

ESTADÍSTICA E INTRODUCCIÓN A LA ECONOMETRÍA
Curso 2022/23
Problemas del Tema 2

OBSERVACIÓN: Los problemas marcados con asterisco deben resolverse utilizando Excel. En todos los problemas, supondremos que la muestra que se utiliza es una muestra aleatoria simple.

***1.-** Para estudiar los salarios mensuales de los trabajadores de diferentes países se han obtenido muestras de trabajadores de estos países. La hoja “Problema 1” del archivo EXCEL adjunto contiene los salarios (en euros) en el último mes de los trabajadores seleccionados. En este problema nos centraremos exclusivamente en los trabajadores del País 3; llamaremos μ y σ^2 a la media y varianza poblacionales de los salarios mensuales de los trabajadores de este país. Calcula:

- a) Una estimación insesgada de μ .
- b) Una estimación insesgada de σ^2 .
- c) Una estimación insesgada de la varianza de la media de la muestra de los trabajadores del País 3.
- d) Una estimación insesgada de la proporción de trabajadores del País 3 que tienen salarios mensuales inferiores a 1500 euros.

2.- Sea X_1, X_2, X_3 una muestra de una variable aleatoria (v.a.) X con media μ y varianza σ^2 . Para estimar μ se consideran los tres estimadores siguientes:

$$\hat{\mu}_1 = \frac{2X_1 + 4X_2 - X_3}{5}; \quad \hat{\mu}_2 = 0.7X_1 + 0.1X_2 + 0.2X_3; \quad \hat{\mu}_3 = \frac{1}{5}\hat{\mu}_1 + \frac{4}{5}\hat{\mu}_2.$$

- a) Demuestra que los tres estimadores son insesgados.
- b) Analiza cuál de los tres estimadores es más eficiente, y calcula la eficiencia relativa del más eficiente con respecto a los otros dos.

3.- En una encuesta sobre una cuestión de actualidad, el encuestado puede responder 1 (en desacuerdo), 2 (indiferente) ó 3 (de acuerdo). Se sabe que la probabilidad de que una persona seleccionada al azar responda 1 es 2θ , la de que responda 2 es θ y la de que responda 3 es $1 - 3\theta$, siendo θ un parámetro desconocido que está dentro del intervalo $(0, \frac{1}{3})$. Para estimar θ disponemos de una muestra con las respuestas de n encuestados.

- a) Una investigadora estima θ utilizando como estimador $(3 - \bar{X})/5$, siendo \bar{X} la media de la muestra. Demuestra que el estimador que utiliza esta investigadora es insesgado.
- b) Otro investigador estima θ utilizando como estimador $\hat{p}/2$, siendo \hat{p} la proporción de personas de la muestra que responden 1. Demuestra que el estimador que utiliza este investigador es insesgado.
- c) Demuestra que el estimador utilizado por la investigadora es más eficiente que el utilizado por el investigador.
- d) Supongamos que $n = 20$, y que en la muestra se han obtenido nueve respuestas 1, cuatro respuestas 2 y siete respuestas 3.

d1) Calcula la estimación de θ que obtendrá la investigadora.

d2) Calcula la estimación de θ que obtendrá el investigador.

d3) Teniendo en cuenta el resultado demostrado en el apartado **c**, la investigadora afirma que el verdadero valor de θ está más cercano a la estimación obtenida en el subapartado **d1** que a la estimación obtenida en el subapartado **d2**; ¿es cierta esta afirmación?

4.- Se sabe que el peso (en kg) de un ladrillo producido en una fábrica es una v.a. con distribución normal con media μ y desviación típica 0.12 kg. Se dispone de una muestra de n ladrillos.

a) Supongamos que $n = 100$, que la media de la muestra es 4.07 kg, y que la varianza de la muestra es 0.022 kg². Calcula un intervalo de confianza para μ con nivel de confianza del 95% y determina su amplitud.

b) Sin efectuar ningún cálculo, indica razonadamente si el intervalo de confianza para μ que obtendríamos en las siguientes situaciones tendría amplitud mayor o menor que el intervalo calculado en el apartado **a**:

b1) Con la misma muestra que en el apartado **a**, pero con un nivel de confianza del 99%.

b2) Con el mismo nivel de confianza que en el apartado **a**, pero con una muestra formada por 50 ladrillos.

b3) Con la misma muestra y el mismo nivel de confianza que en el apartado **a**, pero suponiendo que la desviación típica del peso de los ladrillos es 0.15 kg.

c) Supongamos que podemos determinar el tamaño de la muestra n . ¿Cómo debe ser n si se quiere conseguir que el intervalo de confianza para μ con nivel de confianza del 95% tenga amplitud inferior a 0.02?

