $U(x)$ es positiva se tiene el resultado inverso.

EXTENSION DEL CONCEPTO DE VARIACION REGULAR

Se consideran funciones $U : R^+ \longmapsto R^+$ tales que existe una función $f : R^+ \longmapsto R^+$ y un número real ρ tal que verifican:

$$\lim_{t \to \infty} \frac{U(t.\, x^{f(t)})}{U(t)} = x^{\rho} \quad \text{para todo x positivo}$$

Así para funciones no decrecientes U y en el caso particular de $\rho = 1$ (restricciones que no nos hacen perder generalidad en los resultados) tenemos la clase Γ :

DEFINICION 2.2.4: Una función no decreciente $U: R \longmapsto R^+$ pertenece a la clase Γ si se dá alguno de los apartados siguientes:

a) existe una función positiva f tal que para todo número real x:

$$\lim_{t \to \infty} \frac{U(t + x\, f(t))}{U(t)} = e^{x}$$

b)

$$\lim_{x \to \infty} \frac{U(x) \left(\int_0^x \int_0^y U(t)\, dt\, dy \right)}{\left(\int_0^x U(t)\, dt \right)^2} = 1$$

c) existen funciones reales c(x), a(x), b(x) y un número real c tal que:

$$c(x) > 0 \quad : \quad \lim_{x \to \infty} c(x) = c > 0$$

$$\lim_{x \to \infty} a(x) = 0$$

$$\int_0^x b(t)\, dt > 0 \quad ; \quad \lim_{x \to \infty} b(x) = 0$$

y para todo valor x : $U(x) = c(x) \exp \left(\int_0^x \frac{1 + a(t)}{\int_0^t b(s) ds} dt \right)$

También tiene interes el estudio de la clase Π, compuesta por las imágenes inversas de las funciones de la clase Γ. Para profundizar en estos temas citaremos la publicación de L. de Haan (1974).

APLICACION A LA TEORIA DE EXTREMOS

Si se analizan las condiciones de pertenencia a los diferentes dominios de atracción, estas podran expresarse en términos de variación regular de la cola derecha de la función de distribución.

TEOREMA 2.2.4: i) La función de distribución F pertenece al dominio de atracción de una función de distribución del tipo II i.e. $\Phi_\alpha(x)$ si y solo si $(1-F(x))$ es de variación regular con exponente $-\alpha$, en infinito.

ii) La función de distribución F pertenece al dominio de atracción de alguna función de distribución del tipo III i.e. $\Psi_\alpha(x)$ si y solo si $x_F < \infty$ y además $1 - F(x_F - \frac{1}{x})$ es de variación regular con exponente $(-\alpha)$, en infinito.

TEOREMA 2.2.5: Una función de distribución F pertenece al dominio de atracción de algunas funciones de distribución de tipo I i.e. $\Lambda(x)$ si y solo si la función $U(x) = \frac{1}{1-F(x)}$ está en la clase Γ.

En el caso de la ley debil de los grandes números y la estabilidad relativa, obtenemos las siguientes equivalencias:

TEOREMA 2.2.5: Dada una función de distribución F con punto final x_F, infinito, se tiene que son equivalentes las siguientes afirmaciones:

a) F es relativamente estable

b) $1 - F$ es $-\infty$-variante en infinito

c) La integral $\int_0^\infty (1 - F(t))\, dt$ es finita y

$$\lim_{x\to\infty} \frac{x\,(1-F(x))}{\int_x^\infty (1-F(t)\, dt} = \infty$$

COROLARIO 2.2.2: Si F pertenece al dominio de atracción de $\Lambda(x)$ y su punto final x_F es infinito se verifica que F es relativamente estable.

TEOREMA 2.2.6: Dada una función de distribución F con punto final x_F en infinito, los siguientes apartados son equivalentes:

a) F verifica la ley debil de los grandes números

b) Para todo $x > 0$ $\lim_{t \to \infty} \frac{1 - F(t+x)}{1 - F(t)} = 0$

c) La integral $\int_0^\infty (1 - F(t))\, dt$ es finita y

$$\lim_{x \to \infty} \frac{1 - F(x)}{\int_x^\infty (1 - F(t))\, dt} = \infty$$

El Teorema 2.2.6 se deduce del Teorema 2.2.5 considerando el hecho de que F verifica la ley debil de los grandes números si y solo si la función G definida por:

$$G(x) = \begin{cases} 0 & \text{si} \quad x \leqslant 0 \\ F(\log x) & \text{si} \quad x > 0 \end{cases}$$

es absolutamente estable.

3. GENERALIZACION DE LA CLASIFICACION DE GREEN. RELACION CON LOS DOMINIOS DE ATRACCION

Como es de suponer el considerar distribuciones con punto final finito, es decir $x_F = \sup \{ x \,/\, F(x) < 1 \} < \infty$ nos con ducirá a funciones de distribución resistentes a datos atípicos. Así veremos que pueden suprimirse las hipótesis que llevan a exigir que el punto final sea infinito, para aplicar las definiciones dadas por Green (1976).

Para comenzar recordaremos una proposición de Gnedenko (1943) que trata el caso $x_F < \infty$ aplicado a la ley de los grandes números y a la estabilidad relativa que ya vimos eran equivalentes a las dos definiciones de resistencia.

PROPOSICION 3.1: Si la función de distribución F es tal que su punto final x_F es finito, existe una sucesión de constantes (A_n) ,, $A_n = x_F$ para todo n tales que: la sucesión de los máximos cumple la ley de los grandes números. Si además $x_F > 0$, existe una sucesión (B_n) tal que $B_n = x_F$, para todo n que hace que la sucesión de los máximos sea relativamente estable.

PROPOSICION 3.2: Si la función de distribución F es tal que $x_F < \infty$, entonces $\Pr(X_{n:n} - X_{n-1:n} > \epsilon) \xrightarrow[n \to \infty]{} 0$ para todo $\epsilon > 0$.

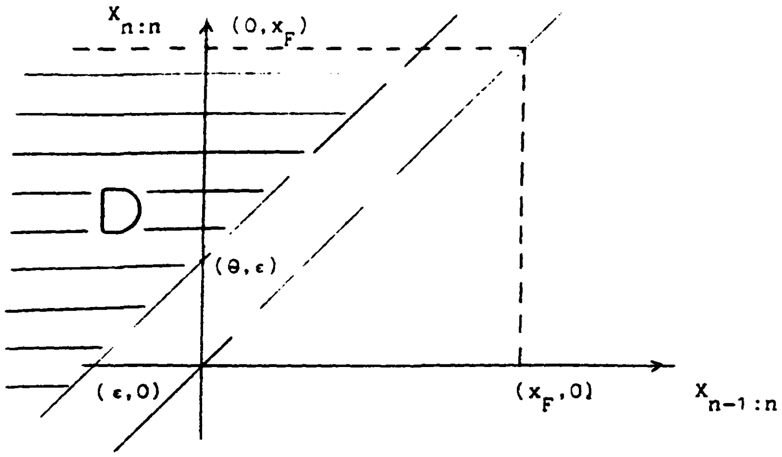
DEMOSTRACION:

Caso 1: Supondremos que F posee una función de densidad f; entonces la función de densidad conjunta de las variables $(X_{n-1:n}, X_{n:n})$ será:

$$f_{X_{n-1:n},X_{n:n}}(x,y) = \frac{n!}{n-2!}(F(x))^{n-2} f(x)\, f(y) \text{ para } x \leq y$$

por lo tanto

$$\Pr(X_{n:n} - X_{n-1:n} > \epsilon) = \iint_D f_{X_{n-1:n},X_{n:n}}(x,y)\,dx\,dy =$$



$$= \int_{-\infty}^{+\infty} n(n-1) \left(\int_{x+\epsilon}^{\infty} f(x)\ f(y)\ (F(x))^{n-2}\ dy \right) dx =$$

$$= \int_{-\infty}^{+\infty} n(n-1)\ (F(x)^{n-2}\ (1 - F(x+\epsilon))\ f(x)\ dx =$$

$$= \int_{-\infty}^{x_F-\epsilon} n(n-1)\ (F(x))^{n-2}\ (1 - F(x+\epsilon))\ \ f(x)\ dx +$$

$$+ \int_{x_F-\epsilon}^{x_F} n(n-1)(F(x))^{n-2}\ (1 - F(x+\epsilon))f(x)\ dx$$

pero la función $1 - F(x+\epsilon) = 0$ para todo x en $(x_F-\epsilon,\ x_F)$, con lo que la segunda integral se anula, y así nos queda:

$$Pr(X_{n:n}-X_{n-1:n} > \epsilon) = \int_{-\infty}^{x_F-\epsilon} n(n-1)(F(x))^{n-2}(1-F(x+\epsilon)\ f(x)\ dx <$$

$$< \int_{-\infty}^{x_F-\epsilon} n(n-1)(F(x))^{n-2}\ f(x)\ dx = n(F(x_F-\epsilon))^{n-1}$$

hemos utilizado que $(1-F(x+\epsilon)) < 1$ para todo $x < x_F-\epsilon$ y por la misma razón $F(x_F-\epsilon) < 1$ con lo cual:

$$Pr(X_{n:n} - X_{n-1:n} > \epsilon) < n(F(x_F-\epsilon))^{n-1} \underset{n \longmapsto \infty}{\longmapsto} 0$$

entonces:

$$Pr(X_{n:n} - X_{n-1:n} > \epsilon) \underset{n \to \infty}{\longmapsto} 0 \quad \text{para todo } \epsilon > 0$$

c.q.d.

<u>Caso 2</u>: En el caso general, se tiene que:

$$Pr(X_{n:n} - X_{n-1:n} > \epsilon) \leq Pr(x_F - X_{n:n} > \epsilon/2) +$$

$$+ Pr(x_F - X_{n-1:n} > \epsilon/2)$$

y sabemos que $Pr(x_F - X_{n:n} > \epsilon/2) \underset{n \to \infty}{\longmapsto} 0$

Por otro lado: $Pr(x_F - X_{n-1:n} < \epsilon) = Pr(X_{n-1:n} > x_F - \epsilon) =$

$=$ Pr("haya al menos dos observaciones, de las n, mayores que

$x_F-\epsilon$") $= 1 - ((F(x_F-\epsilon))^n + n(F(x_F-\epsilon))^{n-1}(1-F(x_F-\epsilon))) \underset{n \to \infty}{\longmapsto} 1$

es decir: $Pr(x_F - X_{n-1:n} > \epsilon/2) \underset{n \to \infty}{\longmapsto} 0$

En resumen $Pr(X_{n:n} - X_{n-1:n} > \epsilon) \underset{n \to \infty}{\longmapsto} 0$ para todo $\epsilon > 0$

c.q.d.

Con los resultados anteriores podemos dar la siguiente definición, sin hacer ninguna hipótesis sobre el punto final x_F, de la distribución básica:

<u>DEFINICION 3.1</u>: La función de distribución F es absolutamente resistente a datos atípicos si para todo $\epsilon > 0$ se tiene:

$$Pr(X_{n:n} - X_{n-1:n} > \epsilon) \longmapsto 0 \quad \text{cuando } n \to \infty$$

Y como consecuencia de la proposición 3.2:

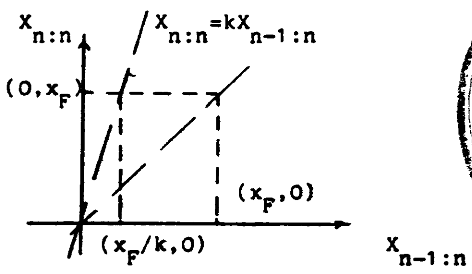
<u>PROPOSICION 3.3</u>: Sea F una función de distribución con punto final x_F, finito, entonces F es absolutamente resistente a datos atípicos.

A continuación efectuamos el mismo tratamiento con la resistencia relativa (Definición 5.5 del Capítulo I) y el punto final finito.

<u>PROPOSICIÒN 3.4</u>: Si el punto final de la función de distribución F es tal que $0 < x_F < \infty$, se verifica que:

$$Pr(X_{n:n} / X_{n-1:n} > k) \underset{n \to \infty}{\longmapsto} 0 \quad \text{para todo } k > 1$$

<u>DEMOSTRACION</u>: Sea $k > 1$, entonces



$$Pr(X_{n:n} / X_{n-1:n} > k) \leqslant Pr(X_{n-1:n} \leqslant x_F/k) =$$

$$= F(x_F/k)^n + n(1-F(x_F/k))\,(F(x_F/k))^{n-1}$$

y como $F(x_F/k) < 1$ para todo $k > 1$, está última expresión con verge a cero cuando $n \to \infty$ por lo que:

$$Pr(X_{n:n}/X_{n-1:n} > k) \underset{n \to \infty}{\longrightarrow} 0 \qquad \text{c.q.d.}$$

PROPOSICION 3.5: Si el punto final de la distribución F, el valor x_F es negativo, se verifica que F es relativamente resistente a datos atípicos.

DEMOSTRACION: Trivialmente si $X_{n-1:n} \leqslant X_{n:n} \leqslant 0$ se verifica que $\dfrac{X_{n:n}}{X_{n-1:n}} \leqslant 1$ por ser variables negativas por tanto para todo $k > 1$ $Pr(X_{n:n}/X_{n-1:n} > k) = 0$ para todo n

y así F es relativamente resistente a datos atípicos

c.q.d.

Definimos nuevamente la resistencia relativa a datos atípicos, igual que en el caso anterior, sin hacer ninguna hipótesis sobre el punto final de la distribución.

DEFINICION 3.2: Diremos que una función de distribución F es relativamente resistente a datos atípicos si para todo $k > 1$ se tiene:

$$Pr(X_{n:n}/X_{n-1:n} > k) \longrightarrow 0 \text{ cuando } n \to \infty$$

y la consiguiente proposición, basada en las anteriores.

PROPOSICION 3.6: Si una función de distribución es tal que su punto final es finito, entonces es relativamente resistente a datos atípicos.

Por lo tanto en el caso de tener punto final finito, la función de distribución correspondiente posee las dos características de resistencia. Veremos a continuación las implicaciones que existen en el caso de punto final infinito.

PROPOSICION 3.7: Sea una función de distribución F con $x_F = \infty$ entonces si F es absolutamente resistente a datos atípicos, se verifica que F es relativamente resistente.

DEMOSTRACION: Por los teoremas de Green (1976) tenemos que suponiendo $x_F = \infty$ se dan las siguientes equivalencias:

$$F \text{ absolutamente resistente} \iff \text{para todo } \epsilon > 0 \lim_{x \to \infty} \frac{1-F(x+\epsilon)}{1-F(x)} = 0$$

$$F \text{ relativamente resistente} \iff \text{para todo } k > 1 \lim_{x \to \infty} \frac{1-F(kx)}{1-F(x)} = 0$$

Es decir que si F es absolutamente resistente $\Rightarrow$ dados $\epsilon > 0$ y $\delta > 0$ existe un x_0 tal que

$$\frac{1-F(x+\epsilon)}{1-F(x)} < \delta \text{ para todo } x \geq x_0$$

Además dados $\epsilon > 0$ y $k > 1$ existe un x_0' tal que:

$$kx_0' = x_0' + \epsilon \quad \text{y} \quad kx > x + \epsilon \quad \text{para todo } x \geq x_0'$$

Sea $x_0'' = \max(x_0, x_0')$, entonces para todo $x \geq x_0''$ se darán simultaneamente las dos condiciones anteriores, o sea:

$$\frac{1-F(kx)}{1-F(x)} < \frac{1-F(x+\epsilon)}{1-F(x)} < \delta$$

entonces: para todo $k > 1$: $\lim_{x \to \infty} \frac{1-F(kx)}{1-F(x)} = 0 \iff$ F es relativamente resistente c.q.d.

La implicación inversa no es cierta en general, por ejemplo la distribución de Poisson es relativamente resistente y no es absolutamente resistente.

En el caso de funciones de distribución propensas a datos atípicos (Definiciones 5.6 y 5.7 del Capítulo I) se dan las siguientes implicaciones:

PROPOSICION 3.8: Si F es una función de distribución con $x_F = \infty$ absolutamente propensa a datos atípicos se verifica que F no es absolutamente resistente.

DEMOSTRACION: Utilizando el Teorema 2.1.3 , si F es absolutamente propensa existen $\epsilon > 0$, $\delta > 0$ y un n_0 tales que:

$$\Pr(X_{n:n} - X_{n-1:n} > \epsilon) \geq \delta \text{ para todo } n \geq n_0 \Rightarrow \text{existe un } \epsilon > 0$$

$$\text{tal que } \lim_{n\to\infty} \Pr(X_{n:n} - X_{n-1:n} > \epsilon) \neq 0 \Leftrightarrow F \text{ no es absolutamente resistente}$$

c.q.d.

Por lo tanto la definición de absolutamente propensa a datos atípicos no tiene ningún sentido en distribuciones con punto final finito.

PROPOSICION 3.9: Si F es una función de distribución con $x_F = \infty$ relativamente propensa a datos atípicos, se verifica que F no es relativamente resistente (y por tanto no es absolutamente resistente, tampoco)

DEMOSTRACION: Análogamente a lo anterior, utilizando la razón

$$R_n = \frac{X_{n:n}}{X_{n-1:n}}$$

<u>PROPOSICION 3.10</u>: Si F es una función de distribución con $x_F = \infty$ y es relativamente propensa a datos atípicos, se verifica que F es absolutamente propensa.

<u>DEMOSTRACION</u>: Tenemos que llegar a que:

Existen $k > 1$, $\delta > 0$ tales que $\dfrac{1 - F(kx)}{1 - F(x)} \geq \delta$ para todo x =>

=> existen $\alpha > 0$, y $\beta > 0$ tal que $\dfrac{1 - F(x+\beta)}{1 - F(x)} \geq \alpha$ para todo x

Lo demostramos con la negación de las condiciones anteriores, así supongamos que:

Dados $\alpha > 0$, $\beta > 0$, existe x'_0 tal que $\dfrac{1 - F(x'_0+\beta)}{1 - F(x'_0)} < \alpha$

Sea un $\beta^* > 1$, entonces existe un x_{β^*} tal que:

$$x_{\beta^*} + \beta^* = \beta^* x_{\beta^*} \quad \text{ó sea} \quad x_{\beta^*} = \frac{\beta^*}{\beta^*-1}$$

$$y \ \beta^* x > x + \beta^* \quad \text{para todo } x > x_{\beta^*}$$

Si x'_0 es tal que verifica: $\dfrac{1 - F(x'_0+\beta^*)}{1 - F(x'_0)} < \alpha$ para todo α

veamos que x'_0 no puede ser menor que x_{β^*}.

Supongamos $x'_0 < x_{\beta^*}$ entonces:

$$\frac{1-F(x'_0+\beta^*)}{1-F(x'_0)} > \frac{1-F(x_{\beta^*}+\beta^*)}{1-F(x'_0)} > (1-F(\beta^{*2}/\beta^*-1)) = \alpha^*$$

por tanto existen $\alpha^{*} > 0$, $\beta^{*} > 0$ tales que: $\dfrac{1-F(x_0' + \beta^{*})}{1-F(x_0')} > \alpha^{*}$

en contradicción con la hipótesis, así que:

$x_0' > x_{\beta^{*}}$, pero si esto ocurre así:

$\beta^{*} x_0' > x_0' + \beta^{*}$ lo cual implica

$$\frac{1-F(\beta^{*} x_0')}{1-F(x_0')} < \frac{1-F(x_0' + \beta^{*})}{1-F(x_0')} < \alpha \text{ para todo } \alpha .$$

es decir F no es relativamente propensa c.q.d.

Como resultado final, vamos a relacionar el comportamiento de las distribuciones respecto a los datos atípicos, considerados como máximos demasiado alejados, con los dominios de atracción de los tipos de distribuciones de valores extremos ya descritos en el Teorema 2.1.8.

TEOREMA 3.1: Dada una función de distribución F, se verifican las siguientes implicaciones:

a) $F \in D(\Lambda) \Rightarrow$ F es relativamente resistente a datos atípicos

b) $F \in D(\Phi_\alpha) \Rightarrow$ F es relativamente propensa a datos atípicos $\Rightarrow$

$\Rightarrow$ F es absolutamente propensa a datos atípicos.

c) $F \in D(\Psi_\alpha) \Rightarrow$ F es absolutamente resistente a datos atípicos ($\Leftrightarrow$ F es relativamente resistente ya que el punto final será finito).

DEMOSTRACION:

a) Si $x_{\omega F} < \infty$ se verifica que F es tanto relativamente como absolutamente resistente.

Supongamos $x_F = \infty$, entonces, por el Corolario 2.2.2., se verifica que F es relativamente estable, equivalente, en este caso, con que F sea relativamente resistente a datos atípicos.

b) Si $F \in D(\Phi_\alpha)$ <=> por el Teorema 2.1.10) $\begin{cases} \lim\limits_{y\to\infty} \dfrac{1-F(ky)}{1-F(y)} = k^{-\alpha} \\ x_F = \infty \end{cases}$
(para algún $\alpha > 0$)

para todo $k > 0$, por tanto para todo $B > 1$, $k > 0$ se tiene:

$$1 / Bk^{\alpha} \leqslant \lim_{y\to\infty} \frac{1-F(ky)}{1-F(y)} \leqslant B / k^{\alpha}$$

así existirá un y_0 tal que para todo $y \geqslant y_0$

$$\frac{1-F(ky)}{1-F(y)} \geqslant \frac{1}{B\,k^{\alpha}} = M_1$$

sea $\dfrac{y_0}{k} \leqslant y \leqslant y_0$ entonces $\dfrac{1-F(ky)}{1-F(y)} \geqslant \dfrac{1-F(ky_0)}{1-F(y_0/k)} = M_2$

sea $y < \dfrac{y_0}{k}$ entonces $\dfrac{1-F(ky)}{1-F(y)} \geqslant 1 - F(y_0) = M_3$

de esta forma seleccionando $M = \max(M_1, M_2, M_3)$ nos queda

$$\frac{1-F(ky)}{1-F(y)} \geqslant M \quad \text{para todo } y .$$

condición necesaria y suficiente para que F sea relativamente propensa, lo que implica, dado que el punto final es infinito, que F es absolutamente propensa por la Proposición

3.10.

c) Si $F \in D(\Psi_\alpha)$ para algún $\alpha > 0$ <=> $\begin{cases} x_F < \infty \\ y \\ \lim\limits_{h \to 0} \dfrac{1-F(x_F - xh)}{1-F(x_F - h)} = x^\alpha \quad \forall\, x > 0 \end{cases}$

entonces por ser el punto final finito la distribución será absolutamente y relativamente resistente a datos atípicos.

COMENTARIOS: Las definiciones de Green (1976) suponen un primer paso muy valioso en todo el planteamiento de la propensión y resistencia, pero adolece de algunos inconvenientes, incluso con nuestra generalización del apartado 3.

Por ejemplo, no discrimina entre las distintas funciones de distribución con punto final finito. También, consideramos que las definiciones de propensión suponen simplemente una negación del concepto de resistencia pero se pierde el significado que tenía ésta, respecto al comportamiento de la diferencia, o razón en cada caso, de los extremos superiores.

Las definiciones que propondremos a continuación, supera estas anomalías, dando una clasificación más completa de las distribuciones.

4. DISTRIBUCIONES PROPENSAS (SP), NEUTRAS (SN) Y RESISTENTES (SR) SEGUN EL COMPORTAMIENTO ASINTOTICO DE $S_n = X_{n:n} - X_{n-1:n}$

Corresponde a este apartado una de las cuestiones fundamentales de nuestro trabajo. Seguimos considerando como dato atípico el extremo superior $X_{n:n}$ demasiado alejado del resto, representado por el siguiente extremo $X_{n-1:n}$, y reflejando la distancia por la diferencia entre ambos $S_n = X_{n:n} - X_{n-1:n}$.

