Problemas de Intervalos de Confianza

1. En un artículo de publicación periódica, se informa que se usó una muestra de tamaño 5 como base para calcular un intervalo de confianza de 95% para la frecuencia natural promedio (Hz) de vigas deslaminadas de cierto tipo. El intervalo resultante fue (229.7, 233.5). Suponemos que la frecuencia natural sigue una ley normal con media μ y desviación estándar σ.

(a) ¿Podemos afirmar que $p\,(\mu \in (229.7, 233.5)) = 0.95$? Justifica tu respuesta.

(b) Determina el valor promedio muestral y la desviación estándar muestral (no olvides las unidades).

(c) Se decide que un nivel de confianza de 99% es más apropiado. Determina el intervalo de confianza correspondiente.

2. Una empresa fabrica fusibles eléctricos. En el control de calidad, se toma una muestra aleatoria de 1000 unidades y se observan 17 fusibles defectuosos.

(a) En este caso, ¿podemos usar una aproximación normal para estimar la proporción de fusibles defectuosos?

(b) Calcula un intervalo de confianza para estimar la proporción de defectos con un nivel de confianza del 99%.

(c) Según las especificaciones de uno de nuestros clientes, la máxima proporción aceptable de elementos defectuosos es del 2%. Considerando la información disponible, ¿estamos en condiciones de cumplir las especificaciones del cliente?

3. Sea Y una variable aleatoria discreta que toma valores 0 o 1, con probabilidades $p(Y=0) = a$ y $p(Y=1) = 1 - a$ donde $0 < a < 1$ es desconocido. Suponemos que se cumple que $na(1-a) \geq 8$.

(a) Para un tamaño de la muestra n dado, calcula la longitud máxima del intervalo de confianza.

(b) Queremos estimar a con un intervalo de confianza de 90%, ¿Cuál tiene que ser, como mínimo, el tamaño de la muestra para que la longitud del intervalo de confianza sea como máximo de 0.02 (equivalente a un radio de 0.01)?

4. La afluencia de visitantes al parque del Tibidabo durante un mes, medida a través de una muestra aleatoria durante 10 días elegidos aleatoriamente, han sido los siguientes:

$$682, \quad 553, \quad 555, \quad 666, \quad 657, \quad 649, \quad 522, \quad 568, \quad 700, \quad 558$$

Suponiendo que los niveles de afluencia siguen una distribución normal, determinar con un 95 por ciento de confianza, un intervalo de confianza para la afluencia media al mes.

5. Una variable física X (amb les unitats corresponents) és aleatòria i sabem que segueix una distribució normal amb σ coneguda. Per estimar el valor esperat de X, amb un nivell de confiança del 90%, s'ha pres una mostra aleatòria simple de mida $n = 9$, i s'ha obtingut l'interval

$$(118.25, 123.55).$$

(a) Calculeu la mitjana aritmètica dels valors observats de la variable.

(b) Calculeu la variància (poblacional) de la variable X.

(c) Si es vol augmentar el nivell de confiança al 97%, determineu l'interval de confiança resultant fent servir la mateixa mostra dels apartats anteriors.

(d) Es vol estudiar la influència que pot tenir la mida de la mostra. Determineu (de forma aproximada), quina hauria de ser la mida de la mostra perquè la longitud de l'interval de confiança del 97% sigui igual a la longitud de l'interval inicial.

6. Un artículo publicado en la revista *"Transactions of the American Fisheries Society"* en el año 1993 informa de los resultados de un estudio que investiga la contaminación por mercurio en la lubina. Se seleccionó una muestra de lubinas de 51 lagos de Florida y se midió la concentración de mercurio en el tejido muscular (ppm). Los valores de concentración de mercurio fueron:

1.230	1.330	0.040	0.044	1.200	0.270	0.490	0.190	0.830	0.810	0.710	0.500
0.490	1.160	0.050	0.150	0.190	0.770	1.080	0.980	0.630	0.560	0.410	0.730
0.590	0.340	0.340	0.840	0.500	0.340	0.280	0.340	0.750	0.870	0.560	0.170
0.180	0.190	0.040	0.490	1.100	0.160	0.100	0.210	0.860	0.520	0.650	0.270
0.940	0.400	0.430									

