Estadística Inferencial

Jose UGARTE - Clifford TORRES

Facultad de Ciencias - UNI (Clase 3)

September 3, 2019

Introducción

- Para saber la bondad de un **estimador puntual** se define el **error de estimación**, el cual se define como la distancia entre el estimador puntual y el parámetro objetivo.
- Debemos de tener en cuenta que dicho error varia aleatoriamente en muestreo repetitivo.

Definición

Dado un **estimador** $\hat{\theta}$ *y su* **parámetro objetivo** θ. *Llamamos* **error de estimación** *a la cantidad:*

$$\epsilon = |\hat{\theta} - \theta|$$

Pregunta

Observación

Siendo el estimador puntual $\hat{\theta}$ una variable aleatoria entonces ϵ es también una variable aleatoria.

Pregunta: Si ϵ es una variable aleatoria como podemos sacar información relevante de esta ?

- Lo que si se puede hacer es verificar si la probabilidad de dicha variable aleatoria es menor que una cierta cantidad, digamos **b**, así, estamos enfocados en :

$$P(|\epsilon| < b) = P(|\hat{\theta} - \theta| < b)$$

- De esta manera tenemos una medida de la bondad de dicho error, de forma probabilista.
- Esta es una medida de cuantas veces el estimador $\hat{\theta}$ cae dentro del intervalo $[\theta - b, \theta + b]$.

Calculo de b y estimación de la probabilidad de ϵ

- Un problema recurrente es determinar **b** tal que $P(|\epsilon| < b) = 0.9$. En el caso de tener la densidad de probabilidad de $\hat{\theta}$ entonces debemos de resolver la ecuación en **b** :

$$\int_{\theta-b}^{\theta+b} f(\hat{\theta}) \mathrm{d}\hat{\theta} = 0.9$$

Calculo de b y estimación de la probabilidad de ϵ

- En el caso de que $\hat{\theta}$ sea *insesgado*($E(\hat{\theta}) = \theta$) podemos utilizar el teorema de Tchebysheff para obtener un *limite inferior* de dicha probabilidad :

$$1 - \frac{1}{k^2} \leq P(|\hat{\theta} - \theta| \leq k\sigma_{\hat{\theta}})$$

- De esta manera, por ejemplo, para $k = 2$ nos dice que la probabilidad de que $\hat{\theta}$ se encuentre entre $[\theta - 2\sigma_{\hat{\theta}}, \theta + 2\hat{\theta}]$ es de *por lo menos* 0.75.
- Si incrementamos k obtenemos un intervalo muy grande lo cual no es de gran ayuda.

Ejemplo: Aplicación de Tshebysheff

Ejemplo

El número de clientes por día en un mostrador de ventas, Y, ha sido observado durante un largo periodo y se encontró que tiene una media de 20 y desviación estándar de 2. La distribución de probabilidad de Y no se conoce. ¿Qué se puede decir acerca de la probabilidad de que, mañana, Y sea mayor que 16 pero menor que 24?

Ejemplo

Suponga que la experiencia ha demostrado que el tiempo Y (en minutos) necesario para realizar una prueba periódica de mantenimiento en una máquina de dictados sigue una distribución gamma con $\alpha = 3,1$ y $\beta = 2$. Un trabajador de mantenimiento de nuevo ingreso tarda $22,5$ minutos en probar la máquina. ¿El tiempo para realizar la prueba está en desacuerdo con la experiencia anterior?

Calculo de b y estimación de la probabilidad de ϵ

- Para poder aplicar los casos anteriores se debe de calcular el valor del error standard de $\hat{\theta}$(es decir ($\sigma_{\hat{\theta}}$)), para esto usaremos la tabla siguiente:

Tabla 1 **Valores esperados y errores estándar de algunos estimadores puntuales comunes**

Parámetro objetivo θ	Tamaño(s) muestral(es)	Estimador puntual $\hat{\theta}$	$E(\hat{\theta})$	Error estándar $\sigma_{\hat{\theta}}$
μ	n	$\overline{Y}$	μ	$\dfrac{\sigma}{\sqrt{n}}$
p	n	$\hat{p} = \dfrac{Y}{n}$	p	$\sqrt{\dfrac{pq}{n}}$
$\mu_1 - \mu_2$	n_1 y n_2	$\overline{Y}_1 - \overline{Y}_2$	$\mu_1 - \mu_2$	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ *†
$p_1 - p_2$	n_1 y n_2	$\hat{p}_1 - \hat{p}_2$	$p_1 - p_2$	$\sqrt{\dfrac{p_1q_1}{n_1} + \dfrac{p_2q_2}{n_2}}$ †

* σ_1^2 y σ_2^2 son las varianzas de las poblaciones 1 y 2, respectivamente.