Analizaremos por separado distribuciones con punto final infinito y finito, solucionando este último caso con la aplicación del primero a un cambio de variable.

4.1. Caso de punto final infinito

Sean funciones de distribución F, tales que su punto final $x_F = \sup\{x \mid F(x) < 1\}$ es $x_F = \infty$.

4.1.1. Definiciones de distribuciones SP, SN y SR

DEFINICION 4.1.1: Diremos que una función de distribución F con $x_F = \infty$ es RESISTENTE con la suma a datos atípicos, y la denominaremos SR, si $S_n \equiv o_p(1)$, es decir para todo $\epsilon > 0$

$$\lim_{n\to\infty} \Pr(S_n > \epsilon) = 0$$

Esta definición coincide con la de resistencia absoluta de Green (1976). Las dos siguientes son las que constituyen la novedad de nuestro planteamiento al desglosar la no resistencia en dos comportamientos diferentes.

DEFINICION 4.1.2: Diremos que una función de distribución F, con $x_F = \infty$, es NEUTRA con la suma, a datos atípicos y la denominaremos SN si

$$\begin{cases} S_n \equiv O_p(1) \\ \text{y} \\ S_n \not\equiv o_p(1) \end{cases}$$

es decir si S_n está acotada en probabilidad pero no converge a cero en probabilidad.

Teniendo en cuenta el significado de la notación O_p y

o_p, podemos enunciar la siguiente proposición:

PROPOSICION 4.1.1: Dada una función de distribución F que posee función de densidad f será SN si:

a) Para todo $\epsilon > 0$, $\exists\, c_\epsilon$ y $\exists\, N_\epsilon$ tales que: $Pr(S_n > c_\epsilon) < \epsilon$

para todo $n \geq N_\epsilon$ y $c \geq c_\epsilon$

b) Existen $\epsilon_0 < 1$ y $c^* > 1$ tales que:

$Pr(S_n > c^*) = \epsilon_0$ para infinitos n's.

DEMOSTRACION:

Para que una función de distribución F sea SN deben cumplirse las condiciones:

a) $S_n = O_p(1)$, es decir para todo $\epsilon > 0$, existen c_ϵ, N_ϵ tales que

$Pr(S_n > c_\epsilon) < \epsilon$ para todo $n > N_\epsilon$ y por supuesto para todo $c > c_\epsilon$

b) $S_n \neq o_p(1)$ es decir negando el que para todo $\epsilon > 0$ el

$\lim_{n \to \infty} Pr(S_n > \epsilon) = 0$ nos queda

Existen $c_0 > 0$, $\epsilon_0 > 0$ tales que $Pr(S_n > c_0) > \epsilon_0$ para infinitos n's.

Si aplicamos a) al ϵ_0 de b), existen c_{ϵ_0} y N_{ϵ_0} tales que

$Pr(S_n > c_{\epsilon_0}) < \epsilon_0$ para todo $n > N_{\epsilon_0}$ y $c > c_{\epsilon_0}$ es decir para infinitos n's.

También por b), existe c_0 tal que $Pr(S_n > c_0) > \epsilon_0$ para infinitos n's

Evidentemente $c_0 < c_{\epsilon_0}$ ya que $k(c) = Pr(S_n > c)$ es monótona no creciente para cualquier n.

Por otra parte:

$$k(c) = Pr(S_n > c) = \int_{-\infty}^{+\infty} \frac{1-F(x+c)}{1-F(x)}\, n(n-1)(F(x))^{n-2}(1-F(x))f(x)dx$$

donde: $n(n-1)(F(x))^{n-2}\,(1-F(x))\,f(x) = f_{X_{n-1:n}}(x)$, la función de densidad del segundo extremo $X_{n-1:n}$.

Como $\frac{1-F(x+c)}{1-F(x)}$ es una función continua en c para todo x y

además

$$\left| \frac{1-F(x+c)}{1-F(x)} \cdot f_{X_{n-1:n}}(x) \right| \leq \left| f_{X_{n-1:n}}(x) \right| \quad \text{para todo c}$$

donde

$$\int_{-\infty}^{+\infty} f_{X_{n-1:n}}(x)\, dx = 1$$

resulta que $k(c)$ es una función continua, luego existe un c^* tal que $Pr(S_n > c^*) = \epsilon_0$ para infinitos n's siendo $c_0 < c^* < c_{\epsilon_0}$

c.q.d.

<u>DEFINICION 4.1.3</u>: Diremos que una función de distribución es PROPENSA con la suma y la denominaremos SP si $S_n \xrightarrow{P} \infty$ cuando $n \to \infty$, es decir si para todo $\epsilon > 0$ $\lim_{n\to\infty} Pr(S_n > \epsilon) = 1$.

Con estas definiciones, se conoce el comportamiento de los extremos en los diferentes casos pero no son aplicables con facilidad a los casos particulares. Por esto hemos establecido una relación con propiedades de la función de distribución básica.

4.1.2.- Caracterización con propiedades de las funciones de distribución

TEOREMA 4.1.1: Una función de distribución F con $x_F = \infty$ será SR si y solo si para todo $\epsilon > 0$ $\lim_{x \to \infty} \frac{1-F(x+\epsilon)}{1-F(x)} = 0$.

DEMOSTRACION: Utilizando el paralelismo entre la definición de SR y que el máximo cumpla la ley de los grandes números, es un resultado de Gnedenko (1943).

Para demostrar los siguientes teoremas de caracterización, necesitamos suponer que existe función de densidad además de un resultado debido a Schuster (1984) que enunciaremos a continuación:

TEOREMA 4.1.2: Sea F una función de distribución absolutamente continua, con densidad f y $x_F = \infty$. Supongamos que el límite $h_1 = \lim_{x \to \infty} \frac{f(x)}{1-F(x)}$ existe y es posiblemente infinito

Entonces

$$h_1 = 0 \iff S_n \xrightarrow{P} \infty$$

$$h_1 = a \quad \text{donde} \quad 0 < a < \infty \iff \begin{cases} S_n = O_p(1) \\ S_n \neq o_p(1) \end{cases}$$

$$h_1 = \infty \iff S_n = o_p(1)$$

A partir de aquí caracterizaremos las distribuciones propensas y resistentes de forma análoga al Teorema 4.1.1.

TEOREMA 4.1.3: Una función de distribución F con $x_F = \infty$ y función de densidad f es SN si y solo si

$$\text{Para todo } \epsilon > 0 \quad \lim_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)} = e^{-c\epsilon} \quad \text{donde } c = \lim_{x\to\infty} \frac{f(x)}{1-F(x)}$$

suponiendo que $0 < c < \infty$ y que F es estrictamente monótona

DEMOSTRACION:

$$=\!>) \text{ Si F es SN entonces } \left\{ \begin{array}{l} S_n \equiv O_p(1) \\ \\ S_n \not\equiv o_p(1) \end{array} \right. \iff \lim_{x\to\infty} \frac{f(x)}{1-F(x)} = c$$

donde $0 < c < \infty$ por el Teorema 4.1.2.

$$\text{Esto equivale a: } \lim_{x\to\infty} \frac{d \log \frac{1}{1-F(x)}}{dx} = c \iff \text{para todo } M > 0$$

existe x_0, tal que:

$$\left| \frac{d \log \frac{1}{1-F(x)}}{dx} - c \right| < M \quad \text{para todo } x > x_0$$

o sea

$$c - M < \frac{d \log \frac{1}{1-F(x)}}{dx} < c + M \quad \text{para todo } x > x_0$$

Tomemos un $\epsilon > 0$ cualquiera, entonces para todo $x' > x_0$, se tiene:

$$\int_{x'}^{x'+\epsilon} (c-M)\ dx < \int_{x'}^{x'+\epsilon} d \log \frac{1}{1-F(x)} < \int_{x'}^{x'+\epsilon} (c+M)\ dx$$

es decir:

$$\epsilon\,(c-M) < \log \frac{1-F(x')}{1-F(x'+\epsilon)} < \epsilon\ (c+M)$$

lo que implica

$$e^{+M\epsilon}\ e^{-\epsilon c} < \frac{1-F(x'+\epsilon)}{1-F(x')} < e^{-c\epsilon}\ e^{-M\epsilon} \quad \text{para todo } x' > x_0$$

y como es también para todo $M > 0$:

$$\lim_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)} = e^{-c\,\epsilon} \quad \text{tal que } 0 < c < \infty\ .$$

$\Leftarrow$) Supongamos que para todo $\epsilon > 0$, $\lim\limits_{x\to\infty} \dfrac{1-F(x+\epsilon)}{1-F(x)} = e^{-c\,\epsilon}$

donde $0 < c < \infty$.

Como $x_F = \infty$ y F es estrictamente monótona, para x's suficientemente grandes, F es cóncava, luego

$$\frac{F(x) - F(x-\epsilon)}{\epsilon} < f(x) < \frac{F(x+\epsilon) - F(x)}{\epsilon} \quad \text{para todo } \epsilon > 0$$

También : $\lim\limits_{x\to\infty} \dfrac{1-F(x+\epsilon)}{1-F(x)} = e^{-c\,\epsilon} \iff \lim\limits_{x\to\infty} \dfrac{1-F(x-\epsilon)}{1-F(x)} = e^{c\,\epsilon}$

de manera que:

$$\frac{F(x) - F(x-\varepsilon)}{\varepsilon(1-F(x))} < \frac{f(x)}{1-F(x)} < \frac{F(x+\varepsilon) - F(x)}{\varepsilon(1-F(x))}$$

que implica: $1/\varepsilon \left(\frac{1-F(x-\varepsilon)}{1-F(x)} - 1\right) < \frac{f(x)}{1-F(x)} < 1/\varepsilon \left(1 - \frac{1-F(x+\varepsilon)}{1-F(x)}\right)$

Como además: $\lim_{x\to\infty} \left(\frac{1-F(x-\varepsilon)}{1-F(x)} - 1\right) = e^{c\varepsilon}-1$

$$\lim_{x\to\infty} \left(1 - \frac{1-F(x+\varepsilon)}{1-F(x)}\right) = 1 - e^{-c\varepsilon}$$

y desarrollando en serie la función exponencial:

$$\frac{e^{c\varepsilon}-1}{\varepsilon} = c + o(\varepsilon) \quad ; \quad \frac{1-e^{-c\varepsilon}}{\varepsilon} = c - o(\varepsilon)$$

cuando $\varepsilon \to 0$ para $0 < c < \infty$.

Lo aplicamos en la cadena de desigualdades anterior

$$c - o(\varepsilon) < \frac{f(x)}{1-F(x)} < c + o(\varepsilon) \text{ para x suficientemente grande}$$

y obtenemos: $\lim_{x\to\infty} \frac{f(x)}{1-F(x)} = c$ c.q.d.

<u>TEOREMA 4.1.4</u>: Una función de distribución F con $x_F = \infty$ y

función de densidad F será SP si y solo si para todo $\epsilon > 0$,

$$\lim_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)} = 1$$, suponiendo F estrictamente monótona.

DEMOSTRACION:

=>) Supongamos que F es SP, entonces $S_n \xrightarrow{P} \infty$ lo que es equivalente a $\lim_{x\to\infty} \frac{f(x)}{1-F(x)} = 0$.

o sea $$\lim_{x\to\infty} \frac{d \log(1-F(x))}{dx} = 0$$

por tanto para todo $M > 0$, existe x_0 tal que:

$$\left| \frac{d \log(1-F(x))}{dx} \right| < M \quad \text{para todo } x > x_0$$

lo que es igual $$-M < \frac{d \log(1-F(x))}{dx} < 0$$

ya que $\log(1-F(x))$ es una función monótona decreciente

Sea un $\epsilon > 0$ cualquiera y $x' > x_0$

$$-M\epsilon < \int_{x'}^{x'+\epsilon} d \log(1-F(x)) < 0$$

así

$$-M\epsilon < \log \frac{(1-F(x'+\epsilon))}{1-F(x')} < 0$$

que nos lleva a que: $\lim\limits_{x \to \infty} \log \dfrac{1-F(x+\varepsilon)}{1-F(x)} = 0$ ó lo que

es igual: $\lim\limits_{x \to \infty} \dfrac{1-F(x+\varepsilon)}{1-F(x)} = 1$

<=) Supongamos que para todo $\varepsilon > 0$, $\lim\limits_{x \to \infty} \dfrac{1-F(x+\varepsilon)}{1-F(x)} = 1 =$

$= \lim\limits_{x' \to \infty} \dfrac{1-F(x')}{1-F(x'-\varepsilon)}$

Como $x_F = \infty$ y F estrictamente monótona, para valores altos la función de distribución F es cóncava, y así necesariamente se tiene:

$$\frac{F(x+\varepsilon) - F(x)}{\varepsilon} < f(x) < \frac{F(x) - F(x-\varepsilon)}{\varepsilon} \quad \text{para todo } \varepsilon > 0$$

y también

$$1/\varepsilon \left(1 - \frac{1-F(x+\varepsilon)}{1-F(x)} \right) < \frac{f(x)}{1-F(x)} < 1/\varepsilon \left(\frac{1-F(x-\varepsilon)}{1-F(x)} - 1 \right)$$

Como los dos extremos convergen a cero cuando $x \to \infty$ nos queda:

$\lim\limits_{x \to \infty} \dfrac{f(x)}{1-F(x)} = 0 \iff S_n \xrightarrow{P} \infty$ por el Teorema 4.1.2., o sea F

es SP c.q.d.

Utilizando la definición de función de variación lenta obtenemos una condición suficiente para la propensión:

TEOREMA 4.1.5: Si F es una función de distribución tal que $U(x) = 1-F(x)$, para $x > 0$ es una función de variación lenta en infinito, se verifica que F es SP.

DEMOSTRACIÓN:

Tendremos que demostrar:

$$\lim_{x\to\infty} \frac{1-F(kx)}{1-F(x)} = 1, \ \forall\, k > 0 \Rightarrow \lim_{x\to\infty} \frac{1-F(x+\varepsilon)}{1-F(x)} = 1 \ , \ \forall \varepsilon > 0$$

sea $k > 1$, entonces existe $x_0 = k/k-1$ tal que

$$kx_0 = k + x_0 \quad y \quad kx > x + k \quad \forall\, x > x_0$$

es decir: $\frac{1-F(kx)}{1-F(x)} < \frac{1-F(x+k)}{1-F(x)} < 1 \quad , \forall\, x > 0$

por tanto $\lim_{x\to\infty} \frac{1-F(kx)}{1-F(x)} = 1 \Rightarrow \lim_{x\to\infty} \frac{1-F(x+k)}{1-F(x)} = 1$

sea $0 < k < 1$, entonces consideramos $y = kx$

$$\lim_{x\to\infty} \frac{1-F(kx)}{1-F(x)} = \lim_{y\to\infty} \frac{1-F(y)}{1-F(\frac{1}{k}y)} = 1$$

equivalente a que:

$$\lim_{y\to\infty} \frac{1-F(\alpha y)}{1-F(y)} = 1 \qquad \text{con } \alpha > 1$$

y por el caso anterior, se verifica:

$$\lim_{y\to\infty} \frac{1-F(y+\alpha)}{1-F(y)} = 1 \quad \text{con } \alpha = 1/k$$

Además si $0 < \epsilon_1 < 1 < \epsilon_2$ se tiene:

$$\frac{1-F(y+\epsilon_2)}{1-F(y)} < \frac{1-F(y+\epsilon_1)}{1-F(y)} < 1$$

por lo que $\lim_{y\to\infty} \frac{1-F(y+\epsilon_2)}{1-F(y)} = 1 \Rightarrow \lim_{y\to\infty} \frac{1-F(y+\epsilon_1)}{1-F(y)} = 1$

c.q.d.

En el caso de distribuciones resistentes habrá que considerar el concepto de variación rápida (DEFINICION 2.2.3.)

TEOREMA 4.1.6: Si F es una función de distribución con $x_F = \infty$ y resistente con la suma a datos atípicos (SR), entonces la función definida como $U(x) = 1-F(x)$ para $x > 0$ es de variación rápida con índice de variación $\rho = -\infty$.

DEMOSTRACION:

Si F es SR esto es equivalente a que:

$$\forall \epsilon > 0 \quad \lim_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)} = 0$$

y como para todo $\epsilon > 0$ y $k > 1$ existe un $x_0(k,\epsilon)$ tal que:

$x + \epsilon \leqslant kx \quad \forall x \geqslant x_0$, se tiene

$$\frac{1-G(x+\epsilon)}{1-G(x)} > \frac{1-G(kx)}{1-G(x)} > 0 , \quad \forall x > x_0(k,\epsilon)$$

y así $\lim\limits_{x\to\infty} \dfrac{1-G(x+\varepsilon)}{1-G(x)} = 0$, $\forall\, \varepsilon>0 \Rightarrow \lim\limits_{x\to\infty} \dfrac{1-G(kx)}{1-G(x)} = 0$, $\forall\, k>1$

Utilizando lo anterior para $k'<1$

$$\lim_{x\to\infty} \frac{1-G(k'x)}{1-G(x)} = \infty \ .$$

es decir $1-G$ es de variación rápida con $\rho = -\infty$

c.q.d.

Ejemplos:

Con estas últimas caracterizaciones se pueden clasificar de un modo sencillo algunas de las distribuciones conocidas:

SR: Normal y Rayleigh.

SN: Exponencial, Logística, Doble exponencial $\Lambda(x)$, y Laplace.

SP: Pareto, la distribución de valores extremos $\Phi_\alpha(x)$ para cualquier $\alpha>0$, Cauchy y t_n-Student.

4.2 Caso de punto final finito

Si tenemos una función de distribución F con punto final x_F finito, donde $x_F = \sup\,(x$ tal que $F(x)<1)$, se verifica que $S_n \xrightarrow{P} 0$ cuando $n\to\infty$, siendo $S_n = X_{n:n} - X_{n-1:n}$ (PROPOSICION 3.2). Por tanto todo este tipo de distribuciones serían resistentes con la suma, es decir SR con la notación del apartado anterior.

Este comportamiento es perfectamente congruente con la intuición, pero nos interesará discriminar dentro de estas distribuciones, según la forma de converger a la unidad en las proximidades del punto final.

Para esto, si X es la variable aleatoria con función de distribución F, pasaremos a la variable aleatoria $Y = \dfrac{1}{x_F - X}$, cuya función de distribución es $G(y) = F(x_F - 1/y)$ tal que su punto final $x_G = \infty$, con lo que podemos aplicar las definiciones del caso anterior.

El paso a la variable Y puede justificarse en la forma de estudiar las variaciones de las diferentes colas de las distribuciones. Según esto como nos interesa el comportamiento de 1-F a la izquierda de x_F, esto será el mismo que el de H, donde $H(z) = 1-F(x_F-z)$, a la derecha del cero.

Pero H, es la función de distribución de $Z = x_F - X$. No obstante interesa trasladar el estudio de la cola de una función de distribución al de la cola de otra, por lo que pasamos a la función G, donde $1-G(y) = H(1/y)$, ya que el comportamiento de 1-G en el infinito coincide con el de H en el cero. Así llegamos a estudiar $G(y) = F(x_F - 1/y)$ que es la función de distribución de la variable aleatoria $Y = \dfrac{1}{x_F - X}$

4.2.1. Definiciones de distribuciones SP^*, SN^* y SR^*

Según la introducción anterior si X es la variable aleatoria con punto final finito, las definiciones sobre su caracter en cuanto a datos atípicos van a surgir de las definiciones que ya tenemos sobre la variable $Y = \dfrac{1}{x_F - X}$. Por tanto veremos en primer lugar la relación que existe entre:

$S_n = X_{n:n} - X_{n-1:n}$ y $S'_n = Y_{n:n} - Y_{n-1:n}$.

Evidentemente si $X_{1:n} \leqslant \ldots \leqslant X_{n:n}$ se verificará $Y_{1:n} \leqslant \ldots$ $\ldots \leqslant Y_{n:n}$ donde $Y_{k:n} = \dfrac{1}{x_F - X_{k:n}}$, así:

$$S'_n = \frac{1}{x_F - X_{n:n}} - \frac{1}{x_F - X_{n-1:n}} = \frac{X_{n-1:n} - X_{n:n}}{(x_F - X_{n:n})\,(x_F - X_{n-1:n})} =$$

$$= \frac{S_n}{(x_F - X_{n:n})\,(x_F - X_{n-1:n})}$$

y como $\dfrac{1}{x_F - X_{n:n}} \leqslant \dfrac{1}{x_F - X_{n:n}}$ podemos acotar:

$$0 \leqslant \frac{S_n}{(x_F - X_{n-1:n})^2} \leqslant S'_n \leqslant \frac{S_n}{(x_F - X_{n:n})^2} \qquad (1)$$

Pero tenemos que: $\begin{cases} S_n \xrightarrow{P} 0 \\ (x_F - X_{n:n}) \xrightarrow{P} 0 \\ (x_F - X_{n-1:n}) \xrightarrow{P} 0 \end{cases}$

Por tanto el límite de S'_n dependerá de las velocidades relativas de convergencia a cero en probabilidad de las variables aleatorias anteriores.

Veamos los diferentes casos que se pueden presentar:

CASO 1: Si $\dfrac{S_n}{(x_F-X_{n:n})^2} \xrightarrow{P} 0$ cuando $n\to\infty$, se verifica que $S'_n \xrightarrow{P} 0$ cuando $n\to\infty$ que equivale a exigir que G, la función de distribución de la variable aleatoria Y sea SR. Así podemos enunciar la correspondiente definición y un resultado sobre el comportamiento de S_n en distribuciones resistentes.

DEFINICION 4.2.1: Dada una función de distribución F con punto final x_F finito, diremos que es resistente con la suma y la denominaremos SR^* si la función de distribución G, definida como $G(y) = F(x_F-1/y)$ cuyo punto final x_G es infinito, es SR.

PROPOSICION 4.2.1: Una función de distribución F, con $x_F<\infty$ será SR^* si $\dfrac{S_n}{(x_F-X_{n:n})^2} \xrightarrow{P} 0$ cuando $n\to\infty$.

DEMOSTRACION: Utilizando la desigualdad de la derecha en (1).

CASO 2: Si en la cadena de desigualdades (1) se tiene:

$\dfrac{S_n}{(x_F-X_{n:n})^2} \equiv O_P(1)$ es lo mismo que exigir:

Para todo $\epsilon>0$, existe c_ϵ, N_ϵ tales que: $Pr(\dfrac{S_n}{(x_F-X_{n:n})^2} > c_\epsilon) < \epsilon$

$\forall\ n \geq N_\epsilon$

pero $Pr(S'_n > c_\epsilon) \leq P(S_n/(x_F-X_{n:n})^2 > c_\epsilon) < \epsilon$

que implica: $S'_n \equiv O_P(1)$

Si exigimos además que $\dfrac{S_n}{(x_F-X_{n:n})^2} \xrightarrow{P}\!\!\!\!\!\!/\;\; 0 \Rightarrow S'_n \neq o_P(1)$

Por lo tanto igual que en el caso anterior nos quedará:

<u>DEFINICION 4.2.2</u>: Dada una función de distribución F con punto final x_F finito, diremos que es neutra* con la suma y la denominaremos SN^* si la función de distribución G, definida como $G(y) = F(x_F-1/y)$ cuyo punto final x_G es infinito, es SN.

