Inferencia-estadistica-Problemas...

S1XTHS3NS3

Estadística

1º Grado en Ingeniería Informática

Escuela Politécnica Superior de Córdoba
Universidad de Córdoba

Inferencia estadística

1. La organización de la Vuelta Ciclista a España ha marcado, para una determinada etapa, un tiempo t de duración de la misma de 5 horas y 40 minutos, y supone que la diferencia entre este tiempo t y el empleado por los ciclistas sigue una variable aleatoria normal X de media cero.
Tomada una muestra de 25 ciclistas elegidos al azar:

-La probabilidad $P\left(\frac{\bar{X}}{S} > 0.57\right)$ es: **0,005**.

2. Supongamos que en una población se ha medido una característica X que se encuentra normalmente distribuida con media 14 y con varianza 16.
Si se toma una muestra (m.a.s.) de tamaño 9,

-¿Están los datos de la media muestral más dispersos que los datos poblacionales? **No**.
-El coeficiente de variación de la media muestral es **0,09524** y el poblacional es **0,2857**.

3. Supongamos que en una población se ha medido una característica X que se encuentra normalmente distribuida con media 12 y con varianza 16.
Si se toma una muestra (m.a.s.) de tamaño 9,

-¿Cuál es la probabilidad de que la media muestral tome valores superiores a 14? **0,0668**.

4. El peso del contenido de las cajas de bombones de una cierta marca puede suponerse que se distribuye de forma normal. Una inspección del Ministerio de Sanidad, rechaza una partida de bombones si el peso medio de los bombones contenidos en las cajas es inferior a 200 g. Para comprobar si es aceptable la partida toma al azar una muestra de 10 cajas, pesando el contenido de las mismas. Los resultados muestrales proporcionan una media de 198.5 g y una cuasi-desviación típica de 2.635 g.
Otra empresa del sector dice proporcionar el mismo producto pero con mayor peso. Para comprobar tal afirmación se toman 15 cajas y se pesa el contenido de las mismas, resultando como media muestral 205.5 g y desviación típica muestral 3.65 g.

-De acuerdo con los resultados muestrales y con una significación del 10%, ¿la variabilidad en el peso de las cajas de bombones en ambas empresas es el mismo? **Sí**.
-Para tomar la decisión anterior, el valor del estadístico muestral usado es **0,4864** y el extremo inferior de la región de aceptación es **0,3326**.

5. Sea X una v.a. con función de densidad $f(x) = k\, e^{kx}$ si $-1 < x < 1$, siendo k un parámetro desconocido. Para estimarlo, se obtiene una muestra de tamaño 10 de la que resulta una media de 0.5. Aplicando el método de máxima verosimilitud,

-¿Qué valor tomaría el parámetro k? **-2**.

6. Para una muestra de una población de media μ y varianza $K\mu^2$ se considera

como estimador de μ: $\hat{\mu} = \frac{X_1 + 4X_2}{5}$

-¿Cuál es el sesgo del estimador? **0**.

7. Se sospecha que una moneda está trucada. Para comprobarlo, se lanzó 1000 veces y se obtuvieron 615 caras.

-La estimación máximo verosímil de la probabilidad de obtener cruz es: **0,3850**.

Mediante intervalos de confianza:

-¿Es cierta la sospecha con una confianza del 95%? **Sí**.
-El extremo inferior del intervalo construido es **0,5848** y el superior **0,6451**.

8. Sea una variable X cuya función de densidad es:

$$f(x) = e^{-x(x-\theta)^2} \quad \forall x \in \Re$$

de la que se conoce que E[X] = θ y Var[X] = 1/2π.
Extraída una muestra de tamaño n, se obtienen los siguientes estimadores para el parámetro θ:

$$\widehat{\theta}_1 = \frac{X_1 + X_2 + \cdots + X_n}{n} \; ; \; \widehat{\theta}_2 = \frac{X_1 + X_2 + \cdots + X_{n-1}}{n-1} \; ; \; \widehat{\theta}_3 = \frac{\widehat{\theta}_1 + \widehat{\theta}_2}{2}$$

-¿Es $\widehat{\theta}_1$ insesgado? **Sí**.

-¿Es $\widehat{\theta}_2$ consistente? **Sí**.

-¿Qué estimador es más eficiente, el primero o el segundo? **Primero**.

-¿Alguno de los estimadores anteriores se corresponde con el de máxima verosimilitud? **Sí**. En caso afirmativo, ¿cuál es, el primero, el segundo o el tercero? **Primero**.

