

1.- Encuestados 169 conductores de la comunidad valenciana se obtuvo que los kilómetros anuales que habían recorrido por término medio fueron 14000, siendo la varianza de 810000 Km². Estimar el kilometraje medio anual de los conductores valencianos con un nivel de confianza del 95%. Supóngase normalidad.

Como el tamaño de la población es muy grande (varios millones) podemos considerar que la muestra es M.A.S. de tamaño n=169 . Por otro lado como la muestra es grande (n> 30) podemos considerar el intervalo para la media con varianza poblacional conocida haciendo $S = \sigma$:

$$\mu \in \left[\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] \quad \text{con } 1 - \alpha \text{ de confianza}$$

$\mu \in \left[14000 \pm 1.96 \frac{900}{\sqrt{169}} \right]$ con 95% de confianza esto es el kilometraje recorrido medio por conductor será de : [13864'31;;14135'69] Km con un 95% de confianza .

2.- El numero total de alumnos matriculados en 2º de ADE es de 250 (100 varones y 150 mujeres).Se ha realizado un muestreo aleatorio entre ellos resultando los siguientes datos:

Alumno	Sexo	Edad	Peso	2A -Suponiendo que los pesos tanto de los alumnos como de las alumnas de 2º de Empresariales siguen una distribución normal, estimar el peso medio de las alumnas (chicas)de Empresariales, con una confianza del 95 % 2B.- Con los mismos supuestos estimar el peso medio de los chicos, con la misma confianza 2C-Si consideramos la muestra anterior como representativa de la población juvenil nacional obtener el intervalo de confianza del 99 % para el peso medio de los jóvenes españoles (sin distinción de sexo)
1	varón	20	72	
2	varón	21	76	
3	varón	22	82	
4	varón	25	79	
5	varón	19	87	
6	mujer	23	72	
7	mujer	22	68	
8	mujer	20	59	
9	mujer	21	62	
10	mujer	22	64	
11	mujer	21	70	
12	mujer	19	67	

El análisis descriptivo arroja como resultado los siguientes estadísticos:

	Edad(todos)	Edad(chicas)	Edad(chicos)	Peso(todos)	Peso(chicas)	Peso(chicos)
número	12	7	5	12	7	5
suma	255	148	107	858	462	396
media	21,25	21,14	21,4	71,5	66	79,2
Des.Tip	1,64	1,25	2,06	7,98	4,24	5,11

2.a) Se trata de la estimación de la media de una normal con varianza desconocida (muestra pequeña , n=7,)(intervalo t-student) y aplicando el factor corrector de poblaciones finitas (M.I.) ya que la población también es pequeña.

$$\mu \in \left[\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n-1}} \sqrt{\frac{N-n}{N-1}} \right]$$

La media del peso de las chicas es 66 Kg y la desviación típica 4,24 , el tamaño de la población es ahora el número de chicas N=150,y el valor crítico para alfa=0.05 de la tabla es $t_{\alpha/2}(6 \text{ g.l.}) = 2,447$ sustituyendo:

$$\mu \in \left[66 \pm 2,447 \frac{4,24}{\sqrt{6}} \sqrt{\frac{150-7}{149}} \right] \Rightarrow \mu \in [61.85, 70.15] \text{ Kg. con el 95\% de confianza}$$

2b) Igual que en el apartado anterior, se trata de la estimación de la media de una normal con varianza desconocida (muestra pequeña , n=5,)(intervalo t-student) y aplicando el factor corrector de poblaciones finitas (M.I.) ya que la población también es pequeña

$$\mu \in \left[\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n-1}} \sqrt{\frac{N-n}{N-1}} \right]$$

La media del peso de los chicos es 79.2 Kg y la desviación típica 5.11, el tamaño de la población es ahora el número de chicos N=100,y el valor crítico para alfa=0.05 de la tabla es $t_{\alpha/2}(4 \text{ g.l.}) = 2,776$ sustituyendo:

$$\mu \in \left[79.2 \pm 2,776 \frac{5,11}{\sqrt{4}} \sqrt{\frac{100-5}{99}} \right] \Rightarrow \mu \in [72.25, 86.15] \text{ Kg. con el 95\% de confianza}$$

2c) El peso de los jóvenes del país puede considerarse una variable normal, la población es muy grande por lo que se puede considerar M.A.S.; la muestra , en cambio, es pequeña así que tomamos el intervalo para varianza poblacional desconocida:

$$\mu \in \left[\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n-1}} \right]$$

Donde el peso medio es 71.5 Kg. con una desviación típica de 7.98 Kg, el tamaño muestral es n=12 y como alfa=0.01 el valor de la t de student, $t_{\alpha/2}(11 \text{ g.l.}) = 3.106$ Sustituyendo:

$$\mu \in \left[71.5 \pm 3,106 \frac{7,98}{\sqrt{11}} \right] \Rightarrow \mu \in [64.02, 78.98] \text{ Kg. con el 99\% de confianza}$$

Habría que comentar que la muestra difícilmente puede considerarse representativa de la juventud española.

