Esercizi Unità 2

Analisi dei dati 2023/24

Cristiano Varin

1. Sia $X_1, \ldots, X_n$ un campione casuale semplice da una variabile casuale normale di media 1 e varianza $\sigma^2 > 0$.

   (a) Si ottenga lo stimatore di σ^2 con il metodo dei momenti.

   (b) Si ottenga lo stimatore di σ^2 con il metodo della massima verosimiglianza.

   Soluzione.

   (a) Non possiamo risolvere l'esercizio con il momento di ordine uno perché è costante, $\mathrm{E}(X) = 1$. Consideriamo, quindi, il momento di ordine due,

   $$\begin{aligned}\mathrm{E}(X^2) &= \mathrm{Var}(X) + \mathrm{E}(X)^2 \\ &= \sigma^2 + 1.\end{aligned}$$

   Perciò lo stimatore con il metodo dei momenti di σ^2 è

   $$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 1.$$

   (b) La funzione di densità di una variabile casuale normale di media uno e varianza σ^2 è

   $$f(x; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-1)^2}{2\sigma^2}\right\}.$$

   La funzione di log-verosimiglianza per σ^2 è

   $$\ell(\sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - 1)^2$$

   con una corrispondente funzione punteggio

   $$\ell'(\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i - 1)^2.$$

   Risolvendo l'equazione di verosimiglianza otteniamo il punto critico

   $$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - 1)^2,$$

che coincide con lo stimatore di massima verosimiglianza perché la derivata seconda della log-verosimiglianza è negativa quando calcolata in $\hat{\sigma}^2$. La derivata seconda della log-verosimiglianza è

$$\ell''(\sigma^2) = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n}(X_i - 1)^2,$$

quindi

$$\begin{aligned}
\ell''(\hat{\sigma}^2) &= \frac{n}{2\hat{\sigma}^4} - \frac{1}{\hat{\sigma}^6}\sum_{i=1}^{n}(X_i - 1)^2, \\
&= \frac{n}{2\hat{\sigma}^4} - \frac{1}{\hat{\sigma}^6}(n\hat{\sigma}^2) \\
&= -\frac{n}{2\hat{\sigma}^4} < 0.
\end{aligned}$$

2. Sia $X_1, \ldots, X_n$ un campione casuale semplice da una variabile casuale continua con funzione di densità

$$f(x;\alpha) = \begin{cases} \dfrac{\alpha}{x}\left(\dfrac{2}{x}\right)^{\alpha}, & \text{se } x > 2,\ \alpha > 1, \\ 0, & \text{altrimenti.} \end{cases}$$

   (a) Si ottenga lo stimatore di α con il metodo dei momenti.

   (b) Si ottenga lo stimatore di α con il metodo della massima verosimiglianza.

   (c) Si verifichi la consistenza dei due stimatori ottenuti ai punti precedenti.

   Soluzione.

   (a) Il valore atteso di X_1 è

$$\begin{aligned}
\mathrm{E}(X_1) &= \int_2^{\infty} x\, f(x;\alpha)\, \mathrm{d}x \\
&= \int_2^{\infty} x\, \frac{\alpha}{x}\left(\frac{2}{x}\right)^{\alpha} \mathrm{d}x \\
&= \alpha 2^{\alpha} \int_2^{\infty} x^{-\alpha}\, \mathrm{d}x \\
&= \alpha 2^{\alpha} \left.\frac{x^{-\alpha+1}}{-\alpha+1}\right|_2^{\infty} \\
&= \frac{2\alpha}{\alpha - 1}
\end{aligned}$$

   Lo stimatore con il metodo dei momenti si ottiene risolvendo l'equazione

$$\bar{X} = \frac{2\hat{\alpha}}{\hat{\alpha} - 1}.$$

   da cui si ottiene

$$\hat{\alpha} = \frac{\bar{X}}{\bar{X} - 2}.$$

(b) La log-verosimiglianza per α è

$$\ell(\alpha) = n \log \alpha + \alpha n \log 2 - \alpha \sum_{i=1}^{n} \log X_i,$$

con corrispondente la funzione punteggio

$$\ell'(\alpha) = \frac{n}{\alpha} + n \log 2 - \sum_{i=1}^{n} \log X_i.$$

Risolvendo l'equazione di verosimiglianza otteniamo il punto critico

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \log X_i - n \log 2}.$$

che corrisponde allo stimatore di massima verosimiglianza perché la derivata seconda della log-verosimiglianza è negativa per qualsiasi valore di α,

