

Tema 5. Distribuciones continuas

En este tema, se presentan las principales distribuciones continuas y sus características. Se abordarán las distribuciones continuas más relevantes, como la **distribución normal**, la **distribución exponencial**, la **distribución gamma** y la **distribución uniforme**, entre otras, destacando sus propiedades y aplicaciones prácticas. La Tabla 2.2 resume la función de densidad, la esperanza matemática, la varianza y la función característica de las distribuciones presentadas en este tema. A través de ejemplos, se facilitará la comprensión de las diferencias entre estas distribuciones, así como la identificación de la más adecuada para resolver problemas en diversos contextos.

Tabla 2.2. Distribuciones continuas

Distribución	Función de densidad	Esperanza matemática	Varianza	Función característica
Normal $X \sim N(\mu, \sigma)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $\forall x \in \mathbb{R}$	μ	σ^2	$e^{it\mu - \frac{\sigma^2 t^2}{2}}$
Uniforme $X \sim U(a, b)$	$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{en otro caso} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{itb} - e^{ita}}{it(b-a)}$
Exponencial $X \sim \text{Exp}(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{1}{1 - \frac{1}{\lambda}it}$
Gamma $X \sim \gamma(\alpha, \lambda)$	$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ $x > 0$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	$\left(\frac{1}{1 - \frac{1}{\lambda}it} \right)^\alpha$
Beta $X \sim \beta(\alpha, \beta)$	$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$ $x \in (0, 1)$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	no existe en forma cerrada

Fuente: elaboración propia.

5.1. Distribución normal

Distribución normal

La variable aleatoria normal es la más importante de las distribuciones continuas. La *distribución normal*, también conocida como *distribución de Gauss*, es una distribución de probabilidad continua que se caracteriza por su forma de campana. Aunque *Abraham de Moivre* (1667-1754) fue el primero en reconocer la distribución normal, fue *Carl Friedrich Gauss* (1777-1855) quien realizó desarrollos más profundos y formuló la ecuación que describe esta curva.

Se dice que una variable aleatoria absolutamente continua X sigue una *distribución normal* de parámetros μ y σ , y se denota como $X \sim N(\mu, \sigma)$, si su función de densidad viene dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}$$

La función de densidad de una variable aleatoria X con distribución normal es una curva simétrica cuya forma se asemeja a una campana, comúnmente conocida como *campana de Gauss*. Cuando X toma valores inferiores a la media, la función de densidad $f(x)$ es creciente; y cuando X toma valores superiores a la media, $f(x)$ es decreciente (ver Figuras 2.9 y 2.10).

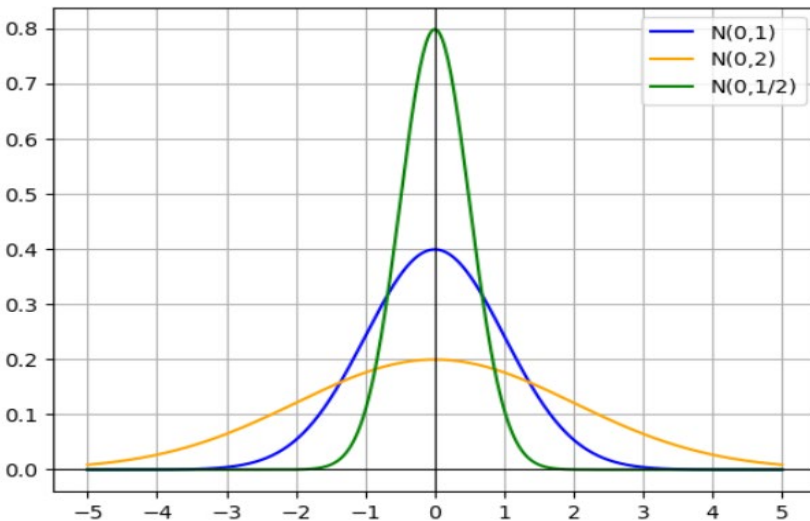


Figura 2.9. Ejemplos de la distribución normal.

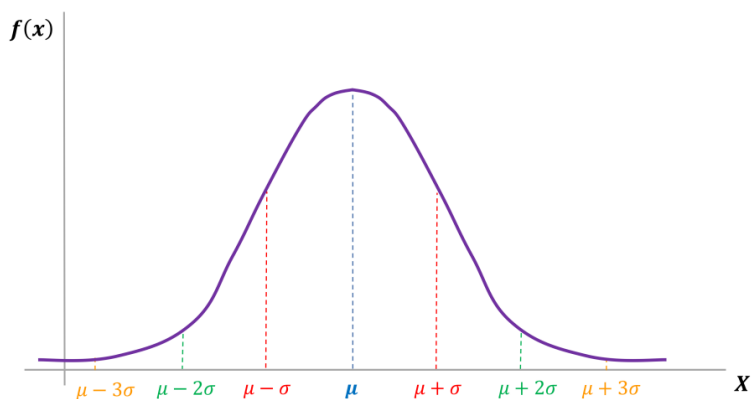


Figura 2.10. Función de densidad de la distribución normal $X \sim N(\mu, \sigma)$.

La simetría de la función de densidad implica que está centrada en su media. En una distribución normal, la media, la mediana y la moda coinciden. La dispersión de los datos se mide a través de la desviación estándar, que define la amplitud de la curva (no se debe confundir dispersión con apuntamiento, el apuntamiento de las variables aleatorias normales siempre es constante y la dispersión es un parámetro que varía de unas a otras).

La *esperanza matemática* de una variable aleatoria X con distribución normal es:

$$E[X] = \mu$$

La *varianza* de una variable aleatoria X con distribución normal es σ^2 .

La *función característica* de una variable aleatoria X con distribución normal de parámetros μ y σ , es:

$$\varphi_X(t) = E[e^{itX}] = e^{it\mu - \frac{\sigma^2 t^2}{2}} \quad \forall t \in \mathbb{R}$$

El parámetro μ es el valor con la mayor densidad de probabilidad, indica su esperanza, y el parámetro σ representa su dispersión.

