Università degli Studi di Roma TreCorso di laurea in Matematica
Tutorato di ST1 - A.A. 2007/2008
Docente: Prof.ssa E.Scoppola - Tutrice: Dott.ssa Katia Colaneri

Tutorato n.6 del 18/04/2008-Soluzioni

Esercizio 1

$X_1, ..., X_n$ è un campione casuale da Bernoulli(θ).
Applico il metodo di fattorizzazione:

$$f(x, \theta) = \theta^x (1-\theta)^{1-x} \mathbf{1}_{\{0;1\}}(x)$$

$$f(x_1, x_2, ..., x_n, \theta) = \prod_{i=1}^n \theta_i^x (1-\theta)^{1-x_i} \mathbf{1}_{\{0;1\}}(x_i) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \prod_{i=1}^n \mathbf{1}_{\{0;1\}}(x_i)$$

Definisco:

$$\begin{array}{rcl} s(x_1, ...x_n) & = & \sum_{i=1}^n x_i \\ g(s(x_1, ..., x_n), \theta) & = & \theta^{s(x_1, ..., x_n)} (1-\theta)^{n-s(x_1 ..., x_n)} \\ h(x_1, ..., x_n) & = & \prod_{i=1}^n \mathbf{1}_{\{0;1\}}(x_i) \end{array}$$

Osserviamo che la funzione g dipende da $x_1, ..., x_n$ solo attraverso la funzione $s(x_1, ..., x_n)$.
Allora

$$f(x_1, x_2, ..., x_n, \theta) = g(s(x_1, ..., x_n), \theta) h(x_1, ..., x_n)$$

Dunque $S = s(X_1, ...X_n) = \sum_{i=1}^n X_i$ è una statistica sufficiente per θ .

Esercizio 2

$X_1, ..., X_n$ è un campione casuale da $f(x, \theta) = \theta^2 x e^{-\theta x} \mathbf{1}_{(0,+\infty)}(x)$.
Anche in questo caso applico il metodo di fattorizzazione:

$$f(x_1, x_2, ..., x_n, \theta) = \prod_{i=1}^n \theta^2 x_i e^{-\theta x_i} \mathbf{1}_{(0,+\infty)}(x_i) = \theta^{2n} e^{\theta \sum_{i=1}^n x_i} \prod_{i=1}^n x_i \mathbf{1}_{(0,+\infty)}(x_i)$$

Definisco:

$$\begin{array}{rcl} s(x_1, ..., x_n) & = & \sum_{i=1}^n x_i \\ g(s(x_1, ..., x_n), \theta) & = & \theta^{2n} e^{\theta s(x_1, ..., x_n)} \\ h(x_1, ..., x_n) & = & \prod_{i=1}^n x_i \mathbf{1}_{(0,+\infty)}(x_i) \end{array}$$

Allora

$$f(x_1, x_2, ..., x_n, \theta) = g(s(x_1, ..., x_n), \theta) h(x_1, ..., x_n)$$

Dunque $S = s(X_1, ..., X_n) = \sum_{i=1}^n X_i$ è una statistica sufficiente per θ .

In alternativa si può osservare che $f(x,\theta) = \theta^2 x e^{-\theta x} \mathbf{1}_{(0,+\infty)}(x)$ appartiene ad una famiglia esponenziale cioè $f(x,\theta) = a(\theta)b(x)e^{c(\theta)d(x)}$ con:

$$\begin{array}{rcc} a(\theta) & = & \theta^2 \\ b(x) & = & x\mathbf{1}_{(0,+\infty)}(x) \\ c(\theta) & = & -\theta \\ d(x) & = & x \end{array}$$

Quindi per un risultato teorico $\sum_{i=1}^n d(X_i) = \sum_{i=1}^n X_i$ è una statistica sufficiente per θ.

