14 Distribución normal

EJERCICIOS PROPUESTOS

1 a 3. Ejercicios resueltos.

4. Sea *X* una variable aleatoria continua cuya función de densidad es:

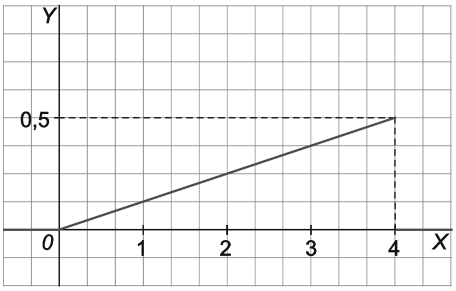
$$f(x)=\begin{cases}\dfrac{x}{8} & \text{si } 0<x<4\\ 0 & \text{en el resto}\end{cases}$$

a) Represéntala gráficamente.

b) Comprueba que es función de densidad.

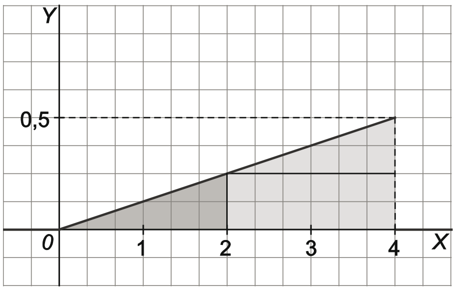
c) Calcula $P(X<2)$; $P(2<X<4)$.

a) La gráfica de esta función es:



b) Se trata de una función de densidad, puesto que cumple las condiciones:
1) $f(x)\geq 0$
2) El área bajo la gráfica de la función es 1 (área del triángulo).

c) En el gráfico adjunto se pueden ver gráficamente las dos probabilidades que se piden: $P(X<2)$ y $P(2<X<4)$, por lo que :



$$P(X<2)=\frac{2\cdot\frac{2}{8}}{2}=\frac{1}{4} \qquad P(2<X<4)=\frac{2\cdot\frac{2}{8}}{2}+2\cdot\frac{2}{8}=\frac{3}{4}$$

5. Utilizando la tabla de la normal estándar, determina σ y p en cada caso:

a) $P(Z \le 0,89) = p$ **c)** $P(Z \ge -1,58) = p$ **e)** $P(-1,97 \le Z \le -0,34) = p$ **g)** $P(Z \le \sigma) = 0,0075$

b) $P(Z \ge 1,74) = p$ **d)** $P(-0,65 < Z < 1,15) = p$ **f)** $P(Z < \sigma) = 0,695$ **h)** $P(-\sigma \le Z \le \sigma) = 0,9676$

a) $p = P(Z \le 0,89) = \Phi(0,89) = 0,8133$

b) $p = P(Z \ge 1,74) = 1 - P(Z < 1,74) = 1 - \Phi(1,74) = 1 - 0,9591 = 0,0409$

c) $p = P(Z \ge -1,58) = P(Z \le 1,58) = \Phi(1,58) = 0,9429$

d) $p = P(-0,65 < Z < 1,15) = P(Z < 1,15) - P(Z < -0,65) = P(Z < 1,15) - P(Z > 0,65) = P(Z < 1,15) - (1 - P(Z < 0,65)) =$
$= P(Z < 1,15) - (1 - P(Z < 0,65)) = \Phi(1,15) - (1 - \Phi(0,65)) = 0,8749 - (1 - 0,7422) = 0,6171$

e) $p = P(-1,97 \le Z \le -0,34) = P(0,34 \le Z \le 1,97) = \Phi(1,97) - \Phi(0,34) = 0,9756 - 0,6331 = 0,3425$

f) $P(Z < \sigma) = 0,695 \Rightarrow \sigma = 0,51$

g) $P(Z \le \sigma) = 0,0075 \Rightarrow P(Z \ge -\sigma) = 0,0075 \Rightarrow 1 - P(Z \le -\sigma) = 0,0075 \Rightarrow 0,9925 = P(Z \le -\sigma)$

$\Rightarrow -\sigma = 2,43 \Rightarrow \sigma = -2,43$

h) $P(-\sigma \le Z \le \sigma) = 0,9676 \Rightarrow P(Z \le \sigma) - P(Z \le -\sigma) = 0,9676 \Rightarrow P(Z \le \sigma) - P(Z \ge \sigma) = 0,9676 \Rightarrow$
$\Rightarrow P(Z \le \sigma) - (1 - P(Z \le \sigma)) = 0,9676 \Rightarrow 2P(Z \le \sigma) = 1,9676 \Rightarrow P(Z \le \sigma) = 0,9838 \Rightarrow \sigma = 2,14$

6 a 8. Ejercicios resueltos.

9. Un fabricante de pilas para linternas asegura que la duración de su producto tiene una distribución normal de media 80 horas de uso con una varianza de 16. Calcula la probabilidad de que una pila elegida al azar dure:

a) Más de 90 horas

b) Entre 70 y 85 horas

Si un comerciante compra un lote de 1000 pilas al fabricante, calcula cuántas pilas tendrán una vida superior a:

c) 100 horas

d) 90 horas

Sea la variable aleatoria *X*: "duración de las pilas (en horas)". El fabricante asegura que $X \sim N(\mu = 80; \sigma = 4)$.
Entonces, las probabilidades que se piden se obtienen tipificando la variable y utilizando las tablas de la distribución normal estándar.

a) $P(X > 90) = P\left(\frac{X-80}{4} > \frac{90-80}{4}\right) = P(Z > 2,5) = 1 - P(Z \le 2,5) = 1 - 0,9938 = 0,0062$

b) $P(70 < X < 85) = P\left(\frac{70-80}{4} < Z < \frac{85-80}{4}\right) = P(-2,5 < Z < 1,25) = P(Z < 1,25) - P(Z < -2,5) =$

$$= P(Z < 1,25) - \left(1 - P(Z < 2,5)\right) = 0,8944 - (1 - 0,9938) = 0,8882$$

c) $P(X > 100) = P\left(Z > \frac{100-80}{4}\right) = P(Z > 5) = 1 - P(Z \le 5) = 0$

No se espera que ninguna tenga una vida superior a 100 horas.

d) Como $P(X > 90) = 0,0062$, se esperará que el 0,62 % de las 1000 pilas, esto es 6,2, superen la 90 horas de duración. Es decir, unas 6 pilas.

10. Una máquina produce tornillos cuya longitud tiene una distribución normal de media 5 cm y desviación típica 5 mm. No se pueden vender los tornillos que se desvíen 6 mm o más de la media. De un lote de 500 tornillos, ¿Cuántos deben ser descartados para la venta?

Sea la variable X: "longitud de los tornillos, en cm". Se sabe que $X \sim N(\mu = 5; \sigma = 0{,}5)$.

Un tornillo es apto para la venta si su longitud está en el intervalo (5 – 0,6; 5 + 0,6) = (4,4; 5,6). La probabilidad de que la longitud de un tornillo elegido al azar esté en este intervalo es

$$P(4{,}4 < X < 5{,}6) = P\left(\frac{4{,}4-5}{0{,}5} < Z < \frac{5{,}6-5}{0{,}5}\right) = P(-1{,}2 < Z < 1{,}2) = 2P(0 < Z < 1{,}2) =$$
$$= 2(\Phi(1{,}2) - \Phi(0)) = 2(0{,}8849 - 0{,}5) = 0{,}7698$$

Por tanto, la probabilidad de que un tornillo elegido al azar no se pueda vender es 1 – 0,7698 = 0,2302. Es decir, el 23,02% serán descartados para la venta.

Si el lote tiene 500 tornillos, el 23,02% de 500 es 115,1.

De modo que entre 115 y 116 tornillos serán descartados para la venta.

11. A una prueba de acceso específica de una universidad se han presentado 2500 aspirantes para 300 plazas. Una vez efectuados los exámenes, las calificaciones tienen una distribución normal de media 6,5 y varianza 4. Calcula la nota de corte para los admitidos.

Sea la variable aleatoria X: "calificaciones de la prueba de acceso". Se tiene que $X \sim N(\mu = 6{,}5;\ \sigma = 2)$.

Si hay 300 plazas para 2500 aspirantes, ello significa que los admitidos serán el $\frac{300}{2500} \cdot 100\ \% = 12\ \%$ de los aspirantes. Por lo que el 88% de los que se presentan será rechazado. Es decir, la nota de corte, c, será aquella calificación tal que $P(X < c) = 0{,}88$.

Tipificando, $P\left(Z < \frac{c-6{,}5}{2}\right) = 0{,}88$.

Y, mediante las tablas de la distribución normal estándar, se tiene que $\frac{c-6{,}5}{2} = 1{,}18 \Rightarrow c = 8{,}86$.

12. Una empresa de teléfonos móviles ha hecho un estudio sobre el tiempo que tardan sus baterías en descargarse, llegando a la conclusión de que dicha duración, en días, tiene una distribución normal de media 2,8 días y desviación típica 1 día. Si se elige al azar un teléfono móvil de esta empresa, calcula la probabilidad de que la duración de su batería:

a) Esté comprendida entre 3,1 y 3,3 días.

b) Sea inferior a 2 días.

Sea la variable aleatoria X: "tiempo que tardan en descargarse las baterías (en días)". Se sabe que $X \sim N(\mu = 2{,}8;\ \sigma = 1)$. Entonces, elegido un teléfono móvil al azar,

a) $P(3{,}1 < X < 3{,}3) = P\left(\frac{3{,}1-2{,}8}{1} < Z < \frac{3{,}3-2{,}8}{1}\right) = P(0{,}3 < Z < 0{,}5) = \Phi(0{,}5) - \Phi(0{,}3) = 0{,}6915 - 0{,}6179 = 0{,}0736$

b) $P(X < 2) = P\left(Z < \frac{2-2{,}8}{1}\right) = P(Z < -0{,}8) = 1 - \Phi(0{,}8) = 1 - 0{,}7881 = 0{,}2119$

13. Ejercicio interactivo.

14 y 15. Ejercicios resueltos.

16. El 40 % de las personas empadronadas en una ciudad vive en urbanizaciones alejadas del centro. De una muestra de 1500 personas, ¿cuál es la probabilidad de que menos de 550 vivan en urbanizaciones?

Sea la variable aleatoria Y: "n.º de personas que vive en urbanizaciones". La variable Y tiene una distribución binomial de parámetros $n = 1500$ y $p = 0,4$.

Como $np = 600$ y $nq = 900$, para el cálculo de la probabilidad pedida, se puede aproximar mediante una distribución normal de media $\mu = np = 1500 \cdot 0,4 = 600$ y varianza $\sigma^2 = npq = 1500 \cdot 0,4 \cdot 0,6 = 360$.

Esto es, $Y \approx X \sim N\left(\mu = 600; \sigma^2 = 360\right)$.

Entonces, aplicando la corrección por continuidad y tipificando:

$$P(Y < 550) \approx P(X \le 549,5) = P\left(Z \le \frac{549,5 - 600}{\sqrt{360}}\right) = P(Z \le -2,66) = 1 - P(Z < 2,66) = 1 - 0,9961 = 0,0039$$

17. En una población, el 45 % de las personas adultas se declara fumadora. Si de la ciudad elegimos una muestra de 250 personas adultas, calcula la probabilidad de que más de la mitad sean fumadoras.

Considera la variable Y: "n.º de personas adultas, de las 250, que se declaran fumadoras". La variable Y tiene una distribución binomial de parámetros $n = 250$ y $p = 0,45$.

Como $np = 112,5$ y $nq = 137,5$, se puede aproximar el cálculo de probabilidades de la variable Y por el de una variable con distribución normal, de media $\mu = 250 \cdot 0,45 = 112,5$ y varianza $\sigma^2 = 250 \cdot 0,45 \cdot 0,55 = 61,875$. Es decir, $Y \approx X \sim N\left(\mu = 112,5;\ \sigma = \sqrt{61,875}\right)$.

Entonces, aplicando la corrección por continuidad y tipificando la variable normal, resulta:

$$P(Y > 125) \approx P(X \ge 125,5) = P\left(Z \ge \frac{125,5 - 112,5}{\sqrt{61,875}}\right) = P(Z \ge 1,65) = 1 - \Phi(1,65) = 1 - 0,9505 = 0,0495$$

18. El primer examen de una oposición es un test que consta de una batería de 100 preguntas, cada una de las cuales tiene 5 posibles respuestas de las que solo una es correcta. Si una persona responde al azar, calcula la probabilidad de que acierte al menos 25 preguntas.

Sea la variable aleatoria Y: "n.º de respuestas correctas de las 100".

La variable Y tiene una distribución binomial $Y \sim \text{Bin}(n = 100;\ p = 0,2)$. Entonces, $np = 20$, $nq = 80$.

En este caso, se puede aproximar el cálculo de probabilidades de la distribución binomial por la de una variable aleatoria normal:

$$Y \approx X \sim N\left(\mu = 100 \cdot 0,2 = 20; \sigma = \sqrt{100 \cdot 0,2 \cdot 0,8} = 4\right)$$

De manera que: $P(Y \ge 25) \approx P(X \ge 24,5) = P\left(Z > \frac{24,5 - 20}{4}\right) = P(Z > 1,13) = 1 - \Phi(1,13) = 1 - 0,8708 = 0,1292$

19. Ejercicio interactivo.

20. Ejercicio resuelto.

21. La tabla recoge las hipotecas concedidas por una entidad en cientos de miles de euros durante el 2012.

Euros	[1; 1,2)	[1,2; 1,4)	[1,4; 1,6)	[1,6; 1,8)	[1,8; 2)
f_i	810	3514	8374	5403	1519

a) Halla la media y la desviación de los datos.

b) Ajusta a una distribución normal a ese conjunto de datos.

c) Valora la bondad del ajuste.

d) ¿Cuál es la probabilidad de que una hipoteca haya sido superior a los 150 000 €?

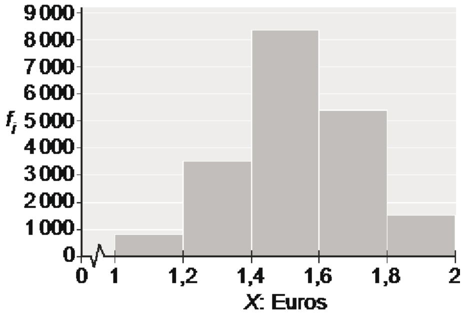
a) Construimos la tabla de frecuencias ampliada:

Clases	f_i	x_i	f_ix_i	$f_ix_i^2$
[1; 1,2)	810	1,1	891	980,1
[1,2; 1,4)	3514	1,3	4568,2	5938,66
[1,4; 1,6)	8374	1,5	12561	18841,5
[1,6; 1,8)	5403	1,7	9185,1	15614,67
[1,8; 2)	1519	1,9	2886,1	5483,59
	19620		30 091,4	46 858,52

$\overline{X} = \dfrac{30\,091{,}4}{19\,620} = 1{,}5337$

$s^2 = \dfrac{46\,858{,}52}{19\,620} - 1{,}5337^2 = 0{,}036 \Rightarrow s = 0{,}190$

Para visualizarlo mejor representamos los datos en un histograma.



b) La distribución de frecuencias de la variable X se aproxima mediante una $N(\mu = 1{,}5337;\ \sigma = 0{,}190)$

c) Calculamos las probabilidades para determinar las frecuencias teóricas.

