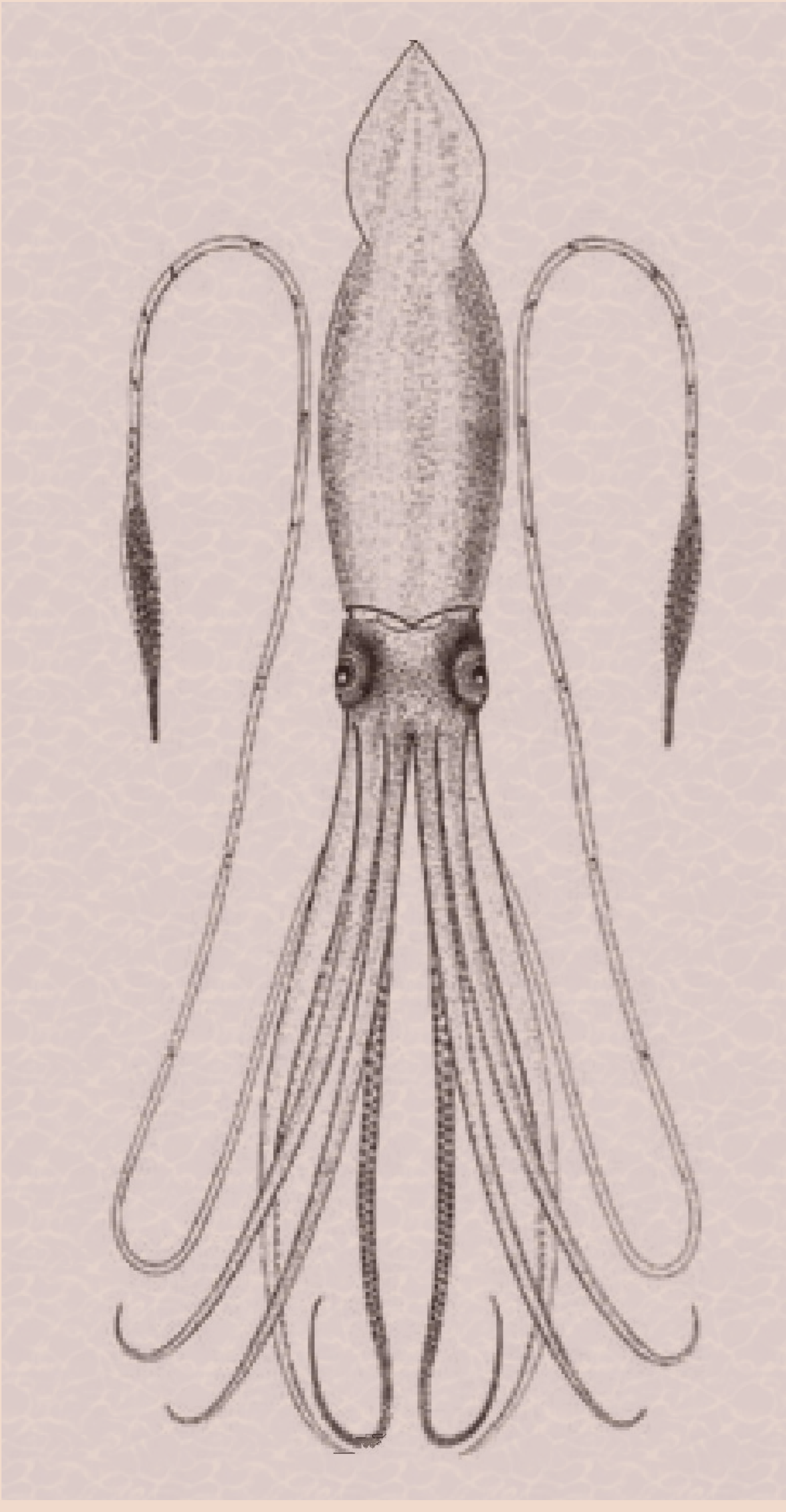
ESERCIZI di STATISTICA (corso progredito)



Gianfranco Adimari
&
Francesco Pauli

Gianfranco Adimari & Francesco Pauli.

Esercizi

di

Statistica

(corso progredito)

Versione provvisoria - aggiornata il 6 luglio 2011

Premessa

Questo volume raccoglie gran parte degli esercizi proposti come temi d'esame per l'insegnamento *Statistica (corso progredito)* -insegnamento obbligatorio dei corsi di laurea di secondo livello offerti dalla Facoltà di Scienze Statistiche di Padova- negli anni accademici dal 2004/2005 al 2009/2010.

La natura degli esercizi presentati in queste pagine ne giustifica l'organizzazione secondo una semplice lista, e suggerisce al lettore che con essi si cimenta lo studio previo di tutti gli argomenti teorici a cui fanno riferimento. Si tratta di argomenti che riguardano gli elementi classici dell'Inferenza Statistica parametrica (problemi di stima puntuale, intervallare, di verifica d'ipotesi, teoria esatta e asintotica della funzione di verosimiglianza, procedure ottime e famiglie esponenziali), con l'aggiunta dei fondamenti dell'approccio Bayesiano all'inferenza e di primi elementi dell'inferenza basata su funzioni non distorte di stima e dell'inferenza robusta. Lo scopo del volume è quello di fornire al lettore uno strumento che lo aiuti ad assimilare questi elementi di teoria della Statistica. Naturalmente, al lettore è richiesta anche la conoscenza delle nozioni trattate usualmente nei corsi di base di Analisi Matematica e Calcolo delle Probabilità.

Come testi teorici di riferimento segnaliamo, in particolare, quelli ancora oggi suggeriti per l'insegnamento citato sopra:

- *Introduzione alla Statistica. II Inferenza, Verosimiglianza, Modelli*, di Luigi Pace e Alessandra Salvan (CEDAM, Padova, 2001), compreso il manoscritto relativo al capitolo 12, reperibile al link `https://www.statistica.unipd.it/insegnamenti/statprog/matdid/dispensa_PaceSalvan.pdf`;

- *Inferenza Statistica. Una presentazione basata sul concetto di verosimiglianza*, di Adelchi Azzalini (Springer, 2000);

- *Introduzione alla Statistica Bayesiana* (capitoli 1, 3, 4 e 6), di Brunero Liseo.

Padova e Trieste, 27 gennaio 2011 G.A. F.P.

1. La tabella sottostante riporta i dati relativi al numero di asteroidi, di diametro superiore o uguale a 20km, entrati in collisione con la superficie emersa della terra negli ultimi 600 milioni di anni (abbreviato $600Ma$). In particolare, la tabella fornisce il numero di collisioni per archi temporali di ampiezza $100Ma$. Si assuma che la variabile Y che descrive il numero di collisioni in $100Ma$ segua una legge di Poisson di parametro λ e che l'osservazione disponibile possa considerarsi come un campione casuale semplice $(y_1, y_2, \ldots, y_6)$ da Y.

Periodo ($100Ma$ da oggi.)	0-1	1-2	2-3	3-4	4-5	5-6
Numero collisioni	16	6	6	2	4	1

(a) Si mostri che la famiglia $Gamma(a, b)$ [1] è coniugata naturale per il modello di Poisson.

(b) Si assuma come distribuzione *a priori* per λ una $Gamma(1, 0.2)$ e si ottenga, sulla base dell'osservazione campionaria, la stima puntuale bayesiana per il numero medio di collisioni in $100Ma$.

(c) Si dica, motivando la risposta, se l'intervallo di credibilità a più alta densità *a posteriori* (HPD) è simmetrico rispetto alla stima puntuale di cui al punto precedente.

(d) Si dia una stima puntuale bayesiana della probabilità che ci sia almeno un impatto nei prossimi $100Ma$ sulla superficie emersa.

Soluzione

(a) Basta far vedere che, scegliendo come distribuzione a priori per λ un elemento della famiglia gamma, la distribuzione a posteriori appartiene ancora alla stessa famiglia. Per la distribuzione a posteriori vale la relazione

$$\pi(\lambda|\mathbf{y}) \propto \frac{\lambda^{\sum_i y_i}}{\prod_i y_i!} e^{-n\lambda} \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda} \propto \lambda^{a+\sum_i y_i - 1} e^{-(b+n)\lambda},$$

dove, nel membro di destra, riconosciamo proprio il nucleo di una densità $Gamma(a', b')$, con $a' = a + \sum_i y_i$ e $b' = b + n$.

(b) Nel caso specifico, con $a = 1$, $b = 0.2$, $\sum_i y_i = 35$ e $n = 6$, si ottiene una distribuzione a posteriori $Gamma(36, 6.2)$. Una stima puntuale bayesiana opportuna è la media della distribuzione a posteriori. Ricordando che per la variabile casuale $Gamma(a, b)$ la speranza matematica è a/b si ha la soluzione $E(\lambda|\mathbf{y}) = 36/6.2 = 5.806$.

(c) Si ottiene un intervallo simmetrico rispetto alla media a posteriori se la funzione di densità della distribuzione a posteriori è simmetrica rispetto alla sua media. La distribuzione gamma è una distribuzione asimmetrica, si otterrà dunque un intervallo HPD asimmetrico.

(d) Osserviamo che l'entità di interesse, ovvero la probabilità (condizionata al valore λ del parametro) dell'evento $Y = 1$, è

$$\Pr\{Y \geq 1|\lambda\} = 1 - \Pr\{Y = 0|\lambda\} = 1 - e^{-\lambda}$$

Conviene considerare il complemento a uno, cioè la probabilità che non vi sia neppure un impatto in $100Ma$, $e^{-\lambda}$. Una stima puntuale di tale quantità è data, allora, dalla media a posteriori

$$\int e^{-\lambda} \pi(\lambda|\mathbf{y}) d\lambda.$$

[1] Se $U \sim Gamma(a, b)$, allora $f(u; a, b) = \frac{b^a}{\Gamma(a)} u^{a-1} \exp(-bu)$, per $u > 0$, $a > 0$ e $b > 0$.

Più precisamente, si ha

$$\begin{aligned}
\Pr\{Y=0\} &= \int \Pr\{Y=0|\lambda\}\pi(\lambda|\mathbf{y})d\lambda \\
&= \int e^{-\lambda}\frac{b'^{a'}}{\Gamma(a')}\lambda^{a'-1}e^{-b'\lambda}d\lambda \\
&= \frac{b'^{a'}}{\Gamma(a')}\int \lambda^{a'-1}e^{-(b'+1)\lambda}d\lambda \\
&= \frac{b'^{a'}}{\Gamma(a')}\frac{\Gamma(a')}{(b'+1)^{a'}} \\
&= \left(\frac{b'}{b'+1}\right)^{a'} \\
&= \left(\frac{b+n}{b+n+1}\right)^{a+\sum_i y_i} = 0.0046 \; .
\end{aligned}$$

Quindi la stima puntuale bayesiana cercata è $1-0.0046=0.9954$.

2. Sia Y la variabile casuale che descrive il numero di veicoli che transitano, in una determinata fascia oraria giornaliera, su un tratto dell'autostrada Serenissima, nel quale è installato un autovelox. Si suppone che Y abbia legge di Poisson di parametro λ e che, per il generico veicolo che transita in quel tratto di autostrada in quella fascia oraria, la probabilità di prendere una multa per eccesso di velocità è $\theta/(\lambda+1)$, con $\theta \in (0,1)$, indipendentemente da quanto possa accadere a qualsiasi altro veicolo.

Un'operazione di monitoraggio del traffico permette di ottenere, per n giorni, i dati $(\{(y_i, x_i),\ \ i = 1, 2, \ldots, n)\}$ sul numero di veicoli transitati (nel tratto e nella fascia oraria considerati) e il numero corrispondente di multe comminate. Le osservazioni relative a giorni distinti si assumono indipendenti.

(a) Si indichi con X la variabile casuale (di cui sono determinazioni le osservazioni $x_1, x_2, \ldots, x_n$) che descrive il numero di multe comminate al giorno. Si dica se la variabile (Y, X) ha distribuzione appartenente a una famiglia esponenziale.

(b) Si ottengano gli stimatori di massima verosimiglianza per λ e θ.

(c) Si mostri che il parametro $\tau = \theta\lambda/(\lambda+1)$ rappresenta il numero medio di multe comminate al giorno e si ottenga la funzione di log-verosimiglianza profilo per τ.

Soluzione

(a) Dalle ipotesi formulate si deduce che la variabile X ha distribuzione, condizionata all'evento $Y = y$, che è binomiale di parametri y e $\theta/(\lambda+1)$. Ne segue che la variabile casuale bidimensionale (Y, X) ha funzione di densità

$$\begin{aligned}
f_{Y,X}(y,x) &= f_Y(y) f_{X|Y=y}(x) \\
&= \frac{e^{-\lambda}\lambda^y}{y!} I_{\{0,1,2,\ldots\}}(y) \\
&\times \binom{y}{x} \left(\frac{\theta}{\lambda+1}\right)^x \left(1-\frac{\theta}{\lambda+1}\right)^{y-x} I_{\{0,1,2,\ldots\}}(x) I_{\{0,1,2,\ldots\}}(y-x).
\end{aligned}$$

Quindi,

$$\begin{aligned}
f_{Y,X}(y,x) &= c(\lambda)h(y,x) \\
&\times \exp\{y[\log(\lambda)+\log(\lambda+1-\theta)-\log(\lambda+1)] \\
&+ x[\log(\theta)-\log(\lambda+1-\theta)]\},
\end{aligned}$$

dove $c(\lambda)=\exp\{-\lambda\}$ e $h(y,x)=\frac{\binom{y}{x}}{y!} I_{\{0,1,2,\ldots\}}(y) I_{\{0,1,2,\ldots\}}(x) I_{\{0,1,2,\ldots\}}(y-x)$. Pertanto, la variabile (Y, X) ha distribuzione appartenente ad una famiglia esponenziale di ordine 2.

