Variables aleatorias: problemas resueltos

BENITO J. GONZÁLEZ RODRÍGUEZ (bjglez@ull.es)
DOMINGO HERNÁNDEZ ABREU (dhabreu@ull.es)
MATEO M. JIMÉNEZ PAIZ (mjimenez@ull.es)
M. ISABEL MARRERO RODRÍGUEZ (imarrero@ull.es)
ALEJANDRO SANABRIA GARCÍA (asgarcia@ull.es)

Departamento de Análisis Matemático
Universidad de La Laguna

Índice

6. Problemas resueltos

6.1. Variables aleatorias discretas

Ejercicio 6.1. *Se venden* 5000 *billetes para una rifa a* 1 *euro cada uno. Si el único premio del sorteo es de* 1800 *euros, calcular el resultado que debe esperar una persona que compra* 3 *billetes.*

RESOLUCIÓN. Consideramos la variable aleatoria discreta

$$\xi = \text{'cantidad de dinero obtenido en el juego'}.$$

Los posibles valores de ξ son dos:

- Si se gana la rifa, se obtiene un beneficio de $1800-3=1777$ euros. Por la Ley de Laplace, la probabilidad de que ocurra este hecho es de $3/5000$.
- Si no se gana la rifa, resulta una pérdida de 3 euros. Nuevamente por la Ley de Laplace, la probabilidad de que esto ocurra es de $4997/5000$.

Por lo tanto, la distribución de probabilidad para la variable aleatoria ξ será:

$\xi = x_i$	1777	-3
$P(\xi = x_i)$	$3/5000$	$4997/5000$

El resultado que debe esperar una persona que compra 3 billetes es

$$E[\xi] = 1777 \cdot \frac{3}{5000} - 3 \cdot \frac{4997}{5000} = -\frac{9660}{5000} = -\frac{483}{250} = -1.932,$$

lo que interpretamos como que, en promedio, cabe esperar una pérdida de 1.93 euros. □

Ejercicio 6.2. *Una variable aleatoria discreta toma todos los valores enteros entre* 0 *y* 4 *con la siguiente función de densidad:*

X	0	1	2	3	4
$f(x)$	0.3	0.25	0.25	0.1	0.1

Calcular su esperanza y varianza.

RESOLUCIÓN. La esperanza matemática de la variable X viene dada por:

$$E[X] = \sum_{i=0}^{4} x_i \cdot f(x_i) = (0 \cdot 0.3) + (1 \cdot 0.25) + (2 \cdot 0.25) + (3 \cdot 0.1) + (4 \cdot 0.1) = 1.45.$$

La varianza viene dada por Var[X]=$E[X^2] - (E[X])^2$. Calculamos la esperanza de la variable X^2:

$$E[X^2] = \sum_{i=0}^{4} x_i^2 \cdot f(x_i^2) = \sum_{i=0}^{4} x_i^2 \cdot f(x_i) = (0^2 \cdot 0.3) + (1^2 \cdot 0.25) + (2^2 \cdot 0.25) + (3^2 \cdot 0.1) + (4^2 \cdot 0.1) = 3.75.$$

Por tanto,

$$\mathrm{Var}[X] = 3.75 - 1.45^2 = 1.6475.$$

□

6.1.1. Distribución binomial

Ejercicio 6.3. *Suponiendo que es equiprobable el tener hijo o hija, determinar el número esperado de varones en una familia de ocho hijos, así como la probabilidad de que efectivamente resulte este número.*

RESOLUCIÓN. Consideramos la variable aleatoria

$$\xi = \text{'número de hijos varones en el total de ocho'}.$$

Al ser equiprobable el tener hijo o hija, la probabilidad de éxito (el tener un hijo) es $p = 0.5$. Además, el sexo de un hijo no influye en el de los restantes. Por tanto, la variable ξ sigue una distribución binomial de parámetros $n = 8$ y $p = 0.5$, y su función de densidad es

$$f(x) = \binom{8}{x} \cdot 0.5^x \cdot (1-0.5)^{8-x} = \binom{8}{x} \cdot 0.5^x \cdot 0.5^{8-x} = \binom{8}{x} \cdot 0.5^8 \quad (x \in \{0, 1, 2, \ldots, 8\}).$$

