PROBLEMES ESTADÍSTICA ENGINYERIA
MOSTREIG i ESTIMACIÓ DE PARÀMETRES

1.- El precio medio del m^2 en la venta de casas nuevas durante el último año en una determinada ciudad fue de 115000 pts. La desviación típica de la población fue de 25000 pts. Se toma una muestra aleatoria de 100 casas nuevas de esta ciudad.

a) ¿Cuál es la probabilidad de que la media muestral de los precios de venta sea menor que 110000 pts?

b) ¿Cuál es la probabilidad de que la media muestral de los precios de venta esté entre 113000 pts. y 117000 pts.??

c) ¿Cuál es la probabilidad de que la media muestral de los precios de venta esté entre 114000 pts. y 116000 pts.?

d) Sin hacer cálculos, razonar en cuál de los siguientes rangos resulta más probable que se encuentre la media muestral de los precios de venta:

113000 pts-	115000 pts
114000 pts-	116000 pts
115000 pts-	117000 pts
116000 pts-	118000 pts

2.- Se ha tomado una muestra de 16 directores de oficina de corporaciones de una gran ciudad, con el fin de estimar el tiempo medio que emplean en desplazarse para ir a su trabajo. Supongamos que la distribución de dichos tiempos en la población sigue una normal con media 87 minutos y desviación típica 22.

a) ¿Cuál es el error estándar de la media muestral de los tiempos de desplazamiento?

b) ¿Cuál es la probabilidad de que la media muestral sea inferior a 100 minutos?

c) ¿Cuál es la probabilidad de que la media muestral sea superior a 80 minutos?

d) ¿Cuál es la probabilidad de que la media muestral esté entre 85 y 95 minutos?

e) Supongamos que se toma una segunda muestra de quince directores, independiente de la anterior. Sin hacer los cálculos, razonar si la probabilidades calculadas en los apartados b), c) y d) serán mayores, menores o iguales para esta segunda muestra. Utilizar gráficos para ilustrar las respuestas.

3.- Una compañía produce cereales para el desayuno. La media del peso que contienen las cajas de estos cereales es de doscientos gramos y su desviación típica de seis gramos. La distribución de los pesos de la población es normal. Se eligen cuatro cajas, que pueden considerarse como una muestra aleatoria del total de la producción.

a) ¿Cuál es el error estándar de la media muestral del peso de las cuatro cajas?

b) ¿Cuál es la probabilidad de que la media del peso de esas cuatro cajas sea inferior que 197 gramos?

c) ¿Cuál es la probabilidad , en media, el peso de estas cuatro cajas esté entre 105 y 195 gramos?

d) ¿Cuál es la probabilidad de que la suma del peso de estas cuatro cajas sea menor de 800 gr.?

e) Se eligen al azar dos de estas cuatro cajas ¿Cuál es la probabilidad de que, en media, el contenido de estas dos cajas pese entre 195 y 200 gramos?

[0]Sol.: a) 0.0228; b) 0.5762; c) 0.3108: d) el intervalo 114000 pts-116000 pts
[0]Sol.: a) 5.5 ; b) 0.9909; c) 0.8980; d) 0.5671; e) es menor en los tres casos.

4.- La tasa de rentabilidad de ciertos tipos de acciones sigue una distribución con desviación típica 3.8. Se extrae una muestra de tales acciones con el fin de estimar el precio medio.

a) ¿Qué tamaño ha de tener la muestra para asegurarnos que la probabilidad de que la media muestral difiera de la media poblacional en una cantidad superior a 1 sea menor que 0.1?

b) Sin realizar los cálculos razonar si será preciso un tamaño muestral mayor o menor que el requerido en el apartado a) para garantizar que la probabilidad de que la media muestral difiera de la media poblacional en más de 1 sea inferior a 0.05.

c) Sin realizar los cálculos razonar si será preciso un tamaño muestral mayor o menor que el requerido en el apartado a) para garantizar que la probabilidad de que la media muestral difiera de la media poblacional en más de 1.5 sea inferior a 0.1.