(b) Per la funzione di verosimiglianza, relativa all'osservazione $(y_1, x_1), (y_2, x_2), \ldots, (y_n, x_n)$, vale la relazione

$$L(\lambda,\theta) \propto \prod_i \lambda^{y_i} e^{-\lambda} \left(\frac{\theta}{\lambda+1}\right)^{x_i} \left(1-\frac{\theta}{\lambda+1}\right)^{y_i-x_i}.$$

Ora, per λ fissato, possiamo considerare la funzione

$$g(\theta) = \prod_i \left(\frac{\theta}{\lambda+1}\right)^{x_i} \left(1-\frac{\theta}{\lambda+1}\right)^{y_i-x_i}.$$

Si ha

$$\log g(\theta) = \sum_i [x_i \log(\theta) + (y_i - x_i)\log(\lambda+1-\theta)] + costante$$

e

$$\frac{d \log g(\theta)}{d\theta} = \sum_{i=1}^n \left[\frac{x_i}{\theta} - \frac{y_i - x_i}{\lambda+1-\theta}\right].$$

Uguagliando quindi a zero e risolvendo in θ si ottiene la stima vincolata

$$\tilde{\theta}_\lambda = (\lambda+1)\frac{\bar{x}}{\bar{y}},$$

dove $\bar{x}$ e $\bar{y}$ indicano le medie campionarie dei valori $x_1, x_2, \ldots, x_n$ e $y_1, y_2, \ldots, y_n$, rispettivamente. Si può quindi calcolare la funzione di verosimiglianza profilo per λ, la quale risulta

$$L_P(\lambda) \propto \prod_{i=1}^n \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}.$$

La massimizzazione di $L_P(\lambda)$ porta ad ottenere $\hat{\lambda} = \bar{y}$. Quindi, in definitiva, $\hat{\lambda} = \bar{y}$ e $\hat{\theta} = \tilde{\theta}_{\hat{\lambda}} = \bar{x}(\bar{y}+1)/\bar{y}$.

(c) Si ha $E[X] = E[E[X|Y=y]] = E[Y\theta/(\lambda+1)] = \lambda\theta/(\lambda+1)$. Per calcolare la log-verosimiglianza profilo per τ, possiamo considerare la riparametrizzazione $(\lambda,\theta) \to (\lambda,\tau)$ definita da

$$\tau = \theta\lambda/(\lambda+1)$$

$$\lambda = \lambda,$$

con trasformazione inversa

$$\theta = (\lambda + 1)\tau/\lambda$$
$$\lambda = \lambda.$$

Quindi,

$$L(\tau, \lambda) \propto \prod_{i=1}^{n} \lambda^{y_i} e^{-\lambda} \left(\frac{\tau}{\lambda}\right)^{x_i} \left(1 - \frac{\tau}{\lambda}\right)^{y_i - x_i}$$

e

$$l(\tau, \lambda) = \log L(\tau, \lambda) = \sum_{i=1}^{n} [-\lambda + x_i \log(\tau) + (y_i - x_i) \log(\lambda - \tau)].$$

Inoltre,

$$\frac{\partial l(\tau, \lambda)}{\partial \lambda} = -n + \sum_{i=1}^{n} \frac{y_i - x_i}{\lambda - \tau}$$

da cui, uguagliando a zero, si ottiene la stima vincolata $\tilde{\lambda}_\tau = \tau + \bar{z}$, dove $\bar{z} = (1/n) \sum_i (y_i - x_i)$. Pertanto, si ha

$$L_P(\tau) = L(\tau, \tilde{\lambda}_\tau) \propto \prod_{i=1}^{n} (\tau + \bar{z})^{y_i} e^{-(\tau + \bar{z})} \left(\frac{\tau}{\tau + \bar{z}}\right)^{x_i} \left(1 - \frac{\tau}{\tau + \bar{z}}\right)^{y_i - x_i}$$

e

$$l_P(\tau) = l(\tau, \tilde{\lambda}_\tau) = \sum_{i=1}^{n} [x_i \log(\tau) - \tau].$$

3. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile X con funzione di densità $f(x; \alpha, \lambda) = \alpha \lambda x^{\alpha - 1} \exp(-\lambda x^\alpha)$, per $x > 0$, $\alpha > 0$ e $\lambda > 0$.

(a) Supposto $\alpha = 1$ noto, si ottenga uno stimatore non distorto per la varianza di X, nella classe degli stimatori che hanno forma $\bar{X}^2/c$, con $c > 0$ opportuna costante e $\bar{X} = (1/n) \sum_{i=1}^{n} X_i$.

(b) Cosa si può dire circa l'efficienza dello stimatore di cui al punto (a) ?

(c) Si supponga ora $\lambda = 1$ noto e si faccia riferimento al modello $\mathcal{F} = \{f(y; \alpha, 1),\ \alpha > 0\}$. Si stabilisca se lo stimatore di massima verosimiglianza $\hat{\alpha}$ è robusto ad $\mathcal{F}$.

(d) Supponendo di disporre di un valore iniziale $\hat{\alpha}_0$, si ottenga l'approssimazione dello stimatore $\hat{\alpha}$ (di cui al punto precedente) fornita dal primo passo dell'algoritmo di Newton-Raphson.

Soluzione

(a) Dato che

$$E[\bar{X}^2] = var(\bar{X}) + (E[\bar{X}])^2 = 1/(n\lambda^2) + 1/\lambda^2 = (1 + n)/(n\lambda^2),$$

pechè risulti $E[\bar{X}^2/c] = 1/\lambda^2$ deve essere $c = (n + 1)/n$.

(b) Per il modello esponenziale (famiglia esponenziale regolare monoparametrica) $\sum_i x_i$ è statistica canonica, quindi statistica sufficiente minimale e completa. Ne segue che lo stimatore trovato al punto precedente, in quanto non distorto e funzione di $\sum_i x_i$, è efficiente.

(c) In questo caso $f(x;\alpha,1) = \alpha x^{\alpha-1}e^{-x^\alpha}$. Quindi, la funzione di verosimiglianza è

$$L(\alpha) = \alpha^n \prod_{i=1}^n x_i^{\alpha-1}e^{-x_i^\alpha}$$

e la funzione di log-verosimiglianza vale

$$l(\alpha) = n\log(\alpha) + \sum_{i=1}^n [(\alpha-1)\log(x_i) - x_i^\alpha].$$

Pertanto, lo score di verosimiglianza risulta essere

$$l_*(\alpha) = n/\alpha + \sum_{i=1}^n \log(x_i)(1-x_i^\alpha).$$

Questa è la funzione di stima che definisce lo stimatore di massima verosimiglianza $\hat{\alpha}$. Dato che la funzione (in x) $\log(x)(1-x^\alpha)$ è non limitata, $\hat{\alpha}$ non è robusto ad $\mathcal{F}$.

(d) Indicando con $j(\alpha)$ l'informazione osservata, si ha

$$\hat{\alpha}_1 = \hat{\alpha}_0 + \frac{l_*(\hat{\alpha}_0)}{j(\hat{\alpha}_0)} = \hat{\alpha}_0 - \frac{l_*(\hat{\alpha}_0)}{l_{**}(\hat{\alpha}_0)} = \hat{\alpha}_0 - \frac{n/\hat{\alpha}_0 + \sum_i \log(x_i)(1-x_i^{\hat{\alpha}_0})}{l_{**}(\hat{\alpha}_0)},$$

con

$$l_{**}(\alpha) = \frac{dl_*(\alpha)}{d\alpha} = -n/\alpha^2 - \sum_i x_i^\alpha[\log(x_i)]^2.$$

4. Per recarsi in ufficio, un pendolare può viaggiare in treno o in corriera. Egli vuole stabilire con quale dei due mezzi ha maggiore probabilità di arrivare in orario al lavoro sulla base dell'esperienza passata: su $n = 7$ volte che ha preso il treno è sempre arrivato in orario, mentre su $m = 9$ volte che ha preso la corriera è arrivato in ritardo 1 volta.

Siano θ e ψ le probabilità di arrivare in orario al lavoro viaggiando in treno e in corriera, rispettivamente. Si assumano come indipendenti le prove e si scelga per (θ,ψ) una distribuzione a priori che prevede che le due leggi marginali siano indipendenti e uniformi su $[0,1]$.

(a) Si ricavi la distribuzione a posteriori per (θ,ψ).
(b) Si fornisca una stima Bayesiana di $\theta - \psi$.
(c) Si trovi un intervallo di credibilità HPD al 95% per θ.
(d) Si verifichi l'ipotesi $\theta > \psi$.

Soluzione

(a) La distribuzione a priori è $\pi(\theta,\psi) \propto I_{[0,1]}(\theta)I_{[0,1]}(\psi)$. Siano X e Y le variabili che descrivono il numero di volte in cui il pendolare arriva in orario, viaggiando in treno e in corriera rispettivamente. Nelle ipotesi formulate, per tali variabili è ragionevole un modello binomiale. Pertanto, la funzione di verosimiglianza associata all'osservazione $(x = n, y = m-1)$ è

$$L(\theta,\psi) \propto \theta^n(1-\theta)^0\psi^{m-1}(1-\psi).$$

Per la distribuzione a posteriori si ha, dunque,

$$\pi(\theta,\psi|x,y) \propto \theta^n \psi^{m-1}(1-\psi) I_{[0,1]}(\theta) I_{[0,1]}(\psi),$$

che corrisponde a $(\theta|x,y) \sim$ Beta$(n+1,1)$, $(\psi|x,y) \sim$ Beta$(m,2)$ indipendenti. Pertanto,

$$\pi(\theta,\psi|x,y) = (n+1)(m+1)m\theta^n \psi^{m-1}(1-\psi) I_{[0,1]}(\theta) I_{[0,1]}(\psi).$$

(b) Calcoliamo la media a posteriori. Si ha

$$E(\theta-\psi|x,y) = E(\theta|x,y) - E(\psi|x,y) = \frac{n+1}{n+2} - \frac{m}{m+2} = 0.889 - 0.818 = 0.071.$$

(c) La distribuzione a posteriori di θ è Beta$(n+1,1)$. La relativa funzione di densità è funzione crescente in θ su $[0,1]$. Di conseguenza, un intervallo di credibilità HPD al 95% ha la forma $[C,1]$, dove C è il quantile di ordine 0.05 della distribuzione Beta, che soddisfa a

$$0.05 = \int_0^C (n+1)\theta^n d\theta = \left[\theta^{n+1}\right]_0^C = C^{n+1}.$$

Qundi $C = (0.05)^{1/(n+1)} = 0.68$

(d) Si ha

$$\Pr\{\theta > \psi|x,y\} = 1 - \int\int_{\theta<\psi} \pi(\theta,\psi|x,y) d\theta d\psi$$

dove

$$\begin{aligned}
\int\int_{\theta<\psi} \pi(\theta,\psi|x,y)d\theta d\psi &= \int\int_{\theta<\psi} (n+1)(m+1)m\theta^n\psi^{m-1}(1-\psi)d\theta d\psi \\
&= \int_0^1 (m+1)m\psi^{m-1}(1-\psi)\int_0^\psi (n+1)\theta^n d\theta d\psi \\
&= \int_0^1 (m+1)m\psi^{m-1}(1-\psi)\psi^{n+1}d\psi \\
&= (m+1)m\left(\int_0^1 \psi^{m+n}d\psi - \int_0^1 \psi^{m+n+1}d\psi\right) \\
&= (m+1)m\left(\frac{1}{m+n+1} - \frac{1}{m+n+2}\right) = 0.29\ .
\end{aligned}$$

L'ipotesi formulata è dunque accettata.

5. Sia Y una variabile casuale discreta con supporto $\{-1,0,1\}$ e distribuzione appartenente alla famiglia caratterizzata dalla legge $p(y;\theta)$, con spazio parametrico $\Theta = \{\theta_1,\theta_2,\theta_3,\theta_4\}$, specificata in tabella.

	θ_1	θ_2	θ_3	θ_4
$y=-1$	0.3	0.2	0.1	0.15
$y=0$	0.1	0.4	0.7	0.15
$y=1$	0.6	0.4	0.2	0.7

Si supponga di disporre di due realizzazioni indipendenti, y_1 e y_2, di Y.

(a) Si ricavi la funzione di verosimiglianza per θ.

(b) Si proponga un test, di livello $\alpha = 0.04$, per risolvere il problema di verifica d'ipotesi $H_0 : \theta = \theta_3$, $H_1 : \theta \neq \theta_3$.

(c) Si calcoli la funzione di potenza relativa al test di cui al punto (b), stabilendo se esso è non distorto.

