ATTENZIONE: SCRIVERE LE RISPOSTE SOLO SU QUESTI FOGLI.
EVENTUALI ALTRI FOGLI A QUADRETTI CONSEGNATI
NON VERRANNO CONSIDERATI

Cognome e Nome: ______________________ **Matricola e Corso Studi:** ______________

Esercizio A. Sia $\underline{X}_n = (X_1, X_2, \dots X_n)$ un campione casuale proveniente da una popolazione con funzione di densità di probabilità

$$f_X(x;\theta) = \frac{\theta^2}{x^3}\ \exp\left\{-\frac{\theta}{x}\right\}, \qquad x \geq 0, \quad \theta > 0$$

con $E[X] = \theta/2$ e $V[X] = \theta^2/4$.

A1. Verificare che l'insieme delle densità $\mathscr{F} = \{f_X(\cdot;\theta),\ \theta \in \Theta\}$ costituisce una famiglia esponenziale.

A2. Determinare la funzione di verosimiglianza di θ e una statistica sufficiente con il criterio di fattorizzazione.

A3. Determinare la stima di massima verosimiglianza del parametro θ.

A4. Determinare l'informazione attesa e l'informazione osservata di Fisher per il modello.

A5. Determinare l'approssimazione normale della funzione di verosimiglianza e l'insieme di verosimiglianza di livello q per un generico campione osservato $\underline{x}_n$.

A6. Determinare l'insieme di verosimiglianza di livello $q = 0.147$ che si ottiene con un campione osservato di 20 unità nel quale la somma dell'inverso delle osservazioni è uguale a 60.

A7. Determinare lo stimatore dei momenti di θ, che indichiamo con $\hat{\theta}_{mom}$, e studiarne la correttezza.

A8. Determinare l'errore errore quadratico medio di $\hat{\theta}_{mom}$, stabilire se si tratta di uno stimatore consistente e determinarne la distribuzione asintotica dello stimatore dei momenti.

A9. Determinare il limite inferiore di Cramer-Rao per gli stimatori non distorti di θ e calcolare l'efficienza dello stimatore $\hat{\theta}_{mom}$.

A10. Stabilire se lo stimatore $\hat{\theta}_{mom}$ è lo stimatore UMVUE (ottimo nella classe degli stimatori non distorti) e studiarne l'ammissibilità di $\hat{\theta}_{mom}$.

Soluzioni

A1. Per verificare che l'insieme delle densità $\mathscr{F} = \{f_X(\cdot;\theta), \theta \in \Theta\}$ costituisca una famiglia esponenziale è necessario individuare le funzioni $h(x), \eta(\theta), T(x), B(\theta)$ che permettano di riscrivere la generica funzione di densità f_X, appartenente al suddetto insieme, come segue

$$f_X(x;\theta) = h(x)\ \exp\{T(x)\ \eta(\theta) - B(\theta)\}.$$

Riscriviamo dunque la funzione di densità come

$$f_X(x;\theta) = \frac{\theta^2}{x^3}\exp\left\{-\frac{\theta}{x}\right\} = \frac{1}{x^3}\exp\left\{-\frac{\theta}{x} - (-2log(\theta))\right\},$$

da cui è possibile identificare le suddette quattro funzioni. Infatti,

$$\begin{aligned} h(x) &= \frac{1}{x^3} \\ \eta(\theta) &= -\theta \\ T(x) &= \frac{1}{x} \\ B(\theta) &= -2log(\theta) \end{aligned}$$

e si può concludere che $\mathscr{F} = \{f_X(\cdot;\theta), \theta \in \Theta\}$ costituisce una famiglia esponenziale.

A2. Determiniamo innanzitutto la funzione di verosimiglianza:

$$L(\theta;\underline{x}_n) = f_n(\underline{x}_n;\theta) \overset{\text{I.I.D.}}{=} \prod_{i=1}^{n} f_X(x_i;\theta) = \frac{\theta^{2n}}{\prod_{i=1}^{n} x_i^3}\exp\left\{-\theta\sum_{i=1}^{n}\frac{1}{x_i}\right\}.$$

Ricordando il criterio di fattorizzazione, secondo cui

$$L(\theta;\underline{x}_n) = h(\underline{x}_n)g(\theta, T(\underline{x}_n)) \propto g(\theta, T(\underline{x}_n)),$$

è possibile identificare

$$h(\underline{x}_n) = \frac{1}{\prod_{i=1}^n x_i^3}$$
$$g(\theta, T(\underline{x}_n)) = \theta^{2n} \ \exp\left\{-\theta \sum_{i=1}^n \frac{1}{x_i}\right\}.$$

Dunque, data la definizione di statistica sufficiente, essa è pari a

$$T(\underline{x}_n) = \sum_{i=1}^n \frac{1}{x_i}.$$

Data l'appartenenza del modello alla famiglia esponenziale, tale statistica è anche minimale e completa.