Probabilidades teóricas	$19\,620P(L_i \le X < L_{i+1})$	Frecuencias teóricas	Frecuencias empíricas
$P(1 \le X < 1{,}2) = P(-2{,}81 \le Z < -1{,}76) = 0{,}0367$	720,05	720	810
$P(1{,}2 \le X < 1{,}4) = P(-1{,}76 \le Z < -0{,}70) = 0{,}2028$	3978,9	3979	3514
$P(1{,}4 \le X < 1{,}6) = P(-0{,}70 \le Z < 0{,}35) = 0{,}3948$	7445,9	7446	8374
$P(1{,}6 \le X < 1{,}8) = P(0{,}35 \le Z < 1{,}40) = 0{,}2883$	5656,4	5656	5403
$P(1{,}8 \le X < 2) = P(1{,}40 \le Z < 2{,}46) = 0{,}0739$	1449,9	1450	1519

La suma de las probabilidades es 0,9965 debido a que las colas de la distribución tienen restos no recogidos en el proceso.
El promedio de la desviación de las frecuencias teóricas respecto de las empíricas es del 8,9 %, lo que indica que la aproximación, no siendo excesivamente mala, no es todo lo fiable que pudiéramos desear.

d) Asumiendo como bueno el ajuste, tenemos que la probabilidad pedida es
$P(X > 1{,}5) \approx P(Z \ge -0{,}18) = P(Z \le 0{,}18) = 0{,}5714$

23 a 29. Ejercicios resueltos.

EJERCICIOS

Variables aleatorias continuas.

30. Considera la siguiente función:

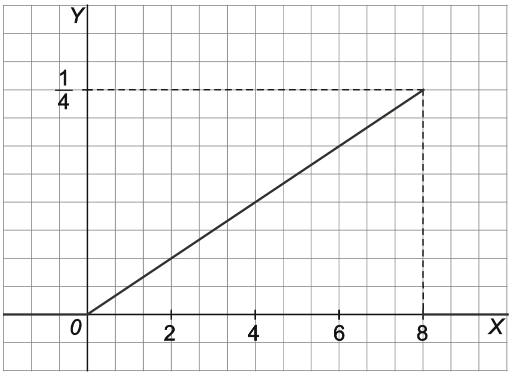
$$f(x)=\begin{cases}\dfrac{x}{32} & \text{si } 0\le x\le 8\\ 0 & \text{en el resto}\end{cases}$$

a) Represéntala gráficamente.

b) Comprueba que es función de densidad de alguna variable aleatoria X.

c) Calcula la probabilidad $P(2<X<9)$.

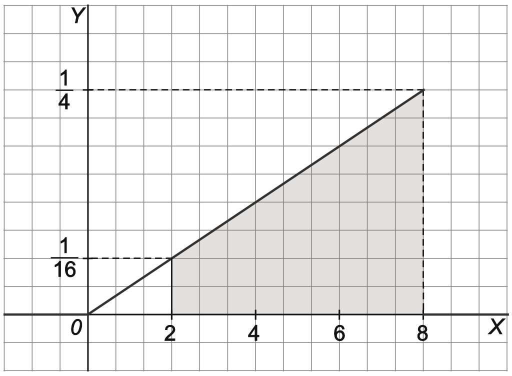
a) La gráfica de la función de densidad es:



b) Se trata, en efecto, de la función de densidad de alguna variable aleatoria, puesto que verifica los dos requisitos:
Los valores de la función son mayores o iguales que cero, $f(x)\ge 0$, cualquiera que sea el valor x.
El área de la región comprendida entre la gráfica de la función, el eje *X*, y la recta vertical $x=8$, es la unidad, ya que: Área $=\frac{1}{2}\left(8\cdot\frac{1}{4}\right)=1$

c) En este caso,

$$P(2<X<9)=P(2<X<8)=1-P(0<X<2)$$
$$=1-\frac{2\cdot\frac{1}{16}}{2}=\frac{15}{16}=0{,}9375$$



31. Sea *X* una variable aleatoria continua con función de densidad:

$$f(x)=\begin{cases}\dfrac{2}{k} & \text{si } -1\le x\le 5\\ 0 & \text{en el resto}\end{cases}$$

a) Halla el valor de k.

b) Representa gráficamente la función de densidad.

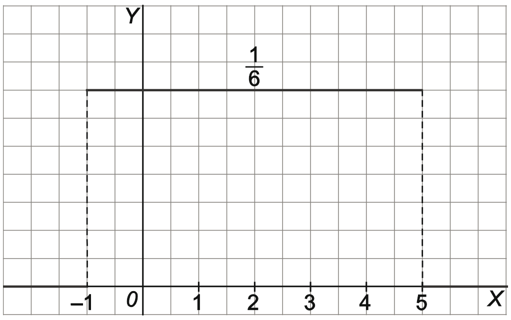
c) Calcula la probabilidad $P(X>1)$ y $P(-0{,}5<X\le 2)$.

a) Para hallar el valor de k, se debe tener en cuenta que el área de la región comprendida entre la gráfica de la función, el eje X y las rectas verticales $x=-1$ y $x=5$, debe ser la unidad. Es decir,

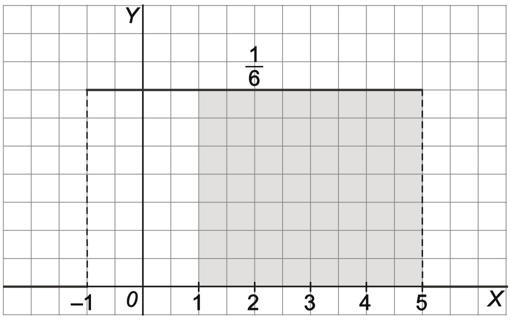
$$\left(5-(-1)\right)\cdot\frac{2}{k}=1\Rightarrow\frac{12}{k}=1\Rightarrow k=12$$

De modo que la función de densidad de la variable aleatoria X, queda: $f(x)=\begin{cases}\dfrac{1}{6} & \text{si } -1\le x\le 5\\ 0 & \text{en el resto}\end{cases}$.

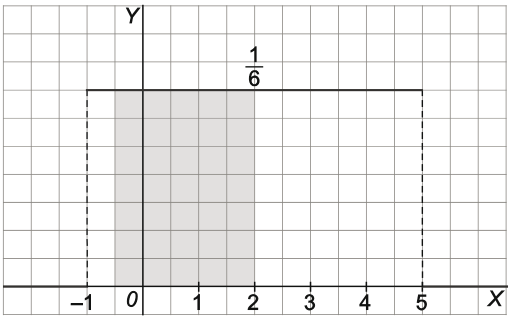
b) La gráfica de la función de densidad es:



c) Las probabilidades se obtienen mediante el área de la región que se indica en cada caso:



$$P(X>1)=(5-1)\cdot\frac{1}{6}=\frac{2}{3}=0{,}66667$$



$$P(-0{,}5<X\le 2)=\left(2-(-0{,}5)\right)\cdot\frac{1}{6}=\frac{5}{12}=0{,}41667$$

32. Considera una variable aleatoria continua *X* con función de densidad:

$$f(x)=\begin{cases} kx+\dfrac{1}{4} & \text{si } 0\le x\le 2 \\ 0 & \text{en el resto}\end{cases}$$

a) Calcula el valor de la constante k.

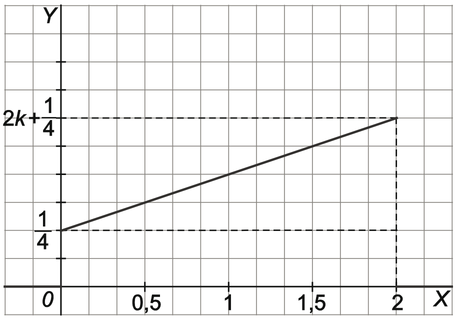
b) Dibuja la gráfica de la función de densidad.

c) Calcula $P(0,5<X\le 1,5)$.

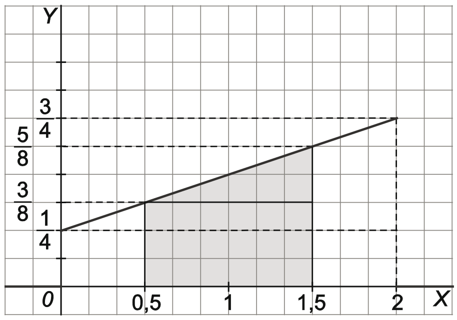
a) Para calcular la constante $k>0$, el área de la región comprendida por la gráfica de la función, el eje X y las rectas verticales $x=0$ y $x=2$, debe ser la unidad.
Como la región mencionada es un trapecio el área se calcula:

$$\frac{\left(2k+\frac{1}{4}+\frac{1}{4}\right)\cdot 2}{2}=1\Rightarrow \frac{1}{2}+2k=1\Rightarrow k=\frac{1}{4}$$

b) La gráfica de la función de densidad de la variable X, tiene la siguiente forma:



c) La gráfica de la función de densidad, con la probabilidad que se pide (coloreada) es:



Y la probabilidad que se pide es la del trapecio de la zona coloreada:

$$P(0,5<X<1,5)=\frac{(1,5-0,5)\left(\frac{5}{8}+\frac{3}{8}\right)}{2}=\frac{1}{2}$$

33. La función de densidad de una variable aleatoria *X* viene dada por:

$$f(x)=\begin{cases} k(x+1) & \text{si } -1\le x\le 5 \\ 0 & \text{en el resto} \end{cases}$$

a) Calcula el valor de la constante *k*.

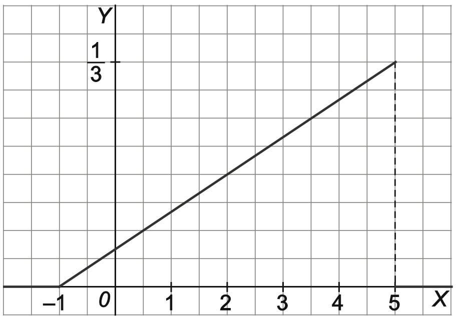
b) Dibuja la gráfica de la función de densidad.

c) Calcula $P(0{,}5<X\le 3)$.

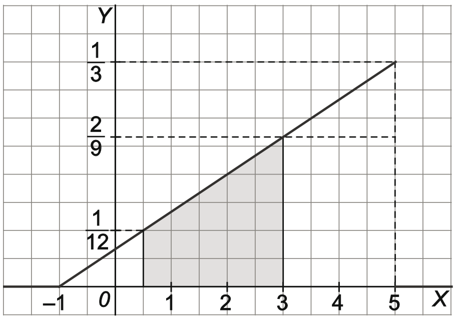
a) Dado que $f(-1)=0$ y $f(5)=6k$, para que sea función de densidad, el área del triángulo que forma la recta con el eje *X* y $x=5$ debe ser

$$\frac{\left(5-(-1)\right)6k}{2}=1\Rightarrow 36k=2\Rightarrow k=\frac{1}{18}$$

b) La gráfica de la función de densidades:



c) La región coloreada del apartado c) es:



La probabilidad pedida es el área de la región coloreada en la gráfica es un trapecio cuya área es:

$$P(0{,}5<X<3)=\frac{(3-0{,}5)\left(\frac{2}{9}+\frac{1}{12}\right)}{2}=\frac{55}{144}=0{,}3819$$

Variables aleatorias con distribución normal.

34. Considera una variable aleatoria $Z \sim N(0, 1)$. Calcula las probabilidades siguientes utilizando las tablas.

a) $P(0,24 \le Z \le 1,73)$ **c)** $P(Z > 2,12)$ **e)** $P(-1,88 \le Z \le -0,67)$ **g)** $P(|Z| \ge 0,7)$

b) $P(Z < -1,12)$ **d)** $P(Z \ge -0,38)$ **f)** $P(-1,64 \le Z \le 1,14)$ **h)** $P(|Z-1| \le 0,83)$

a) $P(0,24 < Z < 1,73) = \Phi(1,73) - \Phi(0,24) = 0,9582 - 0,5948 = 0,3634$

b) $P(Z < -1,12) = P(Z > 1,12) = 1 - P(Z \le 1,12) = 1 - \Phi(1,12) = 1 - 0,8686 = 0,1314$

c) $P(Z > 2,12) = 1 - P(Z \le 2,12) = 1 - \Phi(2,12) = 1 - 0,9830 = 0,017$

d) $P(Z > -0,38) = P(Z \le 0,38) = \Phi(0,38) = 0,6480$

e) $P(-1,88 < Z < -0,67) = P(0,67 < Z < 1,88) = \Phi(1,88) - \Phi(0,67) = 0,9699 - 0,7486 = 0,2213$

f) $P(-1,64 < Z < 1,14) = P(Z < 1,14) - P(Z < -1,64) = P(Z < 1,14) - P(Z > 1,64) = P(Z < 1,14) - (1 - P(Z \le 1,64)) =$

$= \Phi(1,14) - (1 - \Phi(1,64)) = 0,8729 - (1 - 0,9495) = 0,8224$

g) $P(|Z| \ge 0,7) = P(Z > 0,7) + P(Z < -0,7) = 2P(Z > 0,7) = 2(1 - P(Z \le 0,7)) = 2(1 - \Phi(0,7)) = 2(1 - 0,7580) = 0,484$

h) $P(|Z-1| \le 0,83) = P(-0,83 \le Z - 1 \le 0,83) = P(0,17 \le Z \le 1,83) = \Phi(1,83) - \Phi(0,17) = 0,9664 - 0,5675 = 0,3989$

35. Considera una variable aleatoria que sigue una distribución $N(0, 1)$, halla el valor de k en cada caso.

a) $P(Z \le k) = 0,9147$ **c)** $P(Z > k) = 0,6331$ **e)** $P(-k \le Z \le k) = 0,7498$ **g)** $P(1,8 < 3Z < k) = 0,0624$

b) $P(Z \ge k) = 0,0329$ **d)** $P(Z < k) = 0,4013$ **f)** $P(k < Z < 1,3) = 0,5475$ **h)** $P(-k < Z - 1 < 0,37) = 0,189$

a) $P(Z \le k) = 0,9147 \Rightarrow k = 1,37$

b) $P(Z \ge k) = 0,0329 \Rightarrow 1 - P(Z < k) = 0,0329 \Rightarrow P(Z < k) = 0,9671 \Rightarrow k = 1,84$

c) $P(Z > k) = 0,6331 \Rightarrow P(Z < -k) = 0,6331 \Rightarrow -k = 0,34 \Rightarrow k = -0,34$

d) $P(Z < k) = 0,4013 \Rightarrow P(Z \ge k) = 1 - 0,4013 = 0,5987 \Rightarrow P(Z \le -k) = 0,5987 \Rightarrow -k = 0,25 \Rightarrow k = -0,25$

e) $P(-k \le Z \le k) = 0,7498 \Rightarrow P(Z \le k) - P(Z \le -k) = P(Z \le k) - P(Z > k) = P(Z \le k) - (1 - P(Z > k)) = 2P(Z \le k) - 1$

$0,7498 = 2P(Z \le k) - 1 \Rightarrow P(Z \le k) = 0,8749 \Rightarrow k = 1,15$

f) $P(k < Z < 1,3) = 0,5475 \Rightarrow P(Z < 1,3) - P(Z \le k) = 0,5475 \Rightarrow P(Z \le k) = 0,3557 \Rightarrow P(Z < -k) = 0,6443 \Rightarrow$

$\Rightarrow -k = 0,37 \Rightarrow k = -0,37$

g) $P(1,8 < 3Z < k) = 0,0624 \Rightarrow P\left(0,6 < Z < \frac{k}{3}\right) = P\left(Z < \frac{k}{3}\right) - P(Z < 0,6) = 0,0624 \Rightarrow P\left(Z < \frac{k}{3}\right) - 0,7257 = 0,0624 \Rightarrow$

$\Rightarrow P\left(Z < \frac{k}{3}\right) = 0,7881 \Rightarrow \frac{k}{3} = 0,80 \Rightarrow k = 2,4$

h) $P(-k < Z - 1 < 0,37) = 0,189 \Rightarrow P(1 - k < Z < 1,37) = P(Z < 1,37) - P(Z < 1 - k) = 0,189 \Rightarrow P(Z < 1 - k) = 0,7257 \Rightarrow$

$\Rightarrow 1 - k = 0,6 \Rightarrow k = 0,4$

36. Sea *X* una variable aleatoria con distribución normal de media 5 y desviación típica 2. Calcula las probabilidades siguientes.

a) $P(X<3)$ **c)** $P(-2\le X-4<2)$ **e)** $P(1\le X<9)$

b) $P(X<2)$ **d)** $P(|2X|\le 1)$ **f)** $P(X\le 2\mid X\ge 1)$

a) $P(X<3)=P\left(Z\le\frac{3-5}{2}\right)=P(Z<-1)=1-P(Z\le 1)=0{,}1587$

b) $P(X<2)=P\left(Z\le\frac{2-5}{2}\right)=P(Z\le -1{,}5)=1-P(Z<1{,}5)=0{,}0668$

c) $P(-2\le X-4<2)=P(2\le X<6)=P\left(\frac{2-5}{2}\le Z\le\frac{6-5}{2}\right)=P(-1{,}5\le Z\le 0{,}5)=P(Z\le 0{,}5)-P(Z\le -1{,}5)=$

$=P(Z\le 0{,}5)-(1-P(Z\le 1{,}5))=0{,}6915-0{,}0668=0{,}6247$

d) $P(|2X|\le 1)=P(-1\le 2x\le 1)=P(-0{,}5\le X\le 0{,}5)=P\left(\frac{-0{,}5-5}{2}\le Z\le\frac{0{,}5-5}{2}\right)=P(-2{,}75\le Z\le -2{,}25)=$

$=P(2{,}25\le Z\le 2{,}75)=0{,}9970-0{,}9878=0{,}0092$

e) $P(1\le X<9)=P\left(\frac{1-5}{2}\le Z<\frac{9-5}{2}\right)=P(-2\le Z<2)=P(Z<2)-P(Z<-2)=P(Z<2)-P(Z\ge 2)=$

$=P(Z<2)-(1-P(Z<2))=2P(Z<2)-1=2\cdot 0{,}9772-1=0{,}9544$

f) $P(X\le 2\mid X\ge 1)=\dfrac{P(1\le X\le 2)}{P(X\ge 1)}=\dfrac{P\left(\frac{1-5}{2}\le Z\le\frac{2-5}{2}\right)}{P\left(Z\ge\frac{1-5}{2}\right)}=\dfrac{P(-2\le Z\le -1{,}5)}{P(Z\ge -2)}=\dfrac{P(1{,}5\le Z\le 2)}{P(Z\le 2)}==\dfrac{0{,}9772-0{,}9332}{0{,}9772}=0{,}045$

Aproximación de una binomial por una normal.