<u>PROPOSICION 4.2.2</u>: Para que una función de distribución F con $x_F < \infty$ sea SN^* es suficiente que

$$\begin{cases} \dfrac{S_n}{(x_F-X_{n:n})^2} \equiv O_P(1) \\ \text{y} \\ \dfrac{S_n}{(x_F-X_{n-1:n})^2} \neq o_P(1) \end{cases}$$

<u>DEMOSTRACION</u>: Con las desigualdades (1).

<u>CASO 3</u>: Si en (1) se tiene $\dfrac{S_n}{(x_F-X_{n-1:n})^2} \xrightarrow{P} \infty \Rightarrow$

$\Rightarrow S'_n \xrightarrow{P} \infty$ que equivale a exigir que la variable aleatoria

$Y = \dfrac{1}{x_F-X}$ sea propensa con la suma, por lo tanto:

<u>DEFINICION 4.2.3</u>: Sea F una función de distribución con $x_F < \infty$,

diremos que es propensa * con la suma y la denominaremos SP^* si la función de distribución G, definida como $G(y) = F(x_F - 1/y)$, cuyo punto final $x_G = \infty$, es SP.

PROPOSICION 4.2.3: Dada una función de distribución F con $x_F < \infty$, será SP^* si $\dfrac{S_n}{(x_F - X_{n-1:n})^2} \xrightarrow{P} \infty$.

DEMOSTRACION: Utilizando (1).

También podríamos enunciar las correspondientes condiciones necesarias basándonos en (1).

PROPOSICION 4.2.4: Dada una función de distribución F con $x_F < \infty$ se tiene que:

i) Si F es SR^* se verifica que $\dfrac{S_n}{(x_F - X_{n-1:n})^2} \xrightarrow{P} 0$

ii) Si F es SN^* se verifica que
$$\left\{ \begin{array}{l} \dfrac{S_n}{(x_F - X_{n-1:n})^2} = O_P(1) \\ \text{y} \\ \dfrac{S_n}{(x_F - X_{n-1:n})^2} \neq o_P(1) \end{array} \right.$$

iii) Si F es SP^* se verifica que $\dfrac{S_n}{(x_F - X_{n:n})^2} \xrightarrow{P} \infty$.

COMENTARIOS: En resumen las definiciones pueden basarse en la idea de que máximos alejados de la variable $Y = \frac{1}{x_F - X}$ corresponden a máximos próximos a x_F de la variable X y así los distintos comportamientos de Y se traducen en lo mismo para X.

4.2.2. Caracterización con propiedades de las funciones de distribución

Para los siguientes resultados tendremos en cuenta como se traslada el estudio de la variación de una función en infinito a su variación en cero y por último en cualquier valor finito, que es el objetivo de este apartado. Supondremos que $x_F > 0$.

PROPOSICION 4.2.5: Si una función de distribución F con $x_F < \infty$ es SR^* entonces la función definida por $U(x) = 1-F(x)$ para $x > 0$ es de variación rápida con $\rho = -\infty$, a la izquierda de x_F.

DEMOSTRACION: Considerando la función de distribución G, tal que $G(y) = F(x_F - 1/y)$, por ser F, SR^* será G, SR, y por el Teorema 4.1.6 esto implica que la función definida por $H(y) = 1-G(y)$ para $y > 0$ es de variación rápida con $\rho = -\infty$ en infinito.

Pero que $H(y) = 1-G(y)$ sea $(-\infty)$-variante en infinito es equivalente a que $H(1/y) = 1-G(1/y)$ sea de variación rápida con $\rho = \infty$ en cero lo que es equivalente a que $U(y) = H(\frac{1}{x_F - y}) = 1 - G(\frac{1}{x_F - y}) = 1 - F(y)$ sea de variación rápida con $\rho = -\infty$ a la izquierda de x_F .

c.q.d.

Teniendo en cuenta resultados anteriores sabemos que en las condiciones exigidas a la función de distribución F en este apartado el que $\frac{S_n}{(x_F-X_{n:n})^2} \xrightarrow{P} 0$ implica que F es SR^*. No obstante partiendo de que 1-F sea de variación rápida obtenemos una condición más general como es que:

$\frac{S_n}{(x_F-X_{n:n})} \xrightarrow{P} 0$ cuando $n\to\infty$, ya que:

$$\frac{S_n}{(x_F-X_{n:n})^2} \xrightarrow{P} 0 \quad \text{cuando } n\to\infty \Rightarrow \frac{S_n}{(x_F-X_{n:n})} \xrightarrow{P} 0$$

cuando $n\to\infty$, no siendo cierta siempre la implicación inversa pues $(x_F-X_{n:n}) \xrightarrow{P} 0$ para $n\to\infty$.

Antes de enunciar el teorema correspondiente, daremos un lema cuyo resultado utilizaremos en la demostración.

<u>LEMA 4.2.1</u>: Sea una función de distribución F con $x_F<\infty$ y función de densidad f, tal que la función $U(x) = 1-F(x)$ para $x > 0$ es de variación rápida con $\rho = -\infty$, a la izquierda de x_F, entonces se verifica que: $\frac{x_F-X_{n:n}}{x_F-X_{n-1:n}} \xrightarrow{P} 1$ cuando $n\to\infty$.

<u>DEMOSTRACION:</u>

Sea X la variable aleatoria cuya función de distribución es F; si definimos a partir de ella la variable

$Y = \frac{1}{x_F-X} \geqslant 0$, tenemos:

$$\frac{x_F - X_{n:n}}{x_F - X_{n-1:n}} = \frac{Y_{n-1:n}}{Y_{n:n}}$$

por lo que hemos de demostrar que $\dfrac{Y_{n-1:n}}{Y_{n:n}} \xrightarrow{P} 1$

La función de densidad conjunta de $(Y_{n-1:n}, Y_{n:n})$:

$$f_{Y_{n-1:n}}(x,y) = n(n-1)\ (G(y))^{n-2}\ g(x)\ g(y) \quad \text{tal que } 0 \leqslant x \leqslant y$$

donde $G(y) = F(x_F - 1/y)$ es la función de distribución de la variable Y, y g es su función de densidad. Si definimos la variable $Z_n = \dfrac{Y_{n-1:n}}{Y_{n:n}}$, su función de distribución será:

para $0 < z < 1$;

$$H_n(z) = \int_0^{\infty} \int_{x/z}^{\infty} n(n-1)\ (G(x))^{n-2}\ g(x)\ g(y)\ dy =$$

$$= \int_0^{\infty} \frac{1-G(x/z)}{1-G(x)}\ d\ G_{Y_{n-1:n}}(x)$$

donde $G_{Y_{n-1:n}}(x)$, es la función de distribución de la variable aleatoria $Y_{n-1:n}$.

$$G_{Y_{n-1:n}}(x) = (G(x))^n + n(G(x))^{n-1}\ (1-G(x))$$

que por ser $x_G = \infty$, será:

$$G_{Y_{n-1:n}}(x) \longmapsto 0 \quad \text{cuando } n \to \infty, \ \forall\, x > 0$$

Si llamamos $h_z(x) = \dfrac{1-G(x/z)}{1-G(x)}$ será:

$0 < \dfrac{1-G(x/z)}{1-G(x)} < 1$, ó sea $h_z(x)$ acotada para $0 < z < 1$

y $h_z(x)$ continua para cada z, por ser G absolutamente continua.

Además $\lim\limits_{x \to \infty} h_z(x) = \lim\limits_{x \to \infty} \dfrac{1-G(x/z)}{1-G(x)} = 0$, para cada $0 < z < 1$,

por ser G tal que $U(x) = 1-G(x)$ en $x > 0$ es de variación rápida en infinito.

En estas condiciones: (Chow, Teicher (1978))

$$\lim_{n \to \infty} \int_0^{\infty} \frac{1-G(x/z)}{1-G(x)} \, d\, G_{Y_{n-1:n}}(x) = 0 \text{ , para cada } 0 < z < 1$$

Luego:

$$H_n(z) \longmapsto H(z) = \begin{cases} 0 & z < 1 \\ 1 & z \geq 1 \end{cases} \quad \text{cuando } n \to \infty$$

lo que implica: $\dfrac{Y_{n-1:n}}{Y_{n:n}} \xrightarrow{P} 1$ cuando $n \to \infty$

c.q.d.

TEOREMA 4.2.1: Si la función de distribución absolutamente continua F, con $x_F < \infty$ es tal que la función U definida por $U(x) = 1-F(x)$ para $x > 0$ es de variación rápida con $\rho = \infty$ a la izquierda de x_F, entonces

$$\frac{S_n}{(x_F - X_{n:n})} \xrightarrow[n\to\infty]{P} 0$$

DEMOSTRACION:

Como $S_n = X_{n:n} - X_{n-1:n} = (x_F - X_{n-1:n}) - (x_F - X_{n:n})$

se tiene: $$\frac{S_n}{(x_F - X_{n:n})} = \frac{x_F - X_{n-1:n}}{x_F - X_{n:n}} - 1$$

y por el Lema 4.2.1. sabemos que $\dfrac{x_F - X_{n-1:n}}{x_F - X_{n:n}} \xrightarrow{P} 1$

cuando $n \to \infty$, por lo tanto:

$$\frac{S_n}{(x_F - X_{n:n})} \xrightarrow{P} 0 \text{ cuando } n \to \infty \qquad \text{c.q.d.}$$

Para establecer equivalencias con las definiciones iniciales, habrá que exigir alguna condición más restrictiva a la función de distribución F.

TEOREMA 4.2.2: Sea una función de distribución F absolumente continua, con función de densidad f y punto final x_F finito, entonces:

i) F es SR^* $\iff \lim_{x \to x_F^-} \dfrac{1-F(x)}{f(x)\,(x_F-x)^2} = 0$

ii) F es SN^* $\iff \lim_{x \to x_F^-} \dfrac{1-F(x)}{f(x)\,(x_F-x)^2} = c$ donde $0 < c < \infty$

iii) F es SP^* $\iff \lim_{x \to x_F^-} \dfrac{1-F(x)}{f(x)\,(x_F-x)^2} = \infty$

DEMOSTRACION:

Por las definiciones correspondientes sabemos que el comportamiento de F es equivalente al comportamiento de la función de distribución G, definida por $G(y) = F(x_F-1/y)$, con $x_G = \infty$.

Entonces como $\lim_{y \to \infty} \dfrac{1-G(y)}{g(y)} = \lim_{x \to x_F^-} \dfrac{1-F(x)}{f(x)\,(x_F-x)^2}$

se obtienen las condiciones anteriores c.q.d.

Aplicando este teorema se pueden verificar los siguientes casos particulares

EJEMPLOS:

La familia de funciones de distribución $(F_\alpha)_{\alpha>0}$ donde

$$F_\alpha(x) = \begin{cases} 1 - e^{-1/(x_F-x)^\alpha} & x \leq x_F \\ 1 & x \geq x_F \end{cases}$$

será: SR^* para $\alpha > 1$

SN^* para $\alpha = 1$ donde $x_F < \infty$.

SP^* para $\alpha < 1$

La distribución Uniforme es SP^*, al igual que la familia con crecimiento polinomial en un punto finito x_F dado por:

$$F_\alpha(x) = \begin{cases} 0 & x \leq x_F - k^{-1/\alpha} \\ 1 - k(x_F - x)^\alpha & x_F - k^{-1/\alpha} \leq x \leq x_F \\ 1 & x \geq x_F \end{cases}$$

5. DISTRIBUCIONES PROPENSAS (RP), NEUTRAS (RN) Y RESISTENTES (RR) SEGUN EL COMPORTAMIENTO ASINTOTICO DE $R_n = \dfrac{X_{n:n}}{X_{n-1:n}}$

Procederemos de forma análoga al apartado anterior estudiando la razón entre los dos extremos superiores. Tambíen procederemos a relacionarlas con las definiciones relativas a la suma.

5.1. Caso de punto final infinito

Daremos las definiciones, teniendo en cuenta que nuestro concepto de resistencia con la razón coincide con la resistencia relativa de Green (1976) y las demas nociones forman parte de la extensión que también efectuamos en el caso de la suma.

5.1.1. Definiciones de distribuciones RR, RN y RP

DEFINICION 5.1.1: Dada una función de distribución F con $x_F = \infty$, diremos que es RESISTENTE con la razón a datos atípicos y la denominaremos RR si $R_n \xrightarrow{P} 1$ siendo $R_n = \dfrac{X_{n:n}}{X_{n-1:n}}$.

DEFINICION 5.1.2: Dada una función de distribución F con $x_F = \infty$, diremos que es NEUTRA con la razón a datos atípicos y la denominaremos RN si R_n, ya definido, es:

$$\left\{ \begin{array}{l} R_n = O_P(1) \\ \text{y} \\ R_n - 1 \neq o_P(1) \end{array} \right.$$

DEFINICION 5.1.3: Dada una función de distribución F con $x_F = \infty$, diremos que es PROPENSA con la razón a datos atípicos y la denominaremos RP si $R_n \xrightarrow{P} \infty$.

La convergencia en probabilidad a infinito y las notaciones O_P y o_p fueron expuestas en el apartado anterior.

A continuación demostraremos un resultado análogo al de Schuster (1984) que se refiere a S_n y que ya fue utilizado en el apartado anterior, pero relativo a R_n.

Supongamos que F es absolutamente continua, con función de densidad f y llamemos:

$Q(u) = F^{-1}(u) = \inf(x \;/\; F(x) \geq u)$, función cuantílica

$U_{k:n}$ al estadístico de orden k, en una muestra de tamaño n de una $U(0,1)$.

Dado que $x_F = \infty$ y que $\left\{\begin{array}{l} X_{n:n} \xrightarrow{w} x_F \\ y \\ X_{n-1:n} \xrightarrow{w} x_F \end{array}\right.$ cuando $n \to \infty$

podemos considerar que existe un N_0 tal que tanto $X_{n:n} \geq 0$ como $X_{n-1:n} \geq 0$ para todo $n \geq N_0$, lo que justifica el tomar logaritmos antes de efectuar los límites. Así tenemos que:

$$\log R_n = \log X_{n:n} - \log X_{n-1:n} = \log Q(U_{n:n}) - \log Q(U_{n-1:n})$$

si suponemos que en un entorno de uno existe la función cuantílica Q y es diferenciable con derivada continua $1/f(Q)$, siguiendo con lo anterior:

$$\log R_n = \frac{1}{f(Q(U_n^*))\ Q(U_n^*)} \quad (U_{n:n} - U_{n-1:n})$$

donde $U_{n-1:n} \leq U_n^* \leq U_{n:n}$

Llamando $H(u) = \dfrac{1}{f(Q(u))\ Q(u)}$ nos queda:

$$\log R_n = H(U_n^*)\ \frac{(U_{n:n} - U_{n-1:n})}{1 - U_n^*}$$

y por lo tanto

$$H(U_n^*)\ \left(\frac{U_{n:n} - U_{n-1:n}}{1 - U_{n-1:n}}\right) \leq \log R_n \leq H(U_n^*)\ \left(\frac{U_{n:n} - U_{n-1:n}}{1 - U_{n:n}}\right)$$

ea $Z_n = \dfrac{1 - U_{n:n}}{1 - U_{n-1:n}}$, entonces se demuestra que es una variable con distribución Uniforme en (0,1). Luego:

$$H(U_n^*)\ (1 - Z_n) \leqslant \log R_n \leqslant H(U_n^*)\ (Z_n^{-1} - 1)$$

Ambas sucesiones de variables $(1-Z_n)$ y $(Z_n^{-1}-1)$ estan acotadas en probabilidad es decir son $O_p(1)$ por lo que el comportamiento asintótico de $\log R_n$ viene determinado por el de $H(U_n^*)$. Co- com ademas $U_n^* \xrightarrow{P} 1$ cuando $n \to \infty$, no dependerá de los distintos valores del límite, si existe: $k_1 = \lim\limits_{u \to 1^-} \dfrac{1 - u}{f(Q(u))Q(u)} =$

$= \lim\limits_{x \to \infty} \dfrac{1-F(x)}{x\ f(x)}$.

<u>TEOREMA 5.1.1.</u>: Sea una función de distribución absolutamente continua F con $x_F = \infty$, tal que la función cuantílica existe y es diferenciable con derivada continua en un entorno de la unidad. Supongamos que k_1 existe siendo posiblemente infinito, entonces:

i) $k_1 = 0 \Leftrightarrow \log R_n \xrightarrow{P} 0 \Leftrightarrow R_n \xrightarrow{P} 1$ cuando $n \to \infty$

ii) $k_1 = c\ (0 < c < \infty) \Leftrightarrow \left\{ \begin{array}{l} \log R_n \equiv O_p(1) \\ y \\ \log R_n \not\equiv o_p(1) \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} R_n \equiv O_p(1) \\ y \\ R_n - 1 \not\equiv o_p(1) \end{array} \right.$

cuando $n \to \infty$

iii) $k_1 = \infty \Leftrightarrow \log R_n \xrightarrow{P} \infty \Leftrightarrow R_n \xrightarrow{P} \infty$ cuando $n \to \infty$

DEMOSTRACION: Utilizando el razonamiento previo al enunciado del teorema:

i) Supongamos $k_1 = 0$, como $U_n^* \xrightarrow{P} 1$ y H es continua se verifica $H(U_n^*) \xrightarrow{P} 0$ y como $(Z_n^{-1} - 1)$ y $(1 - Z_n)$ estan acotadas en probabilidad,

$$\log R_n \xrightarrow{P} 0 \iff R_n \xrightarrow{P} 1 \quad \text{cuando } n \to \infty$$

Ahora supongamos $\log R_n \xrightarrow{P} 0$; como para n suficientemente grandes $U_n^* \gtrsim 1$, esto hace que en ese caso $H(U_n^*) \geq 0$ y así se verifica $H(U_n^*)\ (1 - Z_n) \xrightarrow{P} 0$ es decir $H(U_n^*) \xrightarrow{P} 0 \Rightarrow k_1 = 0$

ii) Supongamos $k_1 = c$ donde $0 < c < \infty$, entonces

$$\log R_n \leq H(U_n^*)\ (Z_n^{-1} - 1) \Rightarrow \log R_n \equiv O_P(1)$$

$$y \ \log R_n \geq H(U_n^*)\ (1 - Z_n) = c(1-Z_n) + o_P(1) \Rightarrow \log R_n \not\equiv o_P(1)$$

$$\text{Si } \log R_n \equiv O_P(1) \Rightarrow H(U_n^*)\ (1-Z_n) \equiv O_P(1) \Rightarrow H(U_n^*) \equiv O_P(1)$$

$$y \ \log R_n \not\equiv o_P(1) \Rightarrow H(U_n^*)\ (Z_n^{-1} - 1) \not\equiv o_P(1) \Rightarrow H(U_n^*) \not\equiv o_P(1)$$

lo que nos lleva a $k_1 = c$, donde $0 < c < \infty$.

iii) Suponemos $k_1 = \infty$, entonces $H(U_n^*) \xrightarrow{P} \infty$ cuando $n \to \infty$, esto es: $H(U_n^*)\ (1-Z_n) \xrightarrow{P} \infty$ que nos dá $\log R_n \xrightarrow{P} \infty$ cuando $n \to \infty$.

Inversamente si $\log R_n \overset{P}{\longmapsto} \infty \Rightarrow H(U_n^*)\,(Z_n^{-1} - 1) \overset{P}{\longmapsto} \infty \Rightarrow$ $\Rightarrow H(U_n^*) \overset{P}{\longmapsto} \infty$, todas las convergencias cuando $n \to \infty$, Por lo anterior $k_1 = \infty$.

c.q.d.

5.1.2.- Relación con las definiciones de distribuciones SR, SN y SP.

PROPOSICION 5.1.1: i) Si F es una función de distribución SR, se verifica que F es RR.

ii) Si F es una función de distribución SN, se verifica que F es RR.

DEMOSTRACION: Como $\dfrac{S_n}{X_{n:n}} \to (1 - R_n^{-1})$, el que $\dfrac{1}{X_{n:n}} \overset{P}{\longmapsto} 0$ cuando $n \to \infty$, nos lleva a las implicaciones

i) $S_n \overset{P}{\longmapsto} 0 \Rightarrow R_n \overset{P}{\longmapsto} 1$ para $n \to \infty$

ii) $\left.\begin{array}{l} S_n \equiv O_P(1) \\ \text{y} \\ S_n \neq o_P(1) \end{array}\right\} \Rightarrow R_n \overset{P}{\longmapsto} 1$ para $n \to \infty$.

PROPOSICION 5.1.2: Si F es una función de distribución RN se verifica que es tambien SP.

DEMOSTRACION: Partimos de la igualdad:

$$S_n = X_{n:n}\,(1 - R_n^{-1}) = X_{n:n}\left(\frac{R_n - 1}{R_n}\right)$$

y como $X_{n:n} \xrightarrow{P} \infty$ cuando $n \to \infty$

$$\left.\begin{array}{l} R_n \equiv O_P(1) \\ y \\ R_n - 1 \neq o_P(1) \end{array}\right\} \Rightarrow S_n \xrightarrow{P} \infty$$

PROPOSICION 5.1.3: Si F es una función de distribución RP, se tiene que es SP.

DEMOSTRACION: Utilizando de nuevo $S_n = X_{n:n}(1 - R_n^{-1})$ y razonando como en la proposicion anterior.

OBSERVACIONES: De las relaciones expuestas surge la congruencia con la intuición de que si dos observaciones están a una distancia pequeña entonces su razón se acercará a la unidad e inversamente si su cociente es alto se implicará que su distancia también lo es.

En resumen el concepto mas fuerte de propensión se obtiene con la razón y el de resistencia con la suma.

5.1.3. Caracterización con propiedades de las funciones de distribución

TEOREMA 5.1.2: Dada una función de distribución F, con $x_F = \infty$, será RR si y solo si la función $U(x) = 1 - F(x)$ para $x > 0$ es de variación rápida con $\rho = -\infty$ en infinito.

DEMOSTRACION:

Si F es RR $\Leftrightarrow R_n \xrightarrow{P} 1$ y esto segun Geffroy (1959) es equivalente a la estabilidad relativa de la sucesión de los máximos que es equivalente a su vez (Gnedenko (1943)) con que:

$$\lim_{x \to \infty} \frac{1 - F(kx)}{1 - F(x)} = 0 \qquad \forall\, k > 1$$

y esta es la definición de función de variación rápida con $\rho = -\infty$.

Para los siguientes resultados tendremos que hacer algunas hipótesis no demasiado restrictivas sobre las funciones de distribución consideradas.

TEOREMA 5.1.3: Sea una función de distribución F con $x_F = \infty$, absolutamente continua y monótona con derivada continua y positiva para $x's$ suficientemente grandes.

i) Si F es RN se verifica que la función $U(x) = 1-F(x)$ para $x > 0$ es de variación regular con exponente de regularidad $\rho = -\alpha$, para algun $\alpha > 0$.

ii) Si la función $U(x) = 1 - F(x)$ para $x > 0$, es de variación regular con exponente $\rho = -\alpha$, para algun $\alpha > 0$, y la función de densidad f es monótona no creciente, entonces F es RN.

DEMOSTRACION:

i) Si F es RN $\iff \lim_{n \to \infty} \frac{x\, f(x)}{1 - F(x)} = c$ siendo $0 < c < \infty$.

Entonces por los resultados de L de Haan (1970) (Teorema 1.2.1 b) y Lema 1.2.2 b)) se tiene que la función $U(x) = 1 - F(x)$ para $x > 0$ es $(-c)$-variante.

ii) Para demostrar este apartado nos remitimos al Teorema 2.7.1 b) en de Haan (1970).

