
1.- Se ha tomado una muestra de los precios de un mismo producto alimenticio en 16 comercios, elegidos al azar en un barrio de una ciudad, y se han encontrado los siguientes precios:

95, 108, 97, 112, 99, 106, 105, 100, 99, 98, 104, 110, 107, 111, 103, 110

Suponiendo que los precios de este producto se distribuyen según una normal de varianza 25 y media desconocida:

a) ¿Cuál es la distribución de la media muestral?

Resolución

X = precios $\to N\left(\mu,\ \sqrt{25}\right) = N(\mu,\ 5)$. Sabemos, por el teorema central del límite que si $X \to N(\mu, \sigma)$ y $\overline{X}$ = media de las muestras de tamaño n, como la media de la muestra dada, de tamaño 16, es

$$\overline{x} = \frac{95+108+97+112+99+106+105+100+99+98+104+110+107+11+103+110}{16} = 104 \Rightarrow \overline{X} \to N\left(\overline{x}, \frac{\sigma}{\sqrt{n}}\right).$$

n = 16 y σ = 5. Sustituyendo, $\overline{X} \to N\left(104\ ;\ \frac{5}{\sqrt{16}}\right) = N(104\ ; 1,25)$.

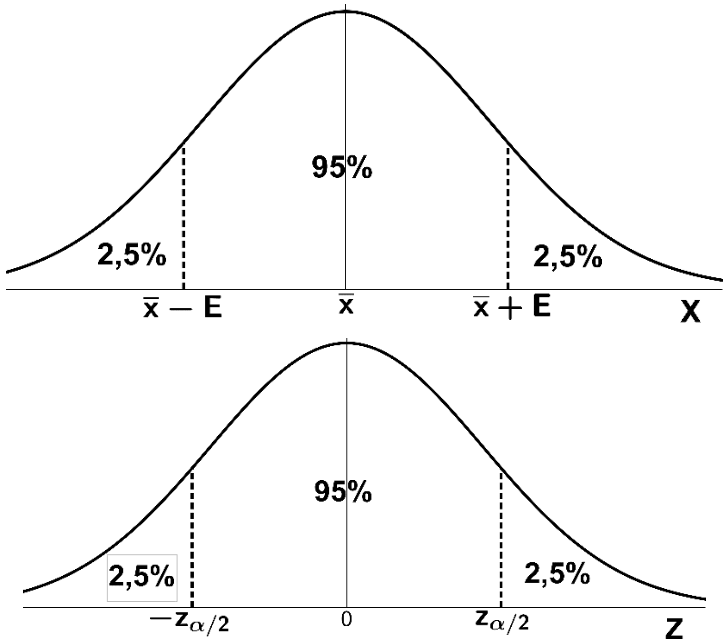
b) Determine el intervalo de confianza, al 95%, para la media poblacional.

Resolución

El intervalo de confianza a nivel de confianza del 95% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$,

siendo $\overline{x} = 104$ la media de la muestra de tamaño n = 16 y $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ es el máximo error de estimación.

Como el área bajo la campana es 100%, entonces 100% – 95% = 5% y 5% : 2 = 2,5%



$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 95\% + 2,5\% = 97,5\% = 0,975$

Esta probabilidad coincide con $\frac{1+n_c}{2} = \frac{1+0,95}{2} = 0,975$

Como $p(Z < z_{\alpha/2}) = 0,975$ *usando la tabla de la* $N(0, 1)$ $\rightarrow z_{\alpha/2} = 1,96.$

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1,96.\frac{5}{\sqrt{16}} = 2,45$; $I_c = (104 - 2,45 ;\ 104 + 2,45) \cong (101,55 ;\ 106,45)$

2.- La variable altura de las alumnas que estudian en una escuela de idiomas sigue una distribución normal de media 1,62 m y desviación típica 0,12 m. ¿Cuál es la probabilidad de que la media de una muestra aleatoria de 100 alumnas sea mayor que 1,60.

Resolución

X = altura $\rightarrow N(1,62 ;\ 0,12)$. Sabemos, por el teorema central del límite que si $X \rightarrow N(\mu, \sigma)$ y $\overline{X}$ = media de las muestras de tamaño n, entonces $\overline{X} \rightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

n = 100. Sustituyendo, $\overline{X} \rightarrow N\left(1,62 ;\ \frac{0,12}{\sqrt{100}}\right) \cong N(1,62 ; 0,012)$. Tipificando, $Z = \frac{\overline{X} - 1,62}{0,012} \rightarrow N(0,\ 1)$

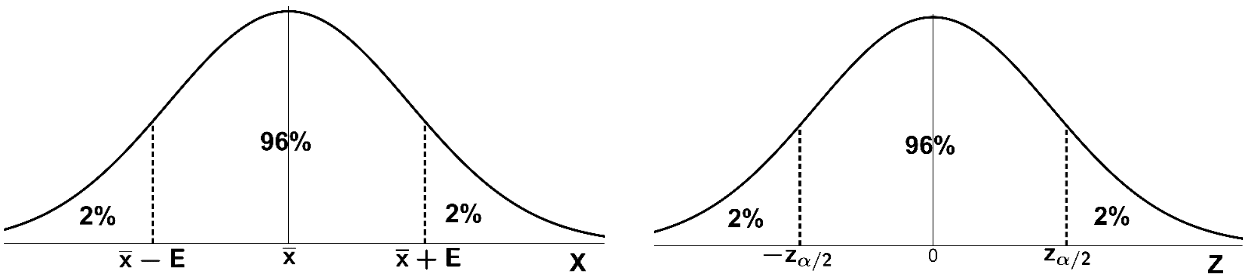
Piden $p(\overline{X} > 1,60) = p\left(\frac{\overline{X} - 1,62}{0,012} > \frac{1,60 - 1,62}{0,012}\right) \cong p(Z > -1,67) = p(Z \leq 1,67) = 0,9525 = 95,25\%$

3.- El peso de los individuos de una ciudad se distribuye según una ley normal de media desconocida y varianza 9 kg^2. Se ha seleccionado, en esa ciudad una muestra aleatoria que ha dado un peso medio de 65 kg. Con una confianza del 96% se ha construido un intervalo para la media poblacional cuyo límite inferior ha resultado ser 62,95 kg.
a) ¿Cuál ha sido el tamaño de la muestra seleccionada?
b) Determine el límite superior del intervalo.

Resolución

X = peso $\rightarrow N(\mu, \sqrt{9}) = N(\mu, 3)$. El intervalo de confianza a nivel de confianza del 96% para estimar μ, es $I_c = (\overline{x} - E,\ \overline{x} + E)$, siendo $\overline{x} = 65$ la media de la muestra de tamaño n y $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$, el máximo error de estimación.