Soluzione

(a) La funzione di verosimiglianza può essere rappresentata dalle quattro tabelle che seguono

$L(\theta_1; y_1, y_2) = p(y_1; \theta_1)p(y_2; \theta_1)$	$y_2 = -1$	$y_2 = 0$	$y_2 = 1$
$y_1 = -1$	0.09	0.03	0.18
$y_1 = 0$	0.03	0.01	0.06
$y_1 = 1$	0.18	0.06	0.36

$L(\theta_2; y_1, y_2) = p(y_1; \theta_2)p(y_2; \theta_2)$	$y_2 = -1$	$y_2 = 0$	$y_2 = 1$
$y_1 = -1$	0.04	0.08	0.08
$y_1 = 0$	0.08	0.16	0.16
$y_1 = 1$	0.08	0.16	0.16

$L(\theta_3; y_1, y_2) = p(y_1; \theta_3)p(y_2; \theta_3)$	$y_2 = -1$	$y_2 = 0$	$y_2 = 1$
$y_1 = -1$	0.01	0.07	0.02
$y_1 = 0$	0.07	0.49	0.14
$y_1 = 1$	0.02	0.14	0.04

$L(\theta_4; y_1, y_2) = p(y_1; \theta_4)p(y_2; \theta_4)$	$y_2 = -1$	$y_2 = 0$	$y_2 = 1$
$y_1 = -1$	0.0225	0.0225	0.105
$y_1 = 0$	0.0225	0.0225	0.105
$y_1 = 1$	0.105	0.105	0.49

Quindi, se fosse $y_1 = 1$ e $y_2 = 0$, si avrebbe $L(\theta_1; y_1, y_2) = 0.06$, $L(\theta_2; y_1, y_2) = 0.16$, $L(\theta_3; y_1, y_2) = 0.14$, $L(\theta_4; y_1, y_2) = 0.105$ e la stima di massima verosimiglianza sarebbe θ_2. Se fosse $y_1 = -1$ e $y_2 = -1$, la stima di massima verosimiglianza sarebbe θ_1. Procedendo così, caso per caso, si ottiene lo stimatore di massima verosimiglianza

$$\hat{\theta} = \begin{cases} \theta_1 & \text{se } (y_1, y_2) \in \{(-1, -1), (-1, 1), (1, -1)\} \\ \theta_2 & \text{se } (y_1, y_2) \in \{(0, -1), (-1, 0), (1, 0), (0, 1)\} \\ \theta_3 & \text{se } (y_1, y_2) \in \{(0, 0)\} \\ \theta_4 & \text{se } (y_1, y_2) \in \{(1, 1)\} \end{cases}$$

(b) Un test adeguato è il test del rapporto di verosimiglianza $\lambda^* = L(\theta_3; y_1, y_2)/[\max_\theta L(\theta; y_1, y_2)]$. La tabella che riporta i valori massimi della verosimiglianza, corrispondenti ad ogni possibile realizzazione campionaria, è la seguente:

$\max_\theta L(\theta; y_1, y_2)$	$y_2 = -1$	$y_2 = 0$	$y_2 = 1$
$y_1 = -1$	0.09	0.08	0.18
$y_1 = 0$	0.08	0.49	0.16
$y_1 = 1$	0.18	0.16	0.49

Quindi, i valori assunti dalla statistica test λ^* sono

$L(\theta_3; y_1, y_2)/[\max_\theta L(\theta; y_1, y_2)]$	$y_2 = -1$	$y_2 = 0$	$y_2 = 1$
$y_1 = -1$	0.1111	0.875	0.1111
$y_1 = 0$	0.875	1	0.875
$y_1 = 1$	0.1111	0.875	0.0816

Ne segue che, al livello 0.04, l'ipotesi nulla è rifiutata se $\lambda^* \leq 0.0816$, ossia se $(y_1, y_2) \in \mathcal{R} = \{(1, 1)\}$.

(c) La funzione di potenza è $\pi(\theta) = \Pr_\theta\{(y_1, y_2) \in \mathcal{R}\}$. Evidentemente, $\pi(\theta_3) = 0.04$. Sotto H_1 si ha : $\pi(\theta_1) = \Pr_{\theta_1}\{(y_1, y_2) \in \mathcal{R}\} = 0.36$; $\pi(\theta_2) = \Pr_{\theta_1}\{(y_1, y_2) \in \mathcal{R}\} = 0.16$; $\pi(\theta_4) = \Pr_{\theta_1}\{(y_1, y_2) \in \mathcal{R}\} = 0.49$. Il test è non distorto.

6. Sia $x_1, \dots, x_n$ un campione casuale semplice da una variabile X. Si suppone che X abbia legge normale di media μ e varianza $\theta^2\mu^2$, con $\mu > 0$ parametro ignoto e $\theta > 0$ costante nota.

(a) Si mostri che la classe parametrica considerata costituisce una famiglia esponenziale non regolare. Si fornisca la statistica sufficiente minimale per l'inferenza su μ.

(b) Si costruisca un intervallo di confidenza per μ di livello esatto 0.9.

(c) Si mostri che la statistica sufficiente minimale di cui al punto (a) non è completa.

(d) Si ottenga lo stimatore di massima verosimiglianza $\hat{\mu}$ per μ.

(e) Si stabilisca se $\hat{\mu}$ è stimatore consistente per la media di X anche quando X ha legge gamma con parametro di forma $\alpha = 1/\theta^2$ e parametro di scala $\lambda = \theta^{-2}/\mu$.

Soluzione

(a) Poiché

$$f(x; \mu) = \frac{1}{\sqrt{2\pi}\theta\mu} \exp\{-\frac{1}{2\theta^2\mu^2}(x - \mu)^2\} = \frac{e^{-1/(2\theta^2)}}{\sqrt{2\pi}\theta\mu} \exp\{\frac{x}{\theta^2\mu} - \frac{x^2}{2\theta^2\mu^2}\},$$

il modello parametrico considerato costituisce una famiglia esponenziale non regolare di odine due (ordine maggiore della dimensione del parametro). Sotto campionamento casuale semplice, la statistica canonica $T = (\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2)$ è statistica sufficiente minimale.

(b) Poiché X/μ ha distribuzione $N(1, \theta^2)$, si ha che $\bar{x}/\mu \sim N(1, \theta^2/n)$ è quantità pivotale. Inoltre, $\frac{(\bar{x}/\mu - 1)}{\theta/\sqrt{n}} \sim N(0, 1)$. Ne segue che un intervallo di confidenza esatto per μ, di livello 0.9, si ottiene invertendo la relazione $-1.64 < \frac{(\bar{x}/\mu - 1)}{\theta/\sqrt{n}} < 1.64$. Risulta

$$\frac{\bar{x}}{1 + 1.64\ \theta/\sqrt{n}} < \mu < \frac{\bar{x}}{1 - 1.64\ \theta/\sqrt{n}}.$$

(c) Dato che X/μ ha distribuzione che non dipende da μ, lo stesso accade per $\sum_{i=1}^n x_i/\mu$ e $\sum_{i=1}^n x_i^2/\mu^2$. Quindi ha distribuzione che non dipende da μ anche la funzione

$$g(T) = \frac{\sum_{i=1}^n x_i/\mu}{\sqrt{\sum_{i=1}^n x_i^2/\mu^2}}.$$

Posto $\gamma = E[g(T)]$, γ non dipende da μ e $E_\mu[g(T)-\gamma]=0$, $\forall \mu > 0$. Abbiamo quindi individuato una funzione della statistica T, non identicamente nulla, la cui media è zero per ogni valore del parametro μ. La statistica T non è completa.

(d) La funzione di log-verosimiglianza vale

$$l(\mu) = \frac{\sum_i x_i}{\theta^2 \mu} - \frac{\sum_i x_i^2}{2\theta^2\mu^2} - n\log(\mu).$$

Quindi,

$$l_*(\mu) = \frac{1}{2\theta^2\mu^3}\left(-2\mu\sum_{i=1}^n x_i + 2\sum_{i=1}^n x_i^2 - 2n\theta^2\mu^2\right).$$

Si ha $l_*(\mu)=0$ se $h(\mu) = -2\mu\sum_{i=1}^n x_i + 2\sum_{i=1}^n x_i^2 - 2n\theta^2\mu^2 = 0$. Tale equazione ha due radici, delle quali solo una, la più grande, è positiva. Infatti, $h(\mu)$ è una parabola con vertice in alto e $h(0) = 2\sum_{i=1}^n x_i^2 > 0$. Pertanto

$$\hat{\mu} = \frac{-\sum_i x_i + \sqrt{(\sum_i x_i)^2 + 4n\theta^2\sum_i x_i^2}}{2n\theta^2}.$$

(e) Se $X \sim Gamma(1/\theta^2, \theta^{-2}/\mu)$, si ha $E(X)=\mu$. Quindi, anche sotto il modello gamma, μ rappresenta la media del generico elemento della classe parametrica. Inoltre, $var(X) = \theta^2\mu^2$ e $E(X^2) = \theta^2\mu^2 + \mu^2$. Pertanto, sotto il modello gamma, per lo score di verosimiglianza ottenuto sotto l'assunto di normalità, vale la relazione

$$E[l_*(\mu)] = \frac{1}{2\theta^2\mu^3}\left[-2n\mu^2 + 2n(\theta^2\mu^2+\mu^2) - 2n\theta^2\mu^2\right] = 0.$$

Lo score di verosimiglianza ottenuto dal modello normale è, dunque, funzione di stima non distorta anche sotto il modello gamma. Ne segue che $\hat{\mu}$ è stimatore consistente per la media anche sotto il modello gamma.

Ciò può anche essere verificato direttamente. Infatti, sotto il modello gamma, si ha

$$\hat{\mu} = -\frac{\bar{x}}{2\theta^2} + \frac{\sqrt{\bar{x}^2 + 4\theta^2(1/n)\sum_i x_i^2}}{2\theta^2} \xrightarrow{p} -\frac{\mu}{2\theta^2} + \frac{\sqrt{\mu^2 + 4\theta^2(\theta^2\mu^2+\mu^2)}}{2\theta^2} = \mu$$

(utilizzando la Legge dei Grandi Numeri e il fatto che $\hat{\mu}$ è funzione continua di $\bar{x}$ e $(1/n)\sum_i x_i^2$).

7. È certo che n individui sono stati infettati da uno di tre virus (lo stesso per tutti): *HRV-A* (rinovirus umano A), *HRV-B* (rinovirus umano B) e *H3N2* (ceppo dell'influenzavirus A). Dovendo stabilire a quale dei tre virus il gruppo sia stato esposto, si rileva, per ciascun individuo, la presenza/assenza di un sintomo: la febbre elevata. È noto che tale sintomo compare con probabilità 0.3 nei soggetti colpiti dal virus *HRV-A*, 0.2 nei soggetti colpiti dal virus *HRV-B* e 0.9 in quelli colpiti dal virus *H3N2*. È noto inoltre che il virus *HRV-A* si riscontra due volte più frequentemente degli altri.

(a) Si individui l'entità di interesse, scegliendo, sulla base delle informazioni disponibili, un'opportuna distribuzione a priori.

(b) Posto $n=7$, sulla base dell'osservazione $\bar{G}_1\ G_2\ \bar{G}_3\ \bar{G}_4\ \bar{G}_5\ \bar{G}_6\ G_7$ (dove G_i indica l'evento "il soggetto i presenta il sintomo" e $\bar{G}_i$ è la negazione di G_i), si fornisca la distribuzione a posteriori.

(c) Si dia una stima puntuale bayesiana per l'entità di interesse.

(d) Si assume che il tempo necessario per la guarigione si distribuisce secondo un'esponenziale la cui media è 3 giorni per il virus *HRV-A*, 6 per il virus *HRV-B* e 7 per il virus *H3N2*. Sia dia una stima puntuale a posteriori per il tempo medio di guarigione per un soggetto esposto.

Soluzione Usiamo la notazione $P(E)$ per indicare la probabilità con cui si verifica l'evento E, cioè $P(E) = \Pr\{E\}$.

(a) L'entità di interesse è chiaramente il virus. Si dovrà quindi considerare una distribuzione discreta con tre modalità (diciamo A, B, C) rappresentanti i tre virus. L'informazione rilevante è che 'il virus A si riscontra due volte più frequentemente degli altri'. Quindi, la distribuzione di probabilità a priori è tale che, se si pone $p = P(C)$, $2p + p + p = 1$. Pertanto, a priori,

$$P(A) = 0.5 \quad P(B) = P(C) = 0.25.$$

(b) Condizionatamente al virus, la funzione di verosimiglianza associata all'osservazione $\bar{G}_1 G_2 \bar{G}_3 \bar{G}_4 \bar{G}_5 \bar{G}_6 G_7$ è proporzionale alla funzione di probabilità di una binomiale di parametri 7 e θ, dove θ indica la probabilità (condizionata al virus) di presenza del sintomo. Quindi, θ è pari a $P(G|A)$, $P(G|B)$ e $P(G|C)$, rispettivamente. La verosimiglianza è perciò proporzionale a

$$\begin{cases} P(G|A)^2(1 - P(G|A))^5 = 0.0151263 \\ P(G|B)^2(1 - P(G|B))^5 = 0.0131072 \\ P(G|C)^2(1 - P(G|C))^5 = 0.0000081 \end{cases}$$

Indicanco con O l'osservazione campionaria $\bar{G}_1 G_2 \bar{G}_3 \bar{G}_4 \bar{G}_5 \bar{G}_6 G_7$, si ha, dunque,

$$\begin{cases} P(A|O) \propto 0.0151263 \times 0.5 \\ P(B|O) \propto 0.0131072 \times 0.25 \\ P(C|O) \propto 0.0000081 \times 0.25 \end{cases}$$

e normalizzando si ottiene la distribuzione a posteriori

$$\begin{cases} P(A|O) = 0.69758 \\ P(B|O) = 0.30223 \\ P(C|O) = 0.0001877. \end{cases}$$

(c) La stima puntuale sensata in questo caso è la moda della distribuzione a posteriori, quindi il virus A.