Teorema 2.4. La distribución normal es reproductiva respecto a ambos parámetros. Es decir, sean X_1, \dots, X_m variables aleatorias independientes con distribución normal, tal que $X_j \sim N(\mu_j, \sigma_j)$ con $j = 1, \dots, m$. Entonces, $S_m = \sum_{j=1}^m X_j$ es una variable aleatoria con distribución normal tal que:

$$S_m \sim N\left(\sum_{j=1}^m \mu_j, \sqrt{\sum_{j=1}^m \sigma_j^2}\right)$$

Demostración 2.4

Para su demostración se utiliza la función característica de una variable aleatoria con distribución normal.

$$X_j \sim N(\mu_j, \sigma_j) \Rightarrow \varphi_{X_j}(t) = E[e^{itX_j}] = e^{it\mu_j - \frac{\sigma_j^2 t^2}{2}} \quad \forall t \in \mathbb{R}$$

La función característica de la nueva variable, S_m , es:

$$\varphi_{S_m}(t) = E[e^{itS_m}] = E[e^{it\sum_{j=1}^m X_j}] = E[e^{itX_1} e^{itX_2} \dots e^{itX_m}]$$

como X_1, \dots, X_m son variables aleatorias independientes,

$$E[e^{itX_1} e^{itX_2} \dots e^{itX_m}] = E[e^{itX_1}] E[e^{itX_2}] \dots E[e^{itX_m}]$$

$$\begin{aligned} &= \prod_{j=1}^m \left(e^{it\mu_j - \frac{\sigma_j^2 t^2}{2}} \right) = e^{\sum_{j=1}^m it\mu_j - \frac{\sigma_j^2 t^2}{2}} \\ &= e^{it\sum_{j=1}^m \mu_j - \frac{t^2 \sum_{j=1}^m \sigma_j^2}{2}} \end{aligned}$$

Entonces,

$$S_m = \sum_{j=1}^m X_j \sim N\left(\sum_{j=1}^m \mu_j, \sqrt{\sum_{j=1}^m \sigma_j^2}\right)$$

Como consecuencia del teorema anterior, si $\bar{X}_m = \frac{1}{m} \sum_{j=1}^m X_j$, donde X_1, \dots, X_m son v.a.i.i.d con distribución normal ($X_j \sim N(\mu, \sigma)$ para $j = 1, \dots, m$), entonces,

$$S_m \sim N\left(\mu, \frac{\sigma}{\sqrt{m}}\right)$$

Algunas propiedades de la distribución normal

1. La distribución normal es *simétrica* alrededor de su media (μ). Es decir, la probabilidad de que una variable tome un valor mayor o menor que la media es la misma, $P(X \leq \mu - c) = P(X \geq \mu + c)$.

2. En la distribución normal, la moda y la mediana son iguales a la media (μ) y se encuentran en el centro de la curva.
3. *Regla Empírica (regla 68-95-99.7)*: En una distribución normal:
 - Aproximadamente el 68% de los valores cae dentro del intervalo $[\mu - \sigma, \mu + \sigma]$.
 - Aproximadamente el 95% de los valores cae dentro del intervalo $[\mu - 2\sigma, \mu + 2\sigma]$.
 - Aproximadamente el 99.7% de los valores cae dentro del intervalo $[\mu - 3\sigma, \mu + 3\sigma]$.
4. Todos los momentos respecto al origen de orden impar de una variable aleatoria con distribución normal estándar, $N(0, 1)$, son cero.
5. Sea X una variable aleatoria con distribución normal $X \sim N(\mu, \sigma)$. Sea Y una transformación lineal de la variable aleatoria X dada por $Y = aX + b$ con $a, b \in \mathbb{R}$. Entonces, la variable aleatoria Y también se distribuye según una Normal de parámetros $a\mu + b$ y $a\sigma$, tal que $Y \sim N(a\mu + b, a\sigma)$.

La demostración se realiza mediante la función característica:

$$\begin{aligned} \varphi_Y(t) &= E[e^{itY}] = E[e^{it(aX+b)}] = e^{itb} E[e^{itaX}] = e^{itb} e^{ita\mu - \frac{\sigma^2 t^2 a^2}{2}} \\ &= e^{it(b+a\mu) - \frac{\sigma^2 t^2 a^2}{2}} \Rightarrow Y \sim N(a\mu + b, a\sigma) \quad \forall t \in \mathbb{R} \end{aligned}$$

Distribución normal estándar

Un caso particular de la distribución normal es la *normal estándar*. Se dice que una variable aleatoria absolutamente continua X sigue una *distribución normal estándar*, y se denota como $X \sim N(0, 1)$, si su función de densidad viene dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \forall x \in \mathbb{R}$$

Como se muestra en la Figura 2.11, la distribución $N(0, 1)$ se caracteriza por centrarse en el cero, es decir $E[X] = \mu = 0$ y tener una desviación típica de 1, tal que $\sigma = 1$. Prácticamente toda la probabilidad de la variable aleatoria $X \sim N(0, 1)$ se concentra en el intervalo $[-3, 3]$ ($F(3) = 0.9987$ y $F(-3) = 0.0013$).

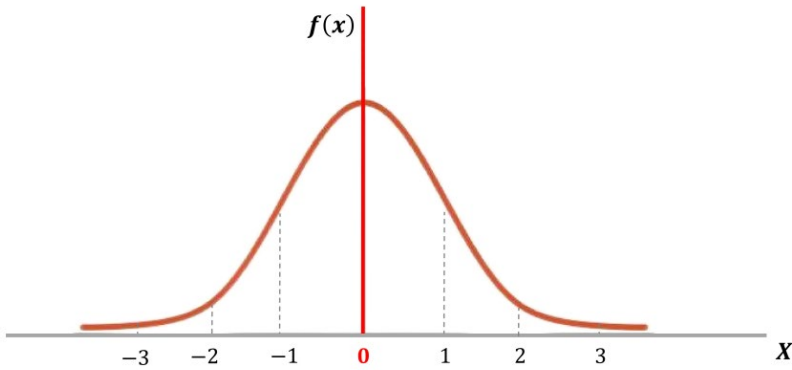


Figura 2.11. Distribución normal estándar $X \sim N(0, 1)$.

Tipificación: cualquier distribución normal puede ser transformada en una distribución normal estándar (con media 0 y desviación estándar 1). Habitualmente se denomina Z a la variable aleatoria tipificada cuya distribución es una normal estándar. La transformación de una distribución $X \sim N(\mu, \sigma)$ en una $N(0, 1)$ se llama tipificación de la variable X y viene dada por la siguiente expresión:

$$Z = \frac{X - \mu}{\sigma}$$

Una de las utilidades que tiene el tipificar la variable aleatoria X es el cálculo de la probabilidad en cualquier intervalo.

La función de distribución de una variable aleatoria $N(\mu, \sigma)$ viene dada por la expresión:

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Esta integral es una función impropia y por lo tanto no existe una expresión algebraica.