El número esperado de hijos varones viene dado por la esperanza matemática de la variable aleatoria ξ, esto es:

$$E[\xi] = np = 8 \cdot 0.5 = 4 \text{ hijos varones.}$$

Por otra parte, la probabilidad de que nazcan exactamente 4 hijos varones es:

$$P(\xi = 4) = \binom{8}{4} \cdot 0.5^8 = \frac{8!}{4! \cdot 4!} \cdot 0.5^8 = \frac{8 \cdot 7 \cdot 6 \cdot 5}{4 \cdot 3 \cdot 2} \cdot 0.5^8 = (7 \cdot 5 \cdot 2) \cdot 0.5^8 = 70 \cdot 0.5^8 \simeq 0.2734,$$

lo que supone un 27.34 %. □

6.1.2. Distribución de Poisson

Ejercicio 6.4. *Un laboratorio farmacéutico encarga una encuesta para estimar el consumo de cierto medicamento que elabora, con vistas a controlar su producción. Se sabe que a lo largo de un año cada persona tiene una posibilidad entre mil de necesitar el medicamento, y que el laboratorio podrá vender una media de cuatro mil unidades del producto al año. Se pide hallar:*

a) *Probabilidad de que el número de enfermos no exceda de cuatro por año.*

b) *Número de enfermos esperado por año.*

c) *Probabilidad de que el número de enfermos sea superior a dos por año.*

d) *Probabilidad de que haya doce enfermos por año.*

RESOLUCIÓN. Consideramos la variable aleatoria

$$\xi = \text{‘número de enfermos por año’},$$

la cual, según los datos del problema, responde a un modelo de Poisson en un proceso con parámetro $1/1000$ (número de enfermos al año por cada 1000 medicamentos vendidos). Si la empresa farmacéutica espera vender una media de 4000 unidades anuales, el parámetro λ = ‘cantidad de enfermos esperada al año’ de la correspondiente variable aleatoria de Poisson ξ es:

$$\lambda = \frac{1}{1000} \cdot 4000 = 4 \qquad \left(\frac{1}{1000} = \frac{\lambda}{4000}, \text{ de donde } \lambda = 4 \right).$$

La función de densidad para la variable ξ será:

$$f(x) = P(\xi = x) = \frac{4^x}{x!} e^{-4} \quad (x \in \{0, 1, 2, \dots\}).$$

a) La probabilidad de que el número de enfermos no exceda de cuatro por año es:

$$\begin{aligned} P(\xi \leq 4) &= \sum_{k=0}^{4} f(k) = f(0) + f(1) + f(2) + f(3) + f(4) = \left(\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!} + \frac{4^3}{3!} + \frac{4^4}{4!} \right) e^{-4} \\ &= \left(1 + 4 + 8 + \frac{32}{3} + \frac{32}{3} \right) e^{-4} = \left(\frac{39 + 64}{3} \right) e^{-4} = \frac{103}{3} e^{-4} \simeq 0.6288. \end{aligned}$$

b) Tal y como dijimos anteriormente, el número de enfermos esperado por año viene dado por el parámetro de la variable aleatoria ξ, esto es:

$$E[\xi] = \lambda = 4.$$

c) La probabilidad de que el número de enfermos sea superior a dos por año es:

$$\begin{aligned} P(\xi > 2) &= 1 - P(\xi \leq 2) = 1 - [f(0) + f(1) + f(2)] \\ &= 1 - (1 + 4 + 8)\, e^{-4} = 1 - 13\, e^{-4} \simeq 0.7619. \end{aligned}$$

d) La probabilidad de que haya doce enfermos por año es:

$$P(\xi = 12) = f(12) = \frac{4^{12}}{12!}\, e^{-4} \simeq 0.0006.$$

OBSERVACIÓN. Podríamos haber resuelto el ejercicio considerando la variable aleatoria

$$\xi = \text{‘número de medicamentos adquiridos de un total de 4000’},$$

la cual sigue un modelo binomial de parámetros $n = 4000$ y $p = 1/1000$ (que es la probabilidad que tiene una persona de caer enferma o, equivalentemente, de necesitar el medicamento). Ahora bien, ya que $n > 50$ y $p = 0.001 < 0.1$ (alternativamente, $np = 4 < 5$), la distribución de Poisson de parámetro $\lambda = np = 4$ es una buena aproximación de la binomial anterior. □

