

Hoja de problemas sobre Estadística

CURSO TEMA

1ºBach Estadística 06

WWW.DANIPARTAL.NET

Colegio Marista "La Inmaculada" de Granada

INFORMACIÓN GENERAL

Hoja de problemas sobre Estadística.

Recuerda que el web del IES Ayala posee multitud de exámenes resueltos de Selectividad de cursos anteriores:

<https://www.iesayala.com/selectividadmatematicas/ficheros/andaluciaccss.html>

PROBLEMA 1

En una clase de 30 alumnos, 2 alumnos han sacado una calificación de 2 en un proyecto. 4 alumnos un 3. 6 alumnos un 4. 4 alumnos un 5. 5 alumnos un 6. 3 alumnos un 7. 3 alumnos un 8. 2 alumnos un 9. Y un solo alumno un 10.

Obtener la media de la nota de la clase y su desviación típica (expresar los valores con dos cifras decimales redondeadas).

Si comparamos esta aula con otra clase donde el coeficiente de variación ha resultado del 30%, ¿cuál de las dos clases presentan menor dispersión de los valores respecto de su media?

Completamos la siguiente tabla estadística.

Nota x_i	Frecuencia absoluta n_i	$n_i \cdot x_i$	$n_i \cdot (x_i - \mu)^2$
2	2	4	23,53
3	4	12	23,62
4	6	24	12,27
5	4	20	0,74
6	5	30	1,62
7	3	21	7,40
8	3	24	19,81
9	2	18	25,49
10	1	10	20,88
Acumulado	N=30	Suma = 163	Suma = 135,37

Donde la media viene dada por la expresión:

$$\mu = \sum_{i=1}^n \frac{n_i \cdot x_i}{N} = 5,43$$

La varianza es igual a:

$$\sigma^2 = \sum_{i=1}^n \frac{n_i \cdot (x_i - \mu)^2}{N} = 4,51$$

La desviación típica es:

$$\sigma = \sqrt{4,51} = 2,12$$

Y el coeficiente de variación se calcula como:

$$CV = \frac{\sigma}{\mu} \cdot 100\% = 39,04\%$$

Un CV del 39,04% es superior al CV del 30% de la otra clase. Por lo que los valores de la otra clase están menos dispersos respecto de su media.

PROBLEMA 2

Tiramos cinco veces una moneda. ¿Cuál es la probabilidad de sacar 2 caras y 3 cruces?

Podemos plantearlo como un problema de distribución binomial de tamaño 5 (al repetirse cinco veces el experimento aleatorio). Si definimos el éxito como sacar cara y el fracaso como sacar cruz, la probabilidad de éxito en una tirada será $p = 1/2$ (hay un 50% de sacar cara al lanzar una moneda). Por lo tanto:

$B(5, 1/2)$

Donde la variable aleatoria X (sacar cara) puede ser igual a 0, 1, 2, 3, 4 o 5.

Para responder al enunciado necesitamos calcular la probabilidad de que el número de éxitos (número de caras totales) sea igual a 2. Por lo tanto:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(X = 2) = \binom{5}{2} (0,5)^2 (1-0,5)^3 = \frac{5!}{2! (5-2)!} (0,5)^2 (0,5)^3 = \frac{5 \cdot 4 \cdot 3 \cdot 2}{2 \cdot 3 \cdot 2} (0,5)^5 = 10 \cdot 0,03125 = 0,3125$$

Es decir, en tanto por ciento, la probabilidad de sacar 2 caras es del 31,25%.

Puedes calcular rápidamente la probabilidad en la distribución binomial con la web:

<https://calculadorasonline.com/distribucion-binomial-probabilidad-binomial>

PROBLEMA 3

Una oficina consta de 10 teléfonos. La probabilidad de que a cierta hora del día un teléfono esté comunicando es del 20%. ¿Cuál es la probabilidad de que a esa hora se marquen los 10 teléfonos y no se encuentren más de dos comunicando?

El enunciado presenta un éxito (estar comunicando) con una probabilidad de éxito del 20% (o lo que es lo mismo, $p=0,2$).

El tamaño de la repetición del experimento es $n=10$. Por lo que planteamos una binomial:

$B(10, 0,2)$

Si no puede haber más de dos teléfonos comunicando, significa que la variable aleatoria X puede tomar los valores 0, 1 y 2. Por lo que debemos calcular la probabilidad de éxito para esos tres valores.

Como el tamaño de la muestra no supera el valor de 30, no estamos en condiciones de aproximar la binomial a la normal.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(X = 0) = \binom{10}{0} (0,2)^0 (1-0,2)^{10} = \frac{10!}{0! (10-0)!} (0,2)^0 (0,8)^{10} = 0,1074$$

$$P(X = 1) = \binom{10}{1} (0,2)^1 (1-0,2)^9 = \frac{10!}{1! (10-1)!} (0,2)^1 (0,8)^9 = 0,2684$$

$$P(X = 2) = \binom{10}{2} (0,2)^2 (1-0,2)^8 = \frac{10!}{2! (10-2)!} (0,2)^2 (0,8)^8 = 0,3020$$

La suma de estas tres probabilidades nos da la solución del ejercicio: $0,1074 + 0,2684 + 0,3020 = 0,6778$

Una probabilidad del 67,78%.

PROBLEMA 4

Una fábrica de bolígrafos establece en un 5% el porcentaje de bolígrafos con punta defectuosa. ¿Cuál es la probabilidad de que en una caja de 8 bolígrafos haya más de 7 defectuosos? ¿Cuál es la media de bolígrafos defectuosos en una caja escogida al azar?

Estamos ante un nuevo experimento binomial, de tamaño de muestra $n=8$ y probabilidad de éxito (que la punta sea defectuosa) de $p=0,05$.

$B(8, 0,05)$

Plantear la probabilidad de que haya más de 7 bolígrafos defectuosos será igual a la probabilidad de que los 8 bolígrafos estén defectuosos. Es decir:

$$P(X = 8)$$

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(X = 8) = \binom{8}{8} (0,05)^8 (1 - 0,05)^0 = \frac{8!}{8! (8-8)!} (0,05)^8 (0,95)^0 = 0,00000 \dots \approx 0$$

Es decir, la probabilidad de que haya más de 7 bolígrafos defectuosos prácticamente tiende a 0%.

La media en la distribución normal se obtiene como el producto del tamaño de la muestra por la probabilidad de éxito (recuerda que, en este caso, el éxito ocurre cuando un bolígrafo es defectuoso).

$$\mu = n \cdot p$$

$$\mu = 8 \cdot 0,05 = 0,4$$

Como estamos ante una variable discreta, aproximando a la baja, la media de bolígrafos defectuosos que encontraremos en cajas de tamaño 8 es un valor que se aproxima a 0.

PROBLEMA 5

Si la probabilidad de que una persona use gafas es del 10%, hallar la media de la distribución de personas que usen gafas al elegir 20 personas al azar y la desviación típica de la distribución.

Nuevo ejercicio binomial, de tamaño $n=20$ y probabilidad de éxito $p=0,1$ (el éxito se considera tener gafas).

La media en la binomial se obtiene como:

$$\mu = n \cdot p$$

$$\mu = 20 \cdot 0,1 = 2$$

Mientras que la desviación típica en la binomial se obtiene como:

$$\sigma = \sqrt{n \cdot p \cdot (1 - p)}$$

$$\sigma = \sqrt{20 \cdot 0,1 \cdot (1 - 0,1)} = 1,34$$

PROBLEMA 6

En una distribución normal tipificada, ¿qué intervalo simétrico $[-k, k]$ alrededor de la media acumula el 68,26% de las observaciones?