Como el área bajo la campana es 100%, entonces 100% – 96% = 4% y 4% : 2 = 2%



$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 96\% + 2\% = 98\% = 0,98$

Esta probabilidad coincide con $\frac{1 + n_c}{2} = \frac{1 + 0,96}{2} = 0,98$

Como $p(Z < z_{\alpha/2}) = 0,98$ *usando la tabla de la* $N(0, 1)$, *por interpolación* $\rightarrow z_{\alpha/2} = 2,055.$

Como el límite inferior del intervalo es 62,95, entonces $65 - E = 62,95 \Rightarrow E = 2,05$

b) El límite superior del intervalo es entonces $65 + 2,05 = 67,05$

a) Piden hallar n si $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 2,5 \Leftrightarrow z_{\alpha/2}.\frac{\sigma}{2,05} = \sqrt{n}$ $elevando\ al\ cuadrado$ $\rightarrow (z_{\alpha/2})^2.\frac{\sigma^2}{2,05^2} = n$

Sustituyendo: $n = 2,055^2.\frac{3^2}{2,05^2} \cong 9,04$. Luego, el tamaño de la muestra ha sido de 10 individuos

4.- La variable X se distribuye según una ley normal de media 10 y desviación típica 3. Determine el tamaño de una muestra extraída de la población, de modo que la probabilidad de la media muestral esté por encima de 12 sea de 0,025.

Resolución

X $\rightarrow N(10,\ 3)$. Sabemos, por el teorema central del límite que si $X \rightarrow N(\mu, \sigma)$
y $\overline{X}$ = media de las muestras de tamaño n, entonces $\overline{X} \rightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

Sustituyendo, $\overline{X} \rightarrow N\left(10,\ \frac{3}{\sqrt{n}}\right)$. Tipificando, $Z = \frac{\overline{X} - 10}{\frac{3}{\sqrt{n}}} \rightarrow N(0,\ 1)$

Nos piden hallar n para que $0,025 = p(\overline{X} > 12) = p\left(\frac{\overline{X} - 10}{\frac{3}{\sqrt{n}}} > \frac{12 - 10}{\frac{3}{\sqrt{n}}}\right) \cong p\left(Z > \frac{2\sqrt{n}}{3}\right) =$

$= 1 - p\left(Z \leq \frac{2\sqrt{n}}{3}\right) \Rightarrow p\left(Z \leq \frac{2\sqrt{n}}{3}\right) = 0,975$. Usando la tabla de la N(0, 1) en sentido inverso $\frac{2\sqrt{n}}{3} = 1,96$.

Despejando, $\sqrt{n} = \frac{3.1,96}{2} = 2,94$. Elevando al cuadrado, n = 8,6436. El tamaño muestral debe ser 9

5.- Sea una población formada sólo por 3 elementos con valores 2, 4 y 6. Consideremos todas las muestras, con reemplazamiento, de tamaño 2. Calcule la media y desviación típica de la población, así como de las medias muestrales. ¿Qué relación hay entre ambas medias?

Resolución

las muestras son {2–2, 2–4, 2–6, 4–2, 4–4, 4–6, 6–2, 6–4, 6–6}.

La desviación típica, σ, de la población es $\sigma = \sqrt{\frac{\Sigma x_i^2}{N} - \mu^2}$ siendo N el tamaño de la población, N = 3

y μ la media de la población, $\mu = \frac{\Sigma x_i}{N}$

x_i 2 4 6 | 12 x_i^2 4 16 36 | 56 ; $\mu = \frac{\Sigma x_i}{N} = \frac{12}{3} = 4$ $\sigma = \sqrt{\frac{\Sigma x_i^2}{N} - \mu^2} = \sqrt{\frac{56}{3} - 4^2} = \sqrt{\frac{8}{3}}$

Por el teorema central del límite la media de las medias muéstrales también es μ = 4. O sea, la relación entre ambas medias es que son iguales.

La desviación típica de las medias muestrales es $\frac{\sigma}{\sqrt{n}}$, en este caso n = 2.

Así, la desviación típica de las medias muestrales es $\frac{\sqrt{\frac{8}{3}}}{\sqrt{2}} = \sqrt{\frac{4}{3}} \cong 1,1547$

6.- Un ascensor admite como peso máximo 300 kg. La población de usuarios tiene un peso que se distribuye según una ley normal de media 70 kg y desviación típica 10 kg. Calcule la probabilidad de que 4 personas cualesquiera de dicha población, que suban al ascensor, superen el peso máximo.

Resolución

X = peso $\rightarrow N(70,\ 10)$. Sabemos, por el teorema central del límite que si $X \rightarrow N(\mu, \sigma)$

y $\overline{X}$ = media de las muestras de tamaño n, entonces $\overline{X} \rightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

n = 4. Sustituyendo, $\overline{X} \rightarrow N\left(70\ ;\ \frac{10}{\sqrt{4}}\right) = N(70,\ 5)$. Tipificando, $Z = \frac{\overline{X} - 70}{5} \rightarrow N(0,\ 1)$

Como 300:4 = 75, la probabilidad que se pide es

$$p(\overline{X} > 75) = p\left(\frac{\overline{X} - 70}{5} > \frac{75 - 70}{5}\right) \cong p(Z > 1) = 1 - p(Z \leq 1) = 1 - 0,8413 = 15,87\%$$

7.- Una variable aleatoria X sobre una población tiene de media 50 y de desviación típica 5. Extraemos, aleatoriamente, de dicha población 1000 muestras, todas ellas de tamaño 64. De cada muestra calculamos su media, y llamamos A al conjunto de números formados con esas medias.
a) Diga, de forma razonada, qué valores se pueden esperar para la media y la desviación típica de A.

Resolución

X $\rightarrow N(50,\ 5)$. Sabemos, por el teorema central del límite, que si $X \rightarrow N(\mu, \sigma)$

y $\overline{X}$ = media de las muestras de tamaño n, entonces $\overline{X} \rightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

n = 64. Sustituyendo, $\overline{X} \rightarrow N\left(50\ ;\ \frac{5}{\sqrt{64}}\right) = N(50\ ;\ 0,625)$.

El teorema central del límite nos dice que se puede esperar para estas 1000 muestras de tamaño 64, una media de 50 y una desviación típica de 0,625.

b) ¿Cuál es la probabilidad de que una de esas muestras tenga una media comprendida entre 48,5 y 50,5?

Resolución

Tipificando, $Z = \frac{\overline{X} - 50}{0,625} \rightarrow N(0,\ 1)$.

Piden $p(48,5 < \overline{X} < 50,5) = p\left(\frac{48,5 - 50}{0,625} < \frac{\overline{X} - 50}{0,625} < \frac{50,5 - 50}{0,625}\right) = p(-2,4 < Z < 0,8) =$

$= p(Z < 0,8) - p(Z > 2,4) = p(Z < 0,8) - 1 + p(Z < 2,4) = 0,7881 - 1 + 0,9918 = 0,7799 = 77,9$

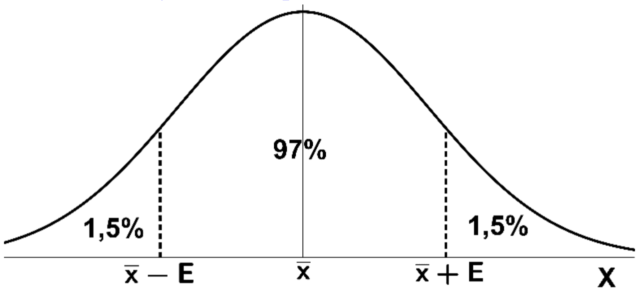
8.- Se sabe que la desviación típica de las tallas de los alumnos de una Universidad es igual a 6 cm. Para estimar la talla media de dichos alumnos se toma una muestra de 64 estudiantes, resultando una media muestral de 173 cm.
a) Determine el intervalo de confianza de la talla media de los alumnos de la Universidad, con un nivel de confianza de 0,97.

