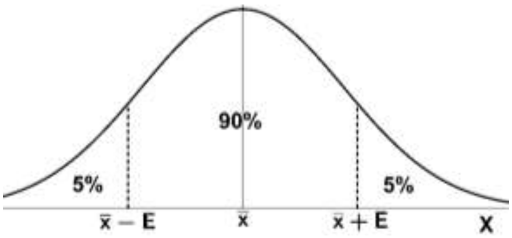
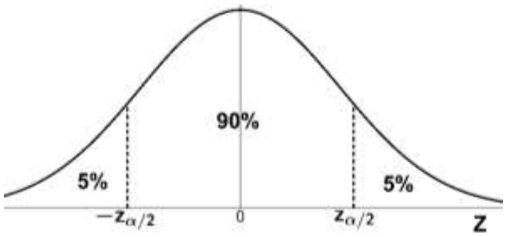

1.- En una población, una variable aleatoria sigue una ley normal de media desconocida y desviación típica 20.
a) Si de una muestra de tamaño 25 se ha observado que la media es 2743, determine un intervalo, con el 90% de confianza, para la media de la población.

Resolución

X → N(μ, 20). El intervalo de confianza a nivel de confianza del 90% para estimar μ,
es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 2743$ 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% – 90% = 10% y 10% : 2 = 5%





$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 90\% + 5\% = 95\% = 0{,}95$

Esta probabilidad coincide con $\frac{1+n_c}{2} = \frac{1+0{,}9}{2} = 0{,}95$

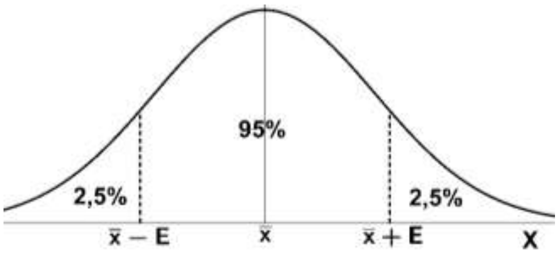
Como $p(Z < z_{\alpha/2}) = 0{,}95 \xrightarrow{\text{usando la tabla de la } N(0,1),\ \text{por interpolación}} z_{\alpha/2} = 1{,}645$

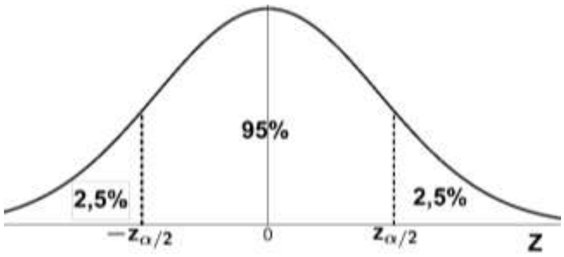
$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1{,}645.\frac{20}{\sqrt{25}} = 6{,}58$; $I_c = (2743 - 6{,}58\,;\ 2743 + 6{,}58) \cong (2736{,}42\,;\ 2749{,}58)$

b) Elegida una muestra, su media ha sido 2740. Se ha construido un intervalo de confianza, al 95%, que ha resultado ser (2736,08, 2743,92). ¿Cuál era el tamaño de la muestra?

Resolución

$\overline{x} = 2740$ la media de la muestra de tamaño n.
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 $p(Z < z_{\alpha/2}) = 0{,}975 \xrightarrow{\text{usando la tabla de la } N(0,1)} z_{\alpha/2} = 1{,}96$.

El error de estimación es la mitad de la longitud del intervalo de confianza $\Rightarrow E = \frac{2743{,}92 - 2736{,}08}{2} = 3{,}92$.

Piden hallar n sabiendo que $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 3{,}92 \Leftrightarrow z_{\alpha/2}.\frac{\sigma}{3{,}92} = \sqrt{n} \xrightarrow{\text{elevando al cuadrado}} (z_{\alpha/2})^2.\frac{\sigma^2}{3{,}92^2} = n$

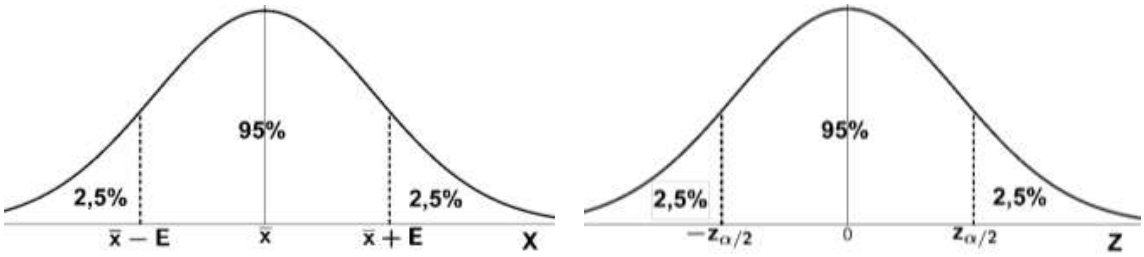
Sustituyendo: $n = 1{,}96^2.\frac{20^2}{3{,}92^2} = 100$. Luego, el tamaño de la muestra es 100

2.- El tiempo de vida de un tipo de insecto sigue una distribución normal con media desconocida y desviación típica 25 días. Para estimar la vida media se hace un seguimiento a la duración de la vida de una muestra de n insectos. Calcule el valor de n para que el intervalo de confianza de esta media, con un nivel de confianza del 95%, tenga una amplitud como máximo de 5 días.

Resolución

X = tiempo de vida → $N(\mu, 25)$. El intervalo de confianza a nivel de confianza del 95% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x}$ la media de la muestra de tamaño n 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 $p(Z < z_{\alpha/2}) = 0,975 \xrightarrow{\text{usando la tabla de la } N(0,1)} z_{\alpha/2} = 1,96$.

El error de estimación es la mitad de la amplitud $\Rightarrow E = \frac{5}{2} = 2,5$.

Piden hallar n sabiendo que $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \leq 2,5 \Leftrightarrow z_{\alpha/2}.\frac{\sigma}{2,5} \leq \sqrt{n} \xrightarrow{\text{elevando al cuadrado}} (z_{\alpha/2})^2.\frac{\sigma^2}{2,5^2} \leq n$

Sustituyendo: $n \geq 1,96^2.\frac{25^2}{2,5^2} = 384,16$. Luego, n = 385

3.- (prueba extraordinaria) Las ventas mensuales de una tienda de electrodomésticos se distribuyen según una ley normal con desviación típica 90000 ptas. En un estudio estadístico de las ventas realizadas en los últimos 9 meses se ha encontrado un intervalo de confianza para la media mensual de las ventas, cuyos extremos son 466300 y 583900 ptas.
a) ¿Cuál ha sido la media de las ventas en estos 9 meses?
b) ¿Cuál es el nivel de confianza de este intervalo?

Resolución

X = valor de las ventas → $N(\mu, 90000)$. El intervalo de confianza a nivel de confianza del $n_c = 1 - \alpha$ para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x}$ la media de la muestra de tamaño n = 9 y $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$ es el máximo error de estimación.

Como la media es el punto medio del intervalo de confianza, $\overline{x} = \frac{466300 + 583900}{2} = 525100$

Como el error de estimación es la mitad de la longitud del intervalo de confianza, entonces

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = \frac{583900 - 466300}{2} = 58800 \Rightarrow z_{\alpha/2} = \frac{58800\sqrt{n}}{\sigma} = \frac{58800\sqrt{9}}{90000} = 1,96$. Luego,

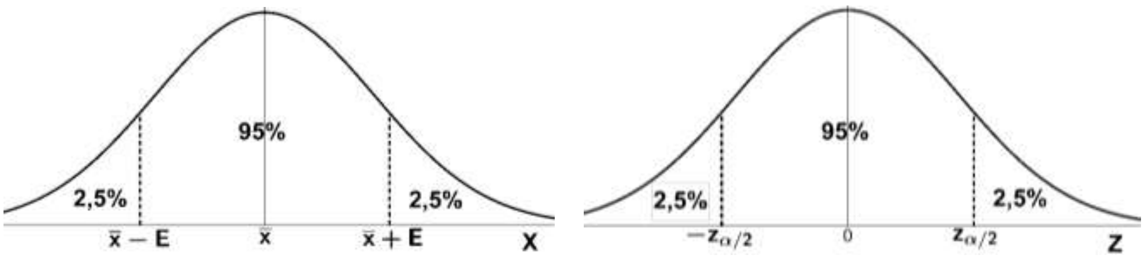
$\frac{1 + n_c}{2} = p(Z < 1,96) = 0,975$. Despejando, el nivel de confianza es $n_c = 2\,.\,0,975 - 1 = 0,95 = 95\%$

4.- (prueba extraordinaria) La media de las estaturas de una muestra aleatoria de 400 personas de una ciudad es 1,75 metros. Se sabe que la estatura de las personas de esa ciudad es una variable aleatoria que sigue una distribución normal con varianza 0,16 m^2.
a) Construya un intervalo, de un 95% de confianza, para la media de las estaturas de la población.

Resolución

X = estaturas $\rightarrow N\left(\mu, \sqrt{0{,}16}\right) = N(\mu, 0{,}4)$. El intervalo de confianza a nivel de confianza del 95% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 1{,}75$ la media de la muestra de tamaño n = 400 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 $\frac{1 + n_c}{2} = \frac{1 + 0{,}95}{2} = 0{,}975$

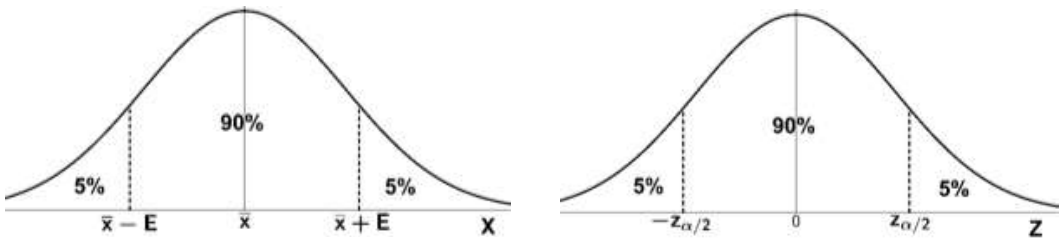
Como $p(Z < z_{\alpha/2}) = 0{,}975 \xrightarrow{\text{usando la tabla de la } N(0,1)} z_{\alpha/2} = 1{,}96.$

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1{,}96.\frac{0{,}4}{\sqrt{400}} = 0{,}0392$; $I_c = (1{,}75 - 0{,}0392\,;\, 1{,}75 + 0{,}0392) \cong (1{,}7108\ ;\ 1{,}7892)$

b) ¿Cuál sería el mínimo tamaño muestral necesario para que pueda decirse que la verdadera media de las estaturas está a menos de 2 cm de la media muestral, con una confianza del 90%?