Ejercicio 6.5. *El recuento de glóbulos blancos de un individuo sano puede presentar un promedio en valor mínimo de hasta* 6000 *por milímetro cúbico de sangre. Para detectar una deficiencia de glóbulos blancos se determina su número en una gota de sangre de* 0.001 *milímetros cúbicos. ¿Cuántos glóbulos blancos cabe esperar en un individuo sano? ¿Cuánto de raro sería encontrar un máximo de* 2 *glóbulos blancos?*

RESOLUCIÓN. Consideramos la variable aleatoria

$$\xi = \text{‘número de glóbulos blancos en 0.001 mm}^3\text{ de sangre’}.$$

El recuento de glóbulos blancos de un individuo sano por milímetro cúbico de sangre puede ser considerado un proceso de Poisson de parámetro 6000. La variable aleatoria ξ del correspondiente proceso de Poisson tiene por parámetro el número de glóbulos blancos esperados en 0.001 mm^3 de sangre:

$$\lambda = 0.001 \cdot 6000 = 6 \qquad \left(\frac{1}{6000} = \frac{0.001}{\lambda}, \text{ de donde } \lambda = 6 \right).$$

Por consiguiente, la función de densidad de probabilidad para la variable aleatoria de Poisson ξ viene dada por:

$$f(x) = P(\xi = x) = \frac{6^x}{x!} e^{-6} \quad (x \in \{0, 1, 2, \ldots\}).$$

El número de glóbulos blancos esperados en un individuo sano coincide con la esperanza de la variable:

$$E[\xi] = \lambda = 6 \text{ glóbulos blancos.}$$

Por otra parte, la probabilidad de encontrar como máximo 2 glóbulos blancos en 0.001 mm^3 de sangre en un individuo sano será

$$\begin{aligned} P(\xi \leq 2) &= f(0) + f(1) + f(2) = \left(\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} \right) e^{-6} \\ &= (1 + 6 + 18) e^{-6} = 25 e^{-6} \simeq 0.0620. \end{aligned}$$

De este resultado cabe inferir que es bastante improbable que se presente dicha situación, al ser la probabilidad de que ocurra muy cercana a cero. Más concretamente, podemos afirmar que esta situación se da en un 6.2 % de los casos. □

Ejercicio 6.6. *Si X es una distribución de Poisson y se sabe que $P(X = 0) = 0.2$, calcular $P(X > 2)$ y $P(1 \leq X < 4)$.*

RESOLUCIÓN. La variable aleatoria X sigue una distribución de Poisson de parámetro λ, en principio desconocido. Su función de densidad de probabilidad viene dada por:

$$f(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad (x \in \{0, 1, 2, \ldots\}).$$

Sin embargo, sabemos que $P(X = 0) = 0.2$, por lo que

$$f(0) = \frac{\lambda^0}{0!} e^{-\lambda} = e^{-\lambda} = 0.2.$$

Tomando logaritmos en la anterior ecuación resulta

$$\ln e^{-\lambda} = -\lambda \ln e = -\lambda = \ln 0.2,$$

esto es, $\lambda = -\ln 0.2 \simeq 1.609$. Consecuentemente

$$\begin{aligned}
P(X>2) &= 1 - P(X \leq 2) = 1 - [f(0)+f(1)+f(2)] \\
&= 1 - \left(\frac{1.609^0}{0!} + \frac{1.609^1}{1!} + \frac{1.609^2}{2!} \right) e^{\ln 0.2} \\
&= 1 - \left(1 + 1.609 + \frac{1.609^2}{2!} \right) 0.2 \\
&\simeq 0.2191,
\end{aligned}$$

y

$$P(1 \leq X < 4) = f(1)+f(2)+f(3) = \left(1.609 + \frac{1.609^2}{2!} + \frac{1.609^3}{3!} \right) 0.2 \simeq 0.7195.$$

