TP: Estimación Puntal y por Intervalos

1. Un fabricante de medicamentos veterinarios, está interesado en conocer la proporción de animales que padecen infecciones locales, cuya condición puede ser controlada por un nuevo producto desarrollado por la empresa. Para satisfacer su inquietud, se realizó un estudio en el que participaron 500 animales que padecen infecciones locales y se encontró que un 80% de los mismos pueden controlar la infección con el medicamento. Suponiendo que los 500 animales son representativos del grupo de animales con infecciones locales en la población total, responder las siguientes preguntas:

a) ¿Cuál es la población?

b) ¿Cuál es la muestra?

c) Identificar el parámetro de interés

d) Identificar el estimador y proporcionar su valor en la muestra dada

e) En este caso, ¿se conoce el valor del parámetro?

a) Población: los animales que padecen infecciones locales.
b) Muestra: 500 animales seleccionados que padecen infecciones locales.
c) Parámetro: Proporción de animales que en la población total pueden controlar la infección con el medicamento
d) Estimador: la Proporción de animales que en la muestra pueden controlar la infección con
el medicamento. Estimación: 80% (400 de los 500 animales de la muestra).

e) No, el verdadero valor del Parámetro es desconocido.

2. Durante el último Censo Económico de 1994, en una localidad de la Provincia de Entre Ríos se censaron 32.000 comercios. A comienzos de este año, a partir de la base de datos construida con las respuestas al censo, se obtuvo una muestra de 400 comercios entre los que se obtuvo la siguiente información:

- ✓ **50** de los censados en 1994 ya **no existían** más.
- ✓ Los **empleados** de los comercios detectados ascendían a **1.000 personas**.
- ✓ Los dedicados a **artículos alimenticios eran 80**.
- ✓ Las **ventas** de estos últimos el año anterior fueron de **16 millones de $**.
- ✓ Las **ventas totales de la muestra** durante el año 2000 alcanzaron **120 millones de $**.

A partir de esta información muestral, **estimar**:

a) ¿ Cuántos de los 32.000 comercios censados en 1994 **se dedicaban al rubro alimentos**?.
b) ¿ Qué **cantidad de empleados** tenían ?.
c) La cantidad de comercios que **ya no existen más**.

d) ¿ Cuál hubiese sido el **monto total de las ventas** de la población de comercios en el año 2000.
e) El **porcentaje** de comercios que **siguen existiendo** en el 2001.
f) El **promedio de ventas** de los comercios **alimenticios**.

a) $\frac{80}{400}32.000 = 6400 \; se \; dedican \; al \; rubro$

Otra manera:

$$\frac{80}{350}28.000 = 6400 \; se \; dedican \; al \; rubro$$

b) $\frac{1000}{400}32.000 \cong 80.000 \; empleados$

Otra manera:

$$\frac{1000}{350}28.000 \cong 80.000 \; empleados$$

c) $\frac{50}{400}32.000 \cong 4000 \; ya \; no \; existen \; mas$

d) $\frac{120}{400}32.000 = \; 96000 \; millones$

e) $\frac{16}{80} = 0.2 \; millones$

3. Las observaciones siguientes corresponden a una muestra aleatoria de tamaño 9 de la variable aleatoria X, consumo de carbón por servicios eléctricos en millones de toneladas, en un año dado: 406 395 400 450 390 410 415 401 408

Hallar una estimación puntual para μ, consumo medio de carbón para servicios eléctricos. ¿Es el valor que ha obtenido igual al consumo medio de carbón para electricidad en el año en cuestión?

$$\bar{x} = \frac{406 + 395 + 400 + 450 + 390 + 410 + 415 + 401 + 408}{9} = 408.3$$

4. Se analizó una muestra de 12 piezas de pan blanco de cierta marca y se determinó el porcentaje de carbohidratos contenido en cada una de las piezas, obteniéndose los siguientes valores:

76.93 76.88 77.07 76.68 76.39 75.09 77.67 76.88 78.15 76.50 77.16 76.42

a) Estimar el esperanza del porcentaje de carbohidratos contenido en las piezas de pan de esta marca.

X: porcentaje de carbohidratos contenido en las piezas de pan de esta marca

E (X)=$\bar{x}$=(76.93+76.88+77.07+⋯.)/12=921.82/12=76.818

b) Estimar la proporción de piezas de pan de esta marca cuyo contenido de carbohidratos no excede el 76.5%.

