Estadística inferencial

1. Introducción

La estadística inferencial trata de la elaboración de conclusiones para una población, partiendo de los resultados de una muestra y el del grado de fiabilidad de las conclusiones.

2. Muestreo

La teoría del muestreo estudia la relación entre una población y las muestras tomadas de ellas.

Consideramos los siguientes ejemplos:

a) Queremos saber el porcentaje de votantes que están a favor de un candidato en unas elecciones.

b) Para la fabricación de los tarros de envasado de espárragos se desea saber cuál es la longitud media de los espárragos.

c) Una determinada marca de chinchetas quiere saber cuál es la proporción de chinchetas buenas o defectuosas de cada caja.

En los tres ejemplos se trata de estudiar un conjunto de objetos muy grande, que llamaremos población. Estos elementos pueden ser personas, espárragos, chinchetas,…

En la práctica es imposible estudiar todos los elementos de la población, bien por ser muy numerosa o ser muy costoso el estudio. Por ello elegimos una parte de la población que llamaremos muestra y realizaremos el estudio sobre los elementos de la muestra para posteriormente inferir estos resultados sobre la población.

El proceso mediante el cual se extrae una muestra de una población se llama muestreo.

La característica más importante que debe tener una muestra es la representatividad, es decir que represente bien a la población.

Ejemplo de no representatividad: “ Las mujeres y el amor: Una revolución cultural” de Share Hite. Se enviaron 1.000.000 de cuestionarios a mujeres. Solamente 4.500 fueron contestados y devueltos. El trabajo se basó en estos cuestionarios.

Tipos de muestreo

a) Muestreo aleatorio simple.

Es el más sencillo. Todos los elementos de la población tienen la misma probabilidad de ser elegidos para la muestra.

Se parte de un listado de todos los elementos de la población a los que se asigna un número. Se introducen en una urna y luego se extraen tantos números como elementos vaya a tener la muestra.

También puede hacerse mediante una tabla de números aleatorios.

b) Muestreo aleatorio estratificado.

La población se divide en grupos homogéneos, llamados estratos, y posteriormente se extrae una muestra aleatoria simple de cada estrato.

Ejemplo: se tiene una población en la que el 60% son mujeres y el 40% hombres. Para escoger una muestra de 2000 personas se divide la población en estratos, hombres y mujeres, y se extraen al azar 1200 mujeres y 800 hombres.

c) Muestreo aleatorio sistemático.

Conocido el tamaño de la población (N) y de la muestra (n), dividir N entre n y el resultado K nos indica que hemos de seleccionar los elementos de la muestra de k en k. Se selecciona entonces un elemento de la población y a partir de él se seleccionan de k en k lo elementos siguientes.

d) Muestreo por conglomerados y áreas.

Se divide la población en distintas secciones o conglomerados. Se eligen al azar unas pocas de estas secciones y se toman todos los elementos de las secciones elegidas para formar la muestra.

Observación:

Se habla de un muestreo con reemplazamiento cuando cada elemento de la población tomado para la muestra, vuelve de nuevo a ella para poder volver a ser elegido. Cada elemento de la población puede ser seleccionado más de una vez. Este tipo de muestreo hace que una población finita pueda ser considerada, al menos teóricamente, como una población infinita ya que se puede tomar cualquier número de nuestras sin agotarla.

3. Teoría de muestras. Distribución en el muestreo.

- Parámetros poblacionales o parámetros. Característica de la población que es objeto de estudio. Son los índices centrales y de dispersión que definen la población.

- Estadísticos maestrales o estadísticos. Información del parámetro que contiene la muestra. Son los índices centrales y de dispersión que definen una muestra.

El estudio de determinadas características de una población se efectúa a través de diversas muestran que pueden extraerse de ella. Los estadísticos (media aritmética, mediana, desviación típica, …) obtenidos de las muestras van a permitir decidir sobre la aproximación apropiada del correspondiente parámetro de la población.