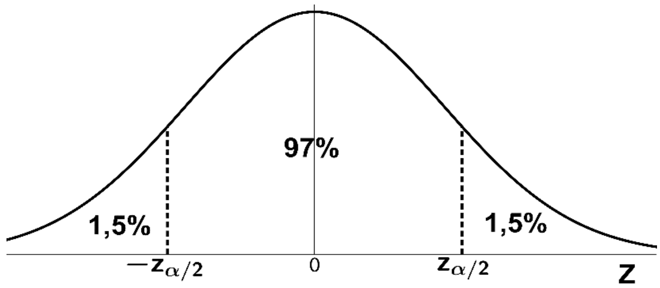
Resolución

X = talla$\rightarrow N(\mu, 6)$. El intervalo de confianza a nivel de confianza del 96% para estimar μ,

es $I_c = (\overline{x} - E,\ \overline{x} + E)$, siendo $\overline{x} = 173$ la media de la muestra de tamaño n = 64 y $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$, el máximo error de estimación.

Como el área bajo la campana es 100%, entonces 100% – 97% = 3% y 3% : 2 = 1,5%





$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 97\% + 1,5\% = 98,5\% = 0,985$

Esta probabilidad coincide con $\frac{1+n_c}{2} = \frac{1+0,97}{2} = 0,985$

Como $p(Z < z_{\alpha/2}) = 0,985$ *usando la tabla de la* $N(0,1)$ $\rightarrow z_{\alpha/2} = 2,17.$

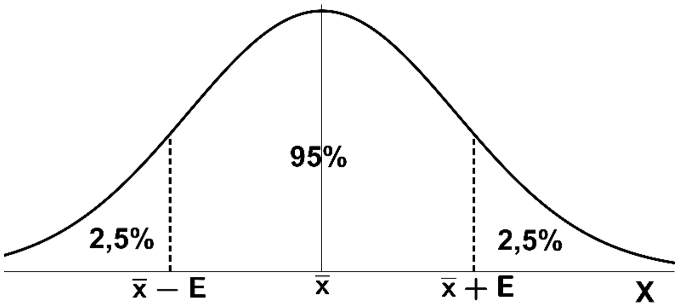
$E = z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} = 2,17\cdot\frac{6}{\sqrt{64}} = 1,6275$; $I_c = (173 - 1,6275 \ ; \ 173 + 1,6275) \cong (171,37 \ ; \ 174,63)$

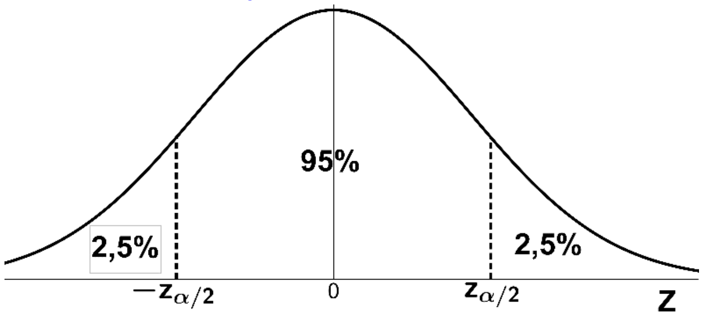
b) Calcule el tamaño muestral necesario para estimar la talla media de los alumnos de la Universidad, con un nivel de confianza del 95% y un error máximo de estimación no superior a 1,2 cm.

Resolución

$E = z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}$ es el máximo error de estimación y σ = 6.

Como el área bajo la campana es 100%, entonces 100% – 95% = 5% y 5% : 2 = 2,5%





$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 95\% + 2,5\% = 97,5\% = 0,975$

Esta probabilidad coincide con $\frac{1+n_c}{2} = \frac{1+0,95}{2} = 0,975$

Como $p(Z < z_{\alpha/2}) = 0,975$ *usando la tabla de la* $N(0,1)$ $\rightarrow z_{\alpha/2} = 1,96.$

Piden hallar n para que $E = z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} \leq 1,2 \Leftrightarrow z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{1,2} \leq \sqrt{n}$ *elevando al cuadrado* $\rightarrow (z_{\alpha/2})^2\cdot\frac{\sigma^2}{1,2^2} = n$

Sustituyendo: $n = 1,96^2\cdot\frac{6^2}{1,2^2} = 96,04$. Luego, el tamaño de la muestra ha sido 97 alumnos

9.- Un contable toma una muestra aleatoria de tamaño n = 36 de una población de 1000 cuentas por cobrar. El valor medio de las cuentas por cobrar es de 2600 ptas, con una desviación típica poblacional de 450 ptas.

a) Calcule la probabilidad de que la media muestral sea inferior a 2500 ptas.

Resolución

X = valor de las cuentas $\rightarrow N(2600,\ 450)$. Sabemos, por el teorema central del límite que si $X \rightarrow N(\mu, \sigma)$ y $\overline{X}$ = media de las muestras de tamaño n, entonces $\overline{X} \rightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

n = 36. Sustituyendo, $\overline{X} \rightarrow N\left(2600\ ,\ \frac{450}{\sqrt{36}}\right) \cong N(2600,\ 75)$. Tipificando, $Z = \frac{\overline{X} - 2600}{75} \rightarrow N(0,\ 1)$

Piden

$$p(\overline{X} < 2500) = p\left(\frac{\overline{X} - 2600}{75} < \frac{2500 - 2600}{75}\right) \cong p(Z < -1,33) = p(Z > 1,33) = 1 - 0,9082 = 9,18\%$$

b) Calcule la probabilidad de que la media muestral se encuentre a no más de 225 ptas de la media de la población

Resolución

Como la media es 2600, la probabilidad que se pide es

$$p(2600 - 225 \le \overline{X} \le 2600 + 225) = p(2375 \le \overline{X} \le 2825) = p\left(\frac{2375 - 2600}{75} \le \frac{\overline{X} - 2600}{75} \le \frac{2825 - 2600}{75}\right) =$$

$$= p(-3 \le Z \le 3) = p(Z \le 3) - 1 + p(Z \le 3) = 2p(Z \le 3) - 1 = 2\ .\ 0,9987 - 1 = 99,74\%$$

10.- Se dispone de una muestra aleatoria de 10 alumnos de una población de alumnos de 3º de E.S.O. Se sabe, por experiencias anteriores, que la altura de los alumnos de ese curso se distribuye según una variable aleatoria normal de media 167 cm. y desviación típica 3,2 cm.

a) Calcule la probabilidad de que la media muestral esté comprendida entre 166 cm y la media poblacional.

