Capítulo 5

Distribuciones especiales

En el presente capítulo se estudiarán algunos modelos de distribuciones que son muy aplicadas en Estadística. Por cuestiones de programación nos restringiremos a dos modelos discretos *Binomial y Poisson* y un modelo continuo, la distribución *normal* que es el más importante en cuanto aplicación; la razón de ello, se verá al abordar el teorema central del límite, que culmina el capítulo, y que , en esencia, nos dice que la suma de n variables aleatorias, tiende a una variable normal, cuando n tiende a infinito.

Las tres distribuciones están tabuladas con lo que los cálculos, que en la práctica suelen ser tediosos, se simplifican notablemente; la consulta de las tablas es sencilla, y en el aula, se aprenderá su manejo.

5.1 Distribución Binomial

5.1.1 Distribución de Bernoulli

Sea un experimento, en el que estamos interesados sólamente en la verificación de un suceso A, al que llamaremos *éxito* o en su contrario $\overline{A}$, al que llamaremos *fracaso*. Supongamos que $\mathrm{P}(A) = p$; entonces $\mathrm{P}\left(\overline{A}\right) = 1 - p = q$. Se define la v.a. X de la forma

$$X(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in \overline{A} \end{cases}$$

La v.a. X sólo puede tomar dos valores: 0 y 1 con probabilidades respectivas q y p. Tal variable se dice que tiene una distribución de *Bernoulli* con parámetro p. Es una v.a. discreta, cuya función de cuantía es

X	0	1
$f(x)$	q	p

Cálculo de parámetros

Es fácil hallar la esperanza y varianza de la distribución de *Bernoulli*, así como la función generatriz; las usaremos posteriormente para deducir la parámetros de la *Binomial* cuyo cálculo directo es más complicado.

- $\mathrm{E}(X) = 1 \cdot p + 0 \cdot q = p$
- $\mathrm{E}\left(X^2\right) = 1^2 \cdot p + 0^2 \cdot q = p$
- $\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \left[\mathrm{E}(X)\right]^2 = p - p^2 = p(1-p) = pq$
- $\psi(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right) = p \cdot \mathrm{e}^t + q, \quad -\infty < t < \infty$

Pruebas de Bernoulli

Si realizamos el experimento n veces con independencia y definimos las siguientes variables (una por realización)

$$X_i = \begin{cases} 1, & \text{si hay éxito en la prueba} - i - \\ 0, & \text{si hay fracaso en la prueba} - i - \end{cases}$$

Tenemos n variables independientes, todas con una distribución de *Bernoulli* de parámetro p. A éstas pruebes se les llama *n pruebas de Bernoulli* con parámetro p.

Por ejemplo, si lanzamos diez veces una moneda, y la variable X_i toma el valor 1 si se obtiene cara en el $i-$simo lanzamiento y 0 en caso contrario, entonces $\{X_i\}_{i=1}^{10}$ constituyen 10 pruebas de Bernoulli con parámetro $\frac{1}{2}$.

5.1.2 Distribución Binomial

La variable X que mide el número de éxitos en n pruebas de *Bernoulli* se llama variable *binomial* con parámetros (n,p), siendo p la probabilidad de un éxito. Se representa $X \sim B(n,p)$ y tiene la siguiente función de cuantía

$$f(x) = \begin{cases} \binom{n}{x} p^x \cdot (1-p)^{n-x}, & x \in \{0,1,2,\ldots,n\} \\ 0, & \text{en otro caso} \end{cases}$$

Obsérvese que la variable Y que mide el número de fracasos en n pruebas es $Y = n - X$ (es decir $X + Y = n$) y tiene una distribución: $Y \sim B(n,q)$ (siendo $q = 1 - p$).

La variable binomial se puede expresar como suma de variables de *Bernoulli* de la siguiente manera: para cada prueba i se define

$$X_i = \begin{cases} 1 & \text{Si hay éxito en la prueba} - i - \\ 0 & \text{Si hay fracaso en la prueba} - i - \end{cases}$$

Cada variable X_i sigue una distribución $b(p)$ y el número total de éxitos (X) es la suma de ellas:

$$X = X_1 + X_2 + \cdots + X_n.$$

Cálculo de parámetros

Usaremos la descomposición anterior para calcular los parámetros de la distribución $B(n,p)$

- $\mathrm{E}(X) = \sum_{i=1}^{n} \mathrm{E}(X_i) = np$
- $\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = npq$
- $\psi(t) = \prod_{i=1}^{n} \mathrm{E}\left(\mathrm{e}^{tX_i}\right) = (p \cdot \mathrm{e}^t + q)^n$

Usaremos de nuevo la función generatriz para demostrar el siguiente

Teorema 5.1 *Si $X_1, X_2, \ldots, X_k$ son v.a. independientes, cada una con distribución binomial $B(n_i, p)$, entonces $X = X_1 + X_2 + \cdots + X_k$ se distribuye binomialmente con parámetros $(n_1 + n_2 + \cdots + n_k, p)$; es decir*

$$X_i \sim B(n_i, p) \to X = X_1 + X_2 + \cdots + X_k \sim B(n_1 + n_2 + \cdots + n_k, p)$$

Bastará demostrar que su función generatriz es $\psi(t) = (p \cdot \mathrm{e}^t + q)^{n_1 + \cdots + n_k}$, lo cual es inmediato.

5.1.3 Tablas

La tablas recogen la función de distribución de la variable Binomial $B(n,p)$ hasta $n = 30$ y para valores de p de 5 en 5 centésimas hasta el valor 0'5 (incluyendo el valor 0'1 y $\frac{1}{3}$); para valores mayores de 30, se aproximará mediante la distribución de Poisson o Normal.
La precision es de cuatro dígitos (en algunos casos más); esto quiere decir que el error absoluto es menor que $\frac{1}{2}10^{-4}$ (media diezmilésima). Cuando aparece un valor 0'0000 (ó 0) quiere decir que es menor que 0'00005; igualmente el valor 1 indica que la probabilidad exacta es mayor que 0'99995. La tabla da el valor:

$$F(k) = \mathrm{P}\left(X \leq k\right) = \sum_{j=0}^{k} \mathrm{P}\left(X = j\right) = \sum_{j=0}^{k} \binom{n}{j} p^j \cdot (1-p)^{n-j}$$

Para valores puntuales (función de cuantía) usaremos que

$$F(k) = f(0) + f(1) + \cdots + f(k)$$

por lo que

$$f(0) = F(0), \qquad f(k) = F(k) - F(k-1) \qquad k > 0.$$

Para valores de $p > 0'5$ se usa el resultado siguiente: si $X \sim B(n,p)$ representando X el número de éxitos en n pruebas de Bernoulli, el número de fracasos es $Y = n - X$ que tendrá una distribución binomial con parámetros n y $q = 1 - p$. Es decir:

$$X \sim B(n,p) \longrightarrow Y = n - X \sim B(n, 1-p)$$

Así, si $p > 0'5$ entonces $1 - p < 0'5$ y la función de distribución que aparece en las tablas es la de Y. Por tanto se sustituye X por $n - Y$

- $$\begin{aligned} \mathrm{P}\left(X \leq k\right) = \mathrm{P}\left(n - Y \leq k\right) &= \mathrm{P}\left(Y \geq n-k\right) \\ &= 1 - \mathrm{P}\left(Y < n-k\right) \\ &= 1 - F_Y(n-k-1) \end{aligned}$$

- $\mathrm{P}\left(X = k\right) = \mathrm{P}\left(Y = n-k\right) = F_Y(n-k) - F_Y(n-k-1)$

- $\mathrm{P}\left(X \geq k\right) = \mathrm{P}\left(n - Y \geq k\right) = \mathrm{P}\left(Y \leq n-k\right) = F_Y(n-k)$.

Ejemplos:

(a) Sea $X \sim B(13, 0'4)$. Calcular $\mathrm{P}(X \leq 7)$ y $\mathrm{P}(X = 10)$.
En la tabla se obtiene directamente:

- $\mathrm{P}(X \leq 7) = F(7) = 0'9023$.
- $\mathrm{P}(X = 10) = F(10) - F(9) = 0'9987 - 0'9922 = 0'0065$.

Este último valor se puede hallar con más precisión con la calculadora

$$\mathrm{P}(X = 10) = \binom{13}{10} 0'4^{10} \cdot 0'6^3 = 0'006477683.$$

(b) Dada $X \sim B(20, 0'8)$, calcular $\mathrm{P}(X \leq 16), \mathrm{P}(X = 14)$ y $\mathrm{P}(X > 8)$. La distribución $Y = 20 - X$ es binomial de parámetros 20 y $0'2$.

- $\mathrm{P}(X \leq 16) = \mathrm{P}(Y \geq 4) = 1 - \mathrm{P}(Y \leq 3) = 1 - F_Y(3) = 1 - 0'4114 = 0'5886$.
- $\mathrm{P}(X = 14) = \mathrm{P}(Y = 6) = F_Y(6) - F_Y(5) = 0'9133 - 0'8042 = 0'1091$.
- $\mathrm{P}(X > 8) = \mathrm{P}(Y < 12) = \mathrm{P}(Y \leq 11) = F_Y(11) = 0'9999$.

(c) Calcúlese la probabilidad de obtener 10, 11, o 12 caras al lanzar 20 monedas. Repetir el problema con 25 monedas.