□

6.2. Variables aleatorias continuas

Ejercicio 6.7. *La función de distribución de la variable aleatoria que representa la duración en minutos de una llamada telefónica es*

$$f(x) = \begin{cases} 1 - \frac{2}{3}e^{-2x/3} - \frac{1}{3}e^{-x/3}, & \text{si } x > 0 \\ 0, & \text{si } x \leq 0. \end{cases}$$

Hallar su función de densidad, así como la probabilidad de que una llamada dure entre 3 *y* 6 *minutos.*

RESOLUCIÓN. Sabemos que la función de densidad de probabilidad coincide con la derivada de la función de distribución. Por tanto, la función de densidad será

$$f(x) = F'(x) = \frac{dF(x)}{dx} = \begin{cases} \frac{4}{9}e^{-2x/3} + \frac{1}{9}e^{-x/3}, & \text{si } x > 0 \\ 0, & \text{si } x \leq 0. \end{cases}$$

La probabilidad de que una llamada dure entre 3 y 6 minutos es:

$$P(3 \leq \xi \leq 6) = F(6) - F(3) \simeq 0.1555,$$

donde ξ denota la variable aleatoria que mide la duración de una llamada en minutos. □

6.2.1. Distribución normal

Ejercicio 6.8. *Una confitura puede ser calificada de «almíbar» si contiene entre* 420 *y* 520 *gramos de azúcar por kilo de confitura. Un fabricante comprueba* 200 *botes de confitura de* 1 *kilogramos encontrando que el peso medio de azúcar es de* 465 *gramos, con una desviación típica de* 30 *gramos. Sabiendo que el contenido de azúcar se distribuye normalmente (porque proviene de frutas con un contenido variable de azúcar), calcular el porcentaje de la producción del fabricante que no debe ser etiquetado como almíbar, considerando la muestra como representativa de la producción total.*

RESOLUCIÓN. Sea la variable aleatoria

$$\xi = \text{'contenido de azúcar (gr) en botes de confitura de 1 kg'}.$$

Sabemos que el contenido de azúcar en dichos botes se distribuye normalmente con un peso medio por bote de 465 gr (esto es, $\mu = 465$ gr) y una desviación típica de 30 gr (esto es, $\sigma = 30$ gr). Simbólicamente: $\xi \sim N(465, 30)$.

Considerando la muestra de 200 botes como representativa de la producción total, el porcentaje de producción que puede ser calificado de almíbar es el $P(420 < \xi < 520) \cdot 100\,\%$. Al tipificar la variable ξ resulta que $\xi = 30Z + 465$, donde Z es la normal estándar. Se tiene entonces:

$$\begin{aligned}
P(420 < \xi < 520) &= P\left(\frac{420-465}{30} < Z < \frac{520-465}{30}\right) = P(-1.5 < Z < 1.83) \\
&= P(Z < 1.83) - P(Z < -1.5) = P(Z < 1.83) - P(Z > 1.5) \\
&= P(Z < 1.83) - [1 - P(Z \leq 1.5)] = 0.9664 - (1 - 0.9332) = 0.8996.
\end{aligned}$$

Encontramos así que el 89.96 % de la producción total puede ser calificada de almíbar. El resto, un 10.04 %, no debe ser calificado como tal. □