37. Una cierta vacuna produce una reacción adversa en el 10% de los casos. Calcula la probabilidad de que, en una campaña de vacunación realizada a 1000 personas,

a) Se produzcan al menos 90 casos de reacción adversa.

b) No se produzcan más de 100 casos de reacción adversa.

Considera la variable aleatoria *X*: “nº de casos con reacción adversa, de los 1000”. La distribución de probabilidad de *X* es binomial $\text{Bin}(n=1000; p=0{,}1)$.

Como $np=100\ge 5$ y $n(1-p)=900\ge 5$, se puede aproximar la distribución de probabilidad de *X* por la de una variable aleatoria normal $Y\sim N\left(\mu=np=100;\sigma=\sqrt{np(1-p)}=\sqrt{90}\right)$.

a) Utilizando la aproximación normal a la binomial y teniendo en cuenta la corrección por continuidad:

$$P(X\ge 90)\approx P(Y\ge 89{,}5)=P\left(Z\ge\frac{89{,}5-100}{\sqrt{90}}\right)=P(Z\ge -1{,}11)=\Phi(1{,}11)=0{,}8665$$

b) En este caso

$$P(X\le 100)\sim P(Y\le 100{,}5)=P\left(Z\le\frac{100{,}5-100}{\sqrt{90}}\right)=P(Z\le 0{,}05)=\Phi(0{,}05)=0{,}5199$$

38. Se lanza 500 veces una moneda equilibrada. Halla:

a) El número esperado de caras.

b) La probabilidad de que el número de caras esté comprendido entre 100 y 200.

c) La probabilidad de obtener más de 300 caras.

Sea la variable aleatoria X: "n.º de caras obtenidas en los 500 lanzamientos". Puesto que se considera que la moneda está equilibrada, la probabilidad de obtener cara en un lanzamiento es $p = 0,5$.

La distribución de probabilidad de $X \sim \text{Bin}(n = 500; p = 0,5)$.

a) El número esperado de caras, o esperanza de la variable X, es

$E[X] = np = 500 \cdot 0,50 = 250$.

Para el cálculo de probabilidades, se puede aproximar la distribución de probabilidad de la variable X por la de una variable Y con distribución normal: $Y \sim N\left(\mu = np = 250;\ \sigma = \sqrt{npq} = \sqrt{125}\right)$.

De manera que:

$$P(100 \le X \le 200) \approx P(99,5 \le Y \le 200,5) = P\left(\frac{99,5-250}{\sqrt{125}} \le Z \le \frac{200,5-250}{\sqrt{125}}\right) = P(-13,46 \le Z \le -4,42) =$$

$$= P(4,42 \le Z \le 13,46) = 1-1 = 0$$

La probabilidad es muy próxima a cero.

b) En este caso,

$$P(300 < X) \approx P(300,5 \le Y) = P\left(\frac{300,5-250}{\sqrt{125}} \le Z\right) = P(4,52 \le Z) = 1 - P(Z < 4,52) = 1-1 = 0$$

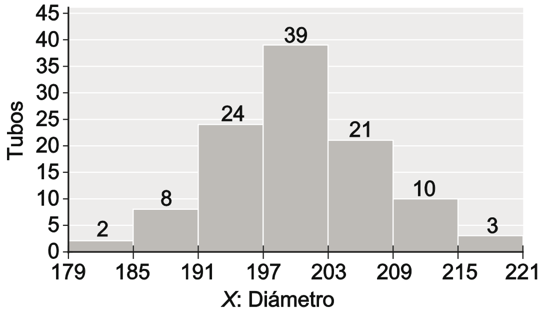
Ajuste a una distribución normal.

39. Una empresa debe producir tubos de acero de 200 mm de diámetro. El proceso de producción está sometido a control de calidad y en una inspección se midió el diámetro de 107 tubos elegidos al azar. Los datos se muestran en la tabla de la derecha.

Ajusta una distribución normal y comprueba la bondad del ajuste.

Diámetro	[179, 185)	[185, 191)	[191, 197)	[197, 203)	[203, 209)	[209, 215)	[215, 221)	
tubos	2	8	24	39	21	10	3	107

En primer lugar se dibuja el histograma de frecuencias, que es aproximadamente simétrico y que, por tanto es susceptible de aproximarse por una distribución normal.



Para obtener los parámetros de la distribución normal se calculan la media y la desviación típica (o la varianza) de la distribución de las frecuencias observadas (distribución empírica). La tabla siguiente facilita los cálculos:

Diámetro X (mm)	tubos f_j	x_j	$f_j x_j$	$f_j x_j^2$
[179, 185)	2	182	364	66248
[185, 191)	8	188	1504	282752
[191, 197)	24	194	4656	903264
[197, 203)	39	200	7800	1560000
[203, 209)	21	206	4326	891156
[209, 215)	10	212	2120	449440
[215, 221)	3	218	654	142572
	107		21424	4295432

De donde,

$\bar{X} = \dfrac{21424}{107} = 200{,}224 \qquad s = \sqrt{\dfrac{4295432}{107} - 200{,}224^2} = 7{,}3793$

De modo que la variable X: "diámetro, en mm, de los tubos de acero", tiene una distribución que puede aproximarse por una normal: $X \sim N(\mu = 200{,}224;\ \sigma = 7{,}3793)$

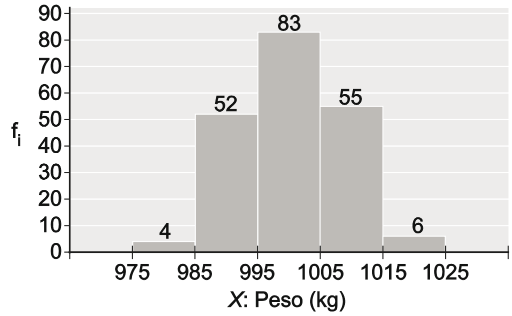
Para comprobar la bondad del ajuste, se calculan las frecuencias teóricas y se comparan con las frecuencias observadas. En la tabla siguiente se incluyen las probabilidades de los intervalos, calculadas con la distribución normal ajustada, y las frecuencias teóricas correspondientes a cada intervalo:

Probabilidades teóricas de cada intervalo	107 ·probabilidad	frecuencias teóricas
$P(179 \le X < 185) = P(-2{,}88 \le Z < -2{,}06) = 0{,}0177$	1,8939	2
$P(185 \le X < 191) = P(-2{,}06 \le Z < -1{,}25) = 0{,}0859$	9,1913	9
$P(191 \le X < 197) = P(-1{,}25 \le Z < -0{,}44) = 0{,}2244$	24,0108	24
$P(197 \le X < 203) = P(-0{,}44 \le Z < 0{,}38) = 0{,}3180$	34,026	34
$P(203 \le X < 209) = P(0{,}38 \le Z < 1{,}19) = 0{,}2350$	25,145	25
$P(209 \le X < 215) = P(1{,}19 \le Z < 2{,}00) = 0{,}0942$	10,0794	10
$P(215 \le X < 221) = P(2{,}00 \le Z < 2{,}82) = 0{,}0204$	2,1828	2
		106

Comparando las frecuencias teóricas con las observadas, puede verificarse que la aproximación realizada es buena.

Síntesis

40. Se han pesado en una balanza 200 paquetes de azúcar de una máquina envasadora, obteniéndose el siguiente histograma de frecuencias:



a) Ajusta una distribución normal a la distribución de frecuencias y comprueba la bondad del ajuste.

b) Con la distribución ajustada, calcula la probabilidad de que un paquete elegido al azar pese menos de 990 g.

c) Si se seleccionan 50 paquetes, ¿cuántos pesarán entre 1000 g y 1010 g?

a) Como el histograma de frecuencias absolutas es aproximadamente simétrico, puede aproximarse razonablemente por un distribución normal.Para obtener los parámetros de la distribución normal, se calculan la media y la desviación típica de la distribución de frecuencias, utilizando la tabla.

X (g)	f_j	x_j	f_jx_j	$f_jx_j^2$
[975, 985)	4	980	3920	3 841 600
[985, 995)	52	990	51 480	50 965 200
[995, 1005)	83	1000	83 000	83 000 000
[1005, 1015)	55	1010	55 550	56 105 500
[1010, 1025)	6	1020	6120	6 242 400
	200		200 070	200 154 700

$$\bar{X} = \frac{200070}{200} = 1000{,}35 \qquad s = \sqrt{\frac{200154700}{200} - 1000{,}35^2} = 8{,}5661$$

De esta manera, el ajuste de la distribución empírica (de frecuencias) se realiza mediante una distribución normal cuyos parámetros (media y varianza) son los de la distribución de frecuencias. Es decir, que a la variable *X*: "peso (en gramos) de los paquetes de azúcar", se le ajusta una distribución:

$$X \sim N(\mu = 1000{,}35; \sigma = 8{,}5661)$$

Para comprobar la "bondad" del ajuste, se comparan las frecuencias observadas con las frecuencias teóricas que corresponderían a cada intervalo si la distribución de frecuencias fuera la ajustada, es decir, la normal. Con la distribución normal ajustada, se calculan las probabilidades de cada uno de los intervalos de la tabla y, multiplicando estas por 200, se obtienen las frecuencias teóricas correspondientes a cada intervalo:

Probabilidades teóricas de cada intervalo	200 ·probabilidad	frecuencias teóricas
$P(975 \le X < 985) = P(-2{,}96 \le Z < -1{,}79) = 0{,}0352$	7,04	7
$P(985 \le X < 995) = P(-1{,}79 \le Z < -0{,}62) = 0{,}2309$	46,18	46
$P(995 \le X < 1005) = P(-0{,}62 \le Z < 0{,}54) = 0{,}4378$	87,56	88
$P(1005 \le X < 1015) = P(0{,}54 \le Z < 1{,}71) = 0{,}2510$	50,2	50
$P(1015 \le X < 1025) = P(1{,}71 \le Z < 2{,}88) = 0{,}0416$	8,32	8
		199

Comparando las frecuencias teóricas con las observadas, puede afirmarse que la aproximación es razonablemente buena.

b) En este caso, se trata de calcular la probabilidad siguiente:

$$P(X < 990) = P\left(Z < \frac{990 - 1000{,}35}{8{,}5661}\right) = P(Z < -1{,}21) = 1 - \Phi(1{,}21) = 1 - 0{,}8869 = 0{,}1131$$

c) En este otro se pide:

$$P(1000 < X < 1010) = P\left(\frac{1000 - 1000{,}35}{8{,}5661} < Z < \frac{1010 - 1000{,}35}{8{,}5661}\right) = P(-0{,}04 < Z < 1{,}13) =$$

$$= \Phi(1{,}13) - (1 - \Phi(0{,}04)) = 0{,}3868$$

41. Una variable aleatoria *X* sigue una distribución normal de media μ y desviación típica σ. Si $P(|X-\mu| \le 1) = 0,383$, halla la desviación típica de *X*.

Como $0,383 = P(|X-\mu| \le 1) = P(-1 \le X-\mu \le 1) = P\left(\frac{-1}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{1}{\sigma}\right) = P\left(\frac{-1}{\sigma} \le Z \le \frac{1}{\sigma}\right) = 2\Phi\left(\frac{1}{\sigma}\right) - 1$,

Se tiene que:

$\Phi\left(\frac{1}{\sigma}\right) = 0,6915 \Rightarrow \frac{1}{\sigma} = 0,5 \Rightarrow \sigma = \frac{1}{0,5} = 2$

42. El retraso con el que llegan los autobuses de una línea que cubre el trayecto entre dos ciudades sigue una distribución normal de media μ = 5 minutos.Si se sabe que el 30% llega con un retraso superior a 7 minutos.

a) Calcula la desviación típica de la distribución.
b) Determina el porcentaje de autobuses que llega antes de la hora fijada.
c) Halla la probabilidad de que el retraso supere los 8 minutos.
d) ¿Cuál es retraso de un autobús que llega antes que el 80% de los autobuses?

a) Sea la variable aleatoria *X*: "retraso, en minutos, con el llegan los autobuses". *X* tiene una distribución $N(\mu = 5;\, \sigma)$, con σ desconocida.

Si el 30% llega con un retraso superior a 7 minutos, ello significa que $P(X > 7) = 0,3$.

Utilizando la información anterior, se tiene que

$$P(X > 7) = 0,3 \Rightarrow P\left(Z > \frac{7-5}{\sigma}\right) = 0,3 \Rightarrow \frac{2}{\sigma} = 0,525 \Rightarrow \sigma = 3,81$$

b) En este caso se trata de calcular $P(X < 0)$.

$P(X < 0) = P\left(Z < \frac{0-5}{3,81}\right) = P(Z < -1,31) = 1 - \Phi(1,31) = 1 - 0,9049 = 0,0951$, es decir, alrededor del 9%.

c) La probabilidad de que el retraso supere los 8 minutos viene dada por

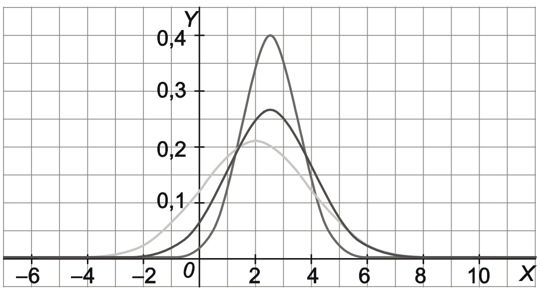
$$P(X > 8) = P\left(Z > \frac{8-5}{3,81}\right) = P(Z > 0,79) = 1 - P(Z \le 0,79) = 1 - 0,7852 = 0,2148$$

d) Ahora se trata de calcular el valor de *c* de tal manera que $P(X < c) = 0,2$, que equivale a:

$$P\left(Z < \frac{c-5}{3,81}\right) = 0,2 \Rightarrow \Phi\left(\frac{c-5}{3,81}\right) = 0,2 \Rightarrow \frac{c-5}{3,81} = -0,842 \Rightarrow c = 1,792 \text{minutos}$$

CUESTIONES

43. Observa las gráficas de las siguientes funciones de densidad



Asigna a cada gráfica la media y desviación típica que les corresponde.

a) $\mu = 2,5$ y $\sigma = 1,5$

b) $\mu = 2,5$ y $\sigma = 1$

c) $\mu = 2$ y $\sigma = 0,5$

a) Curva de color rojo (altura intermedia).

b) Curva de color azul (la que alcanza mayor altura).

c) Curva de color rosa (la que tiene el máximo sobre *x*= 2).

