Tema 10. Muestreo. Intervalos de confianza
Problemas Resueltos

Muestreo

1. En una ciudad se quiere hacer una encuesta para conocer el porcentaje de ciudadanos que aprueban la gestión del ayuntamiento en cuestiones medioambientales (limpieza de calles, contaminación, cuidado de parques…). Se pretende que la muestra sea representativa por sexo y edad; para la edad se establecen tres estratos: 10 a 25 años (jóvenes), 25 a 60 años (adultos) y mayores de 60. El número de personas de cada grupo es: jóvenes, 3000; adultos, 8500; mayores de 60, 2500. Por sexo, la distribución es: 6800 hombres y 7200 mujeres, que se suponen proporcionales a cada grupo de edad.
Si el tamaño de la muestra es de 500 personas, determina, redondeando si es necesario, el tamaño muestral correspondiente a cada estrato.
Solución:
El total de ciudadanos que pueden ser encuestados es de 3000 + 8500 + 2500 = 14000. De ellos, el $\frac{6800}{14000}\cdot 100 = 48,571\%$ son hombres. Por tanto, en cada estrato se tienen:

	Jóvenes	Adultos	Mayores
Hombres	1457	4129	1214
Mujeres	1543	4371	1286

Se pretende encuestar a 500 personas de las 14000 existentes; esto supone 1 por cada 28 (14.000/500 = 28), luego dividiendo entre 28 el número de componentes de cada estrato, se obtienen los tamaños muestrales respectivos, que son los dados en la siguiente tabla.

	Jóvenes	Adultos	Mayores	Total
Hombres	52	147 + 1	43	243
Mujeres	55	156	46	257

En hombres adultos se ha añadido 1 con el fin de completar la muestra, pues en los redondeos se pierde un individuo. Se hace en ese estrato por ser el que deja más resto.

2. Supongamos que en un centro escolar los alumnos y docentes se distribuyen de acuerdo con la tabla:

Alumnos	1 y 2º ESO	3º y 4º ESO	BACH	PROFS
Hombres	110	95	115	20
Mujeres	130	120	130	25

Si se quiere realizar una encuesta entre ellos de tamaño n = 50, por el método de muestreo estratificado por sexo y nivel de trabajo, ¿a cuántas personas de cada *clase* hay que preguntar?
Solución:
El total de personas que pueden ser encuestados es de 240 + 215 + 245 + 45 = 745, que es la suma de los cuatro niveles. Como se pretende encuestar a 50 personas de las 745 posibles, y 745 : 50 = 14,9, hay que preguntar a 1 de cada 14,9 personas de cada grupo.
Los valores que se obtienen son:

Muestra	1 y 2º ESO	3º y 4º ESO	BACH	PROFS
Hombres	7,38	6,38	7,93	1,34
Mujeres	8,72	8,05	8,72	1,68

Hay que redondear atendiendo a los restos.
Puede optarse por la siguiente elección:

Muestra	1 y 2º ESO	3º y 4º ESO	BACH	PROFS
Hombres	7	6	8	1
Mujeres	9	8	9	2

3. Supongamos que en tu ciudad (o en tu comarca) hay 4 institutos o colegios (IES), cuyo alumnado se distribuye como sigue:

	IES 1	IES 2	IES 3	IES 4
Alumnos	370	460	620	520
Alumnas	340	500	650	540

Si se quiere realizar un muestreo de tamaño 100, estratificado por sexo y centro escolar, ¿cuáles serían los tamaños muestrales correspondientes a cada IES y a cada estrato?
Solución:
El total de alumnos/as por IES se indica en la siguiente tabla:

	IES 1	IES 2	IES 3	IES 4	Total
Alumnos	370	460	620	520	1970
Alumnas	340	500	650	540	2030
Total	710	960	1270	1060	4000

Como hay que elegir una muestra de tamaño 100 habrá que tomar 1 de cada 40 estudiantes.
Los valores que se obtienen son:

	IES 1	IES 2	IES 3	IES 4	Total
Alumnos	9,25	11,5	15,5	13	49,25
Alumnas	8,5	12,5	16,25	13,5	50,75
Total	17,75	24	31,75	26,5	100

Hay que redondear atendiendo a los restos. En este caso puede optarse por subir los representantes del IES 1 y 3 y bajar los del IES 4; eligiendo a 49 alumnos y 51 alumnas. Por tanto, el tamaño muestral correspondiente a cada estrato es el siguiente:

	IES 1	IES 2	IES 3	IES 4	Total
Alumnos	9	11	16	13	49
Alumnas	9	13	16	13	51
Total	18	24	32	26	100

4. (Selectividad, Galicia 2016).
a) Se desea tomar una muestra estratificada de las personas mayores de edad de un municipio, cuyos estratos son los siguientes intervalos de edades, en años: de 18 a 30, de 31 a 45, de 45 a 60 y mayores de 60. En el primer intervalo hay 7500 personas, en el segundo 8400, en el tercero 5700 y en el cuarto 3000. Calcula el tamaño de la muestra total y su composición, sabiendo que el muestreo se hace con afijación proporcional y se han elegido 375 personas del primer estrato.
b) Dada la población {2, 4, 6}, construye todas las muestras posibles de tamaño 2, que puedan formar mediante muestreo aleatorio simple, y halla la varianza de las medias muestrales de todas las muestras.
Solución:
a) Como la afijación es proporcional, el peso de cada estrato en la muestra es directamente proporcional a los individuos de la población correspondiente.
Si en el primer estrato, formado por 7500 personal se han elegido $m_1 = 375$;
del segundo estrato: 8400 $\to$ se elegirán m_2 $\Rightarrow m_2 = \dfrac{8400 \cdot 375}{7500} = 420$;

del tercer estrato: 5700 → se elegirán m_3 $\Rightarrow m_3 = \dfrac{5700 \cdot 375}{7500} = 285$;

del cuarto estrato: 3000 → se elegirán m_4 $\Rightarrow m_4 = \dfrac{3000 \cdot 375}{7500} = 150$;

El tamaño total de la muestra es 375 + 420 + 285 + 150 = 1230.
Se elige 1 de cada 20 personas en cada estrato.

b) En el muestreo aleatorio simple se mantiene la probabilidad de extracción en cada caso. Por tanto, las extracciones deben hacerse con reemplazamiento.
El número de muestras de tamaño 2 que pueden obtenerse de la población {2, 4, 6} son 9:
{2, 2}; {2, 4}; {2, 6}; {4, 2}; {4, 4}; {4, 6}; {6, 2}; {6, 4} y {6, 6}
(Son las variaciones con repetición de 3 elementos tomados 2 a 2).

Muestra	Elementos	Media de las muestras: $\bar{x}_i$	$\left(\bar{x}_i - \bar{x}\right)^2$
$M1$	{2, 2}	2	4
$M2$	{2, 4}	3	1
$M3$	{2, 6}	4	0
$M4$	{4, 2}	3	1
$M5$	{4, 4}	4	0
$M6$	{4, 6}	5	1
$M7$	{6, 2}	4	0
$M8$	{6, 4}	5	1
$M9$	{6, 6}	6	4
Sumas		36	12

Recuerda que la media y la varianza valen:

$$\mu = \frac{\sum_{i=1}^{n} x_i}{N}; \quad \sigma^2 = \frac{\sum_{i=1}^{n} \left(x_i - \mu\right)^2}{N}$$

La media y la varianza de la población son:

$$\mu = \frac{2+4+6}{3} = 4;$$

$$\sigma^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3} = \frac{8}{3}.$$

Media de las muestras: $\bar{x} = \dfrac{36}{9} = 4$;

Varianza de las muestras: $\sigma_{\bar{x}}^{\,2} = \dfrac{12}{9} = \dfrac{4}{3}$ → Desviación típica: $\sigma_{\bar{x}} = \sqrt{\dfrac{4}{3}} = \dfrac{2}{\sqrt{3}}$.

Puede comprobarse que: $\mu = \bar{x} = 4$; y que $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{\sqrt{8/3}}{\sqrt{2}} = \dfrac{2}{\sqrt{3}}$.

Distribución Binomial (Repaso)

5. Un dado, cuyas caras están numeradas del 1 al 6, se lanza cinco veces. Halla la probabilidad de que el número 3 salga:

a) Exactamente dos veces. b) Una vez a lo sumo. c) Más de una vez.

Solución:

El número de treses puede medirse a partir de la binomial $B\left(5, \dfrac{1}{6}\right)$.

a) $P(X=2) = \binom{5}{2} \cdot \left(\dfrac{1}{6}\right)^2 \cdot \left(\dfrac{5}{6}\right)^3 = 10 \cdot \dfrac{125}{7776} \approx 0{,}16075$.

b) $P(X \le 1) = P(X=0) + P(X=1) = \binom{5}{0} \cdot \left(\dfrac{1}{6}\right)^0 \cdot \left(\dfrac{5}{6}\right)^5 + \binom{5}{1} \cdot \left(\dfrac{1}{6}\right)^1 \cdot \left(\dfrac{5}{6}\right)^4 =$

$$= 1\cdot\frac{3125}{7776}+5\cdot\frac{625}{7776}=\frac{6250}{7776}\approx 0,80375.$$

c) $P(X>1)=1-P(X\le 1)=1-\dfrac{6250}{7776}=1-0,80375=0,19625$.

6. En un Centro Escolar el 25% de los alumnos son de origen extranjero. Si se eligen 6 estudiantes al azar, ¿cuál es la probabilidad de 4 o más sean de origen extranjero?

Solución:

El número de alumnos de origen extranjero puede estudiarse como una binomial $B(6, 0,25)$.

$$P(X\ge 4)=P(X=4)+P(X=5)+P(X=6) =$$

$$= \binom{6}{4}0,25^4\cdot 0,75^2+\binom{6}{5}0,25^5\cdot 0,75+\binom{6}{6}0,25^6 = 15\cdot 0,25^4\cdot 0,75^2+6\cdot 0,25^5\cdot 0,75+0,24^6 =$$

$$= 0,03296+0,00439+0,00024=0,03759.$$

7. Se lanza una moneda correcta 10 veces y se mide el número de caras y cruces obtenidas.

a) ¿Cuántos resultados forman el espacio muestral? ¿Cuál es la probabilidad de cada uno de los resultados posibles?

b) ¿Cuál es la probabilidad de que salgan 4 caras?

Solución:

a) Si para cada moneda se designa por 0 el suceso cara y por 1 el suceso cruz, el espacio muestral será:

E = {0000000000, 0000000001, 0000000010…, 1111111110, 1111111111}.