Para calcular el valor de la función de distribución en cualquier punto x , es necesario realizar la integral utilizando algún algoritmo de cálculo numérico. Como cualquier variable aleatoria $N(\mu, \sigma)$ se puede transformar en una $N(0, 1)$. Entonces,

$$F_X(x) = P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(Z \leq \frac{x - \mu}{\sigma}\right)$$

Ahora, ya solo es necesario la función de distribución de la $N(0, 1)$, y estos valores se pueden encontrar tabulados o calcularse mediante un algoritmo.

Para determinar la probabilidad en un intervalo, el proceso de tipificación es tal que:

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right) = P\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{b - \mu}{\sigma}\right) - P\left(Z \leq \frac{a - \mu}{\sigma}\right) \end{aligned}$$

Ejemplo 5.1

La capacidad del ascensor de la facultad está limitada a un peso máximo de 320 kilogramos para sus cuatro ocupantes. Dado que el peso de una persona sigue una distribución normal $N(75, 8)$, se desea calcular la probabilidad de que el peso total de cuatro personas supere los 320 kilogramos.

Si el peso de una persona sigue una distribución normal $N(75, 8)$, la muestra de 4 personas sigue una distribución normal $N(300, \sqrt{256}) \equiv N(300, 16)$.

Entonces, la probabilidad de que el peso total de cuatro personas supere los 320 kilogramos se calcula como:

$$\begin{aligned} P(X_1 + X_2 + X_3 + X_4 > 320) &= P(Y > 320) = P\left(\frac{Y - 300}{16} > \frac{320 - 300}{16}\right) \\ &= P(Z > 1.25) = 1 - P(Z \leq 1.25) \\ &= 1 - 0.8944 = 0.1056 \end{aligned}$$

5.2. Relaciones entre algunas distribuciones discretas y la distribución normal

En la práctica, para simplificar los cálculos, algunas distribuciones discretas pueden aproximarse a una distribución normal, porque por el teorema central del límite, la suma de variables aleatorias independientes y con la misma distribución tiende a una distribución normal a medida que aumenta el tamaño de la muestra.

- Sea X una variable aleatoria con distribución binomial, $X \sim B(n, p)$. Cuando el número de experimentos n es grande ($n \geq 30$) y la probabilidad de éxito p no es demasiado cercana a 0 o 1 ($0.1 < p < 0.9$), se puede aproximar la distribución binomial a la distribución normal de la siguiente manera:

$$X \sim B(n, p) \approx N(np, \sqrt{npq}) \text{ siendo } n \geq 30 \text{ y } 0.1 < p < 0.9$$

- Sea X una variable aleatoria con distribución de Poisson, $X \sim P(\lambda)$. Cuando $\lambda > 10$, se puede aproximar la distribución de Poisson a la distribución normal de la siguiente manera:

$$X \sim P(\lambda) \approx N(\lambda, \sqrt{\lambda}) \text{ siendo } \lambda > 10$$

Debemos de tener en cuenta que cuando se aproxima una variable aleatoria discreta a una variable aleatoria continua hay que aplicar el *coeficiente de corrección por continuidad*. Es decir, cuando un punto de masa x se toma como un valor continuo, x es equivalente al intervalo $(x - \frac{1}{2}, x + \frac{1}{2})$. Entonces, al calcular una probabilidad específica de una variable aleatoria discreta, aproximando su distribución a una distribución continua, se tiene que:

$$P(a \leq X \leq b) \approx P\left(a - \frac{1}{2} \leq X \leq b + \frac{1}{2}\right)$$

$$P(a < X < b) \approx P\left(a + \frac{1}{2} \leq X \leq b - \frac{1}{2}\right)$$

Por tanto,

$$P(X \leq b) \approx P\left(X \leq b + \frac{1}{2}\right)$$

$$P(X < b) \approx P\left(X \leq b - \frac{1}{2}\right)$$

$$P(X \geq a) \approx P\left(X \geq a - \frac{1}{2}\right)$$

$$P(X > a) \approx P\left(X \geq a + \frac{1}{2}\right)$$

Ejemplo 5.2

Una evaluación continua de una asignatura consiste en un test de 40 preguntas, que deben ser respondidas con “verdadero” o “falso”. Se aprueba si se contestan correctamente al menos 20 preguntas. Un alumno responde al cuestionario de la evaluación continua lanzando una moneda al aire y contestando “verdadero” si sale cara y “falso” si sale cruz. Se pide:

1. Calcular la probabilidad de aprobar esta evaluación.
2. Determinar la probabilidad de acertar más de 23 preguntas y menos de 30.

1. La variable aleatoria X que representa el número de preguntas aciertas es una variable aleatoria con distribución binomial de parámetros $n = 40$, y $p = \frac{1}{2}$, tal que $X \sim B(40, \frac{1}{2})$.

Como $n = 40$ y $p = \frac{1}{2}$, cumplen los requisitos para realizar la aproximación de la distribución binomial a la distribución normal. Entonces, $X \sim B(40, \frac{1}{2}) \approx N(20, \sqrt{10})$.

La probabilidad de aprobar esta evaluación se calcula como:

$$\begin{aligned} P(X \geq 20) &\approx P(X \geq 19.5) = P\left(Z \geq \frac{19.5 - 20}{\sqrt{10}}\right) = P(Z \geq -0.16) \\ &= 1 - P(Z < -0.16) = 1 - 0.4364 = 0.5636 \end{aligned}$$

2. La probabilidad de acertar más de 23 preguntas y menos de 30 se calcula como:

$$\begin{aligned} P(23 < X < 30) &\approx P(23.5 \leq X \leq 29.5) = P\left(\frac{23.5 - 20}{\sqrt{10}} \leq Z \leq \frac{29.5 - 20}{\sqrt{10}}\right) \\ &= P(1.11 \leq Z \leq 3) = P(Z \leq 3) - P(Z \leq 1.11) \\ &= 0.9987 - 0.8665 = 0.1322 \end{aligned}$$

5.3. Distribución uniforme

La *distribución uniforme* es una distribución de probabilidad continua que asigna la misma probabilidad a todos los intervalos de igual longitud dentro de un intervalo definido. A diferencia de la distribución uniforme discreta (ver la sección 4.8), que solo asigna probabilidades a un conjunto finito de valores específicos, la distribución uniforme permite que cualquier intervalo dentro del intervalo de definición sea posible, lo que implica una función de densidad de probabilidad constante en todo el rango.