44. Sea una variable aleatoria *X* con distribución normal $N(\mu, \sigma)$, calcula el porcentaje de valores de la variable que distan dela media:

a) Menos de 1,5 desviaciones típicas

b) Más de 0,5 y menos de 2 desviaciones típicas.

a) Los valores de *X* que distan de la media μ menos de 1,5 desviaciones típicas son los incluidos en el intervalo $(\mu - 1,5\cdot\sigma; \mu + 1,5\cdot\sigma)$.
Entonces,

$$P(\mu - 1,5\sigma < X < \mu + 1,5\sigma) = P\left(\frac{\mu - 1,5\sigma - \mu}{\sigma} < Z < \frac{\mu - 1,5\sigma - \mu}{\sigma}\right) = P(-1,5 < Z < 1,5) = P(Z < 1,5) - P(Z < -1,5) =$$

$$= \Phi(1,5) - (1 - \Phi(1,5)) = 2\cdot\Phi(1,5) - 1 = 2\cdot 0,9332 - 1 = 0,8664$$

b) En este caso, por la simetría de la densidad de la distribución normal, se tiene que

$$P(0,5\cdot\sigma < |X - \mu| < 2\cdot\sigma) = P\left(0,5 < \left|\frac{X-\mu}{\sigma}\right| < 2\right) = 2\cdot(\Phi(2) - \Phi(0,5)) = 2\cdot(0,9772 - 0,6915) = 0,5714.$$

45. La calificación *X* de un examen es una variable aleatoria que toma cualquier valor del intervalo [0, 10]. Se sabe que *X* tiene una distribución continua con función de densidad:

$$f(x) = \begin{cases} \dfrac{x+1}{60} & \text{si } 0 \le x \le 10 \\ 0 & \text{en el resto} \end{cases}$$

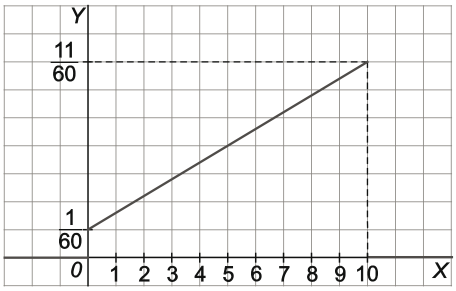
a) Comprueba que es una función de densidad.

b) ¿Cuál es la probabilidad de aprobar este examen?

c) Calcula la probabilidad de obtener nota superior a 7.

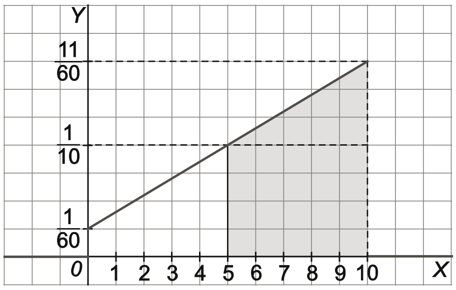
a) La gráfica de esta función de densidad es la mostrada.
En efecto, se trata de una función de densidad, puesto que $f(x) \ge 0$, para cualquier valor de *x*, y el área de la región limitada por la gráfica de la función, el eje *X* y las rectas verticales $x = 0$ y $x = 10$, es la unidad, ya que se puede poner como la suma de las áreas de un rectángulo y un triángulo (o un trapecio):

$$10\cdot\frac{1}{60} + \frac{10\cdot\left(\frac{11}{60} - \frac{1}{60}\right)}{2} = \frac{1}{6} + \frac{5}{6} = 1$$



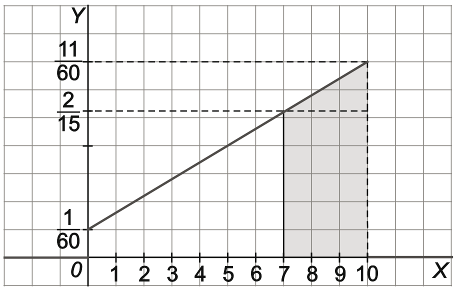
b) La probabilidad de aprobar el examen (obtener nota igual o superior a 5) es el área de la región coloreada
Que, como se ve en la figura, se puede obtener por la suma de las áreas de un rectángulo y un triángulo:

$$P(X \ge 5) = (10-5)\cdot\frac{1}{10} + \frac{(10-5)\left(\frac{11}{60} - \frac{1}{10}\right)}{2} = \frac{17}{24} = 0,7083$$



c) La probabilidad de obtener nota superior a 7 es el área de la región sombreada

$$P(X > 7) = (10-7)\frac{2}{15} + \frac{(10-7)\left(\frac{11}{60} - \frac{2}{15}\right)}{2} = \frac{2}{5} + \frac{3}{40} = \frac{19}{40} = 0,475$$



46. El tiempo empleado por Juan para ir todos los días al instituto sigue una distribución normal de media 20 minutos y desviación típica 4 minutos.

a) En un día elegido al azar ¿cuál es la probabilidad de que tarde menos de 25 minutos?

b) ¿Cuál es porcentaje de días que tarda entre 15 y 25 minutos?

Sea la variable aleatoria *X*: “tiempo, en minutos, empleado por Juan para ir al instituto”, cuya distribución es $X \sim N(\mu = 20; \sigma = 4)$

a) Si se elige un día la azar, la probabilidad de que tarde menos de 25 minutos es:

$$P(X < 25) = P\left(Z > \frac{25-20}{4}\right) = \Phi(1{,}25) = 0{,}8944$$

b) La probabilidad de que un día elegido al azar tarde entre 15 y 25 minutos es:

$$P(15 < X < 25) = P\left(\frac{15-20}{4} < Z < \frac{25-20}{4}\right) = \Phi(1{,}25) - 1 + \Phi(1{,}25) = 0{,}8944 - 1 + 0{,}8944 = 0{,}7888$$

De manera que en torno al 79% de los días Juan tarda entre 15 y 25 minutos en llegar al instituto.

47. La venta diaria, sin contar fines de semana, de teléfonos móviles en un centro comercial sigue una distribución normal de media 15 y desviación típica 4. Calcula, en un día elegido al azar, la probabilidad de que se vendan:

a) Más de 21 teléfonos

b) Menos de 7

c) Al menos 17

d) Entre 11 y 19 teléfonos

La variable *X*: “número de teléfonos móviles vendidos diariamente en el centro comercial” tiene una distribución $N(\mu = 15;\ \sigma = 4)$, luego, tipificando en cada uno de los casos:

a) $P(X > 21) = P\left(Z > \frac{21-16}{4}\right) = P(Z > 1{,}25) = 1 - \Phi(1{,}25) = 1 - 0{,}8944 = 0{,}1056$

b) $P(X < 7) = P\left(Z < \frac{7-15}{4}\right) = P(Z < -2) = 1 - \Phi(2) = 1 - 0{,}9772 = 0{,}0228$

c) $P(X \geq 17) = P\left(Z \geq \frac{17-16}{4}\right) = P(Z \geq 0{,}25) = 1 - \Phi(0{,}25) = 1 - 0{,}5987 = 0{,}4013$

d) $P(11 < X < 19) = P\left(\frac{11-15}{4} < Z < \frac{19-15}{4}\right) = P(-1 < Z < 1) = 2 \cdot \Phi(1) - 1 = 2 \cdot 0{,}8413 - 1 = 0{,}6826$

48. Un test específico para determinar el estado de salud de los trabajadores de una empresa tiene una distribución normal de media 100 y desviación típica 8. Si un trabajador supera los 125 puntos debe ser objeto de una segunda revisión en profundidad.

¿Cuál es el porcentaje de trabajadores que necesitará una segunda revisión?

Sea *X*: “estado de salud de los trabajadores”, cuya distribución es $N(\mu = 100;\ \sigma = 8)$.

Para estimar el porcentaje de trabajadores que necesitará una segunda revisión, debe calcularse:

$$P(X > 125) = P\left(Z > \frac{125-100}{8}\right) = P(Z > 3{,}125) = 1 - \Phi(3{,}125) = 1 - 0{,}9991 = 0{,}0009$$

Luego, se estima que la cantidad de trabajadores que necesitará una segunda revisión es menor que el $1‰$.

49. En una fábrica de botellas se sabe que la capacidad de las mismas sigue una distribución normal de media 2 L y desviación típica 0,015 L.

Si se desechan todas las botellas que tengan una capacidad inferior a 1,98 L o superior a 2,02 L. Calcula el porcentaje de botellas que será rechazado.

La variable aleatoria *X*: "capacidad (litros) de las botellas" tiene una distribución $N(\mu = 2;\ \sigma = 0{,}015)$.
Para estimar las botellas que serán rechazadas, se calcula la probabilidad de que una botella elegida al azar sea aceptada:

$$P(1{,}98 < X < 2{,}02) = P\left(\frac{1{,}98-2}{0{,}015} < Z < \frac{2{,}02-2}{0{,}015}\right) = P(-1{,}33 < Z < 1{,}33) = 2\cdot\Phi(1{,}33) - 1 = 2\cdot 0{,}9082 - 1 = 0{,}8164$$

Por tanto una botella elegida al azar será rechazada con probabilidad $P(\text{rechazar la botella}) = 1 - 0{,}8164 = 0{,}1836$.
Es decir, el 18,36% de las botellas serán rechazadas.

50. El consumo medio de gasolina de un determinado modelo de vehículo es de 6,5 L por cada 100 km. Se sabe que el consumo de gasolina de este modelo tiene una distribución normal con desviación típica 2 L.

a) ¿Cuál es la probabilidad de que un vehículo de este modelo, elegido al azar, consuma más de 7 L/100 km?

b) ¿Cuál es porcentaje de vehículos de este modelo que consume entre 5 y 6 L/100 km?

c) Si se seleccionan 200 vehículos ¿Cuántos consumirán menos de 6 L/100 km?

Considera la variable aleatoria *X*: "consumo de gasolina del modelo en cuestión". Se estima que $X \sim N(\mu = 6{,}5;\ \sigma = 2)$.

a) La probabilidad que se pide es $P(X > 7) = P\left(Z > \frac{7-6{,}5}{2}\right) = P(Z > 0{,}25) = 1 - \Phi(0{,}25) = 1 - 0{,}5987 = 0{,}4013$.

b) La probabilidad de que un vehículo de este modelo elegido al azar consuma entre 5 y 6 L/100 km.

$$P(5 < X < 6) = P\left(\frac{5-6{,}5}{2} < Z < \frac{6-6{,}5}{2}\right) = P(-0{,}75 < Z < -0{,}25) = P(0{,}25 < Z < 0{,}75) =$$
$$= \Phi(0{,}75) - \Phi(0{,}25) = 0{,}7734 - 0{,}5987 = 0{,}1747$$

Es decir, el 17,47% de los vehículos de este modelo consume entre 5 y 6 L/100 km.

c) La probabilidad de que un vehículo de este modelo elegido al azar consuma menos de 6 L/100 km.

$$P(X < 6) = P\left(Z < \frac{6-6{,}5}{2}\right) = P(Z < -0{,}25) = 1 - \Phi(0{,}25) = 1 - 0{,}5987 = 0{,}4013$$

El 40,13 % de los vehículos consumirán menos de 6 L/100 km. Como el 40,13 % de 200 es 80,26, aproximadamente 80 vehículos de los 200 de la muestra consumirán menos de 6 L/100 km.

51. La medida del diámetro interior de las arandelas que produce una máquina sigue una distribución normal de media10 mm y desviación típica 0,1 mm. Las referencias especifican que solo son admisibles las arandelas con un diámetro interior entre 9,85 y 10,2.

Calcula el porcentaje de arandelas producidas que cumplen las especificaciones.

Sea la variable aleatoria *X*: "medida en mm. del diámetro interior de las arandelas". Su distribución de probabilidad es $N(\mu = 10;\ \sigma = 0{,}1)$.
Para estimar el porcentaje de arandelas que cumplen las especificaciones, se calcula la probabilidad de que el diámetro interior de una arandela elegida al azar esté entre 9,85 y 10,2 mm.

$$P(9{,}85 < X < 10{,}2) = P\left(\frac{9{,}85-10}{0{,}1} < Z < \frac{10{,}2-10}{0{,}1}\right) = P(-1{,}5 < Z < 2) = \Phi(2) - \Phi(-1{,}5) = \Phi(2) - (1 - \Phi(1{,}5)) =$$
$$= 0{,}9772 - (1 - 0{,}9332) = 0{,}9104$$

De manera que el 91,04% de las arandelas cumple las especificaciones.

52. El gasto trimestral, en euros, en material de oficina de los empleados de una empresa sigue una distribución normal $N(\mu = 550,\ \sigma = 60)$.

a) Determina el porcentaje de empleados que gasta entre 500 y 625 €.

b) Si se eligen al azar 60 empleados ¿Cuántos de ellos se espera que superen los 650 € de gasto?

Se considera la variable aleatoria X: "gasto corriente en material, en euros". Se sabe que $X \sim N(\mu = 550; \sigma = 60)$.

a) Si se elige un empleado al azar, la probabilidad de que su gasto esté entre 500 € y 625 € es:

$$P(500 < X < 625) = P\left(\frac{500-550}{60} < Z < \frac{625-550}{60}\right) = P(-0,8333 < Z < 1,25) =$$

$$= \Phi(1,25) - 1 + \Phi(0,83) = 0,8944 - 1 + 0,7967 = 0,6911$$

De modo que el 69,11% de los empleados presenta un gasto trimestral en material entre 500 € y 625 €.

b) La probabilidad de que un empleado elegido al azar supere los 650 € en gasto de material al trimestre es:

$$P(X > 650) = P\left(Z > \frac{650-550}{60}\right) = P(Z > 1,67) = 1 - P(Z < 1,67) = 1 - 0,9525 = 0,0475$$

Esto es, el 4,75% de los empleados supera los 650 € en gasto.El 4,75% de 60 es 2,85. Luego, de los 60 empleados elegidos al azar se espera que 3 supere los 650 € en gasto trimestral.

53. El voltaje de la corriente eléctrica que llega a un domicilio tiene una distribución normal de media 220 y desviación típica 1,5 voltios. Calcula la probabilidad de que una medida del voltaje tomada al azar:

a) Sea menor de 221 voltios.

b) Esté comprendida entre 215 y 222 voltios.

c) Supere el límite establecido en 230 voltios para que salte el diferencial.

La variable X: "voltaje de corriente eléctrica" tiene una distribución $N(\mu = 220; \sigma = 1,5)$.

a) La probabilidad pedida es: $P(X < 221) = P\left(Z < \frac{221-220}{1,5}\right) = P(Z < 0,67) = 0,7454$.

b) En este caso:

$$P(215 < X < 222) = P\left(\frac{215-220}{1,5} < Z < \frac{222-220}{1,5}\right) = P(-3,33 < Z < 1,33) = \Phi(1,33) - \Phi(-3,33) =$$

$$= \Phi(1,33) - (1 - \Phi(3,33)) = 0,9082 - (1 - 0,9996) = 0,9078$$

c) Por último, la probabilidad de que salte el diferencial es: $P(X > 230) = P\left(Z > \frac{230-220}{1,5}\right) = P(Z > 6,67) = 0$.