- Si $X =$ número de caras, $X \sim B(20, 0'5)$. Como aparece en tablas

$$f(10) + f(11) + f(12) = F(12) - F(9) = 0'8684 - 0'4119 = 0'4565.$$

- En este caso $X \sim B(25, 0'5)$ que no aparece en tablas; usamos el cálculo directo

$$\begin{aligned} f(10) + f(11) + f(12) &= \binom{25}{10}\left(\frac{1}{2}\right)^{25} \\ &+ \binom{25}{11}\left(\frac{1}{2}\right)^{25} \\ &+ \binom{25}{12}\left(\frac{1}{2}\right)^{25} \\ &= 0'3852. \end{aligned}$$

Ejemplo: *En el lanzamiento de* 20 *monedas ¿Cuál es la probabilidad de obtener*

(a) Exactamente 14 *caras.*

(b) Al menos 14 *caras.*

En cada lanzamiento, llamaremos éxito a obtener cara. Estamos ante 20 pruebas de *Bernoulli* con parámetro $\frac{1}{2}$. El número total de caras sigue una distribución $B(20, \frac{1}{2})$. Consultando las tablas, se tiene

- $\mathrm{P}(X = 14) = F(14) - F(13) = 0'9793 - 0'9423 = 0'0370.$

- $\mathrm{P}(X \geq 14) = 1 - \mathrm{P}(X \leq 13) = 1 - F(13) = 1 - 0'9423 = 0'0577.$

5.2 Distribución de Poisson

La Distribución de Poisson toma el nombre de *Simeón Dennis Poisson* (1781-1840), que desarrolló esta distribución como límite de la binomial. Se emplea para describir procesos, como la distribución de las llamadas telefónicas, las llegadas de automóviles a un determinado parking, el número de accidentes en cierto punto de cruces de carreteras, el número de defectos de una cinta magnética, etc. Estas variables son discretas que toman valores enteros $0, 1, 2, 3, \ldots$ y miden el número de ocurrencias por unidad de magnitud (tiempo, espacio, etc). Se le llamó la distribución de los *sucesos raros* . Un ejemplo es el número de accidentes por unidad de tiempo (por ejemplo una semana) en un determinado cruce o punto negro de alguna carretera; podemos dividir la semana en n intervalos de modo que en cada uno se pueda producir a lo sumo un accidente (tomando n lo suficientemente grande para que el intervalo de tiempo sea lo suficientemente pequeño) y podemos suponer que los accidentes son independientes de un intervalo a otro y que la probabilidad de accidente en cualquier intervalo es la misma p (que será muy pequeña); entonces como máximo se pueden producir n accidentes y la variable $X = \{$ número de accidentes por semana $\}$ corresponde a una binomial $B(n, p)$ con n muy grande lo que dificulta el cálculo; Si hacemos tender n a infinito (y por tanto p a cero) y manteniendo constante el producto $n \cdot p$ (que es la media de la distribució binomial y al que llamaremos λ) tenemos

$$\begin{aligned}
\mathrm{P}\,(X=k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n\cdot(n-1)\cdots(n-k+1)}{k\,!}\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k\,!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n}\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-k}
\end{aligned}$$

Si $n \to \infty$ y el valor $p \to 0$, se tiene

$$\lim_{n\to\infty}\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n}\left(1-\frac{\lambda}{n}\right)^{-k} = 1$$

y puesto que $\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} = \mathrm{e}^{-\lambda}$, obtenemos

$$\mathrm{P}\,(X=K) \approx \mathrm{e}^{-\lambda}\cdot\frac{\lambda^k}{k\,!}$$

Definición 5.1 *Se dice que una v.a. X tiene una distribución de* Poisson *con media $\lambda > 0$, que denotaremos $X \sim P(\lambda)$, si es una v.a. discreta con función de cuantía*

$$f(k) = \begin{cases} \mathrm{e}^{-\lambda}\cdot\dfrac{\lambda^k}{k\,!}, & k\in\{0,1,\ldots\} \\ 0, & \text{en otro caso} \end{cases}$$

La demostración de que la función anterior es de cuantía, viene por el hecho de que λ es positivo, y que aplicando el desarrollo de *Taylor* de la función exponencial se tiene

$$\sum_{k=0}^{\infty}\mathrm{e}^{-\lambda}\cdot\frac{\lambda^k}{k\,!} = \mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k\,!} = \mathrm{e}^{-\lambda}\cdot\mathrm{e}^{\lambda} = 1.$$

Cálculo de parámetros

- Función generatriz
$$\psi(t)=\sum_{x=0}^{\infty}\frac{\mathrm{e}^{tx}\,\mathrm{e}^{-\lambda}\,\lambda^x}{x\,!}=\mathrm{e}^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda\,\mathrm{e}^t)^x}{x\,!}=\mathrm{e}^{-\lambda}\,\mathrm{e}^{\lambda e^t}=\mathrm{e}^{\lambda(e^t-1)}$$
- Media. El valor medio es precisamente λ. En efecto, como $\mathrm{E}\,(X)=\psi'(0)$
$$\psi'(t)=\mathrm{e}^{\lambda(e^t-1)}\cdot\lambda e^t\rightarrow\psi'(0)=\lambda.$$
- Varianza. Calculemos primero $\mathrm{E}\left(X^2\right)=\psi''(0)$
$$\psi''(t)=\mathrm{e}^{\lambda(e^t-1)}\cdot(\lambda e^t)^2+\mathrm{e}^{\lambda(e^t-1)}(\lambda e^t)\rightarrow\psi''(0)=\lambda^2+\lambda.$$
Entonces $\mathrm{Var}\,(X)=\mathrm{E}\left(X^2\right)-[\mathrm{E}\,(X)]^2=\lambda^2+\lambda-\lambda^2=\lambda$

Teorema 5.2 *Si las variables $X_1,X_2,\ldots,X_k$ son independientes, y cada una tiene una distribución de Poisson, $X_i\sim P(\lambda_i)$, $i=1,\ldots,k$, la suma X es una variable de Poisson de parámetro $\lambda_1+\cdots+\lambda_k$*

Demostración: Si la variable X tiene una distribución $P(\lambda_1+\cdots+\lambda_k)$, su función generatriz debe ser:
$$\psi(t)=\mathrm{e}^{(\lambda_1+\lambda_2+\cdots+\lambda_k)(e^t-1)}$$
Como las variables X_i son independientes, la función generatriz de la suma se obtiene multiplicando las funciones generatrices de las variables sumandos
$$\psi(t)=\prod_{i=1}^{k}\psi_i(t)=\prod_{i=1}^{k}\mathrm{e}^{\lambda_i(e^t-1)}=\mathrm{e}^{(\lambda_1+\lambda_2+\cdots+\lambda_k)(e^t-1)}$$
Supongamos que la entrada de personas por hora a un cierto local tiene una distribución $X\sim P(7)$ ¿Cómo se distribuye la variable $Y=\{$ número de personas que entran cada 2 horas $\}$? Podemos descomponer Y en la suma de $X_1=\{$ número de personas que entran en la primera hora $\}$ y $X_2=\{$ número de personas que entran en la segunda hora$\}$. Según el teorema anterior
$$Y=X_1+X_2\sim P(7+7)=P(14).$$
En general si se cambia la unidad de medida (tiempo en este caso) el número de acontecimientos sigue una distribución de Poisson de media proporcional

Tiempo	Distribución
1	$P(\lambda)$
t	$P(\lambda \cdot t)$

Esto es aplicable a otras magnitudes como longitud, superficie, et.

5.2.1 Tablas

La precisión es de 4 dígitos y los valores del parámetro λ aparecen de décima en décima hasta 10 y en intervalos de uno hasta 21. Se recoge el valor

$$F(k) = \mathrm{P}\left(X \leq k\right) = \sum_{i=0}^{k} \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k\,!}$$

Ejemplo: *Si* $X \sim P(6)$ *calcular* $\mathrm{P}\left(X \leq 9\right)$ *y* $\mathrm{P}\left(X = 15\right)$

(a) $\mathrm{P}\left(X \leq 9\right) = F(9) = 0'9161$

(b) $\mathrm{P}\left(X = 15\right) = F(15) - F(14) = 0'9995 - 0'9986 = 0'0009$

Las probabilidades puntuales (funcion de cuantía) se pueden calcular sin usar la tabla, con la ayuda de una calculadora; así, en el ejemplo anterior se podría haber efectuado el cálculo de la forma:

$$\mathrm{P}\left(X = 15\right) = \mathrm{e}^{-6} \cdot \frac{6^{15}}{15\,!} = 0'000891.$$

El valor obtenido con las tablas, 0’0009, es menos exacto; se debe esto a que las tablas se presentan con cuatro dígitos, y al restar dos números próximos, se pierde significación.

Ejemplo: *Si el número de defectos por metro en una cinta magnética sigue una distribución* $P(6)$, *calcúlese la probabilidad de que en* 5 *metros haya* 30 *defectos.*

Puesto que el número X de defectos en 5 metros sigue una distribución $P(30)$, usaremos cálculo directo

$$\mathrm{P}\left(X = 30\right) = \mathrm{e}^{-30} \cdot \frac{30^{30}}{30\,!} = 0'072634.$$

Ejemplo: *El número de llegadas de vehículos a un determinado parking, tiene una distribución de Poisson de media* 6 *cada hora. Calcúlese la probabilidad de que*

(a) En una determinada hora entren 4 *o más vehículos.*

(b) En una hora y media entren exactamente 10 *vehículos.*

(c) Entre algún vehículo en media hora.

La solución es

(a) Como $X \sim P(6)$, $\quad \mathrm{P}(X \geq 4) = 1 - \mathrm{P}(X \leq 3) = 0'8488$

(b) $X \sim P(9)$, $\quad \mathrm{P}(X = 10) = 0'1186$

(c) $X \sim P(3) \to \mathrm{P}(X \geq 1) = 1 - \mathrm{P}(X = 0) = 1 - 0'0498 = 0'9502$

5.2.2 Aproximación de la Binomial por la de Poisson

Las tablas para la distribución *Binomial* suelen ser del orden de $n = 30$ ó 25 (o menos), y en determinados problemas no se pueden utilizar por ser n mayor. Hemos visto en la introducción que se puede aproximar por la de *Poisson*. Si n es grande (> 30), y $X \sim B(n, p)$, entonces si llamemos $\lambda = np$, se tiene

$$\boxed{B(n, p) \approx P(np)}$$

En la práctica, esta aproximación da buenos resultados cuando $p <$0'5 y $np < 5$.

Ejemplo: *La probabilidad de éxito en una determinada prueba es* $0'04$*. Si se realizan* 100 *pruebas con independencia, ¿cuál es la probabilidad de obtener al menos* 4 *éxitos?*

La variable $X =$ número de éxitos en 100 pruebas tiene una distribución binomial $B(100, 0'04)$. Como las tablas no sirven y el cálculo directo es algo engorroso se puede aproximar por la distribución de Poisson.

$$X \sim B(100, 0'04) \approx P(4).$$

Por la tanto

$$\mathrm{P}(X \geq 4) = 1 - \mathrm{P}(X \leq 3) = 1 - F(3) = 1 - 0'4335 = 0'5665.$$

El cálculo directo proporciona el valor 0'56652988. La aproximación es muy buena.

Relación con la exponencial

Imaginemos que una sucesión de acontecimientos se producen en el tiempo según un proceso de Poisson de media λ por unidad de tiempo (u.t.) y queremos medir el tiempo hasta que aparece el primer acontecimiento. Sean

$$\begin{array}{rcl} X & = & \{ \text{ número de acontecimientos por u.t. } \} \quad X \sim P(\lambda) \\ T & = & \{ \text{ tiempo hasta el primer acontecimiento } \} \end{array}$$

Partimos de $t = 0$; el número de acontecimientos en el intervalo $[0, t]$ sigue una distribución $P(\lambda t)$ y la probabilidad de ningún acontecimiento en ese intervalo es

$$\mathrm{P}\,(X = 0) = f(0) = \mathrm{e}^{-\lambda t}\,\frac{(\lambda t)^0}{0\,!} = \mathrm{e}^{-\lambda t}\,.$$

Estudiemos la distribución de T

$$\begin{array}{rcl} \mathrm{P}\,(T \leq t) & = & 1 - \mathrm{P}\,(T > t) \\ & = & 1 - \mathrm{P}\,(\text{ningún acontecimiento en } [0, t]) \\ & = & 1 - \mathrm{e}^{-\lambda t}\,. \end{array}$$

Esta distribución es la *exponencial* (véase el ejemplo de la página 71). Lo visto es válido para el tiempo de espera entre dos acontecimientos consecutivos. Se resume asi

> Si el número de acontecimientos que se producen por unidad de tiempo siguen un proceso de Poisson de media λ, entonces el tiempo entre dos acontecimientos consecutivos, sigue una distribución exponencial de parámetro λ (de media $1/\lambda$).