5.- El beneficio mensual (en miles de euros) de una empresa es una v.a. X con distribución normal. Para estudiar las características de X , se dispone de una muestra de tamaño 6 de X , que ha proporcionado los resultados siguientes:

Beneficios: 23.4 23.2 24.1 19.3 20.1 22.2

a) Calcula un intervalo de confianza para el beneficio mensual medio con nivel de confianza del 95%, y determina su amplitud.

b) Una investigadora indica que, utilizando esta misma muestra pero un nivel de confianza diferente, el intervalo de confianza para el beneficio mensual medio que ella ha obtenido tiene amplitud 4.76.

b1) ¿Cuál es el intervalo de confianza que ha obtenido esta investigadora?

b2) ¿Cuál es el nivel de confianza que ha utilizado esta investigadora para obtener su intervalo de confianza?

6.- Una gran empresa utiliza un test para analizar la aptitud de sus empleados. Se evalúa con este test a una muestra de 352 empleados; se sabe que la media y la desviación típica de las calificaciones obtenidas en el test por los empleados de la muestra fueron 60.41 y 11.28, respectivamente.

a) Calcula un intervalo de confianza para la calificación media del test con nivel de confianza del 99%, si se sabe que las calificaciones siguen una distribución normal con desviación típica 11.

b) Calcula un intervalo de confianza para la calificación media del test con nivel de confianza del 99%, si se sabe que las calificaciones siguen una distribución normal pero se desconoce cuál es su desviación típica.

c) Calcula un intervalo de confianza aproximado para la calificación media del test con nivel de confianza del 99%, si se sabe que las calificaciones siguen una distribución que no es normal y se desconoce cuál es su desviación típica.

d) Supongamos que sabemos que las calificaciones no siguen una distribución normal, pero sí conocemos su desviación típica, que es 11. Si tuvieras que calcular un intervalo de confianza aproximado para la calificación media del test con nivel de confianza del 99% en esta situación, ¿darías la misma respuesta que en el apartado anterior? Si tu respuesta es afirmativa, indica por qué; en caso contrario, calcula el intervalo de confianza aproximado que obtendrías.

***7.-** Se ha tomado una muestra de 324 personas de una población, y a cada persona se le ha pedido que indique su sexo y que califique en una escala de uno a cinco su estado de salud en el último año (5 indica que el estado de salud ha sido muy bueno y 1 que ha sido muy malo). Los resultados obtenidos pueden verse en la hoja “Problema 7” del archivo EXCEL adjunto.

a) Llamemos μ a la media poblacional de la variable “estado de salud”. Calcula un intervalo de confianza aproximado para μ con nivel de confianza del 95%.

b) Llamemos μ_H a la media de la variable “estado de salud” para la población de hombres, y μ_M a la media de la variable “estado de salud” para la población de mujeres. Calcula un intervalo de confianza aproximado para μ_H con nivel de confianza del 95%, calcula un intervalo de confianza aproximado para μ_M con nivel de confianza del 95%, compara los dos intervalos obtenidos, y comenta qué indican estos resultados.

8.- En las próximas elecciones se disputan el puesto de presidente el candidato del partido conservador y el candidato del partido progresista. Llamaremos p a la proporción de votantes de la población que votarán al candidato del partido progresista. Para estimar este valor, se ha encuestado a 120 votantes; 66 de estos votantes votarán al candidato del partido progresista, y los demás votarán al candidato del partido conservador. Basándose en los resultados de esta encuesta, un periodista afirma que el intervalo de confianza aproximado para p con nivel de confianza del 95% sólo contiene valores mayores a 0.5 y, por tanto, el candidato del partido progresista ganará las elecciones, con un nivel de confianza del 95%. ¿Es cierta esta afirmación?

9.- Para saber cuánto se valora en las empresas un título universitario, se ha preguntado a los directores de recursos humanos de 340 empresas si consideraban importante que un candidato a un puesto de trabajo tuviera título universitario; 255 directores respondieron que sí y el resto que no. Llamaremos p a la proporción poblacional de empresas que consideran importante un título universitario.

a) Calcula un intervalo de confianza aproximado para p con nivel de confianza del 98%, e indica por qué podemos considerar como válida esta aproximación.

b) Con estos mismos datos, el intervalo de confianza aproximado para p que ha obtenido una investigadora, utilizando otro nivel de confianza, ha sido (0.7162, 0.7838).

b1) Teniendo en cuenta el intervalo obtenido en el apartado anterior, y sin efectuar ningún cálculo, indica si el nivel de confianza que ha utilizado esta investigadora es inferior o superior al 98%.

b2) Calcula cuál es el nivel de confianza que ha utilizado esta investigadora.

c) Calcula un intervalo de confianza aproximado para $1 - p$ con nivel de confianza del 98%, justificando el procedimiento que utilices para tu cálculo.