5.- De acuerdo con los datos del ministerio de Economía y Hacienda, el 15% de las declaraciones del IRPF del último año darán lugar a una devolución. Se toma una muestra aleatoria de 10 declaraciones.

a) ¿Cuál es la media de la distribución en el muestreo de la proporción muestral de declaraciones que darán lugar a una devolución?

b) ¿Cuál es la varianza de la proporción muestral?

c) ¿Cuál es el error estándar de la proporción muestral?

d) ¿Cuál es la probabilidad de que la proporción muestral sea mayor que 0.8?

6.- El dueño de una portal de ventas de discos por internet ha comprobado que el 20% de los clientes que acceden a su portal realizan una compra. Cierta mañana entraron en el portal 180 personas, que pueden ser consideradas como una muestra aleatoria de todos sus clientes.

a) ¿Cuál será la media de la proporción muestral de clientes que realizaron alguna compra?

b) ¿Cuál es la varianza de la proporción muestral?

c) ¿Cuál es el error estándar de la proporción muestral?

d) ¿Cuál es la probabilidad de que la proporción muestral sea mayor que 0.15?

7.- El administrador de una gran cadena de hospitales opina que, de entre todos sus pacientes, el 30% generará facturas que se pagarán con más de dos meses de retraso. Se toma una muestra aleatoria de 200 pacientes.

a) ¿Cuál es el error estándar de la proporción muestral de pacientes con facturas cuyo pago se retrasará dos meses?

b) ¿Cuál es la probabilidad de que la proporción muestral sea inferior a 0.25?

c) ¿Cuál es la probabilidad de que la proporción muestral sea mayor que 0.33?

---

[0]Sol.: a) 3; b) 0.1587; c) 0.0475; d) 0.5; e) 0.3810
[0]Sol.: a) $n \geq 40$; b) mayor; c) menor
[0]Sol.: a) 0.15; b) 0.01275; c) 0.1129; d) casi nula.
[0]Sol.: a) 0.2; b) ≈ 0.0009; c) 0.03; d) 0.9525

d) ¿Cuál es la probabilidad de que la proporción muestral esté entre 0.27 y 0.33?

e) Sin realizar los cálculos, razonar en cuál de los siguientes intervalos es más probable que se encuentre la proporción muestral: 0.29-0.31; 0.30-0.32; 0.31-0.33; 0.32-0.34.

f) Supongamos que se toma al azar una muestra de 500 pacientes. Sin realizar los cálculos razonar si las probabilidades de los apartados b), c) y d) resultarán en este caso mayores, menores o iguales que las calculadas para la muestra anterior.

8.- Se toma una muestra aleatoria de 100 votantes con el fin de estimar la proporción de los mismos que están a favor de un aumento en los impuestos sobre la gasolina para contar así con un ingreso adicional para reparaciones de las autopistas. ¿Cuál es el mayor valor que puede tomar el error estándar de la proporción muestral de esta medida?

9.- Continuando en la situación del problema anterior, se decide que una muestra de 100 votantes es muy pequeña para obtener una estimación de la proporción poblacional que resulte suficientemente creíble. Se decide exigir que la probabilidad de que la proporción muestral difiera de la proporción poblacional (cualquiera que sea su valor) en más de 0.03 no debe ser superior a 0.05. ¿Qué tamaño ha de tener la muestra para poder garantizar que se cumple este requisito?

10.- Una compañía quiere estimar la proporción de personas que son posibles compradores de máquinas de afeitar eléctricas que ven retransmisiones partidos de La Liga de Campeones. Se toma una muestra de 120 individuos que se identificaron como posibles compradores de afeitadoras eléctricas. Supongamos que la proporción de posibles compradores de afeitadoras eléctricas en la población que ven estas retransmisiones es 0.25.

a) 0.10 es la probabilidad de que la proporción muestral exceda a la proporción poblacional ¿en qué valor?

b) 0.05 es la probabilidad de que la proporción muestral esté por debajo de la proporción poblacional ¿en qué cantidad?

c) 0.30 es la probabilidad de que la proporción muestral difiera de la proporción poblacional ¿en menos de qué cantidad?

11.- Supongamos que el 50% de los españoles adultos opina que es necesaria una revisión del sistema nacional público de hospitales. ¿Cuál es la probabilidad de que más del 56% de los componentes de una muestra de 150 españoles adultos tenga esa opinión?

