Inferencia Estadística

Marisol García Peña

Departamento de Matemáticas
Pontificia Universidad Javeriana

Bogotá, 2022

Estimación por intervalos

Estimación por intervalos

- Estimación puntual $\Longrightarrow$ Un solo valor.
- No hay información sobre la precisión y confiabilidad de la estimación.
- Estimador $\Longrightarrow$ f.d.p. $\Longrightarrow$ prob. que el estimador "igual" al parámetro es cero. (v.a. cotinuas).
- Estimación puntual debe ir acompañada de alguna medida del posible error de la estimación.
- De estimación puntual de θ a estimación de que θ esté contenido en algún intervalo.

Estimación por intervalos (cont.)

- Alternativa $\Longrightarrow$ Intervalo aleatorio que contenga a θ, con cierta probabilidad.
- Nivel de confianza $(1-\alpha)$ $\Longrightarrow$ Probabilidad de que un intervalo de confianza contenga al verdadero valor del parámetro.
- Nivel de confianza $(1-\alpha)$ $\Longrightarrow$ El $100(1-\alpha)\,\%$ de las muestras darían lugar a un intervalo que incluye θ y solo $100\alpha\,\%$ de las muestras producirían un intervalo erróneo.
- Cuanto mayor sea el nivel de confianza $\Longrightarrow$ Mayor probabilidad de que el valor del parámetro esté dentro del intervalo.

Estimación por intervalos (cont.)

Intervalo de confianza

Sea $X_1, \ldots, X_n$ una muestra aleatoria con función de densidad o de probabilidad $f(\bullet;\theta)$. Sean $T_1 = t_1(X_1, \ldots, X_2)$ y $T_2 = t_2(X_1, \ldots, X_2)$ dos estadísticas tales que $T_1 < T_2$ y $P[T_1 < \tau(\theta) < T_2] = \gamma = 1 - \alpha$, el intervalo (T_1, T_2) se conoce como intervalo de cofianza para $\tau(\theta)$ del 100γ por ciento, $\gamma = 1 - \alpha, 0 < \alpha < 1$ es el coeficiente/nivel de confianza y T_1 y T_2 son los límites inferior y superior para $\tau(\theta)$. α se conoce como el nivel de significancia.

Estimación por intervalos (cont.)

- Un valor (t_1, t_2) del intervalo aleatorio (T_1, T_2) también es llamado un IC para $\tau(\theta)$ del $100(1-\alpha)$ por ciento.
- $P[h_1(T_1) \leq \tau(\theta) \leq h_2(T_2)] = 1-\alpha$, h_1 y h_2 funciones de un estadístico T, relacionado con el parámetro a estimar.
- Apmlitud del intervalo de confianza, $A(X_1, \ldots, X_n) = T_2(X_1, \ldots, X_n) - T_1(X_1, \ldots, X_n)$.
- Amplitud esperada, $E[A(X_1, \ldots, X_n)]$.
- Intervalos unilaterales.

Estimación por intervalos (cont.)

- Cuantifica el riesgo. Por ejemplo, 95 %.
- Si son tomadas 100 muestras del mismo tamaño $\Longrightarrow$ en el 95 % de los casos los intervalos deberían contener el parámetro θ o $\tau(\theta)$.
- Coeficiente de confianza mayor $\Longrightarrow$ IC más amplio.
- Mayor confianza $\Longrightarrow$ pérdida de precisión.

Estimación por intervalos (cont.)

Intervalo de confianza unilateral

Sea $X_1, \ldots, X_n$ una muestra aleatoria con con función de densidad o de probabilidad $f(\bullet; \theta)$. Sea $T_1 = t_1(X_1, \ldots, X_n)$ una estadística tal que $P[T_1 < \tau(\theta)] = \gamma = 1 - \alpha$, T_1 se conoce como el límite inferior de confianza unilateral. Similarmente, $T_2 = t_2(X_1, \ldots, X_n)$ es una estadística tal que $P[\tau(\theta) < T_2] = \gamma = 1 - \alpha$, T_2 se conoce como el límite superior de confianza unilateral.

Estimación por intervalos (cont.)