Sin embargo el tiempo de espera (desde $t = 0$) hasta el $-i-$ésimo acontecimiento sigue una distribución no estudiada en este libro; se llama *Gamma*, de la que la exponencial es un caso particular. Naturalmente se aplica también a otras magnitudes aparte del tiempo; si los defectos de una cinta magnética tienen una distribución de Poisson de parámetro λ por unidad de longitud (digamos metros), entonces la distancia (en metros) entre dos defectos consecutivos sigue una distribución exponencial.

5.3 Distribución Normal

El que la distribución *normal* sea la mas importante, se verá posteriormente pero se ha comprobado que mediciones realizadas respecto a muchas variables aleatorias parecen haber sido generadas a partir de frecuencias poblacionales que se aproximan de forma notable por una variable normal.

5.3.1 Definición

Una v.a. continua X tiene una distribución *normal* de parámetros μ y σ, y escribiremos $X \sim N(\mu, \sigma)$ si su f.d.d es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \, \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

Veremos ahora que la media es μ y la desviación típica σ. El caso particular $\mu = 0$ y $\sigma = 1$ se llama *tipificada* y veremos que cualquier otra se puede reducir a ésta. Se acostumbra a representarla por $Z = N(0, 1)$.

Cálculo de parámetros

Asumiendo que $\int\limits_{-\infty}^{\infty} f(x)\, dx = 1$, usaremos este resultado para calcular la función generatriz, y deducir de ésta los momentos primero y segundo.

$$\psi(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right) = \int\limits_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \, \mathrm{e}^{tx - \frac{(x-\mu)^2}{2\sigma^2}} \, dx$$

Manipulando el exponente de e

$$tx - \frac{(x-\mu)^2}{2\sigma^2} = \mu t + \frac{1}{2}\sigma^2 t^2 - \frac{[x - (\mu + \sigma^2 t)]^2}{2\sigma^2}$$

resultando que

$$\psi(t) = \mathrm{e}^{\mu t + \frac{1}{2}\sigma^2 t^2} \int\limits_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \, \mathrm{e}^{-\frac{[x-(\mu+\sigma^2 t)]^2}{2\sigma^2}} \, dx$$

Como la integral vale 1 (es la función de densidad sustituyendo la constante μ por $\mu + \sigma^2 t$), se deduce

$$\psi(t) = \mathrm{e}^{\mu t + \frac{1}{2}\sigma^2 t^2}, \text{ para todo } x$$

Ahora es fácil deducir la media y varianza

- $\psi'(t) = (\mu + t\sigma^2)\,\mathrm{e}^{\mu t + \frac{1}{2}\sigma^2 t^2}$
- $\mathrm{E}\,(X) = \psi'(0) = \mu$
- $\psi''(t) = ((\mu + t\sigma^2)^2 + \sigma^2)\,\mathrm{e}^{\mu t + \frac{1}{2}\sigma^2 t^2}$
- $\mathrm{E}\left(X^2\right) = \psi''(0) = \mu^2 + \sigma^2$
- $\mathrm{Var}\,(X) = \mathrm{E}\left(X^2\right) - (\mathrm{E}\,(X))^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$

Teorema 5.3 *Si* $X \sim N(\mu, \sigma)$, entonces $\dfrac{X - \mu}{\sigma} \sim N(0, 1)$

La demostración queda a cargo del alumno, que debe usar el resultado

$$Y = aX + b \to \psi_Y(t) = \mathrm{e}^{bt}\,\psi_X(at)$$

El proceso de pasar de una normal cualquiera a la normal tipificada se llama *tipificar.*

$$X \xrightarrow{\text{Tipificar}} \frac{X - \mu}{\sigma} = Z.$$

5.3.2 Tabla

Denotaremos por $\phi(x)$ a la función de densidad de la normal tipificada y $\Phi(x)$ a la función de distribución; esta última viene tabulada para valores positivos y menores que 4, pues prácticamente $\Phi(4) \approx 1$. El alumno debe demostrar que $\Phi(x) + \Phi(-x) = 1$ con lo que el cálculo para los valores negativos se pueden conseguir fácilmente. La tabla recoge los valores

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\dfrac{x^2}{2}}\,dx$$

Está tabulada para $z \in [0, 4'09]$ de centésima en centésima. Para valores mayores se toma $\Phi(z) \approx 1$. Para valores negativos, usamos $\Phi(-z) = 1 - \Phi(z)$. La precisión es de 4 dígitos, aunque para $z \geq 3$, al ser valores próximos a 1 se aumenta la precisión dando el número de nueves en forma de exponente; por ejemplo

$\Phi(3'93) = .9^4 58$ quiere decir $0'999958$.

Si el valor aparece en la tabla, por ejemplo, calcular $\Phi(2'76)$, se mira en la intersección de la fila correspondiente a $2'7$ y en la columna $0'06$ obteniéndose $0'9971$. Si el valor no aparece en la tabla, por ejemplo, calcular $\Phi(1'342)$, como el valor se encuentra entre $1'34$ y $1'35$ se usa interpolación lineal: conocida la función en dos puntos $y_0 = \Phi(z_0)$ e $y_1 = \Phi(z_1)$, sustituimos la curva entre los puntos $(z_0, y_0), (z_1, y_1)$ por un segmento lineal; dado $z \in [z_0, z_1]$, el valor $y = \Phi(z)$ se calcula:

$$y = y_0 + \frac{z - z_0}{z_1 - z_0}(y_1 - y_0)$$

En el ejemplo anterior se tiene:

$$y_0 = 0'9099 \qquad y_1 = 0'9115$$

por lo que

$$y = 0'9099 + \frac{0'002}{0'01}(0'9115 - 0'9099) = 0'91022 \approx 0'9102.$$

Búsqueda inversa: hallar z dado $\Phi(z)$. Usaremos:

$$z = z_0 + \frac{y - y_0}{y_1 - y_0}(z_1 - z_0)$$

Hallar z tal que $\Phi(z) =$ 0'675; en este caso conocemos y y buscamos z. El valor 0'675 está comprendido entre los valores 0'6736 y 0'6772 correspondientes a los valores $z_0 = 0'45$ y $z_1 = 0'46$ respectivamente.

$$z = 0'45 + \frac{0'675 - 0'6736}{0'6772 - 0'6736}(0'01) = 0'453888 \approx 0'454.$$

Valores negativos de z. Hállese $\Phi(-1'65)$. Puesto que $\Phi(z) = 1 - \Phi(-z)$, tenemos

$$\Phi(-1'65) = 1 - \Phi(1'65) = 1 - 0'9505 = 0'0495.$$

Hállese z tal que $\Phi(z) = 0'1230$. Un valor para $\Phi(z)$ menor que 0'5, delata que z es negativo; se busca por tanto:

$$\Phi(-z) = 1 - 0'1230 = 0'8770 \to -z = 1'16 \to z = -1'16.$$

Ejemplo: *Dada* $X \sim N(3, 2)$*, calcúlese* $\mathrm{P}(1 < X < 5)$.

Se procede a tipificar la variable X

$$\frac{X - 3}{2} = Z \sim N(0, 1)$$

y hacemos los cálculos siguientes

$$\begin{aligned}
\mathrm{P}\left(1<X<5\right) &= \mathrm{P}\left(\frac{1-3}{2}<\frac{X-3}{2}<\frac{5-3}{2}\right) \\
&= \mathrm{P}\left(-1<Z<1\right) \\
&= \Phi(1)-\Phi(-1) \\
&= 2\Phi(1)-1 \\
&= 2\cdot 0'8413-1 \\
&= 0'6826.
\end{aligned}$$

5.3.3 Combinación lineal de normales

Usando la función generatriz de momentos se demuestran los siguientes resultados

Teorema 5.4 *Si las k v.a. X_i son independientes y normales con media μ_i y varianza σ_i^2, entonces la variable $a_1X_1 + a_2X_2 + \ldots + a_kX_k + b$ tiene una distribución normal con media $a_1\mu_1 + a_2\mu_2 + \ldots + a_k\mu_k + b$ y varianza $a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots + a_k^2\sigma_k^2$*

Teorema 5.5 *Si las n v.a. X_i son una muestra aleatoria de una distribución normal con media μ y varianza σ^2, entonces la variable media muestral $\overline{X_n}$ tiene una distribución normal con media μ y varianza $\frac{\sigma^2}{n}$.*

Ejemplo: *Supongamos que el voltaje medido en un circuito tiene una distribución* normal *con media* 220 *y desviación* 4. *Se toman dos medidas independientes del voltaje. Hallar la probabilidad de que la media muestral esté entre* 218 *y* 221.

Si X_1 y X_2 son las dos medidas, ambas tienen una distribución

$$X_1 \sim N(220,4), \qquad X_2 \sim N(220,4).$$

La media muestral $\overline{X}_2 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ se distribuye normalmente:

$$\mathrm{E}\left(\overline{X}_2\right) = 200, \quad \mathrm{Var}\left(X\right) = \frac{1}{4}\mathrm{Var}\left(X_1\right) + \frac{1}{4}\mathrm{Var}\left(X_2\right) = 8.$$

Por tanto $\overline{X}_2 \sim N(220, \sqrt{8})$ y tipificando

$$\begin{aligned}
\mathrm{P}\left(218 < \overline{X}_2 < 221\right) &= \mathrm{P}\left(\frac{218-220}{\sqrt{8}} < Z < \frac{221-220}{\sqrt{8}}\right) \\
&= \mathrm{P}\left(-0'71 < Z < 0'35\right) \\
&= \Phi(0'35) - \Phi(-0'71) \\
&= \Phi(0'35) + \Phi(0'71) - 1 = 0'3979.
\end{aligned}$$

5.4 Teorema central del límite

Teorema 5.6 (Teorema central del límite) *Si las v. a.* $\{X_i\}_1^n$ *son independientes e idénticamente distribuidas con* $\mathrm{E}(X_i) = \mu$ *y varianza* $\mathrm{Var}(X_i) = \sigma^2$ *(finita), entonces*

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \leq x\right) = \Phi(x)$$

Hay que notar que la variable puede ser *cualquiera*, continua o discreta, y el teorema dice que

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}$$

será aproximadamente normal tipificada. Esto equivale a

$$\overline{X}_n \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

O también

$$X_1 + X_2 + \ldots + X_n \sim N(n\mu, \sigma\sqrt{n})$$

Una consecuencia es que las distribuciones *Binomial* y *Poisson* se pueden aproximar por la *Normal*.