54. La duración de las llamadas de teléfono en una oficina comercial sigue una distribución normal de media40 segundos y varianza 100 s^2. Calcula el porcentaje de llamadas:

a) Que supera el minuto.

b) Cuya duración está comprendida entre 30 y 45 s.

La variable aleatoria X: "duración (en segundos) de las llamadas de teléfono", tiene una distribución $N(\mu = 40; \sigma = 10)$.

a) La probabilidad de que una llamada elegida al azar haya durado más de 1 minuto (60 s) es:

$$P(X > 60) = P\left(Z > \frac{60-40}{10}\right) = P(Z > 2) = 1 - \Phi(2) = 1 - 0,9772 = 0,0228$$

Es decir, el 2,28% de las llamadas en esa oficina supera el minuto.

b) La probabilidad de que una llamada elegida al azar esté comprendida entre 30 s y 45 s es:

$$P(30 \le X \le 45) = P\left(\frac{30-40}{10} \le Z \le \frac{45-40}{10}\right) = P(-1 \le Z \le 0,5) = \Phi(0,5) - \Phi(-1) =$$

$$= \Phi(0,5) - (1 - \Phi(1)) = 0,6915 - (1 - 0,8413) = 0,5328$$

De manera que el 53,28% de las llamadas telefónicas en esa oficina duran entre 30 s y 45 s.

55. En el balance final del último año en una entidad financiera, la morosidad (créditos concedidos y no pagados) ha ascendido al 25%. Si de dicha entidad se eligen al azar 50 clientes a los que se les ha concedido un crédito, calcula probabilidad de que entre los 50:

a) Haya al menos 20 morosos.

b) Los morosos sean exactamente 12

Sea la variable aleatoria *X*: "número de morosos entre los 50 clientes seleccionados". La distribución de probabilidad de esta variable es $X \sim \text{Bin}(n = 50; p = 0,25)$, cuya esperanza y varianza son:

$E[X] = 50 \cdot 0,25 = 12,5$ $\text{Var}(X) = 50 \cdot 0,25 \cdot 0,75 = 9,375 \Rightarrow \sigma = \sqrt{9,375} = 3,062$

Para el cálculo de probabilidades, se puede aproximar la distribución binomial de *X* por la distribución de una variable aleatoria normal $Y \sim N(\mu = 12,5; \sigma = 3,062)$. Entonces:

a) La probabilidad de que haya al menos 20 morosos es:

$$P(X \geq 20) \approx P(Y \geq 19,5) = P\left(Z \geq \frac{19,5 - 12,5}{3,062}\right) = P(Z \geq 2,29) = 1 - \Phi(2,29) = 1 - 0,9890 = 0,0110$$

b) La probabilidad de que haya exactamente 12 morosos se puede calcular también con la distribución normal.

$$P(X = 12) \approx P(11,5 < Y < 12,5) = P\left(\frac{11,5 - 12,5}{3,062} < Z < \frac{12,5 - 12,5}{3,062}\right) = P(-0,33 < Z < 0) = \Phi(0) - \Phi(-0,33) =$$

$$= \Phi(0) - 1 + \Phi(0,33) = 0,5 - 1 + 0,6293 = 0,1293$$

Obsérvese que si se hace directamente por la binomial el resultado obtenido es prácticamente el mismo.

$$P(X = 12) = \binom{50}{20} \cdot 0,25^{12} \cdot 0,75^{38} = 0,1293676...$$

56. Un tribunal debe calificar a los 700 aspirantes a las 25 plazas convocadas para cubrir vacantes en un organismo oficial. Si las calificaciones son de 0 a 10 y su distribución es normal de media $\mu = 5,7$ puntos y desviación típica $\sigma = 1,5$ puntos, contesta:

a) ¿Cuántos opositores han obtenido puntuación superior o igual 5 puntos?

b) ¿Cuál es la nota de corte para ser seleccionado?

Sea la variable aleatoria *X*: "calificación de las pruebas". Se tiene que $X \sim N(\mu = 5,7; \sigma = 1,5)$. Entonces:

a) La probabilidad de que un opositor elegido al azar obtenga calificación igual o superior a 5:

$$P(X \geq 5) = P\left(Z \geq \frac{5 - 5,7}{1,5}\right) = P(Z \geq -0,47) = P(Z \leq 0,47) = 0,6808$$

De manera que el 68,08% de los opositores ha obtenido puntuación superior o igual 5 y, por tanto, $700 \cdot 0,6808 = 476,56$, es decir unos 477 opositores han obtenido nota superior o igual a 5.

b) Para obtener la nota de corte, *c*, se debe tenerse en cuenta que 675 opositores no obtendrán plaza, lo que supone el 96,43% de los 700 aspirantes. Luego planteando el problema, tipificando y utilizando las tablas de la normal estándar:

$$P(X < c) = 0,9643 \Rightarrow P\left(Z < \frac{c - 5,7}{1,5}\right) = 0,9643 \Rightarrow \frac{c - 5,7}{1,5} = 1,80 \Rightarrow c = 8,4$$

La nota de corte es, por tanto, 8,4.

57. La estatura media de los bebés nacidos en una maternidad en el último año ha sido de 45 cm. Además, el 20 % ha medido más de 50 cm. Si se supone que la estatura de los bebés sigue una distribución normal,

a) Calcula la desviación típica.

b) Calcula la probabilidad de que un bebé elegido al azar mida menos de 40 cm.

c) Si en la maternidad nacieron 100 bebés en el último mes ¿Cuántos se espera que hayan sobrepasado los 55 cm?

Sea la variable aleatoria X: "estatura en cm de los bebés nacidos el último año en la maternidad".

La distribución de probabilidad de X es $N(\mu = 45;\ \sigma)$, con σ desconocida.

a) Si el 20% de los bebés ha medido más de 50 cm, entonces $P(X > 50) = 0,2$.

De la información anterior se tiene que:

$$P(X > 50) = P\left(Z > \frac{50-45}{\sigma}\right) = 0,2 \Rightarrow \frac{5}{\sigma} = 0,843 \Rightarrow \sigma = 5,931$$

b) Con la desviación típica obtenida en el apartado anterior:

$$P(X < 40) = P\left(Z < \frac{40-45}{5,931}\right) = P(Z < -0,84) = 1-\Phi(0,84) = 1-0,7995 = 0,2005$$

c) La probabilidad de que un bebé elegido al azar supere los 55 cm es:

$$P(X > 55) = P\left(Z > \frac{55-45}{5,931}\right) = P(Z > 1,69) = 1-\Phi(1,69) = 1-0,9545 = 0,0455$$

Por tanto, de 100 bebés se espera que hayan sobrepasado los 55 cm, el 4,55%, es decir, entre 4 y 5 bebés.

58. El 70 % de los habitantes de una localidad se oponen a que en su municipio se construya un cementerio nuclear. Se elige una muestra de los habitantes.

a) Si la muestra es de 50 personas, calcula la esperanza y la desviación típica de la variable X: "número de personas de las 50, que se oponen a la construcción del cementerio nuclear".

b) Si la muestra es de 8 personas, ¿cuál es la probabilidad de que al menos 5 se opongan al proyecto?

c) Si la muestra es de 80 personas, calcula la probabilidad de que más de 60 se opongan al proyecto.

a) La variable aleatoria X sigue una distribución $\text{Bin}(n = 50;\ p = 0,7)$, por tanto su esperanza y su varianza son:

$E[X] = np = 50 \cdot 0,7 = 35$, $\quad \text{Var}(X) = np(1-p) = 50 \cdot 0,7 \cdot 0,3 = 10,5$

Y la desviación típica: $\sigma = \sqrt{\text{Var}(X)} = \sqrt{10,5} = 3,2404$

b) En este caso, la variable Y: "número de personas, de las 8, que se oponen a la construcción del cementerio nuclear", sigue una distribución binomial $\text{Bin}(n = 8,\ p = 0,7)$, por lo que:

$$P(Y \geq 5) = P(Y = 5) + P(Y = 6) + P(Y = 7) + P(Y = 8) = \binom{8}{6} \cdot 0,7^6 \cdot 0,3^2 + \binom{8}{7} \cdot 0,7^7 \cdot 0,3^1 + \binom{8}{8} \cdot 0,7^8 =$$

$$= 0,2541 + 0,2965 + 0,1977 + 0,0576 = 0,8059$$

c) De forma similar a lo hecho en el apartado a), la distribución de $X \sim \text{Bin}(n = 80,\ p = 0,7)$ se puede aproximar por la de la normal $T \sim N(\mu = 56;\ \sigma = 4,1)$, ya que $\text{Var}(X) = np(1-p) = 80 \cdot 0,7 \cdot 0,3 = 16,8 \Rightarrow \sigma = \sqrt{16,8} = 4,1)$ de manera que la probabilidad de que más de 60 se opongan a la construcción es:

$$P(X > 60) \approx P(T > 60,5) = P\left(Z > \frac{60,5-56}{4,1}\right) = P(Z > 1,10) = 1-\Phi(1,10) = 1-0,8643 = 0,1357.$$

59. Las puntuaciones de un test, que mide la capacidad de cálculo de las personas adultas, siguen una distribución normal de media 6 y varianza 4. El test se pasa a una muestra aleatoria de 1000 personas,

a) ¿Cuántas personas obtendrán una puntuación superior a 8,5?

b) Si para superar el test es preciso obtener más de 5 puntos ¿Cuál será el porcentaje de personas que no lo superará?

a) Considera la variable aleatoria X: "Puntuaciones del test", cuya distribución de probabilidad es: $X \sim N(\mu = 6; \sigma = 2)$. Entonces, la probabilidad de que una persona elegida al azar obtenga una puntuación superior a 8,5 puntos es: $P(X > 8{,}5) = P\left(Z > \dfrac{8{,}5-6}{2}\right) = P(Z > 1{,}25) = 1 - \Phi(1{,}25) = 1 - 0{,}8944 = 0{,}1056$.

Es decir, el 10,56% de las personas obtendrá una puntuación superior a 8,5 puntos, luego si el test se ha pasado a 1000 personas, alrededor de 106 conseguirán una puntuación superior a 8,5 puntos.

b) Se calcula la probabilidad de que una persona elegida al azar obtenga más de 5 puntos. $P(X > 5) = P\left(Z > \dfrac{5-6}{2}\right) = P(Z > -0{,}5) = \Phi(0{,}5) = 0{,}6915$

De modo que el 69,15% obtendrá puntuación superior a 5 puntos.

60. El peso , en gramos, de los recién nacidos en la maternidad de un hospital sigue una distribución normal de media 3000 g. Si el 95 % de los bebés tiene un peso comprendido entre 2600 y 3400 g ¿Cuál es el valor de σ?

Considera la variable aleatoria X: "el peso, en gramos, de los recién nacidos". X tiene una distribución $N(\mu = 3000; \sigma)$. Si el peso del 95% de los bebés está entre 2600 y 3400 g, ello significa que:

$$P(2600 < X < 3400) = 0{,}95 \Rightarrow P\left(\frac{2400-3000}{\sigma} < Z < \frac{3400-3000}{\sigma}\right) = 0{,}95 \Rightarrow P\left(-\frac{600}{\sigma} < Z < \frac{600}{\sigma}\right) = 0{,}95$$

Ahora bien, como $P(-a < Z < a) = 2 \cdot \Phi(a) - 1$ cualquiera que sea $a > 0$, se tiene que

$$2\Phi\left(\frac{600}{\sigma}\right) - 1 = 0{,}95 \Rightarrow \Phi\left(\frac{600}{\sigma}\right) = 0{,}975 \Rightarrow \frac{600}{\sigma} = 1{,}96 \Rightarrow \sigma = 306{,}12$$

61. La producción de trigo por hectárea de terreno en una comarca sigue una distribución $N(\mu, \sigma)$ Los datos históricos indican que solo en el 10 % de los años la producción supera los 4000 kg/ha, mientras que en el 60 % de los años queda por debajo de los 3200 kg/ha.

a) Calcula la media y la desviación típica de la distribución.

b) Calcula la probabilidad de que la producción alcance los 3500 kilos por hectárea en una un año elegido al azar.

a) Considerando la variable aleatoria X: "producción de trigo por hectárea en la comarca", cuya distribución es $N(\mu, \sigma)$ .Que el 10% de los años la producción supere los 4000 kg/ha significa que: $P(X > 4000) = 0{,}1$, Mientras que el 60% de los años la producción queda por debajo de los 3200 kg/ha significa que: $P(X < 3200) = 0{,}6$.

Con la información disponible se tiene que: $P(X > 4000) = 0{,}1 \Rightarrow P\left(Z > \dfrac{4000-\mu}{\sigma}\right) = 0{,}1 \Rightarrow \dfrac{4000-\mu}{\sigma} = 1{,}281$

$$P(X < 3200) = 0{,}6 \Rightarrow P\left(Z < \frac{3200-\mu}{\sigma}\right) = 0{,}6 \Rightarrow \frac{3200-\mu}{\sigma} = 0{,}253$$

El sistema de dos ecuaciones con dos incógnitas, queda: $\begin{cases} \mu + 1{,}281\sigma = 4000 \\ \mu + 0{,}253\sigma = 3200 \end{cases}$

cuya solución es $\mu = 3003{,}11$ y $\sigma = 778{,}21$ kg/ha.

b) Con los parámetros calculados en el apartado a), se tiene que:

$$P(X \geq 3500) = P\left(Z > \frac{3500-3003{,}11}{778{,}21}\right) = P(Z > 0{,}64) = 1 - \Phi(0{,}64) = 1 - 0{,}7389 = 0{,}2611$$

62. Según los datos del organismo correspondiente el 80 % de los incendios que se producen en la época de calor son provocados. Si este verano se han producido 150 incendios en una determinada región, calcula la probabilidad de que:

a) Más de 100 hayan sido provocados.

b) Como mucho 30 hayan sido accidentales.

c) El número de incendios provocados supere el 80% del total de incendios.

a) Sea la variable aleatoria *X*: “número de incendios provocados, de los 150 producidos”. La variable *X* tiene una distribución binomial $\text{Bin}(n = 150;\ p = 0,8)$, cuya esperanza y varianza son respectivamente:

$E[X] = 150 \cdot 0,8 = 120$ $\qquad$ $\text{Var}(X) = 150 \cdot 0,8 \cdot 0,2 = 24 \Rightarrow \sigma = \sqrt{24}$

Y se puede aproximar por la distribución de una variable $Y \sim N(\mu = 120;\ \sigma = \sqrt{24})$. Entonces:
Utilizando la aproximación normal, la probabilidad de que más de 100 incendios hayan sido provocados es:

$$P(X > 100) \approx P(Y \geq 100,5) = P\left(Z \geq \frac{100,5 - 120}{\sqrt{24}}\right) = P(Z \geq -3,98) = \Phi(3,98) = 1$$

b) Que como mucho 30 sean accidentales equivale a que másde 120 sean intencionados, entonces:

$$P(X > 120) \approx P(Y \geq 120,5) = P\left(Z \geq \frac{120,5 - 120}{\sqrt{24}}\right) = P(Z \geq 0,10) = 1 - \Phi(0,10) = 1 - 0,5398 = 0,4602$$

c) Si el número de incendios provocados supera el 80% del total, dado que el 80% de 150 es 120, tenemos que la probabilidad pedida es justo la del apartado anterior.

63. La estatura, en cm, de las alumnas de un centro escolar sigue una distribución $N(\mu = 165,\ \sigma = 12)$ y la estatura de los chicos es $N(\mu = 175,\ \sigma = 12)$. Se eligen al azar un chico y una chica, calcula la probabilidad de que:

a) La chica supere los 170 cm de altura.

b) El chico mida más de 170 cm.

c) Los dos superen los 170 cm.

d) Al menos uno de ellos supere los 170 cm.

e) La chica sea más alta que el chico.