Ordenemos los datos:

75.09 76.39 76.42 76.50 76.68 76.88 76.88 76.93 77.07 77.16 77.67 78.15

Por lo cantidad estimada es 4 / 12 = 1/3 = 0.3333

5. En una experiencia genética se extraen 20 moscas de una caja experimental. Medida la longitud del ala en cada mosca se obtuvieron los siguientes valores: 93, 90, 97, 90, 93, 91, 96, 94, 91, 91, 88, 93, 95, 91, 89, 92, 87, 88, 90, 86 Suponiendo que la longitud del ala sigue una distribución normal, hallar un intervalo de confianza de nivel 0.95 para los parámetros μ.

Datos para realizar la estimación:

Como queremos obtener un intervalo con un 95% de confianza, tenemos:

1 − α = 0,95, y por tanto α = 0,05 y α / 2 = 0,025, así, (1− α / 2) = 0,975.

Al ser n "chica", x~t de Student y para resolver el problema, debe calcularse con la muestra los valores de la media y el desvío estándar muestrales que se indican precedentemente.

n = 20

$$\bar{x} = \frac{93+90+97+90+93+91+96+94+91+91+88+93+95+91+89+92+87+88+90+86}{20} = 91{,}25$$

$$S = \sqrt{\frac{(93-91{,}25)^2+\cdots\ldots}{20-1}} = 2{,}93$$

$$S^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

$\bar{x}$ *se distribuye con los parámatro* $\mu \;\; , \frac{S}{\sqrt{n}}$

Reemplazando en $P(L_I = \bar{x} - t.\, S_{\bar{x}} \leq \mu \leq L_S = \bar{x} + t.\, S_{\bar{x}}) = 1 - \alpha$

6. En un laboratorio de investigación biológica, se selecciona un conjunto de 100 cobayos que padecen una cierta enfermedad para aplicarles un tratamiento. Sobre los mismos se obtuvo un tiempo medio de supervivencia de 46 semanas. Se sabe por experiencias anteriores que la variancia es de 36 semanas^2. Construir e interpretar un intervalo de confianza del 95 % para el tiempo medio de supervivencia.

Los datos disponibles son los siguientes:

x: supervivencia en meses de los cobayos que presentan una cierta enfermedad y reciben un tratamiento ~ ¿? (¿? ; $\sigma^2 = 36$)

$\alpha = 0{,}05$ o 5 %

n = 100 muestra de los cobayos en esas condiciones (grande)

Al ser n “grande”, x ~ N (¿? , 6/ 10); por el Teorema Central del Límite

En base a esa información se pide calcular:

$46 \pm 1.96 * 6/\ 10 \Rightarrow 46 \pm 1{,}18 \Rightarrow 44{,}82 < \mu < 47{,}18$

Es decir que el tiempo medio de supervivencia de los cobayos afectados por la enfermedad y que reciben el tratamiento, se encuentra **entre 44.8 y 47.2 semanas**, con una confianza del 95 %.

7. En una región agrícola en la cual se siembra predominantemente una cierta variedad de trigo, una compañía productora de semillas ha desarrollado una nueva variedad T y desea conocer su rendimiento promedio.

Para ello se siembran ocho lotes experimentales con la nueva variedad y se obtienen los siguientes rendimientos (en toneladas por ha):

3.15 - 3.92 - 4.26 - 3.72 - 4.19 - 3.42 - 4.38 - 4.50

Suponiendo que el rendimiento sigue una distribución normal, estimar el promedio real (μ) a un nivel de significancia (error) del 5%.

La información suministrada indica:

x: rendimiento de la nueva variedad T de semillas de trigo ~ N (¿? ; σ = ¿?)

α = 0,05 o 5 % ; n = 8 lotes de muestra (“chica”)

A partir de la muestra: $\bar{X}$ = 3,94 Tn/Ha ; S = 0,48 Tn/Ha .

Al ser n “chica”, $\bar{x}$ ~ t de Student y para resolver el problema, debe calcularse con la muestra los valores de la media y el desvío estándar muestrales que se indican precedentemente.