Son las variaciones con repetición de 2 elementos (el 0 y el 1) tomados 10 a 10. Su número es $VR_{2,10}=2^0=1024$.

P(cada suceso elemental) = $\dfrac{1}{1024}$.

b) Es un experimento binomial: $B\left(10,\ \dfrac{1}{2}\right)\to p=q=\dfrac{1}{2}$.

Si X cuenta el número de caras,

P(4 caras) = $P(X=4)=\dbinom{10}{4}\dfrac{1}{2^{10}}=\dfrac{210}{2^{10}}=\dfrac{105}{512}$.

8. En una moneda trucada la probabilidad de obtener cara es 0,4. Si se lanza 5 veces, calcula la probabilidad de obtener al menos 3 caras.

Solución:

Se trata de una distribución de probabilidad binomial: $B(5, 0,4)\to p = 0,4$; $q = 0,6$.

$$P(X\ge 3)=P(X=3)+P(X=4)+P(X=5) = \binom{5}{3}0,4^3\cdot 0,6^2+\binom{5}{4}0,4^4\cdot 0,6+\binom{5}{5}0,4^5 =$$

$$= 10\cdot 0,4^3\cdot 0,6^2+5\cdot 0,4^4\cdot 0,6+0,4^5=0,2304+0,0768+0,01024=0,31744.$$

9. Un examen consta de 8 preguntas con 3 posibles respuestas cada una, de las que sólo una de ellas es correcta. Si un estudiante responde al azar marcando las respuestas aleatoriamente, calcula la probabilidad de que:

a) No acierte ninguna respuesta correcta. b) Acierte 6 o más preguntas.

Solución:

Si se contesta al azar, la probabilidad de acertar $p=\frac{1}{3}$; la de fallar, $q=\frac{2}{3}$.

Se trata de una distribución de probabilidad binomial, $B\left(8,\ \frac{1}{3}\right)$.

a) $P(X=0)=\binom{8}{0}\left(\frac{1}{3}\right)^0\cdot\left(\frac{2}{3}\right)^8=\frac{256}{6561}=0,039$.

b) $P(X\geq 6)=P(X=6)+\mathrm{P}(X=7)+P(X=8)=$

$$=\binom{8}{6}\cdot\left(\frac{1}{3}\right)^6\cdot\left(\frac{2}{3}\right)^2+\binom{8}{7}\cdot\left(\frac{1}{3}\right)^1\cdot\left(\frac{2}{3}\right)+\binom{8}{8}\cdot\left(\frac{1}{3}\right)^8\cdot\left(\frac{2}{3}\right)^0=$$

$$=28\cdot\frac{4}{6561}+8\cdot\frac{2}{6561}+\frac{1}{6561}=\frac{129}{6561}=0,0197.$$

Distribución Normal (Repaso)

10. Utilizando la tabla normal *N*(0, 1) calcula:

a) $P(Z<1,2)$ b) $P(Z<1,27)$ c) $P(Z<-1,2)$ d) $P(Z<-1,27)$

Solución:

En la tabla puede leerse directamente:

a) $P(Z<1,2)$ = 0,8849.

b) $P(Z<1,27)$ = 0,8980.

c) $P(Z<-1,2)=1-P(Z<1,2)$ = 1 – 0,8849 = 0,1151.

d) $P(Z<-1,27)=1-P(Z<1,27)$ = 1 – 0,8980 = 0,1020.

11. Utilizando la tabla normal *N*(0, 1) calcula interpolando:

a) $P(Z<1,325)$ b) $P(Z<1,645)$ c) $P(Z<0,666)$ d) $P(Z<1,863)$

Solución:

a) $P(Z<1,325)$ está entre $P(Z<1,32)$ y $P(Z<1,33)$. Como es el punto medio, puede asignársele la media de ambos resultados:

$$P(Z<1,325)=\frac{P(Z<1,32)+P(Z<1,33)}{2}=\frac{0,9066+0,9082}{2}=0,9074.$$

b) Análogamente:

$$P(Z<1,645)=\frac{P(Z<1,64)+P(Z<1,65)}{2}=\frac{0,9495+0,9505}{2}=0,95.$$

c) Para calcular $P(Z<0,666)$ debe dividirse la diferencia de los valores $P(Z<0,66)$ y $P(Z<0,67)$ y sumar 6/10 de ella (lo que corresponde a las 6 milésimas de diferencia entre 0,66 y 0,666) al valor $P(Z<0,66)$.

Como $P(Z<0,67)-P(Z<0,66)=0,7486-0,7454=0,0032$ $\rightarrow \frac{6}{10}\cdot 0,0032=0,00192$, se asignará a $P(Z<0,666)$ el valor:

$$P(Z<0,666)=P(Z<0,66)+0,00192=0,7454+0,00192=0,74732.$$

d) Análogamente, para calcular $P(Z<1,863)$, se hallan los 3/10 de la diferencia $P(Z<1,87)-P(Z<1,86)=0,9693-0,9686=0,0007$, y se le suma a $P(Z<1,86)$.
Se obtiene:

$$P(Z<1,863)=P(Z<1,86)+\frac{3}{10}\cdot 0,0007=0,9686+0,00021=0,96881.$$

<u>Observación</u>: En la práctica, salvo en casos sencillos, y dada la escasa diferencia de los valores de probabilidad, no hay inconveniente en aproximar cada valor de *Z* a las centésimas. Así:

$$P(Z<0,666)\approx P(Z<0,67)=0,7486;\quad P(Z<1,863)\approx P(Z<1,86)=0,9686.$$

12. Utilizando la tabla normal *N*(0, 1), determina el valor de *k* que cumple:

a) $P(Z<k)=0,9115$ b) $P(Z<k)=0,9452$

c) $P(Z<k)=0,1587$ d) $P(Z<k)=0,95$

<u>Solución</u>:

a) El valor 0,9115 de la tabla se corresponde con *Z* = 1,35. Por tanto, *k* = 1,35.

b) El valor 0,9452 de la tabla se corresponde con *Z* = 1,6. Por tanto, *k* = 1,6.

c) Como 0,1587 es menor que 0,5 hay que buscar el valor de *Z* que deja por debajo 1 – 0,1587 = 0,8413. Ese valor es *Z* = 1. Por tanto, *k* = –1.

d) El valor 0,95 no aparece en la tabla. Como está entre 0,9495, correspondiente a *Z* = 1,64, y 0,9505, correspondiente a *Z* = 1,65, el valor de *k* buscado es *k* = 1,645.

13. Para una distribución normal *N*(50, 5), halla:

a) $P(X<56)$ b) $P(X>58)$ c) $P(X<48)$ d) $P(48<X<56)$

<u>Solución</u>:

En todos los casos hay que tipificar la variable: $Z=\frac{X-50}{5}$. Con esto:

a) $P(X<56)=P\left(Z<\frac{56-50}{5}\right)=P(Z<1,2)=0,8849$.

b) $P(X>56)=P\left(Z>\frac{58-50}{5}\right)=P(Z>1,6)=1-P(Z<1,6)=1-0,9452=0,0548$.

c) $P(X<48)=P\left(Z>\frac{48-50}{5}\right)=P(Z<-0,4)=1-P(Z<0,4)=1-0,6554=0,3446$.

d) $P(48 < X < 56) = P\left(\dfrac{48-50}{5} < Z < \dfrac{56-50}{5}\right) = P(-0,4 < Z < 1,2) =$

$= P(Z < 1,2) - P(Z < -0,4) = 0,8849 - 0,3446 = 0,5403$.

14. Supongamos que la estatura media de las alumnas de bachillerato se distribuye normalmente con media μ = 166 cm y desviación típica 9 cm. Si se elige una alumna al azar halla la probabilidad de que su estatura sea:

a) Superior a 175 cm. b) Inferior a 155 cm. c) Esté entre 155 cm y 175 cm.

Solución:

La normal de media μ y desviación típica σ, *N*(μ, σ), se tipifica mediante el cambio $Z = \dfrac{X-\mu}{\sigma}$, (en este caso, para μ = 166 y σ = 9 → $Z = \dfrac{X-166}{9}$), se tendrá:

a) $P(X > 175) = P\left(Z > \dfrac{175-166}{9}\right) = P(Z > 1) = 1 - P(Z < 1) = 1 - 0,8413 = 0,1587$.

b) $P(X < 155) = P\left(Z < \dfrac{155-166}{9}\right) = P(Z < -1,22) = 1 - P(Z < 1,22) = 1 - 0,8888 = 0,1112$.

c) $P(155 < X < 175) = P(X < 175) - P(X < 155) = 0,8413 - 0,1112 = 0,7301$.

15. Para una distribución normal *N*(60, 5), determina el valor de *k* que cumple:

a) $P(X < k) = 0,90$ b) $P(X > k) = 0,95$ c) $P(60-k < X < 60+k) = 0,9544$

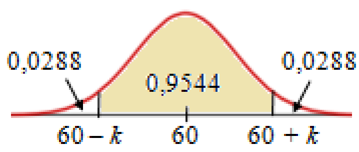
Solución:

En todos los casos hay que tipificar la variable: $Z = \dfrac{X-60}{5}$. Con esto:

a) $P(X < k) = 0,90 \Leftrightarrow P\left(Z < \dfrac{k-60}{5}\right) = 0,90 \Rightarrow \dfrac{k-60}{5} = 1,28 \Rightarrow k = 66,4$.

b) $P(X > k) = 0,95 \Leftrightarrow P\left(Z > \dfrac{k-60}{5}\right) = 0,95 \Rightarrow \dfrac{k-60}{5} = -1,645 \Rightarrow k = 51,775$.

c) $P(60-k < X < 60+k) = 0,9544 \Rightarrow$ la probabilidad que cae fuera de ese intervalo es 1 – 0,9544 = 0,0456; la mitad (0,0228) en la cola de la izquierda de la campana, la otra mitad en la cola derecha.



Por tanto, $P(X < 60+k) = 0,9544 + 0,0288 = 0,9772$.

Luego:

$$P(X < 60+k) = 0,9772 \Leftrightarrow P\left(Z < \dfrac{60+k-60}{5}\right) = 0,9772 \Rightarrow \dfrac{k}{5} = 2 \Rightarrow k = 10.$$

Esto es, en el intervalo (50, 70) = (60 – 2σ, 60 + 2σ) caen el 95,44 % de los valores de *X*, de una distribución *N*(60, 5).

16. Supongamos que los chicos de 15 años de un determinado país tienen una estatura que se distribuye según una normal de media 168 cm y desviación típica 12 cm. Si se quieren seleccionar al 5 % de los chicos más altos, ¿a partir de qué altura debe hacerse?