Ejemplo:

Se tira 50 veces cuatro dados equilibrados, apuntado cada vez los números que han salido. Después se calcula la media de cada una de las muestras de cuatro números: por ejemplo muestra 1: 1, 5, 6, 4 $\Rightarrow \bar{x}_1 = 4$

Así podíamos considerar una muestra de 50 muestras y sus respectivas medias, obteniendo un conjunto de 50 medias $\bar{x}_1, \bar{x}_2, \bar{x}_3, ..., \bar{x}_{50}$

Para solucionar los problemas anteriores es necesario conocer las relaciones existentes entre los parámetros muestrales y los parámetros poblacionales. Será necesario entonces conocer la distribución muestral de los estadísticos (si consideramos todas las posibles muestras de tamaño N de una población dada. Para cada muestra, podemos calcular un estadístico (media, desviación típica,…) que variará de una muestra a otra. De esta manera obtenemos una distribución del estadístico que se llama distribución de muestreo).

La distribución de un estimador al tomar muestras de tamaño n en la población se llama *distribución en el muestreo*.

Distribución en el muestreo de la media

Los fabricantes de tarros para el envasado de espárragos desean saber la longitud media de los espárragos. La longitud media poblacional la representaremos por μ ; y por σ la desviación típica poblacional.

Con el fin de hacernos una idea de cómo puede ser μ , elegiremos una muestra aleatoria formada por 40 espárragos, y se obtiene que:

- La longitud media de la muestra es: $\bar{x}_1 = 17{,}3\ cm$
- La desviación típica de la muestra es: $s_1 = 0{,}8\ cm$

Si elegimos otras muestras de tamaño 40 y calculamos sus medias y sus desviaciones típicas, obtendremos: $\bar{x}_1, \bar{x}_2, \bar{x}_3, ..., \bar{x}_n$ y $s_1, s_2, s_3, ..., s_n$.

Los distintos valores de $\bar{x}_i$ dan lugar a una variable aleatoria que representaremos por $\bar{X}$.

1) Si la población sigue una distribución normal $N(\mu, \sigma)$, la distribución de las medias muestrales sigue también una normal $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ donde n es el tamaño de las muestras.

2) Si la población no sigue una distribución normal se aplica el “Teorema Central del Límite”:

   “Si se toman muestras de tamaño $n > 30$ de una población, con una distribución cualquiera de media μ y una desviación típica σ , la distribución de las medias muestrales se aproxima a una distribución normal $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

$$\boxed{\mu_{\bar{x}} = \mu \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$$

Observación: Como la desviación de la media muestral es $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ su variabilidad disminuye al aumentar el tamaño n de la muestra, dicho de otra manera, la precisión del estimador es mayor al aumentar el tamaño de la muestra.

Ejemplo 1:

Una población está formada por sólo tres elementos, con valores 1, 3, 5. Consideramos todas las posibles muestras de tamaño dos con reemplazamiento que pueden formarse. Se pide calcular:

a) La media de la población.

b) La desviación típica de la población.

c) La media de la distribución muestral de las medias.

d) La desviación típica de la distribución muestral de las medias.

Solución:

Este ejemplo sirve para comprobar lo que ocurre con las medias de las diferentes muestras.

a) $\mu = \frac{\sum x_i f_i}{N} = \frac{1+3+5}{3} = 3$ b) $\sigma = \sqrt{\frac{\sum x_i^2 f_i}{N} - \bar{x}^2} = \sqrt{\frac{1+9+25}{3} - 3^2} = 1{,}63$

Muestra	(1,1)	(1,3)	(1,5)	(3,1)	(3,3)	(3,5)	(5,1)	(5,3)	(5,5)
Medias ($\bar{x}_i$)	1	2	3	2	3	4	3	4	5

$\bar{x}_i$	Nº de muestras f_i	$\bar{x}_i f_i$	$\bar{x}_i^2 f_i$
1	1	1	1
2	2	4	8
3	3	9	27
4	2	8	32
5	1	5	25
Σ	9	27	93

c) $\mu_{\bar{x}} = \frac{\sum \bar{x}_i f_i}{N} = \frac{27}{9} = 3$

d) $\sigma_{\bar{x}} = \sqrt{\frac{\sum \bar{x}_i^2 f_i}{N} - \mu_{\bar{x}}^2} = \sqrt{\frac{93}{9} - 3^2} = 1{,}154$

Cuando la población es infinita o las muestras se extraen con reemplazamiento:

$\mu_{\bar{x}} = \mu$ $\quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ $\quad \left(1{,}15 = \frac{1{,}63}{\sqrt{2}}\right)$

Ejemplo 2:

Los tornillos fabricados por cierta máquina de precisión, que se distribuyen según una normal, tienen un peso medio de 142,32 gramos y una desviación típica de 8,5 gramos. $N(142{,}8'5)$

a) Hallar la probabilidad de que una muestra elegida al azar de 25 tornillos, tomada entre ellos, tenga un peso medio superior a 144,6 gramos.

b) Realizar el mismo cálculo si la muestra es de 100 tormillos.