(a) Teniendo en cuenta que la contaminación de mercurio sigue una distribución normal, que la media de la concentración de mercurio de la muestra dada $\left(\frac{1}{51}\sum_{x=1}^{51} x_i\right)$ es 0.5354 ppm y que su desviación típica $\left(\sqrt{\frac{\sum_{x=1}^{51}(x_i-\bar{x})^2}{51}}\right)$ es 0.3479 ppm, determina un intervalo de confianza para la varianza con un nivel de confianza del 95%.

(b) Las lubinas procedentes de los lagos dónde la concentración de mercurio supera los 0.700 ppm se declaran no aptas para el consumo. Si se estima que en más de la mitad de los lagos de Florida las lubinas no son aptas para el consumo se declara el estado de emergencia (en Florida hay más de 1400 lagos). Estima un intervalo con un nivel de confianza del 99% para la proporción de lagos con lubinas no aptas para el consumo. ¿Recomendarías que se declarase el estado de emergencia?

Problemas de Hipótesis

1. Una medida de la sombra del azulejo de mármol es la cantidad de luz reflejada por éste y se mide en una escala de 0-255. Un azulejo perfectamente negro no refleja luz alguna y mide 0 y un azulejo perfectamente blanco mide 255. Se midió una muestra de nueve azulejos Mezza Perla, obteniéndose los siguientes resultados:

$$204.999, 206.149, 202.150, 207.048, 203.496, 206.343, 203.496, 206.676, 205.831$$

El fabricante asegura que el azulejo Mezza Perla refleja una cantidad de luz superior a 206.5. Asumiendo normalidad, se pide:

(a) Sospechamos que el índice de reflexión és inferior al indicado por el fabricante. Haz un test para contrastar nuestra hipòtesis al 90%. Calcula el p-valor y representa el contraste.

(b) Sabiendo que la varianza poblacional es 4, contrasta la hipótesis al 95% de que la media de luz reflejada por el azulejo Mezza Perla es 206.5. Calcula el p-valor y representa el contraste.

2. Nos han pedido que realicemos un estudio para evaluar si las ventas medias de nuestra compañía en los establecimientos minoristas son mayores o son menores que las ventas medias del sector.

Sabemos que las ventas medias del sector son de 62000 euros por establecimiento. Para llevar a cabo el estudio tomamos la muestra aleatoria siguiente: 53700, 55500, 53000, 52400, 51000, 62000, 75000, 53800, 56600. Asumiendo normalidad, se pide:

(a) Evalúa mediante una prueba de hipótesis si las ventas de nuestra empresa son inferiores a las del sector con un nivel de significancia del 5%.

(b) Calcula el p-valor del contraste y representa el contraste.

(c) Suponiendo que la varianza de las ventas en el conjunto de la población es conocida y tiene un valor de 54760000, estima y representa un intervalo de confianza al 99% para el valor medio de las ventas.

3. Según el fabricante, la resistencia de un cierto tipo de cuerda es de 8000 N. Para verificar la calidad de una partida de este tipo de cuerda, hemos realizado una prueba de resistencia a la rotura con 6 muestras escogidas aleatoriamente, y hemos obtenido una resistencia promedio de 7750 N con desviación estándar de 145 N. Asumiendo normalidad, contesta a las siguientes cuestiones:

(a) ¿Atendiendo a la información disponible, debemos presentar una reclamación al fabricante (nivel de significación $\alpha = 0.01$)?

(b) Si sabemos que el valor de la varianza poblacional es 136161 N^2, y tomamos un nivel de significación $\alpha = 0.05$, ¿cuál será la decisión?

(c) En el contraste del apartado anterior, asumiendo que el verdadero valor del parámetro es $\mu_1 = 7700\,\mathrm{N}$, calcula y representa la probabilidad de error de tipo II (probabilidad de mantener H_0 cuando el verdadero valor del parámetro es μ_1).