Solución:
Si X es la variable que describe la altura de los chicos, seleccionar uno entre el 5% de los más altos tiene una probabilidad de 0,05, es decir, $P(X > k) = 0,05$; o, lo que es lo mismo, $P(X < k) = 0,95$.

Como $P(X < k) = 0,95 \Leftrightarrow P\left(Z < \dfrac{k-168}{12}\right) = 0,95 \Rightarrow \dfrac{k-168}{12} = 1,645 \Rightarrow k = 187,74$.

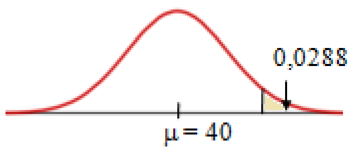
Luego, el 5 % de los chicos más altos miden más de 187,74 cm.

17. El diámetro de las ciruelas de una determina variedad se distribuye normalmente con media 4,5 cm y desviación típica 0,3 cm. Si se desea seleccionar, para su exportación, el 10% de las más grandes, ¿a partir de qué tamaño hay que cogerlas?
Solución:
La medida X de su diámetro se distribuye según la normal: $N(4,5,\ 0,3)$. Esta normal se tipifica haciendo el cambio $Z = \dfrac{X-4,5}{0,3}$.

Se desea encontrar el valor d (de diámetro) tal que $P(X > d) = 0,10 \Rightarrow$

$$P\left(Z > \frac{d-4,5}{0,3}\right) = 0,10 \Rightarrow \frac{d-4,5}{0,3} = 1,28 \Rightarrow d = 0,3 \cdot 1,28 + 4,5 = 4,884 \text{ cm.}$$

18. La edad de los habitantes de cierta ciudad se distribuye normalmente, con una media de 40 años. Se sabe además que el 2,28 % de los habitantes tiene más de 60 años.
a) ¿Cuál es la desviación típica?
b) ¿Cuál es el porcentaje de habitantes con menos de 35 años?
Solución:
La distribución de edad de la población es como se indica en la figura adjunta.



a) Se sabe que $P(X > 60) = 0,0228$.

Como la normal de media μ y desviación típica σ, $N(\mu, \sigma)$, se tipifica mediante el cambio $Z = \dfrac{X-\mu}{\sigma}$, (en nuestro caso, para μ = 40 y σ desconocida), se tendrá:

$$P(X > 60) = P\left(Z > \frac{60-40}{\sigma}\right) = 0,0228 \Leftrightarrow P\left(Z > \frac{20}{\sigma}\right) = 0,0228 \Rightarrow$$

$$\Rightarrow P\left(Z < \frac{20}{\sigma}\right) = 1 - 0,0228 = 0,9772 \Rightarrow \frac{20}{\sigma} = 2 \Rightarrow \sigma = 10.$$

Esto es, la desviación típica vale 10.

b) $P(X < 35) = P\left(Z < \dfrac{35-40}{10}\right) = P(Z < -0,5) = 1 - P(Z < 0,5) = 1 - 0,6915 = 0,3085$.

Esta probabilidad equivale al 30,85 %.

19. La duración de una determinada marca de lavadoras se ajusta a una normal de media 8,4 años y desviación típica 6 meses. El fabricante asegura que sus lavadoras duran más de 7 años, comprometiéndose a: “si una lavadora se estropea antes de 7 años le damos otra nueva”. ¿Cuántas lavadoras nuevas tendrá que reponer por cada 10000 vendidas?

Solución:
Para la normal $N(8,4,\ 0,5)$, la probabilidad de que X sea menor que 7 es:

$$P(X<7)=P\left(Z<\frac{7-8,4}{0,5}\right)=P(Z<-2,8)=1-P(Z<2,8)=1-0,9974=0,0026.$$

Tendrá que reponer 10000 · 0,0026 = 26 lavadoras.

20. Los envases de cartón de una determinada marca de leche contienen 1 litro de media, siendo la desviación típica de 5 ml.
a) ¿Qué porcentaje de envases sobrepasan los 1005 ml?
b) Si el control de calidad rechaza los envases que contengan menos de 990 ml y más de 1010 ml, ¿qué porcentaje de envases habrá que rechazar?
Solución:
El contenido de los envases se ajusta la normal $N(1000,\ 5)$. Se tipifica haciendo $Z=\dfrac{X-1000}{5}$.

a) $$P(X>1005)=P\left(Z>\frac{1005-1000}{5}\right)=P(Z>1)=1-P(Z<1)=1-0,8413=0,1587.$$

El 15,87% de los envases contiene más de 1005 ml.

b) Los envases que se aceptan son los que contienen entre 990 y 1010 ml.

$$P(990<X<1010)=P\left(\frac{990-1000}{5}Z<\frac{1010-1000}{5}\right)=P(-2<Z<2)=$$

$$=P(Z<2)-P(Z<-2)=0,9772-(1-0,9772)=0,9544\to 95,44\%$$

Hay que rechazar el 4,56 % de los envases.

21. Los ingresos anuales de los ejecutivos de una multinacional se distribuyen normalmente con media 45000 € y desviación típica de 3000 €. Si se elige un ejecutivo al azar se pide calcular las siguientes probabilidades:
a) De que sus ingresos anuales sean superiores a 50000 euros.
b) De que sus ingresos anuales estén comprendidos entre 42000 € y 46000 €.
c) De que sus ingresos anuales sean inferiores a 39000 euros.
d) Sabiendo que la probabilidad de que sus ingresos anuales sean superiores a una determinada cantidad es del 1%, ¿cuál es esa cantidad?
Solución:
Se trata de una distribución normal de media μ = 45000 y σ = 3000: $N(45000,\ 3000)$.
Se tipifica haciendo el cambio $Z=\dfrac{X-\mu}{\sigma}=\dfrac{X-45000}{3000}$.

a) $$P(X>50000)=P\left(Z>\frac{50000-45000}{3000}\right)=P(Z>1,67)=1-P(Z<1,67)=$$

= 1 – 0,9525 = 0,0475 → El 4,75% de los ejecutivos gana más de 50000 € anuales.

b) $$P(42000<X<46000)=P\left(\frac{42000-45000}{3000}<Z<\frac{46000-45000}{3000}\right)=P(-1<Z<0,33)=$$

$$=P(Z<0,33)-P(Z<-1)=P(Z<0,33)-(1-P(Z<1))=0,6293-(1-08413)=0,4706.$$

El 47,06% de los ejecutivos gana entre 42000 y 46000 € anuales.

c) $P(X<39000)=P\left(Z<\dfrac{39000-45000}{3000}\right)=P(Z<-2)=1-P(Z<2)$ =

= 1 – 0,9772 = 0,0228 → El 2,28% de los ejecutivos gana menos de 39000 € anuales.

d) Si $P(Z>c)=0,01$ ⇒ c = 2,33 ⇒ $\dfrac{X-45000}{3000}=2,33 \Rightarrow X=45000+3000\cdot 2,33=51990$ €.

Aproximación de la binomial mediante una normal (Repaso)

22. Mediante la aproximación normal de la binomial *B*(50, 0,12) calcula:

a) $P(X=6)$ b) $P(X=12)$ c) $P(6<X\leq 12)$

Solución:

La binomial *B*(50, 0,12) se puede aproximar por la normal de media y desviación típica:

μ = 50 · 0,12 = 6 y $\sigma=\sqrt{50\cdot 0,12\cdot 0,88}\approx 2,3$ → $X' = N(6,\ 2,3)$.

Con esto:

a) $P(X=6)=P(5,5<X'<6,5)=P\left(\dfrac{5,5-6}{2,3}<Z<\dfrac{6,5-6}{2,3}\right)\approx P(-0,22<Z<0,22)$ =

$=P(Z<0,22)-P(Z<-0,22)=0,5871-(1-0,5871)=0,1742$.

b) $P(X=12)=P(11,5<X'<12,5)=P\left(\dfrac{11,5-6}{2,3}<Z<\dfrac{12,5-6}{2,3}\right)=P(2,39<Z<2,83)$ =

$=P(Z<2,83)-P(Z<2,39)=0,9977-0,9916=0,0061$.

c) $P(6<X\leq 12)=P(5,5<X'<12,5)=P\left(\dfrac{5,5-6}{2,3}<Z<\dfrac{12,5-6}{2,3}\right)=P(-0,28<Z<2,83)$ =

= 0,9916 – (1 – 0,5871) = 0,5787.

23. El 42 % de los habitantes de un pueblo pasa cada día por la calle mayor. Elegidos 60 habitantes al azar, ¿qué probabilidad hay de que más de 30 de ellos pasen ese día por la calle mayor?

Solución:

La variable *X* que computa el número de habitantes que pasa por la calle mayor es una variable *B*(60, 0,42), que se aproxima por la normal *X´*: *N*(60 · 0,42, $\sqrt{60\cdot 0,42\cdot 0,58}$) = *N*(25,2, 3,82).

Haciendo la corrección de continuidad y tipificando, se tiene:

$P(X>30)=P(X'>30,5)=P\left(Z>\dfrac{30,5-25,2}{3,82}\right)=P(Z>1,39)=1-P(Z<1,39)$ =

= 1– 0,9177 = 0,0823.

Nota: También puede estudiarse mediante la "distribución muestral de la proporción".

Las muestras de tamaño *n* = 60 extraídas de una población con probabilidad *p* = 0,42, se comportan como una $N\left(0,42,\ \sqrt{\dfrac{0,42\cdot 0,58}{60}}\right)\approx N(0,42,\ 0,064)$.

Que, de una muestra de 60, más de 30… significa que $\hat{p}$ > 0,5. Luego:

$$P(\hat{p} > 0,5) = P\left(Z > \frac{0,5 - 0,42}{0,064}\right) = P(Z > 1,25) = 1 - P(Z < 1,25) = 1 - 0,8944 = 0,1056.$$

→ Puede observarse que los resultados varían: se diferencian en 0,0233. Hay dos motivos: el primero, que el resultado inicial es una "aproximación" de la binomial mediante una normal; el segundo es debido a los redondeos.