- † Se supone que las dos muestras son independientes.

Calculo de b y estimación de la probabilidad

- De otro lado, en casos prácticos: de la vida real, o concretos, o si $\hat{\theta}$ tiene una densidad de probabilidad conocida, podemos utilizar la tabla:
-

Tabla 8.2 Probabilidad de que $(\mu - 2\sigma) < Y < (\mu + 2\sigma)$

Distribución	Probabilidad
Normal	.9544
Uniforme	1.0000
Exponencial	.9502

Ejemplo: Estimación de parámetros

Ejemplo

Una muestra de $n = 1000$ *votantes, seleccionados al azar en una ciudad, mostró* $y = 560$ *a favor del candidato Jones. Estime* p*, la fracción de votantes de la población que están a favor de Jones y precise un límite de error estándar de 2 en el error de estimación.*

Ejemplo: Estimación de parámetros

Ejemplo

Una comparacin de la durabilidad de dos tipos de llantas para automóvil se obtuvo de muestras de pruebas en carretera de $n_1 = n_2 = 100$ *llantas de cada tipo. Se registró el número de millas hasta quedar inútiles, el desgaste se definió como el número de millas hasta que la cantidad restante de superficie de rodamiento llegó a un valor pequeño especificado previamente. Las mediciones para los dos tipos de llantas se obtuvieron de manera independiente y se calcularon las siguientes medias y varianzas:*

- $\bar{y}_1 = 26,400$ *millas,* $\bar{y}_2 =$ *25,100 millas,*
- $s_1^2 = 1,440,000$*,* $s_2^2 = 1,960,000$*.*

Estime la diferencia en la media de millas hasta quedar inútiles y precise un límite de error estándar de 2 en el error de estimación.

Método de las transformaciones

Dadas una variable aleatoria Y con función de densidad f_Y, y una función estrictamente creciente o estrictamente decreciente h. Entonces la pregunta es: Como calcular **la función de densidad de** $U = h(Y)$?

- Veamos el procedimiento constructivo, si h es **creciente**:
- $P(U \leq u) = P(h(Y) \leq u) = P(h^{-1}(h(Y)) \leq h^{-1}(u)) = P(Y \leq h^{-1}(u))$
- $F_U(u) = P(U \leq u) = P(Y \leq h^{-1}(u)) = F_Y(h^{-1}(u))$.
- $f_U(u) = f_Y(h^{-1}(u))\frac{\mathrm{d}h^{-1}(u)}{\mathrm{du}}$.

Método de las transformaciones: Ejemplo

Dada la función

$$f_Y(y) = \begin{cases} 2y, & 0 \leq y \leq 1 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

Determinar la función de densidad de $U = e^{3Y}$.

- El primer paso determinar h^{-1}. En nuestro caso, $h^{-1}(u) = \frac{1}{3}\ln(u)$.
- Determinar $\frac{\mathrm{d}h^{-1}(u)}{\mathrm{d}u} = \frac{1}{3u}$.
- Finalmente, $f_U(u) = f_Y(h^{-1}(u))\frac{\mathrm{d}h^{-1}(u)}{\mathrm{d}u}$

$$f_U(u) = \begin{cases} 2(\frac{1}{3}\ln(u))\frac{1}{3u}, & 0 \leq \frac{1}{3}\ln(u) \leq 1 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

Metodo de las transformaciones:

- Veamos el procedimiento constructivo, si h es **decreciente**:
- $P(U \leq u) = 1 - P(Y \geq h^{-1}(u))$
- $F_U(u) = P(U \leq u) = 1 - F_Y(h^{-1}(u))$.
- $f_U(u) = -f_Y(h^{-1}(u))\frac{\mathrm{d}h^{-1}(u)}{\mathrm{du}} = f_Y(h^{-1}(u))|\frac{\mathrm{d}h^{-1}(u)}{\mathrm{du}}|$.