TEOREMA 5.1.4: Supongamos que la función de distribución F verifica las condiciones del teorema anterior entonces si F es RP equivale a que la función $U(x) = 1-F(x)$ para $x > 0$ es de variación lenta en infinito.

DEMOSTRACION:

$\Rightarrow$ Supongamos que F es RP, entonces $\lim_{x\to\infty} \frac{x\,f(x)}{1-F(x)} = 0$. Por tanto podemos utilizar las demostraciones de de Haan (1970) particularizadas a esta situación.

Llamemos $a(x) = \frac{x\,f(x)}{1-F(x)}$, entonces $a(x) \to 0$ cuando $x \to \infty$

$$\frac{a(x)}{x} = \frac{f(x)}{1-F(x)} \qquad \text{es decir} \qquad \frac{a(x)}{x} = -\frac{d\log(1-F(x))}{dx}$$

como podemos integrar desde algún valor $B > 0$:

$$1 - F(x) = c_0\, e^{-\int_B^x \frac{a(t)}{t}\,dt} \qquad \text{para} \quad x > B$$

Entonces si $k > 1$:

$$\frac{1-F(kx)}{1-F(x)} = \frac{c_0\, e^{-\int_B^{kx} \frac{a(t)}{t}\,dt}}{c_0\, e^{-\int_B^{x} \frac{a(t)}{t}\,dt}} = e^{-\int_1^k \frac{a(xu)}{u}\,du}$$

y asi

$$\lim_{x\to\infty} \frac{1-F(kx)}{1-F(x)} = 1$$

Analogamente para $k < 1$.

- Ahora supongamos que $U(x) = 1-F(x)$ para $x > 0$ es de variación lenta.

Sea $k > 1$; como $x_F = \infty$ y F es absolutamente continua y monótona se verifica que F es cóncava para valores suficientemente grandes. Es decir

$$\frac{F(x) - F((1/k)x)}{x - x/k} < f(x) < \frac{F(kx) - F(x)}{kx - x}$$

por lo que:

$$\frac{F(x) - F((1/k)x)}{(1-1/k)\ (1-F(x))} < \frac{x\ f(x)}{1-F(x)} < \frac{F(kx) - F(x)}{(k-1)\ (1-F(x))}$$

dado que los extremos de las desigualdades convergen a cero cuando $x \to \infty$, se tiene:

$$\lim_{x \to \infty} \frac{x\ f(x)}{1-F(x)} = 0 \iff F \quad \text{es RP}$$

c.q.d.

<u>EJEMPLOS</u>:

RR: Normal; Log-normal; Exponencial; Logística; Weibull; Rayleigh.

RN: Cauchy; Pareto; distribución de valores extremos de tipo II ($\Phi_\alpha(x)$ para cualquier $\alpha > 0$).

RP: La función de distribución $F(x) = \begin{cases} 0 & x < e \\ 1 - 1/\log_e x & x \geq e \end{cases}$

5.2.- Caso de punto final finito

Como en el caso del estudio de $S_n = X_{n:n} - X_{n-1:n}$, tenemos que si F es tal que su punto final es $x_F < \infty$ entonces F es RR; este resultado, intuitivamente valido, no nos permite sin embargo discriminar dentro de esta clase de distribuciones.

Procediendo de forma análoga al caso de la suma, efectuamos un cambio de variable que nos permita aplicar las definiciones del caso anterior. Así definimos la función de distribución $G(y) = F(x_F - 1/y)$, que corresponde a la variable aleatoria $Y = \dfrac{1}{x_F - X}$. Los estadísticos de orden $Y_{k:n} = \dfrac{1}{x_F - X_{k:n}}$

y la razón $R'_n = \dfrac{Y_{n:n}}{Y_{n-1:n}} = \dfrac{x_F - X_{n-1:n}}{x_F - X_{n:n}}$

entonces $R_n = \dfrac{X_{n:n}}{X_{n-1:n}}$,

$$R_n^{-1} = (R'_n - 1)\,\frac{x_F - X_{n:n}}{X_{n-1:n}}$$

$$y \quad R'_n - 1 = (R_n - 1)\,\frac{X_{n-1:n}}{x_F - X_{n:n}}$$

por lo tanto

$$\frac{1}{\left(\dfrac{x_F}{X_{n-1:n}} - 1\right)}(R_n - 1) \leqslant R'_n - 1 \leqslant \frac{1}{\left(\dfrac{x_F}{X_{n:n}} - 1\right)}(R_n - 1)$$

$$\text{donde} \quad \begin{cases} R_n - 1 \xrightarrow{P} 0 \\ \\ \left(\dfrac{x_F}{X_{n-1:n}} - 1\right) \xrightarrow{P} 0 \quad \text{cuando } n \to \infty \\ \\ \left(\dfrac{x_F}{X_{n:n}} - 1\right) \xrightarrow{P} 0 \end{cases}$$

Así el comportamiento asintótico de $R'_n - 1$ dependerá de las velocidades relativas de convergencia a cero de las variables citadas.

5.2.1. Definiciones de distribuciones RR^*, RN^* y RP^*

DEFINICION 5.2.1: Dada una función de distribución F con $x_F < \infty$ diremos que es resistente* con la razón y la denominaremos RR^* si la función de distribución $G(x) = F(x_F - 1/x)$ es RR es decir si $R'_n - 1 \xrightarrow{P} 0$ cuando $n \to \infty$.

DEFINICION 5.2.2: Dada una función de distribución F con $x_F < \infty$ diremos que es neutra* con la razón y la denominaremos RN^* si la función de distribución $G(x) = F(x_F - 1/x)$ es RN es decir

$$\begin{cases} R'_n = O_P(1) \\ y \qquad\qquad \text{cuando } n \to \infty \\ R'_n \neq o_P(1) \end{cases}$$

DEFINICION 5.2.3: Dada una función de distribución F con punto final $x_F < \infty$ diremos que es propensa* con la razón y la denominaremos RP^* si la función de distribución $G(x) = F(x_F - 1/x)$

es RP.

Condiciones necesarias y suficientes vienen recogidas en los siguientes resultados.

PROPOSICION 5.2.1: Dada una función de distribución F con punto final finito, se verifica que si para $n\to\infty$:

i) $$\frac{R_n - 1}{\dfrac{x_F}{X_{n:n}} - 1} \xrightarrow{P} 0 \Rightarrow F \text{ es } RR^*$$

ii) $$\left.\begin{array}{l} \dfrac{R_n - 1}{\dfrac{x_F}{X_{n:n}} - 1} = O_P(1) \\[2em] \dfrac{R_n - 1}{\dfrac{x_F}{X_{n-1:n}} - 1} \neq o_P(1) \end{array}\right\} \Rightarrow F \text{ es } RN^*$$

iii) Si $$\frac{R_n - 1}{\dfrac{x_F}{X_{n-1:n}}} \xrightarrow{P} \infty \Rightarrow F \text{ es } RP^*$$

DEMOSTRACION: En todos los casos utilizando la cadena de desigualdades (2).

PROPOSICION 5.2.2; Dada una función de distribución F con punto final finito se tiene:

i) F es RR* $\Rightarrow \dfrac{R_n - 1}{\dfrac{x_F}{X_{n-1:n}} - 1} \xrightarrow{P} 0$ cuando $n \to \infty$

ii)

$$\text{F es RN}^* \Rightarrow \left\{ \begin{array}{l} \dfrac{R_n - 1}{\dfrac{x_F}{X_{n-1:n}}} = O_P(1) \\ \text{y} \\ \dfrac{R_n - 1}{\dfrac{x_F}{X_{n:n}}} \neq o_P(1) \end{array} \right. \quad \text{cuando } n \to \infty$$

iii) Si F es RP* $\Rightarrow \dfrac{R_n - 1}{\dfrac{x_F}{X_{n:n}}} \xrightarrow{P} \infty$ cuando $n \to \infty$

<u>DEMOSTRACION</u>: Como en la proposición anterior, aplicando las desigualdades (2).

Haciendo uso de la relación con las definiciones relativas a la suma en el caso de punto final infinito, podemos resumir los resultados.

<u>PROPOSICION 5.2.3</u>: Dada F, función de distribución con punto final finito se tiene:

i) Si F es SR* $\Rightarrow$ F es RR*

ii) Si F es SN* $\Rightarrow$ F es RR*

iii) Si F es RN^* => F es SP^*

iv) Si F es RP^* => F es SP^*

DEMOSTRACION: Aplicando las proposiciones 5.1 1, 5.1.2 y 5.1.3. a la función de distribución $G(x) = F(x_F - 1/x)$.

5.2.2.- Caracterización con propiedades de las funciones de distribución.

Estas propiedades, como en las situaciones anteriores, serán relativas a la variación en la cola derecha de la función de distribución. Supondremos $0 < x_F < \infty$.

TEOREMA 5.2.1: Si F, función de distribución con $x_F < \infty$ es RR^* es equivalente a que la función $U(x) = 1 - F(x)$ para $x > 0$ es de variación rápida con exponente de regularidad $\rho = +\infty$.

DEMOSTRACION: Si la función de distribución F es RR^* <=> la función de distribución $G(x) = F(x_F - \frac{1}{x})$ es RR <=> la función $U(x) = 1 - G(x)$ es de variación rápida con $\rho = -\infty$ en infinito <=> $1 - F(x)$ es de variación rápida con $\rho = +\infty$ a la izquierda de x_F.

c.q.d.

TEOREMA 5.2.2: Si F, función de distribución con $x_F < \infty$ es RR^* se verifica que:

$$\lim_{x \to x_F} \frac{(x_F - x)\,(1-F(x))}{\int_x^{x_F} (1-F(u)\,du} = \infty$$

DEMOSTRACION:

Si F es RR^* <=> la función $U(x) = 1 - G(x)$ para $x > 0$ donde $G(x) = F(x_F - 1/x)$ es de variación rápida con $\rho = -\infty$ lo cual es equivalente por el Teorema 2.2.5 a que:

$$\lim_{x\to\infty} \frac{x(1-G(x))}{\int_x^{\infty}(1-G(t))dt} = \infty$$

Pero $$\lim_{x\to\infty} \frac{x(1-G(x))}{\int_x^{\infty}(1-G(t))\,dt} = \lim_{x\to\infty} \frac{x\,(1-F(x_F-1/x))}{\int_x^{\infty}(1-F(x_F-1/t))\,dt} =$$

$$= \lim_{x\to\infty} \frac{x\,(1-F(x_F-1/x))}{\int_{x_F-1/x}^{\infty}\frac{(1-F(u))}{(x_F-u)^2}\,du} = \lim_{x\to x_F} \frac{(1-F(x))}{\int_x^{x_F}\frac{(x_F-x)}{(x_F-u)^2}(1-F(u))du} <$$

$$< \lim_{x\to x_F} \frac{(x_F-x)\,(1-F(x))}{\int_x^{x_F}(1-F(u))\,du}$$

De forma que:

$$\lim_{x\to x_F} \frac{(1-F(x))}{\int_x^{x_F}\frac{(x_F-x)}{(x_F-u)^2}(1-F(u))du} = \infty \Rightarrow \lim_{x\to x_F} \frac{(x_F-x)(1-F(x))}{\int_x^{x_F}(1-F(u))\,du} = \infty$$

c.q.d.

Por último trasladando los conceptos de variación regular en infinito a la correspondiente en un punto finito, obtenemos los resultados siguientes:

<u>TEOREMA 5.2.3</u>: Sea F una función de distribución con $x_F < \infty$ y absolutamente continua con función de densidad $f(x)$:

i) Si $f(x)$ es positiva para $x \geq B$, donde B es algún valor positivo, entonces la función $U(x) = 1-F(x)$ para x 0 es de variación regular a la izquierda de x_F si F es RN^*.

ii) Si $f(x)$ es monótona no creciente el que $U(x) = 1-F(x)$ para $x > 0$ sea de variación regular con exponente $0 < \rho < \infty$ a la izquierda de x_F implica que F es RN^*.

<u>DEMOSTRACION</u>: Aplicando el Teorema 5.1.3 a la función de distribución $G(x) = F(x_F - 1/x)$ que posee punto final infinito.

<u>TEOREMA 5.2.4</u>: Sea F una función de distribución con $x_F < \infty$ y absolutamente continua con derivada positiva para valores suficientemente grandes, entonces si F es RP^* equivale a que la función $U(x) = 1-F(x)$ para $x > 0$ sea de variación lenta a la izquierda de x_F.

<u>DEMOSTRACION</u>: Aplicando el Teorema 5.1.4 a la función de distribución $G(x) = F(x_F - 1/x)$.

EJEMPLOS:

La familia de funciones de distribución $(F_\alpha)_{\alpha > 0}$ donde

$$F_\alpha(x) = \begin{cases} 1 - e^{-1/(x_F - x)^\alpha} & x \leq x_F \\ 1 & x \geq x_F \end{cases}$$

será RR^{*} para todos los valores de $\alpha > 0$.

Por otra parte las funciones de distribución con crecimiento polinomial en un punto finito:

$$F_{\alpha}(x) = \begin{cases} 0 & x \leqslant x_F - k^{-1/\alpha} \\ 1 - k(x_F - x)^{\alpha} & x_F - k^{-1/\alpha} \leqslant x \leqslant x_F \\ 1 & x \geqslant x_F \end{cases}$$

serán RN^{*} así como las distribuciones uniformes.

Y como ejemplo de distribución RP^{*}

$$F(x) = \begin{cases} 0 & x \leqslant x_F - 1/e \\ 1 + \dfrac{1}{\log_e (x_F - x)} & x_F - 1/e \leqslant x \leqslant x_F \end{cases}$$

donde $x_F \geqslant 1/e$.

CAPITULO III

APLICACION A LA INFERENCIA BAYESIANA

CAPITULO III

APLICACION A LA INFERENCIA BAYESIANA

1.SUMARIO

Dedicaremos este capítulo a aplicar las definiciones de propensión y resistencia dadas anteriormente a los elementos que conforman un problema de inferencia con enfoque bayesiano.

Previamente haremos un resumen del tratamiento de datos atípicos desde este punto de vista incidiendo principalmente en las publicaciones relativas a nuestro planteamiento.

A continuación y en los casos de parámetros de localización y escala respectivamente, estudiaremos la influencia del carácter de la distribución a priori en la distribución a posteriori para una observación fija, siguiendo con el análisis de la divergencia entre a priori y a posteriori cuando la observación crece a infinito, mediante la medida de Kullback-Leibler.

2. DISTRIBUCIONES CON PARAMETRO DE LOCALIZACION

2.1.- Resultados previos

Partiendo de la idea bayesiana, ninguna información es rechazable, por lo que es dificil establecer una analogía con los tests de discordancia y la eliminación consecuente de algunas observaciones.

De hecho el desarrollo fundamental en esta linea corresponde a métodos de acomodación, aunque se pueden encontrar también ciertos tratamientos que establecen probabilidades a posteriori del grado de contaminación de los datos.

En uno de los primeros trabajos, de Finetti (1961), considera que una forma de acomodar datos atípicos es utilizar estimadores que sean medias ponderadas de las observaciones, siendo los datos atípicos aquellas observaciones con pesos muy pequeños.

Esta idea, en principio confusa, de asimilar dato atípico con observación poco influyente en el estimador, subyace en los posteriores resultados de Dawid (1973) y O'Hagan (1979), cuando se refieren a verosimilitudes propensas como aquellas en las que la observación extrema anula su efecto en la a posteriori.

La primera publicación, correspondiendo a Dawid, da origen a un teorema que impone condiciones sobre la verosimilitud y la distribución a priori, para que la media a posteriori de una clase bastante amplia de funciones, converja a la media a priori cuando la observación crece hacia infinito.

<u>TEOREMA 2.1.1;</u> Sea una función de densidad con parámetro de localización $f(x-\theta)$, y distribución a priori sobre Ω con medida de probabilidad G. Sea una función $m(\theta)$ tal que la media a priori es finita. Entonces si se verifica:

i) Dado $\epsilon > 0$, $h > 0$ existe un A tal que si $y > A$ se tiene que

$|f(y') - f(y)| < \epsilon f(y)$ siempre que $|y'-y| < h$

ii) Existen algunas constantes B y M tales que:

$$0 < f(y') < M\, f(y)$$

siempre que $y' > y > B$

iii) (a) $\int^{\infty} k(\theta)\, dG < \infty$

(b) $\int^{\infty} |m(\theta)|\, k(\theta)\, dG < \infty$

donde $k(\theta) = \sup_{x} \left(\frac{f(x-\theta)}{f(x)}\right)$

La esperanza a posteriori de m(Ω) dado X = x tiende a la esperanza a priori cuando $x \to \infty$.

Como él mismo señala las condiciones i) y ii) exigen que la verosimilitud tenga una cola derecha sin oscilaciones fuertes, casi uniforme, mientras que con iii) se evitan comportamientos simétricos de la a priori y la verosimilitud que pueden llevar a contradicciones al intercambiarlas en la aplicación del teorema.

Mas tarde O'Hagan, recoge este teorema y con condiciones más fuertes pero más manejables obtiene el mismo resultado.

TEOREMA 2.1.2: Sea una función de densidad con parámetro de localización $f(x-\theta)$ y una distribución a priori sobre Ω con medida de probabilidad G. Sea $m(\theta)$ una función tal que la esperanza a priori es finita. Entonces las condiciones siguientes:

i^{*}) igual a i) del Teorema 2.1.1.

ii^{*}) a) f(y) es continua y positiva para todo $y \in R$.

b) Existe una constante B tal que para todo $y > B$

i) f(y) es decreciente en y

ii) $b(y) = \frac{d \log f(y)}{dy}$ existe y es creciente

en y.

c) Existe una constante $C \leq B$ tal que para todo $y \leq C$ $f(y)$ es creciente en y.

iii*) a) $\int^{\infty} (f(\theta))^{-1}\, dG < \infty$

b) $\int^{\infty} |m(\theta)| \cdot (f(\theta))^{-1}\, dG < \infty$

implican las condiciones i), ii) y iii) del Teorema 2.1.1.

A partir de este resultado define el concepto de propensión de orden n, señalando que como éste implica el de orden n+1, la mayor exigencia corresponde al orden uno.

DEFINICION 2.1.1: Sean $X_1, \ldots, X_{n+1}$ variables aleatorias independiente e identicamente distribuidas dado $\Omega = \theta$, con densidades $f(x_i - \theta)$. La distribución con densidad f es PROPENSA a datos atípicos por la derecha de orden n si cuando $x_{n+1} \to \infty$.

$$Pr(\Omega \leq c / X_1 = x_1, \ldots, X_{n+1} = x_{n+1}) \longmapsto$$

$$\longmapsto Pr(\Omega < c / X_1 = x_1, \ldots, X_n = x_n)$$

para todo c, $x_1, \ldots, x_n$ y para cualquier distribución a priori sobre Ω.

Análogamente se define la propensión por la izquierda. Para poder aplicar los teoremas anteriores a la situación de muestra de tamaño n se tiene en cuenta lo siguiente:

Si P es una medida de probabilidad a priori, la distribución a posteriori dada $X_1 = x_1, \ldots, X_{n+1} = x_{n+1}$, tendrá una densidad respecto a P, como medida dominante:

$$\prod_{i=1}^{n+1} f(x_i - \theta) = f(x_{n+1} - \theta) \prod_{i=1}^{n} f(x_i - \theta)$$

por lo que podemos considerar la distribución a posteriori como obtenida de una única observación x_{n+1} y una medida a priori P^* satisfaciendo:

$$\frac{dP^*}{dP} = \prod_{i=1}^{n} f(x_i - \theta)$$

De esta forma tenemos:

TEOREMA 2.1.3: Una distribución con densidad f, simétrica y verificando las condiciones i) y ii*) es propensa a datos atípicos de orden uno.

Como corolario obtiene que la t-Student es propensa para cualquier valor de los grados de libertad.

Y por último para definir la resistencia a datos atípicos utilizará el concepto de dominancia estocástica (Fishburn y Vickson (1978)), para reflejar la idea de que, en ese caso, la distribución a posteriori debe seguir a la observación, cuando ésta crece.

DEFINICION 2.1.2: Sean $X_1, \ldots, X_{n+1}$ independientes e idénticamente distribuidas dado $\Omega = \theta$, con densidades $f(x_i - \theta)$. La distribución con densidad f, será RESISTENTE a datos atípicos si:

$$Pr(\Omega \leq c / X_1 = x_1, \ldots, X_{n+1} = x_{n+1})$$

es una función decreciente de x_{n+1}, para todo c, $x_1, \ldots x_n$, para todo $n = 0,1,2,\ldots$ y para cualquier distribución a priori sobre Ω ; es decir si $x'_{n+1} > x_{n+1}$ la distribución a posteriori dada $x_1,\ldots,x_n$, x'_{n+1} domina estocasticamente a la dada $x_1,\ldots,x_n$, x_{n+1} .

Como condición suficiente, utiliza la unimodalidad fuerte introducida por Ibragimov (1956) y equivalente a la log--concavidad de la función de densidad.

TEOREMA 2.1 4: Una distribución fuertemente unimodal con densidad f, que tenga derivada acotada es resistente a datos atípicos.

Aplicando el resultado a la distribución Normal se obtiene que es resistente.

2.2. Comportamiento de las distribuciones a priori y a posteriori para una muestra fija

Nos restringiremos al estudio en la cola derecha de la distribución, extendiendose todos los resultados a la cola izquierda haciendo $\theta \to -\infty$.

Sea X una observación de la densidad $f_\theta(x) = f(x-\theta)$ donde $\theta \in \Omega \subset R$, sobre el cual tenemos definida una distriubción a priori con densidad $\pi(\theta)$. Los siguientes resultados podrán aplicarse al caso de muestra con tamaño $n > 1$, haciendo la misma observación que O'Hagan (1979).

Con la notación anterior dado $X = x$, la distribución a posteriori tendrá como función de densidad:

$$\pi(\theta/x) = \frac{f(x-\theta)\ \pi(\theta)}{p(x)}$$

donde $p(x) = \int_{-\infty}^{+\infty} f(x-\theta)\ \pi(\theta)\ d\theta$ será la función de densidad predictiva.

De esta forma la cola derecha de la distribución a posteriori será:

$$1 - \Pi(\theta/x) = \frac{\int_{\theta}^{\infty} f(x-u)\, \pi(u)\, du}{p(x)} =$$

$$= (1 - \Pi(\theta))\, \frac{\int_{-\infty}^{+\infty} f(x-u)\, \pi_{\theta}^{P}(u)\, du}{p(x)}$$

donde $\Pi(\theta)$: función de distribución a priori

$$\text{y} \quad \pi_{\theta}^{D}(u) = \begin{cases} 0 & \text{para} \quad u \leqslant \theta \\ \dfrac{\pi(u)}{1 - \Pi(\theta)} & \text{para} \quad u > \theta \end{cases}$$

es la función de densidad a priori truncada en el valor θ que denominaremos densidad a priori condicionada a la derecha.

