

Introducción a la Estadística y Ciencia de Datos

GUÍA DE ACTIVIDADES - CLASE 6

1. Sea X_1, \dots, X_n una muestra aleatoria de una población con función de probabilidad

x	-1	0	1
$p(x, \theta)$	$(\theta - 1)^2$	$2(\theta - 1)(2 - \theta)$	$(2 - \theta)^2$

Queremos decidir cuál de los dos estimadores, el de momentos o el de máxima verosimilitud, es mejor. Para ello se pide

- Hallar la esperanza de ambos estimadores. ¿Son insesgados?
- Hallar la varianza de ambos estimadores. ¿Qué observa? ¿A qué se deben los resultados observados?

2. Sea X_1, \dots, X_n una muestra aleatoria de una población con densidad de la forma

$$f(x, \theta) = (1 - \theta)\mathbb{I}_{(-\frac{1}{2}, 0)}(x) + (1 + \theta)\mathbb{I}_{(0, \frac{1}{2})}(x), \quad -1 < \theta < 1.$$

Queremos decidir cuál de los dos estimadores, el de momentos o el de máxima verosimilitud, es mejor. Para ello se pide

- Hallar la esperanza de ambos estimadores. ¿Son insesgados?
- Hallar la varianza de ambos estimadores. ¿Cuál tiene menor error cuadrático medio?

3. Sea X_1, \dots, X_n una muestra aleatoria de una población con distribución $\Gamma(\alpha, \lambda)$. Sea el parámetro bivariado $\theta = (\alpha, \lambda)$.

- Determinar el estimador de θ basado en el primer y en el segundo momento.
- Hallar un sistema de ecuaciones cuya solución brinde el estimador de máxima verosimilitud de θ .
- Programar una función en R que a partir de un conjunto de datos, que se puede asumir provienen de una distribución $\Gamma(\alpha, \lambda)$, y una cantidad I de iteraciones del algoritmo de Newton-Raphson¹, devuelva la estimación por el método de máxima verosimilitud de (α, λ) según 3b). *Sugerencias:* explorar los comandos `digamma` y `trigamma` de R, y empezar el algoritmo de Newton-Raphson con la estimación por el método de momentos calculado en 3a).

¹Sea $f : [a, b] \rightarrow \mathbb{R}$ una función derivable y sea $x_0 \in [a, b]$. El algoritmo de Newton-Raphson brinda una aproximación de la solución $f(x) = 0$ en $[a, b]$ y se calcula mediante la expresión $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ con $f'(x_n) \neq 0 \forall n \in \mathbb{N}_0$.

d) Para cada $n \in \{6, 10, 20, 40, 80, 200\}$, hacer lo siguiente:

- i) Simular $N = 1000$ realizaciones de una muestra aleatoria $\Gamma(\alpha_0, \lambda_0)$ siendo $\alpha_0 = 3$ y $\lambda_0 = 4$ de tamaño n y calcular para cada realización de la muestra aleatoria:
 - la estimación de $\theta = (\alpha, \lambda)$ por el método de los momentos propuesto en 3a), que llamaremos $\tilde{\theta}_i = (\tilde{\alpha}_i, \tilde{\lambda}_i)$.
 - la estimación de θ por el método de máxima verosimilitud propuesto en 3c), que llamaremos $\hat{\theta}_i = (\hat{\alpha}_i, \hat{\lambda}_i)$.

Realizar histogramas y boxplots paralelos con los resultados obtenidos para las estimaciones de α y de λ .²

- ii) Computar los ECM muestrales

$$\frac{1}{N} \sum_{i=1}^N (\tilde{\alpha}_i - \alpha_0)^2 \quad , \quad \frac{1}{N} \sum_{i=1}^N (\hat{\alpha}_i - \alpha_0)^2$$

$$\frac{1}{N} \sum_{i=1}^N (\tilde{\lambda}_i - \lambda_0)^2 \quad , \quad \frac{1}{N} \sum_{i=1}^N (\hat{\lambda}_i - \lambda_0)^2$$

Finalmente, graficar n contra cada ECM muestral calculado, todo en un mismo gráfico. Hacer un zoom del gráfico para los tamaños muestrales que no presentan dificultades en su cómputo.

²Notar que para tamaños de muestra chica, el estimador de momentos no es un buen estimador y eso influye en la iteración del método de Newton–Raphson para calcular el EMV.