Laboratorio 7 - Inferencia Estadística

Estimación por intervalos pt. 1 - Conceptos y el método pivotal

Laboratorista: Héctor Lira Talancón

Ago-Dic 2017

1. Sea X el ahorro mensual que hace una persona en la sucursal A, y Y el ahorro mensual que hace una persona en la sucursal B. La dirección del Banco del Ahorro sospecha que hay diferencia entre la variabilidad de X y la de Y. Se planea un estudio en donde se seleccionan aleatoriamente 13 y 16 cuentahabientes de las sucursales A y B, respectivamente. Los resultados fueron:

	N	media	desv. est.
Suc. A (X)	13	2.488	1.6
Suc. B (Y)	16	2.926	1.1

Construir un intervalo de confianza para la diferencia $\mu_X - \mu_Y$ (90% de confianza). ¿Se puede concluir que el ahorro mensual promedio en las dos sucursales es el mismo?

RESPUESTA

Asumimos que X e Y siguen una distribución normal con varianza común.

i) Definimos la cantidad pivotal $Q(\underline{\mathrm{X}}; \mu_X, \mu_Y)$ e identificamos su distribución:

$$Q(\underline{\mathrm{X}}; \mu_X, \mu_Y) = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{s_p\sqrt{\frac{1}{N_X} + \frac{1}{N_Y}}} \sim t_{(N_X+N_Y-2)}$$

ii) Establecemos $P(a \leq Q(\underline{\mathrm{X}}; \mu_X, \mu_Y) \leq b) = 0.90$

$$P(t_{0.05,(27)}) = 0.90$$

2. A partir de una muestra aleatoria de tamaño 17 se construye un intervalo de confianza para el parámetro θ, obteniéndose $\theta \in [38, 45]$ con 96% de confianza. Lo anterior significa que:

a) Si se tomaran otras 50 muestras aleatorias de la misma población y se estimara con cada muestra el valor de θ, es decir, se calcula $\hat{\theta}$, aproximadamente 48 de tales estimaciones pertenecerían al intervalo $[38, 45]$.

b) Si se tomaran otras 50 muestras aleatorias de la misma población y con cada una se construyera un intervalo de confianza, aproximadamente en 48 de tales intervalos se encontraría el valor de θ.

c) La probabilidad de que $\hat{\theta}$ se encuentre fuera del intervalo $[38, 45]$ es igual a 0.04.

d) Ninguna de las anteriores.

3. En un estudio de opinión a nivel nacional realizado el pasado mes de noviembre sobre preferencias para la elección de diputados federales se obtuvieron los siguientes resultados.

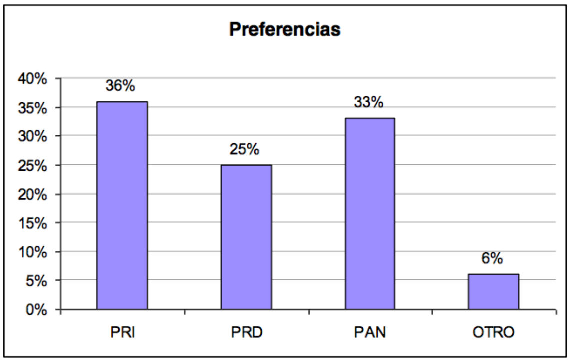

Figura 1: Preferencias elecciones de diputados federales, Nov 2005

Un periodista afirma que la preferencia de las personas en la elección de diputados federales es de 40% para el PRI. ¿La información mostrada confirma lo que dice el periodista? Construya un intervalo de confianza al 95% de confianza para responder esta pregunta.

4. En la siguiente tabla se resume la información de 60 personas entrevistadas, las cuales habían mencionado previamente haber trabajado durante el mes anterior.

	N	media	desv. est.
Hombres	39	5,149.86	5,309.54
Mujeres	21	3,688.69	3,486.60
Total	60	4,624.06	4,786.72

Suponga que los ingresos siguen una distribución normal.

a) Estime mediante un intervalo al 90% el ingreso promedio mensual de las trabajadoras mexicanas.

b) Proporcione un intervalo inferior al 90% para el ingreso mensual de las trabajadoras mexicanas.

5. Sea $X_1, ..., X_n$ una muestra aleatoria de una distribución normal con media desconocida μ pero con varianza σ^2 conocida. Se estima μ mediante un intervalo de confianza de $100(1-\alpha)\%$. Si se duplica el tamaño de la muestra, entonces la longitud del nuevo intervalo:

a) Disminuye a la mitad.

b) Aumenta al doble.

c) Disminuye.

d) Aumenta.

6. Los siguientes datos son los tiempos (en horas) entre las fallas de un equipo de aire acondicionado de un cierto avión: $74, 57, 48, 29, 502, 12, 70, 21, 29, 386, 59, 27, 153, 26, 326$. Asuma que los datos son valores observados de una variable aleatoria cuya distribución es una Exponencial ($X_i \sim Exp(\theta)$). Encuentre un intervalo de confianza del 90% para el tiempo promedio entre las fallas, θ (Sugerencia: se sabe que $\frac{2n\bar{X}}{\theta} \sim \chi^2_{(2n)}$).