Por lo tanto:

$$\frac{1 - \Pi(\theta/x)}{1 - \Pi(\theta)} = \frac{\int_{-\infty}^{+\infty} f(x-u)\, \pi_{\theta}^{D}(u)\, du}{p(x)} = \frac{p_{\theta}^{D}(x)}{p(x)}$$

donde $p_{\theta}^{D}(x)$: **función de densidad predictiva condicionada a la derecha, que corresponde a la función de densidad a priori $\pi_{\theta}^{D}(x)$ descrita anteriormente.**

Para aplicar las definiciones del capítulo II, relati-

vas a la propensión y resistencia con la suma consideraremos, para un valor fijo x:

$$(1) \quad \frac{\pi(\theta/x)}{1-\Pi(\theta/x)} = \frac{\pi(\theta)}{1-\Pi(\theta)} \cdot \frac{f(x-\theta)}{p_\theta^D(x)}$$

De forma que comparar los comportamientos en la cola derecha de las distribuciones a priori y a posteriori depende del valos de:

$$\lim_{\theta\to\infty} \frac{f(x-\theta)}{p_\theta^D(x)}$$

Pero tenemos que:

$$\frac{f(x-\theta)}{p_\theta^D(x)} = \frac{f(x-\theta)}{\int_{-\infty}^{+\infty} f(x-u)\,\pi_\theta^D(u)\,du} = \frac{1-\Pi(\theta)}{\int_\theta^\infty \frac{f(x-u)}{f(x-\theta)}\,\pi(u)\,du}$$

Así suponiendo que la distribución básica es simétrica y con punto final infinito, y que la función de densidad f(x) es monótona decreciente en las colas, para valores de θ suficientemente grandes:

$$f(x-u) < f(x-\theta) \qquad \forall\, u > \theta$$

luego

$$\int_\theta^\infty \frac{f(x-u)}{f(x-\theta)}\,\pi(u)\,du < 1-\Pi(\theta)$$

y

$$\frac{f(x-\theta)}{p_\theta^D(x)} > 1 \quad \text{para valores grandes de } \theta$$

Sustituyendo en (1), obtenemos:

$$(2) \qquad \frac{\pi(\theta/x)}{1-\Pi(\theta/x)} > \frac{\pi(\theta)}{1-\Pi(\theta)} \qquad \text{para valores grandes de } \theta$$

Entonces hay dos situaciones en las que el comportamiento de las distribuciones a priori y a posteriori no dependen del factor $\frac{f(x-\theta)}{p^D_\theta(x)}$ y por tanto de la verosimilitud.

Con las restricciones anteriores sobre la función de densidad f y con distribuciones a priori de punto final infinito enunciaremos los resultados correspondientes.

PROPOSICION 2.2.1: Si la distribución a posteriori para un valor fijo x, es SP se verifica que la distribución a priori es también SP

DEMOSTRACION: Utilizando la desigualdad (2)

$$\lim_{\theta\to\infty} \frac{\pi(\theta/x)}{1-\Pi(\theta/x)} = 0 \Rightarrow \lim_{\theta\to\infty} \frac{\pi(\theta)}{1-\Pi(\theta)} = 0$$

PROPOSICION 2.2.2: Si la distribución a priori es SR se verifica que la distribución a posteriori para un valor x fijo es también SR.

DEMOSTRACION: Utilizando de nuevo la desigualdad (2)

$$\lim_{\theta\to\infty} \frac{\pi(\theta)}{1-\Pi(\theta)} = \infty \Rightarrow \lim_{\theta\to\infty} \frac{\pi(\theta/x)}{1-\Pi(\theta/x)} = \infty$$

PROPOSICION 2.2.3: Si la distribución a priori es SN se verifica que la distribución a posteriori para una observación fija x, no puede ser SP.

<u>DEMOSTRACION</u>: Por la desigualdad (2) y distribución a priori SN

$$\lim_{\theta \to \infty} \frac{\pi(\theta/x)}{1 - \Pi(\theta/x)} > k > 0$$

<u>PROPOSICION 2.2.4</u>: SI para una observacion fija x, la distribución a posteriori es SN, se tiene que la distribución a priori no puede ser SR.

<u>DEMOSTRACION</u>: Por la desigualdad (2) y distribución a posteriori SN

$$\lim_{\theta \to \infty} \frac{\pi(\theta)}{1 - \Pi(\theta)} < k' < \infty \; .$$

<u>Observaciones</u>:

a) Si la distribución a priori tiene punto final finito, se tienen los mismos resultados para SR^*, SN^* y SP^*

b) La hipótesis de monotonía en las colas para la función de densidad no es muy restrictiva dado que la mayoría de las distribuciones usuales la cumplen. Un caso en el que esto no ocurre es la función de distribución: $F(x) = 1-\exp(-x-C \text{ sen } x)$ donde $0{,}5 < C < 1$, a pesar de tener punto final infinito.

c) Los resultados anteriores ponen en evidencia un hecho claro en el planteamiento bayesiano, pues si a priori hay un conocimiento bastante ajustado del parámetro, es decir la distribución a priori es de colas bajas, la distribución a posteriori debe mantener este tipo de comportamiento. Del mismo modo si el conocimiento a posteriori sobre el parámetro es más bien difuso, sólo puede provenir de una situación a priori análoga.

También existe un comportamiento obligado de la verosimilitud y la distribución a posteriori si suponemos que la función de densidad a priori es de cola derecha monótona y punto final infinito, y la verosimilitud es simétrica respecto al origen, se tiene:

TEOREMA 2.2.1: i) Si la verosimilitud es SR se verifica que la distribución a posteriori es SR.

ii) Si la distribución a posteriori es SP se verifica que la verosimilitud es SP.

iii) Si la distribución a posteriori es SN se verifica que la verosimilitud no puede ser SR.

iv) Si la verosimilitud es SN se verifica que la distribución a posteriori no puede ser SP.

DEMOSTRACION:

Si la función de densidad a priori tiene cola derecha monótona se tiene que:

$$\pi(u) < \pi(\theta) \quad \forall \; u > \theta \;,$$

si θ es suficientemente grande, por lo cual,

$$\frac{f(x-\theta)}{p_\theta^D(x)} = \frac{f(x-\theta)}{\displaystyle\int_\theta^\infty f(x-u)\,\frac{\pi(u)}{1-\Pi(\theta)}\,du} > \frac{1-\Pi(\theta)}{\pi(\theta)} \cdot \frac{f(x-\theta)}{\displaystyle\int_\theta^\infty f(x-u)du}$$

luego

$$\frac{\pi(\theta/x)}{1-\Pi(\theta/x)} = \frac{f(x-\theta)}{p_\theta^D(x)} \cdot \frac{\pi(\theta)}{1-\Pi(\theta)} > \frac{f(x-\theta)}{\displaystyle\int_\theta^\infty f(x-u)\,du} = \frac{f(x-\theta)}{F(x-\theta)} =$$

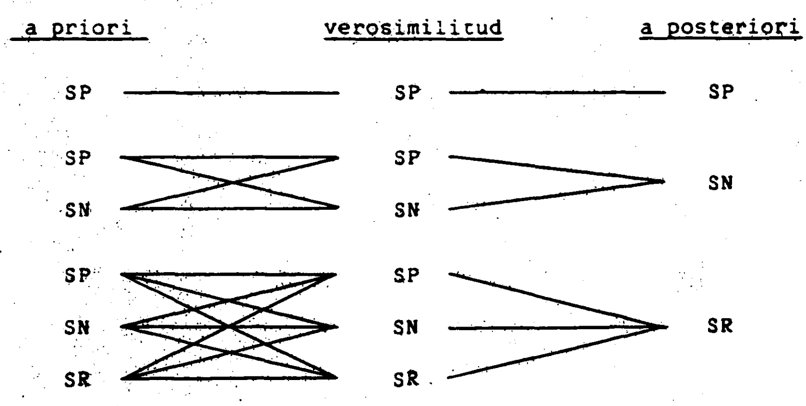
$$= \frac{f(\theta-x)}{1-F(\theta-x)}$$

y así considerando x fijo y tomando limites cuando $\theta \to \infty$ se obtienen los resultados del teorema.

c.q.d.

<u>Observaciones</u>: Se puede resumir este resultado en la idea de que verosimilitudes de colas bajas hacen que la distribución a posteriori sea del mismo tipo, así como que distribuciones a posteriori poco concentradas sólo pueden conseguirse con verosimilitudes de colas altas.

En forma de esquema tenemos que en las hipótesis del comienzo:



Podemos observar que si se mantiene el caracter de la distribución a priori en la distribución a posteriori para cualquier situación es necesario que la verosimilitud sea propensa.

Para que sea suficiente habrá que imponer alguna restricción más sobre dicha verosimilitud propensa, en los casos de distribución a priori SN y SR.

<u>TEOREMA 2.2.2</u>: Sea $f(x)$ una función de densidad continua y positiva para todo x, simétrica y de colas monótonas verificando la condición de que:

Dado $\epsilon_0 > 0$ y $M > 0$ existe un x_0 tal que para todo $x > x_0$

$$\left| \frac{f(x+\epsilon)}{f(x)} - 1 \right| < M \quad \forall\ \epsilon \in (0,\ \epsilon_0)$$

Entonces $\lim\limits_{\theta \to \infty} \dfrac{f(x-\theta)}{p_\theta^D(x)} = 1$ para x fijo

<u>DEMOSTRACION</u>: Dado x tenemos que por ser $f(x)$ simétrica:

$$\frac{p_\theta^D(x)}{f(x-\theta)} = \int_0^\infty \frac{f(v+\theta-x)\ \pi(\theta+v)}{f(\theta-x)\ (1-\Pi(\theta))}\, dv$$

Además

$$\lim_{\theta \to \infty} \int_0^\infty \frac{f(v+\theta-x)\ \pi(\theta+v)}{f(\theta-x)\ (1-\Pi(\theta))}\, dv = \lim_{\theta \to \infty} \int_0^\infty \frac{f(\theta+v)\ \pi(\theta+x+v)}{f(\theta)\ (1-\Pi(\theta+x))}\, dv$$

luego converge a cero cuando ϵ_0 tiende a infinito.

Por otro lado, dada la condición del teorema:

$$\lim_{x\to\infty} \frac{F(x+\epsilon_0)}{F(x)} = 1 \qquad \forall\, \epsilon \in (0,\epsilon_0)$$

y la convergencia es uniforme en cualquier intervalo. Así obtenemos:

$$\lim_{\theta\to\infty} \frac{\int_0^{\epsilon_0} \frac{f(\theta+v)}{f(\theta)}\, \pi(\theta+x+v)\, dv}{\Pi(\theta+x+\epsilon_0) - \Pi(\theta+x)} = 1$$

y de esta forma:

$$\int_0^{\epsilon_0} \frac{f(\theta+v)\, \pi(\theta+x+v)}{f(\theta)\,(1-\Pi(\theta+x)} dv \sim 1 - \frac{1-\Pi(\theta+x+\epsilon_0)}{1-\Pi(\theta+x)}$$

para $\theta \to \infty$, y cualquier $\epsilon_0 > 0$.

Si la distribución a priori es:

$$\text{SN:} \quad \lim_{\theta\to\infty} \frac{1-\Pi(\theta+x+\epsilon_0)}{1-\Pi(\theta+x)} = e^{-c\epsilon_0} \quad \text{para algún } c \;,\, 0 < c < \infty$$

$$\text{SR:} \quad \lim_{\theta\to\infty} \frac{1-\Pi(\theta+x+\epsilon_0)}{1-\Pi(\theta+x)} = 0$$

En ambos casos:

$$\lim_{\theta\to\infty} \int_0^{\infty} \frac{f(\theta+v)\, \pi(\theta+x+v)}{f(\theta)\,(1-\Pi(\theta+x))} dv = 1$$

c.q.d.

Aplicando la regla de L'Hopital: si dado $\epsilon > 0$ se tiene que existe el $\lim\limits_{x\to\infty} \frac{f(x+\epsilon)}{f(x)} = 1$, entonces $\lim\limits_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)} = 1$, equivalente con que F es una función de distribución SP. Para que se dé el resultado inverso es necesario asegurar la existencia de $\lim\limits_{x\to\infty} \frac{f(x+\epsilon)}{f(x)}$, en cuyo caso

$$\lim_{x\to\infty} \frac{f(x+\epsilon)}{f(x)} = \lim_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)}$$

Veremos algunas condiciones que aseguren la existencia de dicho límite.

<u>TEOREMA 2.2.3</u>: Dada una función de distribución F con $x_F = \infty$, simétrica respecto al origen, cuya función de densidad $f(x)$ monótona en las colas verifica que $h(x) = \frac{d \log f(x)}{dx}$ es una función monótona creciente para valores de x suficientemente grandes, entonces como verosimilitud en un planteamiento bayesiano no altera el comportamiento de la distribución a priori al pasar a la distribución a posteriori.

<u>DEMOSTRACION</u>:

Sea $\epsilon_0 > 0$, llamaremos $k_{\epsilon_0}(x) = \frac{f(x+\epsilon_0)}{f(x)}$

Para valores de x suficientemente grandes

$$0 < k_{\epsilon_0}(x) < 1$$

y además

$$k'_{\epsilon_0}(x) = \frac{f(x) \,.\, f'(x+\epsilon_0) - f(x+\epsilon_0) \,.\, f'(x)}{(f(x))^2}$$

pero:

$$f(x)\; f'(x+\epsilon_0) - f(x+\epsilon_0)\; f'(x) = \left(\frac{f'(x+\epsilon_0)}{f(x+\epsilon_0)} - \frac{f'(x)}{f(x)}\right) f(x)f(x+\epsilon_0)$$

luego

$$\text{signo}\;(k'_{\epsilon_0}(x)) = \text{signo}\;\left(\frac{f'(x+\epsilon_0)}{f(x+\epsilon_0)} - \frac{f'(x)}{f(x)}\right) > 0$$

al ser, para estos valores de x en las colas:

$$f(x) \,.\, f(x+\epsilon_0) > 0$$

y la función $h(x) = \frac{d \log f(x)}{dx} = \frac{1}{f(x)} \,.\, \frac{df(x)}{dx}$ monótona creciente

De esta forma la función de densidadd verifica la condición del Teorema 2.2.2.

EJEMPLOS

Dado que la función de densidad de la t-Student con m grados de libertad es:

$$f(x) = k(m+x^2)^{-\frac{m+1}{2}}$$

corresponde al tipo de distribuciones que cumplen el Teorema 2.2.3, puesto que la función:

$$\frac{f'(x)}{f(x)} = \frac{d \log f(x)}{dx} = -k' \frac{x}{(m+x^2)}$$

es monótona creciente para $x > \sqrt{m}$.

2.3. Influencia del tipo de verosimilitud en la divergencia asintótica entre las distribuciones a priori y a posteriori

Teniendo en cuenta el concepto de f-divergencia entre las funciones de densidad a priori y a posteriori (Goel (1983))

$$\mathcal{J}_f(\pi(\theta), \pi(\theta/x)) = \int \pi(\theta/x)\, f\left(\frac{\pi(\theta)}{\pi(\theta/x)}\right) d\theta$$

donde f es cualquier función convexa definida en $(0, \infty)$, para el caso $f(u) = (u-1) \log u$:

$$\mathcal{J}_f(\pi(\theta), \pi(\theta/x)) = \int (\pi(\theta) - \pi(\theta/x)) \log \frac{\pi(\theta)}{\pi(\theta/x)}\, d\theta$$

obtenemos la divergencia de Kullback-Leibler (Kullback (1968)) como medida de la discrepancia entre las dos funciones de densidad.

De este modo analizaremos el comportamiento de la distribución a posteriori respecto a la distribución a priori cuando x crece hasta infinito, para los distintos tipos de verosimilitudes.

En todos los casos supondremos que la función de densidad f es simétrica positiva continua y monótona en las colas. También en los casos de distribuciones SP y SN, supondremos que existe:

$$\lim_{x \to \infty} \frac{f(x+\epsilon)}{f(x)} \qquad \forall \epsilon > 0$$

Por lo tanto, siendo f la función de densidad de F:

i) F es una función de distribución SP equivale a

$$\lim_{x\to\infty} \frac{f(x+\epsilon)}{f(x)} = 1 \qquad \forall \epsilon > 0$$

ii) F es una función de distribución SR equivale a

$$\lim_{x\to\infty} \frac{f(x+\epsilon)}{f(x)} = 0 \qquad \forall \epsilon > 0$$

iii) F es una función de distribución SN equivale a

$$\lim_{x\to\infty} \frac{f(x+\epsilon)}{f(x)} = e^{-c\epsilon} \qquad \forall \epsilon > 0 \text{ , donde } 0 < c < \infty .$$

TEOREMA 2.3.1: Si F es una función de distribución SP y la convergencia de $\lim\limits_{x\to\infty} \dfrac{f(x+\epsilon)}{f(x)} = 1$ es uniforme en intervalos finitos se verifica: $\lim\limits_{x\to\infty} \dfrac{p(x)}{f(x-\theta)} = 1$, $\forall \theta$

para cualquier distribución a priori con densidad positiva $\pi(u)$

DEMOSTRACION: $\left.\begin{array}{l} \lim\limits_{x\to\infty} \dfrac{f(x+\epsilon)}{f(x)} = 1 \\ \\ \lim\limits_{x\to\infty} \dfrac{f(x-\epsilon)}{f(x)} = 1 \end{array}\right\}$ uniformemente en intervalos finitos para ϵ.

Entonces, dado $h > 0$

$$\frac{p(x)}{f(x)} = \int_{-\infty}^{-h} \frac{f(x-u)}{f(x)} \pi(u)\, du + \int_{-h}^{h} \frac{f(x-u)}{f(x)} \pi(u)\, du +$$

$$+ \int_{h}^{\infty} \frac{f(x-u)}{f(x)} \phi(u)\, du.$$

En primer lugar por ser f de colas monótonas y simétrica:

$$I_1 = \int_{-\infty}^{-h} \frac{f(x-u)}{f(x)} \phi(u)\, du < \int_{-\infty}^{-h} \phi(u)\, du = \Phi(-h)$$

Por lo que $I_1 \to 0$ cuando $h \to \infty$.

También por las mismas razones:

$$I_2 = \int_{h}^{\infty} \frac{f(x-u)}{f(x)} \phi(u)\, du < 1 - \Phi(h) \quad \text{para } h > 2x$$

Luego $I_2 \to 0$ cuando $h \to \infty$.

Y por último debido a la convergencia uniforme:

Dado $h > 0$ y $M > 0$ existe x_0 tal que $\forall\, x > x_0$:

$$1 - M < \frac{f(x-u)}{f(x)} < 1 + M \qquad \forall\, u \in (-h, h)$$

De esta forma:

$$\lim_{x \to \infty} \frac{\displaystyle\int_{-h}^{h} \frac{f(x-u)}{f(x)} \phi(u)\, du}{\Phi(h) - \Phi(-h)} = 1$$

y así: $$\lim_{x \to \infty} \frac{p(x)}{f(x)} = 1$$

Además: $$\lim_{x \to \infty} \frac{f(x)}{f(x-\theta)} = 1 \;, \quad \forall\, \theta \text{ por ser SP}$$

En definitiva:

$$\lim_{x\to\infty} \frac{p(x)}{f(x-\theta)} = 1 \qquad \forall\, \theta$$

c.q.d.

Para utilizar el resultado anterior desarrollaremos la expresión de la divergencia de Kullback-Leibler

$$\mathcal{J}_f(\pi(\theta),\ \pi(\theta/x)) = \int_{-\infty}^{+\infty} (\pi(\theta) - \pi(\theta/x)) \log \frac{\pi(\theta)}{\pi(\theta/x)}\, d\theta =$$

$$= \int_{-\infty}^{+\infty} \left(\frac{f(x-\theta)}{p(x)} - 1 \right) \log \left(\frac{f(x-\theta)}{p(x)} \right) \pi(\theta)\, d\theta$$

<u>TEOREMA 2.3.2</u>: Sea F una función de distribución SP y la convergencia de $\lim_{x\to\infty} \frac{f(x+\varepsilon)}{f(x)} = 1$ uniforme en intervalos finitos. Si dada una distribución a priori con densidad positiva $\pi(u)$ se verifica

$$\int_{-\infty}^{+\infty} (c(\theta)-1)\ (\log c(\theta))\ \pi(\theta)\, d\theta < \infty$$

donde $c(\theta) = \sup_x \frac{f(x-\theta)}{p(x)}$,

Entonces:

$$\lim_{x\to\infty} \mathcal{J}_f(\pi(\theta),\ \pi(\theta/x)) = 0$$

DEMOSTRACION: Como tenemos que:

$$\left| \left(\frac{f(x-\theta)}{p(x)} - 1\right) \log \left(\frac{f(x-\theta)}{p(x)}\right) \right| < (c(\theta)-1) \log c(\theta) \quad \forall \theta$$

y esta función es integrable respecto a la distribución a priori, aplicando el Teorema de la Convergencia Dominada resulta:

$$\lim_{x \to \infty} \mathcal{J}_f (\pi(\theta), \pi(\theta/x)) = 0$$

c.q.d.

COMENTARIOS:

Este resultado confirma la idea bayesiana introducida por Dawid (1973), (aunque de Finetti (1961) lo indicó anteriormente), de que en distribuciones propensas a producir datos extremos alejados, el efecto que éstos podrían causar en la distribución a posteriori era el de anularse a si mismos como observaciones y volver al punto de partida, es decir a la distribución a priori.

Nosotros lo hemos traducido en el sentido de que la divergencia de Kullback-Leibler entre las densidades a priori y a posteriori tienen a anularse cuando la observación es extrema si la verosimilitud es propensa con la suma. En estos casos el modelo asimila perfectamente el dato atípico sin modificar las conclusiones erróneamente.

En el caso de distribuciones SN los resultados aparentemente no tienen un significado tan intuitivo como el anterior.

TEOREMA 2.3.3: Sea F una función de distribución SN y tal que:

i) la convergencia en $\lim\limits_{x\to\infty} \dfrac{f(x-\varepsilon)}{f(x)} = e^{c\varepsilon}$ es uniforme en intervalos finitos

ii) si definimos $k(u) = \sup\limits_{x} \dfrac{f(x-u)}{f(x)}$ y se tiene que:

$$\int_h^{\infty} k(u)\,\pi(u)\,du < \infty \qquad \forall\, h$$

siendo $\pi(u)$ la función de densidad a priori.

Entonces:

$$\lim_{x\to\infty} \frac{p(x)}{f(x-\theta)} = e^{-c\theta}\,\psi_{\Omega}(c) \qquad \forall\,\theta \ \text{ tal que } 0<c<\infty$$

con

$$\psi_{\Omega}(c) = \int_{-\infty}^{+\infty} e^{cu}\,\pi(u)\,du$$

DEMOSTRACION:

Tenemos en primer lugar que:

$$\lim_{x\to\infty} \frac{f(x)}{f(x-\theta)} = e^{-c\theta} \qquad \forall\,\theta$$

así que nos faltará calcular $\lim\limits_{x\to\infty} \dfrac{p(x)}{f(x)}$

Sea $\varepsilon_0 > 0$

$$\frac{p(x)}{f(x)} = \int_{-\infty}^{0} \frac{f(x-u)}{f(x)} \pi(u)\, du + \int_{0}^{\epsilon_0} \frac{f(x-u)}{f(x)} \pi(u)\, du +$$

$$+ \int_{\epsilon_0}^{\infty} \frac{f(x-u)}{f(x)} \pi(u)\, du = I_1 + I_2 + I_3$$

Respecto a la primera integral dado que $\forall \epsilon > 0$: $\lim\limits_{x\to\infty} \dfrac{f(x+\epsilon)}{f(x)} =$

$= e^{-C\epsilon}$

como
$$\begin{cases} \dfrac{f(x+\epsilon)}{f(x)} \text{ es monótona en } \epsilon \text{, para x suficiente- mente grande} \\[2ex] 0 < \dfrac{f(x+\epsilon)}{f(x)} < 1 \quad \forall \epsilon > 0 \\[2ex] h(\epsilon) = e^{-C\epsilon} \text{ es continua} \end{cases}$$

se tiene (Apéndice I) que la convergencia es uniforme.