4. Según estudios previos, el tiempo de espera en la cola formada en un semáforo de una travesía urbana, en la franja horaria de 9:00 a 13:00 de los días laborables, seguía una distribución Normal de media 14 minutos y desviación típica 4.8 minutos. En respuesta a las continuas quejas de los vecinos, las autoridades municipales decidieron habilitar una vía de circulación auxiliar, con la intención de descongestionar el tráfico en ese punto y disminuir el tiempo medio de espera. Para verificar la efectividad de las medidas tomadas, transcurrida una semana se realizó un nuevo estudio sobre una muestra aleatoria de 26 vehículos, obteniéndose un tiempo medio de espera de 12.5 minutos y una desviación estándar de 2.7 minutos. Las autoridades municipales mostraron su satisfacción por los resultados obtenidos, afirmando que los datos indicaban una mejora significativa de la fluidez del tráfico; sin embargo, las asociaciones de vecinos mantenían que las medidas tomadas no habían producido ninguna mejora significativa.

(a) Plantea un test para contrastar la hipótesis defendida por las asociaciones de vecinos frente a la de los responsables municipales. Si se concluye que el tiempo medio de espera se había reducido y, realmente, no lo había hecho, ¿cómo se llama el error cometido?

(b) ¿A qué conclusión se llega con un nivel de significación $\alpha = 0.01$?

(c) Si tomamos como válido el valor de la desviación típica obtenido en los estudios previos y un nivel de significación $\alpha = 0.05$, ¿a qué conclusión se llega?

5. Sea X una variable aleatoria normal con esperanza matemática μ y varianza σ^2. Escogemos una muestra aleatoria de X con tamaño $n = 41$. Sabemos que el promedio de la muestra vale 10 y que la desviación estándar (corregida) de la muestra es 1.5.

(a) Determina un intervalo de confianza para μ con nivel de confianza $c = 0.87886$.

(b) Contrasta la hipótesis de que la media poblacional es superior a 10.5, con una significancia del 5%.

6. Los tiempos de reacción, en milisegundos, frente a una matriz de 15 estímulos es una variable aleatoria con distribución normal. Para una muestra aleatoria de 4 sujetos, se han obtenido los siguientes tiempos de reacción: 515, 464, 558, 491.

(a) Un experto afirma que la desviación típica es superior a 100ms, y nosotros sospechamos que es inferior. Contrasta nuestra hipótesis con un nivel de significación del 10%.

(b) Un conocido científico considera que los parámetros de la distribución normal relativa al tiempo de reacción en milisegundos son: esperanza matemática 500 y desviación típica 50. Calcula el tamaño mínimo que debe tener una muestra para que 95 de cada 100 muestras tengan una diferencia inferior a 10 milisegundos entre la media muestral y la media poblacional del tiempo de reacción. O sea calcula la n mínima que consigue que $P\left(\left|\frac{X_1+\ldots+X_n}{n} - \mu\right| < 10\right) = 0.95$.

7. El rendimiento de un nuevo proceso es una variable aleatoria Normal X de media μ y desviación estándar $\sigma = 5$. En la prueba de hipótesis nula $H_0 : \mu \leq 80$ frente a la hipótesis alternativa $H_1 : \mu > 80$, consideramos una muestra aleatoria de tamaño $n = 100$. Si el nivel de significación es de 5.05%, calcular el error de tipo II o sea $\beta(81)$ para $\mu = 81$?

8. El artículo "Refinement of Gravimetric Geoid Using GPS and Leveling Data" (W. Thurston, Journal of Surveying Engineering, $2000 : 27 - 56$) presenta un método para la medición de alturas ortométricas sobre el nivel del mar. Para una muestra de 1225 boyas, 926 dieron resultados que estaban dentro de los límites de tolerancia. ¿Con que nivel de significancia podemos concluir que este método produce resultados dentro de los límites de tolerancia en más de un 75% de las veces?