Considerando las variables aleatorias:
X: "estatura de las alumnas de un centro escolar", cuya distribución es: $X \sim N(\mu_X = 165\ ; \sigma_X = 10)$
Y: "estatura de los alumnos de un centro escolar", con distribución: $Y \sim N(\mu_Y = 175\ ; \sigma_Y = 12)$

a) La probabilidad de que la chica mida más de 170 cm es:

$$P(X > 170) = P\left(Z > \frac{170 - 165}{10}\right) = P(Z > 0,5) = 1 - \Phi(0,5) = 1 - 0,6915 = 0,3085$$

b) La probabilidad de que el chico mida más de 170 cm es:

$$P(Y > 170) = P\left(Z > \frac{170 - 175}{12}\right) = P(Z > -0,4167) = P(Z < 0,42) = \Phi(0,42) = 0,6628$$

c) Entendiendo que la elección del chico es independiente de la elección de la chica, la probabilidad de que ambos superen los 170 cm es:

$$P((X > 170) \cap (Y > 170)) = P(X > 170) \cdot P(Y > 170) = 0,3085 \cdot 0,6628 = 0,2045$$

d) La probabilidad de que al menos uno de ellos supere los 170 cm es:

$$P((X > 170) \cup (Y > 170)) = P(X > 170) + P(Y > 170) - P((X > 170) \cap (Y > 170)) =$$
$$= 0,3085 + 0,6628 - 0,2045 = 0,7668$$

Donde el valor de cada uno de los término ha sido calculado en los apartados anteriores.

e) En este caso se pide la probabilidad de que $P(X > Y) = P(X - Y > 0)$.

Suponiendo que las variables *X* e *Y* son independientes, la distribución de probabilidad de la variable $Y - X$ es normal de media $\mu = \mu_X - \mu_Y = 165 - 175 = -10$ y varianza $\sigma^2 = \sigma_X^2 + \sigma_Y^2 = 10^2 + 12^2 = 244$.

$X - Y \sim N\left(\mu = -10; \sigma = \sqrt{244}\right)$

De manera que la probabilidad de que la chica sea más alta que el chico es:

$$P(X - Y > 0) = P\left(Z > \frac{0 - (-10)}{\sqrt{244}}\right) = P(Z > 0,64) = 0,26109$$

64. En el control de calidad de un gran lote de artículos manufacturados, se estima que el 12% de los artículos tiene algún defecto. Si se examinan 10 artículos del lote, calcula la probabilidad de que no se obtengá más de un artículo defectuoso:

a) Mediante la distribución binomial.

b) Mediante la aproximación normal a la distribución binomial

c) Comenta la diferencia obtenida en los apartados anteriores..

Sea la variable aleatoria *X*: "número de artículos, de los 10 elegidos, que tiene algún defecto".
X tiene una distribución de probabilidad $\text{Bin}(n=10,\ p=0{,}12)$

a) Calculando directamente con las probabilidades de la binomial

$$P(X\le 1)=P(X=0)+P(X=1)=\binom{10}{0}0{,}88^{10}+\binom{10}{1}0{,}12\cdot 0{,}88^{10}=0{,}2785+0{,}3798=0{,}6583$$

b) Utilizando la aproximación por la variable aleatoria $Y \sim N(\mu=1{,}2;\sigma=1{,}028)$ ya que:

$$\mu=np=10\cdot 0{,}12=1{,}2 \quad \text{y} \quad \text{Var}(X)=np(1-p)=10\cdot 0{,}12\cdot 0{,}88=1{,}56\Rightarrow \sigma=\sqrt{1{,}56}=1{,}028$$

$$P(X\le 1)\approx P(Y\le 1{,}5)=P\left(Z\le\frac{1{,}5-1{,}2}{1{,}028}\right)=P(Z\le 0{,}29)=\Phi(0{,}29)=0{,}6141$$

c) Se puede ver que la aproximación obtenida en el aparatado b) no es muy buena, debido a que se tienen pocas observaciones (solo 10) y a que la probabilidad $p=0{,}12$, está alejada de 0,5.

65. El encargado de una plantación de chopos asegura que, en este momento, el diámetro de los árboles sigue una distribución normal de media 20 cm y que el 90 % de ellos tiene un diámetro inferior a 25 cm.

a) Calcula la desviación típica de la distribución.

b) Calcula la probabilidad de que un árbol tenga más de 22 cm de diámetro.

a) Sea la variable aleatoria *X*: "diámetro, en cm, de los árboles de la plantación". Se sabe que $X \sim N(\mu=20;\ \sigma)$.
Como $P(X<25)=0{,}9$, entonces, tipificando:

$$P(X<25)=P\left(Z<\frac{25-20}{\sigma}\right)=0{,}9\Rightarrow\frac{5}{\sigma}=1{,}282\Rightarrow\sigma=3{,}9$$

b) Con la desviación típica calculada en el apartado anterior:

$$P(X>22)=P\left(Z>\frac{22-20}{3{,}9}\right)=P(Z>0{,}5128)=1-P(Z<0{,}51)=1-0{,}6950=0{,}3050$$

66. En cierto hospital se realizan al mes 50 intervenciones quirúrgicas de alto riesgo. La cuarta parte de las mismas, por término medio, presenta complicaciones y, en ese caso, es preciso disponer de plasma adicional cuya vida útil es de un mes.

Calcula la reserva de bolsas de plasma que es preciso tener para cubrir el 90 % de los meses las complicaciones que se pudieran presentar.

Sea la variable aleatoria *X*: "número de bolsas de plasma adicional" (número de operaciones de alto riesgo con complicaciones). Su distribución de probabilidad es $X\sim \text{Bin}(n=50; p=0{,}25)$, y su esperanza y varianza son respectivamente:

$E[X]=np=50\cdot 0{,}25=12{,}5$ $\qquad$ $\text{Var}(X)=npq=50\cdot 0{,}25\cdot 0{,}75=9{,}375$

La distribución de *X* puede aproximarse por la de una variable aleatoria normal: $Y\sim N\left(\mu=12{,}5;\sigma=\sqrt{9{,}375}\right)$.

De modo que si se llama *c* a la cantidad de bolsas de plasma adicionales que se necesitarán para cubrir con probabilidad 0,9 las demandas, se plantea:

$$P(X\ge c)\approx P(Y\ge c-0{,}5)=0{,}9\Rightarrow P\left(Z\ge\frac{c-0{,}5-12{,}5}{\sqrt{9{,}375}}\right)=0{,}9$$

Es decir,

$$1-\Phi\left(\frac{c-13}{3{,}062}\right)=0{,}9\Rightarrow\frac{c-13}{3{,}062}=-1{,}28\Rightarrow c=9{,}08$$

Se precisarán entre 9 y 10 bolsas adicionales para tener cubiertas las necesidades en el 90% de los meses.

67. En un centro educativo, a pesar de los controles rigurosos, un 12 % de los ordenadores resulta infectado por algún tipo de virus informático.

a) Si en un aula hay 10 ordenadores, calcula la probabilidad de que más de un ordenador tenga virus.

b) Si se quiere que la probabilidad de que haya, como máximo, dos ordenadores infectados sea al menos 0,7 ¿cuál tiene que ser el número máximo de ordenadores en el aula?

c) Si en todo el centro el número de ordenadores es 150, ¿cuál es la probabilidad de que por lo menos el 10% de ellos tenga virus?

a) Sea la variable aleatoria X: "número de ordenadores infectados, de los 10". La variable $X \sim \text{Bin}(n=10;\ p=0{,}12)$, entonces:

$$P(X>1)=1-P(X\le 1)=1-P(X=0)-P(X=1)=1-\binom{10}{0}\cdot 0{,}12^0\cdot 0{,}88^{10}-\binom{10}{1}\cdot 0{,}12^1\cdot 0{,}88^9=$$
$$=1-0{,}2785-0{,}3798=0{,}3417$$

b) En este caso, se trata de calcular el mayor valor de n, en la binomial, para que $P(X\le 2)\ge 0{,}7$. Para ello se necesitan las probabilidades:

$$P(X=0)=0{,}88^n \qquad P(X=1)=n\cdot 0{,}12\cdot 0{,}88^{n-1} \qquad P(X=2)=\frac{n\cdot(n-1)}{2}0{,}12^2\,0{,}88^{n-2}$$

Se puede resolver sondeando los posibles valores de n.
Se construye la tabla siguiente, con las probabilidades necesarias para obtener $P(X\le 2)$ .Se puede empezar probando con $n=8$ y se observa que como máximo debe haber $n=11$ ordenadores en el aula para que la probabilidad de tener dos o menos ordenadores infectados sea superior a 0,7.

n	$P(X=0)$	$P(X=1)$	$P(X=2)$	$P(X\le 2)$
8	0,3596	0,3923	0,1057	0,8577
9	0,3164	0,3884	0,1087	0,8136
10	0,2785	0,3798	0,1087	0,7670
11	0,2450	0,3676	0,1063	0,7190
12	0,2157	0,3529	0,1020	0,6706

c) Ahora la variable Y: "número de ordenadores infectados, de los 150". $Y \sim \text{Bin}(n=150;\ p=0{,}12)$.

La esperanza y la varianza de Y son, respectivamente : $E[Y]=150\cdot 0{,}12=18$ y $\text{Var}(X)=150\cdot 0{,}12\cdot 0{,}88=15{,}84$

La distribución de Y se puede aproximar por la de la variable $T \sim N(\mu=18;\ \sigma=\sqrt{15{,}84})$. Entonces, como el 10 % de 150 es 15, se tiene que:

$$P(Y\ge 15)\approx P(T\ge 14{,}5)=P\left(Z\ge\frac{14{,}5-18}{\sqrt{19{,}125}}\right)=P(Z\ge -0{,}88)=\Phi(0{,}88)=0{,}8106$$

68. El cociente de inteligencia es una variable aleatoria con distribución $N(\mu = 100, \sigma = 16)$.

a) Calcula la probabilidad de que un individuo elegido al azar tenga un cociente superior a 115.

b) Si se escoge una muestra de 85 individuos, ¿cuántos se espera que tengan un cociente inferior a 95?

c) Si se sabe que una persona tiene un cociente superior a 100, calcula la probabilidad de que su cociente sea mayor de 120.

a) La probabilidad pedida se obtiene de la siguiente manera:

$$P(X > 115) = P\left(Z > \frac{115-100}{16}\right) = P(Z > 0{,}9375) = 1 - P(Z < 0{,}9375) = 1 - \Phi(0{,}94) = 1 - 0{,}8264 = 0{,}1736$$

b) Calculamos primero la siguiente probabilidad

$$P(X < 95) = P\left(Z < \frac{95-100}{16}\right) = P(Z < -0{,}3125) = 1 - P(Z \geq -0{,}3125) = 1 - \Phi(0{,}32) = 1 - 0{,}6255 = 0{,}3745.$$

Por tanto, aproximadamente el 37,45% de los 85 individuos (31,83) tendrán un cociente inferior a 95, esto es, unos 32 individuos.

c) En este caso, se trata de una probabilidad condicionada.

$$P(X > 120 | X > 100) = \frac{P(X > 120)}{P(X > 100)} = \frac{0{,}1056}{0{,}5} = 0{,}2113$$

Donde la probabilidad del numerador se ha obtenido como sigue

$$P(X > 120) = P\left(Z > \frac{120-100}{16}\right) = P(Z > 1{,}25) = 1 - P(Z < 1{,}25) = 1 - 0{,}8944 = 0{,}1056$$

69. La anchura de las hojas de una especie de árbol siguen una distribución normal de media $\mu = 4$ cm. Además, se ha observado experimentalmente que la anchura del 90 % de las hojas es inferior a 5 cm.

a) Halla la varianza de la distribución.

b) Calcula la probabilidad de que la anchura de una hoja elegida al azar sea mayor de 6 cm.

c) ¿Cuál es la anchura máxima de las hojas que se encuentran entre el 30% con menor anchura?

a) Considerando la variable aleatoria *X*: "anchura de las hojas, en cm", cuya distribución es $N(\mu = 4; \sigma)$, con σ desconocida; se sabe que, elegida una hoja al azar, $P(X < 5) = 0{,}9$.

Para el cálculo de la varianza se plantea:

$$P(X < 5) = 0{,}9 \Rightarrow P\left(Z < \frac{5-4}{\sigma}\right) = 0{,}9 \Rightarrow \Phi\left(\frac{1}{\sigma}\right) = 0{,}9 \Rightarrow \frac{1}{\sigma} = 1{,}281 \Rightarrow \sigma = 0{,}7806$$

Luego, la varianza es $\sigma^2 = 0{,}6093$

b) Con los datos disponibles, $P(X > 6) = P\left(Z > \frac{6-4}{0{,}7806}\right) = P(Z > 2{,}56) = 1 - \Phi(2{,}56) = 1 - 0{,}9948 = 0{,}0052$

c) Necesitamos localizar la anchura a tal que $P(X \leq a) = 0{,}3$, esto es:

$$P(X \leq a) = P\left(Z \leq \frac{a-4}{0{,}7806}\right) = 0{,}3 \Rightarrow P\left(Z \geq \frac{a-4}{0{,}7806}\right) = 0{,}7 \Rightarrow P\left(Z \leq -\frac{a-4}{0{,}7806}\right) = 0{,}7 \Rightarrow \frac{4-a}{0{,}7806} = 0{,}525$$

Por tanto, $a = 4 - 0{,}525 \cdot 0{,}7806 = 3{,}59$.

La anchura máxima que tendrá el 30 % de menor anchura es 3,59 cm.

70. En la segunda vuelta de las elecciones presidenciales, el candidato A obtuvo el 52 % de los votos emitidos. El resto votó a otro candidato o lo hizo en blanco. Si de la población que ha participado en la votación se eligen al azar 2000 personas, calcula la probabilidad de que entre estos:

a) Más del 60% haya votado al candidato A.

b) Menos de la mitad haya votado al candidato A.

c) Más del 60% haya votado al candidato A si se sabe que por lo menos la mitad le votó.

Sea la variable aleatoria X: "número de personas, de las 2000, que han votado al candidato A".

La distribución de X es $\text{Bin}(n = 2000; p = 0,52)$, cuya esperanza y varianza son respectivamente:

$E[W] = 2000 \cdot 0,52 = 1040$ $\qquad$ $\text{Var}(X) = 2000 \cdot 0,52 \cdot 0,48 = 499,2 \Rightarrow \sigma = \sqrt{499,2} = 22,34$

De modo que la distribución de X puede aproximarse por la distribución de la variable $W \sim N(\mu = 1040; \sigma = 22,34)$

a) El 60% de 2000 son 1200 personas, luego:

$$P(X > 1200) \cong P(W > 1200,5) = P\left(Z > \frac{1200,5 - 1040}{22,35}\right) = P(Z > 7,18) = 1 - \Phi(7,18) = 1 - 1 = 0.$$

b) En este caso,

$$P(X < 1000) \approx P(W \le 999,5) = P\left(Z < \frac{999,5 - 1040}{22,35}\right) = P(Z < -1,81) = 1 - \Phi(1,81) = 1 - 0,9649 = 0,0351$$

c) La probabilidad de que al menos la mitad de los 2000 haya votado al candidato A es:

$$P(X \ge 1000) = 1 - P(X < 1000) = 1 - 0,0351 = 0,9649$$

Entonces,

$$P(X > 1200 \mid X \ge 1000) = \frac{P(X > 1200; X \ge 1000)}{P(X \ge 1000)} = \frac{P(X > 1200)}{P(X \ge 1000)} = \frac{0}{0,9649} = 0$$

71. Las calificaciones obtenidas por los estudiantes para acceder a una facultad siguen una distribución normal de media 6,5 y desviación típica 1,5. Si la nota de corte se estableció en 8,2,

a) ¿Cuál es el porcentaje de estudiantes que no pudo acceder a la facultad ese año?

b) Suponiendo que se presentaron 300 solicitudes, ¿Cuántos estudiantes con nota superior a 7,5 se quedaron sin plaza?

