Politecnico di Milano – Facoltà di Ingegneria Informatica

Anno Accademico 2008/2009

Corso di Statistica (2L) per INF e TEL

Docente: Antonio Pievatolo; esercitazioni: Raffaele Argiento

Esercizio 1 - Metodo dei momenti - Appello del 10/07/2007

Il reddito mensile di una certa popolazione è una variabile aleatoria continua X con densità di Pareto data da

$$f(x,a,b) = \frac{ab^a}{x^{a+1}}\mathbf{1}_{(b,\infty}(x), \quad a > 2, \; b > 0$$

I parametri a, b sono entrambi incogniti e per stimarli si analizzano i redditi $X_1, ..., X_n$ di un campione casuale di $n = 100$ individui di questa popolazione, ottenendo un reddito medio campionario pari a $\bar{x} = 1500$ € con varianza campionaria $s^2 = 750000$ €^{2}.

1. Calcolate i primi due momenti $\mu_1(a,b) = \mathbb{E}(X)$ e $\mu_2(a,b) = \mathbb{E}(X_2)$ della densità $f(x,a,b)$.
2. Determinate uno stimatore di a e uno di b usando il metodo dei momenti e calcolatene il valore sulla base delle realizzazioni campionarie fornite.

Soluzione

1.

$$\begin{aligned}
\mu_1(a,b) &= \mathbb{E}_{a,b}(X) = \int_b^\infty x\frac{ab^a}{x^{a+1}}dx = ab^a\int_b^\infty \frac{1}{x^a}dx = ab^a\left[\frac{x^{-a+1}}{-a+1}\right]_b^\infty = \frac{a}{a-1}b \\
\mu_2(a,b) &= \mathbb{E}_{a,b}(X^2) = \int_b^\infty x^2\frac{ab^a}{x^{a+1}}dx = ab^a\int_b^\infty \frac{1}{x^{a-1}}dx = ab^a\left[\frac{x^{-a+2}}{-a+2}\right]_b^\infty = \frac{a}{a-2}b^2
\end{aligned}$$

Si osservi come i due momenti esistono dato che $a > 2$.

2. Sia $M_1 = \bar{X}$ il momento primo campionario, e $M_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$ il momento secondo campionario. Allora lo stimatore per a e b col metodo dei momenti si ottiene risolvendo il sistema:

$$\begin{cases} \mathbb{E}_{a,b}(X) = M_1 \\ \mathbb{E}_{a,b}(X^2) = M_2 \end{cases} \Rightarrow \begin{cases} \frac{ab}{a-1} = M_1 \\ \frac{ab^2}{a-2} = M_2 \end{cases} \Rightarrow \begin{cases} b = \frac{a-1}{a}M_1 \\ \frac{M_1^2(a-1)^2}{a^2}\frac{a}{a-2} = M_2 \end{cases}.$$

La seconda delle due equazioni del sistema si può semplificare in

$$(M_1^2 - M_2)a^2 - 2a(M_1^2 - M^2) + M_1^2 = 0$$

Tale equazione ammette due soluzioni

$$a = 1 \pm \sqrt{\frac{M_2}{M_2 - M_1^2}}$$

(Si osservi che la quantità sotto il segno di radice è positiva!). Dato il vincolo $a > 2$ si ha una sola soluzione ammissibile e dunque gli stimatori per a e b ottenuti col metodo de momenti sono:

$$\hat{a} = 1 + \sqrt{\frac{M_2}{M_2 - M_1^2}}, \qquad \hat{b} = \frac{(\hat{a}-1)}{\hat{a}}M_1.$$

Sostituendo i valori campionari, ricordando che $(M_2 - M_1^2) = \frac{n}{n-1}S^2$, si ottengono le stime

$$\hat{a} = 3.0008, \qquad \hat{b} = 1001.33$$

∎

Esercizio 2

Sia $X_1, \ldots, X_n$ un campione iid da una distribuzione Poisson(θ).

1. Calcolare lo stimatore di massima verosimiglianza di θ.
2. Esso è non distorto?
3. È efficiente?

Soluzione

1. Sia $\underline{x} = (x_1, \ldots, x_n)$ la realizzazione campionaria relativa al campione $X_1, \ldots, X_n$. Allora la funzione di verosimiglianza è:

$$L(\theta; \underline{x}) = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} \mathrm{e}^{-\theta} \frac{\theta_i^x}{x_i!} = \mathrm{e}^{-n\theta} \frac{\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$

La log-verosimiglianza è:

$$l(\theta; \underline{x}) = \log L(\theta; \underline{x}) = -n\theta + (\sum_{i=1}^{n} x_i) \log \theta - \sum_{i=1}^{n} \log x_i!$$

Differenziando rispetto a θ

$$\frac{\partial}{\partial \theta} l(\theta; \underline{x}) = -n + \frac{\sum_{i=1}^n x_i}{\theta}.$$

Quindi

$$\frac{\partial}{\partial \theta} l(\theta; \underline{x}) \geq 0 \Rightarrow \theta \leq \frac{\sum_{i=1}^n x_i}{n}.$$

Ne consegue che lo stimatore di massima verosimiglianza è $\hat{\theta} = \bar{X}$.

2. Lo stimatore $\hat{\theta}$ trovato al punto precedente è banalmente non distorto.

3. Per verificare se $\hat{\theta}$ è efficiente, calcoliamo l'informazione di Fisher relativa ad un modello di Poissoniano:

$$\begin{aligned} I(\theta) &= \mathbb{E}_\theta \left(\frac{\partial}{\partial \theta} \log f(X, \theta) \right)^2 = \mathbb{E}_\theta \left(\frac{X}{\theta} - 1 \right)^2 = \mathbb{E}_\theta \left(1 + \frac{X^2}{\theta^2} - 2\frac{X}{\theta} \right) = \\ &= 1 + \frac{\mathbb{E}_\theta(X)}{\theta^2} - 2\frac{\mathbb{E}_\theta(X)}{\theta} = \frac{1}{\theta}. \end{aligned}$$

Osservando ora che $\mathrm{Var}_\theta(\hat{\theta}) = \frac{\theta}{n}$, concludiamo che lo stimatore ML è efficiente.
Soluzione alternativa: La derivata della log-verosimiglianza si puó fattorizzarecome segue:

$$\frac{\partial}{\partial \theta} = \frac{n}{\theta} \left(\hat{\theta} - \theta \right)$$

dal teorema di FCR (formula (11) pg.14 dispense Stima Puntuale della Prof.ssa Epifani) discende che $\hat{\theta}$ è efficiente.

∎

esercizio 3

Data la famiglia di densità:

$$f(x,\theta) = \frac{1}{\theta} x^{-(1+\frac{1}{\theta})} \mathbf{1}_{[1,\infty)}(x); \quad \theta > 0$$

1. determinare lo stimatore di massima verosimiglianza $\hat{\theta}_n$ di θ basato su un campione casuale di dimensione n estratto da $f(x,\theta)$;
2. dedurre lo stimatore di massima verosimiglianza, chiamiamolo $\hat{\tau}_n$, per $\tau(\theta) = 1/\theta$;
3. mostrare che $\hat{\tau}_n$ è asintoticamente non distorto e consistente $\tau(\theta)$;
4. determinare la f.d.r. asintotica di $\hat{\tau}_n$.

SOLUZIONE Esempio 7.14 pg.30, dispense Stima Puntuale della Prof.ssa Epifani. ▮