Esercizio 3

$X_1, ..., X_n$ è un campione casuale da Uniforme sull'intervallo $(\mu - \sigma, \mu + \sigma)$, quindi

$$f(x, \mu, \sigma) = \frac{1}{2\sigma} \mathbf{1}_{(\mu-\sigma,\mu+\sigma)}(x)$$

Metodo di fattorizzazione:

$$f(x_1, ..., x_n, \mu, \sigma) = \prod_{i=1}^n \frac{1}{2\sigma} \mathbf{1}_{(\mu-\sigma,\mu+\sigma)}(x_i) = \left(\frac{1}{2\sigma}\right)^n \prod_{i=1}^n \mathbf{1}_{(\mu-\sigma,\mu+\sigma)}(x_i)$$

Siano:

$$\begin{array}{rcl} y_1 & = & min\{x_1, ..., x_n\} \\ y_n & = & max\{x_1, ..., x_n\} \end{array}$$

allora possiamo scrivere:

$$f(x_1, ..., x_n, \mu, \sigma) = \left(\frac{1}{2\sigma}\right)^n \mathbf{1}_{(\mu-\sigma,y_n)}(y_1)\mathbf{1}_{(y_1,\mu+\sigma)}(y_n)$$

Definisco:

$$\begin{array}{ccc} s_1(x_1, ..., x_n) & = & y_1 \\ s_2(x_1, ..., x_n) & = & y_n \\ g(s_1(x_1, ..., x_n), s_2(x_1, ..., x_n), \mu, \sigma) & = & \left(\frac{1}{2\sigma}\right)^n \mathbf{1}_{(\mu-\sigma,y_n)}(y_1)\mathbf{1}_{(y_1,\mu+\sigma)}(y_n) \\ h(x_1, ..., x_n) & = & 1 \end{array}$$

Dato che

$$f(x_1, ..., x_n, \mu, \sigma) = g(s_1(x_1, ..., x_n), s_2(x_1, ..., x_n), \mu, \sigma)h(x_1, ..., x_n)$$

possiamo concludere che $s_1(X_1, ..., X_n) = Y_1$ e $s_2(X_1, ..., X_n) = Y_n$ sono due statistiche congiuntamente sufficienti.

Esercizio 4

Osserviamo che $f(x,\theta_1,\theta_2) = \frac{1}{C}x^{\theta_1-1}(1-x)^{\theta_2-1}\mathbf{1}_{(0,1)}(x)$ appartiene ad una famiglia esponenziale a due parametri; infatti se definiamo:

$$\begin{array}{rcl} a(\theta_1,\theta_2) & = & 1/C \\ b(x) & = & \mathbf{1}_{(0,1)}(x) \\ c_1(\theta_1) & = & \theta_1 - 1 \\ d_1(x) & = & \ln x \\ c_2(\theta_2) & = & \theta_2 - 1 \\ d_2(x) & = & \ln(1-x) \end{array}$$

possiamo scrivere

$$f(x,\theta_1,\theta_2) = a(\theta_1,\theta_2)b(x)e^{c_1(\theta_1)d_1(x)+c_2(\theta_2)d_2(x)}$$

Dunque per un risultato teorico $S_1 = \sum_{i=1}^n d_1(X_i) = \sum_{i=1}^n \ln X_i$ e $S_2 = \sum_{i=1}^n d_2(X_i) = \sum_{i=1}^n \ln(1-X_i)$ sono due statistiche congiuntamente sufficienti.

Esercizio 5

1. $f(x,\theta) = \theta x^{\theta-1}\mathbf{1}_{(0,1)}(x)$ Scriviamo la funzione di verosimiglianza:

$$L(x_1,...,x_n,\theta) = \theta^n \prod_{i=1}^n x_i^{\theta-1}\mathbf{1}_{(0,1)}(x_i)$$

$$\log L(x_1,...,x_n,\theta) = \log(\theta^n)+(\theta-1)\log\prod_{i=1}^n x_i(\theta-1) = n\log\theta+(\theta-1)\sum_{i=1}^n \log x_i$$

$$\frac{\partial}{\partial\theta}\log L(x_1,...,x_n,\theta) = \frac{n}{\theta} + \sum_{i=1}^n \log x_i = 0 \Leftrightarrow \theta = -\frac{n}{\sum_{i=1}^n \log x_i}$$