Por esta razón:

$$I_1 \longmapsto \int_{-\infty}^{0} e^{Cu} \pi(u)\, du \quad \text{cuando } x\to\infty .$$

En la segunda por la condicion i), la convergencia en:

$$\lim_{x\to\infty} \frac{f(x-\epsilon)}{f(x)} = e^{C\epsilon} \quad \text{es uniforme para } \epsilon \in (0,\epsilon_0)$$

Así que, analogamente a la situación anterior:

$$I_2 \longmapsto \int_{0}^{\epsilon_0} e^{Cu} \pi(u)\, du \quad \text{cuando } x\to\infty$$

-122-

Y por último en la tercera integral debido a la condición ii):

$$I_3 = \int_{\epsilon_0}^{\infty} \frac{f(x-u)}{f(x)} \pi(u)\, du < \int_{\epsilon_0}^{\infty} k(u)\, \pi(u)\, du < \infty$$

por lo que: $I_3 \to 0$ cuando $\epsilon_0 \to \infty$

En resumen: $\lim\limits_{x\to\infty} \dfrac{p(x)}{f(x)} = \displaystyle\int_{-\infty}^{+\infty} e^{cu}\, \pi(u)\, du$

c.q.d.

TEOREMA 2.3.4: Sea F una función de distribución SN, tal que se verifican las condiciones del Teorema 2.3.3 y además que:

$$\int_{-\infty}^{+\infty} (c(\theta)-1) \log c(\theta)\, \pi(\theta)\, d\theta < \infty$$

donde $c(\theta) = \sup\limits_{x} \dfrac{f(x-\theta)}{p(x)}$

Entonces

$$\lim_{x\to\infty} \mathcal{J}_f(\pi(\theta), \pi(\theta/x)) = \int_{-\infty}^{+\infty} \left(\frac{e^{c\theta}}{\psi_\Omega(c)} - 1\right) (c\theta - \log \psi_\Omega(c))\, \pi(\theta) d\theta$$

DEMOSTRACION: Utilizando el Teorema 2.3.3, y el Teorema de la Convergencia Dominada.

A continuación realizamos el estudio para distribuciones SR, teniendo en cuenta las hipótesis que hicimos sobre la función de densidad al comienzo de este apartado.

<u>TEOREMA 2.3.5</u>: Sea F una función de distribución SR, entonces

$$\lim_{x\to\infty} \frac{p(x)}{f(x-\theta)} = \infty \qquad \forall\, \theta$$

<u>DEMOSTRACION</u>:

Si F es SR $\iff \forall \epsilon > 0 \quad \lim_{x\to\infty} \frac{f(x+\epsilon)}{f(x)} = 0$

También tenemos que $\forall \epsilon > 0$:

$$\lim_{x\to\infty} \frac{f(x-\epsilon)}{f(x)} = \infty$$

es decir que dado $\epsilon_0 > 0$, para todo $k > 0$ existe x_0 tal que:

$$\frac{f(x-\epsilon_0)}{f(x)} > k \qquad \forall\, x > x_0$$

Supongamos que x_0 está en la cola monótona de f por tanto:

$$\forall\, \epsilon > \epsilon_0 \qquad f(x-\epsilon) > f(x-\epsilon_0)$$

es decir:

$$\frac{f(x-\epsilon)}{f(x)} > k \qquad \forall\, x > x_0 \quad y \quad \forall\, \epsilon > \epsilon_0$$

Teniendo en cuenta estas acotaciones calcularemos

$$\lim_{x\to\infty} \frac{p(x)}{f(x-\theta)} = \lim_{y\to\infty} \frac{p(y+\theta)}{f(y)} \qquad \forall\, \theta$$

$$\frac{p(y+\theta)}{f(y)} = \int_{-\infty}^{+\infty} \frac{f(y+\theta-u)}{f(y)} \pi(u)\, du = \int_0^{\infty} \frac{f(y+v)}{f(y)} \pi(\theta-v)\, dv +$$

$$+ \int_0^{\infty} \frac{f(y-v)}{f(y)} \pi(\theta+v)\, dv = I_1 + I_2$$

Respecto a la primera integral, supongamos y en la cola monótona de f:

$$\forall\ v > 0 \quad \frac{f(y+v)}{f(y)} < 1$$

es decir:

$$\frac{f(y+v)}{f(y)} \pi(\theta-v) < \pi(\theta-v) \qquad \forall\ y$$

donde

$$\int_0^{\infty} \pi(\theta-v)\, dv < \infty \qquad \forall\ \theta$$

Aplicando el Teorema de la Convergencia Dominada:

$$I_1 \longmapsto 0 \quad \text{cuando } y \to \infty.$$

Repecto a la segunda, sea $\epsilon_0 > 0$ $k > 0$, entonces existe y_0 tal que para todo $y > y_0$ y además $y-\epsilon_0$ en la cola monótona de f se tiene que:

$$\int_0^{\epsilon_0} \frac{f(y-v)}{f(y)} \pi(\theta+v)\, dv + \int_{\epsilon_0}^{\infty} \frac{f(y-v)}{f(y)} \pi(\theta+v)\, dv >$$

$$> \int_0^{\varepsilon_0} \pi(\theta+v)\, dv + k \int_{\varepsilon_0}^{\infty} \pi(\theta+v)\, dv =$$

$$= \Pi(\varepsilon_0+\theta) - \Pi(\theta) + k(1-\Pi(\varepsilon_0+\theta)) = M(\varepsilon_0,\ k,\ \theta)$$

En resumen, dado $\theta \in \Omega$, $\varepsilon_0 > 0$, $k > 0$ existe $y_0(\theta, \varepsilon_0, k)$ y una constante $M(\theta, \varepsilon_0, k)$ tal que:

$$I_2 = \int_0^{\infty} \frac{f(y-v)}{f(y)}\, \pi(\theta+v)\, dv > M(\varepsilon_0,\ k,\ \theta) \qquad \forall\ y > y_0$$

o sea $\forall\ \theta \quad I_2 \longmapsto \infty$ cuando $y \to \infty$

Por lo tanto:

$$\lim_{x \to \infty} \frac{p(x)}{f(x-\theta)} = \infty \qquad \forall\ \theta$$

c.q.d.

<u>TEOREMA 2.3.6</u>: Sea F una función de distribución SR con función de densidad f que verifica:

i) $\displaystyle\int_{-\infty}^{+\infty} c(\theta)\ (\log c(\theta))\pi(\theta)\ d\theta < \infty$

donde $c(\theta) = \sup\limits_{x} \dfrac{f(x-\theta)}{p(x)}$

y $\pi(\theta)$ es la función de densidad a priori.

ii) $\dfrac{p(x)}{f(x-\theta)} = k(x)$ es monótona creciente $\forall\ \theta$

Entonces

$$\lim_{x \to \infty} \mathcal{J}_f(\pi(\theta),\ \pi(\theta/x)) = \infty$$

DEMOSTRACION:

Desdoblando $\mathcal{J}_f(\pi(\theta),\ \pi(\theta/x))$ obtenemos:

$$\mathcal{J}_f\ (\pi(\theta), \pi(\theta/x)) = \int_{-\infty}^{+\infty} \frac{f(x-\theta)}{p(x)} \log \frac{f(x-\theta)}{p(x)}\ \pi(\theta)\ d\theta\ +$$

$$+ \int_{-\infty}^{+\infty} \log \frac{p(x)}{f(x-\theta)}\ \pi(\theta)\ d\theta$$

Utilizando la condicion i) y además que:

$$\lim_{x \to \infty} \frac{f(x-\theta)}{p(x)} \log \frac{f(x-\theta)}{p(x)} = 0 \qquad \forall\ \theta$$

resulta:

$$\lim_{x \to \infty} \int_{-\infty}^{+\infty} \frac{f(x-\theta)}{p(x)} \log \frac{f(x-\theta)}{p(x)}\ \pi(\theta)\ d\theta = 0$$

Además si consideramos $\{x_n\}_{n \in N}$ sucesión monótona creciente

tal que $\lim_{x \to \infty} x_n = \infty$, como:

$$k(x) = \frac{p(x)}{f(x-\theta)} \quad \text{es monótona creciente } \forall\ \theta$$

tenemos que:

$$h_n(\theta) = \frac{p(x_n)}{f(x_n-\theta)}$$

es tal que la sucesión $\{h_n(\theta)\}_{n\in N}$ es momótona creciente para todo θ.

Además $h_n(\theta) > 0$ y $\lim_{n\to\infty} h_n(\theta) = \infty$

o sea $\lim_{n\to\infty} \log h_n(\theta) = \infty$

y por el Teorema de la Convergencia Monótona:

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} \log h_n(\theta)\, \pi(\theta)\, d\theta = \infty$$

lo que implica

$$\lim_{x\to\infty} \int_{-\infty}^{+\infty} \log \frac{p(x)}{f(x-\theta)}\, \pi(\theta)\, d\theta = \infty$$

En resumen:

$$\lim_{x\to\infty} \mathcal{J}_f\,(\pi(\theta), \pi(\theta/x)) = \infty \qquad \forall\, \theta$$

<u>COMENTARIOS:</u>

En este último caso se refleja el comportamiento de la distribución a posteriori cuando x crece indefinidamente, en una separación total de la distribución a priori, para verosimilitudes SR. De esta forma los datos extremos arrastran consigo la distribución a posteriori.

Con esta conclusión seguimos en la linea de los resultados de O'Hagan (1979), aunque en dicha publicación, como ya expusimos se interpretaba en términos de dominancia estocástica.

3. DISTRIBUCIONES CON PARAMETRO DE ESCALA

Consideramos funciones de distribución $F_\theta(x)$, que dependen de un factor de escala $\theta \in \Omega \subset \mathbb{R}^+$, de manera que existe una función de distribución F, tal que:

$$F_\theta(x) = F(\theta x) \qquad \forall\ x \in R$$

Entonces las correspondientes funciones de densidad serán:

$$f_\theta(x) = \theta f(\theta x) \qquad \forall\ x \in R$$

si ademas tenemos definida sobre Ω una función de distribución a priori Π con función de densidad π, la distribución a posteriori tendrá como función de densidad

$$\pi(\theta/x) = \frac{\theta f(\theta x)\ \pi(\theta)}{p(x)}$$

donde $p(x) = \int_0^\infty u\ f(ux)\ \pi(u)\ du$

3.1. Comportamiento de las distribuciones a priori y a posteriori para una muestra fija

Como en el caso de parámetro de localización, nos restringimos a muestras de tamaño uno y a la cola derecha de la distribución, suponiendo que los puntos finales son infinitos.

Para comparar los comportamientos de las distribuciones a priori y a posteriori calcularemos la cola de la distribución a posteriori.

Dado que la función de distribución a posteriori es:

$$\Pi(\theta/x) = \frac{\int_{\theta}^{\infty} u\ f(ux)\ \pi(u)\ du}{p(x)} = (1-\Pi(\theta)) \int_{0}^{\infty} \frac{u\ f(ux)}{p(x)}\ \pi_{\theta}^{D}(u)\ du$$

donde para cada θ se define:

$$\pi_{\theta}^{D}(u) = \begin{cases} \dfrac{\pi(u)}{1-\Pi(u)} & u > \theta \\[2ex] 0 & u \leqslant \theta \end{cases}$$

función de densidad a priori condicionada a la derecha.

Entonces:

$$\frac{1-\Pi(\theta/x)}{1-\Pi(\theta)} = \frac{\int_{0}^{\infty} u\ f(ux)\ \pi_{\theta}^{D}(u)\ du}{\int_{0}^{\infty} u\ f(ux)\ \pi(u)\ du} = \frac{p_{\theta}^{D}(x)}{p(x)}$$

denominando, como anteriormente, a $p_{\theta}^{D}(x)$: función de densidad predictiva condicionada a la derecha.

De esta forma:

$$\frac{\theta\ \pi(\theta/x)}{1-\Pi(\theta/x)} = \frac{\theta\ f(\theta x)}{p_{\theta}^{D}(x)} \cdot \frac{\theta\ \pi(\theta)}{1-\Pi(\theta)}$$

Y así el factor determinante de las relaciones será:

$$\lim_{\theta \to \infty} \frac{\theta f(\theta x)}{p_{\theta}^{D}(x)} \quad \text{para x fijo}$$

Pero:

$$\frac{p_{\theta}^{D}(x)}{\theta f(\theta x)} = \int_{\theta}^{\infty} \frac{uf(ux)\,\pi(u)}{\theta f(\theta x)\,(1-\pi(\theta))}\,du$$

y sí suponemos que F es UNIMODAL con moda en el origen, es equivalente a que:

$$G(x) = F(x) - xF'(x)$$

sea una función de distribución (Gnedenko, Kolmogorov (1954)).

Por tanto:

$$F \text{ unimodal en el origen} \Rightarrow \lim_{x \to \infty} xF'(x) = \lim_{x \to \infty} xf(x) = 0$$

Si además suponemos que la función de densidad f es monótona en las colas resultará que:

$h(x) = xf(x)$ es monótona no creciente para valores x suficientemente grandes

Aplicándolo al estudio del factor $\dfrac{p_{\theta}^{D}(x)}{\theta f(\theta x)}$:

$uxf(ux) < \theta x f(\theta x) \quad \forall\, u > \theta$, si suponemos un valor θ suficientemente grande

y en este caso se concluye que:

$$\frac{p_\theta^D(x)}{\theta f(\theta x)} = \frac{\int_\theta^\infty \frac{uf(ux)}{\theta f(\theta x)} \pi(u)\, du}{1 - \Pi(\theta)} < 1$$

En resumen para valores θ suficientemente grandes:

$$(1')\ \frac{\theta\pi(\theta/x)}{1-\Pi(\theta/x)} = \frac{\theta f(\theta x)}{p_\theta^D(x)} \frac{\theta\pi(\theta)}{1-\Pi(\theta)} > \frac{\theta\pi(\theta)}{1-\Pi(\theta)}$$

y de esta forma surgen las siguientes relaciones:

<u>PROPOSICION 3.1.1</u>: Si la distribución a priori es RR, se verifica que la distribución a posteriori es RR.

<u>DEMOSTRACION</u>: Partiendo de las desigualdades (1')

$$\lim_{\theta\to\infty} \frac{\theta\pi(\theta)}{1-\Pi(\theta)} = \infty \quad \Rightarrow \quad \lim_{\theta\to\infty} \frac{\theta\pi(\theta/x)}{1-\Pi(\theta/x)} = \infty$$

<u>PROPOSICION 3.1.2</u>: Si la distribución a posteriori es RP se tiene que la distribución a priori debe ser RP

<u>DEMOSTRACION</u>: Con (1') :

$$\lim_{\theta\to\infty} \frac{\theta\pi(\theta/x)}{1-\Pi(\theta/x)} = 0 \quad \Rightarrow \quad \lim_{\theta\to\infty} \frac{\theta\pi(\theta)}{1-\Pi(\theta)} = 0$$

<u>PROPOSICION 3.1.3</u>: Si la distribución a priori es RN, se vevirifca que la distribución a posteriori no puede ser RP.

<u>DEMOSTRACION:</u> Con (1') como $0 < c < \infty$

$$0 < \lim_{\theta\to\infty} \frac{\theta\pi(\theta)}{1-\Pi(\theta)} = c < \lim_{\theta\to\infty} \frac{\theta\pi(\theta/x)}{1-\Pi(\theta/x)}$$

PROPOSICION 3.1.4: Si la distribución a posteriori es RN se verifica que la distribución a priori no puede ser RR.

DEMOSTRACION: Utilizando (1') y analogamente al caso anterior

En los supuestos iniciales considerábamos distribuciones unimodales y así $h(x) = xf(x)$ resultaba monótona no creciente para valores grandes. También obtenemos esta conclusión exigiendo que F sea una función de distribución RN ya que es equivalente a:

$$\lim_{x \to \infty} \frac{xf(x)}{1-F(x)} = c \qquad \text{donde} \quad 0 < c < \infty$$

y por tanto:

$$\lim_{x \to \infty} xf(x) = 0$$

que con la hipótesis de monotonía en la cola de f nos lleva a la misma condición:

De este modo, con distribuciones a priori o bien unimodales o bién RN y densidad monótona en la cola vamos a obtener un cierto tipo de comportamiento en la distribución a posteriori obligado por el de la verosimilitud.

TEOREMA 3.1.1: i) Si la verosimilitud es RR se verifica que la distribución a posteriori es RR.

ii) Si la distribución a posteriori es RP, la verosimilitud no puese ser RR.

iii) Si la distribución a posteriori es RN, la verosimilitud no puede ser RR.

iv) Si la verosimilitud es RN, la distribución a posteriori no puede ser RP.

<u>DEMOSTRACION</u>: Por las hipótesis sobre la distribución a priori tenemos que para valores de θ suficientemente altos:

$$u\,\pi(u) < \theta\pi(\theta) \qquad \forall\, u > \theta$$

y así:

$$\frac{p_{\theta}^{D}(x)}{\theta f(\theta x)} = \int_{\theta}^{\infty} \frac{uf(ux)\;\pi(u)}{\theta f(\theta x)(1-\Pi(\theta))}\,du < \int_{\theta}^{\infty} \frac{\theta f(ux)\pi(\theta)}{\theta f(\theta x)(1-\Pi(\theta))}\,du =$$

$$= \frac{\pi(\theta)}{f(\theta x)(1-\Pi(\theta))} \cdot \frac{1-F(\theta x)}{x}$$

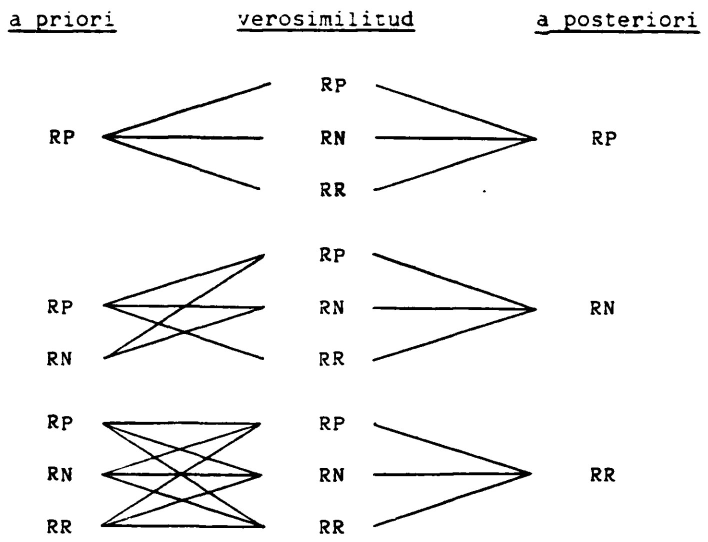
luego:

$$\frac{\theta\pi(\theta x)}{1-\Pi(\theta/x)} > \frac{\theta x f(\theta x)}{1-F(\theta x)}$$

Aplicando esta desigualdad se demuestran los diferentes apartados.

c.q.d.

Como resumen obtenemos el esquema:



OBSERVACIONES:

- Al igual que en el caso de parámetro de localización quedan eliminadas las situaciones por las que con un conocimiento a priori preciso se pasase a un cierto desconocimiento a posteriori, reflejado en colas altas de la distribución. También si a posteriori se tiene una distribución poco concentrada solo puede partir de una distribución a priori del mismo tipo.

- Otra conclusión que surge del esquema, es que en el caso de distribuciones a priori RN, sólo para verosimilitudes RR no podemos obtener distribuciones a posteriori del mismo tipo, es decir RN. Efectivamente, haciendo algunas hipótesis sobre la distribución se demuestra dicho resultado.

TEOREMA 3.1.2: Sea la verosimilitud unimodal con moda en el origen, de colas monótonas y tal que existe $\lim\limits_{x \to \infty} \frac{f(kx)}{f(k)}$, $\forall\ k > 0$

(p. ej. si f es de variación regular)

Entonces si no es RR se tiene que las distribuciones a priori y a posteriori tienen el mismo comportamiento en la cola derecha para un valor x fijo.

DEMOSTRACION: Dado que

$$\frac{\theta\pi(\theta/x)}{1-\Pi(\theta/x)} = \frac{\theta f(\theta x)}{p_\theta^D(x)} \cdot \frac{\theta\pi(\theta)}{1-\Pi(\theta)}$$

tendremos que calcular:

$$\lim_{\theta\to\infty} \frac{\theta f(\theta x)}{p_\theta^D(x)}$$

pero

$$\frac{p_\theta^D(x)}{\theta f(\theta x)} = \int_\theta^\infty \frac{u f(ux)\pi(u)}{\theta f(\theta x)(1-\Pi(\theta))}\, du$$

y por las hipótesis iniciales sobre f:

$$0 < \frac{p_\theta^D(x)}{\theta f(\theta x)} < 1$$

Si además exigimos que F no sea RR; existe $k > 1$ tal que:

$$\lim_{x\to\infty} \frac{1-F(kx)}{1-F(x)} \neq 0 \iff \lim_{x\to\infty} \frac{f(kx)}{f(x)} \neq 0$$

es decir para $k > 1$, existe $\epsilon_0 > 0$ tal que:

$$\frac{f(kx)}{f(x)} > \epsilon_0 \quad \text{para infinitos valores de } x .$$

Entonces:

$$\frac{p_\theta^D(x)}{\theta f(\theta x)} = \int_1^\infty \frac{w f(w\theta x)\pi(w\theta)\,\theta}{f(\theta x)\,(1-\Pi(\theta))}\,dw >$$

$$> \int_1^\infty \frac{w\epsilon_0\,\pi(\theta w)}{1-\Pi(\theta)}\,dw = \epsilon_0 \int_\theta^\infty \frac{z\pi(z)}{\theta(1-\Pi(\theta))}\,dz > \epsilon_0$$

para infinitos valores de θ.

O sea:

$$\lim_{\theta\to\infty} \frac{p_\theta^D(x)}{\theta f(\theta x)} > 0$$

y por tanto

$$0 < \lim_{\theta\to\infty} \frac{\theta f(\theta x)}{p_\theta^D(x)} < \infty$$

c.q.d.

3.2. Influencia del tipo de verosimilitud en la divergencia asintótica entre las distribuciones a priori y a posteri

Seguiremos utilizando como f-divergencia la correspondiente a la medida de Kullback-Leibler.

$$\mathcal{J}_f(\pi(\theta),\pi(\theta/x)) = \int_0^\infty \left(\frac{\theta f(\theta x)}{p(x)} - 1\right)\left(\log \frac{\theta f(\theta x)}{p(x)}\right)\pi(\theta)\, d\theta$$

Por lo tanto el límite de esta expresión dependerá de:

$$\lim_{x\to\infty} \frac{\theta f(\theta x)}{p(x)}$$

y para efectuar su cálculo vamos a utilizar ciertos resultados de la teoría de Karamata (1933) sobre variación de funciones en el infinito.

En primer lugar, recordando los Teoremas 2.2.2 y 2.2.3. del Capítulo II tenemos que si $U:\mathbb{R}^+ \longmapsto \mathbb{R}^+$ es una función integrable Lebesgue en intervalos finitos:

i) $\displaystyle\lim_{x\to\infty} \frac{x\, U(x)}{\int_x^\infty U(t)\, dt} = \lambda$ con $0<\lambda<\infty$, es equivalente a

que U sea de variación regular en infinito con exponente $\rho = (-\lambda-1)$.

ii) Si para algún número real α :

$$\int_1^\infty t^\alpha\, U(t)\, dt < \infty$$

$$\lim_{x\to\infty} \frac{x^{\alpha+1}\, U(x)}{\int_x^\infty t^\alpha\, U(t)\, dt} = \infty$$

se verifica que U(x) es de variación rápida en infinito con exponente $\rho = -\infty$.

