Estadística (Química) - Primer Cuatrimestre - 2020 - Coronavirus
Práctica 5 - Estimación - Intervalos de Confianza - Primera parte

En todos los ejercicios propuestos

a) defina las variables aleatorias y los parámetros involucrados.

b) de ser posible indique:

   i. la distribución de las variables aleatorias
   ii. el significado intuitivo de los parámetros.

c) compare los resultados de hacer las cuentas a mano con las salidas obtenidas con el R, de manera de verificar las primeras y aprender a usar las segundas, en aquellos ejercicios en los que ambas cosas sean posibles.

Ejercicios

1. Sea $X_1, \ldots . X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2.

   (a) Pruebe que $\overline{X}_n^2$ no es un estimador insesgado de μ^2.

   (b) ¿Para qué valores de k es $\widehat{\mu}^2 = \left(\overline{X}_n^2 - ks^2\right)$ un estimador insesgado de μ^2?

2. Sea $X_1, \ldots . X_n$ una muestra aleatoria de una distribución Bernoulli de parámetro p y sea $T_n = \sum_{i=1}^{n} X_i$. Consideremos un nuevo parámetro $\theta = p\,(1-p)$. Muestre que

$$\frac{T_n\,(n-T_n)}{n\,(n-1)}$$

   es un estimador insesgado de θ.

3. Para determinar la constante de un resorte, se puede medir la oscilación de una masa m fijada en uno de sus extremos. Mediante la fórmula $k = 4\pi^2 m/T^2$ donde k es la constante del resorte y T es el período de oscilación de dicha masa. Se realizan mediciones exactas de la masa y del período. Asumamos que cada medición de la masa y del período siguen distribuciones normales, con varianza 0.1 y 0.05, respectivamente. De 5 mediciones de la masa y 4 del período se obtuvieron medias muestrales 0.52 y 1.3. Asumiendo que todas las mediciones son independientes, estime la constante del resorte e indique el valor del error de estimación asociado a la estimación.

4. A partir de un gran número de mediciones, se sabe que un método para determinar la cantidad de manganeso en un mineral comete errores aleatorios con distribución normal de media cero y desviación estándar 0.09.

   (a) Se hicieron 5 mediciones de un mismo mineral y se obtuvo un valor promedio de 7.54, calcule un intervalo de confianza con nivel del 99% para la cantidad de manganeso verdadera que contiene ese mineral.

   (b) Ídem a), pero si se hicieron 10 mediciones.

(c) ¿Cuántas mediciones habría que hacer para que el intervalo de confianza al 99% tenga una longitud ≤ 0.10?

5. Se registró el valor (en Kg) de la reducción del peso, de cada uno de 16 pacientes elegidos al azar, después de una semana de tratamiento. El promedio de esos 16 valores fue de 3.42Kg. Suponga que la pérdida de peso luego de una semana de tratamiento es una variable aleatoria con distribución normal.

   Construya un intervalo de confianza del 99% para el valor medio poblacional de la reducción del peso después de una semana de tratamiento, asumiendo $\sigma = 0.68$Kg. (σ conocido).

6. Considere nuevamente el problema planteado en el ejercicio 5. Construya un intervalo de confianza del 99% para el valor medio poblacional de la reducción del peso después de una semana de tratamiento, pero ahora σ es desconocido y $s = 0.68$Kg. Compare la longitud del intervalo obtenido con el que contruyo en el ejercicio 5.

7. Diez mediciones de recuperación de bromuro potásico por cromatografía de gas-líquido en muestras de tomates de cierta partida arrojaron una media muestral de $782\mu g/g$ y un desvío $s = 16.2\mu g/g$. Suponga que las mediciones tienen distribución normal.

   (a) Halle un intervalo de confianza del 95% para la media (μ) de las mediciones en esta partida de tomates.
   (b) Suponiendo que el error de medición debido al método es despreciable respecto a la variabilidad entre los tomates y que las 10 mediciones se realizan para cada uno de 10 tomates, ¿cómo debe interpretarse la media μ? ¿Cómo debe interpretarse σ^2?
   (c) En cambio, si las 10 mediciones se realizaron sobre el mismo tomate, ¿cómo puede interpretarse μ? ¿Cómo debe interpretarse σ^2?
   (d) Halle un intervalo de confianza del 95% para la varianza $\left(\sigma^2\right)$ de las mediciones.
   (e) Halle un intervalo de confianza del 95% para la desviación estándar (σ) de las mediciones.

