ESERCITAZIONE 5

1) $f_X(x,\vartheta) = 2\vartheta x e^{-\vartheta x^2}$, $x \geq 0$, $\vartheta > 0$

Verifica se $\{f_X(x,\vartheta),\ \vartheta \in \Theta\}$ costituisce una famiglia esponenziale e determina la statistica sufficiente minimale associata al campione casuale $(X_1, \dots, X_n)$

$$f_X(x,\vartheta) = 2x\, e^{\ln\vartheta}\, e^{-\vartheta x^2} = \underbrace{2x}_{h(x)} \exp\{\underbrace{-\vartheta x^2}_{\eta(\vartheta)T(x)} - \underbrace{(-\ln\vartheta)}_{B(\theta)}\} \Rightarrow \text{è famiglia esponenziale}$$

$$L(\theta, \underline{x_n}) = f_n(\underline{x_n}, \vartheta) = \prod_{i=1}^{n} f_X(x_i, \vartheta) = \underbrace{\left(\prod_{i=1}^{n} 2x_i\right)}_{h(\underline{x_n})} \underbrace{\exp\{-\vartheta \sum x_i^2 - (-n\ln(\theta))\}}_{g(\vartheta, T(\underline{x_n}))}$$

$$= h(\underline{x_n})\, g(\vartheta, T(\underline{x_n}))$$

$\Rightarrow$ criterio di fattorizzazione di Fisher ci dice che $T(\underline{x_n}) = \sum x_i^2$ è statistica sufficiente minimale

2) $X_1, \dots, X_n$ iid $X_i \sim N(0,1)$ determina la distribuzione di probabilità di $Y = \sum_{i=1}^{n} X_i^2$ e calcolare $\mathbb{E}[V]$ e $\mathbb{V}[V]$ dove $V = 2Y - 3$.

Si osserva che $\forall\, i = 1 \dots n$ $X_i^2 \sim \chi_1^2$ ($\to$ Chi-quadro con 1 gdl.) (per teoria da campionamento pop. normale)

quindi $Y = \sum_{i=1}^{n} X_i^2 \sim \chi_n^2 = \text{Gamma}\left(\frac{n}{2}, \frac{1}{2}\right)$.

$$\mathbb{E}[V] = \mathbb{E}[2Y-3] = 2\mathbb{E}[Y] - 3 = 2\left(\frac{n}{2}\cdot 2\right) - 3 = 2n - 3$$

$$\mathbb{V}[V] = \mathbb{V}[2Y-3] = 4\mathbb{V}[Y] = 4\left(\frac{n}{2}\cdot 4\right) = 8n$$

3) $\underline{X_n} = (X_1 \dots X_n)$ iid

$f_X(x,\vartheta) = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}}$ $x \geq 0$, $\vartheta > 0$ $\mathbb{E}[X] = \theta\sqrt{\frac{\pi}{2}}$ e $\mathbb{V}[X] = \frac{4-\pi}{2}\theta^2$

i) Determina il modello statistico per l'n-pla campionaria $\underline{X_n} = (X_1 \dots X_n)$

$$\left(\mathcal{X}^n = [0,+\infty)^n,\ f_n(\underline{x_n}, \vartheta),\ (0,+\infty) = \Theta\right)$$

$$f_n(\underline{x_n}, \vartheta) = \prod_{i=1}^{n}\left(\frac{x_i}{\theta^2} e^{-x_i^2/2\theta^2}\right) = \left(\prod x_i\right) \theta^{-2n} e^{-\sum x_i^2/2\theta^2}$$

ii) Scrivere $L(\vartheta, \underline{x}_n)$ ed individuare il nucleo e la statistica sufficiente per il modello.

$$L(\vartheta, \underline{x}_n) = f_n(\underline{x}_n, \vartheta) = \underbrace{(\prod x_i)}_{h(\underline{x}_n)} \underbrace{\vartheta^{-2n} e^{-\sum x_i^2 / 2\vartheta^2}}_{g(\vartheta, T(\underline{x}_n))}$$

$\Rightarrow$ criterio di fattorizzazione di Fisher la statistica sufficiente è $T(\underline{x}_n) = \sum_{i=1}^{n} x_i^2$

iii) Determina $\hat{\vartheta}_{MV}(\underline{x}_n)$

Considero solo il nucleo $g(\vartheta, T(\underline{x}_n))$ e ne faccio il logaritmo

$$\ell(\vartheta, \underline{x}_n) \propto \log\left(\vartheta^{-2n} e^{-\sum x_i^2 / 2\vartheta^2}\right) = -2n \log\vartheta - \frac{\sum x_i^2}{2\vartheta^2}$$

$$\frac{d\ell(\vartheta)}{d\vartheta} = -\frac{2n}{\vartheta} + \cancel{2} \frac{\sum x_i^2}{\cancel{2}\vartheta^3} = 0 \Rightarrow \frac{2n}{\vartheta} = \frac{\sum x_i^2}{\vartheta^3} \Rightarrow \hat{\vartheta} = \sqrt{\frac{\sum x_i^2}{2n}}$$

$$\left.\frac{d^2\ell(\vartheta)}{d\vartheta^2}\right|_{\hat{\vartheta}} = \frac{2n}{\vartheta^2} - \left.\frac{3\sum x_i^2}{\vartheta^4}\right|_{\hat{\vartheta}} = \frac{2n}{\left(\frac{\sum x^2}{2n}\right)} - \frac{3\sum x_i^2}{\left(\frac{\sum x_i^2}{2n}\right)^2} = \frac{4n^2}{\sum x_i^2} - \frac{12n^2}{\sum x^2} < 0$$