Resolución

Observa que 2 cm = 0,02 m. Luego, el error de estimación, E, debe ser E ≤ 0,02.
Como el área bajo la campana es 100%, entonces 100% – 90% = 10% y 10% : 2 = 5%



$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 90\% + 5\% = 95\% = 0{,}95$

Esta probabilidad coincide con $\frac{1 + n_c}{2} = \frac{1 + 0{,}9}{2} = 0{,}95$

Como $p(Z < z_{\alpha/2}) = 0{,}95 \xrightarrow{\text{usando la tabla de la } N(0,1),\ \text{por interpolación}} z_{\alpha/2} = 1{,}645$

Piden hallar n sabiendo que $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \le 0{,}02 \Leftrightarrow z_{\alpha/2}.\frac{\sigma}{2} \le \sqrt{n} \xrightarrow{\text{elevando al cuadrado}} (z_{\alpha/2})^2.\frac{\sigma^2}{0{,}02^2} \le n$

Sustituyendo: $n \ge 1{,}645^2.\frac{0{,}4^2}{0{,}02^2} = 1082{,}4 \Rightarrow$ el mínimo tamaño muestral necesario es n = 1083 personas

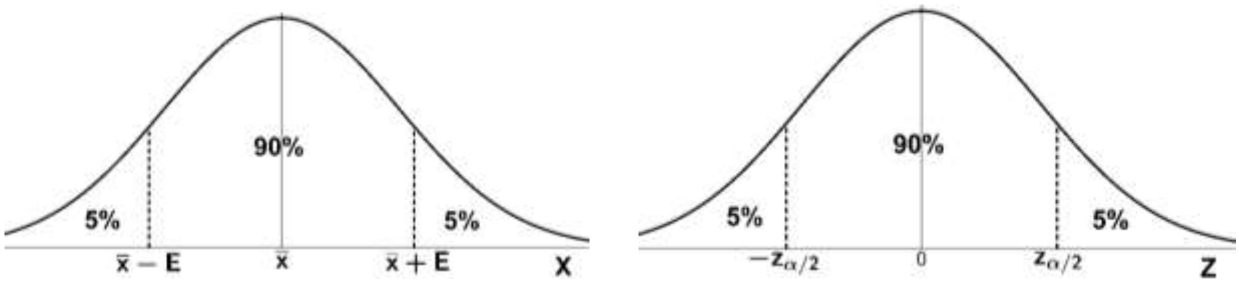
---

5.- (prueba ordinaria) Se ha tomado una muestra aleatoria de 100 individuos a los que se ha medido el nivel de glucosa en sangre, obteniéndose una media muestral de 110 mg/cc. Se sabe que la desviación típica de la población es de 20 mg/cc.
a) Obtén un intervalo de confianza, al 90%, para el nivel de glucosa en sangre en la población.
b) ¿Qué error máximo se comete con la estimación anterior?

Resolución

X = nivel de glucosa $\rightarrow N(\mu, 20)$. El intervalo de confianza a nivel de confianza del 90% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 110$ la media de la muestra de tamaño n = 100 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% – 90% = 10% y 10% : 2 = 5%



$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 90\% + 5\% = 95\% = 0{,}95$

Esta probabilidad coincide con $\frac{1 + n_c}{2} = \frac{1 + 0{,}9}{2} = 0{,}95$

Como $p(Z < z_{\alpha/2}) = 0{,}95 \xrightarrow{\text{usando la tabla de la } N(0,1),\ \text{por interpolación}} z_{\alpha/2} = 1{,}645$

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1{,}645.\frac{20}{\sqrt{100}} = 3{,}29$; $I_c = (110 - 3{,}29\,;\ 110 + 3{,}29) \cong (106{,}71\,; 113{,}29)$

6.- (prueba ordinaria) La media de edad de los alumnos que se presentan a las pruebas de acceso a la Universidad es de 18,1 años y la desviación típica 0,6 años.
(a) De los alumnos anteriores se elige, al azar, una muestra de 100, ¿cuál es la probabilidad de que la media de la edad de la muestra esté comprendida entre 17,9 y 18,2 años?

Resolución

X = edad $\rightarrow N(18{,}1\,;\ 0{,}6)$. 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, $\overline{X} \rightarrow N\left(18{,}1\,;\ \frac{0{,}6}{\sqrt{100}}\right) \cong N(18{,}1\,; 0{,}06)$. Tipificando, $Z = \frac{\overline{X} - 18{,}1}{0{,}06} \rightarrow N(0, 1)$

Nos piden $p(17{,}9 < \overline{X} < 18{,}2)$. Restamos 15 y dividimos entre 0,48 en los dos miembros

$p\left(\frac{17{,}9 - 18{,}1}{0{,}06} < \frac{\overline{X} - 18{,}1}{0{,}06} < \frac{18{,}2 - 18{,}1}{0{,}06}\right) \cong p(-3{,}33 < Z < 1{,}67) = p(Z < 1{,}67) - p(Z > 3{,}33) =$

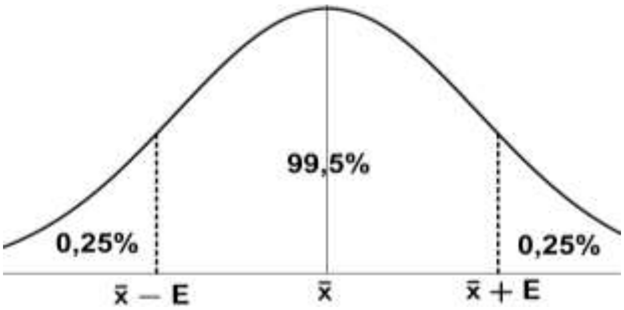
$= p(Z < 1{,}67) - 1 + p(Z < 3{,}33) = 0{,}9525 - 1 + 0{,}99957 = 0{,}95207 \cong 95{,}21\%$

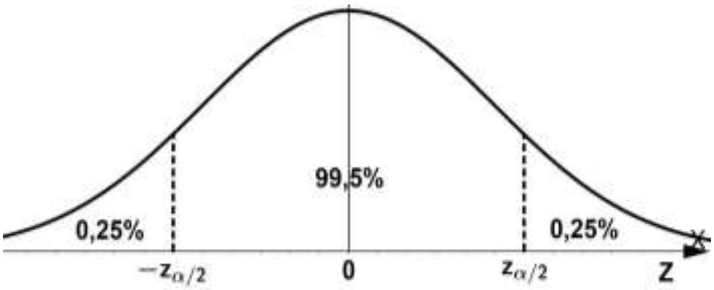
(b) ¿Qué tamaño debe tener una muestra de dicha población para que su media esté comprendida entre 17,9 y 18,3 años, con una confianza del 99,5%?

Resolución

X = edad → $N(18{,}1\ ;\ 0{,}6)$. El intervalo de confianza a nivel de confianza del 99,5% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x}$ la media de la muestra de tamaño n y $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$ es el máximo error de estimación. En este caso, $I_c = (17{,}9\ ;18{,}3)$

Como el área bajo la campana es 100%, entonces 100% – 99,5% = 0,5% y 0,5% : 2 = 0,25%





$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 99{,}5\% + 0{,}25\% = 99{,}75\% = 0{,}9975$

Esta probabilidad coincide con $\frac{1 + n_c}{2} = \frac{1 + 0{,}995}{2} = 0{,}9975$

Como $p(Z < z_{\alpha/2}) = 0{,}9975 \xrightarrow{usando\ la\ tabla\ de\ la\ N(0,1)} z_{\alpha/2} = 2{,}81$

Como el error de estimación es la mitad de la longitud del intervalo de confianza, entonces

$$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = \frac{18{,}3 - 17{,}9}{2} = 0{,}2 \Leftrightarrow z_{\alpha/2}.\frac{\sigma}{0{,}2} = \sqrt{n} \xrightarrow{elevando\ al\ cuadrado} (z_{\alpha/2})^2.\frac{\sigma^2}{0{,}2^2} = n$$

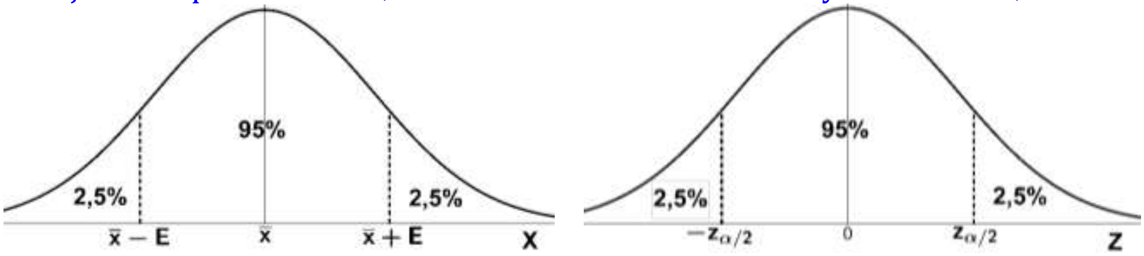
Sustituyendo: $n = 2{,}81^2.\frac{0{,}6^2}{0{,}2^2} \cong 71{,}1 \Rightarrow$ el tamaño muestral mínimo necesario es n = 72 alumnos

7.- Un fabricante de bombillas sabe que la desviación típica de la duración de las bombillas es 90 horas. Tomada una muestra de tamaño 100 se ha encontrado que la media de la duración de las bombillas ha sido 1200 horas. Determine un intervalo, con el 95% de confianza, para la duración media de las bombillas.

Resolución

X = duración → $N(\mu, 90)$. El intervalo de confianza a nivel de confianza del 95% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 1200$ la media de la muestra de tamaño n = 100 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 $p(Z < z_{\alpha/2}) = 0{,}975 \xrightarrow{\text{usando la tabla de la } N(0,1)} z_{\alpha/2} = 1{,}96$.

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1{,}96.\frac{90}{\sqrt{100}} = 17{,}64$; $I_c = (1200 - 17{,}64 \,;\, 1200 + 17{,}64) \cong (1182{,}36 \;;\; 1217{,}64)$

8.- Sea un conjunto de cuatro bolas, marcadas con los números 1, 3, 5 y 7.
a) Escriba todas las muestras de tamaño 2 que podrían formarse con esas bolas si el muestreo se hace sin reposición; calcule las medias de los números de cada muestra y halle la media de todas esas medias.