Y podemos concluir:

<u>TEOREMA 3.2.1:</u> i) Si F es una función de distribución RR se verifica que la cola derecha de la función de densidad f es de variación rápida en infinito con exponente $\rho = -\infty$.

ii) Si F es una función de distribución RN se verifica que la cola derecha de la función de densidad f es de variación regular en infinito con exponente $\rho = -c-1$.

<u>DEMOSTRACION</u>: i) Si F es RR $\Leftrightarrow \lim\limits_{x\to\infty} \dfrac{xf(x)}{1-F(x)} = \infty$ lo cual implica que la función de densidad restringida al intervalo $(0,\infty)$ es de variación rápida con exponente $\rho = -\infty$, en infinito.

ii) Si F es RN $\Leftrightarrow \lim\limits_{x\to\infty} \dfrac{xf(x)}{1-F(x)} = c$ para $0 < c < \infty$ y por tanto la cola derecha de la función de densidad es $(-c-1)$-variante en infinito.

c.q.d.

De esta forma en el caso de parámetro de escala podremos utilizar los resultados relativos al comportamiento asintótico cuando $x\to\infty$ de integrales (supuestamente bien definidas) del tipo:

$$\int_a^b f(t)\, R(xt)\, dt$$

donde R es una función de variación regular en infinito y $0 \leq a < b < \infty$.

Efectivamente la función de densidad predictiva es de esta forma, ya que:

$$p(x) = \int_0^{\infty} \theta f(\theta x)\pi(\theta)\, d\theta$$

Se puede encontrar en Seneta (1976) el estudio para funciones de variación lenta que exponemos y posteriormente ampliamos a variación regular en general, y aplicamos al caso que nos ocupa. En todos los resultados nos referiremos a variación en infinito:

<u>TEOREMA 3.2.2</u>: Sea $L(x)$ una función de variación lenta en $[A,\infty)$ para algún $A > 0$, y la integral: $\int_a^{\infty} t^c f(t)\, dt$ para $0 < a < \infty$ está bien definida para alguna función real dada f y un $c \geqslant 0$. Entonces la integral $\int_a^{\infty} f(t) L(xt)\, dt$ está bien definida

i) en general si $c > 0$

ii) cuando $c = 0$ hay que exigir que $L(t)$ sea finalmente no creciente.

En ambas situaciones:

$$\int_a^{\infty} f(t) L(xt)\, dt \sim L(x) \int_a^{\infty} f(t)\, dt$$

estando bien definida $\int_a^{\infty} f(t)\, dt$, cuando $x \to \infty$.

TEOREMA 3.2.3: Sea $L(x)$ una función de variación lenta en $(0, \infty)$ y acotada en cada subintervalo finito de $(0, \infty)$. Supongamos que la integral

$$\int_0^b t^{-c} f(t)\, dt$$

está bien definida para alguna función real dada f y un $c > 0$. Entonces cuando $x \to \infty$

$$\int_0^b f(t)\, L(xt)\, dt \sim L(x) \int_0^b f(t)\, dt$$

para $c > 0$, siendo necesario añadir que $L(x)$ es no decreciente para c igual a cero, en $(0, \infty)$.

Entonces resumiendo los dos resultados obtenemos:

TEOREMA 3.2.4: Sea $L(x)$ una función de variación lenta en $(0, \infty)$ y acotada en cada subintervalo finito de $(0, \infty)$. Supongamos que las siguientes integrales están bien definidas para algún $b > 0$ y $c > 0$, es decir:

$$\int_0^\infty t^c f(t)\, dt < \infty$$

$$\int_0^b t^{-c} f(t)\, dt < \infty$$

entonces como:

$$\int_b^\infty t^c\, f(t)\, dt < \int_0^\infty t^c\, f(t)\, dt$$

tambien está bien definida: $\int_b^\infty t^c\, f(t)\, dt$

y así:

$$\int_0^\infty f(t)\, L(xt)\, dt = \int_0^b f(t)\, L(xt)\, dt + \int_b^\infty f(t)\, L(xt)\, dt \sim$$

$$\sim L(x) \int_0^\infty f(t)\, dt$$

cuando $x \to \infty$.

Generalizando a funciones de variación regular, dado que si $R(t)$ es ρ-variante tiene la forma:

$$R(t) = x^\rho\, L(t)$$

donde $L(t)$ es una función de variación lenta, el tipo correspondiente de integral será:

$$\int_0^\infty f(t)\, R(xt)\, dt = \int_0^\infty (xt)^\rho\, f(t)\, L(xt)\, dt =$$

$$= x^\rho \int_0^\infty f_1(t)\, L(xt)\, dt\ .$$

Y podemos aplicar los teoremas para funciones de variación lenta con función asociada:

$$f_1(t) = t^{\rho} f(t)$$

TEOREMA 3.2.5: Sea R(t) una función de variación regular con exponente ρ, en $(0,\infty)$ y acotada en subintervalos finitos. Supongamos que:

$$\int_0^{\infty} t^{c+\rho} f(t)\, dt$$

está bien definida para algún $c \geqslant 0$, y que:

$$\int_0^{b} t^{-c+\rho} f(t)\, dt$$

también está bien definida para algún $c > 0$ y $0 < b < \infty$.

Entonces:

$$\int_0^{\infty} f(t)\, R(tx)\, dt \sim R(x) \int_0^{\infty} t^{\rho} f(t)\, dt$$

cuando $x \to \infty$.

Trasladando las hipótesis y conclusiones a nuestro planteamiento quedará de la siguiente forma:

TEOREMA 3.2.6: Sea f una función de densidad positiva, tal que restringida a $(0,\infty)$ es de variación regular con exponente ρ, y acotada en subintervalos finitos.

Supongamos una distribución a priori con densidad $\pi(\theta)$ tal que:

$$\int_0^\infty \theta^{c+\rho}\,\pi(\theta)\,d\theta$$

está bien definida para algún $c \geq 0$, es decir existe momento de orden $(c+\rho)$ para algún $c \geq 0$ de la distribución a priori.

Supongamos que $\int_0^b \theta^{-c+\rho}\,\pi(\theta)\,d\theta$

está bien definida para algún $c > 0$ y $0 < b < \infty$.

Entonces

$$p(x) = \int_0^\infty \theta f(\theta x)\,\pi(\theta)\,d\theta \sim f(x) \int_0^\infty \theta^{c+1}\,\pi(\theta)\,d\theta$$

cuando $x \to \infty$.

Con esta equivalencia conseguimos, en el caso de verosimilitudes con variación regular, calcular:

$$\lim_{x\to\infty} \frac{p(x)}{f(x)} = \int_0^\infty \theta^{\rho+1}\,\pi(\theta)\,d\theta$$

<u>TEOREMA 3.2.7</u>: Sea F una función de distribución RN, y sea una distribución a priori con función de densidad $\pi(\theta)$ tal que:

$$\int_0^\infty \theta^{-c}\,\pi(\theta)\,d\theta < \infty \quad \text{para} \quad 0 < c < \infty .$$

Entonces para cualquier $\theta > 0$:

$$\lim_{x\to\infty} \frac{\theta f(\theta x)}{p(x)} = \frac{\theta^{-c}}{\int_0^\infty \theta^{-c}\,\pi(\theta)\,d\theta}$$

DEMOSTRACION: Si F es RN => f es de variación regular con exponente $\rho = -c-1$ para $0 \leq c < \infty$, es decir:

$$\lim_{x \to \infty} \frac{f(\theta x)}{f(x)} = \theta^{-c-1} \qquad \forall\, \theta > 0$$

Además por el Teorema 3.2.6:

$$\lim_{x \to \infty} \frac{p(x)}{f(x)} = \int_0^\infty \theta^{-c}\, \pi(\theta)\, d\theta$$

por tanto:

$$\lim_{x \to \infty} \frac{\theta f(\theta x)}{p(x)} = \frac{\theta^{-c}}{\int_0^\infty u^{-c}\, \pi(u)\, du}$$

c.q.d.

Pasamos a continuación al cálculo del límite cuando $x \to \infty$ de la divergencia de Kullback-Leibler entre las distribuciones a priori y a posteriori.

TEOREMA 3.2.8: En las hipótesis del Teorema 3.2.7 y suponiendo además que:

$$\int_0^\infty (m(\theta)-1)\,(\log m(\theta))\, \pi(\theta)\, d\theta < \infty$$

donde

$$m(\theta) = \sup_x \frac{\theta f(\theta x)}{p(x)}$$

se verifica que:

$$\lim_{x\to\infty} \mathcal{J}_f\,(\pi(\theta),\pi(\theta/x)) = \int_0^\infty \left(\frac{\theta^{-c}}{M_c} - 1\right) \log \frac{\theta^{-c}}{M_c}\,\pi(\theta)\,d\theta$$

donde $M_c = \int_0^\infty \theta^{-c}\,\pi(\theta)\,d\theta$.

DEMOSTRACION: Utilizando el Teorema 3.2.7 y el Teorema de la Convergencia Dominada.

Para distribuciones RR tenemos que:

$$\lim_{x\to\infty} \frac{xf(x)}{1-F(x)} = \infty$$

lo que implica:

$f(x)$ de variación rápida con $\rho = -\infty$.

es decir

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = \begin{cases} \infty & x < 1 \\ 1 & x = 1 \\ 0 & x > 1 \end{cases}$$

por ser un comportamiento análogo al de $x^{-\infty}$.

Procediendo como en el caso anterior:

TEOREMA 3.2.9: Sea F una función de distribución RR, con función de densidad f, de cola derecha monótona, y sea π la función de densidad a priori, con esperanza finita. Entonces:

$$\lim_{x\to\infty} \frac{p(x)}{\theta f(\theta x)} = \infty \qquad \forall\, \theta > 0$$

DEMOSTRACION: Sea $\theta > 0$:

$$\lim_{x \to \infty} \frac{p(x)}{\theta f(\theta x)} = \lim_{y \to \infty} \frac{p(y/\theta)}{\theta f(y)}$$

Pero:

$$\frac{p(y/\theta)}{\theta f(y)} = \int_0^{\theta} \frac{uf(yu/\theta)\pi(u)}{\theta f(y)}\, du + \int_0^{\infty} \frac{uf(yu/\theta)\pi(u)}{\theta f(y)}\, du =$$

$$= I_1 + I_2 .$$

Respecto a la primera integral:

$$I_1 = \int_0^1 \frac{wf(wy)}{f(y)}\, \theta\pi(\theta w)\, dw$$

y como:

$$\left.\begin{array}{l} \dfrac{f(wy)}{f(y)} \geq 0 \\[2em] \forall\, w \in (0,1)\ ,,\ \lim\limits_{y \to \infty} \dfrac{f(wy)}{f(y)} = \infty \end{array}\right\} \Rightarrow \lim_{y \to \infty} I_1 = \infty$$

aplicando el Lema de Fatou

Para la segunda integral:

$$I_2 = \int_1^{\infty} \frac{wf(wy)}{f(y)}\, \pi(\theta w)\, dw$$

En este caso por ser f de cola monótona, tenemos que para va lores de y suficientemente grandes y para todo w mayor que uno:

$$\left| \frac{wf(wy)}{f(y)} \right| < w$$

y por existir la esperanza de la distribución a priori:

$$\int_1^{\infty} \frac{wf(wy)}{f(y)} \pi(\theta w)\, dw < \int_1^{\infty} w\pi(\theta w)\, dw < \infty$$

Con estas condiciones, aplicando el Teorema de la Convergencia Dominada:

$$\lim_{y \to \infty} \frac{f(wy)}{f(y)} = 0 \qquad \forall\, w > 1$$

implica:

$$\lim_{y \to \infty} I_2 = 0$$

En resumen:

$$\lim_{x \to \infty} \frac{p(x)}{\theta f(\theta x)} = \infty$$

c.q.d.

Y calculando la divergencia asintótica de Kullback-Leibler, conluimos:

TEOREMA 3.2.10: Dada F, función de distribución RR y verificandose las condiciones del teorema anterior:

$$\lim_{x\to\infty} \mathcal{J}_f(\pi(\theta),\pi(\theta/x)) = \infty \qquad \text{para } \theta > 0$$

DEMOSTRACION:

$$\mathcal{J}_f(\pi(\theta),\pi(\theta/x)) = \int_0^\infty \left(\frac{\theta f(\theta x)}{p(x)} - 1\right) \log \frac{\theta f(\theta x)}{p(x)}\, \pi(\theta)\, d\theta$$

si consideramos:

$$h_x(\theta) = \left(\frac{\theta f(\theta x)}{p(x)} - 1\right) \log \frac{\theta f(\theta x)}{p(x)}$$

verifica:

$$\left.\begin{array}{l} h_x(\theta) \geq 0 \\ \\ \lim\limits_{x\to\infty} h_x(\theta) = \infty \end{array}\right\} \Rightarrow \lim_{x\to\infty} \mathcal{J}_f(\pi(\theta),\pi(\theta/x)) = \infty$$

aplicando nuevamente el Lema de Fatou.

c.q.d.

Finalmente nos referiremos a las funciones de distribución RP. En este caso:

$$\lim_{x\to\infty} \frac{xf(x)}{1-F(x)} = 0$$

Vamos a restringirnos a aquellas distribuciones para las que su función de densidad es de variación regular en infinito o sea existe:

$$\lim_{x \to \infty} \frac{f(kx)}{f(x)}$$

para todo k, mayor que cero.

De esta forma, aplicando la regla de L'Hopital:

$$\lim_{x \to \infty} \frac{1-F(kx)}{1-F(x)} = 1 = \lim_{x \to \infty} \frac{kf(kx)}{f(x)}$$

y así necesariamente f debe ser de variación regular con exponente $\rho = -1$:

$$\lim_{x \to \infty} \frac{f(kx)}{f(x)} = k^{-1} \qquad \forall\ k > 0$$

con estas hipótesis de partida, obtenemos los correspondientes resultados:

<u>TEOREMA 3.2.11</u>: Sea F una función de distribución RP y tal que su función de densidad f es de variación regular en infinito. Sea una distribución a priori, con esperanza finita. Entonces:

$$\lim_{x \to \infty} \frac{p(x)}{\theta f(\theta x)} = 1 \qquad \forall\ \theta > 0$$

<u>DEMOSTRACION</u>: Como en casos anteriores, calculamos:

$$\frac{p(y/\theta)}{\theta f(y)} = \int_0^{\theta} \frac{uf(yu/\theta)\ \pi(u)}{\theta f(y)}\, du + \int_{\theta}^{\infty} \frac{uf(yu/\theta)\pi(u)\ du}{\theta f(y)} =$$

$$= I_1 + I_2 .$$

Para la primera integral consideramos que por ser $f(x)$ de variación regular con exponente $\rho = -1$:

$$\lim_{y \to \infty} \frac{wf(wy)}{f(y)} = 1$$

y además la convergencia es uniforme en intervalos finitos.

Entonces:

$$I_1 = \int_0^1 \frac{wf(wy)}{f(y)} \pi(\theta w)\, dw$$

converge a:

$$\int_0^1 \pi(\theta w)\, \theta\, dw = \Pi(\theta) \quad \text{cuando } y \to \infty .$$

La segunda integral será:

$$I_2 = \int_1^\infty \frac{wf(wy)}{f(y)} \pi(\theta w)\, \theta\, dw$$

Por ser f de variación regular en infinito podemos suponer que es de cola derecha monótona, sin perder generalidad.

Además: $\int_0^\infty \theta\pi(\theta)\, d\theta < \infty$

entonces $\forall\, w > 1$: $\frac{f(wy)}{f(y)} < 1$

y así: $\int_1^\infty w\pi(\theta w)\, \theta\, dw < \infty$

Aplicando el Teorema de la Convergencia Dominada:

$$I_2 \longmapsto \int_1^{\infty} \pi(\theta w)\ \theta\ dw = 1 - \Pi(\theta)$$

cuando y crece indefinidamente.

Por tanto: $\lim\limits_{x \to \infty} \dfrac{p(x)}{\theta f(\theta x)} = 1$

c.q.d.

<u>TEOREMA 3.2.12</u>: En las hipótesis del teorema anterior, si F es una función de distribución RP, se tiene que:

$$\lim_{x \to \infty} \mathcal{J}_f\ (\pi(\theta/x),\ \pi(\theta)) = 0 \qquad \forall\ \theta > 0$$

suponiendo que $m(\theta) = \sup\limits_{x} \dfrac{\theta f(\theta x)}{p(\theta)}$ verifica:

$$\int_0^{\infty} (m(\theta)-1)\ (\log\ m(\theta))\ \pi(\theta)\ d\theta < \infty$$

<u>DEMOSTRACION</u>: Utilizando el Teorema anterior y el de la Convergencia Dominada.

<u>OBSERVACIONES</u>:

Como puede deducirse de los resultados obtenidos en el caso de parámetro de escala se extraen conclusiones análogas al caso de parámetro de localización. En particular es destacable el papel que juega el concepto de variación regular, lo que nos permite afirmar la convergencia uniforme en intervalos finitos, necesaria en algunas demostraciones, así como los resultados relativos al tipo de integrales que describen las funciones de densidad predictivas.

CAPITULO IV

RELACION CON OTRAS CLASES DE DISTRIBUCIONES

CAPITULO IV

RELACION CON OTRAS CLASES DE DISTRIBUCIONES

1. SUMARIO

Vamos a considerar ciertas clases de distribuciones que surgen en el análisis de supervivencia, por tanto nos restringiremos a variables no negativas.

En primer lugar plantearemos las relaciones que existen con las definiciones de distribuciones IFR, DFR, IFRA y DFRA según sea la razón de fallo. A continuación trataremos con las distribuciones NBU, NWU, NBUE y NWUE, que se aplican en las denominadas políticas de mantenimiento utilizadas para reducir la incidencia del fallo en el sistema.

2. CLASES DE DISTRIBUCIONES EN TEORIA DE LA FIABILIDAD

Para explicar el origen de las distintas definiciones, es necesario introducir el concepto de envejecimiento:

Supongamos una variable aleatoria X con función de distribución F, tal que F(0) = 0 por tanto la función de supervivencia en un valor x será $\bar{F}(x) = 1-F(x)$.

Así la función de supervivencia condicional de la variable, si ha durado hasta el instante t es:

$$\bar{F}(x/t) = \frac{\bar{F}(t+x)}{\bar{F}(t)} \qquad \text{si } \bar{F}(t) > 0$$

Análogamente la probabilidad condicionada de fallo en un intervalo de amplitud x, para una variable de edad igual a t es:

$$F(x/t) = \frac{F(t+x) - F(t)}{\overline{F}(t)} = 1 - \overline{F}(x/t)$$

A partir de esta distribución se define, cuando exista, la razón de fallo en el instante t:

$$r(t) = \lim_{x \to 0} \frac{1}{x} \frac{F(t+x) - F(t)}{\overline{F}(t)}$$

es decir

$$r(t) = \frac{f(t)}{\overline{F}(t)} = \frac{f(t)}{1-F(t)}$$

cuando exista función de densidad $f(t)$ y sea $\overline{F}(t) > 0$.

Con este planteamiento surgen las siguientes clases de distribuciones (Barlow y Proschan (1975)).

DEFINICION 2.1: Diremos que una función de distribución F es IFR si la función de supervivencia condicionada es una función decreciente de la edad:

$$\overline{F}(x/t) = \frac{\overline{F}(t+x)}{\overline{F}(t)} \text{ es decreciente en } t \geqslant 0 \text{ para cada } x \geqslant 0$$

En el caso de existir la función de densidad, esta definición será equivalente a considerar funciones de distribución con razón de fallo monótona creciente.

De la misma forma:

DEFINICION 2.2: Una función de distribución F será DFR si la función de supervivencia condicionada es creciente con

con la edad, es decir

$$\bar{F}(x/t) = \frac{\bar{F}(t+x)}{\bar{F}(t)} \quad \text{es creciente en } t \geqslant 0 \text{ para cada } x \geqslant 0$$

Para funciones de distribución con función de densidad será equivalente a exigir que la razón de fallo sea monótona decreciente.

EJEMPLOS

- La distribución de Weibull, con función de distribución:

$$F_\alpha(t) = 1 - e^{-(\lambda t)^\alpha} \quad \text{para } t \geqslant 0 \quad \text{tal que} \quad \lambda,\ \alpha > 0$$

es IFR para $\alpha \geqslant 1$ y DFR para $0 < \alpha \leqslant 1$.

- La distribución Gamma: $\gamma(p,a)$ será DFR para $0 < p \leqslant 1$ y IFR para $p \geqslant 1$.

Una generalización de las funciones anteriores la constituyen los conceptos de distribución IFRA y DFRA respectivamente y que enunciamos seguidamente:

DEFINICION 2.3: Una función de distribución F, será IFRA si la función $-\frac{\log(1-F(t)}{t}$ es creciente en $t \geqslant 0$.

Para los casos en que exista la razón de fallo, será equivalente a que $\frac{\int_0^t r(u)\,du}{t}$ sea creciente en $t \geqslant 0$, ya que

$$-\log\,(1-F(t)) = \int_0^t r(u)\,du$$

en dicha situación.

DEFINICION 2.4: Una función de distribución F, será DFRA si la función $\frac{-\log(1-F(t))}{t}$ es decreciente en $t \geq 0$ ó equivalentemente si $\frac{\int_0^t r(u)\,du}{t}$ es decreciente, en $t \geq 0$, cuando exista la razón de fallo.

Si consideramos la definicion de distribución IFR, podríamos interpretarla en términos de propensión y resistencia a extremos alejados. En efecto, la función de supervivencia $\bar{F}(x/t)$, como es la probabilidad de obtener valores que distan x de t suponiendo que la variable es mayor que t, indica, al ser decreciente en t, que la probabilidad de obtener extremos alejados no puede ser alta .

Esta interpretación queda justificada por el siguiente resultado. Suponemos en todo caso que $x_F = \infty$.

PROPOSICION 2.1: Si F es una distribución IFR, se verifica que F no puede ser SP.

DEMOSTRACION: Teniendo en cuenta que, dado $\epsilon > 0$:

$$Pr(X_{n:n} - X_{n-1:n} > \epsilon) = \int_0^\infty \frac{1-F(x+\epsilon)}{1-F(x)}\, dF_{X_{n-1:n}}(x) <$$

$$< \int_0^\infty (1-F(\epsilon))\, dF_{X_{n-1:n}}(x) = 1-F(\epsilon)$$

al ser $\frac{1-F(x+\epsilon)}{1-F(x)}$ decreciente en ϵ , Entonces para todo $\epsilon > 0$, existe $k(\epsilon)$, tal que

$$Pr(X_{n:n} - X_{n-1:n} > \epsilon) < k \quad , \text{ para todo } n$$

es decir:

$$S_n = X_{n:n} - X_{n-1:n} \equiv O_P(1)$$

y por tanto S_n no puede converger en probabilidad hacia infinito.

Si suponemos que existe la razón de fallo para todo x, obtenemos las siguientes implicaciones:

PROPOSICION 2.2: Si F es una función de distribución IFR, se verifica que F es RR.

DEMOSTRACION: Si F es IFR se tiene que :

$$r(x) = \frac{f(x)}{1-F(x)} \quad \text{es una función creciente}$$

por lo tanto,

$$\lim_{x \to \infty} \frac{f(x)}{1-F(x)} \neq 0$$

que implica:

$$\lim_{x \to \infty} \frac{xf(x)}{1-F(x)} = \infty$$

o sea F es RR.

Para el caso de distribuciones DFR nos queda el resultado correspondiente a la resistencia:

PROPOSICION 2.3: Si F es una función de distribución DFR, entonces F no puede ser SR.