A3. Per la determinazione della stima di massima verosimiglianza del parametro θ (i.e. $\widehat{\theta}_{MV}$), si devono attuare i seguenti procedimenti:

$$l(\theta; \underline{x}_n) = \log L(\theta; \underline{x}_n) \propto \log g(\theta, T(\underline{x}_n)) = \log\left[\theta^{2n} \ \exp\left\{-\theta \sum_{i=1}^n \frac{1}{x_i}\right\}\right] = 2n \log(\theta) - \theta \sum_{i=1}^n \frac{1}{x_i}$$
$$l'(\theta; \underline{x}_n) = \frac{2n}{\theta} - \sum_{i=1}^n \frac{1}{x_i} = \frac{2n - \theta \sum_{i=1}^n \frac{1}{x_i}}{\theta}$$
$$l'(\theta; \underline{x}_n) = 0 \Leftrightarrow 2n - \theta \sum_{i=1}^n \frac{1}{x_i} = 0 \Rightarrow \theta^* = \frac{2n}{\sum_{i=1}^n \frac{1}{x_i}}$$
$$l''(\theta; \underline{x}_n) = -\frac{2n}{\theta^2} < 0 \qquad \forall \theta \in \Theta.$$

Dunque la stima di massima verosimiglianza del parametro θ è $\widehat{\theta}_{MV} = \frac{2n}{\sum_{i=1}^n \frac{1}{x_i}}$.

A4. Ricordiamo innanzitutto le definizioni di informazione attesa ed osservata di Fisher, rispettivamente:

$$\mathrm{I}_n(\theta) = E_\theta[I_n(\theta; \underline{X}_n)] = E_\theta[-l''(\theta; \underline{X}_n)]$$
$$I_n(\widehat{\theta}_{MV}; \underline{x}_n) = I_n(\theta; \underline{X}_n)|_{\theta=\widehat{\theta}_{MV}} = -\frac{d^2}{d\theta^2} l(\theta; \underline{x}_n)|_{\theta=\widehat{\theta}_{MV}} = -l''(\widehat{\theta}_{MV}; \underline{x}_n).$$

È importante notare che in tal caso, nonostante le alternative definizioni dell'informazione attesa di Fisher, quella ivi riportata è la più immediata avendo già calcolato al punto precedente la derivata seconda della log-verosimiglianza. Dunque

$$\mathrm{I}_n(\theta) = E_\theta[-l''(\theta; \mathbf{X}_n)] = E_\theta\left[\frac{2n}{\theta^2}\right] = \frac{2n}{\theta^2}$$
$$I_n(\widehat{\theta}_{MV}; \underline{x}_n) = -l''(\widehat{\theta}_{MV}; \underline{x}_n) = \frac{2n}{4n^2}\left(\sum_{i=1}^n \frac{1}{x_i}\right)^2 = \frac{\left(\sum_{i=1}^n \frac{1}{x_i}\right)^2}{2n},$$

ove si ricorda che θ è un parametro incognito, ma <u>fissato</u>.

A5. L'approssimazione normale di una funzione di verosimiglianza ha la seguente forma

$$\widetilde{L}(\theta; \underline{x}_n) = \exp\left\{-\frac{1}{2}(\theta - \widehat{\theta}_{MV})^2 I_n\right\},$$

che nel caso richiesto risulta essere

$$\widetilde{L}(\theta; \underline{x}_n) = \exp\left\{-\frac{1}{2}(\theta - \widehat{\theta}_{MV})^2 I_n(\widehat{\theta}_{MV}; \underline{x}_n)\right\} = \exp\left\{-\frac{1}{2}\left(\theta - \frac{2n}{\sum_{i=1}^n \frac{1}{x_i}}\right)^2 \frac{\left(\sum_{i=1}^n \frac{1}{x_i}\right)^2}{2n}\right\}.$$