Aproximaciones por la normal

- **Binomial**. En la página 169 se vió que la variable $X \sim B(n,p)$ se puede interpretar como la suma de n variables de *Bernoulli* $X = X_1 + X_2 + \cdots + X_n$ siendo

$$X_i = \begin{cases} 1 & \text{Si hay éxito en la prueba} - i - \\ 0 & \text{Si hay fracaso en la prueba} - i - \end{cases}$$

Como $\{X_i\}$ son independientes con media p y varianza $p(1-p) = pq$, y la variable $\overline{X}_n = \dfrac{X}{n}$, según el T.C.L la variable

$$\boxed{\frac{\sqrt{n}(\frac{X}{n} - p)}{\sqrt{pq}} = \frac{X - np}{\sqrt{npq}} \approx N(0,1)}$$

- **Poisson**. Considerando una distribución de *Poisson* $X \sim P(\lambda)$, como límite de una *Binomial* se demuestra que

$$\boxed{\frac{X - \lambda}{\sqrt{\lambda}} \approx N(0,1)}$$

Corrección por continuidad

El paso de una distribución discreta a una continua (en este caso la normal) supone un inconveniente; para una variable discreta pueden ser distintos los números $\mathrm{P}(X < a)$ y $\mathrm{P}(X \leq a)$, y sin embargo, son iguales para una variable continua. Para éstas, es nula la probabilidad $\mathrm{P}(X = a)$ y sin embargo para una discreta puede no serlo. Para corregir este desequilibrio se hace lo siguiente: supongamos que la distribución discreta toma valores enteros; entonces dado $a \in \mathbb{Z}$

$$\mathrm{P}(X < a) = \mathrm{P}\left(X \leq a - \frac{1}{2}\right).$$

A continuación se procede a tipificar. Las correcciones se resumen en

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

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

- $\mathrm{P}\,(a \leq X < b) = \mathrm{P}\left(a - \frac{1}{2} \leq X \leq b - \frac{1}{2}\right)$.

Veamos el siguiente

> **Ejemplo:** *Estímese la probabilidad de obtener más de* 475 *caras al lanzar una moneda* 1000 *veces.*

Aunque el número de caras X, en el lanzamiento de una moneda 1000 veces (o lanzamiento de 1000 monedas) sigue una distribución $X \sim B(1000, 0'5)$, no se puede resolver por tablas como tal. Aproximando por la normal $X \sim N(500, \sqrt{250})$ y, usando la corrección:

$$
\begin{aligned}
\mathrm{P}\,(X > 475) &= 1 - \mathrm{P}\,(X \leq 475) \\
&= 1 - \mathrm{P}\left(X \leq 475'5\right) \\
&= 1 - \mathrm{P}\left(\frac{X - 500}{\sqrt{250}} \leq \frac{475'5 - 500}{\sqrt{250}}\right) \\
&= 1 - \mathrm{P}\left(Z \leq -1'55\right) \\
&= 1 - \Phi(-1'55) \\
&= \Phi(1'55) \\
&= 0'9394
\end{aligned}
$$

Sin usar corrección el resultado sería $0'9429$. Los cálculos directos (con EXCEL) dan $0'9394$ con lo que se ve la ventaja de usar la corrección.

> **Ejemplo:** *Se lanza un dado* 1000 *veces. Hállese la probabilidad de que el número de* seises *esté entre* 150 *y* 170 *(incluidos).*

La variable X = número de seises es $B(1000, \frac{1}{6})$. Los parámetros son $\mu = 500/3$, $\sigma = \sqrt{500/3 \cdot 5/6} = 25\sqrt{2}/3$, así que la variable $\dfrac{X - 500/3}{25\sqrt{2}/3}$ se puede considerar la normal tipificada. Por tanto

$$\begin{aligned}
\mathrm{P}\,(150 \leq X \leq 170) &= \mathrm{P}\left(149'5 \leq X \leq 170'5\right) \\
&= \mathrm{P}\left(\frac{149'5 - \frac{500}{3}}{\frac{25\sqrt{2}}{3}} \leq Z \leq \frac{170'5 - \frac{500}{3}}{\frac{25\sqrt{2}}{3}}\right) \\
&= \mathrm{P}\left(-1'46 \leq Z \leq 0'33\right) \\
&= \Phi(0'33) - \Phi(-1'46) \\
&= 0'5572.
\end{aligned}$$

Sin usar la corrección el resultado es $0'5310$; usando EXCEL, los cálculos directos dan $0'5596$.

Resumen de distribuciones

	Nombre	Repres.	Media	Varianza	Generatriz
Discretas	Bernoulli	$b(p)$	p	pq	$p\,\mathrm{e}^t + q$
	Binomial	$B(n,p)$	np	npq	$(p\,\mathrm{e}^t + q)^n$
	Poisson	$P(\lambda)$	λ	λ	$\mathrm{e}^{\lambda(\mathrm{e}^t - 1)}$
Continuas	Exponencial	$\mathrm{Exp}(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t} \quad (t < \lambda)$
	Normal	$N(\mu,\sigma)$	μ	σ^2	$\mathrm{e}^{\mu t + \frac{1}{2}\sigma^2 t^2}$
	Normal tip.	$Z \sim N(0,1)$	0	1	$\mathrm{e}^{\frac{1}{2}t^2}$

5.5 Problemas

Binomial

5.1 Si $X \sim B(8, 0'35)$, calcular

(a) $\mathrm{P}(X = 3)$ (S: 0’2768)

(b) $\mathrm{P}(X \leq 5)$ (S: 0’9747)

(c) $\mathrm{P}(X \geq 5)$ (S: 0’1061)

5.2 Si $X \sim B(12, 0'7)$ calcular

(a) $\mathrm{P}(X \leq 5)$ (S: 0’0386)

(b) $\mathrm{P}(X > 7)$ (S: 0’7237)

(c) $\mathrm{P}(X = 6)$ (S: 0’0792)

5.3 Una v.a. X se distribuye *binomialmente* con media 3 y varianza 2. Calcular la probabilidad $\mathrm{P}(X = 7)$. (S: 0’0073)

5.4 Se sabe que $X \sim B(8, p)$ y que $\mathrm{P}(X \leq 3) = 0\text{'}1061$. Estimar el valor de p. (S: 0’65)

5.5 ¿Cuántos dígitos se han de tomar para que la probabilidad de que aparezca por lo menos un cinco sea superior a 0’9? (S: 22)

5.6 Un sistema electrónico tiene 20 componentes. Los componentes pueden fallar de forma independiente con probabilidad 0’05, y el sistema falla cuando falla algún componente. Si se sabe que el sistema no funciona ¿cuál es la probabilidad de que hayan fallado más de un componente? (S: 0’4118)

5.7 Lanzamos 15 monedas al aire; probabilidad de que el número de caras esté comprendido entre 5 y 10. (ambas inclusive) (S: 0’8816)

5.8 Una vacuna con el 85 % de eficacia se aplica a 20 personas enfermas. Probabilidad de que sanen más de la mitad. (S: 0’9998)

5.9 Suponiendo que en cada nacimiento la probabilidad de ser chico o chica es la misma, hállese la probabilidad de que en una familia de 10 hijos,

(a) Haya 7 chicas

(b) Haya al menos 4 chicos

(S: 0'1172, 0'8281)

5.10 Una compañía de seguros estima que una de cada mil personas incurre en un cierto accidente a lo largo de un año. Si la compañía tiene aseguradas a diez mil personas ¿cuál es la probabilidad de que en un año no más de tres de sus asegurados incurran en ese accidente? (S: $0'0103$)

5.11 La longitud de cierta pieza se distribuye con función de densidad

$$f(x) = \begin{cases} k(x-1)(3-x), & x \in [1,3] \\ 0, & \text{en otro caso} \end{cases}$$

Sólo es válida una pieza, si su longitud está entre 1'5 y 2'5.

(a) ¿Cuál es la probabilidad de que una pieza sea útil? (S: 0'6875)

(b) Si las piezas se empaquetan en lotes de 5 unidades, y se acepta el lote si contiene menos de 2 piezas defectuosas ¿Cuál es la probabilidad de que un cierto lote sea rechazado? (S: 0'4973)

5.12 Se lanzan 6 dados normales. Llamando éxito a obtener un 5 o un 6, hallar la probabilidad de

(a) 3 éxitos. (S: 0'2195)

(b) Como máximo 3 éxitos. (S: 0'8999)

5.13 Dos personas juegan a cara y cruz y han convenido en continuar la partida hasta que tanto la cara como la cruz se hayan presentado al menos 3 veces. Hallar la probabilidad de que el juego no acabe cuando se han hecho 10 tiradas.
(S: 0'1094)

5.14 Una máquina produce una pieza defectuosa en el 5 % de los casos. De la producción diaria se selecciona una muestra de 12 piezas, y si se encuentran 3 o más defectuosas se inspecciona el 100 % de la producción de ese día. Hállese la probabilidad de que la producción de un día se inspeccione al 100 % (S: 0'0196)

5.15 La ceguera al color es una enfermedad que aparece en un 1 % en una cierta población. Calcular el tamaño n de la muestra con reemplazamiento de la población, para que la probabilidad de que contenga al menos una persona ciega al color sea al menos de 0'95. (S: 59)

Poisson

5.16 Siendo $X \sim P(2'9)$, calcula

(a) $\mathrm{P}(X \leq 5)$ (S: 0'9258)

(b) $\mathrm{P}(X = 7)$ (S: 0'0188)

(c) $\mathrm{P}(X \geq 4)$ (S: 0'3304)

5.17 El número de defectos que tiene una pieza manufacturada sigue una distribución de Poisson con media $0'5$. Si se examinan 10 piezas ¿cuál es la probabilidad de encontrar más de 5 defectos? (S: 0'384)

5.18 Las llamadas a una pequeña central telefónica siguen una distribución de Poisson de media 3 llamadas cada cinco minutos. Calcular la probabilidad de

(a) Que haya 5 llamadas en cinco minutos. (S: 0'1008)

(b) Que haya 5 llamadas en 10 minutos. (S: 0'1606)

(c) Que haya 3 o menos llamadas en un minuto. (S: 0'9966)

(d) Que haya al menos una llamada en un minuto. (S: 0'4512)

5.19 Suponiendo que la demanda de cierto artículo en un mes siga una distribución de Poisson con media 20 ¿qué stock se debe tener para satisfacer toda la demanda de un mes con probabilidad de al menos 0'99? (S: 31)

5.20 En cierta gasolinera, la llegada de vehículos por minuto, sigue una distribución de Poisson de parámetro 1'6. Calcúlese