La distribución de la variable aleatoria X: "calificaciones obtenidas por los alumnos para acceder a la universidad" es $N(\mu = 6,5;\ \sigma = 1,5)$

a) Dado que la nota de corte se estableció en 8,2, la probabilidad de que un alumno elegido al azar no haya obtenido plaza es:

$$P(X < 8,2) = P\left(Z < \frac{8,2 - 6,5}{1,5}\right) = P(Z < 1,133) = 0,8708$$

Luego, el 87,08% de los estudiantes no pudieron acceder a la facultad ese año.

b) La probabilidad de que un alumno elegido al azar haya obtenido nota superior a 7,5 y se haya quedado sin plaza es:

$$P(7,5 < X < 8,2) = P\left(\frac{7,5 - 6,5}{1,5} < Z < \frac{8,2 - 6,5}{1,5}\right) = \Phi(1,13) - \Phi(0,67) = 0,8708 - 0,7486 = 0,1222$$

c) Es decir, el 12,22 % obtuvo nota superior a 7,5 y no accedió a la facultad. El 12,22% de 300 es 36,66. Luego unos 37 alumnos se quedaron sin plaza habiendo obtenido nota superior a 7,5 puntos.

72. El nivel de colesterol total en sangre sigue una distribución normal de media 180 mg/dL y varianza 225mg²/dL² . Se considera que valores del nivel de colesterol mayores de 200 mg/dL son perjudiciales para la salud y debe iniciarse un tratamiento

a) Elegidas 200 personas al azar, ¿cuál es el número esperado de personas que necesitarán tratamiento?

b) Elegida una persona al azar, ¿cuál es la probabilidad de que su nivel de colesterol sea inferior a 170 mg/dL, si se sabe que no precisa tratamiento?

c) Si la varianza se mantiene en su actual valor, calcula qué valor medio debe tener el nivel de colesterol para que solo el 5% de la población deba seguir tratamiento médico.

a) Sea la variable aleatoria X: "nivel de colesterol en sangre, en mg/dL", cuya distribución es $N\left(\mu = 180;\ \sigma = \sqrt{225}\right)$

La probabilidad de que una persona elegida al azar en esta población deba ponerse en tratamiento, por superar los 200 mg/dL es: $P(X > 200) = P\left(Z > \dfrac{200-180}{\sqrt{225}}\right) = P(Z > 1{,}33) = 1 - \Phi(1{,}33) = 1 - 0{,}9082 = 0{,}0918$.

El número esperado de las 200 personas que deben ponerse en tratamiento es $200 \cdot 0{,}0918 = 18{,}36$ .Aproximadamente 18 personas.

b) La probabilidad de que su nivel de colesterol sea inferior a 170 mg/dL, si se sabe que no precisa tratamiento viene dada

por: $P(X < 170 \mid X < 200) = \dfrac{P(X < 170 \cap X < 200)}{P(X < 200)} = \dfrac{P(X < 170)}{P(X < 200)} = \dfrac{P\left(Z < \dfrac{170-180}{\sqrt{225}}\right)}{1 - 0{,}0918} = \dfrac{P(Z < -0{,}67)}{0{,}9082} =$

$= \dfrac{1 - P(Z < 0{,}67)}{0{,}9082} = \dfrac{1 - \Phi(0{,}67)}{0{,}9082} = \dfrac{1 - 0{,}7486}{0{,}9082} = 0{,}2768$

c) Si la varianza sigue siendo 225; y el límite para iniciar tratamiento es de 200 mg/dL, el nivel medio de colesterol,μ,de la población para que solo el 5% deba seguir tratamiento se obtiene a partir de que $P(X > 200) = 0{,}05$, que tipificando resulta: $P\left(Z > \dfrac{200-\mu}{\sqrt{225}}\right) = 0{,}05 \Rightarrow P\left(Z < \dfrac{200-\mu}{\sqrt{225}}\right) = 0{,}95$

Y de la distribución normal estándar resulta: $\dfrac{200-\mu}{\sqrt{225}} = 1{,}645 \Rightarrow \mu = 175{,}325$ mg/dL

73. Las horas de vuelo útiles (en miles) de una determinada clase de aviones comerciales siguen una distribución normal, con una media 200 y desviación típica 20.

a) Elegido un avión al azar, ¿cuál es la probabilidad de que su vida útil supere las 250 000 horas de vuelo?

b) Calcula un intervalo alrededor de 200 de tal manera que el 50% de los aviones de esta clase tenga su número de horas de vuelo útil incluido en el intervalo.

c) ¿Cuántas horas de vuelo tendrá un avión que supere en horas de vuelo al 90% de la flota?

a) Sea la variable aleatoria X: "horas de vuelo útiles (en miles) de una determinada clase de aviones comerciales", con distribución de probabilidad $N(\mu = 200;\ \sigma = 20)$ .Elegido un avión al azar, la probabilidad de que su vida útil supere las 250 mil horas es: $P(X > 250) = P\left(Z > \dfrac{250-200}{20}\right) = P(Z > 2{,}5) = 1 - \Phi(2{,}5) = 1 - 0{,}9938 = 0{,}0062$.

b) El intervalo tiene la forma $(200-\delta, 200+\delta)$, y tipificando: $P(200-\delta < X < 200+\delta) = 0{,}5$ .Simplificando, se tiene: $P\left(\dfrac{-\delta}{20} < Z < \dfrac{\delta}{20}\right) = 0{,}5 \Rightarrow \Phi\left(\dfrac{\delta}{20}\right) - \Phi\left(\dfrac{-\delta}{20}\right) = 0{,}5 \Rightarrow \Phi\left(\dfrac{\delta}{20}\right) - 1 + \Phi\left(\dfrac{\delta}{20}\right) = 0{,}5 \Rightarrow 2\Phi\left(\dfrac{\delta}{20}\right) = 1{,}5$.

Por último, $\Phi\left(\dfrac{\delta}{20}\right) = 0{,}75 \Rightarrow \dfrac{\delta}{20} = 0{,}6745 \Rightarrow \delta = 13{,}49$, luego el intervalo resultante es: $(200 - 13{,}49;\ 200 + 13{,}49) = (186{,}51;\ 213{,}49)$

c) En este caso, el número de horas de vuelo, a, de un avión que supere al 90% de la flota es: $P(X < a) = P\left(Z < \dfrac{a-200}{20}\right) = 0{,}9$ .De las tablas de la normal estándar: $\dfrac{a-200}{20} = 1{,}28 \Rightarrow a = 225{,}6$.

El avión tiene que tener al menos 225 600 horas de vuelo.

74. Una empresa fabrica minas de grafito para portaminas cuya longitud sigue una distribución $N(\mu = 30;\ \sigma = 0,5)$ en mm. Otra empresa produce los envases para las minas con una longitud que sigue una distribución $N(\mu = 31,5;\ \sigma = 0,4)$. Calcula el porcentaje de minas que no caben en el envase.

Sea *X*: "longitud de las minas de grafito, en mm". Su distribución de probabilidad es $N(\mu = 30; \sigma = 0,5)$.

Sea *Y*: "longitud de los envases para las minas, en mm", cuya distribución es $N(\mu = 31,5; \sigma = 0,4)$.

Suponiendo que los envases se fabrican de forma independiente de las minas, la distribución de $X - Y$ es normal de media $\mu = \mu_X - \mu_Y = 30 - 31,5 = -1,5$ y varianza $\sigma^2 = \sigma_X^2 + \sigma_Y^2 = 0,5^2 + 0,4^2 = 0,41 \Rightarrow \sigma = \sqrt{0,41}$
luego: $X - Y \sim N\left(\mu = -1,5; \sigma = \sqrt{0,41}\right)$.

Se debe calcular la probabilidad de que una mina elegida al azar no quepa en un envase también elegido al azar; esto es $P(X > Y)$. O lo que es lo mismo $P(X - Y > 0)$:

$$P(X - Y > 0) = P\left(Z > \frac{0 - (-1,5)}{\sqrt{0,41}}\right) = P(Z > 2,34) = 1 - \Phi(2,34) = 1 - 0,9904 = 0,0096$$

Luego el 0,96% de las minas de grafito fabricadas no cabe en el envase.

75. Una empresa dedica dos máquinas a producir largueros para cama de dos metros de largo. La máquina A produce la tercera parte de los largueros con una longitud que sigue una distribución normal $N(\mu = 200, \sigma = 1)$ en centímetros. La máquina B produce el resto de largueros cuya longitud sigue una distribución $N(\mu = 199, \sigma = 2)$ también en centímetros. De la producción total elegimos al azar un larguero, calcula la probabilidad de que:

a) Su longitud sea menor que 202 cm.

b) El larguero haya sido producido por la segunda máquina, si su longitud es menor de 202 cm.

a) Considerando la variable aleatoria *L*: "longitud, en cm, de los largueros para cama producidos por la empresa".

Sean los sucesos:

A: "el larguero se ha fabricado con la máquina A" y *B*: "el larguero se ha fabricado con la máquina B"

Se sabe que $P(A) = \frac{1}{3}$ y $P(B) = \frac{2}{3}$.

Para calcular la probabilidad $P(L < 202)$ se aplica el teorema de la probabilidad total:

$P(L < 102) = P(L < 202|A) \cdot P(A) + P(L < 202|B) \cdot P(B)$ donde:

$$P(L < 202|A) = P(N(200,1) < 202) == P\left(Z < \frac{202 - 200}{1}\right) = P(Z < 2) = \Phi(2) = 0,9772$$

$$P(L < 202|B) = P(N(199,2) < 202) = P\left(Z < \frac{202 - 199}{2}\right) = P(Z < 1,5) = \Phi(1,5) = 0,9332$$

De manera que: $P(L < 202) = 0,9772 \cdot \frac{1}{3} + 0,9332 \cdot \frac{2}{3} = 0,9479$

b) Aplicando el teorema de Bayes: $P(B \mid L < 202) = \frac{P(L < 202 \mid B)P(B)}{(P(L < 202)} = \frac{0,9332 \cdot \frac{2}{3}}{0,9479} = 0,6563$.

ENTORNO MATEMÁTICO

Probabilidad de que una mujer joven tenga hijos

Esteban ha leído, no sin cierta sorpresa, un estudio que afirma que en los países desarrollados las mujeres (y los hombres) tienen sus hijos cada vez más tarde. Es "raro" que las mujeres tengan hijos antes de los 30 años.

Está extrañado porque nunca se había parado a pensar que su madre, ya que él tiene 17 años y su madre 42, le había tenido con 25 años y a su hermana mayor con 23. Él a esa edad piensa estar todavía en la Universidad y de tener hijos… bueno, de eso, de momento mejor ni hablar.

Como le ha llamado la atención la diferencia de edades a la hora de tener hijos entre su madre y las madres actuales ha decidido investigar un poco para comprobar la veracidad del estudio y de paso poner a prueba lo que él cree son unos extensos conocimientos de estadística y probabilidad.

Ha recogido información del INE con el número de nacimientos (en miles) del primer hijo según la edad de la madre y ha dividido los datos en intervalos (clases) de 5 años, recogiendo el número de nacimientos en cada intervalo para el año en curso y el año en que él nació. Y ha elaborado una tabla como la que se muestra.

Haz los cálculos que haría Esteban y comprueba si el estudio es cierto o no. Para ellotendrás que calcular la media de cada distribución.

¿Cuál era la probabilidad de que una mujer elegida al azar en España fuera madre antes de los 30 años hace 17 años? ¿Y cómo es esta probabilidad en la actualidad?

Para calcular las probabilidades pedidas con esta información, podrías ajustar los datos de las distribuciones de cada año a una distribución normal. ¿Sería adecuado? En caso afirmativo, ¿cuáles serían la media y la desviación típica?

Edad madre	Nacimientos (año actual)	Nacimientos (hace 17 años)
[15,20)	35	20
[20, 25)	93	90
[25, 30)	174	232
[30, 35)	241	184
[35, 40)	122	45
[40, 45)	27	7
[45, 50)	2	0

Se comienza por ajustar una distribución normal a los datos de la tabla.
En primer lugar se calculan la media y la varianza de la distribución de frecuencias observada para el año actual

X Edad madre	f_j Nacimientos (por 1000)	Marca clase (x_j)	f_jx_j	$f_jx_j^2$
[15,20)	35	17,5	612,5	10718,75
[20, 25)	93	22,5	2092,5	47081,25
[25, 30)	174	27,5	4785	131587,5
[30, 35)	241	32,5	7832,5	254556,25
[35, 40)	122	37,5	4575	171562,5
[40, 45)	27	42,5	1147,5	48768,75
[45, 50)	2	47,5	95	4512,5
	694		21140	668787,5

De manera que la media y la varianza son:

$$\bar{X} = \frac{\sum_{j=1}^{7} f_j x_j}{N} = \frac{21140}{694} = 30,461 \qquad \text{Var}(X) = \frac{\sum_{j=1}^{7} f_j x_j^2}{N} - \bar{X}^2 = \frac{668787,5}{694} - 30,461^2 = 35,792$$

Ajustando una distribución normal $N(\mu = 30{,}461;\ \sigma = 5{,}983)$, se pueden calcular las probabilidades de cada uno de los intervalos de la distribución de frecuencias y, a partir de esas probabilidades, las frecuencias teóricas:

X Edad madre	f_j Nacimientos (por 1000)	Probabilidades	Frecuencias teóricas
[15,20)	35	0,0215	15
[20, 25)	93	0,1359	94
[25, 30)	174	0,3413	237
[30, 35)	241	0,3413	237
[35, 40)	122	0,1359	94
[40, 45)	27	0,02115	15
[45, 50)	2	0	0
	694		692

Que puede considerarse un ajuste razonable de la distribución normal (teórica) a la distribución de frecuencias observada.Si hacemos lo mismo para los nacimientos hace 17 años tenemos:

Y Edad madre	f_j Nacimientos (por 1000)	Marca clase (y_j)	f_jy_j	$f_jy_j^2$
[15,20)	20	17,5	350	6125
[20, 25)	90	22,5	2025	45562,5
[25, 30)	232	27,5	6380	175450
[30, 35)	184	32,5	5980	194350
[35, 40)	45	37,5	1687,5	63281,25
[40, 45)	7	42,5	297,5	12643,75
[45, 50)	0	47,5	0	0
	578		16720	497412,5

De manera que la media y la varianza son:

$$\overline{Y} = \frac{\sum_{j=1}^{7} f_j y_j}{N} = \frac{16720}{578} = 28{,}927$$

$$\text{Var}(X) = \frac{\sum_{j=1}^{7} f_j y_j^2}{N} - \overline{Y}^2 = \frac{497412{,}5}{578} - 28{,}927^2 = 23{,}785$$

Ajustando una distribución normal $Y \sim N(\mu = 28{,}927; \sigma = 4{,}877)$, se pueden calcular las probabilidades de cada uno de los intervalos de la distribución de frecuencias y, a partir de esas probabilidades, las frecuencias teóricas:

X Edad madre	f_j Nacimientos (por 1000)	Probabilidades	Frecuencias teóricas
[15,20)	20	0,0215	12
[20, 25)	90	0,1359	79
[25, 30)	232	0,3413	197
[30, 35)	184	0,3413	197
[35, 40)	45	0,1359	79
[40, 45)	7	0,02115	12
[45, 50)	0	0	0
	578		576

Que puede considerarse también un ajuste razonable pero no especialmente bueno de la distribución normal (teórica) a la distribución de frecuencias observada. De esta manera, la probabilidad de que una mujer elegida al azar, haya sido madre antes de los 30 años es: $P(Y < 30) = P\left(Z < \frac{30 - 28{,}927}{4{,}877}\right) = P(Z < 0{,}22) = 0{,}5871.$

El mismo cálculo hecho para el año actual ofrece el siguiente resultado:

$$P(X < 30) = P\left(Z < \frac{30 - 30{,}461}{5{,}983}\right) = P(Z < -0{,}08) = 1 - 0{,}5319 = 0{,}4681$$

Es decir, solo el 46,81 % de las mujeres es madre antes de los 30 años de edad frente al 58,71 % de hace 17 años.