Sfruttando l'approssimazione appena calcolata è possibile ottenere l'insieme di verosimiglianza *approssimato* di livello q, per un generico campione osservato, come segue

$$\widetilde{L}_q = \left[\widehat{\theta}_{MV} - k_q \sqrt{[I_n(\widehat{\theta}_{MV}; \underline{x}_n)]^{-1}}, \widehat{\theta}_{MV} + k_q \sqrt{[I_n(\widehat{\theta}_{MV}; \underline{x}_n)]^{-1}}\right] \qquad k_q = \sqrt{-2ln(q)}.$$

che, nello specifico, risulta essere

$$\widetilde{L}_q = \left[\frac{2n}{\sum_{i=1}^{n}\frac{1}{x_i}} - k_q\sqrt{\frac{2n}{\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)^2}}, \frac{2n}{\sum_{i=1}^{n}\frac{1}{x_i}} + k_q\sqrt{\frac{2n}{\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)^2}}\right] \qquad k_q = \sqrt{-2ln(q)}$$

A6. Dato $q = 0.147, n = 20, \sum_{i=1}^{n}\frac{1}{x_i} = 60$, l'insieme di verosimiglianza è pari a

$$\widetilde{L}_q = \left[\frac{2}{3} - 1.9582\sqrt{\frac{40}{3600}}, \frac{2}{3} + 1.9582\sqrt{\frac{40}{3600}}\right] = [0.4603, 0.8731].$$

A7. La famiglia di densità considerata è parametrizzata da un parametro θ unidimensionale; di conseguenza, il sistema dei momenti si riduce alla sola equazione

$$E_\theta[X] = \overline{X}_n \Rightarrow \frac{\theta}{2} = \overline{X}_n \Rightarrow \widehat{\theta}_{mom}(\underline{X}_n) = 2\overline{X}_n.$$

Calcoliamone il valore atteso:

$$E_\theta[\widehat{\theta}_{mom}(\underline{X}_n)] = E_\theta[2\overline{X}_n] \overset{\text{Linearità}}{=} 2E_\theta[\overline{X}_n] = 2E_\theta[X] = 2\frac{\theta}{2} = \theta \qquad \forall\theta \in \Theta.$$

Dunque lo stimatore dei momenti risulta corretto per il parametro di interesse θ.

A8. Dato che lo stimatore dei momenti è non distorto per il parametro θ, come dimostrato al punto precedente, il suo errore quadratico medio si riduce a

$$MSE_\theta[\widehat{\theta}_{mom}(\underline{X}_n)] = V_\theta[\widehat{\theta}_{mom}(\underline{X}_n)] = V_\theta[2\overline{X}_n] = 4V_\theta[\overline{X}_n] = 4\frac{V_\theta[X]}{n} = 4\frac{\theta^2}{4n} = \frac{\theta^2}{n}.$$

Lo stimatore dei momenti risulta dunque consistente per θ in quanto

$$MSE_\theta[\widehat{\theta}_{mom}(\underline{X}_n)] = \frac{\theta^2}{n} \overset{n\to+\infty}{\to} 0 \qquad \forall\theta \in \Theta.$$

Ricordando la distribuzione asintotica della media campionaria, i.e. $\overline{X}_n$,

$$\overline{X}_n \dot{\sim} N\left(E_\theta[X], \frac{V_\theta[X]}{n}\right)$$

ed essendo lo stimatore dei momenti per θ una trasformazione lineare di $\overline{X}_n$, possiamo facilmente derivare la sua distribuzione asintotica come segue

$$\overline{X}_n \dot{\sim} N\left(\frac{\theta}{2}, \frac{\theta^2}{4n}\right) \Rightarrow \widehat{\theta}_{mom}(\underline{X}_n) = 2\overline{X}_n \dot{\sim} N\left(\theta, \frac{\theta^2}{n}\right).$$

A9. Verificate le condizioni necessarie per la definizione del limite inferiore di Cramer-Rao per gli stimatori non distorti di θ - essendo il modello regolare -, esso può essere calcolato come segue

$$\text{LICR}(\theta) = \frac{1}{\text{I}_n(\theta)} \Rightarrow \text{LICR}(\theta) = \frac{\theta^2}{2n}.$$