24. Un examen de respuesta múltiple consta de 80 preguntas, cada una con 4 opciones, una de ellas correcta y erróneas las otras tres. Si un estudiante contesta al azar, ¿cuál es la probabilidad de que acierte 25 o más preguntas? ¿Y menos de 10?
Solución:
El experimento es de tipo binomial, con *P*(éxito) = *p* = 0,25 y *q* = 0,75. Para *n* = 80, será *B*(80, 0,25).
La binomial *B*(80, 0,25) puede aproximare mediante la normal de media μ = 80 · 0,25 = 20 y $\sigma = \sqrt{80 \cdot 0,25 \cdot 0,75} = \sqrt{15} = 3,87 \to$ *N*(20, 3,87).
Con esto, haciendo la corrección de continuidad y tipificando:

$$P(X \geq 25) = P(X' > 24,5) = P\left(Z > \frac{24,5 - 20}{3,87}\right) = P(Z > 1,16) = 1 - P(Z < 1,16) =$$
$$= 1 - 0,8770 = 0,1230.$$

$$P(X < 10) = P(X' < 9,5) = P\left(Z < \frac{9,5 - 20}{3,87}\right) = P(Z < -2,71) = 1 - P(Z < 2,71) =$$
$$= 1 - 0,9966 = 0,0034.$$

Distribución de las medias muestrales: Intervalo de confianza; Error; Tamaño muestral.

25. (Selectividad, Baleares 2014). El cociente intelectual de unos universitarios se distribuye normalmente con una media de 100 y una desviación típica de 10.
a) Se elige una persona al azar. Busca la probabilidad de que su cociente intelectual se encuentre entre 98 y 103.
b) Se elige una muestra de veinticinco personas al azar. Busca la probabilidad de que la media de sus cocientes intelectuales se encuentre entre 98 y 103.
Solución:
a) La población es *N*(100, 10). Se tipifica haciendo el cambio $Z = \frac{x - 100}{10}$.
Luego:

$$P(98 < x < 103) = P\left(\frac{98 - 100}{10} < Z < \frac{103 - 100}{10}\right) = P(-0,2 < Z < 0,3) =$$
$$= P(Z < 0,3) - P(Z < -0,2) = 0,6179 - (1 - 0,5793) = 0,1972.$$

b) Las medias muestrales de tamaño *n*, obtenidas en una población *N*(μ. σ), se distribuyen según una normal $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.
En este caso: $N\left(100, \frac{10}{\sqrt{25}}\right) \to N(100, 2)$. Se tipifica haciendo el cambio $Z = \frac{\overline{X} - 100}{2}$.

Por tanto,

$$P\left(98<\bar{X}<103\right)=P\left(\frac{98-100}{2}Z<\frac{103-100}{2}\right)=P\left(-1<Z<1,5\right)=$$

$$=P\left(Z<1,5\right)-P\left(Z<-1\right)=0,9332-(1-0,8413)=0,7745.$$

26. (Selectividad, Canarias 2016). En un invernadero que se dedica a la producción de tomates, se ha comprobado que el peso de los tomates sigue una distribución normal con media 100 g y desviación típica 10 g. A la hora de comercializarlos se toman para la clase *A* los comprendidos entre 80 y 120 g. Hallar la probabilidad de que:
a) Elegido un tomate al azar, corresponda a la clase *A*.
b) Elegidos una docena de tomates al azar, su peso medio sea superior a 105 g.
Solución.

a) La población es $N(100, 10)$. Se tipifica haciendo el cambio $Z=\frac{x-100}{10}$.

Luego:

$$P\left(80<x<120\right)=P\left(\frac{80-100}{10}<Z<\frac{120-100}{10}\right)=P\left(-2<Z<2\right)=$$

$$=P\left(Z<2\right)-P\left(Z<-2\right)=0,9772-(1-0,9772)=0,9544.$$

b) Las medias muestrales de tamaño 12, obtenidas en una población $N(100, 10)$, se distribuyen según una normal $N\left(100,\ \frac{10}{\sqrt{12}}\right)\to N\left(100,\ 2,89\right)$. Se tipifica haciendo el cambio $Z=\frac{\bar{X}-100}{2,89}$.

Por tanto,

$$P\left(\bar{X}>105\right)=P\left(Z>\frac{105-100}{2,89}\right)=P\left(Z>1,73\right)=1-P\left(Z<1,73\right)=1-0,9582=0,0418.$$

27. Supongamos que la estatura media de las alumnas de 2º de bachillerato es de 165 cm, con desviación típica 8 cm.
a) Halla los parámetros de las medias muestrales de tamaño $n = 36$ y $n = 64$.
b) ¿Cuál es la probabilidad de que una muestra de 36 alumnas tenga una media de 167 o más cm? ¿Y de que una muestra de 64 alumnas supere esa misma medida?
c) ¿Tiene algo de extraño que una muestra de tamaño 36 dé una media de 170 cm?
Solución:
Los parámetros de la población valen $\mu = 165$ y $\sigma = 8$: $N(165, 8)$.
Los parámetros de la distribución de medias muestrales valen:

$\to$ media: $\bar{X}=\mu$; $\to$ desviación típica: $\sigma_{\bar{X}}=\frac{\sigma}{\sqrt{n}}$.

a) Para $n = 36$, las medias muestrales se distribuyen con media $\bar{X} = 165$ y desviación típica $\sigma_{\bar{X}}=\frac{8}{\sqrt{36}}=\frac{4}{3}$. Esto es: $N\left(165,\ \frac{4}{3}\right)$.

Para $n = 64$: $\bar{X} = 165$ y $\sigma_{\bar{X}}=\frac{8}{\sqrt{64}}=1$. Esto es: $N\left(165,\ 1\right)$.

b) Para n = 36: la $N\left(165,\ \frac{4}{3}\right)$ se tipifica mediante el cambio $Z=\frac{\bar{x}-165}{4/3}$. Por tanto,

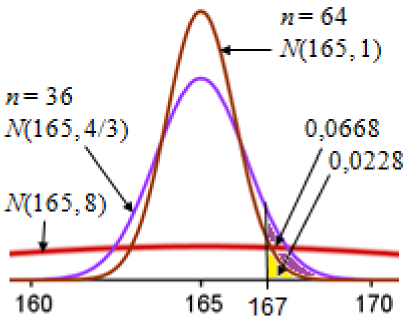
$$P\left(\bar{x}\geq 167\right)=P\left(Z\geq\frac{167-165}{4/3}\right)=P\left(Z\geq 1{,}5\right)=1-P\left(Z<1{,}5\right) = 1 - 0{,}9332 = 0{,}0668.$$

Esto significa que un 6,68 % de las muestras de tamaño 36 tendrá una media de estatura de 167 cm o más.

Para n = 64: la $N(165,\ 1)$ se tipifica mediante el cambio $Z=\frac{\bar{x}-165}{1}=\bar{x}-165$. Por tanto,

$$P\left(\bar{x}\geq 167\right)=P\left(Z\geq 167-165\right)=$$
$$= P\left(Z\geq 2\right)=1-P\left(Z<2\right) = 1 - 0{,}9772 = 0{,}0228.$$

Las distribuciones anteriores se muestran en la adjunta.



c) Para n = 36, $P\left(\bar{x}\geq 170\right)=P\left(Z\geq\frac{170-165}{4/3}\right)=P\left(Z\geq 3{,}75\right)=1-P\left(Z<3{,}75\right) < 0{,}0002.$

(Muchas tablas N(0, 1) solo alcanzan hasta Z = 3,49; indicando que $P\left(Z<3{,}49\right)=0{,}9998$).

Esto indica que menos de 2 de cada 10000 muestras superaría los 170 cm de media. Un resultado tan extraño no puede asumirse, aunque sea probable, pues o bien la muestra está mal elegida o no procede de la población femenina estudiada.

28. A pesar de acudir con "cita previa", el tiempo de espera de los pacientes de una clínica dental sigue una distribución normal de media desconocida y desviación típica 10 minutos. A partir de una muestra aleatoria de 144 pacientes, se obtuvo una media de espera es de 20 minutos.

a) Calcula los intervalos de confianza del 90%, 95% y 99% para la media del tiempo de espera de la población.

b) Explica la relación entre la amplitud de los intervalos y el nivel de confianza.

Solución:

a) El intervalo de confianza para la media poblacional es $IC=\left(\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$,

siendo: n el tamaño muestral, $\bar{x}$ la media muestral, σ la desviación típica y $Z_{\alpha/2}$ el valor correspondiente en la tabla normal para una confianza de $1-\alpha$.

En este caso, se sabe que: n = 144, $\bar{x}$ = 20 y σ = 10.

- Para un nivel de confianza del 90%, $Z_{\alpha/2}$ = 1,645; luego el intervalo de confianza será:

$$\left(20-1{,}645\cdot\frac{10}{\sqrt{144}},\ \ 20+1{,}645\cdot\frac{10}{\sqrt{144}}\right) = (20 - 1{,}37,\ 20 + 1{,}37) = (18{,}63,\ 21{,}37).$$

- Para un nivel de confianza del 95%, $Z_{\alpha/2}$ = 1,96.

El intervalo de confianza será:

$$\left(20-1{,}96\cdot\frac{10}{\sqrt{144}},\ 20+1{,}96\cdot\frac{10}{\sqrt{144}}\right) = (20 - 1{,}63,\ 20 + 1{,}63) = (18{,}37,\ 21{,}63).$$

- Para un nivel de confianza del 99%, $Z_{\alpha/2}$ = 2,575.

Por tanto, el intervalo de confianza será:

$$\left(20-2{,}575\cdot\frac{10}{\sqrt{144}},\ 20+2{,}575\cdot\frac{10}{\sqrt{144}}\right)$$ = (20 – 2,15, 20 + 2,15) = (17,85, 22,15).

b) La amplitud del intervalo es $2\cdot Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$, que está en relación directa con la desviación típica y con el nivel de confianza; y en relación inversa con la raíz cuadrada del tamaño muestral. Si se mantienen fijos σ y *n*, solo depende del nivel de confianza, de $Z_{\alpha/2}$.

Puede observarse que cuando hay menos exigencias en la confianza el intervalo se estrecha; en cambio, si se quiere tener más seguridad en el resultado basta con ampliar el intervalo. Hay menos riesgo en afirmar que la media está entre 17,85 y 22,15, que en decir que está entre 18,63 y 21,37. En el primer caso la confianza es del 99%; en el segundo, del 90%.

29. El número de *tweet* diarios generados por jóvenes de entre 16 y 24 años sigue una distribución normal de media desconocida y desviación típica 6 tweet. Tomada una muestra de 81 usuarios de Twitter, su media diaria ha resultado ser de 12 tweet. Calcula los intervalos de confianza del 80% y 85% para la media de la población.

Solución:

- Para un nivel de confianza del 80%, $Z_{\alpha/2}$ = 1,28; además: *n* = 81, $\bar{x}$ = 12, σ = 6.