Resolución

X = altura $\rightarrow N(167\ ;\ 3,2)$. Sabemos, por el teorema central del límite que si $X \rightarrow N(\mu, \sigma)$ y $\overline{X}$ = media de las muestras de tamaño n, entonces $\overline{X} \rightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

n = 10. Sustituyendo, $\overline{X} \rightarrow N\left(167\ ;\ \frac{3,2}{\sqrt{10}}\right) \cong N(167\ ;\ 1,012)$. Tipificando, $Z = \frac{\overline{X} - 167}{1,012} \rightarrow N(0,\ 1)$

La probabilidad que se pide es

$$p(166 < \overline{X} < 167) = p\left(\frac{166 - 167}{1,012} < \frac{\overline{X} - 167}{1,012} < \frac{167 - 167}{1,012}\right) = p(-0,99 < Z < 0) =$$

$$= p(Z < 0) - 1 + p(Z < 0,99) = 0,5 - 1 + 0,8389 = 0,3389 = 33,89\%$$

b) ¿Cuál es la probabilidad de que la media muestral tenga un valor superior a 169 cm?

Resolución

La probabilidad que se pide es

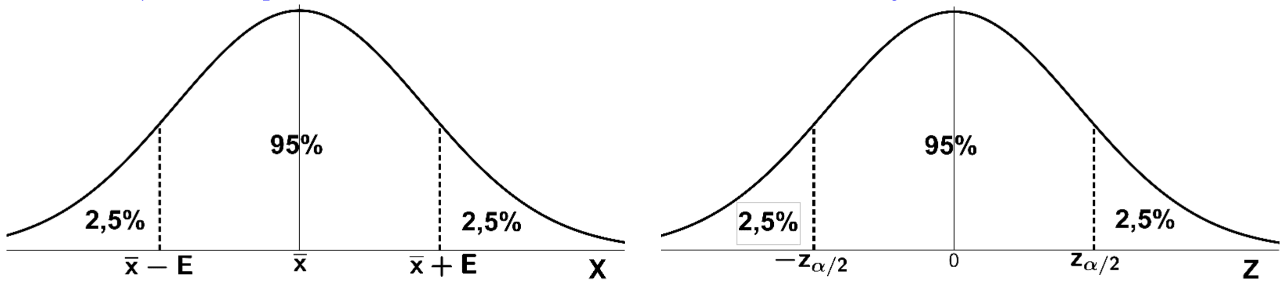
$$p(\overline{X} > 169) = p\left(\frac{\overline{X} - 167}{1,012} > \frac{169 - 167}{1,012}\right) \cong p(Z > 1,98) = 1 - p(Z \le 1,98) = 1 - 0,9761 = 2,39\%$$

11.- Una máquina fabrica clavos cuya longitud sigue una distribución normal con desviación típica 0,5 mm. Se toma una muestra de 25 clavos y se obtiene una longitud media, para los mismos, de 50 mm. Calcule un intervalo de confianza del 95% para la longitud media de la población.

Resolución

X = longitud $\to N(\mu, 0,5)$. El intervalo de confianza a nivel de confianza del 95% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 50$ la media de la muestra de tamaño n = 25 y $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$ es el máximo error de estimación.

Como el área bajo la campana es 100%, entonces 100% – 95% = 5% y 5% : 2 = 2,5%



$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 95\% + 2,5\% = 97,5\% = 0,975$

Esta probabilidad coincide con $\frac{1+n_c}{2} = \frac{1+0,95}{2} = 0,975$

Como $p(Z < z_{\alpha/2}) = 0,975$ *usando la tabla de la* $N(0, 1)$ $\to z_{\alpha/2} = 1,96$.

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1,96.\frac{0,5}{\sqrt{25}} = 0,196$; $I_c = (50 - 0,196\,;\, 50 + 0,196) \cong (49,804\,;\, 50,196)$

12.- Si los alumnos de Preescolar de Andalucía tienen una estatura que es una variable aleatoria de media 95 cm y desviación típica 16 cm y consideramos una muestra aleatoria de 36 de tales alumnos, ¿cuál es la probabilidad de que la media de esa muestra tome valores comprendidos entre 90 cm y 100 cm?

Resolución

X = estatura $\to N(95, 16)$. Sabemos, por el teorema central del límite que si $X \to N(\mu, \sigma)$

y $\overline{X}$ = media de las muestras de tamaño n, entonces $\overline{X} \to N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

n = 36. Sustituyendo, $\overline{X} \to N\left(95, \frac{16}{\sqrt{36}}\right) \cong N(95\,;\, 2,67)$. Tipificando, $Z = \frac{\overline{X} - 95}{2,67} \to N(0, 1)$

La probabilidad que se pide es

$p(90 < \overline{X} < 100) = p\left(\frac{90-95}{2,67} < \frac{\overline{X}-95}{2,67} < \frac{100-95}{2,67}\right) \cong p(-1,87 < Z < 1,87) =$

$= p(Z < 1,87) - 1 + p(Z < 1,87) = 2p(Z < 1,87) - 1 = 2 . 0,9693 - 1 = 0,9386 = 93,86\%$

OTROS DEL 1998 (COU II)

1.-

a) Complete los datos que faltan en la siguiente tabla de distribución de frecuencias, donde n_i, N_i

y f_i representan, respectivamente, la frecuencia absoluta, la frecuencia absoluta acumulada y la frecuencia relativa de la variable X.

X	1	2	3	4	5	6	7	8
n_i	4	4		7	5		7	
N_i				23		38	45	
f_i	0,08		0,16					

b) Calcule la media y la moda de la distribución anterior.

Resolución

Completando las filas n_i y N_i podemos completar el resto de datos

$0,08 = \frac{4}{N} \Rightarrow N = \frac{4}{0,08} = 50$; $0,16 = \frac{n_3}{50} \Rightarrow n_3 = 50 \,.\, 0,16 = 8$

$28 + n_6 = 38 \Rightarrow n_6 = 10$; $45 + n_8 = 50 \Rightarrow n_8 = 5$

x_i	1	2	3	4	5	6	7	8	total
n_i	4	4	8	7	5	10	7	5	50 = N
N_i	4	8	16	23	28	38	45	50	-
f_i	0,08	0,08	0,16	0,14	0,1	0,2	0,14	0,1	1
$x_i n_i$	4	8	24	28	25	60	49	40	238

moda: Mo = 6 (valor de mayor frecuencia absoluta)

media: $\bar{x} = \frac{\Sigma x_i n_i}{N} = \frac{238}{50} = 4,76$

2.- Un examen tipo “test” consta de 8 preguntas con cuatro respuestas posibles cada una, de las cuales sólo una es correcta. Suponiendo que un estudiante responde al azar:
a) ¿Cuál es la probabilidad de que conteste correctamente a más de 6 preguntas?
b) ¿Cuál es la probabilidad de que no acierte ninguna?