Resolución

Las muestras son {1–3, 1–5, 1–7, 3–1, 3–5, 3–7, 5–1, 5–3, 5–7, 7–1, 7–3, 7–5}. Son 12 muestras.

muestras	1–3	1–5	1–7	3–1	3–5	3–7	5–1	5–3	5–7	7–1	7–3	7–5
medias	2	3	4	2	4	5	3	4	6	4	5	6

Por el teorema central del límite la media de las medias muestrales es μ siendo μ la media poblacional.

Luego, la media de las medias muestrales es $m = \mu = \frac{1+3+5+7}{4} = 4$

Sin aplicar el teorema veamos que efectivamente es esa:

$medias = \bar{x}_i$	2	3	4	5	6	*total*
$frecuencia = f_i$	2	2	4	2	2	$n = 12$
$\bar{x}_i f_i$	4	6	16	10	12	48

; $m = \frac{\sum \bar{x}_i f_i}{n} = \frac{48}{12} = 4$

b) Haga lo mismo que en a) pero suponiendo que el muestreo se hace con reemplazamiento.

Resolución

Las muestras son
{1–1, 1–3, 1–5, 1–7, 3–1, 3–3, 3–5, 3–7, 5–1, 5–3, 5–5, 5–7, 7–1, 7–3, 7–5, 7–7}. Son 16 muestras.

1–1	1–3	1–5	1–7	3–1	3–3	3–5	3–7	5–1	5–3	5–5	5–7	7–1	7–3	7–5	7–7
1	2	3	4	2	3	4	5	3	4	5	6	4	5	6	7

Por el teorema central del límite la media de las medias muestrales es μ siendo μ la media poblacional.

Luego, la media de las medias muestrales es $m = \mu = \frac{1+3+5+7}{4} = 4$

Sin aplicar el teorema veamos que efectivamente es esa:

$medias = \bar{x}_i$	1	2	3	4	5	6	7	*total*
$frecuencia = f_i$	1	2	3	4	3	2	1	$n = 16$
$\bar{x}_i f_i$	1	4	9	16	15	12	7	44

; $m = \frac{\sum \bar{x}_i f_i}{n} = \frac{64}{16} = 4$

9.- Al calificar los exámenes de un numeroso grupo de opositores, se ha visto que sus puntuaciones siguen una distribución normal con una media de 72 puntos y una desviación típica de 9 puntos.
¿Cuál es la probabilidad de que, en una muestra de 16 de esos opositores, elegidos al azar, se obtenga una puntuación media superior a 78 puntos?

Resolución

X = puntuación $\rightarrow N(72, 9)$. 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 = 16. Sustituyendo, $\overline{X} \rightarrow N\left(72\,;\ \frac{9}{\sqrt{16}}\right) = N(72\,; 2{,}25)$. Tipificando, $Z = \frac{\overline{X}-72}{2{,}25} \rightarrow N(0,1)$

Nos piden $p(\overline{X} > 78) = p\left(\frac{\overline{X}-72}{2{,}25} > \frac{78-72}{2{,}25}\right) \cong p(Z > 2{,}67) = 1 - p(Z \leq 2{,}67) = 1 - 0{,}9962 = 0{,}38\%$

10.- El tiempo que permanece cada paciente en la consulta de cierto medico es una variable aleatoria que sigue una distribución normal con una desviación típica de 4 minutos.
Se ha tomado una muestra aleatoria de 256 pacientes de este médico y se ha encontrado que su tiempo medio de consulta ha sido de 10 minutos. ¿Cuál es el intervalo de confianza, a un nivel del 95%, para el tiempo medio de consulta que se deduce de esta muestra?

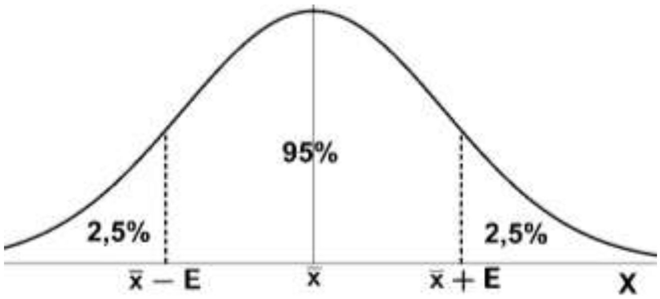
Resolución

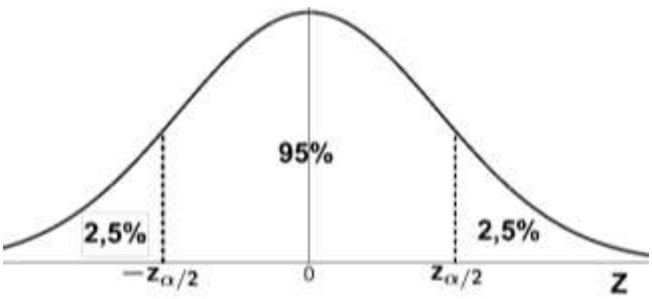
X = tiempo $\rightarrow N(\mu, 4)$. El intervalo de confianza a nivel de confianza del 95% para

estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 10$ la media de la muestra de tamaño n = 256

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

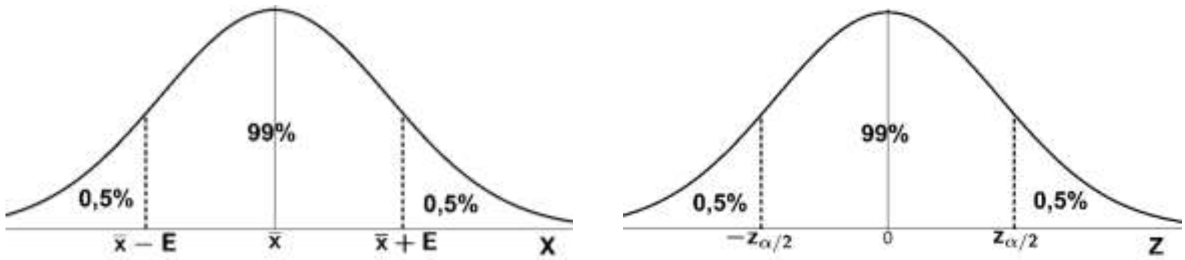
$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1{,}96.\frac{4}{\sqrt{256}} = 0{,}49$; $I_c = (10 - 0{,}49\,;\ 10 + 0{,}49) \cong (9{,}51\,; 10{,}49)$

11.- Se sabe que el tiempo de reacción a un determinado estimulo se distribuye según una ley normal de media desconocida y desviación típica 0,15 segundos.
Observada una muestra de tamaño 9, se ha obtenido una media muestral de 0,85 segundos.
a) Obtenga un intervalo de confianza para la media de la población, con un nivel de confianza del 99%.

Resolución

X = tiempo → $N(\mu\,; 0{,}15)$. El intervalo de confianza a nivel de confianza del 99% para estimar la media, μ es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 0{,}85$ la media de una muestra de tamaño n = 9, $E = z_{\alpha/2} . \frac{\sigma}{\sqrt{n}}$, el máximo error de estimación.

Como el área bajo la campana es 100%, entonces 100% – 99% = 1% y 1% : 2 = 0,5%



$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 99\% + 0{,}5\% = 99{,}5\% = 0{,}995$

Esta probabilidad coincide con $\frac{1 + n_c}{2} = \frac{1 + 0{,}99}{2} = 0{,}995$

Como $p(Z < z_{\alpha/2}) = 0{,}995 \xrightarrow{\textit{usando la tabla de la } N(0,1),\ \textit{por interpolación}} z_{\alpha/2} = 2{,}575$

$E = z_{\alpha/2} . \frac{\sigma}{\sqrt{n}} = 2{,}575 . \frac{0{,}15}{\sqrt{9}} = 0{,}12875$; $I_c = (0{,}85 - 0{,}12875\,;\ 0{,}85 + 0{,}12875) \cong (0{,}72125\ ;\ 0{,}97875)$

b) ¿Con qué nivel de confianza se debería construir un intervalo para la media de manera que los límites de dicho intervalo fuesen 0,768 y 0,932?

Resolución

Como el error de estimación es la mitad de la longitud del intervalo de confianza, entonces

$E = z_{\alpha/2} . \frac{\sigma}{\sqrt{n}} = \frac{0{,}932 - 0{,}768}{2} = 0{,}082 \Rightarrow z_{\alpha/2} = \frac{0{,}082\,\sqrt{n}}{\sigma} = \frac{0{,}082\,\sqrt{9}}{0{,}15} = 1{,}64$. Luego,

$\frac{1 + n_c}{2} = p(Z < 1{,}64) = 0{,}9495$. Despejando, el nivel de confianza es $n_c = 2 . 0{,}9495 - 1 = 89{,}9\%$

12.- En un colegio hay 2000 alumnos distribuidos en 5 cursos así:
400 en 1er curso, 380 en 2°, 520 en 3°, 360 en 4° y 340 en 5°. Se quiere seleccionar una muestra de 100 alumnos, utilizando la técnica de muestreo aleatorio estratificado con afijación proporcional y considerando cada curso como un estrato. ¿Cómo se seleccionaría dicha muestra?

Resolución

Estratos	1º	2º	3º	4º	5º	Total
Nº de elementos de la muestra	x	y	z	t	u	100
Nº de elementos de la población	400	380	520	360	340	2000

Usando la proporcionalidad

$\frac{x}{400} = \frac{y}{380} = \frac{z}{520} = \frac{t}{360} = \frac{u}{340} = \frac{100}{2000} = 0{,}05$. Despejando, x = 20, y = 19, z = 26, t = 18, u = 17.

Luego, se deben seleccionar 20 alumnos de 1º, 19 de 2, 26 de 3º, 18 de 4º y 17 de 5º.

13.- Una población de tamaño 1000 se ha dividido en 4 estratos de tamaño 150, 400, 250 y 200.
a) Utilizando la técnica de muestreo aleatorio estratificado con afijación proporcional se han seleccionado 10 individuos del tercer estrato. ¿Cuál es el tamaño de la muestra?

Resolución

Estratos	1er estrato	2º estrato	3er estrato	4º estrato	Total
Nº de elementos de la muestra	x	y	10	z	n
Nº de elementos de la población	150	400	250	200	1000

Usando la proporcionalidad

$\frac{x}{150} = \frac{y}{400} = \frac{10}{250} = 0{,}04 = \frac{z}{200} = \frac{n}{1000}$. Despejando, x = 6, y = 16, z = 8, n = 40.