La fabricación de prendas de vestir y la distribución normal

Una empresa textil va a producir los pantalones de los uniformes de los 3000 empleados de unos grandes almacenes.

La empresa tiene un problema de decidir cuántos pantalones de cada talla debe producir. ¿Qué crees que necesita saber? ¿Cuáles son los datos que debe pedir a los responsables de los grandes almacenes?

Las medidas corporales suelen ajustarse de manera precisa mediante una distribución normal. Por ello, puedes suponer que la talla de los pantalones que se deben fabricar se puede ajustar por una distribución normal. Infórmate y asigna una talla media y una desviación típica.

Si las tallas van de la 32 a la 50, por ejemplo, ¿cuántos pantalones de cada talla debe fabricar la empresa?

Recuerda que los datos clave son la media, la desviación típica o varianza de la normal y el número de prendas que hay que fabricar.

Advertencia importante: La talla es una variable discreta: 32, 34, 36, etc… ¿Cómo se hace para ajustar una normal?

Si la talla media es 43, la desviación típica 4 y 3000 son las prendas a encargar, para estimar cuántas hay que encargar de cada talla desde la 32 a la 56, se puede proceder de la siguiente manera

Se habilita un intervalo de radio 1 centrado en cada una de las tallas: [31, 33], [33, 35], etc. Como la distribución normal es continua, no importa que los intervalos sean abiertos o cerrados.

Se tipifican todos los valores de los intervalos (Si se utiliza una hoja de cálculo, este paso no sería necesario, porque las probabilidades se pueden calcular directamente).

Se calcula la probabilidad de cada intervalo con la distribución Normal (tipificada o no).

Multiplicando por 3000 se obtiene la frecuencia correspondiente a cada intervalo (o sea, a cada talla).

Se redondea la frecuencia anterior al entero más próximo.

La tabla siguiente contiene el proceso completo.

Talla	Intervalo		Intervalo tipificado		Probabilidad	frec. estimada	Encargo
32	31	33	–3	–2,5	0,0049	14,7	15
34	33	35	–2,5	–2	0,0166	49,8	50
36	35	37	–2	–1,5	0,044	132	132
38	37	39	–1,5	–1	0,0919	275,7	276
40	39	41	–1	–0,5	0,1498	449,4	449
42	41	43	–0,5	0	0,1915	574,5	575
44	43	45	0	0,5	0,1915	574,5	575
46	45	47	0,5	1	0, 1498	449,4	449
48	47	49	1	1,5	0,0919	275,7	276
50	49	51	1,5	2	0,044	132	132
52	51	53	2	2,5	0,0166	49,8	50
54	53	55	2,5	3	0,0049	14,7	15
56	55	57	3	3,5	0,0011	3,3	3
					0,997 611	2992,83	2991,00

NOTA FINAL:
Como la suma de todas las frecuencias será algo inferior a 3000, en el paso 4 se puede multiplicar por un número algo mayor de 3000, por ejemplo 3010.

Si la varianza es mayor, se puede probar con 20 o con 25, de forma que la suma de las frecuencias se aleja más de 3000 que en el caso que hemos mostrado. Ello es debido a que tallas por debajo de 32 y por encima de 56 tienen frecuencias más altas en este caso. Para llegar a 3000, basta con multiplicar en el paso 4 por un número más alto que 3000 y ajustar.

AUTOEVALUACION

Comprueba qué has aprendido

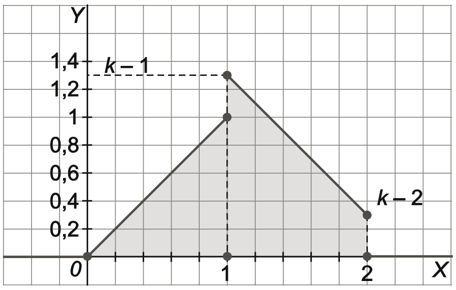
1. Sea X una variable aleatoria con función de densidad $f(x)=\begin{cases} x & \text{si } 0\le x<1 \\ k-x & \text{si } 1\le x\le 2 \\ 0 & \text{en el resto} \end{cases}$.

a) Calcula el valor de k y dibuja la gráfica de $f(x)$.

b) Comprueba que es función de densidad.

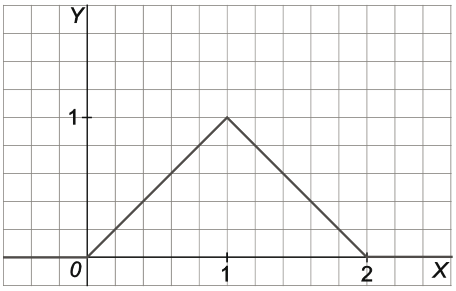
c) Calcula la probabilidad $P(0,5<X\le 1,5)$.

a) Para que sea siempre positiva el valor de k debe ser tal que $k-x>0$ si $1<x<2 \Rightarrow k\ge 2$.

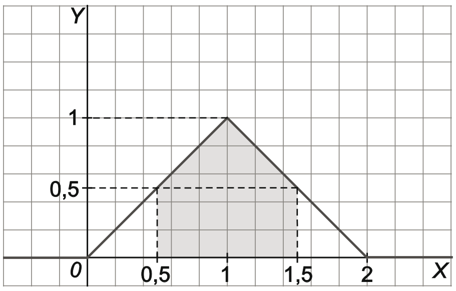


Por otro lado el área bajo la gráfica debe ser 1, esto es, la suma de las áreas del triángulo y el trapecio que pueden verse en la imagen. Por tanto: $\frac{1\cdot 1}{2}+\frac{(k-1+k-2)\cdot 1}{2}=1\Rightarrow \frac{1+2k-3}{2}=1\Rightarrow k=2$.

b) La gráfica resultante es:



c) La probabilidad pedida es el área de la región coloreada en el gráfico siguiente.



Área rectángulo: $1\cdot 0,5=0,5$;Área triángulo: $\frac{1\cdot 0,5}{2}=0,25$. De modo que $P(0,5<X\le 1,5)=0,5+0,25=0,75$.

2. El peso en kg de los bebés nacidos en una maternidad sigue una distribución $N(\mu = 3,\ \sigma = 0,2)$. ¿Cuál será el peso que tiene un bebé si es superior al peso del 90% de los bebés?

Sea a el peso de este bebé, entonces $P(X < a) = 0,9 \Rightarrow P\left(Z < \frac{a-3}{0,2}\right) = 0,9 \Rightarrow \frac{a-3}{0,2} = 1,28 \Rightarrow a = 3,256$ kg

3. En una gran urbanización, el consumo de energía eléctrica por hogar cada dos meses,en kW/h sigue una distribución normal $N(\mu = 60,\ \sigma = 8)$. De la urbanización se elige un hogar al azar.

a) ¿Cuántos kW/h. tendría que consumir para pertenecer al 5% que más consume?

b) ¿Qué proporción de hogares de la urbanización consume menos de 50 kW/h?

c) Si se sabe que un hogar supera la media de consumo, ¿cuál es la probabilidad de que lo supere en un 20%?

a) La variable aleatoria X: "consumo de energía, cada dos meses, en la urbanización" tiene una distribución $N(\mu = 60;\ \sigma = 8)$ en kW/h.

Si c son los kW/h que consume el hogar elegido, para pertenecer al 5% que más consume, debe ser:

$$P(X > c) \geq 0,05 \Rightarrow P(X < c) = 0,95 \Rightarrow P\left(Z \leq \frac{c-60}{8}\right) = 0,95 \Rightarrow c = 73,2 \text{ kW / h}$$

Debe consumir al menos 73,2 kW/h.

b) La probabilidad de que el hogar elegido al azar consuma menos de 50 Kw/h es:

$$P(X < 50) = P\left(Z < \frac{50-60}{8}\right) = P(Z < -1,25) = 1 - \Phi(1,25) = 1 - 0,8944 = 0,1056$$

Es decir, en torno al 10,56 % de los hogares de la urbanización consume menos de 50 kw/h.

c) Si se sabe que un hogar supera la media de consumo, que lo supere en un 20% implica un consumo mayor que 72 kW/h (ya que el 20% de 60 es 12). Por tanto la probabilidad pedida es:

$$P(X > 72 \mid X > 60) = \frac{P(X > 72 \cap X > 60)}{P(X > 60)} = \frac{P(X > 72)}{0,5} = \frac{P\left(Z > \frac{72-60}{8}\right)}{0,5} = \frac{1 - P(Z < 1,5)}{0,5} = \frac{0,0668}{0,5} = 0,1336$$

4. Si $X \sim N(\mu, \sigma)$ y $Z \sim N(0, 1)$, calcula μ y σ sabiendo que $P(X \geq 8) = P(Z \geq 0,5)$ y que $P(X \geq 12) = P(Z \geq 1,5)$.

Teniendo en cuenta el proceso de tipificación se tiene que $P(X \geq 8) = P(Z \geq 0,5) \Rightarrow \frac{8-\mu}{\sigma} = 0,5 \Rightarrow \mu + 0,5\sigma = 8$.

$$P(X \geq 12) = P(Z \geq 1,5) \Rightarrow \frac{12-\mu}{\sigma} = 1,5 \Rightarrow \mu + 1,5\sigma = 12$$

Se resuelve, entonces, el sistema de ecuaciones: $\begin{cases} \mu + 0,5\sigma = 8 \\ \mu + 1,5\sigma = 12 \end{cases} \Rightarrow \sigma = 4;\ \mu = 6$

5. Un tirador olímpico de tiro con arco hace diana con el 70 % de sus flechas. Si en un concurso realiza 30 disparos:

a) ¿Cuál es el número esperado de dianas?

b) ¿Cuál es la probabilidad de que haga más de 25 dianas?

c) ¿Cuál es la probabilidad de que haga exactamente 21 dianas?

a) Considera la variable aleatoria X: "número de dianas conseguidas en los 30 disparos", que tiene distribución binomial de parámetros $n = 30,\ p = 0,7$.

El número esperado de dianas es la esperanza de la variable X: $E[X] = np = 30 \cdot 0,7 = 21$.

b) Para calcular la probabilidad de que haga más de 25 dianas, se aproxima la distribución binomial por la distribución de una variable aleatoria normal Y de media la calculada en el apartado anterior, $\mu = 21$, y varianza $\text{Var}(X) = np(1-p) = 30 \cdot 0,7 \cdot 0,3 = 6,3$. Es decir: $Y \sim N\left(\mu = 21;\ \sigma = \sqrt{6,3}\right)$.

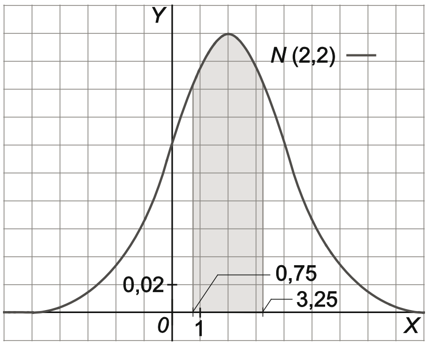
$$P(Y > 25) \approx P(X \ge 25,5) = P\left(Z \ge \frac{25,5-21}{\sqrt{6,3}}\right) = P(Z \ge 1,79) = 1 - \Phi(1,79) = 1 - 0,9633 = 0,0367$$

c) La probabilidad de que haga exactamente 21 dianas viene dada por $P(X = 21) = \binom{30}{21} \cdot 0,7^{21} \cdot 0,3^{9} = 0,1572$

Si se usa la aproximación se tiene

$$P(X = 21) \simeq P(20,5 < Y < 21,5) = P\left(\frac{20,5-21}{\sqrt{6,3}} < Z < \frac{21,5-21}{\sqrt{6,3}}\right) = P(-0,20 < Z < 0,2) = 2 \cdot \Phi(0,2) - 1 = 0,1586.$$

6. Expresa en términos de probabilidad el área sombreada de la figura, sabiendo que la distribución de probabilidad es normal y que su varianza es 4.



En este caso, la distribución de la variable aleatoria es $X \sim N(\mu = 2; \sigma = 2)$, y la probabilidad es $P(0,75 < X < 3,25)$

Se trata de un intervalo simétrico respecto de la media $\mu = 2$ que se puede escribir como:

$$(\mu - k \cdot \sigma, \mu + k \cdot \sigma) = (2 - 2k, 2 + 2k) = (0,75\ ,\ 3,25) \Rightarrow 2 + 2k = 3,25 \Rightarrow k = 0,625$$

Por tanto, ya que:

$$P(\mu - k \cdot \sigma < X < \mu + k \cdot \sigma) = 2\Phi(k) - 1 \Rightarrow P(0,75 < X < 3,25) = 2\Phi(0,625) - 1 = 2 \cdot 0,7341 - 1 = 0,4682$$

donde el valor de $\Phi(0,625)$ se ha calculado interpolando $\Phi(0,62)$ y $\Phi(0,63)$.

Relaciona y contesta

Elige la única respuesta correcta en cada caso

1. **Si $f(x) = \begin{cases} x + c & 0 \le x < 1 \\ 0 & \text{en el resto} \end{cases}$, el valor de c para que f sea una función de densidad de una variable aleatoria es:**

A. $c = 0$ **B.** $c = 1$ **C.** $c = -\frac{1}{2}$ **D.** $c = \frac{1}{2}$.

Solución: D.

2. **La distribución normal con media $\mu = 10$ y varianza $\sigma^2 = 8$ aproxima a la distribución binomial de parámetros:**

A. $n = 30,\ p = 0{,}2$ **B.** $n = 40,\ q = 0{,}8$ **C.** $n = 100,\ p = 0{,}4$ **D.** $n = 50,\ q = 0{,}8$.

Solución: D.

3. **Si X tiene una distribución $N(\mu, \sigma)$, en el intervalo $(-\sigma, \sigma)$ se encuentra el:**

A. 90% de los valores de X
B. 65% de los valores de X
C. 68% de los valores de X
D. Depende de los valores de σ.

Solución: D.

Señala, en cada caso, las respuestas correctas

4. **Los valores $x_1 = 3$ y $x_2 = 19$ de una variable aleatoria X con distribución normal se corresponden, respectivamente, con los valores tipificados $z_1 = -1$ y $z_2 = 3$. Entonces X tiene una distribución normal con:**

A. $\mu = 7$ y $\sigma^2 = 16$ **B.** $\mu = 6$ y $\sigma^2 = 4$ **C.** $\mu = 7$ y $\sigma = 4$ **D.** $\mu = 6$ y $\sigma = 16$.

Soluciones: A. y C.

5. **El porcentaje de observaciones que queda fuera del intervalo (0, 24) si una variable aleatoria tiene una distribución $X \sim N(\mu = 12, \sigma = 3)$ es:**

A. 25 % **B.** 75 % **C.** 0 % **D.** 50 %

Solución: C.

Elige la relación correcta entre las dos afirmaciones

6. **Sea $X \sim \text{Bin}(n, p)$ e $Y \sim N(\mu, \sigma)$ con $\mu = np$ y $\sigma^2 = np(1-p)$. Se consideran las siguientes afirmaciones:**

1. **La distribución de la variable Y aproxima a la de X.**
2. **En la distribución de X, n es grande y p no es muy pequeño.**

A. $1 \Leftrightarrow 2$
B $1 \Rightarrow 2$ pero $2 \not\Rightarrow 1$
C. $2 \Rightarrow 1$ pero $1 \not\Rightarrow 2$
D. 1 y 2 son independientes una de otra

Solución: A.

Señala el dato innecesario para contestar

7. **Para calcular la probabilidad $P(-\sigma < X < 2\sigma)$ con $X \sim N(\mu = 0, \sigma)$ es necesario conocer:**

A. Solo μ **B.** Solo σ **C.** Tanto μ como σ **D.** No hace falta conocer ningún dato

Solución: D.