9. Sea una variable X cuya función de densidad es no nula en:

$$f(x) = (2x)/\theta^2 \quad 0 < x < \theta$$

Se considera el siguiente estimador puntual $\widehat{\theta} = \frac{3}{2}\overline{X}$**:**

-¿Es $\widehat{\theta}$ el estimador de máxima verosimilitud? **No**.

-¿Es $\widehat{\theta}$ insesgado? **Sí**.

-¿Es $\widehat{\theta}$ consistente? **Sí**.

10. Una empresa editora quiere realizar un estudio acerca del Volumen anual de ventas de sus comerciales y del sueldo Mensual de los mismos, ambas medidas en miles de €. Para ello, selecciona una muestra aleatoria entre todos ellos, estudiando también si poseen Titulación Universitaria (T=1) o no (T=2). Los datos obtenidos aparecen en la tabla siguiente:

Vendedor	Sexo	T	V	M	Vendedor	Sexo	T	V	M
1	H	1	82	1.2	14	H	1	50	1.68
2	M	1	75	1.5	15	M	1	46	1.2
3	H	2	92	1.3	16	M	1	42	1.15
4	H	2	90	1.8	17	H	1	65	1.23
5	M	1	70	1.2	18	H	2	81	1.86
6	M	1	65	1.3	19	M	2	83	1.47
7	M	1	60	1.8	20	M	1	72	1.8
8	M	2	90	1.85	21	H	1	75	1.75
9	H	1	58	1.6	22	H	1	80	1.81
10	H	1	50	1.1	23	M	2	76	1.56
11	H	2	86	1.39	24	H	1	90	1.9
12	M	2	85	1.7	25	M	1	85	1.78
13	M	2	83	1.58					

Suponiendo normalidad con parámetros desconocidos en las distribuciones poblacionales,

-La estimación mediante un intervalo de confianza al 99% del Volumen de ventas anual de la empresa tiene como extremo inferior **64,8688** y como superior **81,6112**.
-La estimación mediante un intervalo de confianza al 95% del Volumen de ventas anual de la empresa tiene como extremo inferior **67,0626** y como superior **79,4174**.

11. Se considera una población representada por una v.a. X cuya distribución es N (μ = 1.7 ; σ = 4).

Supuesto extraída una muestra de tamaño 10,

-Determinar un número K tal que P (S^2 > K) = 0.99: **3.34**.

12. El administrador de una compañía de taxis trata de decidir si el uso de llantas radiales en lugar de llantas regulares de cinturón mejora la economía de combustible. Se equipan 12 automóviles con llantas Radiales y se prueban en un recorrido establecido. Otros 12 automóviles se equipan con llantas comunes con Cinturón y se prueban en el mismo recorrido. Los consumos de gasolina en litros cada 100 km son los siguientes:

Automóvil	Llantas Radiales	Llantas con Cinturón
1	4.2	4.1
2	4.7	4.9
3	6.6	6.2
4	7.0	6.9
5	6.7	6.8
6	4.5	4.4
7	5.7	5.7
8	6.0	5.8
9	7.4	6.9
10	4.9	4.7
11	6.1	6.0
12	5.2	4.9

-¿Son iguales los consumos de gasolina en ambos tipos de llantas? **Sí** (Nivel de significación 10%).
-Para tomar la decisión anterior, el valor del estadístico muestral usado es **0,339** y el extremo superior de la región de aceptación es **1,717**.
-¿Qué valor tomaría el mejor estimador de la varianza en el consumo de gasolina con las llantas radiales? **1,1082**.
-¿Y el consumo medio en las llantas con cinturón? **5,6083**.

13. El número de horas trabajadas semanalmente por el maitre de un cierto restaurante es una variable aleatoria X, cuya función de densidad es:

$$f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}\ ;x>0$$

Durante 10 semanas tomadas al azar, el nº de horas trabajadas han sido: 70, 11, 66, 15, 20, 14, 35, 40, 29, 18.

-Aplicando el método de máxima verosimilitud, ¿qué valor de θ es viable a la luz de los datos muestrales? **31,8**.

14. Se sabe que la altura de los alumnos de la Universidad de Córdoba sigue una distribución $X_C \in N(\mu_C ; \sigma^2_C)$, y los de la Universidad de Sevilla $X_S \in N(\mu_S ; \sigma^2_S)$.

Si se eligen al azar 51 alumnos de la Universidad de Córdoba:

-Suponiendo que $\sigma_C = 16$, y que se ha construido un intervalo de confianza para la media poblacional de amplitud 6 cm, ¿cuál es el coeficiente de confianza utilizado? **0,82**.