Dato lo stimatore dei momenti $\widehat{\theta}_{mom}(\underline{X}_n)$ ed essendo esso non distorto per θ, è possibile calcolarne l'efficienza

$$\text{eff}[\widehat{\theta}_{mom}(\underline{X}_n), \theta] = \frac{\text{LICR}(\theta)}{V_\theta[\widehat{\theta}_{mom}(\underline{X}_n)]} = \frac{\frac{\theta^2}{2n}}{\frac{\theta^2}{n}} = \frac{1}{2}.$$

Essendo l'efficienza del suddetto stimatore inferiore ad 1, se ne deduce che la sua varianza è superiore al limite inferiore di Cramer-Rao - che si ricorda essere un limite per la varianza degli stimatori *non* distorti per il parametro di interesse (si veda il punto successivo).

A10. Innanzitutto ricordiamo che nel caso di modelli uniparametrici (i.e. $\theta \subseteq \mathbb{R}$) con distribuzione di probabilità della singola v.a. X che costituisce una famiglia esponenziale, esiste sicuramente uno stimatore nella classe degli stimatori non distorti che abbia varianza pari al limite inferiore di Cramer-Rao[1]. Rientrando tale modello nel suddetto caso possiamo concludere che lo stimatore dei momenti, seppur non distorto, avendo varianza superiore al limite inferiore di Cramer-Rao (i.e. $\text{LICR}(\theta) < V_\theta[\widehat{\theta}_{mom}(\underline{X}_n)]\ \forall n, \forall\theta \in \Theta$) *non* può essere lo stimatore UMVUE. Poiché quest'ultimo stimatore esiste per i modelli che costituiscono una famiglia esponenziale, se ne deduce che lo stimatore dei momenti è inammissibile.

[1]Si veda l'osservazione **4.45 b** a pag. 120 delle dispense di teoria.

Esercizio B. Sia $\underline{X}_n = (X_1, ..., X_n)$ un campione casuale semplice proveniente da una popolazione con distribuzione uniforme nell'intervallo $[-\theta, \theta]$ con θ parametro reale positivo incognito.

B1. Scrivere il modello statistico probabilistico parametrico <u>per la singola osservazione</u> X_i e verificare che il parametro θ è un parametro di scala per il modello.

B2. Calcolare i primi due momenti della distribuzione di X_i.

B3. Verificare che il metodo dei momenti non conduce ad un'equazione utile a determinare uno stimatore di θ se si utilizza il primo momento, ma consente di ottenere uno stimatore utilizzando il secondo momento.

B4. In base al punto precedente, determinare lo stimatore dei momenti di θ, denotandolo con il simbolo $\hat{\theta}_{mom}$.

B5. Mostrare che lo stimatore $\hat{\psi} = \frac{3}{n}\sum_{i=1}^{n} x_i^2$ è uno stimatore corretto per il parametro $\psi = \theta^2$

B6. Verificare che la distribuzione di $Q(X_i, \theta) = \frac{X_i^2}{\theta^2}$ ha distribuzione $Beta(0.5, 1)$ e pertanto ha valore atteso e varianza rispettivamente uguali a $\frac{1}{3}$ e $\frac{4}{45}$.

B7. Studiare la proprietà di consistenza dello stimatore $\hat{\psi}$ per il parametro $\psi = \theta^2$.

B8. Scrivere il modello statistico probabilistico parametrico per l'intera n-upla di osservazioni.

B9. Determinare lo stimatore di massima verosimiglianza per il parametro originario θ.

B10. Fornire un'opportuna argomentazione per verificare che lo stimatore $\hat{\theta}_{mom}$ non è corretto per θ.

Soluzioni

B1. Il modello statistico probabilistico parametrico per la singola osservazione X_i è

$$\left\{\mathscr{X} = [-\theta, \theta], f_X(x;\theta) = \frac{1}{2\theta} I_{[-\theta,\theta]}(x), \Theta = (0, +\infty) = \mathbb{R}^+\right\}.$$

La distribuzione uniforme nell'intervallo $[-\theta, \theta]$ può essere riscritta come segue

$$f_X(x;\theta) = \frac{1}{2\theta} I_{[-\theta,\theta]}(x) = \frac{1}{\theta}\frac{1}{2} I_{[-1,1]}\left(\frac{x}{\theta}\right) = \frac{1}{\theta} g_X\left(\frac{x}{\theta}\right)$$

ove $g_X(x) = \dfrac{1}{2} I_{[-1,1]}(x)$ è la distribuzione uniforme nell'intervallo $[-1, 1]$ e $g_X(\frac{x}{\theta}) = \dfrac{1}{2} I_{[-\theta,\theta]}(x)$. Dunque il parametro θ è un *parametro di scala* per il modello ivi considerato.