Resolución

Como realizamos 8 veces el experimento aleatorio de elegir al azar una respuesta de las cuatro posibles y la probabilidad de que la respuesta sea la correcta es $\frac{1}{4} = 0,25$, entonces la variable aleatoria

X = “nº de preguntas contestadas correctamente” → B(8 ; 0,25)

La ley de probabilidad es $p_k = p(X = k) = (8\ k)0,25^k 0,75^{8-k}$, con k = 0, 1, 2, 3, 4, 5, 6, 7, 8.

a) se pide
$p(X = 7) + p(X = 8) = (8\ 7)0,25^7 0,75^1 + (8\ 8)0,25^8 0,75^0 = 8.\,0,25^7 0,75 + 1.\,0,25^8 \cong 0,0004$

b) se pide $p(X = 0) = (8\ 0)0,25^0 0,75^8 = 1.\,0,75^8 \cong 0,1001$

3.- En cada una de las siguientes tablas se presenta la función de probabilidad de una variable aleatoria X.

x	P(X=x)
1	0'2
2	0'3
3	1'2
4	0'1
Tabla 1	

x	P(X=x)
1	0'4
2	0'3
3	0'2
4	0'1
Tabla 2	

x	P(X=x)
1	0'2
2	0'3
3	0'4
4	0'5
Tabla 3	

a) Elija razonadamente la tabla que representa la función de probabilidad de X.
b) Escriba la función de probabilidad de X.
c) Calcule la media de X.

Resolución

Sea $p_i = p(X = x_i)$. Por ser función de probabilidad debe ser (1) $0 \le p_i \le 1$ y (2) $\sum p_i = 1$

Luego, descartamos la tabla 1 por no cumplirse la propiedad (1) y la tabla 3 por no cumplirse la propiedad (2).

Por tanto, la tabla que representa la función de probabilidad de X es:

x_i	$p_i = p(X = x_i)$	$x_i p_i$
1	0, 4	0,4
2	0, 3	0,6
3	0, 2	0,6
4	0, 1	0,4
total	1	2

La media de X es $\mu = \sum x_i p_i = 2$

4.- Razone si cada una de las siguientes proposiciones es Verdadera o Falsa:
a) Dadas dos variables aleatorias X e Y, el coeficiente de correlación lineal entre ellas tiene el mismo signo que la pendiente de la recta de regresión de Y sobre X.

Resolución

Verdadera: el coeficiente de correlación de ambas variables es $r = \frac{s_{xy}}{s_x s_y}$ y la pendiente de la recta de regresión de Y sobre X es $\frac{s_{xy}}{s_x^2}$. En ambos casos los denominadores son positivos y sólo dependen de la covarianza s_{xy}.

b) El coeficiente de correlación entre una variable aleatoria X y otra Y es −1,5.

Resolución Falsa: el coeficiente de correlación lineal no puede ser menor de −1

c) Si el coeficiente de correlación entre dos variables aleatorias cualesquiera X e Y vale 0,5, entonces el coeficiente de correlación lineal entre Y y X es de $\frac{1}{0,5}$

Resolución

Falsa: el coeficiente de correlación lineal de dos variables X e Y siempre es el mismo $r = \frac{s_{xy}}{s_x s_y}$

5.- Se pasa un cuestionario a un grupo de 40 alumnos para medir su dominio del vocabulario (X) y su capacidad de razonamiento (Y). Con los resultados se han obtenido las siguientes cantidades:

$$\sum_i x_i = 1800 \ ; \ \sum_i x_i^2 = 86000 \ ; \ \sum_i y_i = 1390 \ ; \ \sum_i y_i^2 = 53100 \ ; \ \sum_i x_i y_i = 66500$$

a) Halle la recta de regresión de Y sobre X.
b) ¿Qué capacidad de razonamiento cabe esperar para un alumno que tiene una puntuación de 60 en dominio de vocabulario?

Resolución

Vemos que el nº de datos es N = 40. La recta de regresión de Y sobre X es $r_{YX}: y - \overline{y} = \frac{s_{xy}}{s_x^2}(x - \overline{x})$

Usando el enunciado: $\overline{x} = \frac{\Sigma x_i}{N} = \frac{1800}{40} = 45 \ ; \ s_x^2 = \frac{\Sigma x_i^2}{N} - \overline{x}^2 = \frac{86000}{40} - 45^2 = 125$

$\overline{y} = \frac{\Sigma y_i}{N} = \frac{1390}{40} = 34,75 \ ; \ s_{xy} = \frac{\Sigma x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{66500}{40} - 45.\ 34,75 = 98,75$

La recta de regresión de Y sobre X es $r_{YX}: y - 34,75 = \frac{98,75}{125}(x - 45) \Rightarrow r_{YX}: y = 0,79x - 0,8$.

Para x = 60, y = 0,79.60 – 0,8 = 46,67.

Es decir, cabe esperar una capacidad de razonamiento de 46,67 para una puntuación de 60 en vocabulario.

$s_y^2 = \frac{\Sigma y_i^2}{N} - \overline{y}^2 = \frac{1390}{40} - 34,75^2 = 119,9375$

El coeficiente de correlación lineal es $r = \frac{s_{xy}}{s_x s_y} = \frac{98,75}{\sqrt{125}\,.\,\sqrt{119,9375}} = 0,8065$, positivo y próximo a 1.

Como el coeficiente de correlación es positivo la correlación es directa, cuando aumenta una variable aumenta la otra y es fuerte pues está más próximo a 1 que a 0.

6.- Se ha preguntado a un grupo de deportistas las horas que dedican a entrenamiento durante el fin de semana. Los resultados aparecen en la siguiente distribución de frecuencias:

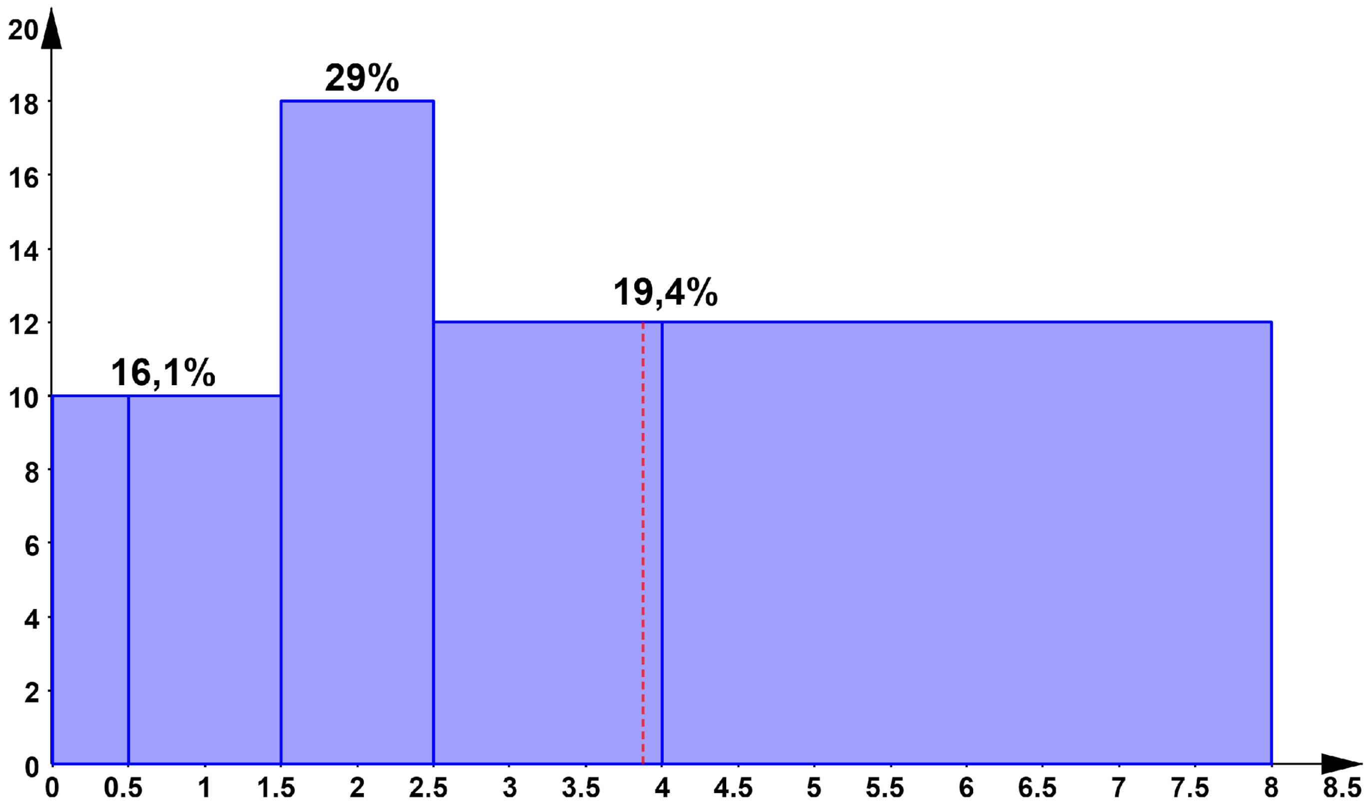
Horas	[0,0'5)	[0'5,1'5)	[1'5,2'5)	[2'5,4)	[4,8)
Personas	10	10	18	12	12

a) Dibuje el histograma de la distribución de frecuencias (absolutas o relativas).
b) Calcule el tercer cuartil.

Resolución

clases (horas)	[0 ; 0,5)	[0,5 ; 1,5)	[1,5 ; 2,5)	[2,5 ; 4)	[4, 8)	total
n_i (frecuencia absoluta)	10	10	18	12	12	62 = N

N_i (frecuencia absoluta acumulada)	10	20	38	50	62	-
f_i (frecuencia relativa)	16,1%	16,1%	29%	19,4%	19,4%	100%



El tercer cuartil, Q3, deja el 75% de los datos a la izquierda. Como $75\%\ de\ N = 75\%\ de\ 62 = 46,5$, la clase que contiene a Q3 es [1,5 ; 2,5).

Aplicando la expresión del tercer cuartil para datos agrupados en intervalos, se tiene:

El tercer cuartil es $Q3 = L_i + \frac{c\left(\frac{3N}{4} - N_{i-1}\right)}{n_i} = 1,5 + \frac{1(46,5 - 38)}{18} \cong 1,972$

Siendo L_i = límite inferior de la clase que contiene a Q3, c = amplitud del intervalo, N = nº total de datos

N_{i-1} = frecuencia absoluta acumulada de la clase anterior a dicha clase.

n_i = frecuencia absoluta de la clase que contiene a Q3

7.- La distribución de una variable aleatoria X es normal con media 6 y desviación típica 4.
a) Calcule $p(X \geq 12)$.
b) Calcule $p(5 \leq X \leq 8)$.

Resolución

$X \rightarrow N(6,\ 4) \Rightarrow Z = \frac{X-6}{4} \rightarrow N(0,\ 1)$.

a) $p(X\geq 12) = p\left(\frac{X-6}{4} \geq \frac{12-6}{4}\right) = p(Z\geq 1,5) = 1 - p(Z < 1,5) = 1 - 0,9332 = 0,0668 = 6,68\%$

b) $p(5\leq X\leq 8) = p\left(\frac{5-6}{4} \leq \frac{X-6}{4} \leq \frac{8-6}{4}\right) = p(-0,25\leq Z\leq 0,5) = p(Z\leq 0,5) - p(Z\geq 0,25) =$

$= p(Z\leq 0,5) - 1 + p(Z < 0,25) = 0,6913 - 1 + 0,5987 = 0,2902 = 29,02\%$.

8.- De una variable aleatoria bidimensional (X, Y) se conoce que la recta regresión de Y sobre X es $y = \frac{x}{2} + 2$ y la recta regresión de X sobre Y es $x = \frac{3y}{2} - \frac{5}{2}$.

a) Calcule, si es posible, el centro de gravedad $(\overline{x}, \overline{y})$ de la distribución.

Resolución

Sabemos que las rectas de regresión se cortan en el punto $(\overline{x}, \overline{y})$. Resolvemos el sistema $\left\{y = \frac{x}{2} + 2 \quad x = \frac{3y}{2} - \frac{5}{2}\right.$

Sustituyendo $x = \frac{3y}{2} - \frac{5}{2}$ en la 1ª ecuación, $y = \frac{1}{2}\left(\frac{3y}{2} - \frac{5}{2}\right) + 2 = \frac{3y}{4} - \frac{5}{4} + 2 = \frac{3y+3}{4}$.

Multiplicando por 4 obtenemos: $4y = 3y + 3 \Rightarrow y = 3$; $x = \frac{3.3}{2} - \frac{5}{2} = 2 \Rightarrow$ el centro de gravedad es (2, 3)

b) Calcule e interprete el coeficiente de correlación.

Resolución

La recta de regresión (de Y sobre X) es $r_{YX}: y - \overline{y} = \frac{s_{xy}}{s_x^2}(x - \overline{x})$

Como $r_{YX}: y = \frac{x}{2} + 2$, entonces $\frac{s_{xy}}{s_x^2} = \frac{1}{2}$

La recta de regresión (de X sobre Y) es $r_{XY}: x - \overline{x} = \frac{s_{xy}}{s_y^2}(y - \overline{y})$

Como $r_{XY}: x = \frac{3y}{2} - \frac{5}{2}$, entonces $\frac{s_{xy}}{s_y^2} = \frac{3}{2}$

Multiplicando, $\frac{s_{xy}^2}{s_x^2 s_y^2} = \frac{1}{2}\frac{3}{2} = \frac{3}{4}$. Luego, el coeficiente de correlación es $r = \frac{s_{xy}}{s_x s_y} = \sqrt{\frac{3}{4}} \cong 0,866$

Como el coeficiente de correlación es positivo la correlación es directa, cuando aumenta una variable aumenta la otra y es fuerte pues está más próximo a 1 que a 0.