Recordamos que en la distribución normal tipificada la media es igual a 0. Por lo tanto, si buscamos un intervalo simétrico alrededor de la media que englobe el 68,26% de las observaciones, significa que en el intervalo mitad $[0, k]$ se acumula el 34,13% (la mitad de 68,26%).

$$P(0 \leq z \leq k) = 0,3413$$

$$P(z \leq k) - P(z \leq 0) = 0,3413$$

Recuerda que, a la izquierda de 0, la gaussiana engloba el 50% de las observaciones. Por lo tanto:

$$P(z \leq k) - 0,5 = 0,3413$$

$$P(z \leq k) = 0,8413$$

En la tabla de probabilidad acumulada debemos aproximar el valor $z=k$ positivo que deja a su izquierda el 84,13% de las observaciones. Y eso ocurre en $z=1$. Por lo tanto, el intervalo solución es $[-1, 1]$.

PROBLEMA 7

En Las calificaciones finales de un colegio (de 0 a 10) siguen una distribución normal de media 6 y desviación típica 1,6. ¿Qué calificación acumula el 45% del alumnado?

Buscamos un valor tipificado que cumpla que acumula a su izquierda un valor igual a 0,45.

La tabla de probabilidades acumuladas que usamos da valores a partir de $z = 0,00$, que acumula el 50% de las observaciones. Por lo que debemos razonar de la siguiente forma: el valor que deja a la izquierda el 45% de las observaciones será un valor negativo, por estar a la izquierda de $z = 0,00$. Vamos a llamar a ese valor $z = -k$.

Si $z = -k$ deja a la izquierda el 45% de las observaciones, su simétrico positivo dejará a la derecha el 45%. Es decir:

$$P(z \geq k) = 0,45$$

Y podemos relacionar la probabilidad superior con la probabilidad inferior con la relación:

$$P(z \geq k) = 1 - P(z \leq k)$$

$$1 - P(z \leq k) = 0,45$$

$$-P(z \leq k) = -0,55$$

$$P(z \leq k) = 0,55$$

Y esta valor sí lo proporciona la tabla. Aproximadamente, $z = 0,13$.

Por lo tanto, el valor inicial que buscamos es el simétrico negativo: $z = -0,13$.

Si usamos la relación de tipificación:

$$Z = \frac{X - \mu}{\sigma}$$

$$-0,13 = \frac{X - 6}{1,6}$$

$$-0,208 = X - 6$$

$$X = 5,79$$

Conclusión: la calificación 5,79 (variable continua) acumula a su izquierda al 45% de los alumnos del instituto.

PROBLEMA 8

En una ciudad una de cada tres familias posee teléfono fijo en casa. Si se eligen al azar 90 familias, calcular la probabilidad de que entre ellas haya más de 30 con teléfono fijo.

En este experimento binomial, una familia puede tener teléfono fijo en casa o puede no tenerlo.

La probabilidad de tener teléfono es $1/3$. Por lo tanto, consideramos:

$$p = 1/3$$

El número de ensayos coincide con el número de familias:

$$n = 90$$

Estamos ante una variable aleatoria binomial $B(90, 1/3)$.

La probabilidad acumulada de tener éxito con más de 30 familias (con teléfono fijo), nos lo da la distribución de probabilidad binomial.

$$P(X \geq 30)$$

¿Problema para calcular esta probabilidad acumulada? Qué tendríamos que calcular $P(X=31) + P(X=32) + \dots + P(X=90)$.

¿Podemos acortar las operaciones? Sí. Aplicando el Teorema central del límite, ya que estamos en condiciones de cumplir:

$$n \cdot p = 90 \cdot \frac{1}{3} = 30 \geq 5$$

$$n \cdot (1 - p) = 90 \cdot \frac{2}{3} = 60 \geq 5$$

De esta forma, la binomial se aproxima a la normal:

$$B(n, p) \sim N(n \cdot p, \sqrt{n \cdot p \cdot (1 - p)})$$

$$B(90, 1/3) \sim N(30, \sqrt{20})$$

Si buscamos que haya más de 30 familias con teléfono, debemos calcular con la distribución normal:

$$P(X \geq 30)$$

Tipificamos la variable:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{30 - 30}{\sqrt{20}} = 0$$

$$P(X \geq 30) = P(Z \geq 0)$$

Al ser la curva normalizada tipificada simétrica respecto al valor $z=0$, la probabilidad acumulada tanto a la derecha como a la izquierda de ese valor es del 50%.

PROBLEMA 9

La probabilidad de que un individuo elegido al azar tenga problemas dermatológicos es de 0,15. Dada una muestra de 50 personas,

a) ¿Cuál es la probabilidad de que ninguna tenga problemas dermatológicos?

b) ¿Cuál es la probabilidad de que al menos cuatro tengan problemas dermatológicos?

a) En este experimento binomial, una persona puede tener problemas de piel o puede no tenerlo.

La probabilidad de tenerlo es $p=0,15$. El tamaño del experimento $n=50$. Nuestra distribución binomial es $B(50, 0,15)$

Si deseamos que, al extraer al azar a una persona, no tenga problemas de piel, deseamos que nuestro experimento aleatorio (que sí tenga problemas de piel) tenga un valor de 0 éxitos.

$$P(X = 0)$$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(X = 0) = \binom{50}{0} 0,15^0 (1 - 0,15)^{50-0} = (0,85)^{50} = 0,0002957$$

Es decir, estamos hablando de una probabilidad del 0,02957%.

b) Que al menos cuatro personas tengan problemas de piel, significa que la pueden tener cuatro, cinco, seis... y hasta 50 personas. Es decir:

$$P(X \geq 4) = P(X = 4) + P(X = 5) + P(X = 6) + \dots + P(X = 50)$$

Agilizamos operaciones aproximando la binomial a la normal, ya que:

$$n \cdot p = 50 \cdot 0,15 = 7,5 \geq 5$$

$$n \cdot (1 - p) = 50 \cdot 0,85 = 42,5 \geq 5$$

De esta forma, la binomial se aproxima a la normal:

$$B(n, p) \sim N(n \cdot p, \sqrt{n \cdot p \cdot (1 - p)})$$

$$B(50, 0,15) \sim N(7,5, 2,52)$$

Si buscamos que haya cuatro o más personas con problemas de piel, debemos calcular con la distribución normal:

$$P(X \geq 4)$$

Tipificamos la variable:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{4 - 7,5}{2,52} = -1,39$$

$$P(X \geq 4) = P(Z \geq -1,39) = P(Z \leq 1,39) = 0,9177$$

La probabilidad es del 91,77%.

PROBLEMA 10

Un modelo de avión tiene capacidad para 260 pasajeros. Sin embargo, la compañía aérea a la que pertenece ha decidido vender más billetes que asientos hay en el avión. La probabilidad de que un pasajero se presente en el aeropuerto el día del vuelo es del 95 %. Si ese día la compañía ha vendido 280 billetes, ¿cuál es la probabilidad de que ese día se presenten 270 pasajeros?

El pasajero puede ir o no ir a recoger el avión. Podemos aplicar variable binomial.

La probabilidad de ir (probabilidad de éxito) es $p=0,95$.

El tamaño del experimento es $n=280$.