El intervalo de confianza será:

$$IC=\left(12-1{,}28\cdot\frac{6}{\sqrt{81}},\ \ 12+1{,}28\cdot\frac{6}{\sqrt{81}}\right)$$ = (12 – 0,85, 12 + 0,85) = (11,15, 12,85).

- Para un nivel de confianza del 85%, $Z_{\alpha/2}$ = 1,44.

Por tanto, el intervalo de confianza será:

$$IC=\left(12-1{,}44\cdot\frac{6}{\sqrt{81}},\ \ 12+1{,}44\cdot\frac{6}{\sqrt{81}}\right)$$ = (12 – 0,96, 12 + 0,96) = (11,04, 12,96).

30. A lo largo de las diferentes pruebas de Selectividad se ha observado que la distribución de las calificaciones en el examen de Matemáticas siguen una ley normal de media 5,7 puntos y desviación típica 1,8.

a) ¿Cuál es la probabilidad de que la nota de un estudiante elegido al azar sea superior a 6,3?

b) ¿Cuál es la probabilidad de que una muestra de 49 alumnos tenga una media superior a 6,5 puntos?

Solución:

a) La población es *N*(5,7, 1,8). Se tipifica haciendo el cambio $Z=\frac{x-5{,}7}{1{,}8}$. Por tanto,

$$P(x>6{,}3)=P\left(Z>\frac{6{,}3-5{,}7}{1{,}8}\right)=P(Z>0{,}33)=1-P(Z<0{,}33)$$ = 1 – 0,6293 = 0,3707.

b) Las medias muestrales de tamaño *n* = 49 se distribuyen como una normal $N\left(5{,}7,\ \frac{1{,}8}{\sqrt{49}}\right)\to$ $N(5{,}7,\ 0{,}257)$. Se tipifica haciendo el cambio $Z=\frac{\bar{x}-5{,}7}{0{,}257}$. Por tanto,

$$P(\bar{x} > 6,5) = P\left(Z > \frac{6,5-5,7}{0,257}\right) = P(Z > 3,11) = 1 - P(Z < 3,11) = 1 - 0,9991 = 0,0009.$$

31. En el último año de Selectividad (2016), en una universidad se tomó una muestra aleatoria de tamaño 49, obteniéndose una nota media en el examen de Matemáticas de 5,5 puntos.
Admitiendo que la desviación típica sigue siendo de 1,8 puntos, se pide:
a) Un intervalo de confianza al 90 % para la nota media de la población.
b) Si para esa misma confianza se quiere un error máximo de 0,5 puntos, ¿cuál debe ser el tamaño de la muestra?
Solución:
a) Para el 90% de confianza (1 − α/2 = 0,9500), $Z_{\alpha/2}$ = 1,645. Luego:

$$IC = \left(5,5 - 1,645\cdot\frac{1,8}{\sqrt{49}},\ \ 5,5 + 1,645\cdot\frac{1,8}{\sqrt{49}}\right) = (5,5 - 0,423,\ 5,5 + 0,423) = (5,077,\ 5,923).$$

b) Como el error viene dado por $E = Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$, si se desea que $E < 0,5$, se tendrá:

$$E = 1,645\cdot\frac{1,8}{\sqrt{n}} < 0,5 \Rightarrow n > \frac{1,645^2\cdot 1,8^2}{0,5^2} = 35,07.$$

Debe tomarse una muestra de, al menos, 36 personas.

32. Se sabe que el número de horas que duermen los habitantes de una ciudad se puede aproximar mediante una distribución normal con desviación típica de 0,64 horas. Si se toma una muestra aleatoria de tamaño 16 y se pregunta por el número de horas que dedican a dormir se obtienen los siguientes datos:

6,5	8,5	6,5	8,5	7,5	7,0	5,5	7,5
7,5	6,5	7,0	7,0	6,5	8,0	7,5	6,0

a) Calcula la media muestral del número de horas que se duermen.
b) Halla el intervalo de confianza para la media de la población con una confianza del 95%
c) Si el que el intervalo de confianza para la media es (6,7, 7,48) horas, ¿con qué nivel de confianza se da?
Solución:
a) La media se halla sumando los 16 valores dados y dividiendo por 16.

$$\bar{x} = \frac{6,5+8,5+6,5+...+7,5+6}{16} = 7,09.$$

b) Como n = 16, σ = 0,64 y $Z_{\alpha/2}$ = 1,96, el intervalo de confianza de la media poblacional es:

$$\left(7,09 - 1,96\cdot\frac{0,64}{\sqrt{16}},\ \ 7,09 + 1,96\cdot\frac{0,64}{\sqrt{16}}\right) \approx (7,09 - 0,31,\ \ 7,09 + 0,31) = (6,78,\ \ 7,4).$$

c) La anchura del intervalo (6,7, 7,48) es 7,48 – 6,7 = 0,78 h. Este valor es igual $2\cdot Z_{\alpha/2}\dfrac{0,64}{\sqrt{16}}$.

Por tanto, $2\cdot Z_{\alpha/2}\cdot 0,16 = 0,78 \Rightarrow Z_{\alpha/2} = 2,4375$.
En la tabla normal, para $Z_{\alpha/2} = 2,4375 \approx 2,44$ se obtiene el valor 0,9927. Esto implica que α/2 = 1 – 0,9927 = 0,0073 ⇒ α = 0,0146.
La confianza es 1 – 0,0146 = 0,9854; esto es, del 98,54%.

33. (Selectividad Castilla la Mancha, 2016). Se estudió el cociente intelectual de 10 estudiantes de 2ª de Bachillerato elegidos aleatoriamente de un determinado centro escolar, siendo estos valores:

80, 96, 87, 104, 105, 99, 112, 89, 90, 100.

Sabiendo que el cociente intelectual se distribuye según una normal con desviación típica 15, se pide:

a) Halla el intervalo de confianza al nivel del 95 % para la media del cociente intelectual de los estudiantes de 2º de Bachillerato de dicho centro escolar.

b) Razona y explica qué se podría hacer para que el intervalo de confianza tuviera menor amplitud con el mismo nivel de confianza.

Solución:

a) La media de la muestra es:

$$\bar{x} = \frac{80+96+87+104+105+99+112+89+90+100}{10} = 96,2 .$$

Luego, el intervalo de confianza de la media poblacional será:

$$\left(96,2-1,96\cdot\frac{15}{\sqrt{10}},\ \ 96,2+1,96\cdot\frac{15}{\sqrt{10}}\right) \approx \left(96,2-9,3,\ \ 96,2+9,3\right) = \left(86,9,\ \ 105,5\right).$$

b) La amplitud del intervalo es $2\cdot Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$, que depende de la desviación típica, de la confianza y del tamaño muestral. En este caso, como σ = 15 no puede cambiarse, la amplitud puede disminuirse aumentando el tamaño muestral, n, o disminuyendo la confianza (el valor de $Z_{\alpha/2}$). Por tanto, si se quiere mantener la confianza, la única manera sería aumentar n.

34. El peso de los paquetes de espagueti de una determinada marca, sigue una distribución normal con desviación típica 20 gramos. Se seleccionan al azar 50 paquetes de esos espaguetis y se observa que tienen un peso medio de 745 gramos.

Halla el intervalo de confianza para el peso medio de los paquetes de espaguetis de esa marca con un nivel de confianza del 97 %.

Solución:

Para $\bar{x}$ = 745 g, σ = 20 g, n = 50 y, para el 97% de confianza ($1-\alpha/2 = 0,9850$), $Z_{\alpha/2}$ = 2,17, el intervalo de confianza de la media poblacional es:

$$\left(\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = \left(745-2,17\cdot\frac{20}{\sqrt{50}},\ 7450+2,17\cdot\frac{20}{\sqrt{50}}\right) =$$

= (745 – 6,14, 745 + 6,14) = (738,86, 751,14).

Es un intervalo de amplitud 12,28 gramos.

35. Se realiza una encuesta a 100 trabajadores menores de 25 años, sobre su sueldo mensual, obteniéndose una media de 900 € con una desviación típica de 120 €.

a) ¿Cuál es el intervalo de confianza para la media de ingresos de los trabajadores de ese sector, con un nivel de confianza del 92 %?

b) ¿Cuál debe ser el tamaño mínimo de la muestra necesario para estimar la media de ingresos mensuales con un error menor de 20 € y con una confianza del 94 %?

Solución:

a) Para $\bar{x}$ = 900, σ = 120, n = 100 y, para el 92 % de confianza, $Z_{\alpha/2}$ = 1,75, el intervalo será:

$$\left(\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = \left(900-1,75\cdot\frac{120}{\sqrt{100}},\ \ 900+1,75\cdot\frac{120}{\sqrt{100}}\right) =$$

$= (900 - 21,\ \ 900 + 21) = (879,\ \ 921)$.

La media de ingresos de los trabajadores menores de 25 años estará entre 879 y 921 €, con una confianza del 92 %.

b) La expresión que permite determinar n es $E = Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$, donde E es el error admitido.

Para una confianza del 94%, $Z_{\alpha/2}$ = 1,88, σ = 120 y E < 20, se tendrá:

$$1,88\cdot\frac{120}{\sqrt{n}} < 20 \Rightarrow n = \left(\frac{1,88\cdot 120}{20}\right)^2 = 127,2.$$

El tamaño muestral mínimo debe ser n = 128.

36. El gasto familiar medio en material escolar (por hijo) sigue una distribución normal con desviación típica σ = 50 €. A partir de una muestra de 100 personas se ha obtenido un gasto medio de 325 euros.

a) Halla un intervalo de confianza al 95 % para el gasto medio por hijo.

b) ¿Qué tamaño deberá tener la muestra para obtener un intervalo de confianza al 99% con una amplitud igual a la anterior?

Solución:

a) En este caso: $\bar{x}$ = 325, σ = 50, n = 100 y $Z_{\alpha/2}$ = 1,96. Por tanto, el intervalo de confianza pedido es,

$$\left(325 - 1,96\cdot\frac{50}{\sqrt{100}},\ 325 + 1,96\cdot\frac{50}{\sqrt{100}}\right) = (325 - 9,8,\ \ 325 + 9,8) = (315,2,\ \ 334,8).$$

b) En el intervalo hallado el error admitido es 9,8. Si se desea mantener el mismo error aumentando la confianza al 99%, el tamaño muestral debe aumentar también.

Para una confianza del 99%, $Z_{\alpha/2}$ = 2,575, σ = 50 y E = 9,8, se tendrá:

$$E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow 2,575\cdot\frac{50}{\sqrt{n}} = 9,8 \Rightarrow n = \left(\frac{2,575\cdot 50}{9,8}\right)^2 = 172,6.$$

El tamaño muestral debe ser 173 o superior.