$$\ell''(\alpha) = -\frac{n}{\alpha^2} < 0.$$

(c) Lo stimatore con il metodo dei momenti è consistente perché la 'Legge dei grandi numeri' assicura che

$$\bar{X} \xrightarrow{p} \frac{2\alpha}{\alpha - 1}, \quad \text{per } n \to \infty,$$

quindi, prendendo la trasformazione continua $g(z) = z/(z-2)$, si ottiene

$$\hat{\alpha} = g(\bar{X}) \xrightarrow{p} g\left(\frac{2\alpha}{\alpha - 1}\right) = \alpha, \quad \text{per } n \to \infty,$$

ovvero $\hat{\alpha}$ è uno stimatore consistente di α. Lo stimatore di massima verosimiglianza di α è, invece, consistente perché ci troviamo nelle condizioni di regolarità sotto le quali gli stimatori di massima verosimiglianza sono consistenti.

3. Sia $X_1, \ldots, X_n$ un campione casuale semplice da una variabile casuale continua con funzione di densità

$$f(x;\theta) = \begin{cases} (\theta+1)x^{-(\theta+2)}, & \text{se } x > 1,\ \theta > 0 \\ 0, & \text{altrimenti.} \end{cases}$$

(a) Si ottenga lo stimatore di θ con il metodo dei momenti.

(b) Trovare lo stimatore di θ con il metodo della massima verosimiglianza.

(c) Si calcolino le stime di θ corrispondenti agli stimatori derivati ai punti precedenti con un campione di dimensione 100 in cui

$$\sum_{i=1}^{100} x_i = 158.39 \quad \text{e} \quad \sum_{i=1}^{100} \log x_i = 32.71.$$

Soluzione

(a) Il valore atteso X di è

$$\int_{-\infty}^{\infty} x\, f(x;\theta)\, \mathrm{d}x = \int_{1}^{\infty} (\theta+1) x^{-(\theta+1)}\, \mathrm{d}x = \frac{\theta+1}{\theta}.$$

Quindi, lo stimatore con il metodo dei momenti di θ è

$$\hat{\theta} = \frac{1}{\bar{X}-1}.$$

(b) La log-verosimiglianza per θ è

$$\ell(\theta) = n \log(\theta+1) - (\theta+2) \sum_{i=1}^{n} \log X_i,$$

con corrispondente funzione punteggio

$$\ell'(\theta) = \frac{n}{\theta+1} - \sum_{i=1}^{n} \log X_i.$$

Risolvendo l'equazione di verosimiglianza otteniamo il punto critico

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log X_i} - 1$$

che corrisponde allo stimatore di massima verosimiglianza perché la derivata seconda della log-verosimiglianza è negativa per qualsiasi valore di θ,

$$\ell''(\theta) = -\frac{n}{(\theta+1)^2} < 0.$$

(c) Con i dati disponibili otteniamo la stima di θ con il metodo dei momenti è $\hat{\theta} = 1/(158.39/100 - 1) = 1.71$, mentre la stima di θ con il metodo della massima verosimiglianza è $\hat{\theta} = 100/32.71 - 1 = 2.06$.

4. Sia $X_1, \dots, X_n$ un campione casuale semplice da una variabile casuale continua con funzione di densità

$$f(x;\theta) = \begin{cases} \frac{1}{\theta^2} \exp\left(-\frac{x}{\theta^2}\right), & x > 0,\ \theta > 0, \\ 0, & \text{altrimenti.} \end{cases}$$

(a) Si ottenga lo stimatore di θ con il metodo dei momenti.

(b) Si ottenga lo stimatore di θ con il metodo della massima verosimiglianza.

(c) Si calcolino le stime di θ ottenute ai punti precedenti con un campione di dimensione $n = 75$ in cui $\sum_{i=1}^{75} x_i = 123.2$.