Nuestra función de probabilidad será $B(n, p)=B(280, 0,95)$. Deseamos que haya 270 casos de éxito:

$$P(X = 270)$$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(X = 270) = \binom{280}{270} 0,95^{270} (0,05)^{280-270} = 0,0655$$

La probabilidad de que se presenten 270 pasajeros es del 6,55%.

PROBLEMA 11

Se sabe que la probabilidad de que un dardo impacte en una diana es 0,4. Si se lanzan 9 dardos, determina:

a) Qué tipo de distribución sigue la variable aleatoria que cuenta el número de dardos que dan en la diana.

b) La media y la desviación típica de esta distribución.

c) La probabilidad de que al menos 5 dardos impacten en la diana.

a) Un dardo puede dar en la diana o puede no impactar en la diana. Solo hay dos opciones posibles. El éxito es dar en la diana, con una probabilidad $p=0,4$. Estamos ante una variable aleatoria discreta, ya que el número de éxitos viene dado por números enteros: 0, 1, 2, 3, 4, ...

La distribución que cuenta el número de dardos que dan en la diana es una Binomial $B(n, p)$, siendo "n" el tamaño del experimento aleatorio y "p" la probabilidad de éxito: $B(9, 0,4)$

b) La media, o esperanza matemática, en la binomial se calcula multiplicando el tamaño del experimento por la probabilidad de éxito:

$$\mu = n \cdot p = 9 \cdot 0,4 = 3,6$$

La desviación típica en la binomial se calcula:

$$\sigma = \sqrt{n \cdot p \cdot (1 - p)} = \sqrt{9 \cdot 0,4 \cdot 0,6} = 1,47$$

c) Si al menos 5 dardos, de un total de 9, impactan en la diana, la probabilidad acumulada será:

$$P(X \geq 5) = P(X = 5) + P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9)$$

No estamos en condiciones de aproximar la binomial a la normal, ya que el producto $n \cdot p$ no supera el valor 5. Por lo que no podemos aplicar el Teorema central del límite.

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(X = 5) = \binom{9}{5} 0,4^5 (0,6)^{9-5} = 0,1672$$

$$P(X = 6) = \binom{9}{6} 0,4^6 (0,6)^{9-6} = 0,0743$$

$$P(X = 7) = \binom{9}{7} 0,4^7 (0,6)^{9-7} = 0,0212$$

$$P(X = 8) = \binom{9}{8} 0,4^8 (0,6)^{9-8} = 0,0035$$

$$P(X = 9) = \binom{9}{9} 0,4^9 (0,6)^{9-9} = 0,0003$$

Sumando las cinco probabilidades, obtenemos una probabilidad acumulada:

$$P(X \geq 5) = 0,2665$$

Obteniendo una probabilidad del 26,65%.

PROBLEMA 12

En un centro de fertilidad, cada intento de inseminación in vitro para cualquier pareja tiene un porcentaje de éxito del 30 %. Esta semana han acudido 10 parejas para realizar el tratamiento. Nos preguntamos por el número de ellas que consiguen tener hijos.

a) ¿De qué tipo de distribución se trata? Calcular su media y su desviación.

b) ¿Qué probabilidad hay de que ninguna pareja conciba? ¿y de que alguna conciba?

a) Estamos en una distribución binomial $B(n, p)$, donde cada pareja puede quedar embarazada (éxito) o no quedar embarazada (fracaso).

La media de una binomial se obtiene como el producto del tamaño de la muestra por la probabilidad de éxito.

$$\mu = n \cdot p = 10 \cdot 0,3 = 3$$

La desviación es la raíz cuadrada del tamaño de la muestra multiplicada por la probabilidad de éxito y por la probabilidad de fracaso.

$$\sigma = \sqrt{n \cdot p \cdot (1 - p)} = \sqrt{10 \cdot 0,3 \cdot 0,7} = 1,45$$

b) La probabilidad de que ninguna pareja quede embarazada es:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(X = 0) = \binom{10}{0} 0,3^0 (1 - 0,3)^{10-0} = 0,7^{10} = 0,0282 \rightarrow 2,82\%$$

La probabilidad de que al menos una pareja quede embarazada es el caso complementario a lo que hemos calculado antes. Es decir:

$$P(X \geq 1) = 1 - P(X = 0) \rightarrow P(X \geq 1) = 1 - 0,0282 = 0,9718 \rightarrow 97,18\%$$

PROBLEMA 13

La probabilidad de que un pez de una determinada especie sobreviva más de 6 meses es del 20 %. Se pide:

a) Si en un acuario tenemos 15 peces de esta especie nacidos este mes, halla la probabilidad de que al menos 2 de ellos sigan vivos dentro de 6 meses.

b) Si en un tanque de una piscifactoría hay 300 peces de esta especie nacidos este mismo mes, halla la probabilidad de que al cabo de 6 meses hayan sobrevivido al menos 50 de ellos.

a) La probabilidad de éxito es $p=0,2$. El tamaño de la muestra es $n=15$. Estamos ante un caso binomial $B(n,p)$.

La probabilidad de éxito de dos o más especies es:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) \\ P(X < 2) &= P(X = 0) + P(X = 1) \\ P(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ P(X = 0) &= \binom{15}{0} 0,2^0 (1-0,2)^{15-0} = 0,8^{15} = 0,0352 \\ P(X = 1) &= \binom{15}{1} 0,2^1 (1-0,2)^{15-1} = 15 \cdot 0,2 \cdot 0,8^{14} = 0,1319 \end{aligned}$$

Por lo tanto:

$$P(X \geq 2) = 1 - (0,0352 + 0,1319) = 0,8329 \rightarrow 83,29\%$$

b) Si consideramos una muestra $n=300$. En este caso podemos aproximar la binomial a la normal, y que se cumple.

$$\begin{aligned} n \cdot p &= 300 \cdot 0,2 = 60 > 5 \\ n \cdot (1-p) &= 300 \cdot 0,8 = 240 > 5 \\ B(n,p) &\approx N(\mu = n \cdot p, \sigma = \sqrt{n \cdot p \cdot (1-p)}) \\ B(300,0,2) &\approx N(60, 6,93) \end{aligned}$$

Si deseamos que sobrevivan al menos 50 parejas, razonamos así en la distribución normal:

$$P(X \geq 50) = 1 - P(X \leq 50)$$

Tipificamos la variable.

$$\begin{aligned} z &= \frac{x - \mu}{\sigma} \\ z &= \frac{50 - 60}{6,93} = -1,44 \\ P(X \leq 50) &= P(Z \leq -1,44) \end{aligned}$$

Por simetría de la distribución gaussiana:

$$\begin{aligned} P(Z \leq -1,44) &= P(Z \geq 1,44) \\ P(Z \geq 1,44) &= 1 - P(Z \leq 1,44) \end{aligned}$$

Mirando la tabla normalizada:

$$\begin{aligned} P(Z \geq 1,44) &= 1 - 0,9251 \\ P(Z \geq 1,44) &= 0,0749 \end{aligned}$$

Por lo tanto:

$$\begin{aligned} P(X \leq 50) &= 0,0749 \\ P(X \geq 50) &= 1 - 0,0749 \\ P(X \geq 50) &= 0,9251 \rightarrow 92,51\% \end{aligned}$$

PROBLEMA 14

Sea una población formada por los elementos 1, 2 y 3. Mediante muestreo aleatorio simple se selecciona una muestra de tamaño 2 con reemplazamiento.