37. (Selectividad, La Rioja 2012). La valoración de las instituciones por parte de los ciudadanos se mide en unas unidades ficticias que denominaremos “u”. Se sabe que, en el caso de los españoles, dicha valoración sigue una normal con desviación típica 25 u.

a) Se elige una muestra de 100 españoles, dando una media de 180 u. Calcula un intervalo de confianza para la media poblacional de la valoración de las instituciones, con una probabilidad del 90%.

b) Si se conoce que la media poblacional es 182 u, calcula la probabilidad de que una muestra de tamaño 100 tenga media inferior a 180 u.

Solución:

a) Para el 90% de confianza, $Z_{\alpha/2}$ = 1,645; como σ = 25, $\bar{x}$ = 180 y n = 100, el intervalo pedido será:

$$\left(180 - 1,645\cdot\frac{25}{\sqrt{100}},\ \ 180 + 1,645\cdot\frac{25}{\sqrt{100}}\right) \approx (180 - 4,1,\ \ 180 + 4,1) = (175,9,\ 184,1).$$

b) Las medias muestrales de tamaño 100, obtenidas en una población $N(182, 25)$, se distribuyen según la normal $N\left(182,\ \frac{25}{\sqrt{100}}\right) \to N(182,\ 2{,}5)$.

Esta normal se tipifica haciendo el cambio $Z = \frac{\overline{X} - 182}{2{,}5}$.

Por tanto,

$$P\left(\overline{X} < 180\right) = P\left(Z < \frac{180-182}{2{,}5}\right) = P(Z < -0{,}8) = 1 - P(Z < 0{,}8) = 1 - 0{,}7881 = 0{,}2119.$$

38. En la lotería navideña, una muestra aleatoria de 10 personas mayores de edad, jugaron las siguientes cantidades (en euros):

74, 72, 65, 75, 80, 81, 82, 84, 87, 90

Sabiendo que el gasto por persona, en la lotería navideña, sigue una distribución normal con desviación típica σ = 20 €, halla un intervalo de confianza para el gasto medio de la población con un nivel de confianza del 95,44 %.

Solución:

La media muestral es:

$$\overline{x} = \frac{74+72+65+75+80+81+82+84+87+90}{10} = 79.$$

Como σ = 20 euros, n = 10 y, para el 95,44% de confianza (1 − α/2 = 0,9772), $Z_{\alpha/2}$ = 2, se tiene:

$$IC = \left(79 - 2\cdot\frac{20}{\sqrt{10}},\ \ 79 + 2\cdot\frac{20}{\sqrt{10}}\right) = (79 - 12{,}65,\ 79 + 12{,}65) = (66{,}35,\ \ 91{,}65).$$

39. La edad de los atletas participantes en la última olimpiada seguía una distribución normal con desviación típica de 5 años. Una muestra aleatoria de 150 atletas dio como resultado una media de edad de 25,4 años.

a) Halla el intervalo de confianza del 94% para la media de edad de todos los atletas olímpicos.

b) ¿Cuál fue el tamaño mínimo de una la muestra si se exigió estimar la media con un nivel de confianza del 92% y con un error máximo de 0,5 años?

Solución:

a) Los datos son: $\overline{x}$ = 25,4, σ = 5, n = 150 y, para el 94% de confianza (1 − α/2 = 0,9700), $Z_{\alpha/2}$ = 1,88. Por tanto:

$$IC = \left(25{,}4 - 1{,}88\cdot\frac{5}{\sqrt{150}},\ \ 25{,}4 + 1{,}88\cdot\frac{5}{\sqrt{150}}\right) = (25{,}4 - 0{,}77,\ \ 25{,}4 + 0{,}77) = (24{,}63,\ \ 26{,}17).$$

b) El error admitido E, viene dado por $E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Para una confianza del 92%, $Z_{\alpha/2}$ = 1,75, σ = 5 y E < 0,5, se tendrá:

$$1{,}75\cdot\frac{5}{\sqrt{n}} < 0{,}5 \Rightarrow \sqrt{n} > \frac{1{,}75\cdot 5}{0{,}5} = 17{,}5 \Rightarrow n > 306{,}25.$$

El tamaño mínimo de la muestra fue de 307 atletas.

40. (Selectividad, Andalucía 2013). El tiempo que los españoles dedican a ver la televisión los domingos es una variable aleatoria que sigue una distribución Normal de media desconocida y desviación típica 75 minutos. Elegida una muestra aleatoria de españoles se ha obtenido, para la media de esa distribución, el intervalo de confianza (188,18, 208,82), con un nivel del 99 %.
a) Calcula la media muestral y el tamaño de la muestra.
b) Calcula el error máximo permitido si se hubiese utilizado una muestra de tamaño 500 y un nivel de confianza del 96 %.
Solución:
a) La media es el punto medio del intervalo de confianza (188,18, 208,82):

$$\bar{x} = \frac{188,18 + 208,82}{2} = 198,5.$$

La amplitud del intervalo es $2 \cdot Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ = 208,82 – 188,18 = 20,64.

Como σ = 75 y $Z_{\alpha/2}$ = 2,575 ⇒ $2 \cdot 2,575 \cdot \frac{75}{\sqrt{n}} = 20,64 \Rightarrow n = \left(\frac{386,25}{20,64}\right)^2 \approx 350$.

El tamaño de la muestra fue de 350 personas.

b) El error máximo es $E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$.

Para σ = 75, n = 500 y 1 − α = 0,96, $Z_{\alpha/2}$ = 2,05, se tendrá: $E = 2,05 \cdot \frac{75}{\sqrt{500}} = 6,88$.

Esto es, la media de la población será: μ = 198,5 ± 6,88.

41. (Selectividad, Cantabria 2014).
a) El tiempo diario que los estudiantes de bachillerato de Cantabria dedican al estudio en las dos semanas previas al inicio de los exámenes de Selectividad de la convocatoria de junio, sigue una distribución normal de media desconocida y desviación típica 15 minutos. Para estimar el tiempo medio se elije una muestra de 300 alumnos. ¿Con qué nivel de confianza debe realizarse la estimación si el error cometido es de 1,88 minutos?
b) Con vistas a la convocatoria de septiembre del mismo año se realiza un análisis similar. El tiempo diario que los estudiantes destinan al estudio las dos semanas anteriores al inicio de los exámenes, sigue una distribución normal con desviación típica 11 minutos. Con una muestra aleatoria de 150 alumnos se ha obtenido un tiempo medio de 173 minutos. Obtener el intervalo de confianza del 93% para el tiempo medio de estudio.
Solución:
a) El error admitido es $E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1,88$.

Como se sabe que σ = 15 y n = 300 ⇒ $Z_{\alpha/2} \frac{15}{\sqrt{300}} = 1,88 \Rightarrow Z_{\alpha/2} = \frac{1,88 \cdot \sqrt{300}}{15} = 2,17$.

Para $Z_{\alpha/2}$ = 2,17 ⇒ 1 – α/2 = 0,9850 ⇒ α = 0,03. La estimación se ha hecho con un nivel de confianza del 97%.

b) Para una confianza del 93%, 1 – α/2 = 0,9650 ⇒ $Z_{\alpha/2}$ = 1,81. También se sabe que σ = 11, $\bar{x}$ = 173 y n = 150. Por tanto, el intervalo pedido es,

$$IC = \left(173 - 1,81\cdot\frac{11}{\sqrt{150}},\ \ 173 + 1,81\cdot\frac{11}{\sqrt{150}}\right) = (173 - 1,63,\ \ 173 + 1,63) = (171,37,\ 174,63).$$

Distribución de la proporción:
Intervalo de confianza; Error, Tamaño muestral.

42. En unas elecciones a alcalde, el 56 % de los votantes optó por el candidato *A* mientras que el 44 % lo hizo por el candidato *B*.
a) Halla la distribución de probabilidad de las muestras de tamaño 50 extraídas de da población. Haz lo mismo para n = 100.
b) Calcula la probabilidad de que en una muestra de 50 votantes haya, al menos, 30 favorables al candidato *A*.
c) Si la muestra es de tamaño 100, ¿cuánto es la probabilidad de que una mayoría apoye al candidato *B*?
Solución:
a) La proporción de la población, para el candidato *A*, es p = 0,56; q = 0,44.
La proporción de las muestras de tamaño 50 se distribuye según la normal:

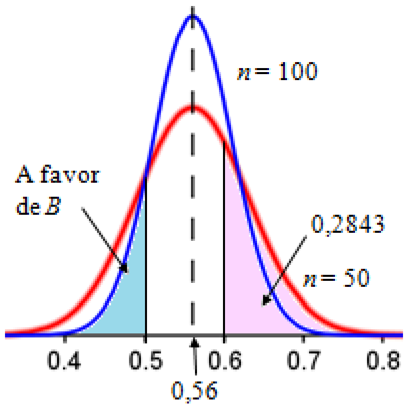
$$N\left(p,\ \sqrt{\frac{pq}{n}}\right) \to N\left(0,56,\ \ \sqrt{\frac{0,56\cdot 0,44}{50}}\right) \approx N(0,56,\ \ 0,07).$$

La proporción de las muestras de tamaño 100 será $N\left(0,56,\ \ \sqrt{\frac{0,56\cdot 0,44}{100}}\right) \approx N(0,56,\ \ 0,05)$.

b) 30 votantes a favor de *A*, entre 50, supone $\hat{p} = \frac{30}{50} = 0,6$; luego hay que calcular,

$$P(\hat{p} \geq 0,6) = P\left(Z \geq \frac{0,6 - 0,56}{0,07}\right) =$$

$$= P(Z \geq 0,57) = 1 - P(Z < 0,57) = 1 - 0,7157 = 0,2843.$$



c) La mayoría para el candidato *B* equivale a que la proporción a favor de *A* sea inferior a 0,5; esto es $\hat{p} < 0,5$, para la N(0,56, 0,05).
Esta probabilidad vale,

$$P(\hat{p} < 0,5) = P\left(Z < \frac{0,5 - 0,56}{0,05}\right) = P(Z < -1,2) = 1 - P(Z < 1,2) = 1 - 0,8849 = 0,1151.$$

43. De una muestra aleatoria de 120 graduados universitarios, 15 de ellos no han conseguido trabajo después de un año de su graduación.
a) Halla el intervalo de confianza, al 99 %, para estimar la proporción de graduados que sí han encontrado trabajo trascurrido un año de su graduación.
b) Para la misma confianza, ¿cuál debe ser el tamaño mínimo de la muestra para estimar la proporción de graduados con trabajo, admitiendo un error máximo del 5 %?
Solución:

La proporción muestral de graduados sin trabajo es: $\hat{q} = \frac{15}{120} = 0{,}125 \Rightarrow$ La proporción de graduados con trabajo es: $\hat{p} = 1 - 0{,}125 = 0{,}875$.

a) El intervalo de confianza para la proporción de la población es:

$$\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}},\ \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) \to IC = \left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right).$$

En este caso: $Z_{\alpha/2} = 2{,}575$; $\hat{p} = 0{,}875$, $\hat{q} = 0{,}125$, y $n = 120$.