B2.

$$E[X_i] = \frac{\theta - \theta}{2} = 0$$
$$E[X_i^2] = V[X_i] + (E[X_i])^2 = \frac{(\theta - (-\theta))^2}{12} = \frac{4\theta^2}{12} = \frac{\theta^2}{3}.$$

B3. Il parametro θ del modello statistico probabilistico parametrico definito al punto B1 è un parametro unidimensionale, dunque il sistema dei momenti è composto in tal caso da una sola equazione. Ricordando che il metodo dei momenti impone l'uguaglianza tra il momento teorico (in questo caso momento primo teorico) ed il momento empirico (in questo caso momento primo empirico) si ottiene

$$E_\theta[X] = \overline{X}_n \Rightarrow 0 = \overline{X}_n$$

da cui non è possibile derivare lo stimatore dei momenti per il parametro θ. Possiamo dunque impostare l'equazione dei momenti mediante il momento secondo:

$$E_\theta[X^2] = \hat{\sigma}_n^2 + \overline{X}_n^2 \Rightarrow \frac{\theta^2}{3} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2 + \overline{X}_n^2 \Rightarrow \widehat{\theta}^* = \sqrt{\frac{3}{n}\sum_{i=1}^{n} X_i^2}.$$

B4. In base a quanto determinato al punto precedente, lo stimatore dei momenti è

$$\widehat{\theta}_{mom} = \sqrt{\frac{3}{n}\sum_{i=1}^{n} X_i^2}.$$

B5. Per mostrare che lo stimatore $\widehat{\psi} = \frac{3}{n}\sum_{i=1}^{n} x_i^2$ è uno stimatore corretto per il parametro $\psi = \theta^2$ è necessario verificare che il suo valore atteso sia pari al parametro di interesse.

$$E_\psi[\widehat{\psi}] = E_\psi\left[\frac{3}{n}\sum_{i=1}^{n} X_i^2\right] \overset{\text{Linearità}}{=} \frac{3}{n}\sum_{i=1}^{n} E_\theta[X_i^2] = \frac{3}{n}\sum_{i=1}^{n}\frac{\theta^2}{3} = \frac{3}{n}\frac{n\theta^2}{3} = \theta^2 = \psi \qquad \forall\theta\in\Theta.$$

B6. Sia $\mathscr{X} = [-1,1]$ il supporto di una distribuzione uniforme nell'intervallo $[-1,1]$. Data la trasformazione quadratica $y = g(x) = x^2$, il supporto delle y è $\mathscr{Y} = [0,1]$ (facile da verificare). Se $0 \le y \le 1$, i valori di $x \in \mathscr{X}$ che soddisfano la condizione $x^2 \le y$ sono

$$D_X(y) = [-\sqrt{y},\ \sqrt{y}].$$

Allora

$$F_Y(y) = Pr(X \in D_X(y)) = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,dx = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Sapendo che per una distribuzione uniforme nell'intervallo $[-1,1]$

$$f_X(x) = \frac{1}{2} \qquad F_X(x) = \frac{x+1}{2},$$

si ha che

$$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) = \frac{\sqrt{y}+1}{2} - \frac{-\sqrt{y}+1}{2} = \frac{2\sqrt{y}}{2} = \sqrt{y}.$$

Per ottenere $f_Y(y)$ basta derivare $F_Y(y)$:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{2\sqrt{y}}.$$