Soluzione

(a) Il valore atteso di X è

$$\mathrm{E}(X) = \int_{0}^{\infty} x \frac{1}{\theta^2} \exp\left(-\frac{x}{\theta^2}\right) \mathrm{d}x.$$

Consideriamo il cambio di variabile $y = -x/\theta^2$ con cui l'integrale diventa

$$\begin{aligned}
\mathrm{E}(X) &= -\theta^2 \int_{-\infty}^{0} y e^y \mathrm{d}y \\
&= -\theta^2 \left(e^y y|_{-\infty}^{0} - \int_{-\infty}^{0} e^y \, \mathrm{d}y \right) \quad \text{(integrazione per parti)} \\
&= \theta^2.
\end{aligned}$$

Lo stimatore di θ con il metodo dei momenti si ottiene risolvendo l'equazione $\hat{\theta}^2 = \bar{X}$, ovvero $\hat{\theta} = \sqrt{\bar{X}}$.

(b) La log-verosimiglianza per θ è

$$\ell(\theta) = -2n \log \theta - \frac{1}{\theta^2} \sum_{i=1}^{n} X_i,$$

con corrispondente funzione punteggio,

$$\ell'(\theta) = -\frac{2n}{\theta} + \frac{2}{\theta^3} \sum_{i=1}^{n} X_i.$$

Risolvendo l'equazione di verosimiglianza otteniamo il punto critico $\hat{\theta} = \sqrt{\bar{X}}$ che corrisponde allo stimatore di massima verosimiglianza poiché la derivata seconda della log-verosimiglianza è negativa in $\hat{\theta}$, infatti

$$\ell''(\theta) = \frac{2n}{\theta^2} - \frac{6}{\theta^4} \sum_{i=1}^{n} X_i$$

da cui segue che

$$\begin{aligned}
\ell''(\hat{\theta}) &= \frac{2n}{\hat{\theta}^2} - \frac{6}{\hat{\theta}^4} \sum_{i=1}^{n} X_i \\
&= \frac{2n}{\bar{X}} - \frac{6}{\bar{X}^2} n \bar{X} \\
&= -\frac{4n}{\bar{X}} < 0.
\end{aligned}$$

In questo esercizio lo stimatore di massima verosimiglianza coincide con lo stimatore basato sul metodo dei momenti. Ricordate che in generale questi due stimatori sono diversi.

(c) La stima di massima verosimiglianza (o con il metodo dei momenti) di θ è $\hat{\theta} = \sqrt{123.2/75} = 1.28$.

5. Sia $X_1, \ldots, X_{10}$ un campione casuale semplice da una variabile casuale discreta con funzione di probabilità

$$\Pr(X = x; \theta) = \begin{cases} 2\theta/3, & \text{se } x = 0, \\ \theta/3, & \text{se } x = 1, \\ 2(1-\theta)/3, & \text{se } x = 2, \\ (1-\theta)/3, & \text{se } x = 3. \end{cases}$$

In un campione casuale di dimensione 10 sono stati osservati i seguenti valori:

$$x_1 = 3, x_2 = 0, x_3 = 2, x_4 = 1, x_5 = 3, x_6 = 2, x_7 = 1, x_8 = 0, x_9 = 2, x_{10} = 1$$

(a) Si calcoli la stima di θ con il metodo dei momenti.

(b) Si calcoli la stima di θ con il metodo di massima verosimiglianza.

Soluzione

(a) Il valore atteso di X è

$$E(X) = 0 \cdot \frac{2\theta}{3} + 1 \cdot \frac{\theta}{3} + 2 \cdot \frac{2(1-\theta)}{3} + 3 \cdot \frac{(1-\theta)}{3} = \frac{7-6\theta}{3}.$$

Lo stimatore di θ con il metodo dei momenti si ottiene risolvendo l'equazione

$$\frac{7-6\hat{\theta}}{3} = \bar{X}$$

ovvero

$$\hat{\theta} = \frac{7-3\bar{X}}{6}.$$

La media dei dati osservati è $\bar{x} = 15/10 = 1.5$ per cui la stima con il metodo dei momenti di θ è $\hat{\theta} = 5/12$.

