Parcial 3

Wednesday, November 17, 2021 9:57 AM

David Alsina

1. (**25 pts**) Suponga que $Y_1, \dots Y_n$ denotan una muestra aleatoria de una distribución uniforme en el intervalo $(\theta, \theta+1)$. Sean: $\hat{\theta}_1 = \bar{Y} - \frac{1}{2}$, $\hat{\theta}_2 = Y_n - \frac{n}{n+1}$

a ▪ Demuestre que $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores insesgados.

b ▪ Encuentre la eficiencia relativa de $\hat{\theta}_1$ respecto a $\hat{\theta}_2$

(a) $\circledast\ E(\hat{\theta}_1) = E\left(\bar{Y} - \frac{1}{2}\right) = E(\bar{Y}) - \frac{1}{2}$

$$= \mu - 1/2$$
$$= \frac{\theta + (\theta+1)}{2} - 1/2$$
$$= \frac{2\theta+1}{2} - 1/2$$
$$= \theta + \frac{1}{2} - 1/2 = \theta$$

$\circledast\ E(\hat{\theta}_2) = E\left(Y_n - \frac{n}{n+1}\right) = E(Y_n) - \frac{n}{n+1}$

$$= \mu - (n/n+1)$$
$$= \left(\theta + \frac{1}{2}\right) - \left(n/n+1\right)$$

$\hat{\theta}_2$ es insesgado cuando $n = 1$.

(b) encuentre $\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}$

Ahora $var(\hat{\theta}_2) = E\left(\left(\hat{\theta}_2 - E(\hat{\theta}_2)\right)^2\right)$

$$= E\left(\left(\left[Y_n - \frac{n}{n+1}\right] - \left[\theta + \frac{1}{2} - \frac{n}{n+1}\right]\right)^2\right)$$
$$= E\left(\left(Y_n - \theta - \frac{1}{2}\right)^2\right)$$
$$= E\left(Y_n^2 - 2Y_n\theta - Y_n + \theta^2 + \theta + \frac{1}{4}\right)$$
$$= E(Y_n^2) - 2\theta \cdot E[Y_n] - E[Y_n] + \theta^2 + \theta + \frac{1}{4}$$
$$= E(Y_n^2) - 2\theta\left(\theta + \frac{1}{2}\right) - \left(\theta + \frac{1}{2}\right) + \theta^2 + \theta + \frac{1}{4}$$
$$= \boxed{-2\theta^2} - \theta\ \boxed{-\theta} - \frac{1}{2} + \boxed{\theta^2} \boxed{+\theta} + \frac{1}{4} + E[Y_n^2]$$
$$= -\theta^2 - \theta - \frac{1}{4} + E[Y_n^2]$$
$$= \frac{(\theta+1)^3 - (\theta)^3}{3(\theta+1) - 3(\theta)} - \theta^2 - \theta - \frac{1}{4}$$
$$= \cancel{\theta^2} + \cancel{\theta} + \frac{1}{3} + \left(-\cancel{\theta^2} - \cancel{\theta} - \frac{1}{4}\right)$$
$$= \frac{1}{12}$$

Calculando $var(\hat{\theta}_1) = var\left(\bar{Y} - \frac{1}{2}\right) = var(\bar{Y})$

$$= \frac{\sigma^2}{n} = \frac{1}{12n}\left((\theta+1) - \theta\right)^2$$

$$= \frac{v}{n} = \frac{1}{12n}(\theta+1)-\theta)$$

$$= \frac{1}{12n}\theta^2$$

Vea que $\dfrac{var(\hat{\theta}_2)}{var(\hat{\theta}_1)} = \dfrac{\frac{1}{12}}{(\frac{1}{12n})\theta^2} = \dfrac{n}{\theta^2}$ si $n \to \infty$ $eff \to \infty$

lo que quiere decir que $\hat{\theta}_1$ es más eficiente.