Solución:

a) Pesos medios: Distribución de la población $N(142'32{,}8'5)$

$\overline{X} : N\left(142'32,\ \dfrac{8'5}{5}\right) = N(142'32,\ 1'7)$ distribución muestral en media.

$$P(\overline{X} \geq 144{,}6) = P\left(\overline{Z} \geq \frac{144{,}6 - 142{,}32}{1{,}7}\right) = P(\overline{Z} \geq 1{,}34) = 1 - P(\overline{Z} \leq 1{,}34) = 1 - 0{,}9099 = 0{,}0901$$

b) $\overline{X} : N\left(142'32,\ \dfrac{8'5}{10}\right) = N(142'32,\ 0'85)$

$$P(\overline{X} \geq 144{,}6) = P\left(\overline{Z} \geq \frac{144{,}6 - 142{,}32}{0{,}85}\right) = P(\overline{Z} \geq 2{,}68) = 1 - P(\overline{Z} \leq 2{,}68) = 1 - 0{,}9963 = 0{,}0037$$

Podemos observar que el tamaño de la muestra influye en la probabilidad de obtener el peso medio ligeramente separado del peso medio de la población.

Ejemplo 3:

Las estaturas de 1200 alumnas de un centro de enseñanza superior se distribuyen normalmente con media 1,72 y desviación típica 0,9 m. Si se toman 100 muestras de 36 estudiantes cada una, se pide:

a) La media y desviación típica esperada de la distribución muestral de medias.

b) ¿En cuántas muestras cabría esperar una media entre 1,68 y 1,73 m?

c) ¿En cuántas muestras cabría esperar una media menor que 1,69 m?

Soluciones:

a) $\overline{X}:N\left(1'72,\ \frac{0'9}{\sqrt{36}}\right)\Rightarrow \mu_{\overline{x}}=1{,}72 \qquad \sigma_{\overline{x}}=\frac{0{,}9}{6}=0{,}15\ m$

b) $P(1'68\le \overline{x}\le 1'73)=P\left(\frac{1'68-1'72}{0'15}\le \overline{z}\le \frac{1'73-1'72}{0'15}\right)=P(-0{,}27\le \overline{z}\le 0{,}07)=$

$=0{,}6064+0{,}5279-1=0{,}1343$

0,1343.100=13,43 aproximadamente en 13 muestras.

c) $P(\overline{x}\le 1'69)=P\left(\overline{z}\le \frac{1{,}69-1{,}72}{0{,}15}\right)=P(\overline{z}\le -0{,}2)=1-P(\overline{z}\le 0{,}2)=1-0{,}5793=0{,}4207$

0,4207.100=42,07 aproximadamente en 42 muestras.

Distribución en el muestreo de una proporción *(no entra en selectividad)*

Las chinchetas de una determinada marca no salen todas buenas. Sea p la proporción de chinchetas buenas (éxito, fracaso)

No sabemos el valor de p, pero podemos aproximarnos de alguna manera. Para ello, tomaremos una muestra aleatoria de 100 chinchetas y observamos que 86 de ellas están bien.

Al valor 86/100 lo llamamos $\hat{p}$ = proporción en la muestra o proporción muestral. Si elegimos otras muestras de tamaño 100, evidentemente el valor de $\hat{p}$ varía. Los distintos valores de $\hat{p}$ dan lugar a una variable aleatoria $\hat{P}$ y que llamaremos estimador de p.

Se puede demostrar que esta variable aleatoria $\hat{P}$ cumple:

[redacted]

Para muestras de tamaño n>30, la distribución muestral de proporciones sigue una distribución normal

$$\boxed{\hat{P} : N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)}$$

Observación: En la práctica puede ocurrir que p sea desconocido. En este caso se aproxima por el de una muestra siempre que n>100.