12.- Las rentabilidades mensuales de cierto tipo de acciones son independientes unas de otras, y siguen una distribución normal con desviación típica 1,7. Se toma una muestra de 12 meses.

a) Hallar la probabilidad de que la desviación típica muestral sea menor que 2.5.

b) Hallar la probabilidad de que la desviación típica muestral sea mayor que 1.

13.- El número de horas que dedican a ver la televisión los estudiantes en la semana anterior a los exámenes finales sigue una distribución normal con una desviación típica de 4.5 horas. Se toma una muestra aleatoria de 30 estudiantes.

---

[0]Sol.: a) 0.0324; b) 0.0618; c) 0.1762; d) 0.6476; e) 0.29-0.31; f) menor, menor, mayor
[0]Sol.: 0.05
[0]Sol.: $n \geq 757$
[0]Sol.: a) 0.0506; b) 0.0648; c) 0.0154
[0]Sol.: 0.0708
[0]Sol.: a) 0.8; b) 0.975

a) La probabilidad de que la desviación típica muestral sea mayor que 3.5 horas, ¿es mayor que 0.95?

b) La probabilidad de que la desviación típica muestral sea menor que seis horas, ¿es mayor que 0.95?

14.- Se extrae una muestra aleatoria de 15 economistas y se les pregunta acerca de su predicción sobre la tasa de inflación para el próximo año. Supongamos que las predicciones para la población completa de economistas sigue una distribución normal con una desviación típica de 1.8.

a) 0.01 es la probabilidad de que la desviación típica sea mayor que ¿qué número?

b) 0.025 es la probabilidad de que la desviación típica sea menor que ¿qué número?

c) Encontrar una par de números, a y b, tales que la probabilidad de que la desviación típica muestral se encuentre entre ellos sea 0.9.

15.- Se toma una muestra de ocho lotes de un producto químico para comprobar la concentración de impurezas. Los niveles porcentuales de impurezas encontrados en la muestra fueron

3.2 4.3 2.1 2.8 3.2 3.6 4.0 3.8

a) Hallar la media y la varianza muestrales. Hallar la proporción muestral de lotes con nivel porcentual de impurezas mayor que 3.75%.

b) ¿Para qué parámetros poblacionales se han hallado en la parte a) estimadores por procedimientos insesgados?

16.- Sea $\hat{\theta}_1$ un estimador insesgado de θ_1, y $\hat{\theta}_2$ un estimador insesgado de θ_2.

a) Probar que $(\hat{\theta}_1 + \hat{\theta}_2)$ es un estimador insesgado de $(\theta_1 + \theta_2)$.

b) Probar que $(\hat{\theta}_1 - \hat{\theta}_2)$ es un estimador insesgado de $(\theta_1 - \theta_2)$.

17.- Sea X_1 y X_2 una muestra aleatoria de dos observaciones independientes de una población con media μ y varianza σ^2. Considerar los siguientes tres estimadores puntuales de μ:

$$\begin{aligned} \overline{X} &= \tfrac{1}{2}X_1 + \tfrac{1}{2}X_2 \\ \hat{\mu}^{(1)} &= \tfrac{1}{4}X_1 + \tfrac{3}{4}X_2 \\ \hat{\mu}^{(2)} &= \tfrac{1}{3}X_1 + \tfrac{2}{3}X_2 \end{aligned}$$

a) Probar que los tres estimadores son insesgados.

b) ¿Cuál de los tres estimadores es más eficiente?

c) Hallar la eficiencia relativa de $\overline{X}$ con respecto a los otros estimadores.

18.- A una clase de estadística asisten estudiantes de Informática de Gestión y de Sistemas. En una muestra de diez estudiantes de Gestión se observaron las siguientes calificaciones en el examen final

62 57 85 59 64 63 71 58 77 72

[0]Sol.: a) Sí; b) Sí

[0]Sol.: a) 2.5969; b) 1.1415; c) 1.2331; 2.341

[0]Sol.: a) $\overline{X} = 3.375$; $S^2_{\overline{X}} = 0.4993$; $\hat{p} = 0.3557$; b) para todos.