10.- Un fabricante quiere estimar la media y la variabilidad de los niveles de impureza (en gramos) de los envíos de materia prima de un determinado proveedor. Extrae para ello una muestra de 15 envíos y comprueba que la media muestral del nivel de impureza es 1.87 gr y que la desviación típica muestral es 0.24 gr. Supondremos que el nivel de impureza de un envío de materia prima de este proveedor es una v.a. con distribución normal con media μ y varianza σ^2 .

a) Calcula un intervalo de confianza para μ con nivel de confianza del 90%.

b) Calcula un intervalo de confianza para σ^2 con nivel de confianza del 95%.

c) Calcula un intervalo de confianza para σ con nivel de confianza del 99%.

d) Calcula un intervalo de confianza para $\frac{1}{\sigma}$ con nivel de confianza del 99%, justificando el procedimiento que utilices para tu cálculo.

***11.-** Para estudiar el consumo (en litros por cada 100 km) de varios modelos nuevos de automóviles se han obtenido muestras del consumo de diferentes automóviles de estos nuevos modelos. Los resultados obtenidos pueden verse en la hoja “Problema 11” del archivo EXCEL adjunto. En este problema nos centraremos exclusivamente en la muestra de automóviles del nuevo modelo de Opel. Supondremos que el consumo de un automóvil del nuevo modelo de Opel es una v.a. con distribución normal con media μ y varianza σ^2 , ambas desconocidas, y llamaremos \bar{X} a la media muestral de los consumos de automóviles del nuevo modelo de Opel. Calcula:

a) Un intervalo de confianza para μ con nivel de confianza del 95%.

b) Un intervalo de confianza para σ^2 con nivel de confianza del 90%.

c) Un intervalo de confianza para la desviación típica de \bar{X} con nivel de confianza del 99%.

12.- Responde **razonadamente** a las siguientes preguntas:

a) Sea X_1, \dots, X_n una muestra de una v.a. X con distribución binomial $\text{Bi}(1; p)$ y sea \bar{X} la media muestral. Si llamamos $\theta = p^2$ y $\hat{\theta} = \bar{X}^2$, ¿es $\hat{\theta}$ un estimador insesgado de θ ?

b) Si, para una muestra dada, el intervalo de confianza para la media poblacional de una población normal con nivel de confianza del 95% ha resultado ser (21.2, 24.5), ¿podemos afirmar que la probabilidad de que la media poblacional quede fuera del intervalo (21.2, 24.5) es 0.05?

c) Sea X una v.a. con distribución uniforme entre 0 y a , siendo a un número real positivo desconocido. Se quiere estimar por intervalos la media de X , a partir de una muestra de tamaño n . Un investigador propone utilizar el siguiente intervalo de confianza aproximado para esta

media, con nivel de confianza del 95%:

$$\left(\bar{x} - 1.96\sqrt{\frac{a^2/12}{n}}, \bar{x} + 1.96\sqrt{\frac{a^2/12}{n}} \right)$$

siendo \bar{x} la media muestral. ¿Te parece razonable la propuesta del investigador?

d) Llamando μ a la media de una v.a. X y σ^2 a su varianza, sabemos que $\sigma^2 = \mu^2/3$. Se quiere estimar por intervalos la media de X , a partir de una muestra de tamaño n . Indica cómo podemos obtener un intervalo de confianza aproximado para μ , con nivel de confianza del 95%, si la única información de la muestra que tenemos es su tamaño n y la media muestral \bar{x} .

e) Si X_1 , X_2 y X_3 son tres v.a. independientes tales que $X_1 \sim N(1; 2)$, $X_2 \sim N(1; 1)$ y $X_3 \sim N(2; 1)$, y llamamos

$$Y = \frac{(X_1 - 1)/\sqrt{2}}{\sqrt{\frac{(X_2 - 1)^2 + (X_3 - 2)^2}{2}}}$$

¿cuál es la probabilidad de que Y sea superior a 1.96?

f) Sea Y una v.a. con distribución t_5 . Determina el valor de c para que sea cierta la igualdad siguiente:

$$P(-1.6493 < Y < c) = 0.90$$