9.- En un proceso de fabricación de lavadoras, la probabilidad de que una lavadora tenga algún defecto es de 0,05. Un operario revisa 4 lavadoras que se acaban de montar.

a) Calcule la probabilidad de que no haya ninguna defectuosa.

b) Calcule la probabilidad de que haya más de una defectuosa.

Resolución

Como realizamos 4 veces el experimento aleatorio de examinar si la lavadora tiene o no algún defecto y la probabilidad de que tenga defecto es 0,05, entonces la variable aleatoria

X = "nº de lavadoras defectuosas" → B(4 ; 0,05)

La ley de probabilidad es $p_k = p(X = k) = \binom{4}{k} 0,05^k 0,95^{4-k}$, con k = 0, 1, 2, 3, 4.

a) se pide $p(X = 0) = \binom{4}{0} 0,05^0 0,95^4 = 1.\, 0,95^4 \cong 0,8145$

b) como

$p(X = 0) + p(X = 1) = \binom{4}{0} 0,05^0 0,95^4 + \binom{4}{1} 0,05^1 0,95^3 = 1.\, 0,95^4 + 4.\, 0,05.\, 0,95^3 \cong 0,986$,

la probabilidad que se pide es $p(X > 1) = 1 - 0,986 \cong 0,014$

10.- Se considera la tabla estadística siguiente

X	2	4	a	3	5
Y	1	2	1	1	3

a) Calcule el valor de "a" sabiendo que la media de X es 3.

Resolución $3 = \bar{x} = \frac{\Sigma x_i}{N} = \frac{a+14}{5}$. Despejando, a = 1

b) Mediante la correspondiente recta de regresión lineal, haga una predicción del valor que se obtiene para Y cuando X = 4,5. Halle el cuadrado del coeficiente de correlación entre X e Y y explique la fiabilidad de la predicción anterior.

Resolución

Sabemos que a = 1 y el nº de datos es N = 5. Elaboramos la tabla de frecuencias adecuada:

x_i	2	4	1	3	5	→ 15
y_i	1	2	1	1	3	→ 8
x_i^2	4	16	1	9	25	→ 55
y_i^2	1	4	1	1	9	→ 16
$x_i y_i$	2	8	1	3	15	→ 29

$\bar{x} = \frac{\Sigma x_i}{N} = \frac{15}{5} = 3$; $\bar{y} = \frac{\Sigma y_i}{N} = \frac{8}{5} = 1,6$; $s_x^2 = \frac{\Sigma x_i^2}{N} - \bar{x}^2 = \frac{55}{5} - 3^2 = 2$

$s_y^2 = \frac{\Sigma y_i^2}{N} - \bar{y}^2 = \frac{16}{5} - 1,6^2 = 0,64$; $s_{xy} = \frac{\Sigma x_i y_i}{N} - \bar{x}\,\bar{y} = \frac{29}{5} - 3.\, 1,6 = 1$

La recta de regresión (de Y sobre X) es $r_{YX}: y - \bar{y} = \frac{s_{xy}}{s_x^2}(x - \bar{x}) \Rightarrow r_{YX}: y - 1,6 = \frac{1}{2}(x - 3)$

Operando y simplificando, la recta de regresión que se pide es $r_{YX}: y = 0,5x + 0,1$

Para x = 4,5, $y = 0,5 .\, 4,5 + 0,1 = 2,35$

El coeficiente de correlación lineal es $r = \frac{s_{xy}}{s_x s_y} = \frac{1}{\sqrt{2}.\, \sqrt{0,64}} \cong 0,8839$, positivo y próximo a 1.

Como el coeficiente de correlación es positivo la correlación es directa, cuando aumenta una variable aumenta la otra y es fuerte pues está más próximo a 1 que a 0.

El cuadrado del coeficiente de correlación es $r^2 = \frac{s_{xy}^2}{s_x^2 s_y^2} = \frac{1}{2 \,.\, 0{,}64} = 0,78125$

11.- Se sabe que la altura de los alumnos varones de una Facultad sigue una ley Normal de media 1,75 m. Sabiendo que el 33% de los alumnos mide más de 1,80 m.
a) Calcule la varianza de la distribución de tales alturas.

Resolución

X = altura $\rightarrow N(1,75\,;\ \sigma) \ \Rightarrow\ Z = \frac{X-1{,}75}{\sigma} \rightarrow N(0,\ 1)$. Según el enunciado, $p(X > 1,8) = 33\% = 0,33$

Luego, $0,33 = p(X > 1,8) = p\left(\frac{X-1{,}75}{\sigma} > \frac{1{,}8-1{,}75}{\sigma}\right) = p\left(Z > \frac{0{,}05}{\sigma}\right) = 1 - p\left(Z \leq \frac{0{,}05}{\sigma}\right)$

Luego, $p\left(Z \leq \frac{0{,}05}{\sigma}\right) = 1 - 0,33 = 0,67$. Usando la N(0, 1) en sentido inverso, $\frac{0{,}05}{\sigma} = 0,44 \Rightarrow \sigma \cong 0,114$.

Por tanto, la varianza es $\sigma^2 \cong 0,013$

b) Calcule el porcentaje de alumnos que miden menos de 1,60 m.

Resolución

$p(X < 1,6) = p\left(\frac{X-1{,}75}{0{,}114} < \frac{1{,}6-1{,}75}{0{,}114}\right) \cong p(Z < -1,32) = 1 - p(Z \leq 1,32) = 1 - 0,9066 = 9,34\%$

X	a	b	1	2
Y	5	1	1	1

12.- Se considera la tabla estadística siguiente

a) Calcule a y b sabiendo que la media de X es 2, la varianza de X es $\frac{3}{2}$ y b > 1.

Resolución

$2 = \bar{x} = \frac{\Sigma x_i}{N} = \frac{a+b+3}{4} \Rightarrow a + b = 5$; $\frac{3}{2} = s_x^2 = \frac{\Sigma x_i^2}{N} - \bar{x}^2 = \frac{a^2+b^2+5}{4} - 2^2 = \frac{a^2+b^2-11}{4}$

Despejando, $12 = 2a^2 + 2b^2 - 22 \Rightarrow a^2 + b^2 = 17$. Como b = 5 – a, sustituyendo, $a^2 + (5 - a)^2 = 17$

Desarrollando y simplificando, $2a^2 - 10a + 8 = 0 \Rightarrow a^2 - 5a + 4 = 0$. Resolviendo, a = 4 ó a = 1

Si a = 4, b = 5 – 4 = 1 (imposible, pues b > 1) ; si a = 1, b = 5 – 1 = 4. Solución: a = 1, b = 4

b) Mediante la correspondiente recta de regresión lineal, haga una predicción del valor que se obtiene para Y cuando X = 3.