- Escribe todas las muestras posibles dentro de la población.
- Calcula la varianza de las medias muestrales.

a) Todas las muestras con reemplazamiento de dos elementos son:

$\{1,1\}, \{1,2\}, \{1,3\}$
 $\{2,1\}, \{2,2\}, \{2,3\}$
 $\{3,1\}, \{3,2\}, \{3,3\}$

b) La media de cada muestra se obtiene sumando los dos números de cada muestra y dividiendo entre dos (tamaño de la muestra).

1	1,5	2
1,5	2	2,5
2	2,5	3

Tenemos una muestra de 9 elementos. La media muestral se obtiene sumando la media de cada muestra y dividiendo entre 9.

$$\mu = \frac{18}{9} = 2$$

Creemos la siguiente tabla estadística para obtener la varianza.

x	N	$n \cdot (x - \mu)^2$
1	1	1
1,5	2	0,5
2	3	0
2,5	2	0,5
3	1	1
Acumulado	9	3

La varianza queda:

$$\sigma^2 = \frac{3}{9} = 0,33$$

PROBLEMA 15

Sea una distribución normal tipificada. Calcula:

- Calcula el valor de k que cumple $P(Z \leq k) = 0,8485$
- Calcula el valor de k que cumple $P(Z \geq k) = 0,9972$
- Calcula el valor de k que cumple $P(1 \leq Z \leq k) = 0,15$
- Calcula el valor de k que cumple $P(Z \leq 2 + k) = 0,9896$

a) $P(Z \leq k) = 0,8485$

Mirando la tabla tipificada:

$$k = 1,03$$

b) La tabla nos da valores acumulados a partir de 0,50. Por lo tanto, el valor de "k" será negativo (a la izquierda del valor medio $z=0$).

$$P(Z \geq k) = 0,9972$$

Por simetría, el valor de "z" que deja a la izquierda un área de 0,9972 es 2,77. Por lo tanto:

$$k = -2,77$$

c) Sabemos que $P(1 \leq Z \leq k) = P(Z \leq k) - P(Z \leq 1)$

Mirando la tabla tipificada: $P(Z \leq 1) = 0,8413$. Por lo tanto:

$$P(Z \leq k) - 0,8413 = 0,15$$

$$P(Z \leq k) = 0,9913$$

Miramos en la tabla el primer valor de la variable Z que acumule el 0,9913 del área a su izquierda.

$$k = 2,38$$

d) El enunciado afirma que $P(Z \leq 2 + k) = 0,9896$.

Mirando la tabla, encontramos el primer valor de Z que acumula a su izquierda un área de 0,9896. Siendo este valor $z = 2,31$. Es decir:

$$2 + k = 2,31$$

$$k = 0,31$$

PROBLEMA 16

Sea una distribución $N(9, 0.5)$. Halla el valor de k que cumple:

- $P(X \leq k) = 0,9699$
- $P(X \geq k) = 0,5199$

a) $P(X \leq k) = 0,9699$

Mirando la tabla tipificada, el valor "z" que acumula a la izquierda un área de 0,9699 es $z=1,88$.

Usando la tipificación de la variable:

$$1,88 = \frac{x - 9}{0,5}$$

$$x = 9,94$$

b) $P(X \geq k) = 0,5199$

El valor "z" que deja a la derecha 0,5199 será negativo, ya que estará a la izquierda del valor medio $z=0$ en la tabla tipificada.

Por simetría, el valor de "z" que deja a la izquierda es 0,5199 es $z=0,05$. Por lo tanto, el valor que deja a la derecha 0,5199 es $z=-0,05$.

Usamos la tipificación de la variable:

$$-0,05 = \frac{x - 9}{0,5}$$

$$x = 8,98$$

PROBLEMA 17

Sea una distribución normal de media μ y desviación típica σ desconocidas. Calcula:

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$$

Tipificamos la variable.

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{\mu - 3\sigma - \mu}{\sigma} = -3$$

$$z = \frac{\mu + 3\sigma - \mu}{\sigma} = 3$$

Por lo tanto:

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq Z \leq 3)$$

$$P(-3 \leq Z \leq 3) = P(Z \leq 3) - P(Z \leq -3)$$

Por simetría de la distribución Gaussiana:

$$P(Z \leq -3) = 1 - P(Z \leq 3)$$

Es decir:

$$P(-3 \leq Z \leq 3) = P(Z \leq 3) - [1 - P(Z \leq 3)]$$

$$P(-3 \leq Z \leq 3) = 2 \cdot P(Z \leq 3) - 1$$

Mirando la tabla normalizada:

$$P(-3 \leq Z \leq 3) = 2 \cdot 0,99865 - 1 = 0,9973 = 99,73\%$$

PROBLEMA 18

En un examen tipo test de 200 preguntas de elección múltiple, cada pregunta tiene una opción correcta y una opción incorrecta. Se aprueba si se contesta a más de 110 preguntas correctas. Suponiendo que se contesta al azar, calcular la probabilidad de aprobar el examen.

Estamos en una distribución binomial. Cada pregunta, respondida al azar, puede ser correcta (éxito) o incorrecta (fracaso).

La probabilidad de éxito es $p=0,5$ ya que de dos opciones, solo una es la correcta.

Si tenemos una muestra de 200 preguntas podemos aproximar la binomial a la normal, ya que:

$$n \cdot p = 200 \cdot 0,5 = 100 > 5$$

$$n \cdot (1 - p) = 200 \cdot 0,5 = 100 > 5$$

Por lo tanto:

$$B(n, p) \approx N(n \cdot p, \sqrt{n \cdot p \cdot (1 - p)})$$

$$B(200, 0.5) \approx N(200 \cdot 0.5, \sqrt{200 \cdot 0.5 \cdot 0.5})$$

$$B(200, 0.5) \approx N(100, \sqrt{50})$$

Para aprobar necesitamos un éxito en al menos 110 preguntas correctas.

$$P(X \geq 110)$$

Tipificamos la variable.

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{110 - 100}{\sqrt{50}} = 1,41$$

$$P(X \geq 110) = P(Z \geq 1,41)$$

Por simetría:

$$P(Z \geq 1,41) = 1 - P(Z \leq 1,41)$$

Mirando la tabla normalizada:

$$P(Z \geq 1,41) = 1 - 0,9207 = 0,0793 \rightarrow 7,93\%$$

PROBLEMA 19

En un barrio el 60% de los hogares tiene al menos dos televisores. Se elige al azar una muestra de 50 hogares.