(d) È data la media del tempo di guarigione condizionata al virus. Per ottenere la media a posteriori è sufficiente farne la media pesando con le probabilità a posteriori dei tre virus

$$3P(A|O) + 6P(B|O) + 7P(C|O) = 3.9.$$

8. La velocità di una reazione chimica dipende dalla dose di una sostanza usata come catalizzatore. In un esperimento, condotto su n prove indipendenti, viene rilevata la velocità della reazione in corrispondenza di diverse dosi (fissate). Siano $x_1, x_2, \ldots, x_n$ le dosi di catalizzatore usate, con $x_1 < x_2 < \ldots < x_n$ e $x_1 = 0$, e sia y_i, $i = 1, \ldots, n$, la velocità rilevata, corrispondente alla dose x_i. Si assume che la variabile casuale Y_i, di cui è determinazione y_i, abbia distribuzione $N(\theta + kx_i, \sigma^2)$, con θ, k e σ^2 parametri ignoti.

(a) Sia $\bar{x} = (1/n)\sum_i x_i$. Si considerino per k gli stimatori

$$T_1 = \frac{y_2 - y_1}{x_2 - x_1}, \quad T_2 = \frac{y_n - y_1}{x_n - x_1}, \quad T_3 = \sum_{i=1}^{n} y_i \frac{x_i - \bar{x}}{\sum_{j=1}^{n}(x_j - \bar{x})^2}$$

e se ne valuti correttezza e consistenza, esplicitando eventuali ipotesi sui valori x_i per garantirne la consistenza.

(b) Sia $\bar{y} = (1/n)\sum_i y_i$. Supposto noto σ^2, si consideri lo stimatore $T = \bar{y} + T_3(v - \bar{x})$ per il parametro $\tau = \theta + kv$ che rappresenta la velocità media della reazione relativa ad una dose pari a v del catalizzatore. Si ottenga una quantità pivotale basata su T, fornendo un intervallo di confidenza per τ di livello esatto 0.95.

(c) Si ponga $\sigma^2 = 1$ e si supponga θ noto. Si mostri che la funzione di stima $q(k; \boldsymbol{y}) = \sum_{i=1}^{n} g(k; y_i, x_i)$, con

$$g(k; y_i, x_i) = \frac{e^{y_i - \theta - kx_i}}{1 + e^{y_i - \theta - kx_i}} - 1/2,$$

è non distorta al modello normale considerato.

(d) Si stabilisca se lo stimatore $\hat{k}$, definito dalla funzione di stima $q(k; \boldsymbol{y})$, è robusto al modello considerato.

Soluzione

(a) Si vede facilmente che i tre stimatori sono corretti. Infatti,

$$E(T_1) = \frac{E(y_2) - E(y_1)}{x_2 - x_1} = \frac{\theta + kx_2 - (\theta + kx_1)}{x_2 - x_1} = k,$$

$$E(T_2) = \frac{E(y_n) - E(y_1)}{x_n - x_1} = \frac{\theta + kx_n - (\theta + kx_1)}{x_n - x_1} = k,$$

$$E(T_3) = \sum_i E(y_i) \frac{x_i - \bar{x}}{\sum_j (x_j - \bar{x})^2} = \sum_i (\theta + kx_i) \frac{x_i - \bar{x}}{\sum_j (x_j - \bar{x})^2} = \sum_i kx_i \frac{x_i - \bar{x}}{\sum_j (x_j - \bar{x})^2} = k.$$

Per verificare la consistenza calcoliamo le varianze.

$$var(T_1) = \frac{var(y_2 - y_1)}{(x_2 - x_1)^2} = \frac{2\sigma^2}{(x_2 - x_1)^2},$$

T_1 non è consistente (è funzione solo delle prime due osservazioni).

$$var(T_2) = \frac{var(y_n - y_1)}{(x_n - x_1)^2} = \frac{2\sigma^2}{(x_n - x_1)^2},$$

T_2 è dunque consistente se $x_n \to \infty$ (che non discende dalle ipotesi fatte).

$$var(T_3) = \sum_i var(y_i) \frac{(x_i - \bar{x})^2}{(\sum_j (x_j - \bar{x})^2)^2} = \sum_i \sigma^2 \frac{(x_i - \bar{x})^2}{(\sum_j (x_j - \bar{x})^2)^2} = \frac{\sigma^2}{\sum_j (x_j - \bar{x})^2}$$

che tende a 0, rendendo consistente T_3, se $\sum_i (x_i - \bar{x})^2 \to \infty$.

(b) Poiché T è combinazione lineare di variabili casuali normali, ha distribuzione normale. Si ha, poi,

$$E(T) = E(\bar{y}) + E(T_3)(v - \bar{x}) = \theta + k\bar{x} + kv - k\bar{x} = \tau.$$

Inoltre

$$\begin{aligned} cov\,(\bar{y}, T_3) &= cov\left(\bar{y}, \sum_h y_h \frac{x_h - \bar{x}}{\sum_j (x_j - \bar{x})^2}\right) \\ &= \sum_h \frac{x_h - \bar{x}}{\sum_j (x_j - \bar{x})^2} cov\,(\bar{y}, y_h) \\ &= \frac{\sigma^2}{n\sum_j (x_j - \bar{x})^2} \sum_h (x_h - \bar{x}) = 0. \end{aligned}$$

Quindi

$$var(T) = var(\bar{y}) + (v - \bar{x})^2 var(T_3) = \frac{\sigma^2}{n} + (v - \bar{x})^2 \frac{\sigma^2}{\sum_j (x_j - \bar{x})^2} = \frac{\sigma^2}{n}\left(1 + \frac{(v - \bar{x})^2}{s_x^2}\right),$$

dove $s_x^2 = (1/n)\sum_j (x_j - \bar{x})^2$. [2] Ne deriva che $(T - \tau)/\sigma_T$, con $\sigma_T = \sqrt{var(T)}$, ha distribuzione normale standard. Pertanto, un intervallo di confidenza di livello esatto 0.95 per τ è l'insieme $\{\tau : T - 1.96\sigma_T < \tau < T + 1.96\sigma_T\}$. Ovviamente, σ_T è quantità nota nelle ipotesi fatte.

(c) Bisogna calcolare la media di $q(k; \boldsymbol{y})$ al modello considerato. Per $z > 0$,

$$\frac{e^{-z}}{1 + e^{-z}} - 1/2 = \frac{1}{e^z + 1} - 1/2 = \frac{1 - e^z}{2(1 + e^z)}$$

e

$$\frac{e^z}{1 + e^z} - 1/2 = \frac{e^z - 1}{2(1 + e^z)}.$$

Quindi la funzione $e^z/(1 + e^z) - 1/2$ è dispari. Pertanto, al modello considerato,

$$E[q(k; \boldsymbol{y})] = \sum_{i=1}^n E[g(k; y_i, x_i)] = \sum_{i=1}^n \int [e^z/(1 + e^z) - 1/2]\phi(z)dz = 0,$$

avendo indicato con $\phi(z)$ la densità della variabile casuale normale standard, che è funzione pari.

(d) Poiché la funzione $e^z/(1 + e^z) - 1/2$ è limitata (varia tra -1/2 e 1/2), lo stimatore definito dalla funzione di stima $q(k; \boldsymbol{y})$ è robusto al modello normale considerato.

9. Sia $y_1, y_2, \ldots, y_n$ un campione casuale semplice da una variabile Y con funzione di densità $f(y; \theta, \gamma) = (\gamma/2)\exp(-\gamma|y - \theta|)$, per $y \in \Re$, con $\theta \in \Re$ e $\gamma > 0$ parametri ignoti.

(a) Si ricavi la funzione di log-verosimiglianza profilo per θ.

[2] In alternativa, si poteva ragionare direttamente sul fatto che

$$T = \frac{1}{n}\sum_i y_i + (v - \bar{x})\sum_i y_i \frac{(x_i - \bar{x})}{\sum_j (x_j - \bar{x})^2} = \sum_i y_i \left[\frac{1}{n} + \frac{(v - \bar{x})(x_i - \bar{x})}{\sum_j (x_j - \bar{x})^2}\right].$$

(b) Sulla base del campione (-0.01, 0.12, 0.28, 0.63, 0.33, 0.19, 0.77, 0.15, -1.22), si risolva, ad un livello di significatività approssimato del 5%, il problema di verifica d'ipotesi $H_0 : \theta = 0$ contro $H_1 : \theta > 0$ [3].

Si supponga $\theta = 0$ noto e γ ignoto.

(c) Si ricavi uno stimatore non distorto per $\beta = 1/\gamma$.

(d) Si costruisca un test, di livello esatto 0.05, per il problema di verifica d'ipotesi $H_0 : \beta = \beta_0$ contro $H_1 : \beta > \beta_0$.

(e) Si fornisca un'espressione per la funzione di potenza del test di cui al punto (d).

Soluzione

(a) La funzione di verosimiglianza è

$$L(\gamma, \theta) = (\gamma/2)^n e^{-\gamma \sum_i |y_i - \theta|}.$$

Quindi la funzione di log-verosimiglianza risulta essere $l(\gamma, \theta) = n \log \gamma - \gamma \sum_i |y_i - \theta|$. Derivando rispetto a γ e uguagliando a zero si ottiene la stima di massima verosimiglianza vincolata $\hat{\gamma}_\theta$

$$\hat{\gamma}_\theta = \frac{n}{\sum_i |y_i - \theta|}.$$

Si ha, dunque,

$$l_P(\theta) = l(\hat{\gamma}_\theta, \theta) = n \log(n) - n \log(\sum_i |y_i - \theta|) - n.$$

(b) La funzione di log-verosimiglianza profilo è massima quando la funzione $\sum_i |y_i - \theta|$ raggiunge il suo valore minimo. Ciò accade quando θ è pari alla mediana campionaria. Quindi $\hat{\theta} = Me = 0.19$. Il problema di verifica d'ipotesi può essere risolto ricorrendo alla statistica radice con segno

$$r_P(0) = \text{sgn}(\hat{\theta} - 0)\sqrt{2(l_P(\hat{\theta}) - l_P(0))},$$

che ha distribuzione asintotica normale standard sotto H_0. Con i dati di cui si dispone risulta $l_P(0) = -1$, $l_P(\hat{\theta}) = 0.978$ e $r_P(0) = 1.989$. Poiché $r_P(0) > 1.64$, l'ipotesi nulla è rifiutata ad un livello approssimato del 5%.

(c) Sia $T = \sum_i |y_i|$. Derivando la funzione di log-verosimiglianza $l(\gamma) = n \log \gamma - \gamma T$ rispetto a γ, si ha $l_*(\gamma) = n/\gamma - T$ e, sfruttando la nota proprietà dello score di verosimiglianza $E_\gamma[n/\gamma - T] = 0$ (prima identità di Bartlett), si ricava che $E_\gamma[T/n] = 1/\gamma$. Quindi T/n è lo stimatore cercato. Si osservi che T/n è lo stimatore di massima verosimiglianza per β.

(d) Il problema di verifica d'ipotesi in questione può essere risolto costruendo un test basato su T/n, essendo quest'ultimo stimatore consistente per β. È facile verificare che $|Y|$ ha distribuzione esponenziale di parametro γ. Quindi T ha distribuzione $Gamma(n, \gamma)$ e $\gamma T = T/\beta$ ha distribuzione $Gamma(n, 1)$. Pertanto, se $F(\cdot; n, 1)$ è la funzione di ripartizione di una variabile casuale $Gamma(n, 1)$ e c è tale che $1 - F(c; n, 1) = 0.05$, allora il test che rifiuta H_0 se $T/\beta_0 > c$ è un test esatto con livello di significatività pari a 0.05.

[3] Si assumano validi i risultati usuali, riguardanti la distribuzione asintotica delle statistiche test basate sulla funzione di verosimiglianza profilo.

(e) Per $\beta > \beta_0$, si ha

$$\pi(\beta) = \Pr_\beta\{T/\beta_0 > c\} = \Pr_\beta\{T/\beta > \beta_0 c/\beta\} = 1 - F(\beta_0 c/\beta; n, 1).$$

10. Secondo i seguaci del culto di Cthulhu, l'adorazione " dei Grandi Antichi che vissero ere prima che comparisse l'uomo, venuti dalle stelle sul mondo giovane", garantirebbe agli adepti una sopravvivenza media di più di 1000 anni. In particolare, si narra di un gruppo di adepti morti all'età di 106, 818, 2837, 227, 426, 122, 3728 e 571 anni, rispettivamente. Si supponga di voler verificare tale teoria assumendo che i dati riportati costituiscano un campione casuale semplice da una variabile X con distribuzione esponenziale di parametro $\lambda > 0$, con densità $f(x; \lambda) = \lambda e^{-\lambda x}$, $x > 0$.