3.- Elegidos al azar 26 alumnos de la Universidad se les preguntó por el número de horas diarias que dedicaban a las (mal llamadas) redes sociales. Resultando que por término medio dedicaban 1.5 horas con una desviación típica de 0.25. Suponiendo normalidad en el número de horas que los alumnos de la Universidad dedican a cultivar sus e-contactos, estimar el tiempo medio que un alumno de la Universidad emplea en este “menester” con una confianza del 99%.

Se trata de estimar la media de una población normal, en la que, dado que no conocemos el tamaño de la población (a pesar de que presumiblemente el muestreo ha sido sin reposición), no podemos hacer otra cosa que suponer que hay varias decenas de miles de alumnos en la Universidad y asumir que el muestreo lo es de una población “infinita” (como si fuera M.A.S.)

Como la muestra es pequeña $n < 30$, tampoco podemos asumir $S \equiv \sigma$, y por lo tanto habrá que considerar el intervalo t de Student.

$$\mu \in \left[\bar{x} \pm t_{\alpha/2} \frac{S}{\sqrt{n-1}} \right] \text{ con } 1 - \alpha \text{ de confianza}$$

como: $\bar{x} = 1.5$, $S = 0.25$, $n = 26$, $t_{0.005}(24 \text{ g.l.}) = 2.787$ obtenemos un intervalo:

$$\mu \in \left[1.5 \pm 2.787 \frac{0.25}{5} \right] \text{ con } 99\% \text{ de confianza esto es:}$$

$\mu \in [1.5 \pm 0.13935]$ con 99% de confianza (es decir entre 1.36 y 1.64 horas diarias por término medio con un 99% de confianza.)

4.- Dada una estimación de la media de una población normal realizada con una confianza del 95% queremos mejorarla aumentando el nivel de confianza al 99 %
a) Manteniendo el error de estimación, ¿Cuánto habrá que aumentar el tamaño muestral para conseguirlo? .

b) Si mantuviéramos el tamaño muestral ¿cómo aumentaría el error de la estimación?

a)

En la estimación de la media de una población normal tenemos que:

$$\mu = \left[\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \quad \text{por lo que el error de estimación es : } E = \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

con un nivel de confianza del 95% $z_{\alpha/2} = 1.96$

con un nivel de confianza del 99 % $z_{\alpha/2} = 2.58$

Si en la segunda estimación el error debe mantenerse tendremos que: $E_1 = E_2$

$$E_1 = \pm 1.96 \frac{\sigma}{\sqrt{n_1}} = E_2 = \pm 2.58 \frac{\sigma}{\sqrt{n_2}} \text{ por lo que } n_2 = \frac{2.58^2 \times n_1}{1.96^2} = 1.7327 n_1$$

b) En cambio, si mantenemos el tamaño muestral al aumentar la confianza el nuevo error pasará a ser:

$$E_2 = \pm Z_{0.005} \frac{\sigma}{\sqrt{n}} = \pm 2.58 \frac{\sigma}{\sqrt{n}} \text{ Mientras que error inicial era : } E_1 = \pm Z_{0.025} \frac{\sigma}{\sqrt{n}} = \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

Por lo que la relación entre el segundo error y el primero será

$$E_2 = \frac{2.58}{1.96} E_1 = 1.3163 E_1 \text{ aprox.}$$

5.- Dada una estimación de una proporción, ¿en cuánto hay que aumentar el tamaño muestral para reducir el error a la cuarta parte, manteniendo el mismo nivel de confianza?