Applichiamo ora quanto dimostrato al nostro caso. Avendo una distribuzione uniforme nell'intervallo $[-\theta,\theta]$, il supporto delle x è $\mathscr{X} = [-\theta,\theta]$. Data la trasformazione quadratica $y = g(x) = \frac{x^2}{\theta^2}$, con *theta* costante reale positiva, $\mathscr{Y} = [0,1]$ (facile da verificare). Per $0 \le y \le 1$, i valori di $x \in \mathscr{X}$ che soddisfano la condizione $\frac{x^2}{\theta^2} \le y \iff x^2 \le y\,\theta^2$ sono

$$D_X(y) = [-\theta\sqrt{y}, \theta\sqrt{y}].$$

Allora

$$F_Y(y) = Pr(X \in D_X(y)) = \int_{-\theta\sqrt{y}}^{\theta\sqrt{y}} f_X(x)\,dx = F_X(\theta\sqrt{y}) - F_X(-\theta\sqrt{y}).$$

Sapendo che per una distribuzione uniforme nell'intervallo $[-\theta,\theta]$

$$f_X(x) = \frac{1}{2\theta} \qquad F_X(x) = \frac{x+1}{2\theta},$$

si ha che

$$F_Y(y) = F_X(\theta\sqrt{y}) - F_X(-\theta\sqrt{y}) = \frac{\theta\sqrt{y}+1}{2\theta} - \frac{-\theta\sqrt{y}+1}{2\theta} = \frac{2\theta\sqrt{y}}{2\theta} = \sqrt{y}.$$

Per ottenere $f_Y(y)$ basta derivare $F_Y(y)$:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{2\sqrt{y}}.$$

Ricordando che la distribuzione Beta di parametro $\underline{\theta} = (\alpha,\beta)$ ha supporto nell'intervallo $[0,1]$ e la sua funzione di densità è

$$f_X(x;\underline{\theta}) = \frac{1}{Beta(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1},$$

è necessario calcolare il coefficiente $Beta(\alpha,\beta)$ e verificare che esso sia pari ad 2 per completare la dimostrazione.

$$B(0.5,1) = \int_0^1 y^{-\frac{1}{2}}dy = 2\left[y^{\frac{1}{2}}\right]_0^1 = 2.$$

Allora, il supporto della variabile $Q(X_i,\theta) = \frac{X^2}{\theta^2}$ è $[0,1]$ (i.e. delle y) e la sua distribuzione è

$$f_Q(x_i;\alpha=0.5,\beta=1) = \frac{1}{2}x_i^{0.5-1}(1-x)^{1-1} = \frac{1}{2\sqrt{x}},$$

che coincide con una distribuzione Beta di parametri $\alpha = 0.5, \beta = 1$. Sfruttando le conoscenze sulla variabile aleatoria Beta,

$$E[Q(X_i,\theta)] = \frac{\alpha}{\alpha+\beta} = \frac{\frac{1}{2}}{\frac{1}{2}+1} = \frac{1}{3}$$

$$V[Q(X_i,\theta)] = \frac{\alpha\,\beta}{(\alpha+\beta)^2\,(\alpha+\beta+1)} = \frac{\frac{1}{2}\cdot 1}{(\frac{1}{2}+1)^2(\frac{1}{2}+1+1)} = \frac{4}{45}.$$

B7. Per studiare la proprietà di consistenza dello stimatore $\widehat{\psi}$ per il parametro $\psi = \theta^2$ è necessario innanzitutto calcolare il suo errore quadratico medio che, data la non distorsione dello stimatore verificata al punto B5, si decompone nella sola varianza dello stimatore stesso. Calcoliamo innanzitutto

$$V_\theta[X_i^2] = E_\theta[X_i^4] - (E_\theta[X_i^2])^2,$$

ove il momento quarto della X_i può essere facilmente ottenuto svolgendo il seguente integrale

$$E_\theta[X_i^4] = \int_{-\theta}^{\theta} x^4 \frac{1}{2\theta}\, dx = \frac{1}{2\theta}\left[\frac{x^5}{5}\right]_{-\theta}^{\theta} = \frac{1}{2\theta}\frac{2\theta^5}{5} = \frac{\theta^4}{5}.$$

Dunque

$$V_\theta[X_i^2] = \frac{\theta^4}{5} - \frac{\theta^4}{9} = \frac{4}{45}\theta^4.$$

Questo stesso risultato poteva essere ottenuto anche ragionando su quanto dimostrato al punto B6. Calcoliamo ora l'errore quadratico medio dello stimatore $\widehat{\psi}$.