8. Se hicieron varias mediciones del contenido de glucosa de una solución. Suponga que estas mediciones siguen un modelo de Gauss sin sesgo, es decir, el modelo descripto en el ejercicio 2 de la Práctica 3, cuando los errores tienen distribución normal.

   (a) Escriba el modelo. Se calculó el intervalo de confianza del 95% para la media, que resultó ser $(10.28, 11.32)$. ¿Qué significa "la media" en este problema?
   (b) Decir si es verdadero o falso, y explicar:
      i. Un 95% de las mediciones observadas pertenecen a ese intervalo.
      ii. Hay una probabilidad de 0.95 de que la próxima medición caiga en el intervalo.
      iii. Alrededor del 95% de las veces que uno realice el ensayo y construya el intervalo de confianza, éste contendrá la verdadera concentración de glucosa de la solución.
      iv. La probabilidad de que el intervalo $(10.28, 11.32)$ contenga a la verdadera concentración de glucosa es de 0.95.

9. Utilizando dos métodos de análisis se hicieron determinaciones del contenido de hierro de una muestra de un mineral. Se asume que las determinaciones correspondientes a ambos métodos tienen

distribución normal cuyo desvío estándar es $\sigma = 0,1\%$. Los resultados obtenidos son los siguientes:

Método 1	$n_1 = 12$	$\overline{x} = 15.22\%$
Método 2	$n_2 = 11$	$\overline{y} = 15.30\%$

Calcule la estimación por intervalo de la diferencia de medias de ambos métodos con nivel de significación del 95%.

10. Este es un ejemplo en el que se desea estudiar si un cambio en las condiciones de un experimento afecta el resultado. Se está estudiando un procedimiento para la determinación de estaño en productos alimenticios. Para ello se tomaron 12 muestras del mismo producto. Se eligieron 6 de estas muestras al azar y se llevaron al punto de ebullición con HCl a reflujo durante 30 minutos. Con las 6 muestras restantes el tiempo de reflujo fue de 70 minutos. Los resultados fueron los siguientes:

Tiempo de reflujo (m)	Estaño encontrado (mg/Kg)	$\overline{x}$	s^2
30	57, 57, 59, 56, 56, 59	57.33	1.867
70	57, 55, 58, 59, 59, 59	57.83	2.57

Suponga que la cantidad de estaño encontrada cuando el tiempo de reflujo es de 30 minutos es una variable aleatoria con distribución $N(\mu_1, \sigma_1^2)$ y dicha distribución es $N(\mu_2, \sigma_2^2)$ cuando el tiempo de reflujo es de 70 minutos.

(a) Calcule la estimación por intervalo del cociente de varianzas de estaño para los dos tiempos de ebullición con nivel de significación del 95%.

(b) Suponga que $\sigma_1 = \sigma_2$. Calcule el intervalo de confianza estimado con nivel de significación del 95% para la diferencia de contenido de estaño para los dos tiempos de ebullición.

11. Una fábrica de calzado desea comparar dos tipos de suela (marca A y marca B). Para ello, se coloca una suela de 2 cm. en cada zapato de 6 individuos y se miden las suelas de cada pie después de 3 meses. La tabla a continuación presenta el grosor de las suelas transcurrido ese tiempo.

Individuo	Suela Izquierda (marca A)	Suela Derecha (marca B)
1	1.43	1.42
2	1.27	1.24
3	1.48	1.39
4	1.53	1.41
5	1.71	1.6
6	1.72	1.61

Se sabe que la diferencia de grosores entre suelas de cada individuo se distribuye normalmente. Calcule la estimación por intervalo para la diferencia media de grosores entre marcas de suela a nivel 0,95 ¿Hay diferencia entre la calidad de las suelas? Justifique