[0]Sol.: b) $\overline{X}$; c) $\frac{Var(\overline{X})}{Var(\hat{\mu}^{(1)})} = 0.8$; $\frac{Var(\overline{X})}{Var(\hat{\mu}^{(2)})} = 0.9$

En una muestra independiente de ocho estudiantes de Sistemas se observaron las siguientes calificaciones en el mismo examen

73 79 85 73 62 51 60 57

a) Utilizar un método de estimación insesgado para obtener una estimación puntual de la diferencia de las calificaciones medias entre los estudiantes de Gestión y los de Sistemas. (Indicación: Utilizar el problema 151)

b) Utilizar un método de estimación insesgado para obtener una estimación puntual de la diferencia entre la proporción poblacional de estudiantes que obtuvieron una calificación mayor que 70 en el grupo de estudiantes de Gestión y el grupo de Sistemas. (Indicación: Utilizar el problema 151)

19.- Se toma una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una población con media μ y varianza σ^2. Se considera el siguiente estimador de μ :

$$\hat{\mu} = \frac{2}{n(n+1)}(X_1 + 2X_2 + 3X_3 + \ldots + nX_n)$$

a) Probar que $\hat{\mu}$ es un estimador insesgado de μ.

b) Hallar la eficiencia relativa de $\hat{\mu}$ respecto a $\overline{X}$, la media muestral.

20.-

a) (Examen junio 2003) Calcular el estimador máximo verosímil (MLE[1]) para el parámetro λ de una población que sigue una ley $Exp(\lambda)$ para una muestra aleatoria simple de tamaño n.

b) (Examen septiembre 2004) Calcular el MLE para el parámetro λ de una población que sigue una ley $Po(\lambda)$ para una muestra aleatoria simple de tamaño n.

c) Calcular el MLE para el parámetro μ de una población que sigue una ley $N(\mu, \sigma^2)$ para una muestra aleatoria simple de tamaño n.

d) Calcular el MLE para el parámetro σ^2 de una población que sigue una ley $N(\mu, \sigma^2)$ para una muestra aleatoria simple de tamaño n.

e) Estudiar si los estimadores MLE de los apartados anteriores son insesgados.

21.- De una población de barras de hierro se extrae una muestra de 64 barras y se calcula la resistencia a la rotura por tracción se obtiene que $\overline{X} = 1012\ Kg/cm^2$. Se sabe por experiencia que en este tipo de barras $\sigma = 25$. Calcular un intervalo de confianza para μ al nivel 0.95.

22.- Para investigar el C.I. medio de una cierta población de estudiantes, se realiza un test a 400 estudiantes. La media y la desviación típica muestrales obtenidas son $\overline{x} = 86$ y $\tilde{s}_X = 10.2$. Calcular un intervalo para μ con un nivel de significación del 98%.

23.- Para investigar un nuevo tipo de combustible para cohetes espaciales, se disparan cuatro unidades y se miden las velocidades iniciales. Los resultados obtenidos, expresados en Km/h, son :19600, 20300, 20500, 19800. Calcular un intervalo para la velocidad media μ con un nivel de confianza del 95%, suponiendo que las velocidades son normales. 20718.3

[0]Sol.: a) 0.2444; b)$-\frac{1}{10}$

[0]Sol.: b) $Var(\hat{\mu}) = \left(\frac{2}{n(n+1)}\right)^2 \frac{n(n+1)(2n+1)}{6}\sigma^2$; $Eff.rel = \frac{Var(\hat{\mu})}{Var(\overline{X})} = \frac{2\,(1+2\,n)}{3\,(1+n)}$.

[1]Del inglés "Maximun Likelihood Estimator"

[0]Sol.: $(1005.88, 1018.13)$

[0]Sol.: $(84.8117, 87.1883)$

[0]Sol.: $(19381.7, 20718.3)$

24.- Un fabricante de cronómetros quiere calcular un intervalo de estimación de la desviación típica del tiempo marcado en 100 horas por todos los cronómetros de un cierto modelo. Para ello pone en marcha 10 cronómetros del modelo durante 100 horas y encuentra que $\tilde{s}_X = 50$ segundos. Encontrar un intervalo para el parámetro σ^2 con $\alpha = 0.01$, suponiendo que la población del tiempo marcado por los cronómetros es normal.

25.- Un auditor informático quiere investigar la proporción de rutinas de un programa que presentan una determinada irregularidad. Para ello observa 120 rutinas, resultando que 30 de ellas presentan alguna irregularidad. Con estos datos buscar unos límites de confianza para la proporción p de rutinas de la población que presentan esa irregularidad con probabilidad del 95%.