Ejemplo 1:

Una población está formada por los elementos 1, 2, 4, 6.

a) Calcular la proporción p de cifras impares.

b) Para cada una de las muestras con reemplazamiento de tamaño dos, calcular la proporción de cifras impares.

c) Calcular la media y la desviación típica de la distribución muestral de proporciones.

Soluciones: a) p=1/4=0,25

Muestras	1,1	1,2	1,4	1,6	2,1	2,2	2,4	2,6	4,1	4,2	4,4	4,6	6,1	6,2	6,4	6,6
proporción	1	0,5	0,5	0,5	0,5	0	0	0	0,5	0	0	0	0,5	0	0	0

b)

$\hat{p}$	Nº muestras	$x_i f_i$	$x_i^2 f_i$
0	9	0	0
0,5	6	3	0,25.6
1	1	1	1
	16	4	2,5

$$\mu_{\hat{p}} = \frac{4}{16} = 0{,}25$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0{,}25.0{,}75}{2}} = 0{,}31$$

$$\sigma = \sqrt{\frac{2{,}5}{16} - 0{,}25^2} = 0{,}306$$

Ejemplo 2:

Una máquina fabrica piezas de precisión. En una producción habitual fabrica un 3% de piezas defectuosas. Un cliente recibe una caja de 500 piezas procedentes de la fábrica.

a) ¿Cuál es la probabilidad de que encuentre más del 5% de piezas defectuosas en la caja?

b) ¿Cuál es la probabilidad de que encuentre menos de un 1% de piezas defectuosas?

Soluciones:

a) p=0,03 n=500 $\hat{P}: N(0'03,\ 0'008)$

$\hat{P}$ = proporción de piezas defectuosas en las diferentes muestras.

$$P(\hat{P} \geq 0{,}05) = P\left(\hat{z} \geq \frac{0{,}05 - 0{,}03}{0{,}008}\right) = P(\hat{z} \geq 2{,}5) = 1 - P(\hat{z} \leq 2{,}5) = 1 - 0{,}9938 = 0{,}0062$$

b) $P(\hat{P} \leq 0{,}01) = P\left(\hat{z} \leq \frac{0{,}01 - 0{,}03}{0{,}008}\right) = P(\hat{z} \leq -2{,}5) = 1 - P(\hat{z} \leq 2{,}5) = 1 - 0{,}9938 = 0{,}0062$

Ejemplo 3:

El 3% de piezas producidas por una máquina son defectuosas. Se toma una muestra aleatoria de 100 piezas. ¿Cuál es la probabilidad de que en la muestra existan menos de 28 piezas defectuosas?.

Solución:

$\hat{P}$ = proporción de piezas defectuosas en las diferentes muestras. $\hat{P}: N(0'03,\ 0'0171)$

$$P\left(\hat{P} \leq \frac{28}{100}\right) = P\left(\hat{z} \leq \frac{0{,}28 - 0{,}03}{0{,}0171}\right) = P(\hat{z} \leq 14{,}62) \approx 1$$

Es decir, es prácticamente seguro, que en una muestra aleatoria de 100 piezas habría menos de 28 piezas defectuosas.

4. Estimación de parámetros: estimación por parámetros de confianza

La estimación por puntos se utiliza poco, puesto que carecemos de datos que nos indiquen el grado de fiabilidad del dato muestral que hemos tomado. Es mucho más interesante obtener un intervalo dentro del cual se tiene cierta confianza de que se encuentre el parámetro que tratamos de estimar.

Llamamos intervalo de confianza al intervalo al que, con cierta probabilidad, contenga al parámetro que estemos estimando. Se denota por I_c

Lamamos nivel de confianza a la probabilidad de que el intervalo de confianza contenga al verdadero valor del parámetro. Se denota por N_c.

Puesto en probabilidad se llama coeficiente de confianza

$$N_c = 95\%;\ coeficiente\ de\ \ confianza = 0{,}95\ se\ denota\ \ por\ 1-\alpha\ \ llamando\ \alpha = nivel\ de\ significación$$

Cuanto más amplio sea el intervalo de confianza determinado, más probable será que incluya el valor estimado y mayor será el nivel de confianza, pero mayor es también el

margen de error = amplitud del intervalo de confianza

Intervalo de confianza para la media poblacional

Nos centraremos en un ejemplo concreto

Ejemplo: Se sabe que las desviación típica de las tallas de los alumnos de una universidad es igual a 5 cm. Se desea estimar la talla media de dichos alumnos, para lo que se escoge una muestra de 100 estudiantes y se obtiene que la media muestral es 172 cm. Hallar un intervalo de confianza de la talla media de los alumnos de la universidad con niveles de confianza del 90% y del 99%.