Luego, el intervalo de confianza para estimar la proporción de graduados con trabajo será:

$$\left(0{,}875 - 2{,}575\cdot\sqrt{\frac{0{,}875\cdot 0{,}125}{120}},\ 0{,}875 + 2{,}575\cdot\sqrt{\frac{0{,}875\cdot 0{,}125}{120}}\right) =$$

$= (0{,}875 – 0{,}078,\ 0{,}875 + 0{,}078) = (0{,}797,\ 0{,}953)$.

b) El error admitido máximo es $E = Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$. Si se desea que $E < 0{,}05 \Rightarrow$

$$\Rightarrow 0{,}05 < 2{,}575\cdot\sqrt{\frac{0{,}875\cdot 0{,}125}{n}} \Rightarrow n > 2{,}575^2\frac{0{,}875\cdot 0{,}125}{0{,}05^2} \Rightarrow n > 290{,}1.$$

El tamaño mínimo debe ser 291 posgraduados.

44. Se quiere estimar la proporción de personas que esperan que su situación económica mejore el año próximo. Para ello se ha preguntado a 500 personas de esa población, de las 175 esperan que su situación económica mejore: son optimistas.
a) Calcula un intervalo de confianza para la proporción de personas optimistas en esta población, con un nivel de confianza del 94 %.
b) Si antes de conocer el resultado se quiere determinar un intervalo de confianza con el mismo nivel (94 %) y un error máximo de 0,02, ¿cuál debe ser el tamaño mínimo de la muestra?
Solución:

a) La proporción de optimistas en la muestra es $\hat{p} = \frac{175}{500} = 0{,}35$. Por tanto, debe estudiarse como una binomial $B(500,\ 0{,}35)$.
En este caso, para el 94% de confianza, $Z_{\alpha/2} = 1{,}88$; $\hat{p} = 0{,}35$, $\hat{q} = 0{,}65$ y $n = 500$, el intervalo de confianza para estimar la proporción de optimistas será:

$$\left(0{,}35 - 1{,}88\cdot\sqrt{\frac{0{,}35\cdot 0{,}65}{500}},\ 0{,}35 + 1{,}88\cdot\sqrt{\frac{0{,}35\cdot 0{,}65}{500}}\right) =$$

$= (0{,}35 – 0{,}04,\ 0{,}35 + 0{,}04) = (0{,}31,\ 0{,}39)$.

b) Si no se conoce la proporción de optimistas puede suponerse $p = q = 0{,}50$.

El error admitido E, viene dado por $E = Z_{\alpha/2}\sqrt{\frac{p\cdot q}{n}}$.

Si se desea que $E < 0{,}02 \Rightarrow 0{,}02 > 1{,}88\cdot\sqrt{\frac{0{,}5\cdot 0{,}5}{n}} \Rightarrow n > 1{,}88^2\frac{0{,}5\cdot 0{,}5}{0{,}02^2} \Rightarrow n > 2009$.

El tamaño mínimo debe ser 2010 personas.

45. Para estimar el grado de satisfacción de sus clientes, una compañía de reparto realiza una encuesta aleatoria entre 122 de sus clientes. De ellos, 103 declararon estar satisfechos.
a) ¿Cuál es la estimación de la proporción de clientes satisfechos?
b) Halla el intervalo de confianza al 99 % para la estimación de la proporción de clientes satisfechos.
Solución:

a) La proporción de clientes satisfechos es: $\hat{p}=\dfrac{103}{122}\approx 0,844$.

b) Para el 99% de confianza, $Z_{\alpha/2}=2,575$; $\hat{p}=0,844$ y $n=122$, se obtiene el intervalo:

$$IC=\left(0,844-2,575\cdot\sqrt{\frac{0,844\cdot 0,156}{122}},\ \ 0,844+2,575\cdot\sqrt{\frac{0,844\cdot 0,156}{122}}\right)=$$

$$=\left(0,844-0,085,\ \ 0,844+0,085\right)=\left(0,759,\ \ 0,929\right).$$

46. (Selectividad, Canarias 2014). En una zona escolar, para una muestra de 200 alumnos, 30 son repetidores.
a) Construir un intervalo de confianza con un nivel del 95 %, para estimar la proporción de alumnos repetidores.
b) Si se ignoran los datos iniciales y con un nivel de confianza del 90 %, ¿cuál es el tamaño mínimo muestral para estimar la proporción de alumnos repetidores con un error máximo del 2 %?
Solución:

a) La proporción de repetidores es: $\hat{p}=\dfrac{30}{200}=0,15$.

Para el 95% de confianza, $Z_{\alpha/2}=1,96$; $\hat{p}=0,15$, $\hat{q}=0,85$; y $n=200$, el intervalo de confianza para la proporción p de la población es:

$$IC=\left(0,15-1,96\cdot\sqrt{\frac{0,15\cdot 0,85}{200}},\ \ 0,15+1,96\cdot\sqrt{\frac{0,15\cdot 0,85}{200}}\right)=$$

$= (0,15 – 0,025,\ 0,15 + 0,025) = (0,125, 0,175)$.

b) Si se ignoran los datos anteriores, la mayor garantía se obtiene partiendo que de $p = q = 0,50$.

El error viene dado por $E=Z_{\alpha/2}\sqrt{\dfrac{p\cdot q}{n}}$. Si se desea que $E < 0,02$, como para el 90%, $Z_{\alpha/2}=1,645 \Rightarrow E=1,645\cdot\sqrt{\dfrac{0,5\cdot 0,5}{n}}<0,02\Rightarrow n>\left(\dfrac{1,645\cdot 0,5}{0,02}\right)^2=260$.

Debe tomarse una muestra de, al menos, 260 alumnos.

47. En una encuesta realizada a 70 jóvenes, 14 se declararon contrarios a la Unión Europea. Halla un intervalo de confianza del 92 % para determinar el porcentaje de la proporción de jóvenes que es contrario a la UE.
Solución:

La proporción muestral de jóvenes contrarios a la UE es: $\hat{p} = \dfrac{14}{70} = 0,2 \Rightarrow \hat{q} = 0,8$. Para una confianza del 92%, $Z_{\alpha/2}$ = 1,75. Por tanto, el intervalo de confianza para la proporción de la población es:

$$\left(0,2 - 1,75\cdot\sqrt{\frac{0,2\cdot 0,8}{70}},\ \ 0,2 + 1,75\cdot\sqrt{\frac{0,2\cdot 0,8}{70}}\right) = (0{,}2 – 0{,}048,\ \ 0{,}2 + 0{,}048) = (0{,}152,\ \ 0{,}248).$$

El porcentaje de jóvenes contrarios a la UE está entre el 15,2% y el 24,48%.

48. En una encuesta telefónica realizada a 125 personas, 35 de ellas declaran que el problema del paro juvenil es el más preocupante.
a) Halla un intervalo de confianza, con una significación α = 0,06, para obtener el porcentaje de la población que sitúan el problema del paro juvenil como el más preocupante.
b) Si la encuesta se hubiese realizado a 900 personas y se hubiese obtenido un valor de $\hat{p} = 0,28$, para la misma confianza, ¿qué margen de error se estaría asumiendo?
<u>Solución</u>:
a) La proporción muestral de personas preocupadas por el paro juvenil es $\hat{p} = \dfrac{35}{125} = 0,28 \Rightarrow \hat{q} = 0,72$. Para α = 0,06, 1 – α/2 = 0,97, luego $Z_{\alpha/2}$ = 1,88. Por tanto, el intervalo de confianza para la proporción de la población es:

$$\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}},\ \ \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) = \left(0,28 - 1,88\cdot\sqrt{\frac{0,28\cdot 0,72}{125}},\ \ 0,28 + 1,88\cdot\sqrt{\frac{0,28\cdot 0,72}{125}}\right) =$$

= (0,28 – 0,076, 0,28 + 0,076) = (0,204, 0,356).
El porcentaje oscila entre el 20,4% y el 35,6%.

b) El error es $E = Z_{\alpha/2}\sqrt{\dfrac{p\cdot q}{n}}$ = $1{,}88\sqrt{\dfrac{0,28\cdot 0,72}{900}} = 0,028$.
El error máximo es del 2,8 %.

49. (Selectividad, Murcia 2016). Para estimar la proporción de individuos de una población que utilizan el comercio electrónico se ha realizado una encuesta a una muestra aleatoria de 200 individuos, de los cuales 90 han respondido que utilizan el comercio electrónico. Con estos datos, hallar un intervalo de confianza del 95 % para la proporción de individuos de la población que utilizan el comercio electrónico.
<u>Solución</u>:
La proporción muestral es $\hat{p} = \dfrac{90}{200} = 0,45 \Rightarrow \hat{q} = 0,55$.
Para el 95% de confianza, $Z_{\alpha/2}$ = 1,96. Por tanto, el intervalo de confianza para la proporción de la población es:

$$\left(0,45 - 1,96\cdot\sqrt{\frac{0,45\cdot 0,55}{200}},\ \ 0,45 + 1,96\cdot\sqrt{\frac{0,45\cdot 0,55}{200}}\right) =$$

= (0,45 – 0,069, 0,45 + 0,069) = (0,381, 0,519).

Otros problemas

50. Una variable aleatoria se distribuye normalmente con desviación típica conocida σ. Halla el valor de σ sabiendo que a partir de una muestra aleatoria simple de tamaño 100, se ha obtenido una media de 25 y se dice que (24,02, 25,98) es el intervalo de confianza para la media poblacional, con un nivel de significación 0,05.
Solución:
El intervalo de confianza de la media poblacional, para las muestras de tamaño muestral n, de media $\bar{x}$ y desviación típica σ es $\left(\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$.
Para $\bar{x}$ = 25, n = 100 y, para el 95 % de confianza, $Z_{\alpha/2}$ = 1,96, el intervalo será:

$$\left(25-1{,}96\cdot\frac{\sigma}{\sqrt{100}},\ 25+1{,}96\cdot\frac{\sigma}{\sqrt{100}}\right)=(24{,}02,\ 25{,}98)\Rightarrow 1{,}96\cdot\frac{\sigma}{\sqrt{100}}=0{,}98\Rightarrow \sigma=5.$$

La desviación típica σ = 5.