Luego, el tamaño de la muestra es de 40 individuos.

b) Utilizando la técnica del apartado anterior se desea seleccionar una muestra de tamaño 20. ¿Cómo debe tomarse dicha muestra?

Resolución

Estratos	1er estrato	2º estrato	3er estrato	4º estrato	Total
Nº de elementos de la muestra	x	y	z	t	20
Nº de elementos de la población	150	400	250	200	1000

Usando la proporcionalidad

$\frac{x}{150} = \frac{y}{400} = \frac{z}{250} = \frac{t}{200} = \frac{20}{1000} = 0{,}02$. Despejando, x = 3, y = 8, z = 5, t = 4.

Luego, se deben seleccionar 3 individuos del 1er estrato, 8 del 2º, 5 del 3º y 4 del 4º

14.- Si una variable aleatoria X obedece a una ley normal de media 10 y desviación típica 2, contesta:
a) Si se toman distintas muestras, todas de tamaño 100, de esa variable X y se considera la nueva variable $\overline{X}$ media muestral, obtener su media y su desviación típica.

Resolución

X $\rightarrow N(10, 2)$. 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 = 100. Sustituyendo, $\overline{X} \rightarrow N\left(10\,;\ \frac{2}{\sqrt{100}}\right) = N(10\,; 0{,}2)$.

b) ¿Cuál es la probabilidad de que X tome valores fuera del intervalo [9, 11]?

Resolución

Sabemos que $Z = \frac{X-10}{2} \rightarrow N(0,1)$. Hallemos primero $p(X \in [9,11]) = p(9 \leq X \leq 11) =$

$= p\left(\frac{9-10}{2} \leq \frac{X-10}{2} \leq \frac{11-10}{2}\right) = p(-0{,}5 \leq Z \leq 0{,}5) = p(Z \leq 0{,}5) - p(Z \geq 0{,}5) =$

$= p(Z \leq 0{,}5) - 1 + p(Z \leq 0{,}5) = 2p(Z \leq 0{,}5) - 1 = 2\,.\,0{,}6915 - 1 = 0{,}383$

Luego, la probabilidad que se pide es $p(X \notin [9,11]) = 1 - 0{,}383 = 0{,}617 = 61{,}7\%$

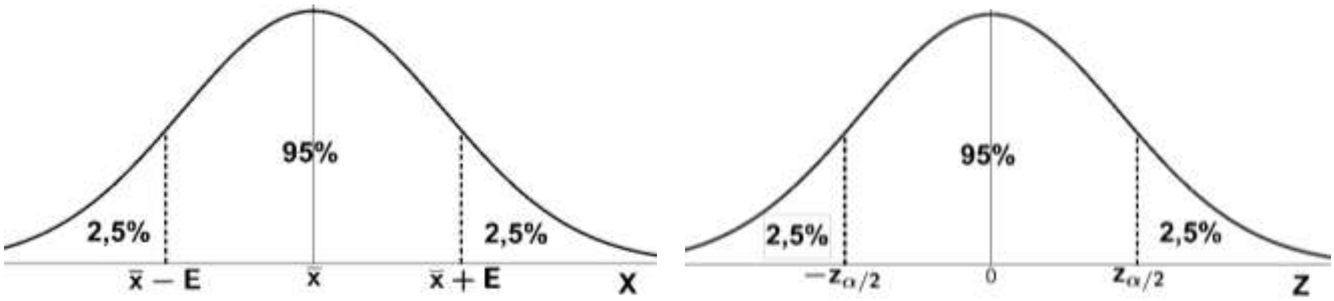
15.- Se desea estimar la cantidad media de dinero que los estudiantes de un centro se gastan para el día de San Valentín. Se sabe que la desviación típica de toda la población es de 700 ptas y que una muestra de 49 estudiantes se gastó una media de 2300 ptas.
a) Hallar un intervalo de confianza con un nivel del 95% para la media poblacional.

Resolución

X = cantidad de dinero → $N(\mu, 700)$. El intervalo de confianza a nivel de confianza del 95% para estimar μ, es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x} = 2300$ la media de la muestra de tamaño n = 49

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

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 1,96.\frac{700}{\sqrt{49}} = 196$; $I_c = (2300 - 196,\ 2300 + 196) = (2104, 2496)$

b) ¿Con que nivel de confianza se puede afirmar que la media de todo el centro está entre 2050 ptas y 2550 ptas?

Resolución

Como el error de estimación es la mitad de la longitud del intervalo de confianza, entonces

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = \frac{2550 - 2050}{2} = 250 \Rightarrow z_{\alpha/2} = \frac{250\sqrt{n}}{\sigma} = \frac{250\sqrt{49}}{700} = 2,5$. Luego,

$\frac{1 + n_c}{2} = p(Z < 2,5) = 0,9938$. Despejando, el nivel de confianza es $n_c = 2 . 0,9938 - 1 = 98,76\%$

16.- Los pesos de 4500 estudiantes de Bachillerato de una ciudad están distribuidos normalmente con una media de 56,5 kg y una desviación típica de 2,5 kg. Si se selecciona al azar una muestra de 16 estudiantes, ¿Cuál es la probabilidad de que tenga una media superior a 58 kg?

Resolución

X = peso → $N(56,5\ ;\ 2,5)$. Sabemos, por el teorema central del límite que si $X \to N(\mu, \sigma)$

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

n = 16. Sustituyendo, $\overline{X} \to N\left(56,5\ ;\ \frac{2,5}{\sqrt{16}}\right) = N(56,5\ ; 0,625)$. Tipificando, $Z = \frac{\overline{X} - 56,5}{0,625} \to N(0, 1)$

Nos piden $p(\overline{X} > 58) = p\left(\frac{\overline{X} - 56,5}{0,625} > \frac{58 - 56,5}{0,625}\right) \cong p(Z > 2,4) = 1 - p(Z \leq 2,4) = 1 - 0,9918 = 0,82\%$

17.- El coeficiente intelectual de los alumnos y alumnas de un centro se distribuye según una ley normal de media 110 y desviación típica 16. Nos proponemos extraer una muestra aleatoria de 25 alumnos.
a) ¿Cuál será la distribución de las medias de las muestras que pueden extraerse?
b) ¿Cuál será la probabilidad de que el coeficiente intelectual medio de los 25 alumnos de la muestra sea superior a 115?

Resolución

X = coeficiente intelectual → $N(110, 16)$. 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 $\overline{X} \to N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

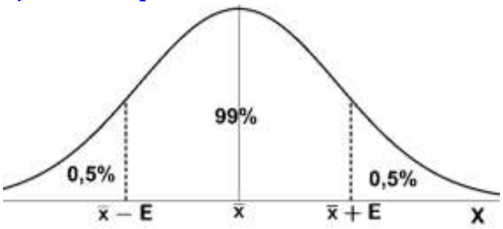
n = 25. Sustituyendo, $\overline{X} \to N\left(110\ ;\ \frac{16}{\sqrt{25}}\right) = N(110\ ; 3{,}2)$. Tipificando, $Z = \frac{\overline{X} - 110}{3{,}2} \to N(0, 1)$

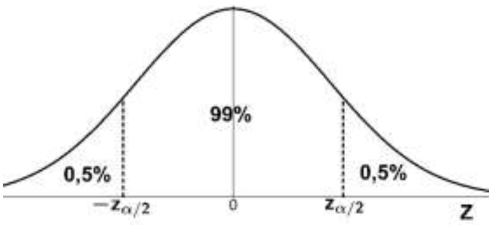
Piden $p(\overline{X} > 115) = p\left(\frac{\overline{X} - 110}{3{,}2} > \frac{115 - 110}{3{,}2}\right) \cong p(Z > 1{,}56) = 1 - p(Z \leq 1{,}56) = 1 - 0{,}9406 = 5{,}94\%$

c) Dar el intervalo para la media de los coeficientes intelectuales de la muestra con un nivel de confianza del 99%.

El intervalo de confianza a nivel de confianza del 99% para estimar la media, μ es $I_c = (\overline{x} - E, \overline{x} + E)$, siendo $\overline{x}$ la media de una muestra de tamaño n = 25, $E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$, el máximo error de estimación.

Como el área bajo la campana es 100%, entonces 100% – 99% = 1% y 1% : 2 = 0,5%





$z_{\alpha/2}$ es el valor de la N(0, 1) que cumple $p(Z < z_{\alpha/2}) = 99\% + 0{,}5\% = 99{,}5\% = 0{,}995$

Esta probabilidad coincide con $\frac{1 + n_c}{2} = \frac{1 + 0{,}99}{2} = 0{,}995$

Como $p(Z < z_{\alpha/2}) = 0{,}995 \xrightarrow{usando\ la\ tabla\ de\ la\ N(0,1),\ \ por\ interpolación} z_{\alpha/2} = 2{,}575$

$E = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 2{,}575.\frac{16}{\sqrt{25}} = 8{,}2432\ ;\ I_c = (\overline{x} - 8{,}2432\ ;\ \overline{x} + 8{,}2432)$

18.- El número de horas semanales de ocio que los estudiantes de Bachillerato dedican a la lectura sigue una Normal de media 7 horas y varianza 1,44. Tomamos una clase de 36 estudiantes para estudiar su media muestral.
a) ¿Cuál es la probabilidad de que esta media esté entre 7,3 y 7,6 horas?

Resolución

X = nº horas → $N(7, \sqrt{1{,}44}) = N(7, 1{,}2)$. 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, $\overline{X} \to N\left(7\ ;\ \frac{1{,}2}{\sqrt{36}}\right) = N(7\ ; 0{,}2)$. Tipificando, $Z = \frac{\overline{X} - 7}{0{,}2} \to N(0, 1)$

Piden $p(7{,}3 < \overline{X} < 7{,}6) = p\left(\frac{7{,}3 - 7}{0{,}2} < \frac{\overline{X} - 7}{0{,}2} < \frac{7{,}6 - 7}{0{,}2}\right) = p(1{,}5 < Z < 3) =$

$= p(Z < 3) - p(Z < 1{,}5) = 0{,}9987 - 0{,}9332 = 0{,}0655 = 6{,}55\%$

---

b) Si tomáramos 150 muestras de 36 estudiantes cada una, ¿En cuántas de ellas cabría esperar una media muestral inferior a 7,1 horas?