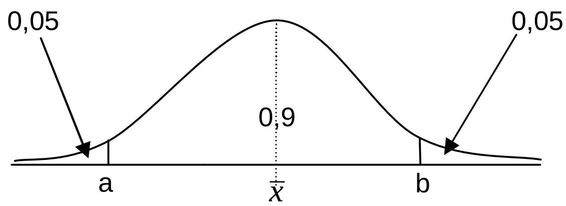
$\sigma = 5 \quad n = 100 \quad \bar{x} = 172$

Ya vimos que la distribución en el muestreo de la $\bar{X}$ es una normal $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$.

En este caso sigue una normal $N(172,\ 0'5)$.

- Para un $N_c = 90\%$ el coeficiente de confianza es 0’9.

Habrá que hallar un intervalo (a b) que cumpla $p(a \le \bar{x} \le b) = 0,9$ tipificando



$$p\left(\frac{a-172}{0,5} \le \bar{z} \le \frac{b-172}{0,5}\right) = 0,9$$

$$p(\bar{x} \le b) = 0,95 \Leftrightarrow p\left(\bar{z} \le \frac{b-172}{0,5}\right) = 0,95$$

consultando la tabla de la normal N(0,1) obtenemos:

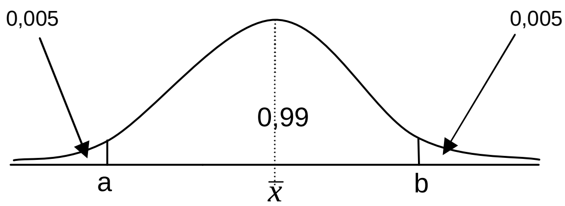
$$\begin{cases} p(\bar{z} \le 1,645) = 0,95 \\ p\left(\bar{z} \le \dfrac{b-172}{0,5}\right) = 0,95 \end{cases} \Leftrightarrow \frac{b-172}{0,5} = 1,645$$

$\Rightarrow b = 172 + 1,645.0,5 = 172,82$ por simetría o repitiendo el proceso:

$$a = 172 - 1,645 \cdot 0,5 = 171,18$$

Así el intervalo de confianza es I.C.=(171,18 172,82)

- Para un $N_c = 99\%$ el coeficiente de confianza es 0,99 y el intervalo de confianza buscado será IC=(a b) tal que $p(a \le \bar{x} \le b) = 0,99$ tipificando



$$p\left(\frac{a-172}{0,5} \le \bar{z} \le \frac{b-172}{0,5}\right) = 0,99$$

$$p(\bar{x} \le b) = 0,995 \Leftrightarrow p\left(\bar{z} \le \frac{b-172}{0,5}\right) = 0,995$$

consultando la tabla de la normal N(0,1) obtenemos:

$$\begin{cases} p(\bar{z} \le 2,585) = 0,995 \\ p\left(\bar{z} \le \dfrac{b-172}{0,5}\right) = 0,995 \end{cases} \Leftrightarrow \frac{b-172}{0,5} = 2,585$$

$\Rightarrow b = 172 + 2,585 \cdot 0,5 = 173,2875$ por simetría o repitiendo el proceso:

$$a = 172 - 2,585 \cdot 0,5 = 170,7125$$

Así el intervalo de confianza es I.C.=(170,71 173,29)

No vamos a repetir este proceso cada vez que queramos calcular un intervalo de confianza, podemos deducir que en general, a cada nivel de confianza le corresponde un z_c llamado <u>valor crítico</u> correspondiente a la distribución normal N(0, 1) y que cumple $p(-z_c \le \bar{z} \le z_c) = N_c$

En el ejemplo anterior: para $N_c = 0{,}90 \rightarrow z_c = 1{,}645$

para $N_c = 0{,}99 \rightarrow z_c = 2{,}585$

Entonces el intervalo de confianza para la media tiene la forma:

[redacted]

Otro ejemplo:

Se ha extraído una muestra de 145 alumnos de una escuela de artes, a los que se les ha propuesto un test de habilidad. La media y la desviación típica de la muestra son 82 y 14 respectivamente. A partir de estos datos, calcular el intervalo de confianza en el cual se hallará la media de la población al nivel de confianza del 95%. Lo mismo para el 99%.