Ejercicio 6.9. *Se sabe que la concentración media de NH_3 en sangre venosa de individuos normales de la población es de* 110 *microgramos por mililitro, y que la concentración de NH_3 del* 99 % *de los individuos se encuentra entre* 85 *y* 135 *microgramos por mililitro. Se pide calcular la desviación típica de dicha población* *normal y los límites del intervalo que comprende al 70 % de los valores de la misma, así como el porcentaje de población que tiene:*

a) *más de* 135 *microgramos por mililitro;*

b) *menos de* 9 *microgramos por mililitro;*

c) *entre* 90 *y* 125 *microgramos por mililitro;*

d) *entre* 85 *y* 100 *microgramos por mililitro.*

RESOLUCIÓN. Consideramos la variable aleatoria

$$\xi = \text{‘concentración media de } NH_3 \text{ } (\mu\text{gr/mL}) \text{ en sangre venosa de individuos normales’}.$$

Sabemos que ξ se distribuye normalmente, con media $\mu = 110$ μgr/mL. Además, en el 99 % de los individuos se encuentra entre 85 y 135 μgr/mL, esto es:

$$P(85 \leq \xi \leq 135) = 0.99.$$

Para tipificar la variable escribimos $\xi = \sigma Z + 110$, donde σ es la desviación típica de ξ y Z la distribución normal estándar. Ahora,

$$\begin{aligned} P(85 \leq \xi \leq 135) &= P(85 \leq \sigma Z + 110 \leq 135) = P\left(\frac{85-110}{\sigma} \leq Z \leq \frac{135-110}{\sigma}\right) \\ &= P\left(-\frac{25}{\sigma} \leq Z \leq \frac{25}{\sigma}\right) = 0.99. \end{aligned}$$

Si denotamos $a = \dfrac{25}{\sigma}$ resulta que

$$P(-a \leq Z \leq a) = 2P(Z \leq a) - 1 = 0.99,$$

de donde

$$P(Z \leq a) = \frac{0.99+1}{2} = 0.9950.$$

En la tabla de distribución para la normal estándar encontramos que $25/\sigma = a = 2.575$. Consecuentemente, la desviación típica de ξ es

$$\sigma = \frac{25}{2.575} = 9.7 \text{ } \mu\text{gr/mL}.$$

El intervalo simétrico respecto a la media que comprende el 70 % de los individuos de la población tiene extremo inferior $\mu - k\sigma = 110 - 9.7k$ y extremo superior $\mu + k\sigma = 110 + 9.7k$, para cierto valor de k. A fin de determinar tal valor, imponemos que

$$P(\mu - k\sigma \leq \xi \leq \mu + k\sigma) = 2P(Z \leq k) - 1 = 0.7.$$

Entonces $P(Z \leq k) = 1.7/2 = 0.85$. Inspeccionando la tabla de distribución para la normal estándar encontramos que $k = 1.04$, de donde

$$110 - 9.7 \cdot 1.04 \simeq 100, \quad 110 + 9.7 \cdot 1.04 \simeq 120.$$

Se concluye que el 70 % de los individuos de la población poseen una concentración media de NH_3 en sangre venosa comprendida entre 100 y 120 μgr/mL.

Finalmente:

a) El porcentaje de población con concentración superior a 135 μgr/mL es del 0.51 %:

$$\begin{aligned} P(\xi > 135) &= 1 - P(\xi \leq 135) = 1 - P(9.7Z + 110 \leq 135) \\ &= 1 - P\left(Z \leq \frac{135 - 110}{9.7}\right) = 1 - P(Z \leq 2.57) \\ &= 1 - 0.9949 = 0.0051. \end{aligned}$$

b) El porcentaje de población con concentración inferior a 95 μgr/mL es el 6.18 %:

$$\begin{aligned} P(\xi < 95) &= P\left(Z < \frac{95 - 110}{9.7}\right) = P(Z < -1.54) \\ &= P(Z > 1.54) = 1 - P(Z \leq 1.54) \\ &= 1 - 0.9382 = 0.0618. \end{aligned}$$

c) El porcentaje de población con concentración comprendida entre 90 y 125 μgr/mL asciende al 91.85 %:

$$\begin{aligned} P(90 \leq \xi \leq 125) &= P\left(\frac{90 - 110}{9.7} \leq Z \leq \frac{125 - 110}{9.7}\right) = P(-2.06 \leq Z \leq 1.54) \\ &= P(Z \leq 1.54) - P(Z \leq -2.06) = P(Z \leq 1.54) - P(Z \geq 2.06) \\ &= P(Z \leq 1.54) - [1 - P(Z < 2.06)] = 0.9382 - (1 - 0.9803) = 0.9185. \end{aligned}$$

d) El porcentaje de población con concentración comprendida entre 85 y 100 μgr/mL es del 14.64 %:

$$
\begin{aligned}
P(85 \leq \xi \leq 100) &= P(-2.57 \leq Z \leq -1.03) = P(Z \leq -1.03) - P(Z \leq -2.57) \\
&= P(Z \geq 1.03) - P(Z \geq 2.57) = [1 - P(Z < 1.03)] - [1 - P(Z < 2.57)] \\
&= P(Z < 2.57) - P(Z < 1.03) = 0.9949 - 0.8485 = 0.1464.
\end{aligned}
$$

□

Ejercicio 6.10. *Se ha comprobado que la distribución del índice de colesterol para un gran número de personas es la siguiente: inferior a* 165 *centigramos,* 58 %*; comprendido entre* 165 *y* 180 *centigramos,* 38 %*. Se sabe que dicha distribución sigue una ley normal.*

a) *Calcular el valor medio del índice de colesterol y su desviación típica.*

b) *Se admite que las personas cuyo índice es superior a* 183 *centigramos deben ser sometidas a tratamiento. ¿Cuál es el número de personas a tratar en una población de* 100000 *individuos?*

RESOLUCIÓN. Consideremos la variable ξ = 'índice de colesterol (cg)'. Esta variable sigue una distribución normal de media μ y desviación típica σ, desconocidas. Sin embargo, el 58 % de la población tiene índice de colesterol inferior a 165 cg:

$$P(\xi < 165) = 0.58,$$

mientras que un 38 % tiene un índice de colesterol comprendido entre 165 y 180 cg:

$$P(165 \leq \xi \leq 180) = 0.38.$$

Por tanto,

$$P(\xi \leq 180) = P(\xi < 165) + P(165 \leq \xi \leq 180) = 0.58 + 0.38 = 0.96.$$

a) Sabemos que $\xi = \sigma Z + \mu$, donde Z es la normal estándar. Luego:

$$
\begin{aligned}
P(\xi < 165) &= P(\sigma Z + \mu < 165) = P\left(Z < \frac{165 - \mu}{\sigma}\right) = 0.58, \\
P(\xi < 180) &= P\left(Z < \frac{180 - \mu}{\sigma}\right) = 0.96.
\end{aligned}
$$

Denotando $a = \dfrac{165-\mu}{\sigma}$ y $b = \dfrac{180-\mu}{\sigma}$, en la tabla de distribución de probabilidad para la normal estándar encontramos que $a = 0.2$ y $b = 1.75$. Consecuentemente:

$$\begin{aligned} 0.2 = \frac{165-\mu}{\sigma} \quad &\Rightarrow \quad \mu + 0.2\,\sigma = 165, \\ 1.75 = \frac{180-\mu}{\sigma} \quad &\Rightarrow \quad \mu + 1.75\,\sigma = 180. \end{aligned}$$

Restando ambas ecuaciones: $1.55\,\sigma = 15$, de donde $\sigma \simeq 9.7$ cg. Por otra parte,

$$\mu = 165 - 0.2\,\sigma = 165 - 0.2 \cdot 9.7 \simeq 163.1 \text{ cg}.$$

Concluimos así que $\xi \sim N(163.1, 9.7)$.

b) Para obtener el porcentaje de población que ha de ser sometida a tratamiento calculamos

$$\begin{aligned} P(\xi > 183) = 1 - P(\xi \leq 183) &= 1 - P\left(Z \leq \frac{183-163.1}{9.7}\right) \\ &= 1 - P(Z \leq 2.05) = 1 - 0.9798 = 0.0202. \end{aligned}$$

El porcentaje buscado es entonces el 2.02 %. En una población de 100000 individuos, este porcentaje se corresponde con $100000 \cdot 2.02/100 = 2020$ de ellos. □