- ¿Cuál es la probabilidad de que al menos 20 de los citados hogares tengan cuando menos dos televisores?
- ¿Cuál es la probabilidad de que entre 35 y 40 hogares tengan al menos dos televisores?

a) Nuevo caso binomial. La probabilidad de éxito es $p=0,6$. El tamaño de la muestra es $n=50$. Por lo tanto estamos en condiciones de aproximar la binomial a la normal, ya que:

$$n \cdot p = 50 \cdot 0,6 = 30 > 5$$

$$n \cdot (1 - p) = 50 \cdot 0,4 = 20 > 5$$

Por lo tanto:

$$B(n, p) \approx N(n \cdot p, \sqrt{n \cdot p \cdot (1 - p)})$$

$$B(50, 0.6) \approx N(50 \cdot 0.6, \sqrt{50 \cdot 0.6 \cdot 0.4})$$

$$B(30, 0.5) \approx N(30, \sqrt{12})$$

Para que al menos 20 hogares posean dos o más televisores, debemos obtener la siguiente probabilidad acumulada:

$$P(X \geq 20)$$

Como la media es $\bar{\mu} = 30$, por simetría, podemos razonar que la probabilidad acumulada a la derecha de $x=20$ será igual a la probabilidad acumulada a la izquierda de $x=40$, ya que 20 y 40 se encuentran a 10 unidades del valor medio. Es decir:

$$P(X \geq 20) = P(X \leq 40)$$

Si tipificamos la variable:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{40 - 30}{\sqrt{12}} = 2,89$$

$$P(X \leq 40) = P(Z \leq 2,89)$$

Mirando la tabla tipificada:

$$P(Z \leq 2,89) = 0,99807 \rightarrow 99,81\%$$

b) La probabilidad de contar entre 35 y 40 hogares con al menos dos televisores es:

$$P(35 \leq X \leq 40) = P(X \leq 40) - P(X \leq 35)$$

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{35 - 30}{\sqrt{12}} = 1,44$$

Mirando la tabla:

$$P(X \leq 40) - P(X \leq 35) = P(Z \leq 2,89) - P(Z \leq 1,44) = 0,9979 - 0,9251 = 0,0728 \rightarrow 7,28\%$$

PROBLEMA 20

Los resultados de un examen tiene una media de 78 y desviación típica 36. Se supone distribución normal.

- ¿Cuál es la probabilidad de que una persona se presente al examen y obtenga al menos una puntuación igual a 72?
- Se declaran como "No Aptos" el 25% de los alumnos con las notas más bajas. ¿Qué proporción de estudiantes obtiene en su examen una nota que supera en al menos 5 puntos a la nota que marca la frontera entre "No Apto" y "Apto"?

a) Tipificamos la variable.

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{72 - 78}{36} = -0,166 \dots \approx -0,17$$

$$P(X > 72) = P(Z > -0,17)$$

Por simetría, el área de la curva tipificada acumulada a la derecha de $z = -0,17$ es lo mismo que el área acumulada a la izquierda de $z = 0,17$.

$$P(Z > -0,17) = P(Z < 0,17)$$

Mirando la tabla normalizada:

$$P(Z < 0,17) = 0,5671 \rightarrow 56,71\%$$

b) Buscamos el valor de la variable Z que acumula a su izquierda el 25% de las observaciones. Este valor, será negativo, ya que la media $z=0$ acumula el 50% de las observaciones.

Además, este valor negativo de Z es simétrico con otro valor positivo Z que deja a su derecha el 25%. O lo que es lo mismo, que deja a su izquierda el 75%.

$$P(Z \leq k) = 0,75$$

Miramos la tabla, el primer valor de Z que alcanza 0,75 es $k=0,68$.

Por lo tanto, el valor negativo que buscamos es $z=-0,68$.

Usamos la tipificación de la variable, para obtener el valor de la variable X.

$$z = \frac{x - \mu}{\sigma}$$

$$-0,68 = \frac{x - 78}{36}$$

$$x = 53,52$$

Buscamos los alumnos que superen en al menos 5 unidades esta nota. Es decir:

$$x = 53,52 + 5 = 58,52$$

Y ahora buscamos la proporción de alumnos que quedan a la derecha de 58,52. Para ello, volvemos a usar la relación tipificada.

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{58,52 - 78}{36} = -0,54$$

Por simetría, el área que queda a la derecha de $z = -0,54$ es la misma que queda a la izquierda de $z = 0,54$.

$$P(Z \geq -0,54) = P(Z \leq 0,54)$$

Miramos la tabla tipificada.

$$P(Z \leq 0,54) = 0,7054 \rightarrow 70,54\%$$

PROBLEMA 21

Los resultados de un test de inteligencia siguen una distribución normal $N(100,15)$.

- Determinar el porcentaje de la población que obtendrá un coeficiente entre 95 y 110.
- ¿Qué intervalo centrado en 100 contiene al 50% de la población?
- En una población de 2500 individuos, ¿cuántos individuos se esperan que obtengan un coeficiente superior a 125?

a) Normalizamos la variable.

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{95 - 100}{15} = -0,33$$

$$z = \frac{110 - 100}{15} = 0,66$$

$$P(95 \leq X \leq 110) = P(-0,33 \leq Z \leq 0,66)$$

La probabilidad del intervalo en cuestión se obtiene con la siguiente diferencia:

$$P(-0,33 \leq Z \leq 0,66) = P(Z \leq 0,66) - P(Z \leq -0,33)$$

Como la tabla no da probabilidades acumuladas para valores negativos de variable tipificada, razonamos de la siguiente forma (por simetría):

$$P(Z \leq -0,33) = 1 - P(Z \leq 0,33)$$

Por lo tanto:

$$P(Z \leq 0,66) - P(Z \leq -0,33) = P(Z \leq 0,66) - [1 - P(Z \leq 0,33)]$$

Miramos la tabla:

$$P(Z \leq 0,66) - [1 - P(Z \leq 0,33)] = 0,7454 - [1 - 0,6293] = 0,3747 \rightarrow 37,47\%$$

b) El intervalo centrado en 100 será simétrico respecto a este valor, que además es el valor medio de la distribución. Por lo tanto, considerando la variable tipificada de media $z=0$, el intervalo será $[-k, +k]$. Y se cumple lo siguiente:

$$P(-k \leq Z \leq k) = 0,50$$

Esto es lo mismo que afirmar que entre 0 y k obtenemos el 25% de las observaciones.

$$P(0 \leq Z \leq k) = 0,25$$

Y esto es lo mismo que afirmar que la probabilidad acumulada a la izquierda de k es del 75% (el 50% a la izquierda de la media $z=0$, más el 25% acumulado en el intervalo $[0, k]$).

$$P(Z \leq k) = 0,75$$

Mirando la tabla, el primer valor de la variable Z que acumula el 75% de las observaciones es $z=0,68$.

Por lo tanto, el intervalo tipificado solución es $[-0,68, +0,68]$.

Usamos la relación de tipificación para sacar el intervalo con la variable X.

$$z = \frac{x - \mu}{\sigma}$$

$$-0,68 = \frac{x - 100}{15} \rightarrow x = 89,8$$

$$0,68 = \frac{x - 100}{15} \rightarrow x = 110,2$$

Siendo el intervalo solución centrado alrededor de 100 el siguiente: $[89.8, 110.2]$

c) En el tercer apartado debemos obtener el porcentaje de alumnos con un coeficiente mayor o igual que 125. Eso implica obtener el área acumulada a la derecha de 125. Y una vez tengamos esa área, multiplicamos el porcentaje por 2500 para obtener el número total de individuos.

Tipificamos la variable:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{125 - 100}{15} = 1,66$$

$$P(X \geq 125) = P(Z \geq 1,66)$$

El área acumulada a la derecha de 1,66 coincide con el total del área de la curva normalizada menos la parte acumulada a la izquierda de 1,66.

$$P(Z \geq 1,66) = 1 - P(Z \leq 1,66)$$

Miramos la tabla.

$$1 - P(Z \leq 1,66) = 1 - 0,9515 = 0,0485 \rightarrow 4,85\%$$

Multiplicamos 2.500 por el porcentaje en tanto por uno.

$$2.500 \times 0,0485 \approx 121 \text{ individuos}$$

PROBLEMA 22

En un pueblo habitan 700 hombres adultos, 800 mujeres adultas y 500 menores. De él se quiere seleccionar una muestra de 80 personas, utilizando, para ello, muestreo estratificado con afijación proporcional. ¿Cuál será la composición que debe tener dicha muestra?