Se dice que una variable aleatoria absolutamente continua X sigue una *distribución uniforme* de parámetros a y b , y se denota como $X \sim U(a, b)$, cuando la probabilidad de que la variable aleatoria X pertenezca a cualquier intervalo (x, y) incluido en (a, b) es la misma si tienen igual longitud. De hecho, por ser X una variable aleatoria continua, el intervalo sobre el que se define X puede ser un intervalo abierto, semiabierto o cerrado, es decir:

$$X \sim U(a, b) = U[a, b] = U(a, b] = U[a, b)$$

La *función de densidad* de la variable aleatoria X con distribución uniforme es:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{en otro caso} \end{cases}$$

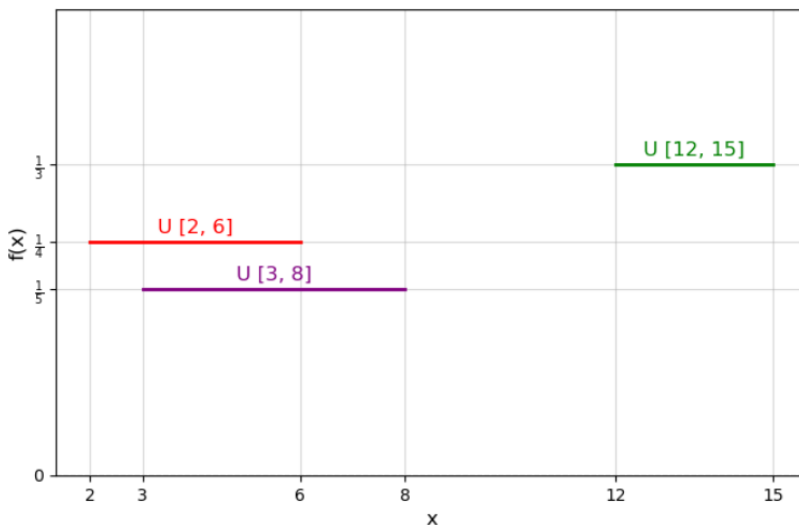


Figura 2.12. Ejemplos de la distribución uniforme.

A través de la función de densidad de la variable aleatoria X , se obtiene la *función de distribución* de la variable aleatoria $X \sim U(a, b)$ como:

$$F_X(x) = \int_a^x \frac{1}{b-a} dt = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

La *esperanza matemática* de una variable aleatoria X con distribución uniforme es:

$$E[X] = \frac{a+b}{2}$$

La *varianza* de una variable aleatoria X con distribución uniforme es:

$$V(X) = \frac{(b-a)^2}{12}$$

La *función característica* de una variable aleatoria X con distribución uniforme de parámetros a y b , es:

$$\varphi_X(t) = E[e^{itX}] = \frac{e^{itb} - e^{ita}}{it(b-a)} \quad \forall t \in \mathbb{R}$$

Teorema 2.5. Sea X una variable aleatoria absolutamente con función de distribución $F_X(x)$. Sea $Y = F_X(x)$. Entonces, $Y \sim U(0, 1)$.

Demostración 2.5

La variable aleatoria $Y = F_X(x)$ tomará valores en $(0, 1)$. Entonces,

$$P(Y \leq y) = P(F_X(x) \leq y) = P(X \leq F_X(y)^{-1}) = F_X(F_X(y)^{-1}) = y \text{ con } y \in (0, 1)$$

Ejemplo 5.3

El consumo familiar de manzanas sigue una distribución uniforme, con una media de 10 kilogramos y una varianza de 3 kilogramos al cuadrado. Se desea calcular la probabilidad de que dicho consumo oscile entre 5 y 10 kilogramos.

La variable aleatoria X que representa el peso de las manzanas consumidas es una variable aleatoria con distribución uniforme en el intervalo (a, b) .

$$\text{Sabido que } E[X] = \frac{a+b}{2} = 10 \text{ y } V(X) = \frac{(b-a)^2}{12} = 3$$

Entonces,

$$\begin{cases} a + b = 20 \\ (b - a)^2 = 36 \end{cases} \mapsto [b - (20 - b)]^2 = 36 \mapsto [2b - 20]^2 = 36 \\ \Rightarrow b = 7 \text{ ó } b = 13$$

Considerando que, para la distribución uniforme $X \sim U(a, b)$, donde a es el límite inferior y b es el límite superior, ha de cumplirse $a < b$. Por eso, $b = 13 \Rightarrow a = 7$. Por tanto, $X \sim U(7, 13)$, y la función de distribución de esta variable es:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 7 \\ \frac{x-7}{13-7} = \frac{x-7}{6} & \text{si } 7 \leq x \leq 13 \\ 1 & \text{si } x > 13 \end{cases}$$

La probabilidad de que el consumo de manzanas oscile entre 5 y 10 kilogramos se calcula como:

$$\begin{aligned} P(5 \leq X \leq 10) &= 0 + P(7 \leq X \leq 10) = P(X \leq 10) - P(X \leq 7) \\ &= F_X(10) - F_X(7) = \frac{10-7}{6} - 0 = \frac{3}{6} = 0.5 \end{aligned}$$

5.4. Distribución exponencial

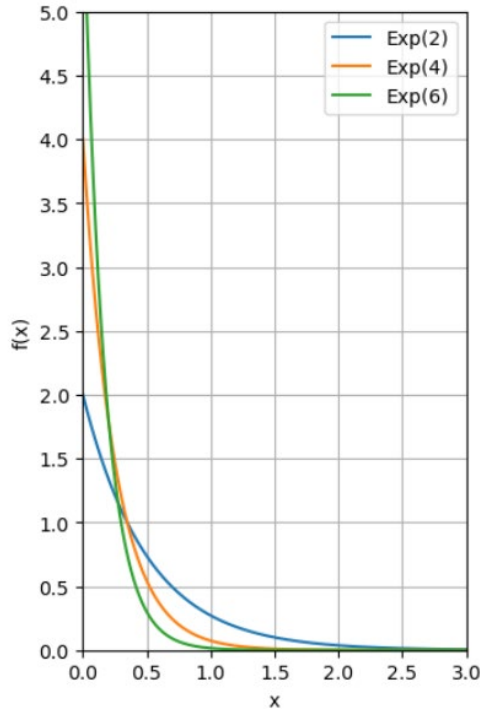


Figura 2.13. Ejemplos de la distribución exponencial.

La *distribución exponencial* es una distribución de probabilidad continua que se puede describir como el tiempo transcurrido entre dos sucesos consecutivos de Poisson. Su función de densidad decae exponencialmente a medida que el tiempo avanza. Además, está completamente definida por su parámetro λ , que representa el número de sucesos de Poisson por unidad de tiempo.

Se dice que una variable aleatoria absolutamente continua X sigue una *distribución exponencial* de parámetro λ , y se denota como $X \sim \text{Exp}(\lambda)$, si la variable aleatoria X describe el tiempo transcurrido entre dos sucesos consecutivos de Poisson, o el tiempo de espera hasta que ocurre el primer suceso de Poisson.