(b) La verosimiglianza per θ con i dati osservati è

$$L(\theta) = \left(\frac{2\theta}{3}\right)^2 \left(\frac{\theta}{3}\right)^3 \left\{\frac{2(1-\theta)}{3}\right\}^3 \left\{\frac{(1-\theta)}{3}\right\}^2,$$

con corrispondente log-verosimiglianza

$$\ell(\theta) = 5\log(\theta) + 5\log(1-\theta),$$

e funzione punteggio

$$\ell'(\theta) = \frac{5}{\theta} - \frac{5}{1-\theta}.$$

La soluzione dell'equazione di verosimiglianza è $\hat{\theta} = 1/2$ e corrisponde alla stima di massima verosimiglianza di θ poiché la derivata seconda della log-verosimiglianza è negativa per qualsiasi valore di θ,

$$\ell''(\theta) = -\frac{5}{\theta^2} - \frac{5}{(1-\theta)^2} < 0.$$

6. Sia $X_1, \ldots, X_n$ un campione casuale semplice da una variabile casuale normale con media zero e varianza σ^2.

(a) Si ottenga lo stimatore di σ^2 con il metodo della massima verosimiglianza.

(b) Si ottengano le approssimazioni dell'errore standard dello stimatore ottenuto al punto precedente utilizzando l'informazione osservata e l'informazione attesa.

(c) Supponendo che un campione casuale di dimensione 60 abbia dato $\sum_{i=1}^{60} x_i^2 = 164.83$, si calcolino la stima di massima verosimiglianza di σ^2 e le stime degli errori standard approssimati ottenuti al punto precedente.

Soluzione

(a) La log-verosimiglianza per σ^2 è

$$\ell(\sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} X_i^2,$$

con corrispondente funzione punteggio

$$\ell'(\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n} X_i^2.$$

La soluzione dell'equazione di verosimiglianza è

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

e corrisponde allo stimatore di massima verosimiglianza poiché la derivata seconda della log-verosimiglianza è negativa in $\hat{\sigma}^2$, infatti

$$\ell''(\sigma^2) = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n} X_i^2$$

da cui segue

$$\begin{aligned}\ell''(\hat{\sigma}^2) &= \frac{n}{2\hat{\sigma}^4} - \frac{1}{\hat{\sigma}^6}\sum_{i=1}^{n} X_i^2\\ &= \frac{n}{2\hat{\sigma}^4} - \frac{1}{\hat{\sigma}^6}(n\hat{\sigma}^2)\\ &= -\frac{n}{2\hat{\sigma}^4} < 0.\end{aligned}$$

(b) L'informazione osservata è

$$J(\sigma^2) = -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{i=1}^{n} X_i^2.$$

L'informazione attesa è

$$I(\sigma^2) = -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{i=1}^{n} \mathrm{E}(X_i^2),$$

per cui dobbiamo calcolare $\mathrm{E}(X_i^2)$. Sappiamo che la variabile casuale X ha media $\mathrm{E}(X) = 0$ e varianza $\mathrm{var}(X) = \sigma^2$ e quindi otteniamo immediatamente che

$$\sigma^2 = \mathrm{var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \mathrm{E}(X^2).$$

L'informazione attesa è, quindi, pari a

$$\begin{aligned}I(\sigma^2) &= -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{i=1}^{n} \mathrm{E}(X_i^2)\\ &= -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{i=1}^{n} \sigma^2\\ &= -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}(n\sigma^2)\\ &= \frac{n}{2\sigma^4}.\end{aligned}$$

Al crescere della dimensione campionaria l'errore standard di $\hat{\sigma}^2$ è approssimativamente pari a

$$\begin{aligned}\mathrm{SE}(\hat{\sigma}^2) &\approx I(\sigma^2)^{-1/2} \\ &= \sigma^2\sqrt{\frac{2}{n}}\end{aligned}$$

oppure, usando l'informazione osservata, a

$$\begin{aligned}\mathrm{SE}(\hat{\sigma}^2) &\approx J(\sigma^2)^{-1/2} \\ &= \left(-\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{i=1}^{n} X_i^2\right)^{-1/2}\end{aligned}$$

Al crescere della dimensione campionaria l'errore standard di $\hat{\sigma}^2$ stimato è approssimativamente pari a

$$\begin{aligned}\widehat{\mathrm{SE}}(\hat{\sigma}^2) &\approx I(\hat{\sigma}^2)^{-1/2} \\ &= \hat{\sigma}^2\sqrt{\frac{2}{n}}.\end{aligned}$$

Lo stesso stimatore si ottiene con l'informazione osservata, infatti

$$\begin{aligned}\widehat{\mathrm{SE}}(\hat{\sigma}^2) &\approx J(\hat{\sigma}^2)^{-1/2} \\ &= \left(-\frac{n}{2\hat{\sigma}^4} + \frac{1}{\hat{\sigma}^6}\sum_{i=1}^{n} X_i^2\right)^{-1/2} \\ &= \left\{-\frac{n}{2\hat{\sigma}^4} + \frac{1}{\hat{\sigma}^6}(n\hat{\sigma}^2)\right\}^{-1/2} \\ &= \hat{\sigma}^2\sqrt{\frac{2}{n}}.\end{aligned}$$

Si ricordi che, però, in generale i due stimatori dell'errore standard possono differire.