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

2. $f(x,\theta) = \theta x^{\theta-1}\mathbf{1}_{(0,1)}(x) = \theta\mathbf{1}_{(0,1)}(x)e^{(\theta-1)\ln x}$.
Allora $f(x,\theta)$ appartiene ad una famiglia esponenziale con:

$$\begin{array}{rcl} a(\theta) & = & \theta \\ b(x) & = & \mathbf{1}_{(0,1)}(x) \\ c(\theta) & = & \theta - 1 \\ d(x) & = & \ln x \end{array}$$

$\sum_{i=1}^n d(X_i) = \sum_{i=1}^n \ln X_i$ è una statistica sufficiente minimale e completa per θ.

3.

$$\ln f(x,\theta) = \ln\theta + (\theta-1)\ln x$$

$$\frac{\partial}{\partial\theta}\ln f(x,\theta) = \frac{1}{\theta} + \ln x$$

$$\mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln(X,\theta)\right)^2\right] = \mathbb{E}\left[\left(\frac{1}{\theta}+\ln X\right)^2\right]$$

Calcoliamo $\mathbb{E}\ln X$:

$$\mathbb{E}(\ln X) = \int_{-\infty}^{+\infty}\ln x f(x,\theta)dx = \int_0^1 \theta x^{\theta-1}\ln x dx = \Big[\ln x x^{\theta}\Big]_0^1 - \int_0^1 x^{\theta-1}dx = \left[-\frac{x^{\theta}}{\theta}\right]_0^1 = -\frac{1}{\theta}$$

Allora $\mathbb{E}\left[\left(\frac{1}{\theta}+\ln X\right)^2\right] = Var(\ln X)$.

Determiniamo la distribuzione di $Y = -\ln X$:

$$P(Y<y) = P(x>e^{-y}) = \int_{e^{-y}}^1 \theta x^{\theta-1} = 1-e^{-\theta y}$$

$$f_Y(y,\theta) = \frac{d}{dy}P(Y<y) = \theta e^{-\theta y}\mathbf{1}_{(0,+\infty)}(y)$$

Y è una esponenziale di parametro θ, allora $VarY = \frac{1}{\theta^2}$.

L'informazione di Fischer è $\frac{n}{\theta^2}$.

Sia $\tau(\theta) = \frac{1}{\theta}$ uno stimatore non distorto di $\frac{1}{\theta}$.

$\tau'(\theta) = -\frac{1}{\theta^2}$ e $(\tau'(\theta))^2 = \frac{1}{\theta^4}$

Allora il limite inferiore di Cramer Rao é: $\frac{1}{n\theta^2}$

4. Abbiamo detto che $\widehat{\theta}_v = -\frac{n}{\sum_{i=1}^n \log X_i}$, dunque uno stimatore di $\frac{1}{\theta}$ è

$$\widehat{\left(\frac{1}{\theta}\right)} = -\frac{\sum_{i=1}^n \log X_i}{n}$$

Questo stimatore è non distorto perchè:

$$\mathbf{E}\left(-\frac{\sum_{i=1}^n \log X_i}{n}\right) = -\frac{1}{n}\left(-\frac{n}{\theta}\right) = \frac{1}{\theta}$$

In più

$$Var\widehat{\left(\frac{1}{\theta}\right)} = Var\left(-\frac{\sum_{i=1}^n \log X_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^n Var(\ln X_i) = \frac{1}{n^2}\frac{n}{\theta^2} = \frac{1}{n\theta^2} = lim.infC\text{–}R$$

Allora $-\frac{\sum i=1^n \log X_i}{n}$ è un UMVUE di $\frac{1}{\theta}$.