Solución:

$\bar{x} = 82$ como no se conoce la desviación típica de la población tomamos la de la muestra

$\hat{s} = \sigma = 14$

n=145

$\bar{X} : N(82,\ 1'16)$

a) Nc=0,95 I.C.=(82-Zc.1'16 82+Zc.1'16)=(79,73 84,27)

$p(z \le z_c) = 0{,}975 \Rightarrow z_c = 1{,}96$

b) Nc=0,99 I.C.=(82- z_c .1'16 82+ z_c .1'16)=(79,006 84,99)

$p(z \le z_c) = 0{,}995 \Rightarrow z_c = 2{,}575$

Intervalo de confianza para la proporción poblacional (no entra en selectividad)

Ya vimos que la distribución en el muestreo de la proporción $\hat{P}$ sigue una distribución normal

$$N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

Siguiendo un argumento similar al apartado anterior:

[redacted]

Donde z_c es el valor crítico correspondiente al nivel de confianza prefijado.

Ejemplo 1:

Para estimar la proporción de estudiantes de una universidad que está a favor de la reinserción social del delincuente, se entrevistó aleatoriamente a 500 estudiantes. El 58% estaba a favor. Calcular el intervalo de confianza, al nivel de confianza del 95%, en el cual se hallará la población universitaria que se encuentra a favor.

Solución: P=0,58 n=500 Nc=0,95

$\hat{P} : N(0'58,\ 0'02)$

$p(\hat{p} \le z_c) = 0{,}975 \Rightarrow z_c = 1{,}96$ $\quad I.C. = (0{,}58 - z_c \cdot 0{,}02 \quad 0{,}58 + z_c \cdot 0{,}02) = (0{,}5408 \quad 0{,}6192)$

Ejemplo 2:

Se ha estudiado una muestra formada por 40 niños de 6 años y se ha observado que 15 de ellos dan positivo en una prueba de agresividad. Hallar el intervalo de confianza al nivel 95% para el parámetro proporción de positivos ante el test de agresividad para la población formada por todos los niños españoles de 6 años.

Solución: $p = \frac{15}{40} = 0{,}375$ n=40 Nc=0,95

$\hat{p} : N(0'37,\ 0'08)$ $\quad I.C. = (0{,}375 - z_c \cdot 0{,}08 \quad 0{,}375 + z_c \cdot 0{,}08) = (0{,}2182 \quad 0{,}5318)$

Al nivel de confianza del 95% la proporción de niños españoles de 6 años que dan positivo en la prueba d agresividad está entre el 21,82% y el 53,18%

5. Tamaño de las muestras.

Se denomina

- margen de error a la diferencia entre el extremo superior y el extremo inferior del intervalo de confianza.
- E=Error al radio del intervalo de confianza.

Hemos visto que una forma de aumentar el nivel de confianza es ampliar el tamaño del intervalo de confianza, pero esto tiene el inconveniente de que aumenta el margen de error.

Otra forma, es aumentar el tamaño de la muestra, ya que el ancho del intervalo depende de n.

Ejemplo 1. (para medias)

Se desea hacer una estimación sobre la edad media de una determinada población. Calcular el tamaño de la muestra para poder hacer dicha estimación con un error menor de medio año a un nivel de confianza del 99,73%. Se conoce de estudios previos que la edad media de dica población tiene una desviación típica de 3 años.

Solución:

E=0,5

n=?

z_c=99,73%

$$I.C.=\left(\mu - z_c\cdot\frac{\sigma}{\sqrt{n}},\ \mu + z_c\cdot\frac{\sigma}{\sqrt{n}}\right)$$

Teniendo en cuenta que el error es el radio del intervalo:

$$0{,}5 = z_c\cdot\frac{3}{\sqrt{n}}$$

Hallamos el valor crítico: $p(z \le z_c) = 0{,}99865 \ \Rightarrow \ z_c = 2{,}995$

De ambas expresiones: $\sqrt{n} = \dfrac{2{,}995\cdot 3}{0{,}5}$ $\qquad n = 322{,}92 \cong 323$

La muestra debe estar compuesta por 323 personas.