(a) Probabilidad de que en un minuto lleguen más de tres vehículos. (S: 0'0788)

(b) Probabilidad de que en un minuto lleguen entre dos y cinco vehículos (ambos inclusive). (S: 0'4691)

(c) Probabilidad de que en un minuto llegue algún vehículo. (S: 0'7981)

5.21 Supóngase que el número de defectos en un rollo de tela de cierta fábrica tiene una distribucion de Poisson con media 0'4. Si se inspecciona una muestra aleatoria de cinco rollos de tela, ¿cuál es la probabilidad de que el número total de defectos en los cinco rollos sea al menos 6? (S: 0'0166)

5.22 Un cierto tipo de cinta magnética contiene de promedio 3 defectos cada 1000 metros. Hallar la probabilidad de que un cinta de 1500 metros, no contenga defectos. (S: 0'0111)

5.23 Si el número de fallos en el tiempo de un cierto componente electrónico sigue un proceso de Poisson, y en promedio cada componente falla una vez cada 2000 horas, ¿cúal es la probabilidad de que en 400 horas de funcionamiento un determinado componente falle dos o más veces? (S: 0'0175)

5.24 El número de hombres que llegan a un comercio sigue un proceso de Poisson, a razón media de 1 por minuto. El número de mujeres que llegan al mismo comercio sigue otro proceso de Poisson a razón media de 2 por minuto. Suponiendo la independencia de los dos procesos, calcular

(a) Probabilidad de que lleguen menos de 3 clientes en un minuto. (S:0'4232)

(b) Probabilidad de que hayan llegado 5 hombres en 10 minutos, si en ese tiempo han llegado 10 clientes. (S:0'1366)

5.25 En una fábrica, el número de accidentes sigue un proceso de Poisson con media de 2 accidentes por semana. Hallar:

(a) Probabilidad de que en una semana ocurra algún accidente. (S:0'8647)

(b) Probabilidad de que en 2 semanas ocurran 4 accidentes. (S:0'1953)

(c) Es lunes y ya ha ocurrido un accidente. Probabilidad de que en esta semana no ocurran más de 3 accidentes. (S:0'8347)

Normal

5.26 Siendo $Z \sim N(0,1)$, calcular

(a) $\mathrm{P}\left(Z > 0'37\right)$ (S: 0'3557)

(b) $\mathrm{P}\left(Z < -1'12\right)$ (S: 0'1315)

(c) $\mathrm{P}\left(0'21 < Z < 0'82\right)$ (S: 0'2107)

(d) $\mathrm{P}\left(Z < 0'123\right)$ (S: 0'5490)

5.27 Si $X \sim N(55, 12)$ calcúlese

(a) $\mathrm{P}\left(X \leq 58\right)$ (S: 0'5987)

(b) $\mathrm{P}\left(X > 50'8\right)$ (S: 0'6368)

(c) $\mathrm{P}\left(49 < X < 61\right)$ (S: 0'383)

5.28 Sabiendo que $X \sim B(1500, 0'002)$, calcúlese, $\mathrm{P}\left(X = 2\right)$ y $\mathrm{P}\left(X \leq 2\right)$ usando

(a) Valores exactos (S: 0'2241, 0'4230)

(b) Aproximación de Poisson (S: 0'2240, 0'4232)

(c) Aproximación por Normal (S: 0'1910, 0'3859)

5.29 Si $X \sim N(\mu, \sigma)$, calcula $\mathrm{P}\left(|X - \mu| < \sigma\right)$ y $\mathrm{P}\left(|X - \mu| < 2\sigma\right)$ (S: 0'6826 y 0'9544)

5.30 Sea $X \sim N(\mu, \sigma)$.

(a) Determinar el valor de $k > 0$ tal que $\mathrm{P}\left(\mu - k\sigma < X < \mu + k\sigma\right) = 0'95$ (S: 1'96)

(b) Determinar el valor de $h > 0$ tal que $\mathrm{P}\left(X > \mu - h\sigma\right) = 0'99$ (S: 2'33)

5.31 Sabiendo que hay un 1 % de zurdos en una determinada población, calcular la probabilidad de encontrar al menos cuatro zurdos entre 200 personas utilizando: valores exactos, aproximación por Poisson y por la normal. (S: 0'1420, 0'1429,0'1423)

5.32 Encontrar una aproximación a la probabilidad de que el número de *unos* obtenido en 12.000 lanzamientos de un dado esté comprendido entre 1.950 y 2.050 (S: 0'785)

5.33 Hallar la probabilidad de que entre 100.000 cifras al azar, la cifra 6 salga al menos 9.971 veces. (S: 0'6217)

5.34 El cociente de inteligencia es una variable aleatoria que se distribuye según una *normal* $N(100, 16)$. Calcular

(a) Probabilidad de que un individuo elegido al azar, tenga un cociente inferior a 120. (S: 0'8944)

(b) Idem entre 118 y 122. (S: 0'0461)

(c) Suponiendo que una persona que estudia Informática tiene un cociente superior a 110, hallar la probabilidad de que un ingeniero informático tenga un cociente superior a 120. (S: 0'3971)

5.35 Dadas 15 variables aleatorias independientes X_i, distribuidas $N(i, i)$, se define la variable X

$$X = \sum_{i=1}^{15} \frac{X_i}{i}$$

Calcular a para que $\mathrm{P}\left(X \leq a\right) = 0'23$ (S: 12'15)

5.36 Si se selecciona una muestra aleatoria de 25 observaciones de una distribución normal con media m y desviación típica 2, ¿cuál es la probabilidad de que la media muestral difiera (en valor absoluto) de m menos de una unidad? (S: 0'9876)

5.37 En la distribución del problema anterior, hállese el mínimo n para que

$$\mathrm{P}\left(\left|\overline{X}_n - m\right| < 0'1\right) \geq 0'9. \quad \text{(Sol: 1089)}$$

5.38 En un cierto examen, las notas de los estudiantes de la universidad A se distribuye normalmente con media 6'25 y desviación típica 1. Las notas de los estudiantes de la universidad B se distribuyen con media 6 y desviación 1'25. Dos estudiantes de A y tres de B hacen dicho examen; calcúlese la probabilidad de que la media de las notas de los estudiantes de A sea mayor que la media de los estudiantes de B. (S: 0'5987)

5.39 Los errores aleatorios que se cometen en las pesadas de una balanza, siguen una normal de media 0 y desviación típica 2 gramos. Hállese

(a) Error máximo en una pesada con probabilidad 0'95. (S: 3'92 gr.)

(b) Idem en 10 pesadas y tomando el promedio. (S: 1'24 gr.)

5.40 Una máquina fabrica tornillos cuyas longitudes se distribuyen *normalmente* con media 20 mm y desviación 0'25. Un tornillo se considera defectuoso si su longitud no está entre 19'5 y 20'5 mm. Los tornillos se fabrican de forma independiente.

(a) Probabilidad de fabricar un tornillo defectuoso. (S: 0'0456)

(b) Probabilidad de que en los diez primeros tornillos, no haya más de dos defectuosos. (S: 0'9911)

5.41 Los bolígrafos de cierta marca son empaquetados en cajas de 50 por una máquina automática. El peso de los bolígrafos se distribuye *normalmente* con media 5 gramos y desviación 0'4. El peso de la caja se distribuye *normalmente* con media 10 gramos y desviación 0'5. Una caja se considera que contiene 50 bolígrafos, si su peso está entre 255 y 265 gramos. Calcular la probabilidad de que

(a) Una caja con 50 bolígrafos, no sea considerada como tal. (S: 0'0818)

(b) Una caja con 52 bolígrafos sea considerada como si tuviera 50. (S: 0'0436)

5.42 La demanda aleatoria de gasolina durante un cierto período de tiempo, se comporta de arreglo a la ley *normal* de media 150.000 litros y desviación típica igual a 10.000 litros. Determinar la cantidad mínima que hay que tener dispuesta a la venta en dicho período, para poder satisfacer la demanda con una probabilidad de al menos 0'95. (S: 166.500)

5.43 Supongamos que el peso de una persona en kg. sigue una distribución normal de media 75 kg. y de desviación típica 10 kg. ¿Cuál es la probabilidad de que al subir 4 personas en un ascensor, su peso supere la carga máxima del mismo que es de 350 kg? (S: 0'0062)

5.44 La distribución de las estaturas de 1000 estudiantes tiene una media de 1'75 m. y una desviación típica de 0'16 m. Si estas estaturas siguen una distribución normal ¿cuántos estudiantes miden 1'80 m? ¿y más de 1'90 m? (S: 23 y 166)

5.45 El número de erratas por página de un libro sigue una distribución de Poisson de media 1. Si el libro tiene 800 páginas ¿cuál es la probabilidad de que el número total de erratas esté entre 750 y 850?

Solución: Llamando $X_i = \{$ número de erratas de la página $-i\}$, el número total de erratas es

$$X = X_1 + X_1 + X_2 + X_1 + \cdots + X_{800}, \quad \mathrm{E}\,(X) = 800, \quad \mathrm{Var}\,(X) = 800$$

Usando el teorema central del límite

$$\frac{X-800}{\sqrt{800}} \approx N(0,1)$$

Entonces (sin usar la corrección por continuidad)

$$\begin{aligned}
\mathrm{P}\left(750 \leq X \leq 850\right) &= \mathrm{P}\left(\frac{750-800}{\sqrt{800}} \leq Z \leq \frac{850-800}{\sqrt{800}}\right) \\
&= \mathrm{P}\left(-1'77 \leq Z \leq 1'77\right) \\
&= \Phi(1'77) - \Phi(-1'77) \\
&= 2\Phi(1'77) - 1 \\
&= 2 \cdot 0'9616 - 1 = 0'9232
\end{aligned}$$

Problemas de examen

5.46 El peso en Kg. de las personas adultas de cierta población tiene una distribución normal con media 65 y desviación 10. Cinco personas entran en un ascensor de carga máxima 320 Kg ¿Cuál es la probabilidad de que no superen ésta carga máxima?

Solución: Sean X_1, X_2, X_3, X_4, X_5 los pesos de las 5 personas. El peso total T es la suma, que suponiendo independencia, tiene una distribución normal de media 325 y desviación $\sqrt{500}$

$$T = X_1 + X_2 + X_3 + X_4 + X_5 \sim N(325, \sqrt{500}).$$

Por tanto

$$\begin{aligned}
\mathrm{P}\left(T \leq 320\right) &= \mathrm{P}\left(Z \leq \frac{320-325}{\sqrt{500}}\right) \\
&= \mathrm{P}\left(Z \leq -0'22\right) \\
&= \Phi(-0'22) = 1 - \Phi(0'22) \\
&= 1 - 0'5871 = 0'4129.
\end{aligned}$$

5.47 El diámetro de las manzanas de cada partida que llega a un almacén se distribuye normalmente. Se sabe que el 10 % de una partida tiene un diámetro inferior a 40 mm y que un 15 % de la misma rebasa los 80 mm. Calcular el porcentaje de manzanas de la partida con un diámetro entre 60 y 70 mm.