La estimación de una proporción p , para un nivel de confianza dado, $1 - \alpha$ será:

$$p \in \left[\hat{p} \pm z_{\alpha/2} \frac{0.5}{\sqrt{n}} \right] \text{ con } 1 - \alpha \text{ de confianza}$$

por lo que el error de estimación será: $E = \pm z_{\alpha/2} \frac{0.5}{\sqrt{n}}$ por lo tanto si queremos reducir a

la cuarta parte este error, el nuevo error de estimación, E_2 , deberá ser: $E_2 = E/4$, y por otra parte, manteniendo el mismo nivel de confianza, E_2 deberá obedecer a la

expresión: $E_2 = \pm z_{\alpha/2} \frac{0.5}{\sqrt{n_2}}$ donde n_2 será el nuevo tamaño muestral necesario para

conseguirlo, por lo que:

$$E_2 = \pm z_{\alpha/2} \frac{0.5}{\sqrt{n_2}} = \frac{E}{4} = \frac{\pm z_{\alpha/2} \frac{0.5}{\sqrt{n}}}{4} \Rightarrow \frac{1}{\sqrt{n_2}} = \frac{1}{4\sqrt{n}} \Rightarrow n_2 = 16n$$

Esto es; habrá que multiplicar por 16 el tamaño muestral para reducir a la cuarta parte el error, manteniendo el mismo nivel de confianza

6.- Se desea determinar un intervalo de confianza con nivel de confianza del 99% para la proporción de ciudadanos que votarán NO en un próximo referéndum. Si se sabe que en una muestra aleatoria de 400 ciudadanos 100 afirmaron su intención de votar NO, obtener dicho intervalo

El intervalo de confianza para una proporción es

$$p \in \left[\hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}} \right] \text{ con un nivel de confianza de } 1 - \alpha$$

Que asumiendo el caso de máxima holgura con $p=q=0.5$, para un nivel de del 99% y considerando que : $\hat{p} = \frac{100}{400} = 0.25$ $z_{0.005} = 2.576$ $n = 400$ resultaría:

$$p \in \left[0.25 \pm 2.576 \frac{0.5}{\sqrt{400}} \right] \text{ con un 99\% de confianza} \Rightarrow$$

$$p \in [0.25 \pm 0.064] \text{ con un 99\% de confianza}$$

$$p \in [0.1856, 0.3144] \text{ con un 99\% de confianza}$$

7.- Para la estimación de la proporción de familias con ingresos superiores a 80000 Euros al año, se han realizado dos muestreos distintos, en ambos el tamaño muestral es el mismo , así como la forma de muestrear ; en ambos , también, el nivel de confianza es idéntico (95,5%) . En la ficha técnica del muestreo A se nos indica que $p=q=0,5$. En el muestreo B se nos indica que se utiliza como p la proporción de familias con ingresos superiores a 80000 euros que se obtuvo en un sondeo anterior. Nos preguntamos por: ¿Cuál de los dos muestreo nos dará un intervalo para dicha proporción de familias con menor amplitud ? ¿Por qué ? ¿Cuál de los dos muestreos es más riguroso?

En el primer caso la amplitud del intervalo de estimación , al considerar $p=q=0.5$, es la máxima que puede llegar a tener un intervalo de esa confianza y ese tamaño muestral, por lo que la segunda estimación tendrá una amplitud menor o a lo sumo igual.

La razón es que la expresión $pq=p(1-p)$ alcanza un máximo para $p=0.5$, (valiendo $0.5^2=0.25$) por lo que para cualquier otro valor de p , $pq \leq 0.25$ y, en consecuencia para cualquier otro valor de p (como el p_a de un estudio anterior) el error de estimación (semi-amplitud del intervalo) será:

$$E = \pm z_{\alpha/2} \sqrt{\frac{p_a q_a}{n}} \leq \pm z_{\alpha/2} \frac{0.5}{\sqrt{n}}$$

El muestreo que asume $p=q=0.5$ da como solución una estimación de igual o mayor amplitud que la auténtica estimación (con el p poblacional real, que es desconocido) con esa confianza.

Y en ese sentido, al dar la estimación de máxima amplitud para confianza del 95.5 %, podemos **asegurar** que la estimación resultante tendrá esa confianza o más confianza aún. Mientras que el segundo muestreo, al trabajar con una mera aproximación de p , puede que acabara dando una estimación que “en realidad” tuviera menos confianza de la pedida. Por eso es más riguroso en ese sentido considerar $p=q=0.5$.

Sin embargo se asimilamos rigor a precisión en la estimación el intervalo más preciso es el que considera la p del estudio previo. Pero hay que tener en cuenta que este resultado “más preciso” quizá no goce de la confianza requerida, sino de algo menos.