Se elige una muestra aleatoria en cada universidad, cuyos datos son:

UCO: 170.2, 160.4, 165.3, 157.8, 178.6, 170.3
US: 180.1, 162.3, 157.1, 163.5, 168.9

-¿La altura de los universitarios es la misma en ambas universidades con una confianza del 90%? **Sí**. Para tomar esta decisión, el extremo inferior de la región de aceptación es **-1,833** y el superior **1,833**, siendo el valor del estadístico usado **0,1466**.
-¿La altura de los universitarios es la misma en ambas universidades con una confianza del 90%? **Sí**. Para tomar esta decisión, el extremo inferior del intervalo construido es **-8,2848** y el superior **9,7248**.

15. Un investigador sospecha que el consumo de ciertas sustancias estimulantes entre jóvenes influye negativamente en su rendimiento académico. Para comprobarlo ha diseñado el siguiente experimento: A dos grupos de 15 alumnos se les puso dos pruebas similares, una tras Consumir (variable C) y otra Sin dicho Consumo (variable SC). Los resultados de las calificaciones son:

Nº de individuo	Sin Consumo (SC)	Con Consumo (C)
1	6.5	6.2
2	7	7.1
3	5	4.7
4	6.3	6.1
5	4.3	4.5
6	3.5	3.7
7	7	7
8	3.7	3.2
9	8	7.5
10	2.7	2.1
11	6	5.7
12	6.7	7
13	3	2.7
14	5	4.8
15	6	6.1

Considerando que las calificaciones poblacionales en ambos casos siguen distribuciones normales de medias desconocidas y varianzas desconocidas e iguales,

-El valor estimado de la varianza poblacional es: **2,7175**.
-Se desea contrastar si las calificaciones medias en ambos grupos son iguales (Consumiendo y Sin Consumir) para un nivel de significación α. Para este objetivo, el valor del estadístico del contraste es: **-0,2479**.
-¿Se podrían considerar iguales al 5% de significación? **Sí**. ¿Cual sería el extremo inferior de la región de aceptación? **-2,048**.

16. Los precios del menú en un restaurante durante un cierto periodo de tiempo siguen una ley Normal (variable M), con media 6 y desviación típica de 3 €. La empresa decide abrir otro restaurante en otra ciudad y piensa que el comportamiento de los clientes no va a cambiar de una ciudad a otra. Para determinar una estimación en el precio de los menús en su nuevo restaurante, realiza una encuesta en 61 clientes sobre el precio que estarían dispuestos a pagar, obteniendo como media 5.70 € y desviación típica 2.5 €.

-¿Qué tamaño muestral necesitaría tomar la empresa, con una confianza del 0.99, si desea realizar una estimación del precio de los menús con una precisión de ±0.5? **178** (Redondear al entero superior).

17. Una variable aleatoria X tiene por función de densidad:

$$f(x) = (\theta + 1)\, x^{\theta} \quad ; \quad 0 < x < 1 \quad ; \quad \theta > -1$$

Se considera la siguiente realización muestral: 0.3, 0.6, 0.7, 0.2, 0.3, 0.1

-Aplicando el método de máxima verosimilitud, ¿qué valor de θ es viable a la luz de los datos muestrales? **-0,1652**.

18. Un fabricante de una cierta pieza afirma en su catálogo que las dimensiones de la citada pieza se asocian a una variable X ∈ D (μ ; σ = 2.116). Un comprador decide rechazar la compra a este fabricante si la dimensión media de una muestra de piezas difiere de en más de ±1 unidades. Se pide:

-¿Cuantas piezas debería haber en la muestra para que la probabilidad de admitir la compra fuese 0.99? **30**.

19. Después de su fabricación y embalaje, la vida media (en días) de un cierto medicamento es N (μ = 1200 ; σ = 40).
Se desea enviar un lote de los citados medicamentos de modo que la vida media muestral no sea inferior a 1180 días con probabilidad 0.95.

-¿Cual es el tamaño del lote? **11**.

20. Se toma una muestra aleatoria simple de 10 alumnos de entre una población escolar correspondiente a un curso académico. Por experiencia en cursos pasados se sabe que la altura de los citados alumnos se distribuye según una ley Normal de media 167 cm con desviación típica de 3.2 cm.

-¿Cuál es la probabilidad de que la media muestral supere los 165 cm? **0,9761**.
-¿Cuál es la probabilidad de que la media muestral sea inferior a 165 cm? **0,0239**.
-¿Cuál es la probabilidad de que la cuasivarianza muestral sea superior a 15.9? **0,1**.

21. Supuesta una característica poblacional representada por una v.a. X cuya distribución es N ($\mu = 0$; σ^2), con σ^2 desconocida de donde se ha extraído una muestra aleatoria de tamaño 6.

-La probabilidad $P\left(\frac{\overline{X}}{S} > 1.8\right)$ es: **0,005**.