5.48 Una máquina fabrica tornillos cuyas longitudes se distribuyen normalmente con media 20 mm y desviación típica 0.25 mm. Un tornillo se considera defectuoso si su longitud no está entre 19,5 y 20,5 mm. Los tornillos se fabrican de forma independiente.

(a) Probabilidad de fabricar un tornillo defectuoso.

(b) Probabilidad de que en los diez primeros tornillos no haya más de 2 defectuosos.

5.49 Encontrar el mínimo número natural $k \geq 490$, tal que la probabilidad de que el número de caras obtenida en 1000 lanzamientos de una moneda esté comprendido entre 490 y k sea mayor que 0'5.

Solución: La variable X que mide el número de caras en 1000 lanzamientos tiene una distribución $B(1000, 0'5)$ o bien, $N(500, \sqrt{250})$. Se tiene

$$
\begin{aligned}
&\mathrm{P}\left(490 \leq X \leq k\right) > 0'5 \\
&\mathrm{P}\left(\frac{490-500}{\sqrt{250}} \leq Z \leq \frac{k-500}{\sqrt{250}}\right) > 0'5 \\
&\Phi\left(\frac{k-500}{\sqrt{250}}\right) - \Phi(-0'6325) > 0'5 \\
&\Phi\left(\frac{k-500}{\sqrt{250}}\right) > 0'5 + 1 - 0'7357 = 0'7643 \\
&\frac{k-500}{\sqrt{250}} > 0'73 \\
&k > 511{,}54 \rightarrow k = 512.
\end{aligned}
$$

5.50 Calcúlese el número mínimo de lanzamientos de una moneda para que la probabilidad de obtener al menos dos caras sea mayor que 0'9 y calcúlese esta probabilidad de forma exacta.

Solución: Sea $X = \{\text{número de caras en } n \text{ lanzamientos}\}$. Se pide

$$\mathrm{P}\left(X \geq 2\right) > 0'9 \rightarrow \mathrm{P}\left(X \leq 1\right) < 0'1 \rightarrow F(1) < 0'1$$

Puesto que X se distribuye binomialmente $B(n, \frac{1}{2})$, buscando en las tablas, el mínimo n que se obtiene es $n = 7$ y la probabilidad exacta es

$$\mathrm{P}\left(X \geq 2\right) = 1 - \mathrm{P}\left(X \leq 1\right) = 1 - \left(\frac{1}{2}\right)^7 - 7\left(\frac{1}{2}\right)^7 = \frac{120}{128} = 0'9375.$$

5.51 Las alturas X en cm. de los hombres de una cierta población siguen una distribución $N(170, 7)$ y la de las mujeres $Y \sim N(160, 6)$. Se eligen al azar un hombre y una mujer. Probabilidad de que el hombre sea mas alto que la mujer.

Solución: Se busca $\mathrm{P}(X > Y) = \mathrm{P}(X - Y > 0)$. Supuesta la independencia de las variables:

$$X - Y \sim N(10, \sqrt{6^2 + 7^2}) = N(10, \sqrt{85})$$

Por tanto

$$\begin{aligned} \mathrm{P}(X - Y > 0) &= \mathrm{P}\left(Z > \frac{-10}{\sqrt{85}}\right) \\ &= \mathrm{P}\left(Z > -1'08\right) \\ &= \mathrm{P}\left(Z < 1'08\right) \\ &= 0'8599. \end{aligned}$$

5.52 Se supone que la longitud X en centímetros de los tornillos fabricados por cierta máquina sigue una distribución $N(20, \sigma)$. Si el 75 % de la producción tienen una longitud entre 18 y 22 cm, hallar el porcentaje de tornillos cuya longitud es menor que 19 cm.

Solución: Se sabe que

$$0'75 = \mathrm{P}(18 \leq X \leq 22) = \mathrm{P}\left(\frac{-2}{\sigma} \leq Z \leq \frac{2}{\sigma}\right) = 2\Phi\left(\frac{2}{\sigma}\right) - 1$$

De donde $\Phi\left(\frac{2}{\sigma}\right) = 0'875$ y por tanto $\frac{2}{\sigma} = 1'15$. Despejando se obtiene el valor $\sigma = 1'7391$. Para terminar

$$\begin{aligned} \mathrm{P}(X < 19) &= \mathrm{P}\left(Z < \frac{19 - 20}{1'7391}\right) \\ &= \Phi(-0'575) \\ &= 1 - \Phi(0'575) \\ &= 0'2843. \end{aligned}$$

5.53 Se sabe que la probabilidad de que un ordenador de la E.P.S. tenga un virus es de 0'15.

(a) Si un laboratorio tiene 10 ordenadores, hallar la probabilidad de que haya más de un ordenador con virus.

(b) ¿Cuál ha de ser el número máximo de ordenadores instalados en un laboratorio para que la probabilidad de que haya menos de dos ordenadores con virus sea de al menos 0.65?

(c) Suponiendo que en toda la Politécnica haya 100 ordenadores, calcúlese la probabilidad de que el número de ordenadores infectados esté entre 10 y 18.

5.54 Una Facultad recibe solicitudes de ingreso para el siguiente curso. Los aspirantes se someten a pruebas de selección puntuadas de 0 a 1000 siguiendo las calificaciones una variable X normal: $N(550, 100)$. Se sabe que hay 350 personas con puntuaciones entre 400 y 450.

(a) ¿Cuántas personas han solicitado el ingreso?

(b) Si la Facultad decide admitir al 25 % de las aspirantes (los de puntuación mas alta), ¿cúal es la mínima calificación necesaria para ser admitido?

(c) ¿Cuántas personas han obtenido entre 620 y 740 puntos?

Solución:

(a) Puesto que $\mathrm{P}(400 \leq X \leq 450) = 0'0919$, si N es el número de aspirantes, es $N \cdot 0'0919 \approx 350$ por lo que $N = 3808$.

(b) Si llamamos a a la calificación mínima para aprobar, la probabilidad de superar ésta es 0'25.

$$\begin{aligned}
&\mathrm{P}(X \geq a) = 0'25 \\
&\mathrm{P}\left(Z \geq \frac{a-550}{100}\right) = 0'25 \\
&\mathrm{P}\left(Z < \frac{a-550}{100}\right) = 0'75 \\
&\frac{a-550}{100} = 0'67 \\
&a = 617.
\end{aligned}$$

(c) La probabilidad de obtener entre 620 y 740 puntos es

$$\mathrm{P}(620 \leq X \leq 740) = 0'2133.$$

Por tanto la solución es $3808 \cdot 0'2133 \approx 812$.

5.55 Un examen de Estadística es de tipo test. Tiene 100 preguntas y cada una tiene 4 alternativas. El alumno debe contestar eligiendo sólo una. La puntuación viene dada por $X = A - \frac{F}{3}$, siendo A el número de aciertos y F el de fallos. Supongamos que un alumno contesta de forma aleatoria a todas las preguntas

(a) ¿Cuál es la distribución de A?

(b) Hallar la esperanza y varianza de X

(c) Calcular la probabilidad de que el alumno anterior obtenga más de 10 puntos en el examen.

Solución:

(a) $A \sim B(100, 0'25)$. Por tanto $\mathrm{E}(A) = 25$ y $\mathrm{Var}(A) = 18'75$.

(b) Como $F = 100 - A$, entonces $X = A - \frac{100 - A}{3} = \frac{4A - 100}{3}$. La variable X **no** es binomial, pero sus parámetros se pueden obtener en función de los de A.

$$\mathrm{E}(X) = \frac{4\mathrm{E}(A) - 100}{3} = \frac{0}{3} = 0, \quad \mathrm{Var}(X) = \frac{16}{9} \cdot \mathrm{Var}(A) = 33'33.$$

(c) Para que $X \geq 10$ ha de ser $A \geq 37'5$ o sea, $A \geq 38$. Como las tablas no sirven, tipificamos

$$\begin{aligned}
\mathrm{P}(A \geq 38) &= \mathrm{P}\left(Z \geq \frac{38 - 25}{\sqrt{18'75}}\right) \\
&= \mathrm{P}(Z \geq 3) \\
&= 1 - \Phi(3) \\
&= 0'0013.
\end{aligned}$$

5.56 Sea X una variable aleatoria que se distribuye $N(10, 2)$

(a) Tomando una muestra aleatoria de tamaño 4, X_1, X_2, X_3, X_4, consideremos la media muestral $\overline{X_4} = \frac{1}{4}(X_1 + X_2 + X_3 + X_4)$. Calcúlese $\mathrm{P}\left(\left|\overline{X_4} - 10\right| < 1\right)$.

(b) Hallar el mínimo n para que $\mathrm{P}\left(\left|\overline{X_n} - 10\right| < 0'1\right) \geq 0'99$

5.57 La media de los pesos de 500 estudiantes es de $68'5$ kg. y la desviación típica de $6'8$ kg. Suponiendo que la distribución de pesos X es normal, calcular cuántos estudiantes pesan:

(a) Entre 54 y 70 kg.

(b) Más de 84 kg.

Solución:

(a)

$$\begin{aligned} \mathrm{P}\left(54 \leq X \leq 70\right) &= \mathrm{P}\left(\frac{54-68'5}{6'8} \leq Z \leq \frac{70-68'5}{6'8}\right) \\ &= \mathrm{P}\left(-2'13 \leq Z \leq 0'22\right) \\ &= \Phi(0'22) - \Phi(-2'13) \\ &= \Phi(0'22) + \Phi(2'13) - 1 \\ &= 0'5871 + 0'9834 - 1 = 0'5705. \end{aligned}$$

El $57'05\,\%$ de los 500 cumplen los requisitos; esto es: 285.

(b)

$$\begin{aligned} \mathrm{P}\left(X > 84\right) &= \mathrm{P}\left(Z > \frac{84-68'5}{6'8}\right) \\ &= \mathrm{P}\left(Z > 2'28\right) \\ &= 1 - \Phi(2'28) \\ &= 1 - 0'9887 = 0'0113. \end{aligned}$$

Lo que supone, aproximadamente, 6 estudiantes.

5.58 Se sabe que el $95\,\%$ de las piezas fabricadas por cierta máquina son correctas. De cada 5 defectuosas, 2 son demasiado grandes y 3 demasiado pequeñas. Si suponemos que el tamaño sigue una distribución $N(10,1)$, se pide:

(a) ¿Cuál es el tamaño máximo y mínimo de una pieza correcta?

(b) Probabilidad de que en un lote de 300 piezas, sean admisibles más de 290.