DEMOSTRACION: Como en el caso anterior consideramos:

$$\text{dado } \epsilon > 0\text{: } Pr(X_{n:n} - X_{n-1:n} > \epsilon) = \int_0^{\infty} \frac{1-F(x+\epsilon)}{1-F(x)} \, dF_{X_{n-1:n}}(x) >$$

$$> 1-F(\epsilon), \text{ por ser } F \text{ de tipo DFR.}$$

Si ϵ es tal que $1-F(\epsilon) > 0$ tenemos que:

existe ϵ, tal que para $k(\epsilon) = 1-F(\epsilon) > 0$

$$Pr(X_{n:n} - X_{n-1:n} > \epsilon) > k(\epsilon) > 0$$

es decir, $S_n = X_{n:n} - X_{n-1:n}$ no puede converger a cero en probabilidad, y por tanto F no puede ser SR.

Para las siguientes relaciones, vamos a utilizar un resultado que describe el comportamiento en la cola de la función de distribución según sea IFRA ó DFRA.

TEOREMA 2.1: (Barlow y Proscham (1975)). Sea F una función de distribución IFRA (DFRA) con cuantil de orden p, x_p. Entonces:

$$1 - F(t) \begin{cases} \geq (\leq)\ e^{-at} & \text{para } 0 \leq t \leq x_p \\ \\ \leq (\geq)\ e^{-at} & \text{para } t \geq x_p \end{cases}$$

donde: $a = -\frac{1}{x_p} \log(1-p)$.

Si ademas tenemos en cuenta que IFR implica IFRA y que DFR implica DFRA, podremos demostrar las siguientes proposiciones que siguen asimilando razón de fallo creciente con resistencia a datos atípicos y razón de fallo decreciente con propensión.

<u>PROPOSICION 2.4</u>: Si F es una función de distribución IFR, se verifica que F es SR, suponiendo $x_F = \infty$.

<u>DEMOSTRACION</u>: Por ser F, tipo IFR:

$$\frac{1-F(x+\epsilon)}{1-F(x)} \quad \text{es decreciente en } x, \quad \forall \epsilon > 0$$

Por tanto

$$\lim_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)} \leqslant \frac{1-F(x+\epsilon)}{1-F(x)} < 1 - F(\epsilon) < 1$$

si suponemos que existe c, $0 < c < \infty$, tal que:

$$\text{dado } \epsilon > 0\text{: } \lim_{x\to\infty} \frac{1-F(x+\epsilon)}{1-F(x)} = e^{-c\epsilon} > 0 \text{ (o sea F, SN)}$$

entonces:

$$e^{-c\epsilon} < 1 - F(\epsilon)$$

y esto para cualquier $\epsilon > 0$, por lo que F no puede ser IFRA por el Teorema 2.1 y por tanto F tampoco puede ser IFR, en contradicción con la hipótesis, así que:

$$\forall\ \epsilon > 0 \qquad \lim_{x \to \infty} \frac{1-F(x+\epsilon)}{1-F(x)} = 0$$

equivalente a que F es SR.

<u>PROPOSICION 2.5:</u> Si F es una función de distribución DFR, entonces F es SP, suponiendo $F(x) > 0$, para todo x.

<u>DEMOSTRACION</u>: Si F es DFR:

$$\frac{1-F(x+\epsilon)}{1-F(x)} \quad \text{es creciente en x, para cada } \epsilon > 0$$

por lo que:

$$0 < 1-F(\epsilon) < \frac{1-F(x+\epsilon)}{1-F(x)} \leqslant \lim_{x \to \infty} \frac{1-F(x+\epsilon)}{1-F(x)}$$

Como en el caso anterior, si suponemos que F es SN:

$$\text{dado } \epsilon > 0 \quad \lim_{x \to \infty} \frac{1-F(x+\epsilon)}{1-F(x)} = e^{-c\epsilon} \quad \text{donde} \quad 0 < c < \infty .$$

y llegamos a contradicción por lo tanto F es SP.

También podemos enunciar los resultados inversos aunque con carácter negativo

<u>PROPOSICION 2.</u>6: Si F es una función de distribución RN se verifica que F no puede ser IFRA por tanto no puede ser IFR.

<u>DEMOSTRACION</u>: Por ser F una función de distribución RN se tiene que la función:

$$U(x) = 1 - F(x)$$

es de variación regular con exponente $\rho = -c$ donde $0 < c < \infty$.

Por tanto:

$$\lim_{x \to \infty} \frac{\log(1-F(x))}{\log x} = \rho$$

como

$$\lim_{x \to \infty} \frac{x}{\log x} = \infty$$

tenemos:

$$\lim_{x \to \infty} \frac{-\log(1-F(x))}{x} = 0$$

por lo que la función: $\frac{-\log(1-F(x))}{x} \geqslant 0$

no puede ser creciente, es decir F no puede ser IFRA.

<u>PROPOSICION 2.7</u>: Si F es RP, entonces F no puede ser IFRA, por tanto no puede ser tampoco IFR.

<u>DEMOSTRACION</u>: Por ser F una función de distribución RP, se verifica que la función:

$$U(x) = 1 - F(x)$$

es de variación lenta, por tanto:

$$\lim_{x \to \infty} \frac{\log(1-F(x))}{\log x} = 0$$

es decir:

$$\lim_{x \to \infty} \frac{\log (1-F(x))}{x} = 0$$

y entonces se tiene que la función $\frac{-\log (1-F(x))}{x} \geq 0$ no puede ser creciente, así F no puede ser IFRA.

PROPOSICION 2.8: Si F es una función de distribución SP, entonces F no puede ser IFRA.

DEMOSTRACION: Como F es SP:

$$\lim_{x \to \infty} \frac{f(x)}{1-F(x)} = 0 = \lim_{x \to \infty} \frac{-d \log(1-F(x))}{dx}$$

y así aplicando L'Hopital:

$$\lim_{x \to \infty} \frac{-\log (1-F(x))}{x} = 0$$

es decir F no puede ser IFRA.

PROPOSICION 2.9: Si F es una función de distribución SR se verifica que F no puede ser DFRA y por tanto no puede ser DFR.

DEMOSTRACION: Si F es SR:

$$\lim_{x \to \infty} \frac{f(x)}{1-F(x)} = \infty$$

y análogamente a la demostración anterior, aplicando L'Hopital

$$\lim_{x \to \infty} \frac{-\log (1-F(x))}{x} = \infty$$

así F no puede ser DFRA.

Y para concluir este capítulo, relacionaremos nuestras definiciones con una nueva generalización de los conceptos anteriores (Barlow y Proscham (1975)).

DEFINICION 2.5: Una función de distribución F sera NBU si:

$$\bar{F}(x+y) \leqslant \bar{F}(x) \, . \, \bar{F}(y) \text{ para } x \geqslant 0 \quad y \geqslant 0$$

DEFINICION 2.6: Una función de distribución F será NWU si:

$$\bar{F}(x+y) \geqslant \bar{F}(x) \, . \, \bar{F}(y) \text{ para } x \geqslant 0 \quad y \geqslant 0$$

DEFINICION 2.7: Una función de distribución F será NBUE si:

a) F tiene esperanza ya sea finita ó infinita μ .

b) $\int_t^\infty \bar{F}(x)\, dx \leqslant \mu \bar{F}(t) \quad \text{para } t \geqslant 0$.

DEFINICION 2.8: Una función de distribución F será NWUE si:

a) F tiene esperanza ya sea finita ó infinita μ.

b) $\int_t^\infty \bar{F}(x)\, dx \geqslant \mu \bar{F}(t) \quad \text{para } t \geqslant 0$

En todos los casos se tiene la igualdad para la distribución Exponencial, ya que es la única distribución con razón de fallo constante.

Las cadenas de implicaciones son:

IFR => IFRA => NBU => NBUE

DFR => DFRA => NWU => NWUE

y así razonando como en las proposiciones anteriores obtenemos unos resultados más generales.

PROPOSICION 2.10: Si F es una función de distribución NBU se tiene que F no puede ser SP.

DEMOSTRACION: Igual que la Proposición 2.1.

PROPOSICION 2.11: Si F es una función de distribución NWU, se tiene que F no puede ser SR.

DEMOSTRACION: Como la Proposición 2.3.

Las demostraciones coinciden con las correspondientes a las funciones de distribución IFR y DFR porque en ellas no utilizábamos el crecimiento y decrecimiento respectivamente de $\bar{F}(x/t)$ sino las acotaciones que implicaban y que coinciden con las definiciones de NBU y NWU.

Seguidamente suponiendo que existe la razón de fallo:

$$\lim_{t\to\infty} \frac{f(t)}{1-F(t)} = \lim_{x\to\infty} \frac{1-F(t)}{\int_t^{\infty} (1-F(x))\,dx}$$

aplicando L'Hopital. Por tanto se tienen las relaciones:

PROPOSICION 2.12: Si F es una función de distribución NBUE y la media de la distribución μ es finita se tiene que F no puede ser SP.

<u>DEMOSTRACION</u>: Por ser F NBUE.

$$\frac{\int_t^\infty (1-F(x))\,dx}{1-F(t)} \leqslant \mu < \infty \qquad \forall\, t \geqslant 0$$

es decir:

$$\lim_{t\to\infty} \frac{\int_t^\infty (1-F(x))\,dx}{1-F(t)} \leqslant \mu < \infty$$

y entonces aplicando la regla de L'Hopital:

$$\lim_{t\to\infty} \frac{1-F(t)}{f(t)} < \infty$$

que implica:

$$\lim_{t\to\infty} \frac{f(t)}{1-F(t)} \neq 0$$ y así F no puede ser SP.

<u>PROPOSICION 2.13</u>: Si F es NWUE y tal que F(0) = 0 se tiene que F no puede ser SR.

<u>DEMOSTRACION:</u> Por ser F, NWUE

$$\int_t^\infty \frac{(1-F(x))\,dx}{1-F(t)} \geqslant \mu > 0 \qquad \text{para } t \geqslant 0$$

lo que implica

$$\lim_{t\to\infty} \frac{\int_t^\infty (1-F(x))\,dx}{1-F(t)} \geqslant \mu > 0$$

y por tanto, como en el caso anterior:

$$\lim_{t \to \infty} \frac{f(t)}{1-F(t)} \neq \infty \quad \text{ó sea } F \text{ no puede ser SR.}$$

<u>Observaciones</u>: Incluso con las sucesivas ampliaciones del concepto se sigue manteniendo la incompatibilidad de las definiciones NBUE y NWUE con la propensión y la resistencia respectivamente.

Es interesante resaltar también que en el caso de distribuciones DFR y DFRA se mantiene el carácter bajo la operación de mixtura en general y para NWU y NWUE sólo en distribuciones que no se cruzan. Esto no ocurre en ningún caso para las demás definiciones. Es decir que el comportamiento de resistencia no tiene por qué mantenerse al construir mixturas ya que pueden ser modelos para explicar datos atípicos, y por tanto correspondientes a distribuciones propensas.

APENDICE I

APENDICE I

<u>TEOREMA:</u> Sea $\{u_n(x)\}$ una sucesión de funciones de $\bar{\mathbb{R}} \longrightarrow \bar{\mathbb{R}}$ monótonas y acotadas que convergen puntualmente a una función $u(x)$, continua (por tanto monótona), entonces la convergencia es uniforme.

<u>DEMOSTRACION:</u> Supongamos que $\{u_n(x)\}$ es una sucesión de funciones monótonas no crecientes y acotadas, por lo tanto:

$$u(x) = \lim_{n \to \infty} u_n(x) \quad \text{monótona no creciente}$$

$$|u(x)| \leqslant M \quad , \text{ para algún } M > 0$$

Entonces, dado $x \in \bar{\mathbb{R}}$ y $\varepsilon > 0$, existe $n_0(x, \varepsilon)$ tal que:

$$| u_n(x) - u(x)| < \varepsilon \qquad \forall\, n \geqslant n_0$$

es decir

$$u(x) - \varepsilon \leqslant u_n(x) \leqslant u(x) + \varepsilon \qquad \forall\, n \geqslant n_0$$

Sean $\{x_1, x_2, \ldots, x_k\}$ tales que:

$$x_1 \leqslant x_2 \leqslant \cdots \leqslant x_k \quad \text{con} \quad x_1 = -\infty \text{ y } x_k = +\infty$$

y además:

$$u(x_i) - u(x_i+1) < \varepsilon \quad \forall\, i = 1, \ldots, k.$$

Sea i tal que $x \in (x_i, x_{i+1})$, entonces:

existe n_i' tal que, $\forall\, n \geqslant n_i$ se verifica

$$u(x_i) - \varepsilon \leqslant u_n(x_i) \leqslant u(x_i) + \varepsilon$$

existe n_i'' tal que $\forall\ n \geq n_i''$ se verifica

$$u(x_{i+1}) - \epsilon \leq u_n(x_{i+1}) \leq u(x_{i+1}) + \epsilon$$

Luego si $n_i = \max\{n_i', n_i''\}$, $\forall\ n \geq n_i$

$$u(x_{i+1}) - \epsilon \leq u_n(x_{i+1}) \leq u_n(x) \leq u_n(x_i) \leq u(x_i) + \epsilon$$

y así

$$-2\epsilon \leq u(x_{i+1}) - u(x) - \epsilon \leq u_n(x) - u(x) \leq u(x_i) - u(x) + \epsilon \leq 2\epsilon$$

Por tanto, dado ϵ', para todo $x \in (x_i, x_{i+1})$, existe n_i tal que:

$$|u_n(x) - u(x)| \leq \epsilon' \qquad \forall\ n \geq n_i$$

Considerando $i = 1, \ldots, k$ obtenemos $n_1, n_2, \ldots n_k$; entonces si:

$$n_o = \max\{n_1, \ldots, n_k\}$$

será:

$$|u_n(x) - u(x)| \leq \epsilon' \qquad \forall\ n \geq n_o\ , \quad \forall\ x$$

es decir, la convergencia es uniforme.

c.q.d.

BIBLIOGRAFIA

BIBLIOGRAFIA

1.- ANDERSON, T W. and DARLING, D.A. (1954). A test of goodness of fit, J.A.S.A., 49, 765-769

2.- ANSCOMBE, F.J. (1960). Rejection of outliers. Technometrics, 2, 123-147.

3 - BARLOW, R.E and PROSCHAM, F. (1975). Statistical Theory of Reliability and Life Testing. Holt, Rinehart and Winston Inc.

4.- BARNETT, V. and LEWIS, T.(1984). Outliers in Statistical Data (2nd ed), Wiley.

5.- BECKMAN, R J. and COOK, R.D.(1983). Outlier...s. Technometrics, 25, 119-163.

6.- BERNOUILLI, D. (1977). Dijudicatio maxime probabilis plurium observatiorum discrepatium atque verisimillima inductio inde formanda. Acta Academiae Scientorum Petropolitanae, 1, 3-33. English translation by C.G. Allen (1961), Biometrika, 48, 3-13.

7.- BOSCOVICH (1957). De literaria expeditione per pontificiam ditionem, et synopsis amplioris operis, ac habentur plura ejus ex exemplaria etiam sensorum impressa. Bononiensi Scientiarum et Artum Instuto atque Academia Commentarii, 4, 353-396.

8.- BOX, G.E.P., and TIAO (1968). A Bayesian approach to some outlier problems. Biometrika, 55, 119-129.

9.- COLLET, D. and LEWIS, T. (1976). The subjective nature of outlier rejection procedures. Applied Statistics, 25, 228-237.

10.- CHAUVENET, W. (1863). Method of least squares. Appendix to Manual of Spherical and Practical Astronomy, Vol. 2, Lippincott, Philadelphia, pp. 469-566; tables 593-599. Reprinted (1960) 5th edn. Dover, New York.

11.- CHOW, Y.S. and TEICHER, H. (1978). Probability Theory. Springer-Verlag.

12.- DAVID, H.A., (1981). Order Statistics, 2nd ed. Wiley, New York.

13.- DAVID, H.A., HARTLEY, H.O, and PEARSON, E.S. (1954). The distribution of the ratio, in a single normal sample, of range to standard deviation. Biometrika, 41, 482-493.

14.- DAVID, H.A., and PAULSON, A.S. (1965). The performance of several tests for outliers. Biometrika, 52, 429-436.

15.- DAWID, A.P. (1973). Posterior expectations for large observation Biometrika, 60, 664-667.

16.- DE FINETTI, B. (1961). The Bayesian approach to the rejection of outliers. Proceedings of the Fourth Berkeley Sumposium on Probability and Statistics, Volume 1, pp. 199-210. Berkeley, Calif.: University Press.

17.- DE HAAN, L. (1970). On regular variation and its application to the weak convergence of sample extremes. Amsterdam Math. Centre Tracts 32.

18.- DE HAAN, L. (1971). A form of regular variation and its application to the domain of attraction of the dou ble exponential distribution. Z. Wahrsch. verw. Ge biete 17, 241-258.

19.- DE HAAN, L. (1974). Equivalence classes of regularly varying functions, Stochastic processes and their applications, 3, 243-259.

20.- DE HAAN, L. (1976). Sample extremes: an elementary intro duction. Statist. Neerlandica 30, 161-172.

21.- DIXON, W.J. (1950). Analysis of extreme values. Ann. Statist., 21, 488-506.

22.- DIXON, W.J. (1951). Ratios involving extreme values. Ann Math. Statist. 22, 68-78.

23.- DOOD, E.L. (1923). The greatest and the least variate under general laws of error. Trans. Amer. Math. Soc. 25, 525-539.

24.- FELLER, W. (1971). An Introduction to Probability Theory and its Applications, Vol. 2, 2nd ed. New York, Wiley.

25.- FERGUSON, T.S. (1961). On the rejection of outliers. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 253-287.

26.- FIELLER, N.R.J. (1976). Some Problems related to the Rejection of Outlying Observations. Ph. D. Thesis, University of Hull.

27.- FISHBURN, P.C. and VICKSON, R.G. (1978). Theoretical foundations of stochastic dominance. Stochastic Dominance: An Approach to Decision Making under

Risk (G.A. Whitmore and M.C. Findley, eds.). Lexington, Mass: D C. Heath.

28.- FISHER, R.A. and TIPPETT, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Cambridge Phil. Soc. 24, 180-190.

29.- FRECHET, M. (1927). Sur la loi de probabilité de l'écart maximum. Ann. Soc. Math. Polon. 6, 93-116.

30.- GEFFROY, J. (1958/1959). Contributions a la théorie des valeurs extrémes. Publ. Inst. Stat. Univ. Paris.

31.- GLAISHER, J.W.L. (1872-73). On the rejection of discordant observations. Monthly Notices Roy. Astr. Soc. 33, 391-402.

32.- GNEDENKO, B.V. (1943). Sur la distribution limite du terme maximum d'une serie aléatoire. Ann. Math. 44, 423-453.

33. GNEDENKO, B.V. and KOLMOGOROV, A.N. (1954). Limit distributions for Sums of Independent Random variables. Addison-Wesley Pb. Co.

34.- GOEL, P.K. (1983). Information Measures and Bayesian Hierarchical Models. J.A.S.A., 78, 408-410.

35.- GREEN, R.F. (1974). A note on outlier-prone families of distributions. Ann. Statist. 2, 1293-1295.

36.- GREEN, R.F. (1976). Outlier-prone and outlier-resistant distributions. J. Amer. Statis. Ass. 71, 502-505.

37.- GRUBBS, F.E. (1950). Sample criteria for testing outlying observations. Ann. Math. Statist. 21, 27-58.

38.- HALPERIN, M. et al. (1955). Tables of percentage points for the stundentized maximum absolute deviate in normal samples. J. Amer. Statist. Ass. 50, 185-195.

39.- HAWKINS, D.M. (1980). Identification of Outliers. Chapman and Hall, London.

40.- HUBER, P.J. (1981). Robust Statistcs Wiley, New York.

41.- IBRAGIMOV, I.A. (1956). On the composition of unimodal distributions. Theory of Probability and Its Appli cations, 1, 255-260.

42.- IRWIN, J.O. (1925). On a criterio for the rejection of outlying observations. Biometrika 17, 238-250.

43.- KALE, B.K. (1977). A note on outlier resistant families an mixtures of distributions. J. Ind. Statist.Assn. 15, 119.

44.- KALE, B.K. and SINHA, S.K. (1971). Estimation of expected life in the presence of an outlier observation. Technometrics, 13, 755-759.

45.- KARAMATA, J. (1930). Sur un mode de croissance réguliére des fonctions. Mathematica (Cluj) 4, 38-53.

46.- KARAMATA; J. (1933). Sur un mode de croissance réguliére. Thèorèmes fondamentaux. Bull. Soc. Math. France, 61, 55-62.

47.- KIMBER, A.C. (1982). Tests for many outliers in an exponential sample. Applied Statistics, 31, 263-271.

48.- KITAGAWA, G. (1977). On the use of AIC for the detection of outliers. Technometrics, 21, 193-199.

49.- KULLBACK, S. (1968). Information Theory and Statistics, New York: Dover Publications.

50.- LEADBETTER, M.R., LINDGREN, G., ROOTZEN, H. (1983). Extremes and Related Properties of Random Sequences and Processes. Springer Verlag. New York Inc.

51.- MENDELEEV, D.I. (1895). Course of work on the renewal of prototypes or standard measures of lengths and weights. (Russian). Vremennik Glavnoi Palaty Mer i Vesov, 2, 157-185.

52.- NEWCOMB, S. (1886). A generalized theory of the combination of observations so as to obtain the best result. Amer, J. Math. 8, 343-366.

53.- NEYMAN, J. and SCOTT, E.L. (1971). Outlier proneness of phenomena and of related distributions.Optimizing Methods in Statistics (J.S. Rustagi, ed.) New York: Academic Press.

54.- O'HAGAN, A (1979). On outlier rejection phenomena in Bayes inference. J. Roy. Statist. Soc. B, 41, 358--367.

55.- PARZEN, E. (1980). Quantile Functions, Convergence in Quantile, and Extreme Value Distribution Theory. Technical Report No. B-3, Texas A & M University, Institute of Statistics.

56.- PEARSON, E.S. and CHANDRA SEKAR, C. (1936). The efficiency of statistical tcls and a criterion for the rejection of outlying observations. Biometrika, 28, 308-320.

57.- PEARSON, E.S. and STEPHENS, M.A. (1964). The ratio of range to standard deviation in the same normal sam ple. Biometrika, 51, 484-487.

58.- PEIRCE, B. (1852). Criterion for the rejection of doubt ful observations. Astr. J., 2, 161-163.

59.- SCHUSTER,E.F. (1984). Classification of Probability Laws by Tail Behaviour. J. Am. Stat. Assn.79, 936-939.

60.- SENETA, E. (1976). Regularly Varying Functions. Lecture Notes in Mathematics 508, Springer-Verlag.

61.- STONE, E.J. (1868). On the rejection of discordant observations. Monthly Notices Roy. Astr. Soc. 28, 165-168.

62.- THOMPSON, W.R. (1935). On a criterion for the rejection of observations and the distribution of the ratio of the deviation to the sample standard deviation. Ann. Math. Statist 6, 214-219.

63.- TIAGO DE OLIVEIRA, J. ed. (1984). Statistical Extremes and Applications. D. Reidel Publishing Co.

64.- TIKU, M.L. (1975). A new Statistic for testing suspectec outliers. Commun. Statist. A, 4, 737-752.

65.- VON MISSES, R.V. (1936). La distribution de la plus grande de n valeurs. Rev. Math. Union Interbalkan, 1, 141-160.