(a) Si mostri che la famiglia di distribuzioni $Gamma(\alpha, \beta)$ [4] è coniugata naturale alla verosimiglianza per λ.
(b) Si fornisca uno stimatore bayesiano per la durata media di vita di un adepto.
(c) Si fornisca uno stimatore bayesiano per la probabilità che un adepto possa vivere 1000 o più anni.
(d) Scegliendo come distribuzione a priori per λ l'elemento della famiglia coniugata di media 1 e varianza 1.44, si verifichi l'ipotesi secondo la quale i seguaci del culto di Cthulhu vivrebbero in media 1000 anni o più.

Soluzione

(a) Indicando con $\boldsymbol{x}$ il campione e con n la sua dimensione, la funzione di verosimiglianza è

$$L(\lambda; \boldsymbol{x}) \propto \lambda^n e^{-\lambda \sum_i x_i}$$

e per la densità a posteriori si ha

$$\begin{aligned}\pi(\lambda|\boldsymbol{x}) &\propto \lambda^n e^{-\lambda \sum_i x_i} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \\ &\propto \lambda^{n+\alpha-1} e^{-(\beta+\sum_i x_i)\lambda}.\end{aligned}$$

La distribuzione a posteriori è quindi una $Gamma$ di parametri $\alpha^* = n+\alpha$ e $\beta^* = \beta+\sum_i x_i$.

(b) Il numero medio di anni di vita è $1/\lambda$. Si tratta dunque di trovare uno stimatore per $1/\lambda$. Si può ricorrere alla media a posteriori:

$$\begin{aligned}E(\lambda^{-1}|\boldsymbol{x}) &= \int \lambda^{-1} (\beta^*)^{\alpha^*} \Gamma(\alpha^*)^{-1} \lambda^{\alpha^*-1} e^{-\beta^*\lambda} d\lambda \\ &= (\beta^*)^{\alpha^*} \Gamma(\alpha^*)^{-1} \int \lambda^{\alpha^*-1-1} e^{-\beta^*\lambda} d\lambda \\ &= (\beta^*)^{\alpha^*} \Gamma(\alpha^*)^{-1} (\beta^*)^{-\alpha^*+1} \Gamma(\alpha^*-1) \\ &= (\beta^*)/(\alpha^*-1),\end{aligned}$$

essendo $\alpha^* > 1$ e quindi $\Gamma(\alpha^*) = (\alpha^*-1)\Gamma(\alpha^*-1)$.

[4] La funzione di densità di una variabile $Y \sim Gamma(\alpha, \beta)$ è $f(y; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}, \quad y > 0, \ \alpha > 0, \ \beta > 0$. Sia $F(y; \alpha, \beta)$ la funzione di ripartizione di Y. Allora, $F(0.001; 8.694, 8835.694) = 0.5638$, $F(1000; 8.694, 8835.694) = 1$, $F(0.001; 8835.694, 8.694) = 0$ e $F(1000; 8835.694, 8.694) = 0.06587$.

(c) Condizionatamente al valore di λ, sia τ la probabilità che un adepto possa vivere 1000 o più anni. Allora, $\tau = \Pr\{X \geq 1000\} = e^{-k\lambda}$, con $k = 1000$, e

$$\begin{aligned} E(\tau|\boldsymbol{x}) &= \int e^{-k\lambda}(\beta^*)^{\alpha^*}\Gamma(\alpha^*)^{-1}\lambda^{\alpha^*-1}e^{-\beta^*\lambda}d\lambda \\ &= \int (\beta^*)^{\alpha^*}\Gamma(\alpha^*)^{-1}\lambda^{\alpha^*-1}e^{-(\beta^*+k)\lambda}d\lambda \\ &= \frac{(\beta^*)^{\alpha^*}}{(\beta^*+k)^{\alpha^*}}\int (\beta^*+k)^{\alpha^*}\Gamma(\alpha^*)^{-1}\lambda^{\alpha^*-1}e^{-(\beta^*+k)\lambda}d\lambda \\ &= \left(\frac{\beta^*}{\beta^*+k}\right)^{\alpha^*}. \end{aligned}$$

(d) Essendo la media di una $Gamma(\alpha, \beta)$ pari a α/β e la varianza pari a α/β^2, si ha

$$\beta = E(\lambda)/var(\lambda) = 0.694$$

$$\alpha = E(\lambda)\beta = 0.694.$$

Pertanto, $\alpha^* = n + \alpha = 8.694$, $\beta^* = \beta + \sum_i x_i = 8835.694$ e

$$\Pr\{1/\lambda \geq 1000|\boldsymbol{x}\} = \Pr\{\lambda \leq 1/1000|\boldsymbol{x}\} = F(0.001; 8.694, 8835.694) = 0.5638.$$

L'ipotesi è dunque accettata.

11. Sia $x_1, x_2, \ldots, x_n$, un campione casuale semplice da una variabile casuale X con distribuzione normale di media μ_0 nota e varianza σ^2 ignota.

(a) Si stabilisca se lo stimatore $T = (1/n)\sum_{i=1}^n x_i^2 - \mu_0^2$ è stimatore efficiente (a varianza minima tra i non distorti) per σ^2.

(b) Si ottenga la funzione d'influenza (al modello normale) per lo stimatore T, stabilendo se esso è robusto.

(c) Posto $\mu_0 = 0$, si costruisca un test esatto (basato su T), di livello $\alpha = 0.05$, per il problema di verifica d'ipotesi $H_0 : \sigma^2 = \sigma_0^2$ contro $H_1 : \sigma^2 > \sigma_0^2$.

(d) Si fornisca un'espressione per la funzione di potenza del test di cui al punto precedente. Si può dire che il test è non distorto?

Soluzione

(a) Il modello parametrico normale considerato costituisce una famiglia esponenziale monoparametrica con statistica canonica (sotto campionamento casuale semplice) pari a $\sum_{i=1}^n (x_i - \mu_0)^2$. Tale statistica è sufficiente, minimale e completa. Essendo

$$E(T) = (1/n)\sum_{i=1}^n E(X_i^2) - \mu_0^2 = E(X^2) - \mu_0^2 = \sigma^2,$$

T è stimatore non distorto per σ^2. Non è però funzione della statistica sufficiente completa. Quindi T non è stimatore efficiente.

(b) La funzione di stima che definisce T come stimatore per σ^2 è

$$q(\sigma^2, \boldsymbol{x}) = (1/n)\sum_{i=1}^{n} g(\sigma^2, x_i),$$

con $g(\sigma^2, x_i) = x_i^2 - \mu_0^2 - \sigma^2$. Quindi, la funzione d'influenza associata a T, al modello parametrico considerato, è

$$IF(x; T, F_{\sigma^2}) = \frac{g(\sigma^2, x)}{-E_{F_{\sigma^2}}[\partial\, g(\sigma^2, x)/\partial\sigma^2]} = \frac{x^2 - \mu_0^2 - \sigma^2}{-E_{F_{\sigma^2}}[-1]} = x^2 - \mu_0^2 - \sigma^2.$$

Essendo la funzione di stima g non limitata in x, lo stimatore T non è robusto al modello normale considerato.

(c) Essendo T stimatore (consistente) per σ^2, una regola sensata per risolvere il problema di verifica d'ipotesi in questione potrebbe essere quella di rifiutare H_0 per valori grandi di T. Essendo $\mu_0 = 0$, si ha che X/σ_0 ha distribuzione normale standard sottoH_0. Quindi, sotto H_0, X^2/σ_0^2 ha distribuzione χ_1^2 e $(1/\sigma_0^2)\sum_{i=1}^n X_i^2$ ha distribuzione χ_n^2. Ne deriva che nT/σ_0^2 ha distribuzione χ_n^2 sotto H_0. Pertanto, il test esatto di livello $\alpha = 0.05$ basato su T rifiuta l'ipotesi nulla se

$$\frac{nT}{\sigma_0^2} > c,$$

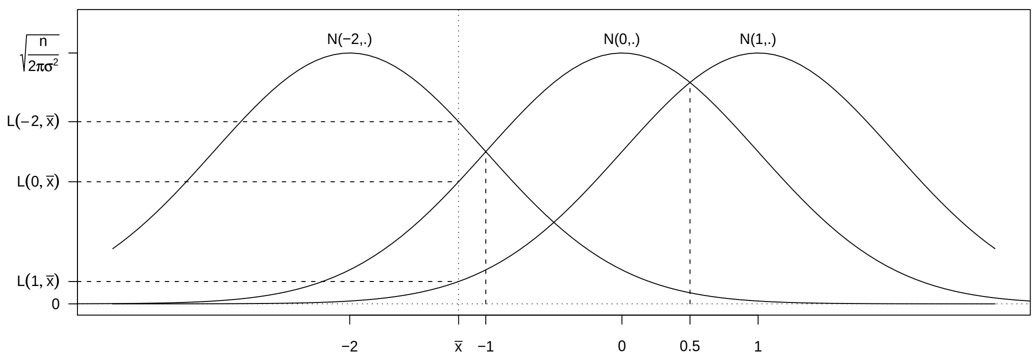
con c tale che $\Pr\{\chi_n^2 > c\} = 0.05$.

(d) La funzione di potenza associata al test del punto precedente è

$$\begin{aligned}\pi(\sigma^2) &= \Pr_{\sigma^2}\{nT/\sigma_0^2 > c\} = \Pr_{\sigma^2}\{(\sigma^2/\sigma^2)nT/\sigma_0^2 > c\} = \Pr_{\sigma^2}\{nT/\sigma^2 > c\sigma_0^2/\sigma^2\}\\ &= \Pr\{\chi_n^2 > c\sigma_0^2/\sigma^2\}.\end{aligned}$$

Pertanto, $\pi(\sigma^2)$ è funzione non decrescente. Essendo sotto H_1 $\sigma^2 > \sigma_0^2$, il test è non distorto.

12. Sia $x_1, \ldots, x_n$ un campione casuale semplice da una variabile casuale X. Si consideri per X il modello parametrico $\mathcal{N}(\theta, \sigma^2)$, con σ^2 noto e spazio parametrico $\Theta = \{-2, 0, 1\}$.



La figura data sopra riporta le funzioni di densità della statistica sufficiente minimale $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$ per i tre valori del parametro. Sull'asse delle ordinate, inoltre, sono indicati anche i livelli della funzione di verosimiglianza, corrispondenti ad un fissato valore $\bar{x}$.

(a) Si ottenga lo stimatore di massima verosimiglianza $\hat{\theta}$ per θ.

(b) Si fornisca la distribuzione di $\hat{\theta}$.

(c) Si stabilisca se $\hat{\theta}$ è stimatore non distorto o, eventualmente, asintoticamente non distorto.

(d) Si stabilisca se $\hat{\theta}$ è stimatore consistente.

Soluzione

(a) La funzione di densità congiunta è

$$f(\boldsymbol{x};\theta) = (2\pi)^{-n/2}(\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_i (x_i-\theta)^2\right\}$$

e quindi la verosimiglianza è

$$L(\theta;\boldsymbol{x}) \propto \exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\theta)^2\right\},$$

definita, ovviamente, solo su $\Theta = \{-2, 0, 1\}$. Si ha dunque (facendo riferimento al grafico)

$$\hat{\theta} = \begin{cases} -2 & \text{se } \bar{x} < -1 \\ 0 & \text{se } -1 < \bar{x} < 0.5 \\ 1 & \text{se } 0.5 < \bar{x}\ . \end{cases}$$

(b) La distribuzione di $\hat{\theta}$, se θ_0 è il vero valore del parametro, è

$$\Pr(\hat{\theta} = t; \theta_0) = \begin{cases} \Phi((-1-\theta_0)\sqrt{n}/\sigma) & \text{se } t = -2 \\ \Phi((0.5-\theta_0)\sqrt{n}/\sigma) - \Phi((-1-\theta_0)\sqrt{n}/\sigma) & \text{se } t = 0 \\ 1 - \Phi((0.5-\theta_0)\sqrt{n}/\sigma) & \text{se } t = 1 \end{cases}$$

dove $\Phi(\cdot)$ indica la funzione di ripartizione della normale standardizzata. Quindi, per i tre valori possibili per θ_0,

t	$\Pr(\hat{\theta}=t;-2)$	$\Pr(\hat{\theta}=t;0)$	$\Pr(\hat{\theta}=t;1)$
-2	$\Phi(\sqrt{n}/\sigma)$	$\Phi(-\sqrt{n}/\sigma)$	$\Phi(-2\sqrt{n}/\sigma)$
0	$\Phi(2.5\sqrt{n}/\sigma) - \Phi(\sqrt{n}/\sigma)$	$\Phi(0.5\sqrt{n}/\sigma) - \Phi(-\sqrt{n}/\sigma)$	$\Phi(-0.5\sqrt{n}/\sigma) - \Phi(-2\sqrt{n}/\sigma)$
1	$1 - \Phi(2.5\sqrt{n}/\sigma)$	$1 - \Phi(0.5\sqrt{n}/\sigma)$	$1 - \Phi(-0.5\sqrt{n}/\sigma)$