Ejemplo

Sea $X_1, \ldots, X_n$ una muestra aleatoria con $f(x, \theta) = \phi_{\theta,9}(x)$. Sea $T_1 = t_1(X_1, \ldots, X_n) = \overline{X} - \frac{6}{\sqrt{n}}$ y $T_2 = t_2(X_1, \ldots, X_n) = \overline{X} + \frac{6}{\sqrt{n}}$.

(T_1, T_2) constituye un IC para $\tau(\theta) = \theta$ con nivel de confianza $1 - \alpha$

$$\begin{aligned}
1 - \alpha &= P\left[\overline{X} - \frac{6}{\sqrt{n}} < \theta < \overline{X} - \frac{6}{\sqrt{n}}\right] \\
&= P\left[-2 < (\overline{X} - \theta)/((3)/\sqrt{n}) < 2\right] \\
&= \Phi(2) - \Phi(-2) = 0.9772 - 0.0228 \\
&= 0.9544
\end{aligned}$$

Cantidad pivotal

Sea $X_1, \ldots, X_n$ una muestra aleatoria con función de densidad o de probabilidad $f(\bullet, \theta)$. Sea $Q = q(X_1, \ldots, X_n, \theta)$, una función de las medidas muestrales y de θ, donde el único elemento desconocido es θ y si la distribución de Q no depende de θ, entonces Q es definida como *cantidad pivotal*.

Distribución de prob. cantidades pivotales $\Longrightarrow$ normal estándar, t, χ^2 o F.

Estimación por intervalos (cont.)

Algunas cantidades pivotales si $X_1, \ldots, X_n$ es $N(\mu, \sigma^2)$.

- Con μ desconocido y σ conocido, sea $\overline{X}$ la media muestral. La cantidad pivotal es $\dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$.

- Con μ desconocido y σ desconocido. El pivote es $\dfrac{\overline{X} - \mu}{S / \sqrt{n}} \sim t_{(n-1)}$. Si n es "grande" usando el TCL la distribución del pivote es $N(0, 1)$.

- Si σ^2 es desconocida, la cantidad pivotal es $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

Estimación por intervalos (cont.)

Método pivotal

Si $Q = q(X_1, \ldots, X_n, \theta)$ es una cantidad pivotal y tiene una f.d.p, entonces para un valor fijo $0 < \gamma < 1, \gamma = 1 - \alpha$ existen q_1 y q_2 que dependen de γ tal que $P[q_1 < Q < q_2] = \gamma$. Si para cada posible valor muestral $(x_1, \ldots, x_n)$, $q_1 < q(x_1, \ldots, x_n, \theta) < q_2$ si y sólo si $t_1(x_1, \ldots, x_n) < \tau(\theta) < t_2(x_1, \ldots, x_n)$ para funciones t_1 y t_2 (que no dependen de θ), entonces (T_1, T_2) es un IC para $\tau(\theta)$ del 100γ por ciento, donde $T_i = t_i(X_1, \ldots, X_n), i = 1, 2$.

$P[q_1 < Y < q_2]$ no es afectada por un cambio de escala o una traslación de la v.a. Y.

Estimación por intervalos (cont.)

- q_1 y q_2 son independientes de θ pues la distribución de Q lo es.
- Para un valor fijo de $\gamma = 1 - \alpha$ existen muchos pares de valores para q_1 y q_2 que se pueden seleccionar tal que $P[q_1 < Q < q_2] = \gamma$.
- Diferentes valores de q_1 y q_2 producirán distintos valores de t_1 y t_2.
- La idea es seleccionar a q_1 y q_2 de tal que t_1 y t_2 estén "cerca".
- Por ejemplo, reduciendo la amplitud del intervalo.

Estimación por intervalos (cont.)

Uso del método del método pivotal.

- Encuentre un estimador de θ o $\tau(\theta)$. Generalmente, el EMV.
- Encuentre una función de θ y el estimador (cantidad pivotal, $Q(\boldsymbol{X}, \theta)$).
- Encuentre q_1 y q_2 tales que $P[q_1 \leq Q \leq q_2] = 1 - \alpha$. Escoja q_1 y q_2 tal que $P[Q \leq q_1] = \alpha/2$ y $P[Q \geq q_2] = \alpha/2$.
- Transforme el IC pivotal al IC para el parámetro θ tal que $P[T_1 \leq \theta \leq T_2] = 1 - \alpha$.

Ejemplo

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una $N(\mu, 1)$. Construya un intervalo del 95 % para μ.