Sea N_x el número de sucesos de Poisson hasta el momento x , esta variable aleatoria es una Poisson de parámetro λx ($N_x \sim P(\lambda x)$). La probabilidad de que el tiempo de espera hasta que ocurre el primer suceso de Poisson sea mayor que x ($P(X > x)$) es tal que:

$$P(X > x) = P(N_x < 1) = P(N_x = 0) = e^{-\lambda x} \quad \forall x > 0$$

Por tanto, la *función de distribución* de la variable aleatoria $X \sim \text{Exp}(\lambda)$ es:

$$F_X(x) = P(X \leq x) = 1 - P(X > x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

La *función de densidad* de la variable aleatoria X con distribución exponencial es:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

La *esperanza matemática* de una variable aleatoria X con distribución exponencial es:

$$E[X] = \frac{1}{\lambda}$$

La *varianza* de una variable aleatoria X con distribución exponencial es:

$$V(X) = \frac{1}{\lambda^2}$$

La *función característica* de una variable aleatoria X con distribución exponencial de parámetro λ , es:

$$\varphi_X(t) = E[e^{itX}] = \frac{1}{1 - \frac{1}{\lambda}it} \quad \forall t \in \mathbb{R}$$

Una propiedad muy importante de la distribución exponencial es la *propiedad de la pérdida de memoria*, que se refleja mediante la siguiente expresión:

$$P(X \geq x + h | X \geq x) = \frac{P(X \geq x + h)}{P(X \geq x)} = \frac{e^{-\lambda(x+h)}}{e^{-\lambda x}} = e^{-\lambda h} = P(X \geq h)$$

Ejemplo 5.4

El tiempo de revisión del ordenador del aula informática de la Facultad de Estudios Estadísticos sigue una distribución exponencial con una media de 20 minutos. Se pide:

1. *Calcular la probabilidad de que el tiempo de revisión sea menor de 10 minutos.*
 2. *El costo de la revisión es de 25 euros por cada media hora o fracción. ¿Cuál es la probabilidad de que una revisión cueste 50 euros?*
1. Sabiendo que $E[X] = \frac{1}{\lambda} = 20 \Rightarrow \lambda = \frac{1}{20} \Rightarrow X \sim \text{Exp}\left(\frac{1}{20}\right)$

La función de distribución de la variable aleatoria $X \sim \text{Exp}(\frac{1}{20})$ es:

$$F_X(x) = \begin{cases} 1 - e^{-\frac{1}{20}x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

La probabilidad de que el tiempo de revisión sea menor de 10 minutos se calcula como:

$$P(X < 10) = F_X(10) = 1 - e^{-\frac{1}{20} \cdot 10} = 1 - e^{-\frac{1}{2}} = 1 - 0.6065 = 0.3935$$

2. Dado que el costo de la revisión es de 25 euros por cada media hora o fracción, un costo de 50 euros corresponde a un tiempo de revisión de 1 hora (ser inferior o igual a 60 minutos). Entonces, la probabilidad de que una revisión cueste 50 euros es:

$$\begin{aligned} P(30 < X \leq 60) &= P(X \leq 60) - P(X < 30) = F_X(60) - F_X(30) \\ &= \left(1 - e^{-\frac{1}{20} \cdot 60}\right) - \left(1 - e^{-\frac{1}{20} \cdot 30}\right) \\ &= e^{-\frac{30}{20}} - e^{-\frac{60}{20}} = 0.2231 - 0.0498 = 0.1733 \end{aligned}$$

5.5. Distribución gamma

La *distribución Erlang* es una distribución de probabilidad continua que describe el tiempo transcurrido hasta el α -ésimo suceso de Poisson ($\alpha \in \mathbb{N}$).

Se dice que una variable aleatoria absolutamente continua X sigue una *distribución Erlang* de parámetros α y λ , con $\alpha > 0$ y $\lambda > 0$, y se denota como $X \sim \text{Erlang}(\alpha, \lambda)$, si la variable aleatoria X describe el tiempo transcurrido hasta el α -ésimo suceso de Poisson. El parámetro λ representa el número medio de sucesos de Poisson por unidad de tiempo.

La *función de densidad* de la variable aleatoria X con distribución Erlang es:

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \quad x > 0$$

La *distribución gamma* es una generalización de la distribución Erlang con $\alpha \in \mathbb{R}$. La función $\Gamma(\alpha)$ se denomina *función gamma de Euler* y se define como la integral impropia para todo $\alpha > 0$, dada por:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

Entre otras, esta función cumple las siguientes propiedades:

- $\Gamma(1) = 1$
- Si $\alpha > 1 \Rightarrow \Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
- Si $\alpha \in \mathbb{N} \Rightarrow \Gamma(\alpha) = (\alpha - 1)!$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

La esperanza matemática de una variable aleatoria X con distribución gamma es:

$$E[X] = \frac{\alpha}{\lambda}$$

La *varianza* de una variable aleatoria X con distribución gamma es:

$$V(X) = \frac{\alpha}{\lambda^2}$$

La *función característica* de una variable aleatoria X con distribución gamma es:

$$\varphi_X(t) = E[e^{itX}] = \left(\frac{1}{1 - \frac{1}{\lambda}it} \right)^\alpha \quad \forall t \in \mathbb{R}$$

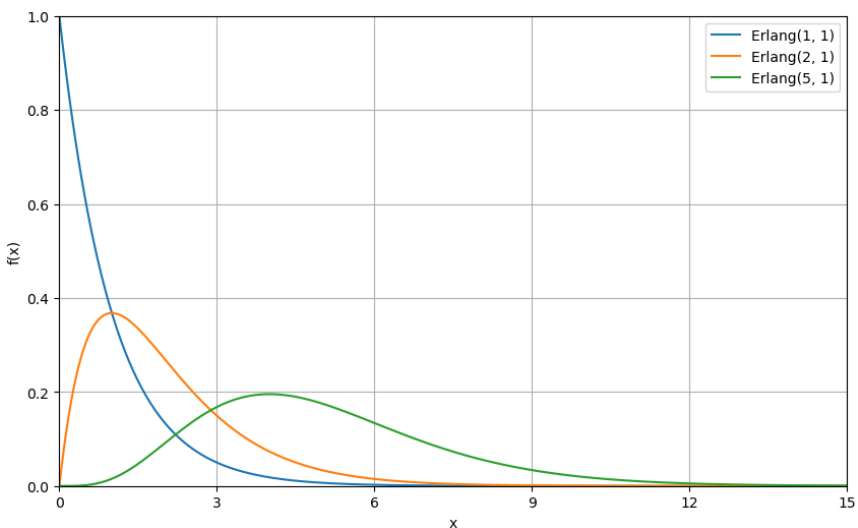


Figura 2.14. Ejemplos de la distribución Erlang.