(c) Lo stimatore non è corretto. Basti notare che

$$\begin{aligned} E_{\theta_0=-2}(\hat{\theta}) &= -2\Phi(\sqrt{n}/\sigma) + 1 - \Phi(2.5\sqrt{n}/\sigma) \\ &= -2\Phi(\sqrt{n}/\sigma) + \Phi(-2.5\sqrt{n}/\sigma) \\ &> -2\Phi(\sqrt{n}/\sigma) > -2\ . \end{aligned}$$

poiché $0 < \Phi(\cdot) < 1$. Lo stimatore è peraltro asintoticamente corretto

$$\begin{aligned} E_{\theta_0=-2}(\hat{\theta}) &= -2\Phi(\sqrt{n}/\sigma) + 1 - \Phi(2.5\sqrt{n}/\sigma) \underset{n\to\infty}{\longrightarrow} -2 \\ E_{\theta_0=0}(\hat{\theta}) &= -2\Phi(-\sqrt{n}/\sigma) + 1 - \Phi(0.5\sqrt{n}/\sigma) \underset{n\to\infty}{\longrightarrow} 0 \\ E_{\theta_0=1}(\hat{\theta}) &= -2\Phi(-2\sqrt{n}/\sigma) + 1 - \Phi(-0.5\sqrt{n}/\sigma) \underset{n\to\infty}{\longrightarrow} 1\ . \end{aligned}$$

(d) Risulta che

$$\begin{aligned}
\Pr_{\theta_0=-2}(\hat\theta=-2) &= \Phi(\sqrt{n}/\sigma) \underset{n\to\infty}{\longrightarrow} 1 \\
\Pr_{\theta_0=0}(\hat\theta=0) &= \Phi(0.5\sqrt{n}/\sigma)-\Phi(-\sqrt{n}/\sigma) \underset{n\to\infty}{\longrightarrow} 1 \\
\Pr_{\theta_0=1}(\hat\theta=1) &= 1-\Phi(-0.5\sqrt{n}/\sigma) \underset{n\to\infty}{\longrightarrow} 1\ .
\end{aligned}$$

Quindi lo stimatore è consistente.

13. In un parco nazionale africano, all'inizio del periodo della migrazione, si effettua un'operazione di monitoraggio delle cicogne nere che vi risiedono. L'operazione consiste nel registrare, in un fissato giorno della settimana, per n settimane, il numero di cicogne avvistate da una postazione della guardia forestale. Sia X_i, $i=1,2,\dots,n$, la variabile casuale che descrive il numero di cicogne avvistate l'i-esima settimana e sia $x_1,x_2,\dots,x_n$ l'osservazione campionaria. Si suppone che X_1 abbia distribuzione di Poisson di parametro $\lambda>0$ e che, per $i=2,\dots,n$, la distribuzione della variabile $X_i|(X_{i-1}=x_{i-1},\dots,X_1=x_1)$ sia quella di Poisson di parametro ρx_{i-1}, con $\rho\in(0,1)$. A priori, siano λ e ρ indipendenti, con distribuzione, rispettivamente, esponenziale di parametro $\alpha>0$ e uniforme sull'intervallo $(0,1)$. [5]

(a) Si ottenga la distribuzione a posteriori per la coppia (λ,ρ).

(b) Si fornisca un intervallo di credibilità al 95% per λ. Si faccia lo stesso per ρ.

(c) Si mostri che la quantità $\psi_h=\rho^{h-1}\lambda$ rappresenta il numero atteso di avvistamenti nell'h-esima settimana di osservazione, per $h=1,2,\dots,n$.

(d) Si dia una valutazione puntuale bayesiana del numero atteso di avvistamenti nell'h-ma settimana di osservazione.

Soluzione

(a) La funzione di verosimiglianza è

$$\begin{aligned}
L(\rho,\lambda) &\propto \frac{\lambda^{x_1}e^{-\lambda}}{x_1!}\prod_{i=2}^{n}\frac{(\rho x_{i-1})^{x_i}}{x_i!}e^{-\rho x_{i-1}} \\
&\propto \lambda^{x_1}e^{-\lambda}\rho^{\sum_{i=2}^n x_i}e^{-\rho\sum_{i=2}^n x_{i-1}},
\end{aligned}$$

per $\rho\in[0,1]$ e $\lambda>0$. Pertanto per la densità a posteriori vale la relazione

$$\begin{aligned}
\pi(\rho,\lambda|\boldsymbol{x}) &\propto \lambda^{x_1}e^{-\lambda}\rho^{\sum_{i=2}^n x_i}e^{-\rho\sum_{i=2}^n x_{i-1}}\alpha e^{-\alpha\lambda}I_{[0,1]}(\rho) \\
&\propto e^{-(1+\alpha)\lambda}\lambda^{x_1}\rho^{\sum_{i=2}^n x_i}e^{-\rho\sum_{i=2}^n x_{i-1}}I_{[0,1]}(\rho).
\end{aligned}$$

E' evidente, quindi, che vi è indipendenza a posteriori tra i due parametri. Il parametro λ è distribuito secondo una $Gamma(x_1+1,\alpha+1)$ (x_1+1 è il parametro di forma, nella notazione del Pace-Salvan, p. 381). Il parametro ρ ha una densità a posteriori che è proporzionale a

[5] Per rispondere alle domande, si supponga nota la funzione di ripartizione della distribuzione gamma, $F(x;a,b)=\int_0^x \frac{b^a}{\Gamma(a)}t^{a-1}e^{-bt}dt$ (cioè la funzione $F(x;a,b)$ rimane indicata nelle risposte).

quella di una $Gamma\,(\sum_{i=2}^{n} x_i + 1, \sum_{i=2}^{n} x_{i-1})$ tra 0 e 1, mentre è 0 altrove. La costante di normalizzazione è, dunque, $1/F(1; s_2, s_1)$ e la funzione di densità a posteriori è

$$\pi(\rho|\boldsymbol{x}) = \frac{s_1^{s_2}}{\Gamma(s_2)F(1; s_2, s_1)}\rho^{s_2-1}e^{-s_1\rho}I_{[0,1]}(\rho),$$

con $s_2 = \sum_{i=2}^{n} x_i + 1$ e $s_1 = \sum_{i=2}^{n} x_{i-1}$.

(b) L'intervallo di credibilità per λ ha estremo inferiore $F^{-1}(0.025; x_1 + 1, 1 + \alpha)$ e estremo superiore $F^{-1}(0.975; x_1 + 1, 1 + \alpha)$. Per quanto riguarda ρ, si può osservare che la sua funzione di ripartizione a posteriori è

$$H(\rho|\boldsymbol{x}) = \frac{F(\rho; s_2, s_1)}{F(1; s_2, s_1)}$$

per $\rho \leq 1$ e 1 per $\rho > 1$. Quindi, l'intervallo ha estremi $H^{-1}(q) = F^{-1}(qF(1; s_2, s_1); s_2, s_1)$ per $q = 0.025$ (estremo inferiore) e $q = 0.975$ (estremo superiore).

(c) Si ha, evidentemente, che $E(X_1) = \lambda$. Inoltre, si ha

$$E(X_2) = E(E(X_2|X_1)) = E(\rho X_1) = \rho E(X_1) = \rho\lambda,$$

$$E(X_3) = E(E(X_3|X_2)) = E(\rho X_2) = \rho E(X_2) = \rho^2\lambda,$$

e, in generale quindi, per induzione,

$$E(X_h) = E(E(X_h|X_{h-1})) = E(\rho X_{h-1}) = \rho E(X_{h-1}) = \rho^{h-1}\lambda.$$

(d) La stima cercata è la media a posteriori che, in virtù dell'indipendenza a posteriori tra i parametri, è il prodotto delle medie: $E(\rho^{h-1}\lambda|\boldsymbol{x}) = E(\rho^{h-1}|\boldsymbol{x})E(\lambda|\boldsymbol{x})$. Si ha poi $E(\lambda|\boldsymbol{x}) = (x_1 + 1)/(\alpha + 1)$, mentre

$$\begin{aligned}
E(\rho^{h-1}|\boldsymbol{x}) &= \int_0^1 \rho^{h-1}\frac{s_1^{s_2}}{\Gamma(s_2)F(1; s_2, s_1)}\rho^{s_2-1}e^{-s_1\rho}d\rho \\
&= \frac{s_1^{s_2}}{\Gamma(s_2)F(1; s_2, s_1)}\int_0^1 \rho^{h-1}\rho^{s_2-1}e^{-s_1\rho}d\rho \\
&= \frac{s_1^{s_2}}{\Gamma(s_2)F(1; s_2, s_1)}F(1; s_2 + h - 1, s_1)\frac{\Gamma(s_2 + h - 1)}{s_1^{s_2+h-1}} \\
&= \frac{F(1; s_2 + h - 1, s_1)\Gamma(s_2 + h - 1)}{\Gamma(s_2)F(1; s_2, s_1)s_1^{h-1}}
\end{aligned}$$

14. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile casuale X con distribuzione uniforme sull'intervallo con estremi a e b, con $b > a$.

(a) Si ottenga lo stimatore di massima verosimiglianza per la coppia (a, b).

(b) Si stabilisca se esiste una statistica sufficiente minimale.

(c) Posto b noto, si consideri il problema di verifica d'ipotesi $H_0 : a = a_0$ contro $H_1 : a < a_0$. Si proponga un test, di livello $\alpha = 0.05$, basato sullo stimatore del metodo dei momenti per a.

Soluzione

(a) La funzione di verosimiglianza è

$$\begin{aligned} L(a,b) &= \prod_{i=1}^{n} \frac{1}{b-a} I_{[a,b]}(x_i) = \frac{1}{(b-a)^n} I_{[a,b]}(x_{(1)}) I_{[a,b]}(x_{(n)}) \\ &= \frac{1}{(b-a)^n} I_{(-\infty, x_{(1)}]}(a) I_{[x_{(n)}, +\infty)}(b), \end{aligned}$$

dove $x_{(1)}$ e $x_{(n)}$ sono, rispettivamente, il più piccolo e il più grande valore osservato. Per b fissato, $L(a,b)$ è funzione continua in a e strettamente crescente per $a \in (-\infty, x_{(1)}]$, a valori positivi. È zero altrove. Quindi il suo punto di massimo è $x_{(1)}$, che non dipende dal valore b fissato. In maniera analoga, si mostra che, per a fissato, $L(a,b)$ è massima in $x_{(n)}$, indipendentemente dal valore a fissato. Quindi $(x_{(1)}, x_{(n)})$ è lo stimatore di massima verosimiglianza.

(b) Poiché

$$L(a,b;\boldsymbol{x}) = \frac{1}{(b-a)^n} I_{(-\infty, x_{(1)}]}(a) I_{[x_{(n)}, +\infty)}(b),$$

è evidente che la verosimiglianza associata all'osservazione $\boldsymbol{x_0}$ è equivalente a quella associata all'osservazione $\boldsymbol{x_1}$ (cioè le due funzioni di verosimiglianza $L(a,b;\boldsymbol{x_0})$ e $L(a,b;\boldsymbol{x_1})$ sono proporzionali) se e solo se $x_{0(1)} = x_{1(1)}$ e $x_{0(n)} = x_{1(n)}$ (i più piccoli e i più grandi valori nei due campioni coincidono). Quindi la partizione di verosimiglianza coincide con la partizione indotta dalla statistica $(x_{(1)}, x_{(n)})$ che è dunque sufficiente minimale.

(c) Una variabile uniforme sull'intervallo $[a,b]$ ha media $(a+b)/2$ e varianza $(b-a)^2/12$. Posto $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$, lo stimatore del metodo dei momenti per a (con b noto) è $\hat{a} = 2\bar{x} - b$. In base al Teorema del Limite Centrale, $\bar{x}$ ha distribuzione approssimabile con quella $N((a+b)/2, (b-a)^2/(12n))$. Quindi, $\hat{a}$ ha distribuzione approssimata $N(a, (b-a)^2/(3n))$. Ne segue che la statistica test per il problema di verifica d'ipotesi in questione è

$$V = \frac{\sqrt{3n}(\hat{a} - a_0)}{b - a_0}$$

che ha distribuzione (approssimata) normale standard, sotto H_0. Il test rifiuta H_0 a livello approssimato del 5% se $V < -1.64$.

15. Le lunghezze (misurate in millimetri dalla punta della coda alla fine dei tentacoli) di 10 esemplari di calamari del genere Architeutis – ossia calamari giganti – sono:

$$12000, 10000, 6800, 4500, 2700, 6096, 3980, 4230, 2980, 4700.$$

Si assume che tali valori (relativi ai 10 ritrovamenti più recenti) costituiscano un campione casuale semplice da una distribuzione lognormale con parametri μ ignoto e σ^2 noto e pari a 0.23 [6].

(a) Si mostri che la classe delle distribuzioni normali costituisce la famiglia di a priori coniugata naturale alla verosimiglianza per μ.