7. Sea $X_1, ..., X_n$ una muestra aleatoria de una población cuya densidad está dada por

$$f(x;k) = k(1+x)^{-k-1}; x > 0, k > o$$

Encuentre un intervalo de confianza de 95% para k a partir de los siguientes valores que tomaron las variables de nuestra muestra aleatoria: $0.85, 1.08, 0.35, 3.28, 1.24, 2.58, 0.02, 0.13, 0.22, 0.52$.

8. En una empresa se obtuvo mediante una encuesta a 36 de sus empleados de la parte operativa y un intervalo de confianza del 95% que el salario promedio estaba en el intervalo $(4500, 6000)$. Si se desea cambiar el nivel de confianza al 99%, el intervalo sería:

   a) Más estrecho e involucraría un riesgo más grande de ser incorrecto.

   b) Más ancho e involucraría un riesgo más pequeño de ser incorrecto.

   c) Más estrecho e involucraría un riesgo más pequeño de ser incorrecto.

   d) Más ancho e involucraría un riesgo más grande de ser incorrecto.

9. Suponga que una mejor medida sobre la efectividad de una campaña publicitaria para determinado artículo es el aumento en las ventas de dicho artículo, es decir, $D_i = Y_i - X_i, i = 1, ..., 101$, donde Y_i es la venta del artículo en la $i-$ésima tienda medida después de la campaña y X_i es la venta del artículo antes de la campaña. Si la media muestral del cambio en ventas es 5.3 (en millones de pesos) y la varianza muestral del cambio en ventas es 144.0, determine el intervalo de confianza al 99% para el cambio promedio en ventas y determine a partir de éste si la campaña fue efectiva.

10. La gerencia de un banco desea comparar dos sistemas de cómputo para atención a clientes. El sistema 1 es el actual y el sistema 2 es el nuevo. En condiciones similares se obtienen dos muestras aleatorias independientes de tamaño 70 de los tiempo de servicio con cada uno de los dos sistemas. Calcula el intervalo al 95% de confianza para la diferencia de medias si $\bar{X}_1 = 3.1, S_1 = 1.8, \bar{X}_2 = 2.3, S_2 = 2.1$.

11. Sea $Y_1, ..., Y_n$ una muestra aleatoria de una población con media μ y varianza σ^2, ambos parámetros desconocidos. De las siguientes opciones, ¿cuál es siempre un pivote exacto o asintótico? Definimos $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$.

    a) $\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}}$

    b) $\frac{(n-1)s^2}{\sigma^2}$

    c) $\frac{\bar{Y}-\mu}{s/\sqrt{n}}$

    d) Ninguna de las anteriores.

12. Se tiene una muestra aleatoria de n personas de una cierta población. Se mide lo siguiente: X = longitud del brazo derecho y Y = longitud del brazo izquierdo. Supongamos que se desea obtener un intervalo de confianza para la diferencia en la longitud media de los brazos en la población. Se consideran las siguientes variables aleatorias:

$$t_A = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{n}}}$$

$$t_B = \frac{\bar{D} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_D^2}{n}}}$$

    donde $D_i = X_i - Y_i$ para $i = 1, ..., n$.

    Entonces,

    a) t_A es apropiado y t_B no es apropiado para obtener el intervalo.

    b) Tanto t_A como t_B son apropiados para obtener el intervalo.

    c) t_B es apropiado y t_A no es apropiado para obtener el intervalo.

    d) Ni t_A ni t_B son apropiados para obtener el intervalo.

13. Sea $X_1, ..., X_n$ una muestra aleatoria de una densidad $Exp(\beta)$. Se sabe que el estimador de máxima verosimilitud para β es $\hat{\beta} = \bar{X}$ y que $\frac{\hat{\beta}-\beta}{\hat{\beta}/\sqrt{n}} \xrightarrow{d} N(0,1)$ cuando $n \to \infty$. Use este resultado para obtener un intervalo de confianza al 90% de confianza para β si se sabe que $X_1 + ... + X_{75} = 120$.

14. Considera $X_1, ..., X_n$ una muestra aleatoria de una densidad $f(x|\theta) = \frac{x}{\theta}\exp(-x^2/2\theta)\mathbb{1}_{(0,\infty)}(x)$ con $\theta > 0$. Se sabe que el estimador de máxima verosimilitud para θ es $\hat{\theta} = \frac{\sum_{i=1}^{n} X_i^2}{2n}$ y que $\hat{\theta} \mathrel{\dot\sim} N(\theta, \frac{\theta^2}{n})$. Utiliza este resultado para encontrar un pivote asintótico y construye un intervalo al 99% de confianza para θ si $n = 120$ y $\sum_{i=1}^{n} X_i^2 = 340$.

15. Se toman muestras independientes de tamaño n_j de cada una de cinco poblaciones. Supón que la población j tiene una distribución Normal con media μ_j y varianza σ, para $j = 1, 2, ..., 5$. Es decir, todas las poblaciones están distribuidas normalmente con la misma varianza, pero (tal vez) con diferentes medias. Los valores de $s_j^2 = \frac{\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2}{n_j-1}$ y n_j son:

j	1	2	3	4	5
s_j^2	40	30	20	42	50
n_j	6	4	3	7	8

Encuentra un intervalo de confianza del 95% para σ^2.