8.- Para conocer la diferencia entre los salarios de hombres y mujeres de cierto sector económico obtenemos información aleatoria de 100 hombres, y de 180 mujeres (trabajadores todos ese sector). Resultando un salario medio de 2000 €, con una desviación típica de 800, para los hombres y un salario medio de 1800 € con una desviación típica de 600 € para las mujeres. Si para conocer la diferencia entre los salarios medios construimos un intervalo de confianza con nivel de confianza del 95%, ¿Qué error podemos cometer al intentar conocer dicha diferencia?

Se trata de estimar la diferencia entre las dos medias procedentes de 2 poblaciones normales, con tamaños muestrales grandes. Por lo que el intervalo de confianza para la estimación será:

$$\mu_A - \mu_B \in \left[(\bar{X}_A - \bar{X}_B) \pm z_{\alpha/2} \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} \right] \text{ con } 1 - \alpha$$

$$\text{por lo que el error será } E = \pm z_{\alpha/2} \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = \pm 1.96 \sqrt{\frac{800^2}{100} + \frac{600^2}{180}}$$

$$\boxed{E = \pm 179.637 \text{ €}}$$

9.- Un instituto oficial de estadística ha realizado una estimación de la tasa de paro a partir de una encuesta sobre la población activa. El número de encuestas válidas fue de 6400. Y el informe estima una tasa de paro del 15 % de la población activa, con un error de $\pm 1\%$, con un nivel de confianza del 95.5 %. ¿Cómo es posible este resultado si el error teórico para la estimación de una proporción, trabajando con un nivel de confianza del 95.5 % es habitualmente $E = \pm (1/\sqrt{n})$; que en este caso sería $\pm 1.25\%$?

Ciertamente si se hace el supuesto de $p=q=0.5$ el error

$$E = \pm z_{\alpha/2} \frac{0.5}{\sqrt{n}} = \pm 2 \frac{0.5}{80} = 0.0125 = \pm 1.25\%$$

pero presumiblemente se ha debido hacer la consideración de que la tasa de paro tomará un valor alrededor de una p más baja que 0.5, por lo que el error acabará siendo más bajo.

$$\text{en concreto: } E = 0.01 = 2 \cdot [p \cdot (1-p)/6400]^{1/2}$$

de donde $0.0001 \cdot 6400/4 = p - p^2$: $p^2 - p + 0.16 = 0$ cuyas dos soluciones son:

$p=0.2$ y $p=0.8$ (obviamente se debe haber considerado que la tasa real de paro "no excede al 20 %".

10.- Se quiere estimar la proporción de clientes insatisfechos con nuestros servicios, se sabe de estudios anteriores que tal proporción es inferior al 30% y queremos realizar el

estudio gastándonos para ello un total de 24000 de u.m. Si la estimación ha de tener una confianza del 95 %, realizar una encuesta cuesta 12 um y la tasa de no respuesta es del 25% ¿ Con qué error podremos realizar nuestra estimación?

Partiendo de que el intervalo de confianza para la estimación de una proporción es:

$$p \in \left[\hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}} \right] \text{ con } 1-\alpha \text{ de confianza}$$

y considerando que $z_{\alpha/2} = 1.96$ y que $p.q \leq 0.3 \times 0.7$

y que con 24000 u.m. se pueden realizar $24000/12 = 2000$ encuestas de las que sólo un 75% serán válidas (25 % de no respuesta)

el valor de n será: $n=1500$ (2000×0.75)

$$\text{por lo que el error será: } E = \pm z_{\alpha/2} \sqrt{\frac{0.3 \times 0.7}{1500}}$$

$$E = \pm 1.96 \times 0.0118321 = \pm 0.02319$$

es decir un error del $\pm 2.32\%$

11.- Estimar , con una confianza del 99 %, la proporción de fallos de conexión de una red informática si se conoce que en una selección aleatoria de 150 conexiones fallaron 10.

partiendo de que el intervalo para la estimación de una proporción es:

$$p \in \left[\hat{p} \pm z_{\alpha/2} \frac{0.5}{\sqrt{n}} \right] \text{ con } 1-\alpha \text{ de confianza}$$

y a partir de la información :

$n = 150$, puede considerarse M.A.S. por ser N muy grande

$$\hat{p} = \frac{10}{150} = 0.066$$

$$z_{\alpha/2} = 2.58 \quad \text{ya que N.C.} = 99\%$$

nos quedaría:

$$p \in \left[0.066 \mp \frac{2.58 \cdot 0.5}{\sqrt{150}} \right] \equiv [0.066 \mp 0.1053]$$

$$p \in [-0.393, 0.1713]$$

y como, obviamente una proporción no puede ser negativa la ***estimación final sería que la proporción es inferior al 17.13 % con una confianza del 99 %***