Ejercicio 6.11. *La anchura X en milímetros de una población de coleópteros sigue una distribución normal $N(\mu, \sigma)$, dándose las siguientes probabilidades: $P(X \leq 12) = 0.77$; $P(X > 7) = 0.84$. Se pide:*

a) *Valores de μ y σ.*

b) *Proporción de individuos con anchura entre 8 y 10 milímetros.*

c) *Calcular x e y tales que $P(X > x) = 0.95$ y $P(X < y) = 0.33$.*

RESOLUCIÓN. Sea la variable aleatoria

$$X = \text{‘anchura de un coleóptero (mm)’}.$$

Sabemos que X sigue una distribución normal para una determinada población de coleópteros. Además $P(X \leq$

$12) = 0.77$ y $P(X > 7) = 0.84$, por lo que

$$P(X \leq 7) = 1 - P(X > 7) = 1 - 0.84 = 0.16.$$

a) Sean μ y σ la media y la desviación típica de la variable X. Se tiene que $X = \sigma Z + \mu$, donde Z es la distribución normal estándar. Por tanto:

$$P(X \leq 7) = P(\sigma Z + \mu \leq 7) = P\left(Z \leq \frac{7-\mu}{\sigma}\right) = 0.16,$$
$$P(X \leq 12) = P(\sigma Z + \mu \leq 12) = P\left(Z \leq \frac{12-\mu}{\sigma}\right) = 0.77.$$

Escribiendo $a = \dfrac{7-\mu}{\sigma}$ y $b = \dfrac{12-\mu}{\sigma}$ e inspeccionando la tabla de la distribución normal tipificada encontramos que $b = 0.74$.

Ahora queremos determinar a. Nótese que este valor ha de ser negativo, por cuanto la probabilidad correspondiente es menor que 0.5. Como

$$P(z \leq a) = P(z \geq |a|) = 1 - P(z < |a|) = 0.16,$$

tenemos

$$P(z < |a|) = 1 - 0.16 = 0.84.$$

Entonces $|a| = 0.99$, de donde $a = -0.99$.

Llegamos así al siguiente sistema de ecuaciones:

$$-0.99 = \frac{7-\mu}{\sigma} \quad \Rightarrow \quad \mu - 0.99\,\sigma = 7,$$
$$0.74 = \frac{12-\mu}{\sigma} \quad \Rightarrow \quad \mu + 0.74\,\sigma = 12.$$

Restamos miembro a miembro las ecuaciones anteriores para obtener $1.73\,\sigma = 5$, y de aquí

$$\sigma = \frac{5}{1.73} \simeq 2.89 \text{ mm}.$$

Sustituimos σ en la primera de las ecuaciones del sistema y despejamos μ:

$$\mu = 7 + 0.99 \cdot 2.89 = 9.86 \text{ mm}.$$

En consecuencia, $X \sim N(9.86, 2.89)$. Escribimos $X = 2.89Z + 9.86$, donde $Z \sim N(0,1)$.

b) La proporción de coleópteros con anchura entre 8 y 10 mm es del 25.88 %, ya que

$$\begin{aligned} P(8 \leq X \leq 10) &= P(8 \leq 2.89Z + 9.86 \leq 10) = P\left(\frac{8-9.86}{2.89} \leq Z \leq \frac{10-9.86}{2.89}\right) \\ &= P(-0.64 \leq Z \leq 0.05) = P(Z \leq 0.05) - P(Z \leq -0.64) \\ &= P(Z \leq 0.05) - P(Z \geq 0.64) = P(Z \leq 0.05) - [1 - P(Z < 0.64)] \\ &= 0.5199 - (1 - 0.7389) = 0.2588. \end{aligned}$$