Resolución

Hallemos primero $p(\overline{X} < 7{,}1) = p\left(\frac{\overline{X} - 7}{0{,}2} < \frac{7{,}1 - 7}{0{,}2}\right) = p(Z < 0{,}5) = 0{,}6915 = 69{,}15\%$

Luego, el nº de muestras que se pide es 69,15% de 150 = 103,725 ≅ 104 muestras

19.- El coeficiente intelectual de una población sigue una distribución normal con media 100 y desviación típica 12. Se elige al azar una muestra de 16 personas.
a) ¿Qué distribución sigue la media muestral? Justifícalo. ¿Cuál es su media y su desviación típica?
b) Encontrar la probabilidad de que la media de sus coeficientes intelectuales este comprendida entre 98 y 102.

Resolución

X = coeficiente intelectual → $N(100, 12)$. 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 = 16. Sustituyendo, $\overline{X} \to N\left(100, \frac{12}{\sqrt{16}}\right) = N(100, 3)$. La media es 100 y la desviación típica 3

Para el b), tipificando, $Z = \frac{\overline{X} - 100}{3} \to N(0, 1)$

Piden $p(98 < \overline{X} < 102) = p\left(\frac{98 - 100}{3} < \frac{\overline{X} - 100}{3} < \frac{102 - 100}{3}\right) \cong p(-0{,}67 < Z < 0{,}67) =$

$= p(Z < 0{,}67) - p(Z > 0{,}67) = p(Z < 0{,}67) - 1 + p(Z < 0{,}67) = 2\,.\,0{,}7486 - 1 = 0{,}4972 = 49{,}72\%$

OTROS DEL 1999 (COU II)

1.- Las notas de 10 alumnos de COU en dos asignaturas X e Y vienen recogidas en la siguiente tabla.

X	5	3	6	5	9	4	1	2	6	7
Y	7	6	1	8	7	2	2	3	6	5

a) Halle la varianza de X y la de Y.
b) Halle la ecuación de la recta de regresión Y sobre X.
c) ¿Qué nota, en la asignatura Y, cabe esperar sabiendo que la nota obtenida en la asignatura X es 3,5? Comente la fiabilidad de la predicción.

Resolución

El nº de datos es N = 10. Elaboramos la tabla de frecuencias adecuada:

x_i	5	3	6	5	9	4	1	2	6	7	→ 48
y_i	7	6	1	8	7	2	2	3	6	5	→ 47
x_i^2	25	9	36	25	81	16	1	4	36	49	→ 282
y_i^2	49	36	1	64	49	4	4	9	36	25	→ 277
$x_i y_i$	35	18	6	40	63	8	2	6	36	35	→ 249

$\overline{x} = \frac{\sum x_i}{N} = \frac{48}{10} = 4,8$; varianza de X $= s_x^2 = \frac{\sum x_i^2}{N} - \overline{x}^2 = \frac{282}{10} - 4,8^2 = 5,16$

$\overline{y} = \frac{\sum y_i}{N} = \frac{47}{10} = 4,7$; varianza de Y $= s_y^2 = \frac{\sum y_i^2}{N} - \overline{y}^2 = \frac{277}{10} - 4,7^2 = 5,61$

$s_{xy} = \frac{\sum x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{249}{10} - 4,8\,.\,4,7 = 2,34$

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

$r_{YX}: y - 4,7 = \frac{2,34}{5,16}(x - 4,8) \Rightarrow r_{YX}: y = 0,453488x + 2,5232558$

Para x = 3,5, sustituyendo en la recta $y = 0,453488\,.\,3,5 + 2,5232558 \cong 4,11$

El coeficiente de correlación es $r = \frac{s_{xy}}{s_x\, s_y} = \frac{2,34}{\sqrt{5,16}\,.\,\sqrt{5,61}} \cong 0,434922$

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

2.- Calcule la desviación típica de una variable aleatoria X que sigue una normal de media 110, sabiendo que p(85 ≤ X ≤ 135) = 0,99.

Resolución

X → $N(110, \sigma)$. Tipificando, $Z = \frac{X - 110}{\sigma} \to N(0,1)$

$0,99 = p(85 \le X \le 135) = p\left(\frac{85-110}{\sigma} \le \frac{X-110}{\sigma} \le \frac{135-110}{\sigma}\right) = p\left(\frac{-25}{\sigma} \le Z \le \frac{25}{\sigma}\right) =$

$= p\left(Z \le \frac{25}{\sigma}\right) - p\left(Z \ge \frac{25}{\sigma}\right) = p\left(Z \le \frac{25}{\sigma}\right) - 1 + p\left(Z \le \frac{25}{\sigma}\right) = 2p\left(Z \le \frac{25}{\sigma}\right) - 1$

Despejando, $p\left(Z \le \frac{25}{\sigma}\right) = \frac{0,99+1}{2} = 0,995$. Usando la tabla de la N(0, 1) en sentido inverso, por interpolación, $\frac{25}{\sigma} = 2,575$. Luego, la desviación típica es $\sigma = \frac{25}{2,575} \cong 9,71$

3.- De una muestra de 50 bombillas se han obtenido los siguientes datos sobre su duración en horas:

Duración	[80, 85)	[85, 90)	[90, 95)	[95, 100)	[100, 105)	[105, 110)
Nº de bombillas	2	7	14	16	9	2

a) Calcule la media y la desviación típica de la duración.
b) Calcule el cuartil primero y el porcentaje de bombillas que duran 95 horas o más.

Resolución

X = duración	[80, 85)	[85, 90)	[90, 95)	[95, 100)	[100, 105)	[105, 110)	total
x_i (marca de clase)	82,5	87,5	92,5	97,5	102,5	107,5	-
frec. absoluta = n_i	2	7	14	16	9	2	50 = N
frec. abs. acumulada	2	9	23	39	48	50	-
$x_i\, n_i$	165	612,5	1295	1560	922,5	215	4770
$x_i^2 n_i$	13612,5	53593,75	119787,5	152100	94556,25	23112,5	456762,6

Media: $\overline{x} = \frac{\sum x_i\, n_i}{N} = \frac{4770}{50} = 95{,}4$; desviación típica: $s = \sqrt{\frac{\sum x_i^2\, n_i}{N} - \overline{x}^2} = \sqrt{\frac{456762{,}6}{50} - 95{,}4^2} \cong 5{,}83884$

El primer cuartil, Q1, deja el 25% de los datos a la izquierda. Como $25\%\ de\ N = 25\%\ de\ 50 = 12{,}5$, la clase que contiene a Q1 es [90, 95).

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

El primer cuartil es $Q1 = L_i + \frac{c\left(\frac{N}{4} - N_{i-1}\right)}{n_i} = 90 + \frac{5(12{,}5 - 9)}{14} = 91{,}25$

Siendo L_i = límite inferior de la clase que contiene a Q1, 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 Q1

El percentil 95, P_{95}, deja el 95% de los datos a la izquierda. Como $95\%\ de\ N = 95\%\ de\ 50 = 47{,}5$, la clase que contiene al percentil 95 es [100, 105).

Aplicando una expresión análoga a la del primer cuartil para datos agrupados en intervalos, se tiene:

El percentil 95 es $P_{95} = L_i + \frac{c\left(\frac{95N}{100} - N_{i-1}\right)}{n_i} = 100 + \frac{5(47{,}5 - 39)}{9} \cong 104{,}72$

Luego, 110 – 104,72 ≅ 5,28%, es el porcentaje de bombillas que duran 95 horas o más

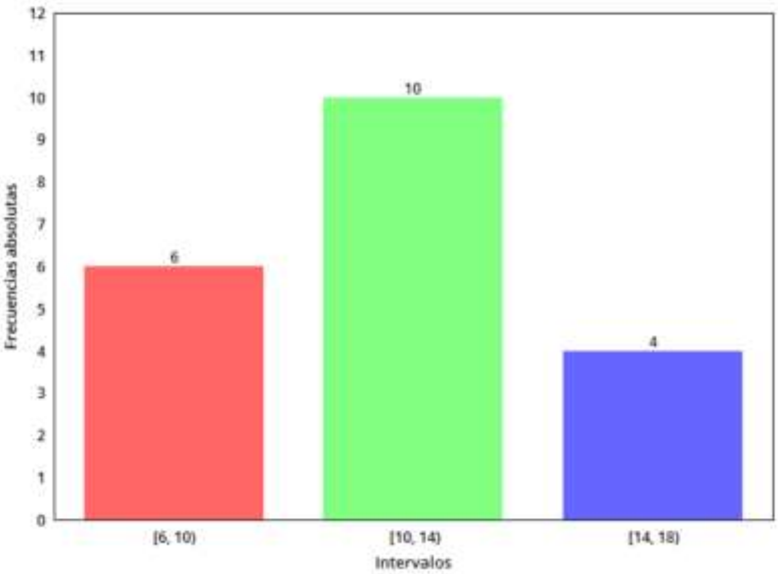
---

4.- Los valores agrupados en intervalos de una determinada variable estadística X vienen reflejados en la tabla siguiente:

Intervalos	Frec. Absol.
[6, 10)	6
[10, 14)	10
[14, 18)	4

a) Construya el histograma.

Resolución



b) Determine la mediana.
c) Calcule el coeficiente de variación.

Resolución

X	[6, 10)	[10, 14)	[14, 18)	total
x_i (marca de clase)	8	12	16	-
frecuencia absoluta = n_i	6	10	4	20 = N
frecuencia absoluta acumulada	6	16	20	-
$x_i\ n_i$	48	1120	64	232
$x_i^2 n_i$	384	1440	1024	2848

La mediana Me deja la mitad de los datos a la izquierda. Como $\frac{20}{2} = 10$, el intervalo mediano es [10, 14).

La mediana es $Me = L_i + \frac{c\left(\frac{N}{2} - N_{i-1}\right)}{n_i}$

Siendo L_i = límite inferior de la clase mediana, c = amplitud del intervalo, N = número total de datos.

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

n_i = frecuencia absoluta de la clase mediana. Luego, $Me = 10 + \frac{4(10-8)}{10} = 10{,}8$

El coeficiente de variación de Pearson es el cociente entre la desviación típica y la media aritmética

El nº total de datos es N = 20. La media es $\overline{x} = \frac{\sum x_i\, n_i}{N} = \frac{232}{20} = 11{,}6$

La desviación típica es $s = \sqrt{\frac{\sum x_i^2\, n_i}{N} - \overline{x}^2} = \sqrt{\frac{2848}{20} - 11{,}6^2} = 2{,}8$

$$= \frac{s}{\overline{x}} = \frac{2{,}8}{11{,}6} \cong 0{,}2414$$

5.- La distribución de frecuencias absolutas f_i de las notas x_i obtenidas por 100 alumnos es:

x_i	0	1	2	3	4	5	6	7
f_i	5	16	23	21	17	9	6	3

Calcule la moda, la media y la desviación típica de estos datos.