26.- (Examen septiembre 2003) Una infección por un virus puede haber perjudicado a muchos ordenadores con *Windwos*. Desde el Centro de Alerta Temprana (CAT) se quiere calcular la proporción de ordenadores infectados. El jefe del centro os pide que calculéis el tamaño de una muestra para que el intervalo de confianza de la proporción muestral de ordenadores infectados tenga amplitud de a lo sumo 0.01 con una probabilidad del 90%.

27.- (Examen junio 2003) Se han medido los siguientes valores (en miles de personas) para la audiencia de un programa de televisión en distintos días (supuestos igualmente distribuidos e independientes):

$$521, 742, 593, 635, 788, 717, 606, 639, 666, 624.$$

Construir un intervalo de confianza del 90%, para la audiencia poblacional media y otro para la varianza poblacional, bajo la hipótesis de que la población de audiencias sigue una ley normal.

Nota Suma de las audiencias=6531, Suma de los cuadrados de las audiencias=4320401.

28.- Supongamos que la empresa para la que trabajamos está en un proyecto de investigación, financiado con fondos de la Comunidad Europea, que pretende extender una nueva aplicación de las TIC. Una de las tareas del proyecto es realizar una encuesta de opinión sobre el grado de aceptación que tendría esta nueva tecnología en el mercado europeo. De entre todas las universidades y empresas participantes en el proyecto, es a tu empresa a la que le toca hacer el protocolo de la encuesta, llevarla a cabo y redactar esta parte del informe final. Como eres el último que llegó a la empresa y el resto de miembros del equipo no se acuerda de la estadística que vio en la carrera, te toca a ti cargar con la responsabilidad. Claro que el coste de la encuesta depende del número n de entrevistas que se realicen y el error de las proporciones de las contestaciones disminuye cuando n aumenta. Como no sabes cuánto dinero está dispuesto a gastar tu jefe, tabula los tamaños muestrales para los errores $\pm 5\%, \pm 3\%, \pm 2\%, \pm 1\%$, y para niveles de confianza 0.95 y 0.99, suponiendo el peor caso. Añade un comentario para que el equipo de dirección del proyecto, en el que hay componentes ignorantes en materia de encuestas, vea como quedarían redactado los datos técnicos de la encuesta, y pueda decidir el tamaño muestral leyendo tu informe.

29.- El número de reservas semanales de billetes de cierto vuelo de una compañía aérea sigue una distribución aproximadamente normal. Se toma un muestra aleatoria de 81 observaciones de números de reservas de este vuelo: el número medio de reservas muestral resulta ser 112, mientras que la desviación típica muestral es 36. Además de estos 81 vuelos, 30 llegaron a su destino con un retraso de más de 15 minutos.

a) Calcular un intervalo de confianza del 95% para el número medio poblacional de reservas en este vuelo.

b) Calcular un intervalo de confianza de 95% para la varianza poblacional de las reservas.

c) Calcular un intervalo de confianza del 95% para la proporción poblacional de vuelos que llegan con un retraso de más de 15 minutos.

d) Calcular el tamaño muestral que asegura un intervalo de confianza de amplitud 0.1 para la proporción de vuelos que llegan con un retraso de más de 15 minutos al nivel de confianza 95%.

[0]Sol.: $(953.834, 12968.3)$
[0]Sol.: $(0.1725, 0.3275)$

30.- Una empresa cervecera sabe que las cantidades de cerveza que contienen sus latas sigue una distribución normal con desviación típica poblacional 0.03 litros.

a) Se extrae una muestra aleatoria de 25 latas y, a partir de la misma, un experto en estadística construye un intervalo de confianza para la media poblacional del contenido en litros de las latas que discurre entre 0.32 y 0.34 ¿Cuál es el nivel de confianza de este intervalo?

b) Un gerente de esta empresa exige un intervalo de confianza del 99% que tenga una amplitud máxima de 0.03 litros a cada lado de la media muestral ¿Cuántas observaciones son necesarias, como mínimo, para alcanzar este objetivo?

[0]Sol.: a) $(104.16, 119.84)$); b) $(972.343, 1814.08)$); c) $(0.265, 0.475)$); d) $n = 385$

[0]Sol.: a) 90.3%; b) $n = 7$