Se dan las condiciones para adoptar que $\bar{X}$ ~ t de Student, con lo cual:

$P(L_I = \bar{x} - t.S_{\bar{x}} \leq \mu \leq L_S = \bar{x} + t.S_{\bar{x}}) = 1-\alpha$ y reemplazando los valores:

$P(L_I = 3{,}94 - 2{,}365.\ 0{,}48 / 8^{1/2} < \mu < L_S = 3{,}94 + 2{,}365.\ 0{,}48 / 8^{1/2}) =$

$P(L_I = 3{,}94 - 0{,}4 < \mu < L_S = 3{,}94 + 0{,}4) = P(3{,}54 < \mu < 4{,}34) = 0{,}95$

Lo que quiere decir que el rendimiento promedio de la nueva semilla se encontrará **entre los 3,54 y 4,34 Tn por Ha** con un 95 % de confianza.

8. En un estudio sobre desnutrición infantil, se tomaron aleatoriamente 16 niños con las condiciones necesarias para el estudios, cuya altura promedio dio 0.85 m. con una variancia de 0.09 m2. Si ya se sabe que la altura se distribuye en forma normal, determinar un intervalo de confianza para la altura promedio de los niños con un nivel de confianza del 95 %.

Se dispone de la siguiente información:

x = altura de los niños ~ N(μ = ¿? ; σ desconocida) (distribución Normal)

n = 16 Tamaño de muestra (pequeño)

α = 0,05

N = ¿?

$\bar{x}$ = 0.85 m Altura promedio muestral

S = 0,3 m Dispersión muestral

A partir de esta información se solicita: Intervalo de confianza para μ

Por tener: x ~ N(¿?,¿?) ;n = 16 (pequeña) ; y variancia poblacional desconocida y trabajarse con la dispersión muestral. Siendo:

$P(L_I = \bar{x} - t.S_{\bar{x}} \leq \mu \leq L_S = \bar{x} + t.S_{\bar{x}}) = 1 - \alpha$ y reemplazando los valores:

$L_I = 0{,}85 - 2{,}131 \cdot 0{,}3 / 4 = 0{,}85 - 0{,}16 =$ **0,69 m**

$L_I = 0{,}85 + 2{,}131 \cdot 0{,}3 / 4 = 0{,}85 + 0{,}16 =$ **1,01 m**

o sea que: **P(0,69 < μ < 1,01) = 0,95**

Esto quiere decir que con una probabilidad del 95 %, la altura promedio de los niños bajo estudio se encontrarán entre 69 y 101 cm. O bien puede interpretarse que de cada 100 muestras de 16 en esas condiciones, en 95 de los intervalos de confianza que se calculen se cubrirá al verdadero promedio de altura desconocido, y en los 5 restantes el verdadero promedio quedará fuera del intervalo calculado.

9. Un grupo de investigadores está interesado en conocer la concentración media de una enzima en cierta población de algas. Se sabe por experiencias previas que la varianza de la concentración de esta enzima es de 35. Si se obtiene una muestra de tamaño 15 dándonos un nivel de concentración media de 18 calcular un intervalo de confianza al nivel 0.95 (se supone normalidad).

$IC_{0.95}$= (15,0061, 209939)

10. Se obtiene una muestra aleatoria de 1000 adultos aparentemente sanos con el fin de establecer un patrón con respecto al que se considerará una lectura «normal» de calcio. Se extrae una muestra de sangre de cada adulto. La variable estudiada es X, número de miligramos de calcio por decilitro de sangre. Se han obtenido una media muestral de 9.5 y una desviación típica muestral de 0.5. Supóngase que X presenta una distribución aproximadamente normal. Hallar un intervalo de confianza de μ del 95 %.

11. La concentración media de dióxido de carbono en el aire en una cierta zona no es habitualmente mayor que 335 p.p.m.v. (partes por millón en volumen). Se sospecha que esta concentración es mayor en la capa de aire más próxima a la superficie. Se ha analizado el aire en 20 puntos elegidos aleatoriamente a una misma altura cerca del suelo, resultando una media muestral de 580 p.p.m.v. y una desviación típica de 180. Suponiendo normalidad para las mediciones, dar una estimación puntual para la concentración media de dióxido de carbono cerca del suelo y calcular un intervalo de confianza al nivel de confianza 0.95.

12. Contesta razonadamente a estas preguntas:
a) ¿Cuál tiene mayor amplitud, un intervalo de confianza al nivel 0.95 ó 0.99?
b) ¿Cuál es el efecto del tamaño muestral sobre la amplitud de un intervalo de confianza?

RESPUESTA/RESULTADOS:

a) El que tiene un nivel de confianza de 0.99;
b) Mayor tamaño muestral menor amplitud, con el mismo nivel de confianza