Resolución

x_i	0	1	2	3	4	5	6	7	total
f_i	5	16	23	21	17	9	6	3	100 = N
$x_i f_i$	0	16	46	63	68	45	36	21	295
$x_i^2 f_i$	0	16	92	189	272	225	216	147	1157

moda: 2 ; media: $\overline{x} = \frac{\sum x_i f_i}{N} = \frac{295}{100} = 2{,}95$

desviación típica: $s = \sqrt{\frac{\sum x_i^2 f_i}{N} - \overline{x}^2} = \sqrt{\frac{1157}{100} - 2{,}95^2} \cong 1{,}693369$

6.- Se sabe que el 10% de las piezas producidas por una máquina son defectuosas. Se elige al azar un lote de 5 de esas piezas y se considera la variable aleatoria X "número de piezas defectuosas en el lote"

a) ¿Cuál es la ley de probabilidad de la variable X?

Resolución

Como realizamos 5 veces el experimento aleatorio de elegir al azar una pieza y la probabilidad de que la pieza sea defectuosa es 10% = 0,1, entonces $X \rightarrow B(5\,;\,0{,}1)$

Luego, la ley de probabilidad es $p_k = p(X = k) = \binom{5}{k} 0{,}1^k 0{,}9^{5-k}$, con k = 0, 1, 2, 3, 4, 5.

b) ¿Cuál es la probabilidad de que al menos 4 piezas no sean defectuosas?

Resolución

Si Y "número de piezas no defectuosas en el lote", entonces $Y \rightarrow B(5\,;\,0{,}9)$, la probabilidad que se pide es

$p(Y \geq 4) = p(Y = 4) + p(Y = 5) = \binom{5}{4} 0{,}9^4 0{,}1^1 + \binom{5}{5} 0{,}9^5 0{,}1^0 = 5.0{,}9^4.\,0{,}1 + 1.0{,}9^4 \cong 0{,}9185 = 91{,}85\%$

7.- En una ciudad de 50000 habitantes se estudia el número de cigarrillos diarios que se fuman. Sabiendo que 20000 habitantes no fuman y que los demás fuman por término medio 15 cigarrillos diarios, halle el número medio de cigarrillos que se fuman en dicha ciudad.

Resolución

El número de personas que fuman en la ciudad es de 50000 – 20000 = 30000 y como fuman por término medio 15 cigarrillos diarios cada uno, el número medio de cigarrillos que se fuman en dicha ciudad es 15·30000 = 450000 cigarrillos.

8.- La cantidad de litros de lluvia que cae en una localidad durante el otoño, es una variable aleatoria que sigue una distribución normal de media $\mu = 100$ l y varianza $\sigma^2 = 25$ l^2.

a) Halle la probabilidad de que la cantidad de litros de lluvia esté en el intervalo $(\mu - 2\sigma, \mu + 2\sigma)$.

Resolución

X = cantidad de litros $\rightarrow N(100, \sqrt{25}) = N(100, 5) \;\Rightarrow\; Z = \frac{X-100}{5} \rightarrow N(0,1)$.

$(\mu - 2\sigma, \mu + 2\sigma) = (100 - 2.5, 100 + 2.5) = (90, 110)$

Nos piden $p(90 < X < 110)$. Restamos 100 y dividimos entre 5 en los tres miembros:

$$p\left(\frac{90-100}{5} < \frac{X-100}{5} < \frac{110-100}{5}\right) = p(-2 < Z < 2) = p(Z < 2) - p(Z > 2) =$$

$$= p(Z < 2) - 1 + p(Z < 2) = 2p(Z < 2) - 1 = 2\,.\,0{,}9772 - 1 = 0{,}9544 = 95{,}44\%.$$

b) Halle el primer cuartil de esta variable.

Resolución

Nos piden el primer cuartil, k, que es el valor que cumple $p(X < k) = 0{,}25$.

$0{,}25 = p(X < k) = p\left(\frac{X-100}{5} < \frac{k-100}{5}\right) = p\left(Z < \frac{k-100}{5}\right) \Rightarrow p\left(Z > \frac{k-100}{5} = 0{,}75\right)$. Usando la tabla de la N(0, 1) en sentido inverso con valores de z negativos, obtenemos por interpolación $\frac{k-100}{5} = -0{,}675$.

Despejando, el primer cuartil es k = 100 – 5 . 0,675 = 96,625

9.- La información estadística de una muestra de tamaño 6 sobre la relación existente entre el número de horas de estudio dedicado a la preparación de un examen y la puntuación obtenida en dicho examen, se refleja en la tabla siguiente:

Horas de estudio (X)	15	20	25	32	18	34
Puntuación (Y)	50	60	70	92	70	84

a) Halla la ecuación de la recta de regresión de Y sobre X.
b) Obtenga el coeficiente de correlación lineal entre X e Y, e interprételo.

Resolución

El nº de datos es N = 6. Elaboramos la tabla de frecuencias adecuada:

x_i	15	20	25	32	18	34	$\to 144$
y_i	50	60	70	92	70	84	$\to 426$
x_i^2	225	400	625	1024	324	1156	$\to 3754$
y_i^2	2500	3600	4900	8464	4900	7056	$\to 31420$
$x_i y_i$	750	1200	1750	2944	1260	2856	$\to 10760$

$\overline{x} = \frac{\sum x_i}{N} = \frac{144}{6} = 24$; $\overline{y} = \frac{\sum y_i}{N} = \frac{426}{6} = 71$; $s_x^2 = \frac{\sum x_i^2}{N} - \overline{x}^2 = \frac{3754}{6} - 24^2 \cong 49{,}666\ldots$

$s_y^2 = \frac{\sum y_i^2}{N} - \overline{y}^2 = \frac{31420}{6} - 71^2 \cong 195{,}666\ldots$; $s_{xy} = \frac{\sum x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{10760}{6} - 24\,.\,71 \cong 89{,}333\ldots$

El coeficiente de correlación lineal es $r = \frac{s_{xy}}{s_x\, s_y} \cong \frac{89{,}333\ldots}{\sqrt{49{,}666\ldots}\,.\,\sqrt{195{,}666\ldots}} \cong 0{,}9062$, 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 muy fuerte pues está mucho más próximo a 1 que a 0.

La recta de regresión (de Y sobre X) es $r_{YX}: y - \overline{y} = \frac{s_{xy}}{s_x^2}(x - \overline{x}) \Rightarrow r_{YX}: y - 71 = \frac{89{,}333\ldots}{49{,}666\ldots}(x - 24)$

Operando y simplificando, la recta de regresión que se pide es $r_{YX}: y = 1{,}7987x + 27{,}8322$

10.- En un pueblo, donde residen 600 hombres y 400 mujeres, se eligen 5 personas al azar. ¿Cuál es la probabilidad de que haya, exactamente 2 hombres entre ellas?

Resolución

Como realizamos 5 veces el experimento aleatorio de elegir al azar una persona y la probabilidad de que la persona sea hombre es $\frac{600}{1000} = 0{,}6$, entonces X = "nº de hombres que se eligen" → B(5 ; 0,6)

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

La probabilidad que se pide $p(X = 2) = \binom{5}{2} 0{,}6^2 0{,}4^3 = 10 . 0{,}6^2 . 0{,}4^3 = 0{,}2304 = 23{,}04\%$

11.- De la distribución bidimensional (X, Y), se toma una muestra de 5 elementos cuyos valores son los que se reflejan en la tabla:

X	4	6	8	10	A
Y	12	14	18	19	B

Se sabe que $\overline{x} = 8$ y que la recta de regresión de Y sobre X es y = 1,25x + 7.
a) Determine los valores de A y B.

Resolución

El nº de datos es N = 5 ; $8 = \overline{x} = \frac{\sum x_i}{N} = \frac{28 + A}{5} \Rightarrow A = 12$

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

Como $r_{YX}: y = 1{,}25x + 7$, entonces $\frac{s_{xy}}{s_x^2} = 1{,}25$ y $7 = \overline{y} - \frac{\overline{x}\, s_{xy}}{s_x^2} = \overline{y} - 8\,.\,1{,}25 \Rightarrow \overline{y} = 17$

Luego, $17 = \overline{y} = \frac{\sum y_i}{N} = \frac{63 + B}{5} \Rightarrow B = 22$

b) Calcule la covarianza de las variables X e Y.

Resolución

El nº de datos es N = 5. Elaboramos la tabla de frecuencias adecuada:

x_i	4	6	8	10	12	→ 40
y_i	12	14	18	19	22	→ 85
$x_i y_i$	48	84	144	190	264	→ 730

$\overline{x} = \frac{\sum x_i}{N} = \frac{40}{5} = 8$; $\overline{y} = \frac{\sum y_i}{N} = \frac{85}{5} = 17$. Covarianza: $s_{xy} = \frac{\sum x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{730}{5} - 8\,.\,17 = 10$

12.- Sabiendo que el peso de una cierta variedad de naranjas sigue una distribución normal de media 200 gramos y desviación típica 20 gramos.
a) Calcule la probabilidad de que una naranja de esta variedad elegida al azar tenga un peso comprendido entre 175 y 225 gramos.

Resolución

X = peso → $N(200, 20) \Rightarrow Z = \frac{X - 200}{20} \rightarrow N(0, 1)$. Nos piden $p(175 < X < 225) =$

$= p\left(\frac{175 - 200}{20} < \frac{X - 200}{20} < \frac{225 - 200}{20}\right) = p(-1{,}25 < Z < 1{,}25) = p(Z < 1{,}25) - p(Z > 1{,}25) =$

$= p(Z < 1{,}25) - 1 + p(Z < 1{,}25) = 2p(Z < 1{,}25) - 1 = 2\,.\,0{,}8944 - 1 = 0{,}7888 = 78{,}88\%$.

b) Determine a partir de qué peso se encuentra el 20% de las naranjas de mayor peso de esa variedad.

Resolución

Las naranjas de mayor peso son las que pesan más de la media, 200 g.

$p(X > 200) = p\left(\frac{X - 200}{20} > \frac{200 - 200}{20}\right) = p(Z > 0) = 1 - p(Z \leq 0) = 1 - 0{,}5 = 0{,}5$

20% de 0,5 = 0,1. Si k es el nº mínimo peso pedido,

$0{,}1 = p(X > k) = p\left(\frac{X - 200}{20} > \frac{k - 200}{20}\right) = p\left(Z > \frac{k - 200}{20}\right) = 1 - p\left(Z \leq \frac{k - 200}{20}\right) \Rightarrow p\left(Z \leq \frac{k - 200}{20}\right) = 0{,}9$

Usando la tabla de la N(0, 1) en sentido inverso obtenemos que $\frac{k - 200}{20} \cong 1{,}28$. Despejando, k ≅ 225,6.