El total de personas es: $700 + 800 + 500 = 2.000$

El porcentaje de hombres, en tanto por uno, es $700/2.000 = 7/20$

El porcentaje de mujeres: $800/2.000 = 2/5$

Y el porcentaje de menores: $500/2.000 = 1/4$

Aplicamos estos porcentajes a la muestra de 80 personas.

$$80 \cdot 7/20 = 28 \text{ hombres}$$

$$80 \cdot 2/5 = 32 \text{ mujeres}$$

$$80 \cdot 1/4 = 20 \text{ menores}$$

PROBLEMA 23

El porcentaje de familias españolas con un solo hijo es del 20%.

Tomamos una muestra de 1.000 familias. ¿Cuál es la probabilidad de que al menos el 21% de las familias de la muestra tengan un solo hijo?

Contamos con variable binomial (o tiene un solo hijo, o no tiene un solo hijo). La proporción de la población es $p=20\%$ (0,20 en tanto por uno). El tamaño de la muestra es $n=1.000$.

Puede haber muestras donde la proporción sea del 20%, o del 19%, o del 21%, etc. Es decir, ahora tendremos como variable aleatoria la proporción. Y la media de esas proporciones se aproxima a una distribución normal, porque se cumple:

$$n \cdot p = 200 \geq 5$$

$$n \cdot (1 - p) = 800 \geq 5$$

En esta aproximación a la binomial, la media de las proporciones es igual a la proporción de la población (0,2) y la desviación de las proporciones es igual a la desviación de la población binomial dividido por la raíz cuadrada del tamaño de la muestra.

$$N\left(p, \sqrt{\frac{p \cdot (1-p)}{n}}\right) = N(0,20; 0,0126)$$

El porcentaje deseado para la muestra es de, al menos, el 21% (0,21 en tanto por uno). Es decir, la variable aleatoria debe ser mayor o igual que 0,21.

$$P(\bar{P} \geq 0,21)$$

Tipificamos como de costumbre:

$$P(\bar{P} \geq 0,21) = P\left(Z \geq \frac{0,21 - 0,20}{0,0126}\right) = P(Z \geq 0,79)$$

Recordamos que la tabla tipificada que utilizamos nos da valores acumulados a la izquierda. Por lo tanto:

$$P(Z \geq 0,79) = 1 - P(Z \leq 0,79) = 0,2148$$

Es decir, la probabilidad de que al menos el 21% de las 1.000 familias de la muestra tengan un solo hijo es del 21,48%. Y la aproximación es coherente porque:

PROBLEMA 24

Se ha obtenido una muestra de 25 alumnos de una Facultad. Se desea estudiar la calificación media de los expedientes de los alumnos de la facultad. Se sabe, de cursos anteriores, que la desviación típica de toda la Facultad es 2,01.

La media de la muestra fue de 4,9.

Obtener la media de la población en el intervalo de confianza del 90% y al 99%.

Ahora partimos de los datos de una muestra (25 alumnos) y deseamos generalizar (inferir) conclusiones válidas para toda la población (cuyo tamaño es mayor que el de la muestra).

Deseamos conocer, en un intervalo de confianza, la media poblacional partiendo de los datos de la media de la muestra (4,9) y conociendo el valor de la desviación de toda la población (2,01).

Para un intervalo de confianza del 90% existe un riesgo del 10%. Por lo que el nivel de significación es $\alpha = 10\%$. O lo que es lo mismo, $\alpha = 0.1$ en tanto por uno.

Con estos datos, podemos sacar el intervalo de confianza (al 90%) de encontrar el valor de la media de la población.

$$\left(\bar{x} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

Conocemos la desviación de la población:

$$\sigma = 2,01$$

El tamaño de la muestra es $n = 25$.

Conocemos la media de la muestra:

$$\mu_{\bar{X}} = \bar{x} = 4,9$$

El nivel de significación es:

$$1 - 0,90 = 0,10$$

El intervalo de confianza de la media de la población resulta:

$$\left(4,9 - z_{0,05} \cdot \frac{2,01}{\sqrt{25}}, 4,9 + z_{0,05} \cdot \frac{2,01}{\sqrt{25}} \right)$$

Sabemos que:

$$P(Z \geq z_{0,05}) = 0,05$$

$$P(Z \leq z_{0,05}) = 1 - 0,05 = 0,95$$

$$z_{0,05} = 1,65$$

Por lo que el intervalo de confianza para la media de la población será:

$$\left(4,9 - 1,65 \cdot \frac{2,01}{\sqrt{25}}, 4,9 + 1,65 \cdot \frac{2,01}{\sqrt{25}} \right) = (4,24, 5,56)$$

Repetimos el razonamiento para el nuevo nivel de significación:

$$1 - 0,99 = 0,01$$

El intervalo de confianza de la media de la población resulta:

$$\left(4,9 - z_{0,005} \cdot \frac{2,01}{\sqrt{25}}, 4,9 + z_{0,005} \cdot \frac{2,01}{\sqrt{25}} \right)$$

Sabemos que:

$$P(Z \geq z_{0,005}) = 0,005$$

$$P(Z \leq z_{0,005}) = 1 - 0,005 = 0,995$$

$$z_{0,005} = 2,58$$

Por lo que el intervalo de confianza para la media de la población será:

$$\left(4,9 - 2,58 \cdot \frac{2,01}{\sqrt{25}}, 4,9 + 2,58 \cdot \frac{2,01}{\sqrt{25}} \right) = (3,86, 5,94)$$

PROBLEMA 25

Sea X una variable aleatoria Normal de media 50 y desviación típica 4. Se toman muestras de tamaño 16.

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

b) ¿Cuál es la probabilidad de que la media muestral esté comprendida entre 47,5 y 52,5?

a) Cada muestra tiene una media. La media de estas medias se llama media muestral.

La distribución de las medias de las muestras es una distribución normal (gaussiana) donde la media ($\mu_{\bar{X}}$) coincide con la media de la población μ :

$$\mu_{\bar{X}} = \mu = 50$$

La desviación de las medias de las muestras ($\sigma_{\bar{X}}$) coincide con la desviación de la población dividido por la raíz cuadrada del tamaño de las muestras ($\frac{\sigma}{\sqrt{n}}$).

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{16}} = 1$$

Es decir: la distribución de la media muestral es una distribución normal de media 50 y desviación 1: $N(50, 1)$.

b) Buscamos: $P(47,5 \leq \bar{X} \leq 52,5)$

Tipificamos:

$$z = \frac{47,5 - 50}{1} = -2,5$$

$$z = \frac{52,5 - 50}{1} = 2,5$$

$$P(-2,5 \leq Z \leq 2,5)$$

Por simetría:

$$P(-2,5 \leq Z \leq 2,5) = 2 \cdot P(0 \leq Z \leq 2,5) = 2 \cdot [P(Z \leq 2,5) - P(Z \leq 0)] = 2 \cdot [0,9938 - 0,5] = 0,9876 \rightarrow 98,76\%$$

PROBLEMA 26

El cociente intelectual de los alumnos de un centro educativo se distribuye según una ley Normal de media 110 y desviación típica 15. Se extrae una muestra aleatoria simple de 25 alumnos.

a) ¿Cuál es la probabilidad de que la media del cociente intelectual de los alumnos de esa muestra sea superior a 113?

b) Razone cómo se vería afectada la respuesta a la pregunta anterior si el tamaño de la muestra aumentase.

a) Al extraer una muestra de 25 alumnos, pasamos a una distribución donde la media es 110 y donde la desviación es $\frac{15}{\sqrt{25}} = 3$.

$$N(110, 3)$$

$$P(X \geq 113)$$

Tipificamos la variable:

$$z = \frac{113 - 110}{3} = 1$$

$$P(Z \geq 1) = 1 - P(Z \leq 1) = 1 - 0,8413 = 0,1587 \rightarrow 15,87\%$$

b) Si el tamaño de la muestra aumenta, disminuye la desviación muestral ya que crece el denominador $\sqrt{n}$, donde n es el tamaño de la muestra. Por lo tanto, menor será el grado de dispersión. Y eso significa que los valores de la variable aleatoria de la media muestral estarán más cercanos al valor marcado por la media muestral.