(c) Probabilidad de que en un lote de 200 piezas, todas sean correctas.

5.59 Se han utilizado dos tipos de pruebas A y B para medir los conocimientos sobre cierta materia en una misma población. Los resultados en ambas tienen una distribución normal y la media de los resultados en la prueba A es $78'3$ con una desviación típica de $4'2$ puntos; en la prueba B la media es $85'1$ y la desviación $3'2$. Una persona ha obtenido $83'1$ puntos en la prueba A, y otra ha conseguido $87'5$ en la prueba B. ¿Cuál está en mejor posición? es decir, ¿cuál tiene por detrás mayor porcentaje de notas?

5.60 El peso X en Kg. de un determinado tipo de peces sigue una distribución normal com media 2'5 y varianza 0'25. Si el peso del pez es inferior a 2 Kg no se puede comercializar.

(a) Se pescan 20 peces ¿cuál es la probabilidad de que menos de tres no se comercialicen?

(b) Contestar a la misma pregunta para 50 peces.

5.61 Supóngase que el voltaje medido en cierto circuito eléctrico tiene una distribución normal con media 120 y desviación 2. Se toman tres medidas independientes del voltaje.

(a) Probabilidad de que las tres medidas estén entre 116 y 118.

(b) Probabilidad de que la media aritmética de las tres medidas esté entre 116 y 118.

5.62 Un hospital tiene asignada una población de 100.000 habitantes. La probabilidad de que una persona deba ser hospitalizada es de $0'001$. ¿Cuántas camas debe tener el hospital para que la probabilidad de que falten sea inferior a $0'01$?

Solución: El número de pacientes sigue una distribución $B(100000, 0'001)$ que se puede considerar $N(100, \sqrt{99'9})$. Llamamos a al número de camas disponibles

$$\begin{aligned}
\mathrm{P}\left(X > a\right) &= \mathrm{P}\left(Z > \frac{a-100}{\sqrt{99'9}}\right) < 0'1 \rightarrow \mathrm{P}\left(Z \leq \frac{a-100}{\sqrt{99'9}}\right) > 0'99 \\
&\rightarrow \Phi\left(\frac{a-100}{\sqrt{99'9}}\right) > 0'99 \rightarrow \frac{a-100}{\sqrt{99'9}} \geq 2'33 \\
&\rightarrow a \geq 123'29.
\end{aligned}$$

Por lo que tiene que haber al menos 124 camas.

5.63 Los resultados de un examen tienen una distribución $N(78, 6)$.

(a) ¿Cuál es la probabilidad de que una persona que se presenta al examen obtenga una calificación mayor que 72?

(b) Si el examinador pretende aprobar sólo al $28'1\,\%$ de los estudiantes ¿cuál es la nota mínima para aprobar?

(c) Si se sabe que la calificación de un estudiante es mayor que 72 ¿cuál es la probabilidad de que sea mayor que 84?

5.64 Calcúlese de forma aproximada la probabilidad de obtener al menos 10 *seises* al lanzar 72 dados.

Solución: La variable X que mide el número de seises es $B\left(72, \frac{1}{6}\right)$ que se puede considerar normal $N(12, \sqrt{10})$

$$\begin{aligned}
\mathrm{P}\,(X \geq 10) &= 1 - \mathrm{P}\,(X < 10) \\
&= 1 - \mathrm{P}\left(X < 9'5\right) \\
&= 1 - \mathrm{P}\left(Z < \frac{9'5 - 12}{\sqrt{10}}\right) \\
&= 1 - \Phi(-0'79) = \Phi(0'79) \\
&= 0'7852.
\end{aligned}$$

5.65 La longitud del cable de cierto conector producido en una fábrica se distribuye normalmente. Sabiendo que un $2\,\%$ de los cables son rechazados en el control porque su longitud no alcanza los 450 mm. (mínimo requerido) y que otro $5\,\%$ es rechazado por superar los 550 mm, (máximo requerido), calcular la probabilidad de que un determinado cable mida entre 490 y 510 mm.

5.66 El número de defectos en el tejido de cierta tela tiene una distribución de *Poisson* con una media de 4 por metro cuadrado.

(a) Probabilidad de que una muestra de un metro cuadrado contenga al menos un defecto.

(b) Probabilidad de que una muestra de tres metros cuadrados contenga al menos un defecto.

Solución: Llamando $X = \{$ número de defectos por m$^2\}$ es $X \sim P(4)$; Si $Y = \{$ número de defectos cada 3 m$^2\}$ es $Y \sim P(12)$.

(a) $\mathrm{P}(X \geq 1) = 1 - \mathrm{P}(X = 0) = 1 - 0'0183 = 0'9817$

(b) $\mathrm{P}(Y \geq 1) = 1 - \mathrm{P}(Y = 0)$. Como las tablas no dan suficiente precisión acudimos al cálculo exacto

$$1 - \mathrm{P}(Y = 0) = 1 - \mathrm{e}^{-12}\frac{12^0}{0!} = 1 - \mathrm{e}^{-12} = 1 - 0'0000061442 = 0'9999939.$$

5.67 Los promedios de las calificaciones de una gran población de estudiantes tiene una distribución *normal* de media $2'4$ y desviación típica $0'8$. ¿Que porcentaje de estudiantes tendría un promedio mayor de 3?

Solución: Si X es la calificación

$$\begin{aligned} \mathrm{P}(X > 3) &= 1 - \mathrm{P}(X \leq 3) \\ &= 1 - \mathrm{P}\left(Z \leq \frac{3 - 2'4}{0'8}\right) \\ &= 1 - \Phi(0'75) \\ &= 1 - 0'7734 = 0'2266. \end{aligned}$$

Por tanto superan la nota de 3 un $22'66\,\%$.

5.68 Se observó durante un largo periodo que la cantidad semanal gastada en el mantenimiento y reparaciones en cierta fábrica tiene aproximadamente una distribución *normal* con una media de 400 euros y una desviación de 20 euros.

(a) Si el presupuesto para la próxima semana es de 450 euros, ¿cuál es la probabilidad de que los costos reales sean mayores que la cantidad presupuestada?

(b) ¿Cuál tendría que ser el presupuesto semanal para que esta cantidad solamente se rebasara con probabilidad de $0'1$?

Solución: Llamando X a la cantidad gastada y tipificando $Z = \dfrac{X - 400}{20}$

(a)

$$\mathrm{P}(X > 450) = \mathrm{P}\big(Z > 2'5\big) = 1 - 0'9938 = 0'0062$$

(b) Sea a la cantidad por lo que $\mathrm{P}(X > a) = 0'1 \to \mathrm{P}\left(Z > \dfrac{a-400}{20}\right) = 0'1$ de donde

$$\Phi\left(\frac{a-400}{20}\right) = 0'9 \to \frac{a-400}{20} = 1'28 \to a = 425'6.$$

5.69 La presencia de un cierto antibiótico en un fármaco viene determinado por una variable X. Este fármaco se consigue mediante la unión de otros tres compuestos que también contienen antibiótico y que se distribuyen independientemente de la siguiente manera: $X_1 \sim N(80, 12)$, $X_2 \sim N(120, 15)$ y $X_3 \sim N(96, 9)$. La fórmula del fármaco, unión de los tres compuestos es la siguiente:

$$X = \frac{3X_1 + X_2 + 2X_3}{6}$$

Calcular la probabilidad de que la concentración de antibiótico en el fármaco esté entre 70 y 90

Solución: La variable X es una combinación lineal de normales, luego es normal con parámetros:

(a) $\mathrm{E}(X) = \dfrac{3 \cdot 80 + 120 + 2 \cdot 96}{6} = 92$

(b) $\mathrm{Var}(X) = \dfrac{9 \cdot 144 + 225 + 4 \cdot 81}{36} = 51'25.$

Por tanto $X \sim N(92, 7'16)$. Luego

$$\mathrm{P}(70 \leq X \leq 90) = \mathrm{P}\left(-\frac{22}{7'16} \leq Z \leq -\frac{2}{7'16}\right) = \Phi(3'07) - \Phi(0'28) = 0'3886.$$

5.70 Un juego consiste en lanzar dos dados; el jugador gana 3 euros si la suma de los dos dados es 7 ó 9, y paga uno en otro caso. Tras 200 lanzamientos ¿cuál es la probabilidad de que el saldo de ganancias sea positivo?

Solución: La probabilidad de sacar 7 ó 9 al lanzar dos dados es $\dfrac{5}{18}$. Sea X el número de éxitos en 200 lanzamientos y G la ganancia; se tiene

$$X \sim B\left(200, \frac{5}{18}\right), \qquad G = 3X - (200 - X) = 4X - 200.$$

Para que $G > 0$ ha de ser $X > 50$; por tanto hemos de calcular $\mathrm{P}(X > 50)$. Aproximamos por la normal

$$Z \approx \frac{X - 200 \cdot \dfrac{5}{18}}{\sqrt{200 \cdot \dfrac{5}{18}\dfrac{13}{18}}} = \frac{X - 55'56}{6'33}$$

Por tanto,

$$\begin{aligned} \mathrm{P}(X > 50) = 1 - \mathrm{P}(X < 50) &= 1 - \mathrm{P}\left(Z < \frac{50 - 55'56}{6'33}\right) \\ &= 1 - \Phi(-0'88) \\ &= \Phi(0'88) \\ &= 0'8106. \end{aligned}$$

5.71 La contaminación acústica de los locales de copas en una ciudad es medida en decibelios (dB) por una variable que se distribuye normalmente con media 46 y varianza 144. El límite permitido por la legislación es de 55 dB, a partir de los cuales la sanción va en función de la cantidad sobrepasada del límite, pudiendo llegar al cierre del local si el límite sobrepasa en 20 o más dB. Si se elige un local al azar y ha sido sancionado ¿cuál es la probabilidad de que sea cerrado?

Solución: Llamando X a la variable, se tiene que $X \sim N(46, 12)$. Se pide la probabilidad de un suceso condicionado:

$$\mathrm{P}(X > 75 \mid X > 55) = \frac{\mathrm{P}(X > 75)}{\mathrm{P}(X > 55)}.$$

Tipificamos la variable $\dfrac{X - 46}{12} \sim N(0,1)$ y, los cálculos son:

$$\frac{\mathrm{P}(X > 75)}{\mathrm{P}(X > 55)} = \frac{\mathrm{P}(Z > 2'42)}{\mathrm{P}(Z > 0'75)} = \frac{1 - \Phi(2'42)}{1 - \Phi(0'75)} = \frac{0'00776}{0'226627} = 0'0342.$$

5.72 El número de colonias de bacterias de cierto tipo en unas muestras de aguas contaminadas tiene una distribución de Poisson con una media de 2 por cm^3.

(a) Si se toman de forma independiente 4 muestras de un cm^3 de esa agua, encuéntrese la probabilidad de que al menos una muestra tenga una o más colonias de bacterias.