---

2. (**25 pts**) Se ha estimado que la proporción de personas que son aficionados a los deportes es una variable aleatoria X con función de densidad:

$$f_X(x|\theta) = \begin{cases} \theta x^{\theta-1} & si \quad 0 < x < 1 \\ 0 & d.l.c. \end{cases}$$

Determine el MLE de θ

Suponiendo que hay $x_1, \dots, x_n$ muestras aleatorias

$$L(x_1, \dots, x_n \mid \theta) = \theta^n \cdot \prod_{i=1}^{n} x_i^{\theta-1}$$

Ahora si buscamos maximizar $L(x_1, \dots, x_n | \theta)$

Vemos que es más fácil maximizar $Ln(L(x_1, \dots, x_n|\theta))$:

$$Ln\left(\theta^n \cdot \prod_{i=1}^{n} x_i^{\theta-1}\right)$$

$$= Ln(\theta^n) + \sum_{i=1}^{n} \ln(x_i^{\theta-1})$$

$$= n\ln(\theta) + \sum_{i=1}^{n} (\theta-1)\cdot\ln(x_i)$$

$$= n\ln(\theta) + (\theta-1)\cdot\sum_{i=1}^{n}\ln(x_i)$$

derivando $\frac{d}{d\theta} Ln(L(\theta))$ e igualando a cero para maximizar:

$$\frac{d}{d\theta}\ln(L(\theta)) = \frac{d}{d\theta}\left(n\ln(\theta) + (\theta-1)\sum_{i=1}^{n}\ln(x_i)\right)$$

$$= \frac{n}{\theta} + (1)\sum_{i=1}^{n}\ln(x_i)$$

$$= \frac{n}{\theta} + \sum_{i=1}^{n}\ln(x_i) = 0$$

de aquí sale que θ es:

de aquí sale que θ es:

$$\frac{n}{\theta} = -\sum_{i=1}^{n} \ln(x_i)$$

$$n = -\theta \sum_{i=1}^{n} \ln(x_i)$$

$$\theta = -\frac{n}{\sum_{i=1}^{n} \ln(x_i)}$$

$$\frac{-n}{\sum_{i=1}^{n} \ln(x_i)}$$ es el MLE de θ.

---

3. (**25 pts**) Suponga que $Y_1, \ldots Y_n$ denotan una muestra aleatoria de una distribución Poisson con media λ. Encuentre un estimador de λ a través del método de momentos.

queremos estimar la Poisson con:

$$E(Y) = \lambda$$

ahora $M_1 = \frac{1}{n} \cdot \sum_{i=1}^{n} (Y_i)^1 = \overline{Y}$

igualando $E(Y) = \overline{Y}$ (Por método de momentos)

$$\lambda = \overline{Y}$$

---

4. (**25 pts**) Un grupo de investigadores quiere verificar si el día con más riesgo para sufrir un infarto en personas que trabajan es el lunes. Para esto, toman una muestra de 200 personas que sufrieron un infarto durante el último año y marcan el día en el que se dio el infarto. Los resultados obtenidos son los siguientes:

- (1) Lunes: 36
- (2) Martes: 27
- (3) Miércoles: 26
- (4) Jueves: 32
- (5) Viernes: 26
- (6) Sábado: 29
- (7) Domingo: 24

¿Los datos muestran evidencia para decir que hay diferencias significativas en el número de infartos que ocurren cada día? Use un $\alpha = 0.05$

Primero note que tenemos varias categorías por lo que será conveniente usar una prueba χ^2 y dado que el número de datos.

Sean $x_1, x_2, \ldots, x_7$ el # de infartos que ocurren en el día i de la semana día i de la semana

$K: 7$ (# de categorías), $H_0: p_1 = p_2 = p_3 = \dots = p_7$

dada la H_0 $p_i = \frac{1}{7}$

luego las cantidades esperadas para cada día

$$E(x_i) = np = \frac{200}{7} = 28.57$$

Ahora el χ^2 es

# Día	$n_i - E(x_i)$
1	7.43
2	−1.57
3	−2.57
4	3.43
5	−2.57
6	0.43
7	−4.57

$$\frac{\sum (n_i - E(x_i))^2}{E(x_i)} = \frac{103.7143}{28.57} = 3.6301$$

valor del estadístico

para obtener χ^2_α necesitamos los $g.l = K - \left(\begin{smallmatrix}\text{restricciones} \\ + \text{ aprox.}\end{smallmatrix}\right)$

la única restricción es $\sum_{i=1}^{7} p_i = 1$ luego $g.l = 7 - 1 = 6$

$\chi^2_{0.05,6} = 12.59$ así $RR := \{ \chi^2 > 12.59 \}$

Vea que el valor del estadístico no está en RR. por lo que no hay suficiencia de evidencias para rechazar H_0, con un nivel de 0.05.