USP – ICMC – SME0806 - Estatística Computacional
3º trabalho – 1º/2017

A distribuição gama é utilizada em diversas aplicações envolvendo variáveis aleatórias positivas. Sua função densidade de probabilidade é dada por

$$f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} \exp(-x/\beta), \ x > 0, \tag{1}$$

em que $\alpha > 0$ é o parâmetro de forma, $\beta > 0$ é o parâmetro de escala e $\Gamma(\cdot)$ denota a função gama.

$X_1, \dots, X_n$, $n > 1$, representa uma amostra aleatória da distribuição em (1). É bem conhecido que os estimadores de máxima verossimilhança de α e β não têm forma fechada. Os estimadores de momentos[1] são dados por

$$\widehat{\alpha}_M = \frac{(\sum_{i=1}^n X_i)^2}{n\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \quad \text{e} \quad \widehat{\beta}_M = \frac{n\sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}{n\sum_{i=1}^n X_i}. \tag{2}$$

Pode ser provado que os estimadores

$$\begin{aligned} \widehat{\alpha} &= \frac{n\sum_{i=1}^n X_i}{n\sum_{i=1}^n X_i \log(X_i) - \sum_{i=1}^n \log(X_i)\sum_{i=1}^n X_i} \quad \text{e} \\ \widehat{\beta} &= \frac{1}{n^2}\left(n\sum_{i=1}^n X_i \log(X_i) - \sum_{i=1}^n \log(X_i) \sum_{i=1}^n X_i\right), \end{aligned} \tag{3}$$

são fortemente consistentes para (α, β) e tais que

$$\sqrt{n}(\widehat{\alpha} - \alpha) \xrightarrow{\text{d}} N\big(0, \alpha^2\{1 + \alpha\psi_1(1+\alpha)\}\big) \quad \text{e} \quad \sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{\text{d}} N\big(0, \beta^2\{1 + \alpha\psi_1(\alpha)\}\big), \tag{4}$$

quando $n \to \infty$, em que $\psi_1(\cdot)$ denota a função trigama.

Desenvolva um estudo de simulação com os seguintes objetivos:

1. Comparar os estimadores de (α, β) em (2) e (3).

2. Avaliar os erros padrão obtidos da aproximação das distribuições em (4).

3. Avaliar a aproximação pela distribuição normal em (4).

Como em (4) as variâncias assintóticas de $\widehat{\alpha}$ e $\widehat{\beta}/\beta$ não dependem de β, considere $\beta = 1$ e diferentes valores de $\alpha \in [0,2;5]$, sendo pelo menos 15 valores nos itens 1 e 2 e dois no item 3. Os tamanhos amostrais devem ser 20 e 50. Os resultados devem ser apresentados somente em forma de gráficos.

[1] Expressão de $\widehat{\beta}_M$ em (2) corrigida em 25/4/2017.