(b) ¿Cuántas muestras de un cm^3 deben seleccionarse para tener una probabilidad de $0'95$ (aproximadamente), de encontrar al menos una colonia de bacterias?

Solución:

(a) Sea $Y = \{$ número de colonias en cuatro muestras independientes $\}$. Entonces es $Y = X_1 + X_2 + X_3 + X_4$ donde $X_i \sim P(2)$ por lo que $Y \sim P(8)$; así

$$\mathrm{P}\,(Y \geq 1) = 1 - \mathrm{P}\,(Y = 0) = 1 - 0'0003 = 0'9997.$$

Otra forma de resolver este apartado es la siguiente: el suceso contrario es que ninguna muestra tenga colonias de bacterias; ésta probabilidad para cada muestra es (véase la tabla de $P(2)$) 0'1353. Como son independientes, para 4 muestras la probabilidad es $0'1353^4 = 0'0003$ y por tanto la probabilidad pedida es $1 - 0'0003 = 0'9997$.

(b) Llamando Y_n al número de colonias en n muestras independiente, se tendrá que $Y \sim P(2n)$, y de la relación $\mathrm{P}\,(Y_n \geq 1) \approx 0'95$ se deduce que $\mathrm{P}\,(Y_n = 0) \approx 0'05$. Por tablas (o cálculos) se deduce que $2n \approx 3$ por lo que se deben tomar 2 muestras.

5.73 La probabilidad de que una persona se manifieste de forma violenta en un espectáculo deportivo es de $0'001$. Si en un estadio de fúbol hay 80,000 personas, de las cuales se comportan de forma violenta al menos 75, ¿cuál es la probabilidad de que esa cifra no pase de 90?

Solución: Sea X la variable que mide el número de personas que se manifiestan de forma violenta; entonces $X \sim B\,(80000, 0'001)$. El problema pide

$$\mathrm{P}\,(X \leq 90 \mid X \geq 75\,) = \frac{\mathrm{P}\,(75 \leq X \leq 90)}{\mathrm{P}\,(X \geq 75)}$$

Como no podemos trabajar con la distribución binomial porque en las tablas no aparece, lo que hacemos es aproximarlo por la distribución normal con parámetros $\mu = 80$ y $\sigma = 8'94$.

$$Z = \frac{X - 80}{8'94} \sim N\,(0, 1)$$

Los cálculos que hay que hacer son

$$\begin{aligned}
\mathrm{P}\left(75 \leq X \leq 90\right) &= \mathrm{P}\left(\frac{75-80}{8'94} \leq Z \leq \frac{90-80}{8'94}\right) \\
&= \Phi\left(\frac{90-80}{8'94}\right) - \Phi\left(\frac{75-80}{8'94}\right) \\
&= \Phi(1'12) - \Phi(-0'56) \\
&= = \Phi(1'12) - 1 + \Phi(0'56) \\
&= 0'8665 - 1 + 0'7123 \\
&= 0'5788.
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}\left(X \geq 75\right) = \mathrm{P}\left(Z \geq \frac{75-80}{8'94}\right) &= 1 - \mathrm{P}\left(Z \leq \frac{-5}{8'94}\right) \\
&= 1 - \mathrm{P}\left(Z \leq -0'56\right) \\
&= 1 - \Phi(-0'56) \\
&= = \Phi(0'56) \\
&= 0'7123.
\end{aligned}$$

Por lo tanto, la probabilidad pedida es:

$$\frac{\mathrm{P}\left(75 \leq X \leq 90\right)}{\mathrm{P}\left(X \geq 75\right)} = \frac{0'5788}{0'7123} = 0'8126.$$

5.74 Supongamos que el número de mensajes que entran en un canal de comunicación en un intervalo de t segundos es una variable de Poisson con media $0'3 \cdot t$. Calcular las probabilidades de los siguientes sucesos:

(a) En un intervalo de 10 segundos llegan exactamente 3 mensajes.

(b) En un intervalo de 20 segundos llegan a lo sumo 20 mensajes.

(c) El número de mensajes que llegan en un intervalo de 5 segundos esté comprendido entre 3 y 7 (ambos incluidos).

Solución: Las variables X_1, X_2 y X_3 que miden el número de mensajes en 10, 20 y 5 segundos tienen distribuciones de Poisson de parámetros 3, 6 y $1'5$ respectivamente

(a) $\mathrm{P}(X_1 = 3) = F_{X_1}(3) - F_{X_1}(2) = 0'6472 - 0'4232 = 0'2240$

(b) $\mathrm{P}(X_2 \leq 20) = F_{X_2}(20) = 1$

(c) $\mathrm{P}(3 \leq X_3 \leq 7) = F_{X_3}(7) - F_{X_3}(2) = 0'9998 - 0'9394 = 0'0604$

5.75 Una de las pruebas de acceso a la Universidad para mayores de 25 años consiste en un test con 100 preguntas, cada una de las cuales tiene 4 posibles respuestas y sólo una correcta. Para superar esta prueba deben obtenerse, al menos, 30 respuestas correctas. Si una persona contesta al azar:

(a) ¿cuál es el número esperado de respuestas correctas?

(b) ¿qué probabilidad tendrá de superar la prueba?

Solución: Sea $X = \{$número de respuestas correctas $\}$; sigue una distribución binomial $B(100, 0'25)$.

(a) $\mathrm{E}(X) = n \cdot p = 100 \cdot 0'25 = 25$.

(b) Hay que hallar $\mathrm{P}(X \geq 30)$. Aproximando por la normal

$$\mu = 25, \quad \sigma^2 = 100 \cdot 0'25 \cdot 0'75 = 18'75 \rightarrow \sigma = \sqrt{18'75}.$$

Ahora hay que tipificar

$$\begin{aligned} \mathrm{P}(X \geq 30) = 1 - \mathrm{P}(X \leq 29) &= 1 - \mathrm{P}\left(Z < \frac{29 - 25}{\sqrt{18'75}}\right) \\ &= 1 - \Phi(0'92) \\ &= 1 - 0'821214 \\ &= 0'178786. \end{aligned}$$

5.76 En una población, el ayuntamiento pretende dar acceso inalámbrico a Internet a 5000 hogares. La probabilidad de que al solicitar una página web se obtenga una denegación de servicio por saturación de la red, es de $0'01$. Se realiza la primera prueba y observa que ya se han producido 40 denegaciones de servicio ¿Cuál es la probabilidad de que este número no supere las 70?

Solución: Sea $X = \mathrm{n}^{\mathrm{o}}$ de denegaciones de servicio $\sim B(5000, 0'01)$. Se pide

$$\mathrm{P}(X \leq 70 \mid X \geq 40) = \frac{\mathrm{P}(40 \leq X) \leq 70}{\mathrm{P}(X \geq 40)}.$$

Aproximando por la normal $\dfrac{X-50}{7'04} \sim N(0,1)$. Los cálculos son

$$\begin{aligned}
\mathrm{P}\left(X \geq 40\right) = 1 - \mathrm{P}\left(X < 40\right) &= 1 - \mathrm{P}\left(X \leq 39\right) \\
&= 1 - \mathrm{P}\left(Z < \frac{39-50}{7'04}\right) \\
&= 1 - \Phi(-1'56) = 0'9406. \\
\mathrm{P}\left(40 \leq X \leq 70\right) &= \mathrm{P}\left(X \leq 70\right) - \mathrm{P}\left(X \leq 39\right) \\
&= \mathrm{P}\left(Z \leq \frac{70-50}{7'04}\right) - (1 - \Phi(1'56)) \\
&= \Phi(2'84) - 1 + \Phi(1'56) \\
&= 0'9383.
\end{aligned}$$

Por tanto la solución es $\dfrac{0'9383}{0'9406} = 0'9976$.

5.77 Una empresa fabrica cierto tipo de chips con un promedio del 1 % de defectuosos. Si tomamos una muestra de 500 chips.

(a) Identifíquese la variable

$$X = \{\text{ número de chips defectuosos de la muestra}\}.$$

(b) ¿Cuál es el valor esperado (media) de X?

(c) Probabilidad de que haya dos o más defectuosos.

Solución:

(a) Es una binomial: $X \sim B(500, 0'01)$

(b) $\mathrm{E}\left(X\right) = np = 500 \cdot 0'01 = 5$

(c) $\mathrm{P}\left(X \geq 2\right) = 1 - \mathrm{P}\left(X \leq 1\right) = 1 - F(1)$. Como n es grande no podemos usar las tablas de la binomial. El valor exacto es

$$1 - f(0) - f(1) = 1 - 0'99^{500} - 500 \cdot 0'01 \cdot 0'99^{499} = 0'9537.$$

Aproximación por Poisson
Puesto que $X \approx P(5)$ se obtiene

$$\mathrm{P}\left(X \geq 2\right) = 1 - \mathrm{P}\left(X \leq 1\right) = F(1) = 1 - 0'0404 = 0'9596$$

Aproximación por la normal
Los parámetros son $\mu = np = 5$ y $\sigma = \sqrt{npq} = \sqrt{4'95} = 2'22$

$$\begin{aligned}
\mathrm{P}\left(X \geq 2\right) &= \mathrm{P}\left(X \geq 1'5\right) \\
&= \mathrm{P}\left(Z \geq \frac{1'5-5}{2'22}\right) \\
&= \mathrm{P}\left(Z \geq -1'58\right) \\
&= 1-\mathrm{P}\left(Z < -1'58\right) \\
&= 1-\Phi(-1'58) \\
&= \Phi(1'58) = 0'9429
\end{aligned}$$

Bibliografía

[1] Morris H. DeGroot. *Probabilidad y Estadística*. Addison-Wesley Iberoamericana, 1988

[2] J. Durá y J. López Cuñat. *Fundamentos de Estadística.* Ariel Economía.

[3] Wonnacott & Wonnacott. *Introducción a la Estadística.* Ed. Limusa.

[4] Mendenhall, Scheaffer y Wackerly. *Estadística Matemática con Aplicaciones.* Grupo Editorial Iberoamérica.

[5] José Requena. *Cálculo de probabilidades. Problemas resueltos.* Ed E.C.U.

[6] Lebart, Morineau, Fenelon. *Tratamiento estadístico de datos.* Ed. Marcombo. Boixareu editores.

[7] Daniel Peña. *Estadística, modelos y métodos. Tomo I - Fundamentos.* Ed. Alianza Universidad.

[8] S. Lipschuzt y J Schiller. *Introducción a la Probabilidad y Estadística.* Mc-Graw Hill 2000.

[9] Ruiz-Maya. *Problemas de Estadística.* Ed. AC

[10] C.M. Cuadras. *Problemas de Probabilidad y Estadística. Volumen 1 y 2.* Ed. Colección Laboratorio de Cálculo. PPV.