Esercizio 6

$f(x,\theta) = \frac{3x^2}{\theta^3}\mathbf{1}_{(0,\theta)}(x)$

1. $$L(x_1,...,x_n,\theta) = \prod_{i=1}^{n} \frac{3x_i^2}{\theta^3}\mathbf{1}_{(0,\theta)}(x_i) = \left(\frac{3}{\theta^3}\right)^n \prod_{i=1}^{n} x_i^2 \mathbf{1}_{(0,\theta)}(x_i)$$

   Chiamo:
   $y_1 = min\{x_1,...,x_n\}$
   $y_n = max\{x_1,...,x_n\}$.

   $$L(x_1,...,x_n,\theta) = \left(\frac{3}{\theta^3}\right)^n \prod_{i=1}^{n} x_i^2 \mathbf{1}_{(0,y_n)}(y_1)\mathbf{1}_{(y_n,+\infty)}(\theta)$$

   La funzione è decrescente in θ, dunque il massimo si ha per $\theta = y_n$
   Allora $\widehat{\theta}_v = Y_n$.

2. Applico il metodo di fattorizzazione:

   $$f(x_1,...,x_n,\theta) = \prod_{i=1}^{n} \frac{3x_i^2}{\theta^3}\mathbf{1}_{(0,\theta)}(x_i)$$

   possiamo scrivere:

   $$f(x_1,...,x_n,\theta) = \left(\frac{3}{\theta^3}\right)^n \prod_{i=1}^{n} x_i^2 \mathbf{1}_{(0,y_n)}(y_1)\mathbf{1}_{(y_n,+\infty)}(\theta)$$

   Definisco:

   $$\begin{array}{ccc} s(x_1,...,x_n) & = & y_n \\ g(s(x_1,...,x_n),\theta) & = & \left(\frac{3}{\theta^3}\right)^n \mathbf{1}_{(s(x_1,...,x_n),+\infty)}(\theta) \\ h(x_1,...,x_n) & = & \prod_{i=1}^{n} x_i^2 \mathbf{1}_{(0,y_n)}(y_1) \end{array}$$

   Dato che

   $$f(x_1,...,x_n,\mu,\sigma) = g(s(x_1,...,x_n),\theta)h(x_1,...,x_n)$$

   possiamo concludere che $s(X_1,...,X_n) = Y_n$ è una statistica sufficiente.
   La funzione di densità di Y_n è:

   $$f_{Y_n}(y,\theta) = \frac{3ny^{3n-1}}{\theta^{3n}}\mathbf{1}_{(0,\theta)}(y).$$

   Y_n è una statistica completa se $\mathbb{E}\left(Z\left(Y_n\right)\right) = 0 \Rightarrow P\left(Z\left(Y_n\right) = 0\right) = 1$
   Verifichiamolo:

   $\mathbb{E}\left(Z\left(Y_n\right)\right) = \int_0^\theta z(y)\frac{3ny^{3n-1}}{\theta^{3n}}dy = \frac{3n}{\theta^{3n}}\int_0^\theta z(y)y^{3n-1}dy = 0\ \forall\theta \Leftrightarrow$

$\frac{\partial}{\partial\theta}\int_0^\theta z(y)y^{3n-1}dy = 0\ \forall\theta \Leftrightarrow$

$z(\theta)\theta^{3n-1}\ \forall\theta \Leftrightarrow z \equiv 0$

Allora $P\left(Z\left(Y_n\right)=0\right)=1 \Rightarrow$ la statistica è completa.

3. Non possiamo calcolare il limite inferiore di Cramer Rao perchè

$$\frac{d}{d\theta}\int_0^\theta \frac{3x^2}{\theta^3}dx \neq \int_0^\theta \frac{d}{d\theta}\frac{3x^2}{\theta^3}dx$$

4. Per calcolare l'UMVUE applichiamo Lehmann-Scheffè:
consideriamo $T = Y_n$, stimatore di θ;

$$\mathbf{E}\left(Y_n\right) = \frac{3n}{3n+1}\theta$$

Questo stimatore è distorto, dunque $\widetilde{T} = \frac{3n+1}{3n}T$ è uno stimatore corretto di θ, ed è funzione di una statistica sufficiente e completa, allora è un UMVUE.