Relaciones entre las distribuciones exponencial y gamma

La distribución exponencial $X \sim \text{Exp}(\lambda)$ es un caso particular de la distribución gamma con parámetro $\alpha = 1$, describe el tiempo hasta la *primera* ocurrencia de un evento, tiene una aplicación importante en situaciones donde se aplica el proceso de Poisson. Es decir, $X \sim \text{Exp}(\lambda) = \gamma(1, \lambda)$. La suma de n variables aleatorias independientes e idénticamente distribuidas de distribución exponencial sigue una distribución Erlang, tales que $X_i \sim \text{Exp}(\lambda)$ es una variable aleatoria *Erlang* (n, λ) .

Ejemplo 5.5

Se supone que el tiempo, en horas, empleado diariamente en transporte público para ir a trabajar en Madrid, sigue una distribución gamma con parámetros $\alpha = 2$ y $\lambda = 2$. Se pide:

1. Calcular el tiempo medio que tarda una persona en transporte público.
2. Calcular la probabilidad de que una persona emplee más de media hora en transporte público.

1. El tiempo medio que tarda una persona en transporte público se calcula como:

$$E[X] = \frac{2}{2} = 1 \text{ hora}$$

2. Como $X \sim \gamma(2, 2)$, su función de densidad es:

$$f(x) = \frac{2^2}{\Gamma(2)} x^{2-1} e^{-2x} = 4x e^{-2x} \quad x > 0$$

Entonces, $F_X(x) = 4 \int_0^x x e^{-2x} dx = 1 - e^{-2x}(1 + 2x)$

La probabilidad de que una persona emplee más de media hora en transporte público se calcula como:

$$\begin{aligned} P\left(X > \frac{1}{2}\right) &= 1 - P\left(X \leq \frac{1}{2}\right) = 1 - F_X\left(\frac{1}{2}\right) \\ &= 1 - \left(1 - e^{-2 \cdot \frac{1}{2}} \left(1 + 2 \cdot \frac{1}{2}\right)\right) \\ &= 2e^{-1} \approx 0.7358 \end{aligned}$$

5.6. Distribución beta

La *distribución beta* es una distribución de probabilidad continua que se utiliza para modelar variables aleatorias acotadas entre 0 y 1. Como ejemplo de distribución beta es el cálculo de la fracción de tiempo que un servidor está ocupado, o la fiabilidad de un sistema. Su forma está determinada por dos parámetros, α y β , que aparecen como exponentes de la variable aleatoria y controlan la forma de la distribución.

Se dice que una variable aleatoria absolutamente continua X sigue una *distribución beta* de parámetros α y β , con $\alpha > 0$ y $\beta > 0$, y se denota como $X \sim \beta(\alpha, \beta)$, si la *función de densidad* de la variable aleatoria X viene dada por:

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad x \in (0, 1)$$

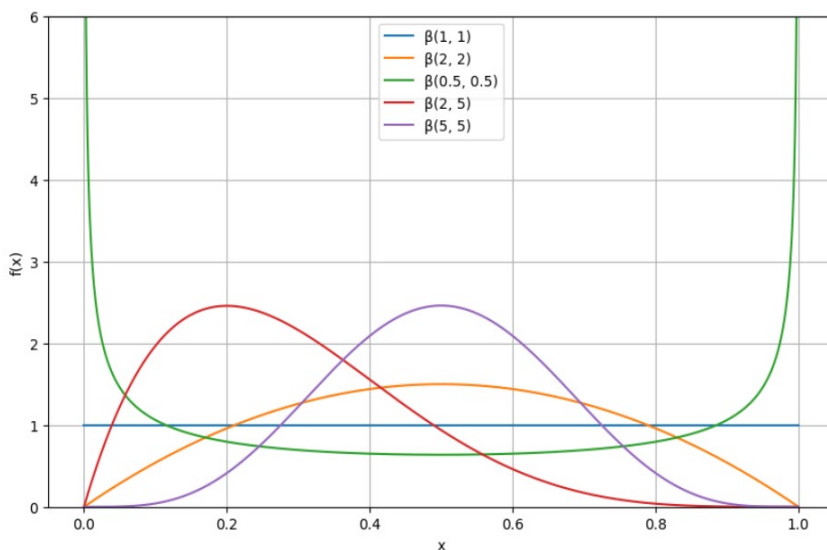


Figura 2.15. Ejemplos de la distribución beta.

La *esperanza matemática* de una variable aleatoria X con distribución beta es:

$$E[X] = \frac{\alpha}{\alpha + \beta}$$

La *varianza* de una variable aleatoria X con distribución beta es:

$$V(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

Como casos particulares de la distribución beta se tiene que:

- Si $\alpha = 1$ y $\beta = 1 \Rightarrow f(x) = 1$ si $0 < x < 1 \Rightarrow X \sim U[0,1]$
- Sea X una variable aleatoria con distribución beta, tal que $X \sim \beta(\alpha, \beta)$. Entonces, $1 - X \sim \beta(\beta, \alpha)$.

Ejemplo 5.6

El almacén de la cafetería de la Facultad de Estudios Estadísticos llena su stock de patatas cada lunes. Se ha observado que la cantidad de patatas que necesitan cada semana para los platos del menú se puede modelar con una distribución beta, con parámetros $\alpha = 4$ y $\beta = 2$. Se pide:

1. Calcular la media de la cantidad de patatas que se espera necesitar cada semana.
 2. Calcular la probabilidad de que la cantidad de patatas que se necesite la próxima semana sea de al menos el 90% de la capacidad del almacén.
1. La media de la cantidad de patatas que se espera necesitar cada semana se calcula como:

$$E[X] = \frac{4}{4 + 2} = \frac{2}{3}$$

2. Como $X \sim \beta(4, 2)$, su función de densidad es:

$$f(x) = \frac{\Gamma(4 + 2)}{\Gamma(4)\Gamma(2)} x^{4-1}(1-x)^{2-1} = 20x^3(1-x) \quad x \in (0, 1)$$

La probabilidad de que la cantidad de patatas que se necesite la próxima semana sea de al menos el 90% de la capacidad del almacén se calcula como:

$$P(X > 0.9) = 20 \int_{0.9}^1 x^3 (1-x) dx = 0.081$$

5.7. Ejercicios

Ejercicios resueltos

Ejercicio R. 5.1

La cantidad de café que una persona consume semanalmente (en centilitros) sigue una distribución uniforme entre 0 y 100. Se desea calcular la probabilidad de que el consumo de café este mes esté entre 20 y 35 centilitros. Además, se deben determinar la esperanza matemática y la varianza de dicha distribución.

