Triparametrična gama porazdelitev

Definirajte nepopolno gama funkcijo za $x > 0$ z

$$\Gamma(a;x) = \frac{1}{\Gamma(a)} \int_0^x t^{a-1}\, e^{-t}\, \mathrm{d}t$$

in logaritemski odvod gama funkcije z

$$\psi(x) = \frac{\mathrm{d}\, \log \Gamma(x)}{\mathrm{d}x} = \frac{\Gamma'(x)}{\Gamma(x)}\,.$$

Slučajna spremenljivka X, ki ima za $a > 0$, $\lambda > 0$ in $\tau > 0$ na $(0, \infty)$ porazdelitveno funkcijo

$$F_X(x) = \Gamma(a; (\lambda x)^\tau)\,,$$

je porazdeljena po transformirani gama porazdelitvi s parametri a, λ in τ, kar bomo označevali z $X \sim \mathrm{T}\Gamma(a, \lambda, \tau)$.

Neznane parametre a, λ in τ bomo ocenili po metodi največega verjetja. Predpostavljamo, da je vzorec enostavni slučajni s ponavljanjem, torej vzorčne vrednosti so neodvisne, enako porzadeljene slučajne spremenljivke $X_1, \ldots, X_n$. Logaritemska funkcija verjetja je

$$\begin{aligned}
\ell(a, \lambda, \tau) &= a\,\tau\, n\, \log \lambda + n\, \log \tau - n\, \log \Gamma(a) + \\
&\qquad + (a\tau - 1) \sum_{i=1}^n \log x_i - \lambda^\tau \sum_{i=1}^n x_i^\tau \\
&= n\,(a\,\tau\, \log \lambda + \log \tau - \log \Gamma(a) + (a\,\tau - 1)\,\overline{\log x} - \lambda^\tau\, \overline{x^\tau}).
\end{aligned}$$

a. Pokažite, da morajo ocene po metodi največjega verjetja zadoščati enačbi

$$\frac{1}{n}\frac{\partial \ell}{\partial \lambda} = \frac{a\,\tau}{\lambda} - \tau\, \lambda^{\tau-1}\, \overline{x^\tau} = 0\,,$$

torej

$$\lambda = \left(\frac{\overline{x^\tau}}{a}\right)^{-\frac{1}{\tau}}$$

in

$$\log \lambda = -\frac{1}{\tau}\, \log \overline{x^\tau} + \frac{1}{\tau}\, \log a\,.$$

b. Pokažite, da velja

$$\frac{1}{n}\frac{\partial \ell}{\partial a} = \tau\,\log\lambda - \psi(a) + \tau\,\overline{\log x} = 0$$

in

$$\psi(a) - \log a - \tau\,\overline{\log x} + \log\overline{x^\tau} = 0.$$

c. Iz parcialnega odvoda po τ sklepajte, da velja

$$a = \frac{\overline{x^\tau}}{\tau\,(\overline{x^\tau\,\log x} - \overline{x^\tau}\;\overline{\log x}}).$$

d. Desno stran enačbe v c. označite z $g(\tau)$. Pokažite, da velja

$$\psi(g(\tau)) - \log(g(\tau)) - \tau\,\overline{\log x} + \log\overline{x^\tau} = 0\,,$$

kar je enačba, v kateri nastopa le neznani parameter τ. Generirajte $n = 1000$ vzorčnih vrednosti in narišite graf desne strani zgornje enačbe. Lahko kaj sklepate o enoličnosti rešitev?

e. Izračunajte Fisherjevo matriko informacije $I(a, \lambda, \tau)$.

f. Generirajte vzorce iz posplošene gama porazdelitve velikosti $n = 1000$ in ocenite parametre. Ponovite postopek $m = 10000$-krat. Narišite histograme vzorčnih ocen in izračunajte njihove empirične standardne napake. Primerjajte te standardne napake z tistimi, ki jih dobite po teoriji za velike vzorce.

g. Kako bi preizkusili domnevo $H_0\colon \tau = 1$ proti $H_1\colon \tau \neq 1$? S simulacijo generirajte porazdelitev testne statistike, če H_0 drži. Za velikost vzorca si izberite $n = 1000$. Komentar?