Quindi $\hat{\vartheta}_{MV} = \sqrt{\frac{\sum x_i^2}{2n}}$

iv) Determina $I_n^{on}(\underline{x}_n)$ e $\tilde{L}_q$ (insieme di verosimiglianza approssimato di livello q).

$$I_n^{on}(\hat{\vartheta}_{MV}, \underline{x}_n) = -\left.\frac{d^2}{d\vartheta^2}\ell(\vartheta, \underline{x}_n)\right|_{\vartheta = \hat{\vartheta}_{MV}} = -\left(\frac{4n^2}{\sum x_i^2} - \frac{12n^2}{\sum x_i^2}\right) = \frac{8n^2}{\sum x_i^2}$$

Per calcolare $\tilde{L}_q$ applichiamo la formula che deriva dall'approssimazione normale della funzione di verosimiglianza

$$\tilde{L}_q = \left(\hat{\vartheta}_{MV} - k_q\sqrt{I_n^{on}(\underline{x}_n)^{-1}}, \ \hat{\vartheta}_{MV} + k_q\sqrt{I_n^{on}(\underline{x}_n)^{-1}}\right) \quad \text{con} \quad k_q = \sqrt{-2\ln q}$$

$$\tilde{L}_q = \left(\sqrt{\frac{\sum x_i^2}{2n}} \pm k_q\sqrt{\frac{\sum x_i^2}{8n^2}}\right)$$

v) Determina lo stimatore dei momenti $\hat{\vartheta}_M(\underline{x}_n)$ di ϑ e stabilire se è non distorto per ϑ.

$$\mathbb{E}(X) = \bar{X}_n \Rightarrow \vartheta\sqrt{\frac{\pi}{2}} = \bar{X}_n \Rightarrow \hat{\vartheta}_M = \sqrt{\frac{2}{\pi}}\bar{X}_n \quad \mathbb{E}(\hat{\vartheta}_M) = \mathbb{E}\left[\sqrt{\frac{2}{\pi}}\bar{X}_n\right]$$

$$= \sqrt{\frac{2}{\pi}}\mathbb{E}[\bar{X}_n] = \sqrt{\frac{2}{\pi}}\mathbb{E}[X] = \sqrt{\frac{2}{\pi}}\vartheta\sqrt{\frac{\pi}{2}} = \vartheta \quad \forall \vartheta \in \Theta$$ è quindi non distorto per ϑ.

vi) Determina $MSE(\hat{\theta}_M(\underline{X}_n))$ e stabilire se è uno stimatore consistente di θ.

$$MSE(\hat{\theta}_M(\underline{X}_n)) = V(\hat{\theta}_M) + (\mathbb{E}(\hat{\theta}_M) - \theta)^2 = V(\hat{\theta}_M)$$

poiché non distorto per θ, l'elemento di distorsione è pari a 0.

$$= V(\hat{\theta}_M) = V\left(\sqrt{\frac{2}{\pi}}\bar{X}_n\right) = \frac{2}{\pi}V(\bar{X}_n)$$

$$= \frac{2}{\pi}\frac{V(X)}{n} = \frac{2}{\pi n}\left(\frac{4-\pi}{2}\theta^2\right) = \frac{4-\pi}{\pi n}\theta^2 \xrightarrow{n\to\infty} 0$$ è quindi consistente in MSE.

vii) Fornire la distribuzione asintotica di $\hat{\theta}_M(\underline{X}_n) = \sqrt{\frac{2}{\pi}}\bar{X}_n$

TLC $$\frac{(\bar{X}_n - \mathbb{E}(X))\sqrt{n}}{\sqrt{V(X)}} \dot{\sim} N(0,1) \Rightarrow \bar{X}_n \dot{\sim} N\left(\mathbb{E}(X), \frac{V(X)}{n}\right)$$

$$\bar{X}_n \dot{\sim} N\left(\theta\sqrt{\frac{\pi}{2}}, \frac{4-\pi}{2n}\theta^2\right) \Rightarrow \sqrt{\frac{2}{\pi}}\bar{X}_n$$ è combinazione di $\bar{X}_n \dot{\sim} N(.,.)$

allora $\sqrt{\frac{2}{\pi}}\bar{X}_n \dot{\sim} N(.,.) \Rightarrow \mathbb{E}\left(\sqrt{\frac{2}{\pi}}\bar{X}_n\right) = \sqrt{\frac{2}{\pi}}\theta\sqrt{\frac{\pi}{2}} = \theta$

$$V\left(\sqrt{\frac{2}{\pi}}\bar{X}_n\right) = \frac{4-\pi}{\pi n}\theta^2$$

Quindi $\hat{\theta}_M(\underline{X}_n) = \sqrt{\frac{2}{\pi}}\bar{X}_n \dot{\sim} N\left(\theta, \frac{4-\pi}{\pi n}\theta^2\right)$

viii) Verifica se $\hat{\theta}_M(\underline{X}_n)$ è o meno lo stimatore UMVUE.