PROBLEMA 27

El peso de los individuos de una población se distribuye según una ley Normal de desviación típica 6 kg. Calcule el tamaño mínimo de la muestra para estimar, con un nivel de confianza del 95%, el peso medio en la población con un error no superior a 1 kg.

Conocemos la desviación de la población: $\sigma = 6$.

El nivel de confianza es del 95%, por lo que el nivel de significación es: $\alpha = 1 - 0,95 = 0,05$.

El error máximo es 1. La expresión para calcular ese error máximo resulta:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

El tamaño de la muestra es "n" y es el valor que debemos calcular.

$$1 = z_{0,025} \cdot \frac{6}{\sqrt{n}}$$

Como hemos calculado en ejercicios anteriores, el valor $Z_{0,025}$ es el valor de la variable tipificada que deja a la derecha un área igual a 0,025 en la curva normalizada. O lo que es lo mismo, el valor que deja a su izquierda un área igual a 0,975. Mirando la tabla normalizada:

$$z_{0,025} = 1,96$$

Sustituyendo:

$$1 = 1,96 \cdot \frac{6}{\sqrt{n}}$$

$$\sqrt{n} = 11,76$$

$$n = 138,30$$

Como la variable es discreta, debemos aproximar a un número entero. ¿Elegimos 138 o 139?

Si el tamaño de la muestra disminuye, el error máximo aumenta. Como no queremos superar un el valor 1 en el error, redondeamos al alza. Por lo que el tamaño mínimo de la muestra debe ser 139.

PROBLEMA 28

En los individuos de una población, la concentración de una proteína en sangre se distribuye según una ley Normal de media desconocida y desviación típica 0.42 g/dl. Se toma una muestra aleatoria de 49 individuos y se obtiene una media muestral de 6,85 g/dl.

- a) Obtenga un intervalo de confianza, al 96%, para estimar la concentración media de la proteína en sangre de los individuos de esa población.
- b) ¿Es suficiente el tamaño de esa muestra para obtener un intervalo de confianza, al 98%, con un error menor que 0.125 g/dl?

a) Conocemos la desviación de la población: $\sigma = 0.42$.

La media de la muestra es $\bar{x} = 6.85$. El tamaño de la muestra es $n = 49$.

Si el intervalo de confianza es del 96% significa que el nivel de significación es: $\alpha = 1 - 0.96 = 0.04$.

Deseamos estimar (inferir) la media poblacional con estos datos y con el intervalo de confianza del 96%.

$$\left(\bar{x} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

El valor $z_{0,04/2} = z_{0,02}$ es el valor de la curva normalizada que deja a la derecha un área igual a 0.02.

$$P(Z \geq z_{0,02}) = 0,02$$

$$P(Z \leq z_{0,02}) = 1 - 0,02 = 0,98$$

$$z_{0,02} = 2,06$$

Por lo que el intervalo de confianza para la media de la población será:

$$\left(6,85 - 2,06 \cdot \frac{0,42}{\sqrt{49}}, 6,85 + 2,06 \cdot \frac{0,42}{\sqrt{49}} \right) = (6,72, 6,97)$$

b) Un nivel de confianza del 98% genera un nivel de significación $\alpha = 1 - 0.98 = 0.02$.

Conocemos la desviación poblacional: $\sigma = 0.42$. Y el tamaño: $n=49$. Con estos datos, podemos despejar el valor del error máximo:

$$E = Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$E = z_{0,01} \cdot \frac{0,42}{\sqrt{49}}$$

Mirando la tabla de tipificada:

$$z_{0,01} = 2,33$$

$$E = 2,33 \cdot \frac{0,42}{\sqrt{49}}$$

$$E = 0.1398$$

Es un valor superior al fijado por el enunciado (0.125), por lo que deberíamos aumentar el tamaño de la muestra para reducir el valor máximo del error.

PROBLEMA 29

El peso neto de las tabletas de chocolate de una determinada marca es una variable aleatoria Normal con media μ y desviación típica 7 gramos. Se sabe que 36 tabletas, elegidas al azar, han dado un peso total de 5274 gramos.

- a) Calcule un intervalo con un nivel de confianza del 94% para la media μ .
- b) Con el mismo nivel de confianza, ¿cuántas tabletas, como mínimo, habrá que tomar como muestra para que la amplitud del intervalo que se obtenga sea, como máximo, de 3 gramos?

a) Necesitamos estimar la media de la población con un intervalo de confianza del 94%. Eso genera un nivel de significación de $\alpha = 1 - 0.94 = 0.06$.

$$\left(\bar{x} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

Sustituimos los valores del enunciado: desviación de la población igual a 7 gramos, tamaño de la muestra igual a 36, y peso total 5274 gramos. **Por lo que el peso medio es: $5274/36 = 146,5$ gramos.**

$$\left(146.5 - z_{0.06} \cdot \frac{7}{\sqrt{36}}, 146.5 + z_{0.06} \cdot \frac{7}{\sqrt{36}}\right)$$

El valor $z_{\frac{0.06}{2}} = z_{0.03}$ es el que deja a la derecha un área 0.03 en la curva tipificada. O lo que es lo mismo, que deja a la izquierda un área 0.97.

$$z_{0.03} = 1.89$$

$$\left(146.5 - 1.89 \cdot \frac{7}{\sqrt{36}}, 146.5 + 1.89 \cdot \frac{7}{\sqrt{36}}\right) = (144.30, 148.69)$$

b) Nos piden el tamaño mínimo para que el error en la estimación de la media poblacional no supere los 1,5 gramos. **Este error se obtiene de operar $3/2 = 1,5$. Porque el error máximo es la mitad de la amplitud del intervalo.**

$$n = \left(\frac{Z_{\alpha} \cdot \sigma}{E}\right)^2$$

$$n = \left(\frac{z_{0.03} \cdot 7}{1,5}\right)^2$$

$$n = \left(\frac{1.89 \cdot 7}{1,5}\right)^2$$

$$n = 77,79$$

Debemos redondear al alza, para garantizar que el error no supere el valor de 1,5 gramos. Por lo tanto, el tamaño muestral debe ser 78.