c) En primer lugar, calculamos la anchura x tal que

$$P(X > x) = 1 - P(X \leq x) = 0.95,$$

o, equivalentemente, $P(X \leq x) = 1 - 0.95 = 0.05$. Tipificando,

$$P(2.89Z + 9.86 \leq x) = P\left(Z \leq \frac{x-9.86}{2.89}\right) = 0.05.$$

Como esta última probabilidad es menor que 0.5, se verifica que $a = \dfrac{x-9.86}{2.89} < 0$. Escribimos entonces

$$P(z \leq a) = P(z \geq |a|) = 1 - P(z < |a|) = 0.05,$$

esto es:

$$P(z < |a|) = 0.95.$$

Utilizando ahora la tabla de distribución para Z vemos que $|a| = 1.64$, de donde $a = -1.64$. Así pues

$$-1.64 = \frac{x-9.86}{2.89},$$

con lo que $x = 9.86 - 1.64 \cdot 2.89 = 5.2$ mm.

Finalmente, calculamos la anchura y tal que $P(X < y) = 0.33$. Tipificando,

$$P(X < y) = P\left(Z < \frac{y-9.86}{2.89}\right) = 0.33.$$

Análogamente que en el caso anterior, al ser esta última probabilidad menor que 0.5, necesariamente $b = \frac{y-9.86}{2.89} < 0$. Como

$$P(z < b) = P(z > |b|) = 1 - P(z \leq |b|) = 0.33,$$

sigue que

$$P(z \leq |b|) = 0.67$$

y, por tanto, $|b| = 0.44$, de donde

$$-0.44 = \frac{y-9.86}{2.89},$$

resultando así el valor $y = 9.86 - 0.44 \cdot 2.89 = 8.59$ mm para la anchura buscada. □

Ejercicio 6.12. *En una investigación sobre los efectos teratogénicos del tabaquismo se estudió una muestra de embarazadas de la cual el 40 % fumaba y el 60 %, no. Cuando nacieron los niños se encontró que 20 de ellos tenían algún tipo de tara de nacimiento. Sea ξ el número de niños cuya madre fumaba durante el embarazo. Si no hay relación entre el hecho de que la madre fumara y los defectos de nacimiento, entonces ξ es una binomial con $n = 20$ y $p = 0.4$. ¿Cuál es la probabilidad de que 12 ó más niños afectados tengan madres que fumaban?*

RESOLUCIÓN. Consideramos la variable aleatoria

$$\xi = \text{‘número de niños cuya madre fumaba durante el embarazo'}.$$

Los datos disponibles indican que la variables ξ es binomial, con parámetros $n = 20$ y $p = 0.4$.

La probabilidad de que 12 ó más niños afectados tengan madres que fumaban viene dada por

$$P(\xi \geq 12) = \sum_{k=12}^{20} \binom{20}{k} \cdot 0.4^k \cdot 0.6^{20-k},$$

o bien

$$\begin{aligned} P(\xi \geq 12) &= 1 - P(\xi < 12) = 1 - P(\xi \leq 11) \\ &= 1 - \sum_{k=0}^{11} \binom{20}{k} \cdot 0.4^k \cdot 0.6^{20-k}. \end{aligned}$$

El elevado número de operaciones que conlleva cualquiera de estos cálculos sugiere aproximar la variable ξ

por una distribución normal. Como $n = 20$ es suficientemente grande y $p = 0.4$ es un valor intermedio, el Teorema de De Moivre garantiza que, efectivamente, ξ puede ser aproximada por una distribución normal X de media $\mu = np = 20 \cdot 0.4 = 8$ y desviación típica $\sigma = \sqrt{np(1-p)} = \sqrt{20 \cdot 0.4 \cdot (1-0.4)} = 2.19$.

Tipificando la nueva variable normal ($X = 2.19Z + 8$, con $Z \sim N(0,1)$), y aplicando además corrección de continuidad, la probabilidad buscada resulta ser:

$$\begin{aligned} P(\xi \geq 12) \simeq P(X \geq 11.5) = 1 - P(X < 11.5) &= 1 - P\left(Z < \frac{11.5-8}{2.19}\right) \\ &= 1 - P(Z < 1.6) = 1 - 0.9452 = 0.0548. \end{aligned}$$

$\square$