(c) La stima di massima verosimiglianza con i dati osservati è $\hat{\sigma}^2 = 164.83/60 = 2.75$ con un errore standard stimato approssimativamente pari a $2.75\sqrt{2/60} = 0.50$.

7. Partendo dai risultati ottenuti nell'esercizio precedente:

   (a) Si ottenga lo stimatore di σ con il metodo della massima verosimiglianza.

   (b) Si ottenga un'approssimazione dell'errore standard dello stimatore ottenuto al punto precedente.

   (c) Supponendo che un campione casuale di dimensione 60 abbia dato $\sum_{i=1}^{60} x_i^2 = 164.83$, si calcoli la stima di massima verosimiglianza di σ e il corrispondente errori standard approssimati.

   Soluzione

(a) Per la proprietà dell'invarianza, lo stimatore di σ è semplicemente $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$ ovvero

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}.$$

(b) Usando i risultati sulle trasformazioni riportati a pagina 31 dell'Unità 1 abbiamo che la varianza di $g(\hat{\sigma}^2)$ al crescere della dimensione campionaria è approssimativamente pari a

$$\widehat{\mathrm{Var}}\{g(\hat{\sigma}^2)\} \approx \{g'(\hat{\sigma}^2)\}^2 \widehat{\mathrm{Var}}(\hat{\sigma}^2)$$

e quindi l'errore standard di $g(\hat{\sigma}^2)$ al crescere della dimensione campionaria è approssimativamente pari a

$$\widehat{\mathrm{SE}}\{g(\hat{\sigma}^2)\} \approx g'(\hat{\sigma}^2)\widehat{\mathrm{SE}}(\hat{\sigma}^2)$$

sempre che $g'(\hat{\sigma}^2)$ sia non nulla. Nel nostro caso abbiamo che $g(z) = \sqrt{z}$ con $g'(z) = 1/(2\sqrt{z})$ e quindi

$$\begin{aligned}
\widehat{\mathrm{SE}}(\hat{\sigma}) &\approx \frac{1}{2\sqrt{\hat{\sigma}^2}}\widehat{\mathrm{SE}}(\hat{\sigma}^2) \\
&= \frac{1}{2\sqrt{\hat{\sigma}^2}}\hat{\sigma}^2\sqrt{\frac{2}{n}} \\
&= \frac{\hat{\sigma}}{\sqrt{2n}}.
\end{aligned}$$

Il risultato vale a meno del caso (degenere) in cui $\hat{\sigma}^2 = 0$.

(c) La stima di massima verosimiglianza con i dati osservati è $\hat{\sigma} = \sqrt{164.83/60} = 1.66$ con un errore standard stimato approssimativamente pari a $1.66/\sqrt{2(60)} = 0.15$.

8. Sia $X_1, \ldots, X_n$ un campione casuale semplice da una distribuzione discreta con funzione di probabilità

$$Pr(X = x; \theta) = \begin{cases} \theta/2, & \text{se } x = -1, \\ (1-\theta), & \text{se } x = 0, \\ \theta/2, & \text{se } x = 1. \end{cases}$$

In un campione di dimensione $n = 261$ è stato osservato 109 volte il valore -1, 49 volte il valore 0 e 103 volte il valore 1.

(a) Si calcoli la stima di θ con il metodo dei momenti .

(b) si calcoli la stima di θ con il metodo della massima verosimiglianza.

(c) Si calcoli una stima dell'errore standard della stima di massima verosimiglianza di θ.