Solución:

Sabiendo que la distribución es uniforme entre 0 y 100, los parámetros de la distribución son $a = 0$ y $b = 100$.

La función de densidad para esta variable aleatoria es:

$$f(x) = \begin{cases} \frac{1}{100} & \text{si } x \in [0, 100] \\ 0 & \text{en otro caso} \end{cases}$$

La función de distribución es:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x-0}{100-0} = \frac{x}{100} & \text{si } 0 \leq x \leq 100 \\ 1 & \text{si } x > 100 \end{cases}$$

La probabilidad de que el consumo de café esté entre 20 y 35 centilitros se calcula como:

$$\begin{aligned} P(20 \leq X \leq 35) &= P(X \leq 35) - P(X \leq 20) \\ &= F_X(35) - F_X(20) \\ &= \frac{35}{100} - \frac{20}{100} = \frac{15}{100} = 0.15 \end{aligned}$$

La esperanza matemática es $E[X] = \frac{a+b}{2} = \frac{0+100}{2} = 50$

La varianza es $V(X) = \frac{(b-a)^2}{12} = \frac{(100-0)^2}{12} = 833.33$

Ejercicio R. 5.2

La duración en minutos de un examen final sigue una distribución exponencial con una media de 100 minutos. ¿Cuál es la probabilidad de que el examen final dure menos de 90 minutos? Si 100 alumnos realizan el examen final, ¿cuál es la probabilidad de que, tras 200 minutos, al menos 5 alumnos estén haciendo el examen?

Solución:

Sabiendo que, $E[X] = \frac{1}{\lambda} = 100 \Rightarrow \lambda = \frac{1}{100} \Rightarrow X \sim \text{Exp} \left(\frac{1}{100} \right)$

La función de distribución de la variable aleatoria $X \sim \text{Exp} \left(\frac{1}{100} \right)$ es:

$$F_X(x) = \begin{cases} 1 - e^{-\frac{1}{100}x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

La probabilidad de que el examen final dure menos de 90 minutos se calcula como:

$$P(X < 90) = F_X(90) = 1 - e^{-\frac{1}{100} \cdot 90} = 1 - e^{-\frac{9}{10}} = 1 - 0.4066 = 0.5934$$

La *probabilidad* de que un alumno aún esté haciendo el examen tras 200 minutos es:

$$\begin{aligned} P(X \geq 200) &= 1 - P(X < 200) = 1 - F_X(200) = \\ &= 1 - (1 - e^{-\frac{1}{100} \cdot 200}) = e^{-2} = 0.1353 \end{aligned}$$

Como nos preguntan por la probabilidad de que al menos 5 alumnos estén realizando el examen, tenemos que utilizar una variable aleatoria discreta, en concreto, una binomial con $n = 100$, y $p = 0.1353$, es decir, $X \sim B(100, 0.1353)$.

En este caso, aunque n es grande, $p = 0.1353 > 0.1$, por lo que no es posible aproximar la distribución binomial a la de Poisson. Lo que se puede hacer es utilizar la aproximación de la distribución binomial a la normal, porque n es grande, $p = 0.1353 \in (0.1, 0.9)$.

$$\Rightarrow X \sim B(100, 0.1353) \approx N(13.53, \sqrt{11.70})$$

La probabilidad de interés es:

$$\begin{aligned} P(X \geq 5) &\approx 1 - P(X \leq 4.5) \\ &= 1 - P\left(Z \leq \frac{4.5 - 13.53}{\sqrt{11.70}}\right) = 1 - P(Z \leq -2.64) \\ &= 1 - 0.0041 = 0.9959 \end{aligned}$$

Ejercicio R. 5.3

El diámetro de una pila debe ser de 50 cm, sin embargo, debido a un error que comete la máquina, este sigue una distribución $N(50, 0.5)$. Calcular:

1. Probabilidad de que al inspeccionar 5 pilas, 3 tengan un diámetro de más de 50.4 cm (la elección es **con reemplazamiento**).
2. Probabilidad de que la medida del diámetro de una pila cualquiera difiera de la media menos de 2 veces la desviación típica.
3. Probabilidad de tener que elegir al menos 5 pilas para que una de ellas tenga un diámetro menor de 49 cm.

Solución:

1. Probabilidad de que al inspeccionar 5 pilas, 3 tengan un diámetro de más de 50.4 cm. (la elección es con reemplazamiento). Lo primero que se debe hacer es definir la variable aleatoria $X \equiv \{\text{medida del diámetro de la pila}\}$ y calcular la probabilidad de que el diámetro sea de más de 50.4 cm.

$$\begin{aligned} P(X > 50.4) &= P\left(Z > \frac{50.4 - 50}{0.5}\right) = 1 - P(Z \leq 0.8) = \\ &= 1 - 0.7881 = 0.2119 \end{aligned}$$

Ahora se define la variable aleatoria $Y \equiv \{\text{número de pilas defectuosas}\}$ y se calcula la probabilidad pedida.

$$Y \sim B(5, 0.2119) \rightarrow P(Y = 3) = \binom{5}{3} 0.2119^3 \cdot (1 - 0.2119)^2 = 0.0591$$

2. Probabilidad de que la medida del diámetro de una pila cualquiera difiera de la media menos de 2 veces la desviación típica.

$$\begin{aligned} P(X \in (50 - 1; 50 + 1)) &= P(X \in (49, 51)) = P\left(\frac{49-50}{0.5} \leq Z \leq \frac{51-50}{0.5}\right) = \\ &= P(Z \leq 2) - P(Z \leq -2) = 0.9772 - (1 - 0.9772) = 0.9544 \end{aligned}$$

3. Probabilidad de tener que elegir al menos 5 pilas para que una de ellas tenga un diámetro menor de 49 cm. Para la resolución de este apartado es necesario calcular $P(X < 49)$ y definir la variable aleatoria $Z \equiv \{\text{número de extracciones hasta obtener el primer éxito}\}$ (éxito es que el diámetro de la pila sea menor de 49 cm.). Entonces,

$$\begin{aligned} P(X < 49) &= P\left(Z < \frac{49 - 50}{0.5}\right) \\ &= P(Z \leq -2) = 1 - 0.9772 = 0.0228 \end{aligned}$$

$$\begin{aligned}
 Z \sim \gamma(p), P(Z \geq 5) &= \sum_{k=5}^{\infty} 0.0228 \cdot (1 - 0.0228)^{k-1} = \\
 &= 0.0228 \sum_{k=5}^{\infty} (1 - 0.0228)^{k-1} = \\
 &= 0.0228 \cdot \frac{(1-0.0228)^4}{0.0228} = (1 - 0.0228)^4 = 0.912
 \end{aligned}$$

Ejercicio R. 5.4

La proporción de errores que se cometen en determinado identificador sigue una distribución beta de parámetros (3, 2). Calcule la probabilidad de que la proporción sea menor de 0.4.