Es decir, a partir de un peso aproximado de 225,6 g se encuentra el 20% de las naranjas de mayor peso de esa variedad.

13.- En una muestra de 10 individuos se han observado dos características, la estatura en metros (X), y el peso en kg (Y), obteniéndose que la media de las estaturas es 1,5 m y su varianza 0,81 m^2, mientras que la media de los pesos es 60 kg y su varianza es 4 kg^2. Además, el coeficiente de correlación lineal entre ambas es $\frac{8}{9}$.

a) Determine la covarianza de X e Y.

b) Suponiendo que entre ambas variables existe una relación de tipo lineal, ¿qué peso se podrá predecir para una persona de 1,77 m de estatura?

Resolución

Nos dicen: $\overline{x} = 1{,}5$; $s_x^2 = 0{,}81$; $\overline{y} = 60$; $s_y^2 = 4$; coeficiente de correlación lineal: $r = \frac{s_{xy}}{s_x\, s_y} = \frac{8}{9}$

Despejando, la covarianza sería $s_{xy} = \frac{8\, s_x\, s_y}{9} = \frac{8\,\sqrt{0{,}81}\,\sqrt{4}}{9} = \frac{8\ .\ 0{,}9\ .\ 2}{9} = 1{,}6$

La recta de regresión (de Y sobre X) es $r_{YX}: y - \overline{y} = \frac{s_{xy}}{s_x^2}(x - \overline{x}) \Rightarrow r_{YX}: y - 60 = \frac{1{,}6}{0{,}81}(x - 1{,}5)$

Operando y simplificando, la recta de regresión es $r_{YX}: y = 1{,}97531x + 57{,}03704$

Para una estatura de 1,77 m, sustituimos y el peso sería $y = 1{,}97531\,.\,1{,}77 + 57{,}03704 \cong 60{,}53$ kg

14.- Sea X una variable aleatoria que toma los valores 0, 1, 2 y 3 con probabilidad no nula.
Sabiendo que: p(X ≤ 0,5) = 0,4, p(X ≤ 2) = 0,7 y E(X) = 1,4. (Nota: E(X) indica media de la variable X)

a) Obtenga la función de probabilidad de X.

b) Calcule la varianza de X.

Resolución

Sea $p_i = p(X = x_i)$, i = 0, 1, 2, 3. Donde, $x_0 = 0, x_1 = 1, x_2 = 2\ , x_3 = 3$

$0{,}4 = p(X \leq 0{,}5) = p(X = 0) = p_0$

$0{,}7 = p(X \leq 2) = p(X = 0) + p(X = 1) + p(X = 2) = 0{,}4 + p_1 + p_2$. Luego, $p_1 + p_2 = 0{,}3$

Por ser X una función de probabilidad, $p_0 + p_1 + p_2 + p_3 = 1$. Luego, $p_3 = 1 - 0{,}7 = 0{,}3$

La media es 1,4 $\Rightarrow 1{,}4 = \mu = \sum x_i\, p_i = 0.\,p_0 + 1.\,p_1 + 2.\,p_2 + 3.\,p_3 = p_1 + 2p_2 + 3\,.\,0{,}3 \Rightarrow p_1 + 2p_2 = 0{,}5$

Restando las ecuaciones, $p_2 = 0{,}2$; $p_1 = 0{,}3 - p_2 = 0{,}3 - 0{,}2 = 0{,}1$

La función de probabilidad de X es

x_i	$p_i = p(X = x_i)$	$x_i p_i$	$x_i^2 p_i$
0	0,4	0	0
1	0,1	0,1	0,1
2	0,2	0,4	0,8
3	0,3	0,9	2,7
total	1	$\mu = \sum x_i\, p_i = 1{,}4$	$\sum x_i^2 p_i = 3{,}6$

La varianza de X es $\sigma^2 = \sum x_i^2 p_i - \mu^2 = 3{,}6 - 1{,}4^2 = 1{,}64$

15.- De una distribución bidimensional (X, Y) se sabe que la recta de regresión de Y sobre X es $y = \frac{2x}{13} + \frac{53}{13}$ y la de X sobre Y es $x = \frac{32y}{13} - \frac{82}{13}$.

a) Halle el coeficiente de correlación lineal entre X e Y, e interprételo.

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 $r_{YX}: y = \frac{2x}{13} + \frac{53}{13}$, entonces $\frac{s_{xy}}{s_x^2} = \frac{2}{13}$

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

Como $r_{XY}: x = \frac{32y}{13} - \frac{82}{13}$, entonces $\frac{s_{xy}}{s_y^2} = \frac{32}{13}$

Multiplicando, $\frac{s_{xy}^2}{s_x^2\, s_y^2} = \frac{2}{13}\frac{32}{13} = \frac{64}{169}$. Luego, el coeficiente de correlación es $r = \frac{s_{xy}}{s_x\, s_y} = \sqrt{\frac{64}{169}} = \frac{8}{13}$

b) Calcule la media de la variable X y la de la variable Y.

Resolución

Sabemos que las rectas de regresión se cortan en el punto $(\overline{x},\ \overline{y})$. Resolvemos el sistema $\begin{cases} y = \frac{2x}{13} + \frac{53}{13} \\ x = \frac{32y}{13} - \frac{82}{13} \end{cases}$

Sustituyendo $x = \frac{32y}{13} - \frac{82}{13}$ en la 1ª ecuación, $y = \frac{2}{13}\left(\frac{32y}{13} - \frac{82}{13}\right) + \frac{53}{13} = \frac{64y}{169} - \frac{164}{169} + \frac{53}{13} = \frac{64y}{169} + \frac{525}{169}$.

Multiplicando por 169 obtenemos: $169y = 64y + 525 \Rightarrow 105y = 525 \Rightarrow y = \frac{525}{105} = 5$

Sustituyendo en $x = \frac{32y}{13} - \frac{82}{13}$ se obtiene $x = \frac{32\,.\,5}{13} - \frac{82}{13} = 6$. Por tanto, $\overline{x} = 6$ e $\overline{y} = 5$

16.- En un concurso de pintura infantil, dos jurados han valorado de 0 a 5 los trabajos de 10 finalistas, como refleja la tabla siguiente:

Concursante	A	B	C	D	E	F	G	H	I	J
Notas jurado 1	2	1	1	4	3	2	2	2	2	1
Notas jurado 2	4	2	1	5	5	3	4	3	2	1

a) Calcule la moda de las notas del jurado 1.

Resolución

notas jurado 1	1	2	3	4
frecuencia absoluta	3	5	1	1

La moda es la nota de mayor frecuencia absoluta. Luego, la moda es 2

b) Calcule la mediana de las notas del jurado 2.

Resolución

Sabemos que la mediana (Me) de una serie estadística es el valor que se encuentra en el centro de la misma, cuando esta ha sido ordenada en sentido creciente o decreciente. La mediana es única.
Si el número de datos es par hay dos términos centrales y se suele tomar como mediana la semisuma de los dos valores centrales.
Ordenados de menor a mayor tenemos: 1 (C), 1 (J), 2 (B), 2 (H), 3 (F), 3 (H), 4 (A), 4 (G), 5 (D) y 56 (E).
En nuestro caso los datos centrales corresponden a la nota 3 es decir a los de los concursantes F y H. Luego, la mediana es 3

---

c) Calcule el coeficiente de correlación lineal entre las notas de ambos jurados e interprételo.

Resolución

El nº de datos es N = 10. Elaboramos la tabla de frecuencias adecuada:

x_i	2	1	1	4	3	2	2	2	2	1	→ 20
y_i	4	2	1	5	5	3	4	3	2	1	→ 30
x_i^2	4	1	1	16	9	4	4	4	4	1	→ 48
y_i^2	16	4	1	25	25	9	16	9	4	1	→ 110
$x_i y_i$	8	2	1	20	15	6	8	6	4	1	→ 71

$\overline{x} = \frac{\sum x_i}{N} = \frac{20}{10} = 2$; varianza de X $= s_x^2 = \frac{\sum x_i^2}{N} - \overline{x}^2 = \frac{48}{10} - 2^2 = 0{,}8$

$\overline{y} = \frac{\sum y_i}{N} = \frac{30}{10} = 3$; varianza de Y $= s_y^2 = \frac{\sum y_i^2}{N} - \overline{y}^2 = \frac{110}{10} - 3^2 = 2$

$s_{xy} = \frac{\sum x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{71}{10} - 2\,.\,3 = 1{,}1$

El coeficiente de correlación es $r = \frac{s_{xy}}{s_x\, s_y} = \frac{1{,}1}{\sqrt{0{,}8}\,.\,\sqrt{2}} \cong 0{,}86963$

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.

17.- Una variable aleatoria X toma los valores {1, 2, 3, 4, 5}, con las siguientes probabilidades:
$p(X=1) = \frac{1}{6}, p(X=2) = \frac{1}{4}, p(X=3) = \frac{1}{6}, p(X=4) = \frac{1}{4}$ y $p(X=5) = \frac{1}{6}$

a) ¿Qué tipo de variable aleatoria es?

b) Calcule la media de X.

Resolución

Es una variable aleatoria discreta porque toma valores aislados: {1, 2, 3, 4, 5}

La función de probabilidad de X es

x_i	$p_i = p(X = x_i)$	$x_i p_i$
1	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{4}$	$\frac{1}{2}$
3	$\frac{1}{6}$	$\frac{1}{2}$
4	$\frac{1}{4}$	1
5	$\frac{1}{6}$	$\frac{5}{6}$
total	1	1,4

La media de X es $\mu = \sum x_i\, p_i \; = 3$

c) Represente gráficamente la función de distribución de X.