$$MSE_\psi[\widehat{\psi}] = V_\psi[\widehat{\psi}] = V_\psi\left[\frac{3}{n}\sum_{i=1}^{n} X_i^2\right] \overset{\text{Indip.}X_i}{=} \frac{9}{n^2}\sum_{i=1}^{n} V_\psi[X_i^2] = \frac{9}{n^2}\sum_{i=1}^{n}\frac{4}{45}\theta^4 = \frac{9}{n^2}\frac{4\,n\,\theta^4}{45} = \frac{4}{5n}\theta^4.$$

Per verificare la consistenza dello stimatore $\widehat{\psi}$ è necessario analizzare il comportamento al limite del suo errore quadratico medio, i.e.

$$\lim_{n\to+\infty} MSE_\psi[\widehat{\psi}] = \lim_{n\to+\infty}\frac{4}{5n}\theta^4 = \lim_{n\to+\infty}\frac{4}{5n}\psi^2 = 0 \qquad \forall\theta\in\Theta.$$

Dunque, lo stimatore $\widehat{\psi}$ è consistente per il parametro $\psi = \theta^2$.

B8. Il modello statistico probabilistico parametrico per l'intera n-upla di osservazioni è

$$\left\{\mathscr{X}^n = [-\theta,\theta]^n,\, f_{X_1,\dots,X_n}(\underline{x}_n;\theta) \overset{\text{I.I.D.}}{=} \frac{1}{2^n\theta^n}\prod_{i=1}^{n} I_{[-\theta,\theta]}(x_i),\, \Theta = (0,+\infty) = \mathbb{R}^+\right\}.$$

B9. Scriviamo innanzitutto la funzione di verosimiglianza. Per poterla ricavare correttamente è necessario riscrivere l'indicatrice in funzione di θ e non di x_i.

$$\begin{cases} x_i \geq -\theta \\ x_i \leq \theta \end{cases} \Rightarrow \begin{cases} \theta \geq -x_i \\ \theta \geq x_i \end{cases}.$$

Estendendo questo risultato alla produttoria della funzione indicatrice otteniamo $I_{[\max(-x_{(1)},\ x_{(n)}),+\infty)}(\theta)$. La funzione di verosimiglianza si ottiene dunque come segue

$$L(\theta;\underline{x}_n) \overset{I.I.D.}{=} \frac{1}{2^n\theta^n}\, I_{[\max(-x_{(1)},\ x_{(n)}),+\infty)}(\theta) \propto \frac{1}{\theta^n}\, I_{[\max(-x_{(1)},\ x_{(n)}),+\infty)}(\theta).$$

Essendo il modello considerato non regolare (i.e. con supporto dipendente dal parametro θ), lo stimatore di massima verosimiglianza può essere ottenuto mediante analisi grafica della funzione di verosimiglianza. Notando che $\frac{1}{\theta^n}$ è una funzione decrescente di θ in $[\max(-x_{(1)},\ x_{(n)}),+\infty)$ (e vale zero altrimenti), il suo punto di massimo, i.e. lo stimatore di massima verosimiglianza per θ, è

$$\widehat{\theta}_{\text{MV}}(\underline{X}_n) = \max(-X_{(1)},\ X_{(n)}).$$

B10. Si noti innanzitutto che il valore atteso è un operatore lineare e che lo stimatore dei momenti $\widehat{\theta}_{mom}$ risulta essere una funzione <u>non</u> lineare di θ. Si ricordi la *disuguaglianza di Jensen* secondo cui, data una v.a. Y per cui $\exists\ E[Y]$ e presa una funzione convessa g per cui esista il valore atteso,

$$E[g(Y)] \geq g(E[Y]).$$

Posto $g(\cdot) = \sqrt{(\cdot)}$ (funzione convessa) e $Y = \frac{3}{n}\sum_{i=1}^{n} X_i^2$, allora

$$E_\theta[\widehat{\theta}_{mom}] = E_\theta[\sqrt{Y}] \geq \sqrt{E_\theta[Y]} = \sqrt{\frac{3}{n}\sum_{i=1}^{n} E_\theta[X_i^2]} = \sqrt{\frac{3}{n}n\frac{\theta^2}{3}} = \theta.$$

Dunque, poiché l'uguaglianza vale per qualche $\theta\in\Theta$ ma non per ogni θ nello spazio dei parametri (come richiesto dalla proprietà di correttezza), si può concludere che lo stimatore dei momenti non è coretto per il parametro di interesse.