Soluzione

(a) Il valore atteso di X è nullo,

$$\mathrm{E}(X) = -1\left(\frac{\theta}{2}\right) + 0(1-\theta) + 1\left(\frac{\theta}{2}\right) = 0.$$

Consideriamo, quindi, il momento di ordine due,

$$\mathrm{E}(X^2) = (-1)^2\left(\frac{\theta}{2}\right) + 0(1-\theta) + (1)^2\left(\frac{\theta}{2}\right) = \theta.$$

Il corrispondente stimatore di θ con il metodo dei momenti è

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Con i dati osservati otteniamo la stima di θ

$$\hat{\theta} = \frac{(109 + 103)}{261} = 0.81.$$

(b) La funzione di log-verosimiglianza è, a meno di termini costanti, pari a

$$\ell(\theta) = (n_{-1} + n_1)\log\theta + n_0 \log(1-\theta),$$

dove n_{-1}, n_0 e n_1 sono il numero di osservazioni pari a -1, 0 e 1. Nel caso del campione osservato abbiamo

$$\ell(\theta) = (109 + 103)\log\theta + 49\log(1-\theta)$$

con una corrispondente funzione punteggio pari a

$$\ell'(\theta) = \frac{212}{\theta} - \frac{49}{1-\theta}.$$

Risolvendo l'equazione di verosimiglianza troviamo il punto critico

$$\hat{\theta} = \frac{212}{212 + 49} = 0.81,$$

che coincide con la stima di massima verosimiglianza poiché

$$\ell''(\theta) = -\frac{212}{\theta^2} - \frac{49}{(1-\theta)^2} < 0.$$

In questo caso la stima di massima verosimiglianza è uguale alla stima con il metodo dei momenti.

(c) L'informazione osservata calcolata in $\hat{\theta}$ è

$$\begin{aligned} J(\hat{\theta}) &= \frac{212}{\hat{\theta}^2} + \frac{49}{(1-\hat{\theta})^2} \\ &= \frac{212}{0.81^2} + \frac{49}{(1-0.81)^2} \\ &= 1680.46. \end{aligned}$$

Quindi, una stima dell'errore standard di $\hat{\theta}$ è $1/\sqrt{1680.46} = 0.024$.

9. Sia $X_1, \ldots, X_n$ un campione casuale da una variabile casuale continua con funzione di densità

$$f(x;\theta) = \frac{\pi}{2}\sin\{\pi(x-\theta)\}, \quad x \in [\theta, \theta+1].$$

(a) Si ottenga lo stimatore di θ con il metodo dei momenti.

(b) Si ottenga una stima dell'errore standard dello stimatore ottenuto al punto precedente.

Soluzione

(a) Il valore atteso di X si calcola con la regola dell'integrazione per parti ottenendo

$$\begin{aligned}
\mathrm{E}(X) &= \frac{\pi}{2}\int_\theta^{\theta+1} x\sin\{\pi(x-\theta)\}\ \mathrm{d}x \\
&= \frac{\pi}{2}\left[-(\theta+1)\frac{\cos\pi}{\pi}+\theta\frac{\cos 0}{\pi}-\int_\theta^{\theta+1}\cos\{\pi(x-\theta)\}\ \mathrm{d}x\right] \\
&= \frac{1}{2}\left(\theta+1+\theta-\frac{\sin\pi}{\pi}+\frac{\sin 0}{\pi}\right) \\
&= \theta+1/2.
\end{aligned}$$

Quindi, lo stimatore di θ con il metodo dei momenti è

$$\hat\theta = \bar X - 1/2.$$

(b) L'errore standard di $\hat\theta$ è

$$\mathrm{SE}(\hat\theta) = \mathrm{SD}(\bar X) = \sqrt{\frac{\mathrm{Var}(X_1)}{n}}.$$

La varianza di X_1 è $\mathrm{Var}(X_1) = \mathrm{E}(X_1^2) - \mathrm{E}(X_1)^2$, per cui dobbiamo calcolare il momento secondo usando due volte la regola dell'integrazione per parti. Il primo utilizzo della regola dell'integrazione per parti ci porta al risultato

$$\begin{aligned}
\mathrm{E}(X_1^2) &= \frac{\pi}{2}\int_\theta^{\theta+1} x^2\sin\{\pi(x-\theta)\}\ \mathrm{d}x \\
&= \frac{\pi}{2}\left[-(\theta+1)^2\frac{\cos\pi}{\pi}+\theta^2\frac{\cos 0}{\pi}+\frac{2}{\pi}\int_\theta^{\theta+1} x\cos\{\pi(x-\theta)\}\ \mathrm{d}x\right] \\
&= \frac{(\theta+1)^2}{2}+\frac{\theta^2}{2}+\int_\theta^{\theta+1} x\cos\{\pi(x-\theta)\}\ \mathrm{d}x.
\end{aligned}$$

