Inferenza statistica – Esonero 1, Parte B2 – 6 aprile 2024

Cognome, nome e n. di matricola: ______________________________

- Si corregge solo quanto riportato in questi fogli
- Gli studenti che hanno diritto alla riduzione del compito possono saltare i quesiti 5 e 7

Problema.
Sia $\mathbf{X}_n = (X_1, \dots X_n)$ un campione i.i.d. da una popolazione X con funzione di densità

$$f_X(x;\theta) = \frac{2^{-\frac{1}{2}}}{\Gamma(0.5)} \frac{1}{x} \exp\left\{-\frac{1}{2}(\log x - \log\theta)^2\right\}, \qquad x \in \mathbb{R}, \quad \theta > 0$$

1. Verificare se il modello per la singola osservazione X_i appartiene alla classe delle famiglie esponenziali.

   Risp.

   $f_X(x;\theta) = c \times \frac{1}{x} \exp\theta\left\{-\frac{1}{2}(\ln x)^2 - \frac{1}{2}(\ln\theta)^2 + \ln x \ln\right\}$ da cui si ottiene che $h(x) = c \times \frac{1}{x} \times \exp\left\{-\frac{1}{2}(\ln x)^2\right\}$, $T(x) = \ln x$, $\eta(\theta) = \ln\theta$, $B(\theta) = -\frac{1}{2}(\ln\theta)^2$. Il modello è quindi una famiglia esponenziale.

2. Determinare il modello statistico per l'n-pla campionaria $\mathbf{X}_n = (X_1, \dots X_n)$.

   Risp.

   Il modello è dato da $\left\{\mathbb{R}^n, c^n \times \frac{1}{\prod x_i} \times \exp\left\{-\frac{1}{2}\sum_{i=1}^n(\ln x_i - \ln\theta)^2\right\}, \Theta = \mathbb{R}^+\right\}$.

3. Individuare una statistica sufficiente e minimale per il parametro θ fornendo l'opportuna argomentazione.

   Risp.

   $T_n(\mathbf{x}_n) = \sum_{i=1}^n T(x_i) = \sum_{i=1}^n \ln x_i$ è statistica sufficiente minimale per il criterio di fattorizzazione.

4. Determinare la stima di massima verosimiglianza $\widehat{\theta}_{mv}(\mathbf{x}_n)$ in corrispondenza del seguente campione osservato $(x_1 = 4.2, x_2 = 5.9, x_3 = 35.1)$.

   Risp.

   $L(\theta) = c \times \exp\left\{-\frac{1}{2}(\ln\theta)^2 + \ln\theta T_n\right\}$. Quindi $\ell(\theta) = c - \frac{n}{2}(\ln\theta)^2 + \ln\theta T_n$. Pertanto $\ell'(\theta) = -n\ln\theta\frac{1}{\theta} + \frac{1}{\theta}T_n = 0 \iff \ln\theta = \frac{1}{n}T_n$ da cui si ottiene $\widehat{\theta}_{mv} = e^{\frac{1}{n}\sum \ln x_i} = \left[\prod_{i=1}^n x_i\right]^{1/n}$. Sostituendo i valori numerici si ottiene $\widehat{\theta}_{mv} = 9.5$.

5. Verificare che l'informazione osservata di Fisher corrisponde alla seguente espressione:

   $$I_{oss} = \frac{n}{\widehat{\theta}_{mv}(\mathbf{x}_n)^2}$$

   Risp.

   $\frac{d^2}{d\theta^2}\ell(\theta) = -\frac{1}{\theta^2}T_n - \frac{n}{\theta^2} + \frac{n\ln\theta}{\theta^2}$. Si ha quindi che $\left.\frac{d^2}{d\theta^2}\ell(\theta)\right|_{\theta=\widehat{\theta}_{mv}} = \frac{1}{\widehat{\theta}^2_{mv}}[T_n - n\frac{T_n}{n} + n] = \frac{n}{\widehat{\theta}^2_{mv}}$.

6. Determinare l'espressione dell'insieme approssimato di verosimiglianza di livello q per θ (basato su approssimazione normale di $\overline{L}(\theta; \mathbf{x}_n)$).

   Risp.

   $\tilde{L}(\theta) = \exp\{-\frac{1}{2}(\theta - \widehat{\theta}_{mv})^2 I_n^{oss}\}$ e $\tilde{L}_q = \widehat{\theta}_{mv} \pm k_q[I_n^{oss}]^{-1/2}$ con $k_q = \sqrt{-2\ln q}$, e con $\widehat{\theta}_{mv}$ e I_n^{oss} ottenuti nei punti che precedono.

7. Verificare se la seguente statistica

$$S_n = S(x_1, ..., x_n) = \prod_{i=1}^{n} x_i$$

è sufficiente e minimale per θ.

Risp.

Poichè si ha $S_n = \prod x_i = e^{T_n}$, ed essendo la funzione esponenziale strettamente monotona, allora S_n è funzione biunivoca di T_n e quindi anch'essa statistica sufficiente minimale.