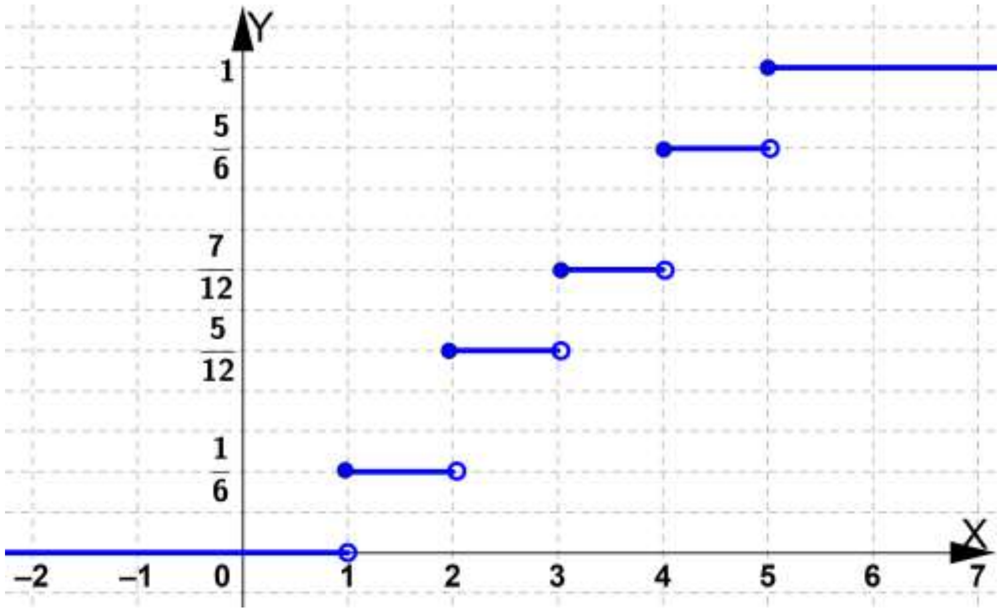
Resolución

Sabemos que la función de distribución de una variable aleatoria discreta X, que se escribe $F(x_i)$, es la suma de todas las probabilidades de la variable aleatoria X hasta el valor x_i, es decir

$F(x_i) = p(X \le x_i) = p(X = x_1) + p(X = x_2) + p(X = x_3) + \ldots + p(X = x_i) = p_1 + p_2 + p_3 + \ldots + p_i$.

Luego, la función de distribución es $F(x) = \begin{cases} 0, \ si \ \ x < 1 \\ \frac{1}{6}, \ si \ \ 1 \leq x < 2 \\ \frac{5}{12}, \ si \ \ 2 \leq x < 3 \\ \frac{7}{12}, \ si \ \ 3 \leq x < 4 \\ \frac{5}{6}, \ si \ \ 4 \leq x < 5 \\ 1, \ si \ \ 5 \leq x \end{cases}$

Su gráfica es

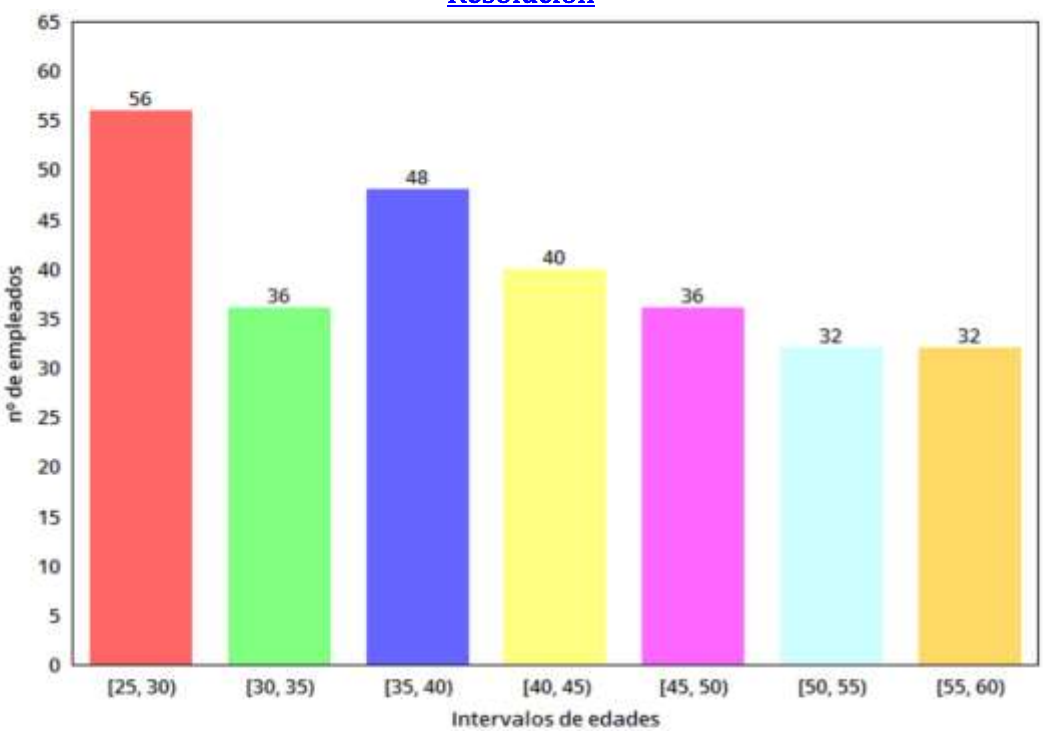


18.- Las edades de 280 empleados de una empresa se han agrupado en los siete intervalos que se indican en la siguiente tabla, en la cual, para cada intervalo, se ha anotado su frecuencia absoluta:

Intervalos por edades	[25, 30)	[30, 35)	[35, 40)	[40, 45)	[45, 50)	[50, 55)	[55, 60)
Nº de empleados	56	36	48	40	36	32	32

a) Dibuje el histograma de la distribución de edades que indica esta tabla.

Resolución



---

b) Calcule la media y la mediana.

Resolución

X	[25, 30)	[30, 35)	[35, 40)	[40, 45)	[45, 50)	[50, 55)	[55, 60)	total
x_i (marca de clase)	27,5	32,5	37,5	42,5	47,5	52,5	57,5	-
frecuencia absoluta = n_i	56	36	48	40	36	32	32	280 = N
frecuencia absoluta acumulada	56	92	140	180	216	248	280	-
$x_i\, n_i$	1540	1170	1800	1700	1710	1640	1840	11440

El nº total de datos es N = 280. La media es $\overline{x} = \frac{\sum x_i\, n_i}{N} = \frac{11440}{280} = 40{,}857$

La mediana Me deja la mitad de los datos a la izquierda. Como $\frac{280}{2} = 140$, el intervalo mediano

es [35, 40). La mediana es $Me = L_i + \frac{c\left(\frac{N}{2} - N_{i-1}\right)}{n_i}$

Siendo L_i = límite inferior de la clase mediana, c = amplitud del intervalo, N = número total de datos.

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

n_i = frecuencia absoluta de la clase mediana. Luego, $Me = 35 + \frac{5(140 - 92)}{48} = 40$

19.- Los estudios realizados sobre una población de recién nacidos, se ha determinado que la talla se distribuye según una ley distribución normal de media 52 cm y desviación típica 2 cm.

a) Halle la probabilidad de que un recién nacido tenga una talla superior a 56 cm.

Resolución

X = talla $\rightarrow N(52, 2) \quad \Rightarrow \quad Z = \frac{X - 52}{2} \rightarrow N(0, 1).$

Piden $p(X > 56) = p\left(\frac{X - 52}{2} > \frac{56 - 52}{2}\right) = p(Z > 1) = 1 - p(Z \leq 2) = 1 - 0{,}9772 = 0{,}0228 = 2{,}28\%$

b) Halle la probabilidad de que un recién nacido tenga una talla comprendida entre 50 cm y 53 cm.

Resolución

Se pide $p(50 < X < 53) = p\left(\frac{50 - 52}{2} < \frac{X - 52}{2} < \frac{53 - 52}{2}\right) = p(-1 < Z < 0{,}5) = p(Z < 0{,}5) - p(Z > 1) =$

$= p(Z < 0{,}5) - 1 + p(Z \leq 1) = 0{,}6915 - 1 + 0{,}8413 = 0{,}5328 = 53{,}28\%.$

20.- En determinada población las temperaturas mínimas X y máximas Y alcanzadas a lo largo de los seis primeros meses del año fueron las siguientes:

X	7,6	7,1	8,7	13,5	11,8	14,3
Y	16,4	17,2	17,8	22,8	20,7	26,9

a) Decida cuál de las dos distribuciones marginales es más dispersa, utilizando para ello los coeficientes de variación.
b) Determine el coeficiente de correlación lineal y la recta de regresión Y sobre X. Comente la fiabilidad de las predicciones basadas en esa recta.

Resolución

El nº de datos es N = 6. Elaboramos la tabla de frecuencias adecuada:

temperaturas mínimas = x_i	7,6	7,1	8,7	13,5	11,8	14,3	$\to$ 63
temperaturas máximas = y_i	16,4	17,2	17,8	22,8	20,7	26,9	$\to$ 121,8
x_i^2	57,76	50,41	75,69	182,25	139,24	204,49	$\to$ 709,84
y_i^2	268,96	295,84	316,84	519,84	428,49	723,61	$\to$ 2553,58
$x_i y_i$	124,64	122,12	154,86	307,8	244,26	384,67	$\to$ 1338,35

$\overline{x} = \frac{\sum x_i}{N} = \frac{63}{6} = 10{,}5$; varianza de X $= s_x^2 = \frac{\sum x_i^2}{N} - \overline{x}^2 = \frac{709{,}84}{6} - 10{,}5^2 \cong 8{,}05667$

$\overline{y} = \frac{\sum y_i}{N} = \frac{121{,}8}{6} = 20{,}3$; varianza de Y $= s_y^2 = \frac{\sum y_i^2}{N} - \overline{y}^2 = \frac{2553{,}58}{6} - 20{,}3^2 = 13{,}50667$

a) El coeficiente de variación de Pearson es el cociente entre la desviación típica y la media aritmética.
Los coeficientes de variación son $\rho_X = \frac{s_x}{\overline{x}} = \frac{\sqrt{8{,}05667}}{10{,}5} \cong 0{,}27$ y $\rho_Y = \frac{s_y}{\overline{y}} = \frac{\sqrt{13{,}50667}}{20{,}3} \cong 0{,}18$
Como vemos el coeficiente de variación de la variable Y es más pequeño. Luego, las temperaturas máximas están menos dispersas.

b) $s_{xy} = \frac{\sum x_i y_i}{N} - \overline{x}\,\overline{y} = \frac{1338{,}35}{6} - 10{,}5 \,.\, 20{,}3 = 9{,}90833$

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

$r_{YX}: y - 20{,}3 = \frac{9{,}90833}{8{,}05667}(x - 10{,}5) \Rightarrow r_{YX}: \mathrm{y} = 1{,}22983\mathrm{x} + 7{,}38678$

Coeficiente de correlación: $r = \frac{s_{xy}}{s_x\, s_y} = \frac{9{,}90833}{\sqrt{8{,}05667}\,.\,\sqrt{13{,}50667}} \cong 0{,}94984$

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