51. (Selectividad, Madrid 2012). Se supone que el gasto que hacen los individuos de una determinada población en regalos de Navidad se puede aproximar por una variable aleatoria con distribución normal de media μ y desviación típica igual a 45 euros.
a) Se toma una muestra aleatoria simple y se obtiene el intervalo de confianza (251,6, 271,2) para μ, con un nivel de confianza del 95 %. Calcula la media muestral y el tamaño de la muestra elegida.
b) Se toma una muestra aleatoria simple de tamaño 64 para estimar μ. Calcula el error máximo cometido por esa estimación con un nivel de confianza del 90 %.
Solución:
a) La media es el punto medio del intervalo: $\bar{x}=\frac{251{,}6+271{,}2}{2}=261{,}6$.

La amplitud del intervalo es $2\cdot Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ = 271,2 – 251,6 = 19,6.

Como σ = 50 y $Z_{\alpha/2}$ = 1,96 ⇒ $2\cdot 1{,}96\cdot\frac{50}{\sqrt{n}}=19{,}6\Rightarrow n=\left(\frac{196}{19{,}6}\right)^2=100$.

b) El error máximo es $E=Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Para σ = 50, n = 64 y 1 − α = 0,90, $Z_{\alpha/2}$ = 1,645, se tendrá: $E=1{,}645\cdot\frac{50}{\sqrt{64}}=10{,}28$.

52. (Selectividad, Castilla La Mancha 2014). En un aeropuerto, el tiempo de espera de un viajero frente a la cinta transportadora hasta que sale su maleta sigue una distribución normal de media desconocida y desviación típica σ = 3 minutos. Se tomó una muestra aleatoria de 50 viajeros, y se observó que el tiempo medio de espera era de 17 minutos.
a) Halla un intervalo de confianza para la media poblacional del tiempo de espera de la maleta en ese aeropuerto con un nivel de confianza del 95 %.
b) ¿Se puede admitir que la media poblacional sea μ = 16 con un nivel de confianza del 95 %? ¿Cómo podríamos disminuir la amplitud del intervalo de confianza sin variar el nivel de confianza? Razona tus respuestas.
Solución:

a) El intervalo de confianza de la media poblacional es: $IC = \left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right)$.

Se conoce: $\bar{x}$ = 17, σ = 3, n = 50 y, para el 95% de confianza, $Z_{\alpha/2}$ = 1,96. Por tanto, el intervalo pedido es,

$$IC = \left(17 - 1,96\cdot\frac{3}{\sqrt{50}},\ 17 + 1,96\cdot\frac{3}{\sqrt{50}} \right) = (17 - 0,83,\ 17 + 0,83) = (16,17,\ 17,83).$$

b) Es muy improbable que la media poblacional sea μ = 16, pues queda fuera del intervalo de confianza. La probabilidad de que μ = 16 es inferior a 0,025.

La amplitud es $2\cdot Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. Si no se varia la confianza, se disminuye aumentando n.

<u>Observación</u>. Podría recordarse que las medias muestrales, de tamaño n, obtenidas en una población $N(\mu.\ \sigma)$, se distribuyen según una normal $N\left(\mu,\ \frac{\sigma}{\sqrt{n}} \right)$.

En este caso: $N\left(17,\ \frac{3}{\sqrt{50}} \right) \to N(17,\ 0,424)$. Se tipifica haciendo el cambio $Z = \frac{\bar{X} - 17}{0,424}$.

Por tanto,

$$P(\bar{X} \le 16) = P\left(Z < \frac{16-17}{0,424} \right) = P(Z < -2,35) = 1 - P(Z < 2,35) = 1 - 0,9906 = 0,0094.$$

53. El nivel medio de colesterol (en mg/dl) en individuos sanos depende de la edad y el sexo; para los hombres con menos de 21 años su distribución es normal con media μ = 160 y desviación típica σ = 10. Un nivel fuera de $\mu \pm 2\sigma$ resulta extraño: indica que puede haber alguna anomalía. Lo mismo cabe decir de las muestras: un nivel muestral fuera de $\mu \pm 2\sigma_{\bar{X}}$ resulta extraño.

a) ¿Cuál es el intervalo de probabilidad admisible (no extraño) para las muestras de tamaño: 1; 9 y 100?

b) ¿Qué porcentaje de individuos o muestras caen en los intervalos hallados?

<u>Solución</u>:

a) Para n = 1, las muestras son de individuos, y su nivel de colesterol es N(160, 10), de donde el intervalo pedido será:

(160 – 2 · 10, 160 + 2 · 10) = (140, 180).

→ Las muestras de tamaño 9 se distribuyen según la normal $N\left(160,\ \frac{10}{\sqrt{9}} \right) \approx N(160,\ 3,3)$.

El intervalo será:

$$(160 - 2\cdot 3,3,\ 160 + 2\cdot 3,3) = (153,4,\ 166,6).$$

→ Las muestras de tamaño 100 se ajustan a la normal $N\left(160,\ \frac{10}{\sqrt{100}} \right) \approx N(160,\ 1)$.

Siendo el intervalo:

$$(160 - 2\cdot 1,\ 160 + 2\cdot 1) = (158,\ 162).$$

b) El valor de probabilidad correspondiente a $Z_{\alpha/2}$ = 2 es 0,9544, luego el 95,44% de las muestras de tamaño 1, 9 o 100 de individuos sanos deben tener unos niveles de colesterol entre 150 y 170, entre 153,4 y 166,6 o entre 158 y 162, respectivamente.

54. (De un examen MIR). Se desea conocer la media de la colesterolemia basal de una población, con una seguridad del 95 % y una precisión de ± 4 mg/dl (error), y se tiene información por un estudio piloto o revisión bibliográfica de que la varianza es de 300 mg/dl. ¿Cuál debe ser el tamaño muestral?
Solución:
Para un nivel de confianza del 95%, $Z_{\alpha/2}$ = 1,96; el error E < 4. Se parte de que $\sigma^2 = 300$.
Como el error viene dado por $E = Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$. Si se desea que E < 4, se tendrá:

$$E = 1,96\cdot\frac{\sqrt{300}}{\sqrt{n}} < 4 \Rightarrow n > \frac{1,96^2\cdot 300}{4^2} = 72,03.$$

Debe tomarse una muestra de, al menos, 73 personas.

55. El perímetro torácico de los individuos adultos (hombres) en una población se distribuye normalmente con desviación típica σ = 6 cm. Si a partir de una muestra de tamaño n se afirma que el intervalo de confianza al 88% es (87, 91), ¿cuál fue el tamaño de la muestra estudiada?
Solución:
La amplitud el intervalo, que es 4 cm, viene dada por $2\cdot Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$.
Como para el 88% de confianza, $Z_{\alpha/2}$ = 1,555 (1 – α/2 = 0,94), se tiene

$$2\cdot 1,555\cdot\frac{6}{\sqrt{n}} = 4 \Rightarrow n = \left(\frac{2\cdot 1,555\cdot 6}{4}\right)^2 = 21,76.$$

El tamaño muestral debe ser de 22 hombres.

56. La ficha técnica de una encuesta indica:
Universo: Mayores de 18 años. *Ámbito*: Nacional. *Muestra*: 1000 entrevistas con un margen de error ±3,16 para datos globales, con un nivel de confianza del 95,5 % (dos sigma) y un p/q = 50/50. *Selección*: Estratificada, aleatoria. *Entrevista*; Telefónica. *Fecha de trabajo de campo*: del 16 al 18 de febrero de 2016. *Realización*: SIGMA DOS. *Dirección*: José Miguel de Elías.
Comprueba si el margen de error es el indicado
Solución:
El error viene dado por $E = Z_{\alpha/2}\sqrt{\dfrac{p\cdot q}{n}}$. Sustituyendo los valores indica dos:

$$E = 2\cdot\sqrt{\frac{0,5\cdot 0,5}{1000}} = 0,03162.... \to 3,162\ \%$$

Efectivamente, el error está dado correctamente.
Observaciones:
1. La indicación “dos sigma”, $Z_{\alpha/2}$ = 2, se corresponde con un nivel de confianza del 95,44%.
2. La indicación p/q = 50/50 expresa que inicialmente no se tienen datos de los valores de p y q; por eso, para tener dar más amplitud al intervalo se toma $p = q = 0,5$.
3. También se podría objetar sobre el método de las entrevistas telefónica: da una mayor agilidad a la recogida de datos, y es más barato; pero puede generar dudas sobre su fiabilidad.

57. Si a la empresa encuestadora le exigen un margen de error de ±3, ¿cuál debe ser el tamaño muestra? (Los demás parámetros permanecen iguales).
Solución:

Hay que resolver la inecuación $E = 2\cdot\sqrt{\frac{0,5\cdot 0,5}{n}} = 0,03 \Rightarrow n = \left(\frac{2\cdot 0,5}{0,03}\right)^2 = 1111,11$. Deberá encuestarse a 1112 personas.

Nota: En la ficha técnica de encuestas reales es frecuente un tamaño muestral de 1111.

58. (Selectividad, Canarias 2016). Un estudio sobre los kilogramos de residuos no minerales que genera cada español al año, ha dado, para una muestra de 100 personas, el intervalo de confianza [1470,6; 1529,4]. Si la desviación típica es de 150 kilogramos, suponiendo que la generación de residuos sigue una distribución normal:

a) ¿Cuál es la media muestral?

b) ¿Cuál es el nivel de confianza utilizado?

c) ¿Cuál sería el correspondiente intervalo con la misma información muestral, pero con un nivel de confianza igual a 0,9?

Solución:

a) La media es el punto medio del intervalo: $\bar{x} = \frac{1470,6 + 1529,4}{2} = 1500$ kg.

b) El error admitido es $E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$, que en este caso vale 29,4.

Como σ = 150 y n = 100 $\Rightarrow$ $29,4 = Z_{\alpha/2}\frac{150}{\sqrt{100}}$ $\Rightarrow$ $Z_{\alpha/2}$ = 1,96.

El nivel de confianza utilizado ha sido del 95%.

c) Para un nivel de confianza igual a 0,9, del 90 %, $Z_{\alpha/2}$ = 1,645. El intervalo pedido será:

$$\left(1500 - 1,645\cdot\frac{150}{\sqrt{100}},\ \ 1500 + 1,645\cdot\frac{150}{\sqrt{100}}\right) \approx$$

= (1500 – 24,7, 1500 + 24,7) = (1475,3, 1524,7).