Ejemplo 2: (para proporciones) ***(no entra en selectividad)***

Una empresa dedicada a la venta de palomitas compra el maíz directamente a los agricultores. Antes de efectuar la compra, un agente de la compañía quiere estimar la probabilidad p de que el gramo de maíz se abra al freírlo. ¿Cuántos gramos deberá examinar para estar seguro al nivel del 90% de que el error máximo que cometa es 0,01? Se ha realizado un estudio sobre una pequeña muestra de 60 gramos, en la que obtuvo 48 que se abrían.

Solución:

P=48/60=0,8

E=0,01

Nc=0,90

$$I.C. = \left(p - z_c \cdot \sqrt{\frac{p(1-p)}{n}} \quad p + z_c \cdot \sqrt{\frac{p(1-p)}{n}} \right)$$

$$0{,}01 = z_c \cdot \sqrt{\frac{p(1-p)}{n}} \qquad 0{,}01 = 1{,}645 \cdot \frac{0{,}4}{\sqrt{n}} \quad \Rightarrow \quad n = \left(\frac{1{,}645 \cdot 0{,}4}{0{,}01} \right)^2 = 432964$$

$$p(z \le z_c) = 0{,}95 \;\Rightarrow\; z_c = 1{,}645$$

Ejercicios:

Distribuciones en el muestreo

1.- En una urna hay 3 bolas con los números 1, 2 y 3.

a) Calcular la media y la desviación típica de esta población

b) Formar todas las muestras posibles que podemos extraer con devolución de esta población de tamaño 2.

c) Formar la distribución de las medias de las muestras y hallar la media y la desviación típica de esta distribución.

2.- Consideramos la población formada por los cinco elementos 0, 3, 4, 6, 8. Hallar:

a) El número de muestras aleatorias de tamaño 2 con devolución.

b) La media y la desviación típica poblacionales.

c) La distribución muestral de las medias muestrales.

d) La media y la desviación típica de las medias muestrales.

Distribuciones en media

3.- Las notas de un grupo de alumnos de aproximadamente normal con media $\mu=5{,}5$ y desviación típica $\sigma=0{,}8$.

a) Hallar la media y la desviación típica de las medias muestrales para muestras de tamaño 4.

b) Calcular la probabilidad de que una media muestral de 4 alumnos elegidos al azar sea mayor que 5,2.

4.- Se sabe que los niño españoles de enseñanza primaria ante una prueba de discriminación visual se distribuyen según una N(4, 2). Extraemos una muestra aleatoria formada por 39 niños y les pasamos la prueba. Hallar la probabilidad de que la media muestral:

a) Sea menor que 3,5

b) Sea mayor que 3,9

c) Esté comprendido entre3,8 y 4,1.

5.- En una universidad se sabe que las tallas de los alumnos se distribuyen normalmente con media 172 cm y desviación tipica 17,5 cm. Se toman muchas muestras de 35 estudiantes.

a) ¿Cuál es la media y la desviación típica de la distribución de las medias muestrales?

b) Hallar la probabilidad de que la media muestral sea inferior a 171 cm.

c) Si se eligen 150 muestras de 35 alumnos ¿en cuántas de ellas cabe esperar que la media muestral sea mayor que 170 cm y menor que 171,5 cm?

6.-El peso de los toros de una determinada ganadería se distribuye según una normal de media 500 kg y 45 kg de desviación típica. Se toman muestras de 35 toros y se calcula el peso medio. Hallar la probabilidad de que la media muestral:

a) Sea mayor de 540 kg.

b) Sea menor de 480 kg.

c) Esté entre los 480 kg y 495 kg.

7.- El peso de las truchas de una piscifactoría sigue una ley N(22, 50). Se toman muestras de 60 truchas y se calcula su peso medio. Hallar las probabilidades de que la media muestral.

a) Sea mayor que 210 gr

b) Sea menor que 185 gr

c) Esté entre 210 y225 gr.

Intervalos de confianza

8.- Varios test de inteligencia dieron una puntuación que sigue una ley normal con media 100 y desviación típica…………..