Calcoliamo ora l'integrale che appare nell'ultima equazione utilizzando di nuovo l'integrazione per parti,

$$\begin{aligned}
\int_\theta^{\theta+1} x\cos\{\pi(x-\theta)\}\ \mathrm{d}x &= (\theta+1)\frac{\sin\pi}{\pi}-\theta\frac{\sin 0}{\pi}-\frac{1}{\pi}\int_\theta^{\theta+1}\sin\{\pi(x-\theta)\}\ \mathrm{d}x \\
&= -\frac{1}{\pi}\int_\theta^{\theta+1}\sin\{\pi(x-\theta)\}\ \mathrm{d}x \\
&= \frac{\cos\pi}{\pi^2}-\frac{\cos 0}{\pi^2} \\
&= -\frac{2}{\pi^2}.
\end{aligned}$$

Unendo il tutto abbiamo che

$$\begin{aligned}
\mathrm{E}(X^2) &= \frac{(\theta+1)^2}{2}+\frac{\theta^2}{2}-\frac{2}{\pi^2} \\
&= \frac{2\theta^2+2\theta+1}{2}-\frac{2}{\pi^2} \\
&= \theta^2+\theta+\frac{1}{2}-\frac{2}{\pi^2}
\end{aligned}$$

Quindi, la varianza di X_1 è

$$\begin{aligned}\mathrm{Var}(X_1) &= (\theta^2 + \theta + 1/2 - 2/\pi^2) - (\theta + 1/2)^2 \\ &= (\theta^2 + \theta + 1/2 - 2/\pi^2) - (\theta^2 + \theta + 1/4) \\ &= 1/4 - 2/\pi^2.\end{aligned}$$

L'errore standard di $\hat{\theta}$ è $\sqrt{(1/4 - 2/\pi^2)/n}$.

10. Sia $X_1, \ldots, X_n$ un campione casuale da una variabile casuale continua con funzione di densità

$$f(x; \theta) = \frac{\theta}{x^2}, \quad 0 < \theta \leq x.$$

(a) Si ottenga lo stimatore di θ con il metodo dei momenti.

(b) Si ottenga lo stimatore di θ con il metodo della massima verosimiglianza.

Soluzione

(a) Il valore atteso di X non esiste finito

$$\mathrm{E}(X) = \int_\theta^\infty x \frac{\theta}{x^2} \, \mathrm{d}x \to \infty,$$

così come tutti i momenti di ordine superiore $\mathrm{E}(X^r)$, $r > 1$. Consideriamo, quindi, i momenti di ordine inferiore ad uno. Il momento di ordine -1 è

$$\mathrm{E}(\mathrm{X}^{-1}) = \int_\theta^\infty \frac{\theta}{x^3} \, \mathrm{d}x = \frac{1}{2\theta}.$$

Quindi, lo stimatore di θ con il metodo dei momenti è

$$\hat{\theta} = \frac{n}{2 \sum_{i=1}^n \mathrm{X}_i^{-1}}.$$

(b) La funzione di verosimiglianza per θ è

$$L(\theta) = \prod_{i=1}^n \frac{\theta}{x_i^2}, \quad \text{se } 0 < \theta < x_i \text{ per } i = 1, \ldots, n,$$

ovvero

$$L(\theta) = \prod_{i=1}^n \frac{\theta}{x_i^2}, \quad \text{se } 0 < \theta < \min_i x_i.$$

La corrispondente log-verosimiglianza è

$$\ell(\theta) = n \log(\theta) - 2 \sum_{i=1}^n \log X_i$$

sempre che $0 < \theta < \min_i x_i$ altrimenti la log-verosimiglianza non è finita. La funzione punteggio è

$$\ell'(\theta) = \frac{n}{\theta} > 0,$$

ovvero la log-verosimiglianza è una funzione crescente e lo stimatore di massima verosimiglianza è $\hat{\theta} = \min_i X_i$.