$$I_n(\theta) = -n\mathbb{E}_\theta\left[\frac{d^2}{d\theta^2}\ln f_X(x,\theta)\right] \Rightarrow$$

$$\ln f_X(x,\theta) = \log(x) - 2\log\theta - \frac{x^2}{2\theta^2}$$

$$\frac{d}{d\theta}\ln f_X(x,\theta) = -\frac{2}{\theta} + \frac{2x^2}{2\theta^3} = -\frac{2}{\theta} + \frac{x^2}{\theta^3}$$

$$\frac{d^2}{d\theta^2}\ln f_X(x,\theta) = \frac{2}{\theta^2} - \frac{3x^2}{\theta^4}$$

$$\mathbb{E}_\theta\left[\frac{d^2}{d\theta^2}\ln f_X(x,\theta)\right] = \frac{2}{\theta^2} - \frac{3\,\mathbb{E}[x^2]}{\theta^4}$$

$$\mathbb{E}[x^2] = V(X) + \mathbb{E}(X)^2 = \frac{4-\pi}{2}\theta^2 + \frac{\pi}{2}\theta^2 = \frac{4\theta^2 - \pi\theta^2 + \pi\theta^2}{2} = 2\theta^2$$

$$= \frac{2}{\theta^2} - \frac{3\times 2\theta^2}{\theta^4} = \frac{2}{\theta^2} - \frac{6}{\theta^2} = -\frac{4}{\theta^2}$$

$$I_n(\theta) = -n\mathbb{E}_\theta\left[\frac{d^2}{d\theta^2}\ln f_X(x,\theta)\right] = \frac{4n}{\theta^2} \rightarrow CR(\theta) = \frac{1}{I_n(\theta)} = \frac{\theta^2}{4n} \neq R(\theta, \hat{\theta}_M(\underline{X}_n)) = V(\hat{\theta}_M(\underline{X}_n)) = \frac{4-\pi}{\pi n}\theta^2$$

si verifica facilmente che il modello è f.exp e quindi è garantito il raggiungimento di $cr(\theta)$ da parte dello stimatore UMVUE di θ, poiché $\hat{\theta}_M$ è tale che $V(\hat{\theta}_M) \neq cr(\theta)$ allora non può essere UMVUE.

ix) Verifica che, per il modello in esame, il parametro θ rappresenta il parametro di scala

- Pongo $\theta = 1$ ottengo la funzione di densità standard $f(x) = x\, e^{-x^2/2}$
- Calcolo $g_1(x,\theta) = \frac{1}{\theta} f\left(\frac{x}{\theta}\right) = \frac{1}{\theta}\left(\frac{x}{\theta}\right) e^{-\left(\frac{x}{\theta}\right)^2/2} = \frac{x}{\theta^2} e^{-x^2/2\theta^2} = f_X(x,\theta)$

poiché ho ottenuto la densità iniziale allora θ è parametro di scala.

x) Determina $r(\theta) = \mathbb{E}(X_i^2) = V(X) - \mathbb{E}(X)^2 = 2\theta^2$

calcolato prima

xi) Determina lo stimatore UMVUE di $r(\theta)$

Considero la statistica sufficiente individuata tramite CFF, ovvero $T(\underline{X}_n) = \sum_{i=1}^{n} X_i^2$ poiché il modello è famiglia esponenziale allora $T(\underline{X}_n)$ è anche completa. Quindi per ottenere UMVUE dobbiamo costruire uno stimatore $d = f(T(\underline{X}_n))$ e tale che $\mathbb{E}(d) = r(\theta)$

$$\Rightarrow \mathbb{E}\left[\sum_{i=1}^{n} X_i^2\right] = \sum_{i=1}^{n} \mathbb{E}(X_i^2) = \sum_{i=1}^{n} 2\theta^2 = 2n\theta^2 \Rightarrow \frac{1}{n}\mathbb{E}\left[\sum X_i^2\right] = 2\theta^2$$

$$\Rightarrow \mathbb{E}\left[\frac{\sum X_i^2}{n}\right] = 2\theta^2$$ quindi poiché $\mathbb{E}\left[\frac{\sum X_i^2}{n}\right] = 2\theta^2 = r(\theta)$

e $\frac{\sum X_i^2}{n}$ è $f(T(\underline{X}_n))$ allora $d = \frac{\sum X_i^2}{n} = d_{UMVUE}$