PROBLEMA 30

Una variable aleatoria X se distribuye de forma Normal, con media μ y desviación típica $\sigma = 0.9$.

a) Una muestra aleatoria de tamaño 9 ha proporcionado los siguientes valores de X : 7.0, 6.4, 8.0, 7.1, 7.3, 7.4, 5.6, 8.8, 7.2

Obtenga un intervalo de confianza para la media μ , con un nivel de confianza del 97%.

b) Con otra muestra, se ha obtenido que un intervalo de confianza para μ , al 95%, es el siguiente (6.906, 7.494). ¿Cuál es el tamaño de la muestra utilizada?

a) En primer lugar debemos obtener la media muestral, sumando los 9 valores de la muestra y dividiendo entre 9.

$$\bar{x} = 7.2$$

Conocemos la desviación poblacional (0.9). El nivel de confianza del 97% genera un nivel de significación:

$$\alpha = 1 - 0.97 = 0.03$$

$$\left(\bar{x} - Z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

$$\left(7.2 - z_{0.03} \cdot \frac{0.9}{\sqrt{9}}, 7.2 + z_{0.03} \cdot \frac{0.9}{\sqrt{9}}\right)$$

$$z_{0.015} = 2.17$$

$$\left(7.2 - 2.17 \cdot \frac{0.9}{\sqrt{9}}, 7.2 + 2.17 \cdot \frac{0.9}{\sqrt{9}}\right) = (6.55, 7.85)$$

b) El error máximo de la estimación de la media poblacional coincide con la mitad del tamaño del intervalo de confianza:

$$E = (7.494 - 6.906)/2 = 0.588/2 = 0.294$$

El nivel de confianza del 95% genera un nivel de significación:

$$\alpha = 1 - 0.95 = 0.05$$

$$z_{0.05/2} = z_{0.025} = 1.96$$

Con estos datos, podemos calcular el tamaño de la muestra:

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \cdot \sigma}{E} \right)^2$$

$$n = \left(\frac{1.96 \cdot 0.9}{0.294} \right)^2 = 36$$

El tamaño de la muestra es 36.

PROBLEMA 31

En un estudio de mercado del automóvil en una ciudad se ha tomado una muestra aleatoria de 300 turismos, y se ha encontrado que 75 de ellos tienen motor diésel. Para un nivel de confianza de 94%:

a) Determine un intervalo de confianza de la proporción de turismos que tienen motor diésel en esa ciudad.

b) ¿Cuál es el error máximo de la estimación de la proporción?

a) Ahora estamos en un experimento binomial: el turismo puede tener motor diésel o no tenerlo.

El tamaño de la muestra es $n=300$. La proporción de la muestra es $75/300$, es decir, $\bar{p} = 0.25$. A partir de estos datos de la muestra, debemos inferir el valor de la proporción de la población con un nivel de confianza del 94%, que genera un nivel de significación:

$$\alpha = 1 - 0.94 = 0.06$$

$$z_{0,06/2} = z_{0,03} = 1.89$$

$$\left(\hat{p} - Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}}, \hat{p} + Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \right)$$

$$\left(0.25 - 1.89 \cdot \sqrt{\frac{0.25 \cdot (1 - 0.25)}{300}}, 0.25 + 1.89 \cdot \sqrt{\frac{0.25 \cdot (1 - 0.25)}{300}} \right) = (0.202, 0.297)$$

b) El error máximo en la estimación de la proporción poblacional:

$$E = z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

$$E = z_{0,03} \cdot \sqrt{\frac{0.25 \cdot (1 - 0.25)}{300}}$$

$$E = 1.89 \cdot \sqrt{\frac{0.25 \cdot (1 - 0.25)}{300}} = 0.04725$$

Este valor coincide con la mitad del tamaño del intervalo (0.202, 0.297).

PROBLEMA 32

Una empresa consultora quiere estudiar algunos aspectos de la vida laboral de los trabajadores de una ciudad. Para ello selecciona una muestra aleatoria de 500 trabajadores, de los que 118 afirman residir en otra ciudad. Con un nivel de confianza del 93%,

a) Calcule un intervalo de confianza para la proporción de trabajadores que residen fuera.

b) Calcule el error cometido en el intervalo anterior.

a) Otra vez un experimento binomial: el trabajador puede residir en otra ciudad o no hacerlo.

El tamaño de la muestra es $n=500$. La proporción de la muestra es $118/500$, es decir, $\bar{p} = 0.236$. A partir de estos datos de la muestra, debemos inferir el valor de la proporción de la población con un nivel de confianza del 93%, que genera un nivel de significación:

$$\alpha = 1 - 0.93 = 0.07$$

$$z_{0,07/2} = z_{0,035} = 1.82$$

$$\left(\hat{p} - Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}}, \hat{p} + Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \right)$$

$$\left(0.236 - 1.82 \cdot \sqrt{\frac{0.236 \cdot (1 - 0.236)}{500}}, 0.236 + 1.82 \cdot \sqrt{\frac{0.236 \cdot (1 - 0.236)}{500}} \right) = (0.201, 0.271)$$

b) El error máximo en la estimación de la proporción poblacional:

$$E = z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

$$E = z_{0.035} \cdot \sqrt{\frac{0.236 \cdot (1 - 0.236)}{500}}$$

$$E = 1.82 \cdot \sqrt{\frac{0.25 \cdot (1 - 0.25)}{500}} = 0.0346$$

Este valor coincide con la mitad del tamaño del intervalo (0.201, 0.271).

PROBLEMA 33

Se realizan dos muestreos estratificados con afijación proporcional para una población dividida en cuatro estratos E1, E2, E3, E4. En la primera muestra se han seleccionado 25 individuos de E1 y 30 de E2. En la segunda muestra se han seleccionado 80 individuos de E3 y 100 de E4. Sabiendo que el estrato E1 tiene 500 individuos y que el E3 tiene 400, determine el tamaño de cada estrato de la población y el tamaño de las muestras de cada estrato.

$$\text{Primera Muestra : } \frac{N}{n} = \frac{500}{25} = \frac{N_2}{30} = \frac{400}{n_3} = \frac{N_4}{n_4}$$

$$\text{Segunda Muestra : } \frac{N}{m} = \frac{500}{m_1} = \frac{N_2}{m_2} = \frac{400}{80} = \frac{N_4}{100}$$

$$\text{Primera : } \frac{500}{25} = \frac{N_2}{30} \rightarrow N_2 = 600$$

$$\text{Primera : } \frac{500}{25} = \frac{400}{n_3} \rightarrow n_3 = 20$$

$$\text{Segunda : } \frac{500}{m_1} = \frac{400}{80} \rightarrow m_1 = 100$$

$$\text{Segunda : } \frac{400}{80} = \frac{N_4}{100} \rightarrow N_4 = 500$$

$$\text{Primera Muestra : } \frac{N}{n} = \frac{500}{25} = \frac{600}{30} = \frac{400}{20} = \frac{500}{n_4}$$

$$\text{Segunda Muestra : } \frac{N}{m} = \frac{500}{100} = \frac{600}{m_2} = \frac{400}{80} = \frac{500}{100}$$

$$\text{Primera : } \frac{400}{20} = \frac{500}{n_4} \rightarrow n_4 = 25$$

$$\text{Segunda : } \frac{500}{100} = \frac{600}{m_2} \rightarrow m_2 = 120$$

$$\text{Primera Muestra : } \frac{N}{n} = \frac{500}{25} = \frac{600}{30} = \frac{400}{20} = \frac{500}{25}$$

$$\text{Segunda Muestra : } \frac{N}{m} = \frac{500}{100} = \frac{600}{120} = \frac{400}{80} = \frac{500}{100}$$

$$N = 500 + 600 + 400 + 500 \rightarrow N = 2.000$$

$$n = 25 + 30 + 20 + 25 \rightarrow n = 100$$

$$m = 100 + 120 + 80 + 100 \rightarrow m = 400$$