Solución:

Sabiendo que la función de densidad de la variable aleatoria $\beta(3,2)$ viene dada por:

$$f(x) = \frac{\Gamma(5)}{\Gamma(3)\Gamma(2)} x^2(1-x) \quad x \in (0, 1)$$

La probabilidad de que la proporción sea menor de 0.4 se calcula como:

$$\int_0^{0.4} \frac{\Gamma(5)}{\Gamma(3)\Gamma(2)} x^2(1-x) dx = 4 \cdot 0.4^3 - 3 \cdot 0.4^4 = 0.179$$

Ejercicios propuestos**Ejercicio P. 5.1**

La vida útil de una pieza de un portátil sigue una distribución exponencial, sabiendo que la probabilidad de que funcione más de 1000 horas es de 80%. Se pide:

1. Calcular la probabilidad de que funcione más de 1500 horas.
2. Determinar cuántas horas puede mantenerse en funcionamiento con una probabilidad de 90%.

Ejercicio P. 5.2

El contenido de un café solo que se vende en la cafetería de la facultad se distribuye normalmente, con una media de 200 ml y una desviación típica de 30 ml. Se pide:

1. Calcular la probabilidad de que el contenido de un café esté entre 180 ml y 220 ml.
2. Determinar el contenido mínimo de un café para que esté en el 10% superior de los cafés servidos.

Ejercicio P. 5.3

Sabemos que la esperanza de una variable aleatoria continua X es 10 y su varianza es 12. Si todos los intervalos de igual longitud son equiprobables. Se pide:

1. Calcular el valor mínimo y el valor máximo que puede tomar la variable.
2. Calcular la probabilidad de que X sea mayor que 10.

Ejercicio P. 5.4

El tiempo (en horas) que tardan en poner un café es una variable aleatoria cuya función de densidad es: $f(x) = \begin{cases} 22e^{-22x} & \forall x > 0 \\ 0 & \text{en otro caso} \end{cases}$. Se pide:

1. Calcular la probabilidad de que en una hora pongan 20 cafés.
2. Calcular la probabilidad de tener que esperar entre 5 y 10 minutos.
3. Calcular el coeficiente de simetría de la variable aleatoria *tiempo* y explicar el resultado.
4. Calcular la probabilidad de que tarden más de 5 minutos en servir dos cafés.

Ejercicio P. 5.5

El número de horas que un estudiante necesita para aprender un tema de programación es una variable aleatoria con distribución $N(\mu, \sigma)$. Si el 84.13% de los alumnos emplea más de tres horas y sólo el 2.28% más de nueve, ¿cuánto valen μ y σ ?

Ejercicio P. 5.6

Supongamos que un banco recibe, en promedio, 6 cheques sin fondo por día. Se pide:

1. Calcular la probabilidad de que el banco reciba 4 cheques sin fondo durante una semana (7 días).
2. Calcular la probabilidad de que el banco reciba menos de 32 cheques sin fondo en un mes (30 días), aproximando mediante la distribución normal.

Ejercicio P. 5.7

Se sabe que una variable aleatoria X sigue una distribución uniforme en el intervalo $[0,3]$, mientras que otra variable aleatoria Y sigue una distribución exponencial con parámetro λ . Se pide: Determinar el valor de λ para que $P(X < 1) = P(Y < 1)$.

Ejercicio P. 5.8

El tiempo durante el cual una cierta marca de batería funciona de manera efectiva hasta que falla (tiempo de fallo) sigue una distribución exponencial con un tiempo promedio de fallas igual a 320 días. Se pide: Calcular la probabilidad de que el tiempo de fallo sea mayor a 500 días.

Ejercicio P. 5.9

En una ciudad, el consumo de energía, medido en millones de kilovatios-hora, es una variable aleatoria que sigue una distribución gamma. Se sabe que la media es 5 millones de kilovatios-hora y la varianza es 10 millones de kilovatios-hora al cuadrado. Se pide encontrar los valores de los parámetros α y λ de la distribución gamma.

Ejercicio P. 5.10

En el presupuesto familiar, la porción que se dedica a salud sigue una distribución beta (2,2). Se pide:

1. Calcular la probabilidad de que se gaste más del 25% del presupuesto familiar en salud.
2. ¿Cuál será el porcentaje medio que las familias dedican a la compra de productos y servicios de salud?

5.8. Evaluación

Todos los estudiantes del Grado en Estadística Aplicada y del Grado en Ciencia de los Datos Aplicada de la UCM, matriculados en la asignatura de Azar y Probabilidad, tienen acceso al Campus Virtual para responder una serie de preguntas seleccionadas aleatoriamente del banco de preguntas, con el fin de obtener la calificación de la evaluación continua.

Este manual está disponible en el repositorio de la UCM, por lo que se ha dispuesto una autoevaluación para cualquier persona interesada en la asignatura, utilizando el mismo banco de preguntas del Campus Virtual, accesible en Google Forms a través del siguiente enlace: <https://forms.gle/YKpy8eJSNQZ7WFEr6>.

Bibliografía

- Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P., & Meester, L. E. (2005). *A Modern Introduction to Probability and Statistics*. Springer.
- Johnson, J. L. (2003). *Probability and Statistics for Computer Science*. John Wiley & Sons.
- Juárez, I. U., Moreno, J. S. M., & Perucha, V. T. (2003). *Lecciones de cálculo de probabilidades: Curso teórico-práctico*. Thomson.
- Juárez, I. U., Moreno, J. S. M., & Perucha, V. T. (2009). *Cálculo de probabilidades*. Ibergaceta.
- Llinas, H., & Llinas, H. (2015). *Estadística descriptiva y distribuciones de probabilidad*. Universidad del Norte.
- Mukhopadhyay, P. (2012). *An Introduction to the Theory of Probability*. World Scientific.
- Shu, Z., & Medina Sánchez, M. Á. (2024). *Development of Practical Skills in Probability: A Teaching Innovation Project to Make Applied Economics More Fun with Games of Chance*. In M. del C. Valls Martínez & J. Montero (Eds.), *Teaching Innovations in Economics: Towards a Sustainable World* (pp. 479–490). Springer. https://doi.org/10.1007/978-3-031-72549-4_23
- Susi García, R., & Espínola Vilchez, R. (2012). *Azar y Probabilidad*. Cersa.