Resolución

Sabemos que a = 1, b = 5 y el nº de datos es N = 4. Elaboramos la tabla de frecuencias adecuada:

	1	4	1	2	$\rightarrow 8$
y_i	5	1	1	1	$\rightarrow 8$
x_i^2	1	16	1	4	$\rightarrow 22$
y_i^2	25	1	1	1	$\rightarrow 28$
$x_i y_i$	5	4	1	2	$\rightarrow 12$

$\overline{x} = \frac{\Sigma x_i}{N} = \frac{8}{4} = 2$; $\overline{y} = \frac{\Sigma y_i}{N} = \frac{8}{4} = 2$; $s_x^2 = \frac{\Sigma x_i^2}{N} - \overline{x}^2 = \frac{22}{4} - 2^2 = 1,5 = \frac{3}{2}$

$s_{xy} = \frac{\Sigma x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{12}{4} - 2 . 2 = -1$

La recta de regresión (de Y sobre X) es $r_{YX}: y - \overline{y} = \frac{s_{xy}}{s_x^2}(x - \overline{x}) \Rightarrow r_{YX}: y - 2 = \frac{-1}{3/2}(x - 2)$

Operando y simplificando, la recta de regresión que se pide es $r_{YX}: y = \frac{-2}{3}x + \frac{10}{3} \Rightarrow r_{YX}: y = \frac{10 - 2x}{3}$

Para $x = 3$, $y = \frac{10 - 2.3}{3} = \frac{4}{3} \cong 1,33$

13.- Los almendros de una plantación tienen una producción cada uno cuyo peso se distribuye normalmente con media 50 kg. Además, se sabe que la probabilidad de que un almendro produzca entre 50 y 60 kg es 0,3413.
a) Calcule la probabilidad de que un almendro produzca entre 40 y 50 kg.
b) Calcule la probabilidad de que un almendro produzca entre 45 y 55 kg.

Resolución

X = peso $\rightarrow N(50\,;\,\sigma) \Rightarrow Z = \frac{X - 50}{\sigma} \rightarrow N(0,\ 1)$. Según el enunciado, $p(50 < X < 60) = 0,3413$

Luego, $0,3413 = p(50 < X < 60) = p\left(\frac{50 - 50}{\sigma} < \frac{X - 50}{\sigma} < \frac{60 - 50}{\sigma}\right) = p\left(0 < Z < \frac{10}{\sigma}\right) =$

$= p\left(Z < \frac{10}{\sigma}\right) - p(Z < 0) = p\left(Z < \frac{10}{\sigma}\right) - 0,5$. Despejando, $p\left(Z < \frac{10}{\sigma}\right) = 0,8413$

Usando la tabla de la N(0, 1) en sentido inverso, $\frac{10}{\sigma} = 1 \Rightarrow \sigma = 10$.

a) Se pide $p(40 < X < 50)$
$= p\left(\frac{40 - 50}{10} < \frac{X - 50}{10} < \frac{50 - 50}{10}\right) = p(-1 < Z < 0) = p(Z < 0) - p(Z > 1) =$

$= p(Z < 0) - 1 + p(Z \leq 1) = 0,5 - 1 + 0,8413 = 0,3413 = 34,13\%$.

b) $p(45 < X < 55)$
$= p\left(\frac{45 - 50}{10} < \frac{X - 50}{10} < \frac{55 - 50}{10}\right) = p(-0,5 < Z < 0,5) = p(Z < 0,5) - p(Z > 0,5) =$

$= p(Z < 0,5) - 1 + p(Z \leq 0,5) = 2p(Z < 0,5) - 1 = 2 . 0,6915 - 1 = 0,383 = 38,3\%$.

14.- Dada la distribución bidimensional

X	0	1	2	3	4
Y	8	7	4	1	0

a) Si una de las tres ecuaciones: y = –2,2x + 8,4, y = 3,6x – 4,3, x = 3,4y + 7,2, corresponde a la recta de regresión de Y sobre X de dicha distribución, ¿cuál de ellas es?
b) Si uno de los tres números –0,98, 0,97, –1,8 corresponde al coeficiente de correlación lineal entre las variables X e Y ¿cuál de ellos es?

Resolución

Elaboramos la tabla de frecuencias adecuada:

	0	1	2	3	4 → 10

y_i	8	7	4	1	0	$\to 20$
x_i^2	0	1	4	9	16	$\to 30$
y_i^2	64	49	16	1	0	$\to 130$
$x_i y_i$	0	7	8	3	0	$\to 18$

$\overline{x} = \frac{\Sigma x_i}{N} = \frac{10}{5} = 2$; $\overline{y} = \frac{\Sigma y_i}{N} = \frac{20}{5} = 4$; $s_x^2 = \frac{\Sigma x_i^2}{N} - \overline{x}^2 = \frac{30}{5} - 2^2 = 2$

$s_y^2 = \frac{\Sigma y_i^2}{N} - \overline{y}^2 = \frac{130}{5} - 4^2 = 10$; $s_{xy} = \frac{\Sigma x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{18}{5} - 2 \,.\, 4 = -4,4$

La recta de regresión (de Y sobre X) es $r_{YX}: y - \overline{y} = \frac{s_{xy}}{s_x^2}(x - \overline{x}) \Rightarrow r_{YX}: y - 4 = \frac{-4,4}{2}(x - 2)$

Operando y simplificando, la recta de regresión es $r_{YX}: y = -2,2x + 8,4$.

El coeficiente de correlación lineal es $r = \frac{s_{xy}}{s_x s_y} = \frac{-4,4}{\sqrt{2}\,.\,\sqrt{10}} \cong -0,98$

15.-

a) Complete la siguiente tabla de frecuencias, sabiendo que la media de la variable X es 3.
(n_i representan las frecuencias absolutas y N_i las frecuencias absolutas acumuladas).

X	1	2	3	4	5
n_i		16		8	
N_i	9			52	

b) Calcule la media y la desviación típica de la distribución anterior.

Resolución

x_i	1	2	3	4	5	total
n_i	a	16	b	8	c	
N_i	9			52		-

Como $N_1 = 9 \Rightarrow$ a = 9

x_i	1	2	3	4	5	total
n_i	9	16	b	8	c	
N_i	9	25	25 + b	52	52 + c	-

Como 25 + b + 8 = 52, entonces b = 19

x_i	1	2	3	4	5	total
n_i	9	16	19	8	c	c + 52 = N
$x_i n_i$	9	32	57	32	5c	5c + 130
N_i	9	25	44	52	52 +c	-

$3 = \overline{x} = \frac{\Sigma x_i n_i}{N} = \frac{5c+130}{c+52} \Rightarrow 3c + 156 = 5c + 130$; 2c = 26 ; c = 13.

Queda la tabla:

x_i	1	2	3	4	5	total
n_i	9	16	19	8	13	65 = N
$x_i n_i$	9	32	57	32	65	195
$x_i^2 n_i$	9	64	171	128	325	697
N_i	9	25	44	52	65	-

b) La media ya sabemos que es 3

La desviación típica es $s = \sqrt{\frac{\Sigma x_i^2 n_i}{N} - \overline{x}^2} = \sqrt{\frac{697}{65} - 3^2} \cong 1,31266$