[6] Se $X \sim N(\mu, \sigma^2)$, allora $Z = \exp(X)$ è una v.c. lognormale di parametri μ e σ^2, che ha densità $f(z;\mu,\sigma^2) = \frac{1}{\sigma z \sqrt{2\pi}} \exp\{-\frac{1}{2\sigma^2}(\log z - \mu)^2\}$, per $z > 0$, e media $\exp(\mu + \sigma^2/2)$.

Si adotti poi come distribuzione a priori per μ l'elemento della famiglia coniugata di media 9 e varianza 0.5.

(b) Si trovi una stima bayesiana per μ.

(c) Si ottenga un intervallo di credibilità al 95% per la lunghezza media di un calamaro.

(d) Si racconta che nel 1875 sia stato ritrovato, nello stretto tra il Canada e la penisola del Labrador, un calamaro lungo 52 piedi (circa 15.85 metri), poi ingloriosamente fatto a pezzi e usato come cibo per cani. Alla luce delle conclusioni tratte sopra, quanto è credibile questo rinvenimento? Cioè, quanto è probabile trovare un calamaro di lunghezza pari o superiore a 15.85 metri?

Soluzione

(a) Sia Z la variabile che descrive la lunghezza del calamaro. Per stabilire che la coniugata naturale è la famiglia di distribuzioni normali, basta osservare che se Z è lognormale di parametri μ e σ^2 allora $X = \log(Z)$ ha distribuzione $N(\mu, \sigma^2)$. Si può allora invocare la proprietà di invarianza della funzione di verosimiglianza rispetto a trasformazioni biunivoche dei dati. Pertanto, se per μ si assume una distribuzione a priori $N(\mu_0, \sigma_0^2)$ si ottiene (formule usuali per il modello normale-normale) come distribuzione a posteriori la $N(\mu^*, \sigma^{2*})$ con

$$\mu^* = \frac{\frac{n}{\sigma^2}\frac{1}{n}\sum_i \log z_i + \frac{1}{\sigma_0^2}\mu_0}{\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}} = \frac{\mu_0\sigma^2 + \bar{x}n\sigma_0^2}{\sigma^2 + n\sigma_0^2}, \qquad \sigma^{2*} = \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}\right)^{-1} = \frac{\sigma^2\sigma_0^2}{\sigma^2 + n\sigma_0^2},$$

dove si è posto $\bar{x} = \frac{1}{n}\sum_i \log z_i$. Naturalmente, il risultato può anche essere ottenuto partendo dalla funzione di verosimiglianza relativa ai dati non trasfomati,

$$\begin{aligned} L(\mu) &\propto \prod_i \frac{1}{\sqrt{2\pi}\sigma z_i} \exp\left\{-\frac{1}{2\sigma^2}(\log z_i - \mu)^2\right\} \\ &\propto \exp\left\{-\frac{1}{2\sigma^2}\sum_i (\log z_i - \mu)^2\right\} \\ &\propto \exp\left\{-\frac{1}{2\sigma^2}(n\mu^2 - 2\mu\sum_i \log z_i)\right\}, \end{aligned}$$

e calcolando esplicitamente la distribuzione a posteriori. Si ottiene

$$\begin{aligned} \pi(\mu|\boldsymbol{z}) &\propto L(\mu)\pi(\mu) \\ &\propto \exp\left\{-\frac{n}{2\sigma^2}(\mu^2 - 2\mu\bar{x})\right\}\exp\left\{-\frac{1}{2\sigma_0^2}(\mu - \mu_0)^2\right\} \\ &\propto \exp\left\{-\frac{n}{2\sigma^2}\mu^2 - \frac{1}{2\sigma_0^2}\mu^2 + \frac{\mu\bar{x}n}{\sigma^2} + \frac{\mu\mu_0}{\sigma_0^2}\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mu^2 + \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right)\mu\right\}, \end{aligned}$$

da cui la conclusione, tenendo presente che per una densità normale di parametri μ e σ^2 risulta $f(x) \propto \exp\{-\frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}\}$.

(b) Dato che $\bar{x} = 8.55499$, dalle espressioni date sopra risulta

$$\sigma^{2*} = \frac{0.23 \times 0.5}{0.23 + 10 \times 0.5} = 0.02198853$$

$$\mu^* = \sigma^{2*}\left(\frac{n}{\sigma^2}\bar{x} + \frac{1}{\sigma_0^2}\mu_0\right) = 0.02198853\left(\frac{10}{0.23}8.55499 + \frac{1}{0.5}9\right) = 8.574561.$$

La stima bayesiana per μ è la media a posteriori, ossia $\mu^* = 8.574561$.

(c) Essendo la distribuzione a posteriori per μ una $N(8.574561, 0.02198853)$, un intervallo di credibilità al 95% per μ è quello di estremi

$$8.574561 \pm \Phi^{-1}(0.025)\sqrt{0.02198853}$$

ossia $[8.28, 8.86]$. La lunghezza media di un calamaro (in termini dei parametri della lognormale) è $e^{\mu+\sigma^2/2}$, cioè una traformazione monotona di μ. Un intervallo di credibilità al 95% si ottiene pertanto trasformando gli estremi dell'intervallo al 95% per μ. Quindi l'intervallo per la lunghezza media di un calamaro è

$$[e^{8.28+\sigma^2/2}, e^{8.86+\sigma^2/2}] = [3988, 7122].$$

(d) Condizionatamente al valore di μ, il logaritmo della lunghezza, X, ha distribuzione $N(\mu, \sigma^2)$ dove $\sigma^2 = 0.23$. Inoltre, a posteriori, μ è una v.c. $N(\mu^*, \sigma^{2*})$ con $\mu^* = 8.574561$ e $\sigma^{2*} = 0.02198853$. Pertanto, $X \sim N(\mu^*, \sigma^{2*} + \sigma^2)$ [7]. Quindi $X \sim N(8.574561, 0.2519885)$. Ne segue che la probabilità di osservare un calamaro di 15.85 metri o più è

$$1 - \Phi\left(\frac{\log(15850) - 8.574561}{\sqrt{0.2519885}}\right) = 1 - \Phi(2.184059) = 0.01447896.$$

16. Un certo sistema (apparecchio) può trovarsi nello stato di funzionamento o essere in riparazione. Siano X e Y le variabili che descrivono, rispettivamente, la durata dello stato di funzionamento e la durata dell'operazione di riparazione. Si assume che X e Y seguano leggi esponenziali indipendenti, di medie (ignote) θ e λ rispettivamente. Sia $\psi = \theta/(\theta + \lambda)$.

(a) Sulla base della sequenza di n coppie (indipendenti) di osservazioni $\{(x_i, y_i),\ i = 1, 2, \ldots, n\}$, si ottenga la funzione di log-verosimiglianza profilo per ψ.

(b) Si mostri che il parametro ψ rappresenta la *disponibilità* del sistema, cioè che risulta essere $\psi = \Pr\{Y \leq X\}$.

(c) Supponendo di osservare solo i valori $\{w_1, w_2, \ldots, w_i, \ldots, w_n\}$, con $w_i = 1$ se $y_i \leq x_i$ e $w_i = 0$ altrimenti, si ottenga la regione critica più potente, di livello α, per il sistema d'ipotesi $H_0 : \psi = 0.5$ contro $H_1 : \psi > 0.5$.

(d) Quanto deve essere grande, approssimativamente, n perchè il test di cui al punto (c) abbia potenza almeno pari a 0.9 quando $\alpha = 0.05$ e $\psi = 0.6$?

[7] Si ricordi che, se $X|\mu \sim N(\mu, \sigma^2)$ e $\mu \sim N(\delta, \tau)$, allora $X \sim N(\delta, \sigma^2 + \tau)$ (si veda, ad esempio, il Lemma 1 del capitolo 6 - Brunero Liseo).

Soluzione

(a) Nelle ipotesi fatte, la funzione di verosimiglianza è

$$L(\theta, \lambda) = \prod_{i=1}^{n} \frac{1}{\theta\lambda} \exp\left(-\frac{x_i}{\theta} - \frac{y_i}{\lambda}\right).$$

Possiamo considerare la riparametrizzazione $(\theta, \lambda) \to (\theta, \psi)$ definita da

$$\psi = \theta/(\theta + \lambda)$$
$$\theta = \theta,$$

con trasformazione inversa

$$\lambda = (\theta - \theta\psi)/\psi$$
$$\theta = \theta.$$

Quindi,

$$L(\theta, \psi) = \prod_{i=1}^{n} \frac{\psi}{\theta^2(1-\psi)} \exp\left(-\frac{x_i}{\theta} - \frac{y_i\psi}{\theta(1-\psi)}\right)$$

e

$$l(\theta, \psi) = \sum_{i=1}^{n} \left[\log(\psi) - 2\log(\theta) - \log(1-\psi) - \frac{1}{\theta}\left(x_i + \frac{y_i\psi}{1-\psi}\right)\right].$$

Pertanto,

$$\frac{\partial}{\partial\theta} l(\theta, \psi) = -\frac{2n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}\left(x_i + \frac{y_i\psi}{1-\psi}\right)$$

e uguagliando a zero si ottiene la stima vincolata

$$\hat{\theta}_\psi = \frac{1}{2n}\sum_{i=1}^{n}\left(x_i + \frac{y_i\psi}{1-\psi}\right) = \frac{1}{2}\left(\bar{x} + \frac{\bar{y}\psi}{1-\psi}\right),$$

dove con $\bar{x}$ e $\bar{y}$ si sono indicate le due medie campionarie. Sostituendo θ con $\hat{\theta}_\psi$ nella espressione di $l(\theta, \psi)$, si ottiene la log-verosimiglianza profilo per ψ: $l_P(\psi) = l(\hat{\theta}_\psi, \psi)$.

(b) La funzione di densità congiunta per la coppia (X, Y) è

$$f_{X,Y}(x, y) = \frac{1}{\theta\lambda} \exp\left(-\frac{x}{\theta} - \frac{y}{\lambda}\right).$$

Quindi,

$$\begin{aligned}
\Pr\{Y \le X\} &= \int_0^\infty \left[\frac{1}{\theta}e^{-\frac{x}{\theta}} \int_0^x \frac{1}{\lambda}e^{-\frac{y}{\lambda}} dy\right] dx \\
&= \int_0^\infty \frac{1}{\theta}e^{-\frac{x}{\theta}} \left[1 - e^{-\frac{x}{\lambda}}\right] dx \\
&= 1 - \int_0^\infty \frac{1}{\theta}e^{-x\left(\frac{1}{\theta}+\frac{1}{\lambda}\right)} dx \\
&= 1 - \frac{1}{\theta}\frac{\theta\lambda}{(\theta+\lambda)} = \frac{\theta}{\theta+\lambda}
\end{aligned}$$

(c) Sia $S = \sum_{i=1}^{n} w_i$. Allora S ha distribuzione binomiale di parametri n e ψ. Per il sistema d'ipotesi

$$H_0 : \psi = \psi_0 \quad \text{contro} \quad H_1 : \psi = \psi_1$$

con $\psi_1 > \psi_0$, il test più potente di livello α è il test del rapporto di verosimiglianza $\Lambda = L(\psi_0)/L(\psi_1)$, che rifiuta H_0 per valori piccoli. Ora, $L(\psi) \propto \psi^s(1-\psi)^{n-s}$, quindi

$$\frac{L(\psi_0)}{L(\psi_1)} = \frac{\psi_0^s(1-\psi_0)^{n-s}}{\psi_1^s(1-\psi_1)^{n-s}} = \left(\frac{1-\psi_0}{1-\psi_1}\right)^n \left[\frac{\psi_0(1-\psi_1)}{\psi_1(1-\psi_0)}\right]^s .$$

Poiché $\psi_1(1-\psi_0) > \psi_0(1-\psi_1)$, Λ è funzione monotona decrescente di s. Quindi rifiutare H_0 per valori piccoli di Λ equivale a rifiutare per valori grandi di s. Questo a prescindere dal particolare valore ψ_1 fissato dall'ipotesi alternativa. Pertanto, il test più potente di livello α cercato ha regione critica $\{s : s > c_\alpha\}$, con c_α numero naturale scelto in modo che

$$\sum_{s=0}^{c_\alpha} \binom{n}{s} 0.5^s(1-0.5)^{n-s} = \sum_{s=0}^{c_\alpha} \binom{n}{s} 0.5^n = 1-\alpha$$

(approssimativamente). Si osservi che, per rispondere al quesito, si sarebbe potuto anche invocare un risultato notevole sulle famiglie esponenziali regolari.

(d) Utilizzando l'approssimazione normale alla binomiale, $S \dot\sim N(n\psi, (n\psi(1-\psi))$, si ha (ricorrendo anche alla correzione per continuità)

$$1 - \Phi\left(\frac{c + 1/2 - 0.5n}{\sqrt{n0.5^2}}\right) = 0.05$$

e

$$1 - \Phi\left(\frac{c + 1/2 - 0.6n}{\sqrt{n0.6 \times 0.4}}\right) = 0.9,$$

da cui

$$\frac{c + 1/2 - 0.5n}{\sqrt{n0.5^2}} = 1.645 \quad \text{e} \quad \frac{c + 1/2 - 0.6n}{\sqrt{n0.6 \times 0.4}} = -1.282.$$

Quindi, deve essere

$$0.5n + 1.645\sqrt{n0.5^2} = 0.6n - 1.282\sqrt{n0.6 \times 0.4}\ ,$$

da cui si ricava $n = 210.4$. Ne segue che deve essere $n = 211$.

17. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile X per la quale si considera la classe parametrica $\mathcal{F} = \{f(x;\lambda),\ \lambda > 0\}$, con

$$f(x;\lambda) = kx^{k-1}\lambda^{-k}\exp(\{-(x/\lambda)^k\}, \qquad x > 0,$$

dove k è un'opportuna costante positiva nota.

(a) Si individui una statistica sufficiente minimale per l'inferenza su λ.

(b) Si ottenga la funzione di stima che definisce lo stimatore di massima verosimiglianza $\hat\lambda$ per λ, stabilendo se tale stimatore è robusto al modello $\mathcal{F}$.

(c) Si calcoli l'informazione attesa per λ.

(d) Supponendo $\hat\lambda = 4$, $k = 2$ e $n = 100$, si ottenga un intervallo di confidenza per λ, di livello approssimato 0.95.

Soluzione

(a) La funzione di densità $f(x;\lambda)$ è evidentemente elemento di una famiglia esponenziale regolare monoparametrica. Ne segue che in questo caso una statistica sufficiente minimale per λ è

$$T = \sum_{i=1}^{n} x_i^k.$$

(b) Risulta essere

$$\log f(x;\lambda) = \log k + (k-1)\log x - (x/\lambda)^k - k\log\lambda$$

e $\frac{d}{d\lambda}\log f(x;\lambda) = kx^k/\lambda^{k+1} - k/\lambda$. Quindi lo score di verosimiglianza è

$$l_*(\lambda) = \sum_{i=1}^{n} \frac{k}{\lambda}\left(\frac{x_i^k}{\lambda^k} - 1\right).$$

Essendo la funzione $g(x;\lambda) = \frac{k}{\lambda}\left(\frac{x^k}{\lambda^k} - 1\right)$ non limitata (in x), lo stimatore di massima verosimiglianza non è robusto al modello $\mathcal{F}$.

(c) Sfruttando la proprietà di non distorsione dello score di verosimiglianza, si ricava che $E\{X^k\} = \lambda^k$. Inoltre si ha

$$l_{**}(\lambda) = -\frac{k(k+1)}{\lambda^{k+2}}\sum_{i=1}^{n} x_i^k + \frac{nk}{\lambda^2}.$$

Quindi,

$$i(\lambda) = E\{-l_{**}(\lambda)\} = \frac{k(k+1)}{\lambda^{k+2}}n\lambda^k - \frac{nk}{\lambda^2} = \frac{nk^2}{\lambda^2}.$$

(d) Usando $\hat{i} = i(\hat{\lambda})$, una stima della varianza di $\hat{\lambda}$ è $\hat{\lambda}^2/nk^2$. Un intervallo di confidenza per λ, di livello approssimato 0.95, è dato allora da $\hat{\lambda} \pm \frac{1.96\hat{\lambda}}{k\sqrt{n}}$, ovvero, con i dati a disposizione, 4 ± 0.39.

18. Un grossista di proiettili si rifornisce di modelli calibro 7.65 da due produttori A e B. È noto che il primo, A, produce i proiettili con un tasso di esemplari difettosi -nel senso che non esplodono- pari al 5%, mentre per il secondo, B, tale tasso è pari al 10%. Un venditore al dettaglio acquista un grosso lotto dal grossista ma non sa quanti dei proiettili che riceve siano fabbricati da A e quanti da B. Volendo valutare il livello di qualità del lotto, effettua un esperimento: sceglie a caso 10 proiettili dal lotto e li esplode. Sia X la variabile che conta il numero di proiettili difettosi nell'esperimento e sia θ la probabilità che un proiettile del lotto risulti difettoso.

(a) Supponendo che il venditore voglia effettuare la sua valutazione in un contesto bayesiano e che sia informato sulle referenze dei fornitori del grossista, quale distribuzione a priori sceglierebbe per l'inferenza su θ, tra una $Beta(52, 640)$, una $Beta(640, 640)$, una $Beta(640, 52)$ e una $Beta(51, 52)$? [8]

(b) Supponendo che il risultato sperimentale sia $X = 1$, sulla base dell' a priori scelta si fornisca la distribuzione a posteriori per θ.

(c) Si dia una stima puntuale bayesiana per θ.

[8] Si tenga presente che, per una $Beta(a, b)$ la media è $a/(a + b)$ e la varianza è $ab/[(a + b)^2(a + b + 1)]$.

(d) Si supponga ora che sia noto che l'intero lotto proviene da uno dei due produttori, A o B. Quale sarebbe in questo caso un' a priori opportuna?

(e) Nell'ipotesi di cui al punto precedente, si verifichi (sulla base dello stesso risultato sperimentale $X = 1$) l'ipotesi $H_0 :\ \theta > 0.075$.

Soluzione

(a) In base alle informazioni sui fornitori, si deve ritenere che θ sia compreso tra 0.05 e 0.1. Le distribuzioni mensionate hanno media, varianza e scarto quadratico medio riportati nella tabella che segue.

	$\mathrm{Beta}(52, 640)$	$\mathrm{Beta}(640, 640)$	$\mathrm{Beta}(640, 52)$	$\mathrm{Beta}(51, 52)$
media	0.07514451	0.5	0.9248555	0.4903846
var	0.0001002854	0.0001951600	0.0001002854	0.002403620
s.q.m.	0.01001426	0.01396997	0.01001426	0.04902672

E' quindi ragionevole scegliere la $\mathrm{Beta}(52, 640)$ che meglio riflette le informazioni.

(b) La funzione di verosimiglianza è quella binomiale, $L(\theta) \propto \theta^X (1-\theta)^{10-X}$. Com'è noto, la classe di distribuzioni beta è, in questo caso, coniugata naturale. La distribuzione a posteriori è una $\mathrm{Beta}(52 + 1, 640 + 9)$.

(c) Una stima puntuale bayesiana è la media della distribuzione a posteriori, quindi

$$E(\theta|X) = \frac{52 + 1}{52 + 1 + 640 + 9} = 0.075.$$

(d) Se si suppone che l'intero lotto proviene da uno dei produttori, i valori possibili per θ si riducono a due: $\theta = 0.05$ e $\theta = 0.1$. Non essendoci ulteriori informazioni, l'a priori adeguata è tale che $\Pr\{\theta = 0.05\} = \Pr\{\theta = 0.1\} = 0.5$.

(e) A posteriori,

$$\begin{aligned}\Pr\{\theta = 0.05|X = 1\} &\propto \Pr\{\theta = 0.05\} \Pr\{X = 1|\theta = 0.05\} = 0.5 \times 0.05 \times 0.95^9 = 0.01575624 \\ \Pr\{\theta = 0.1|X = 1\} &\propto \Pr\{\theta = 0.1\} \Pr\{X = 1|\theta = 0.1\} = 0.5 \times 0.1 \times 0.9^9 = 0.01937102\ .\end{aligned}$$

e quindi

$$\begin{aligned}\Pr\{\theta = 0.05|X = 1\} &= \frac{0.01575624}{0.01575624 + 0.01937102} = 0.4485474 \\ \Pr\{\theta = 0.1|X = 1\} &= \frac{0.01937102}{0.01575624 + 0.01937102} = 0.5514526\ .\end{aligned}$$

Si accetta dunque l'ipotesi H_0, in quanto

$$\frac{\Pr\{H_0|X = 1\}}{1 - \Pr\{H_0|X = 1\}} = \frac{0.5514526}{0.4485474} = 1.229419 > 1.$$

19. Sia Y una variabile casuale con funzione di densità

$$f(y; a, \theta) = \frac{\theta^a}{\Gamma(a)} y^{-a-1} e^{-\frac{\theta}{y}}, \qquad y > 0, \tag{1}$$

dove $a > 0$ e $\theta > 0$ sono due parametri.

(a) Si mostri che la variabile $X = 1/Y$ ha distribuzione gamma di parametri a e θ.

(b) Supponendo di conoscere $a > 2$ e di disporre di un campione casuale semplice $y_1, y_2, \ldots, y_n$ da Y, si ottenga lo stimatore per θ basato sul metodo dei momenti. [9]

(c) Utilizzando lo stimatore di cui al punto precedente e supponendo n grande, si ottenga un intervallo di confidenza per θ, di livello approssimato 0.95.

Si supponga ora di usare il modello (1) per descrivere la velocità sul giro -ottenuta come rapporto tra la lunghezza del circuito di prova e il tempo di percorrenza- raggiunta da alcune automobili sportive. I valori riportati rappresentano le velocità sul giro (in km/h) rilevate per due automobili, A e B, per $n = 10$ giri.

A	247.58	239.20	243.96	240.51	243.84	250.90	245.47	247.79	249.94	245.04
B	250.77	247.93	248.21	253.98	244.64	244.11	245.68	246.07	246.24	250.78

Si indichino con Y_A e Y_B le variabili che descrivono la velocità delle due automobili e si assuma per entrambe una distribuzione (1) di parametri $a = 3600$ e, rispettivamente, θ_A e θ_B ignoti. Si supponga, infine, di poter trattare le realizzazioni di Y_A e di Y_B come indipendenti.

(d) Al fine di stabilire se una delle due automobili è più veloce dell'altra, si utilizzi il test del rapporto di verosimiglianza per risolvere il problema di verifica d'ipotesi $H_0 : \theta_A = \theta_B$ contro $H_1 : \theta_A \neq \theta_B$, al livello approssimato del 5%.

Soluzione

(a) Lo jacobiano della trasformazione è

$$\left|\frac{dy(x)}{dx}\right| = x^{-2};$$

quindi la funzione di densità di X è

$$f(x) = \frac{\theta^a}{\Gamma(a)} \frac{1}{x^{-a-1}} e^{-\theta x} x^{-2} = \frac{\theta^a}{\Gamma(a)} x^{a-1} e^{-\theta x}, \quad \text{per} \quad x > 0,$$

ossia quella di una variabile casuale gamma di parametri a e θ.

(b) Lo stimatore si ricava eguagliando la media teorica alla media campionaria $\bar{y}$,

$$\frac{\theta}{a-1} = \bar{y} \quad \Rightarrow \quad \tilde{\theta} = (a-1)\bar{y} \ .$$

(c) In base al Teorema del Limite Centrale, la media campionaria ha distribuzione approssimabile con una $N\left(E(Y), \frac{var(Y)}{n}\right)$. Quindi $\tilde{\theta}$ ha distribuzione $N\left((a-1)E(Y), \frac{(a-1)^2 var(Y)}{n}\right)$ cioè $N\left(\theta, \frac{(a-1)^2 var(Y)}{n}\right)$. Alla luce di ciò, l'intervallo di confidenza richiesto ha estremi

$$\tilde{\theta} \pm \Phi^{-1}(0.025) \frac{(a-1)\sqrt{v\hat{a}r(Y)}}{\sqrt{n}} \ ,$$

dove $v\hat{a}r(Y) = \frac{\tilde{\theta}^2}{(a-1)^2(a-2)}$.

[9] Quando è $a > 2$, la distribuzione (1) ha media $\theta/(a-1)$ e varianza $\theta^2/[(a-1)^2(a-2)]$.

(d) Occorre calcolare la quantità $W = 2(l(\hat{\theta}_A, \hat{\theta}_B) - l(\hat{\theta}))$ dove $l(\hat{\theta})$ è il massimo della log-verosimiglianza calcolata nell'ipotesi H_0 di eguaglianza dei parametri $\theta_A = \theta_B = \theta$, mentre $l(\hat{\theta}_A, \hat{\theta}_B)$ è il massimo della log-verosimiglianza nell'ipotesi H_1, cioè con parametri differenti per ciascuno gruppo di osservazioni.

Calcoliamo, in termini generali, la stima di massima verosimiglianza -e il valore corrispondente della log-verosimiglianza- per un generico campione $x_1, \ldots, x_n$ da una variabile con densità (1). Si ha

$$L(\theta) \propto \frac{\theta^{na}}{\Gamma(a)^n} \prod x_i^{-a-1} e^{-\theta \sum_i x_i^{-1}} \propto \theta^{na} e^{-\theta \sum_i x_i^{-1}}.$$

La log-verosimiglianza è quindi

$$l(\theta) = na \log(\theta) - \theta \sum_i x_i^{-1}$$

ed è massima per θ che soddisfa

$$0 = l_*(\theta) = \frac{na}{\theta} - \sum_i x_i^{-1}.$$

Pertanto, $\hat{\theta} = \frac{na}{\sum_i x_i^{-1}}$. Il massimo della log-verosimiglianza vale

$$l(\hat{\theta}) = na \log\left(\frac{na}{\sum_i x_i^{-1}}\right) - \frac{na}{\sum_i x_i^{-1}} \sum_i x_i^{-1} = na \log(na) - na \log\left(\sum_i x_i^{-1}\right) - na.$$

Ora, nell'ipotesi H_0, l'osservazione campionaria è equivalente ad un'unico campione casuale semplice di dimensione $n_A + n_B = 20$. Ne segue che

$$l(\hat{\theta}) = 20 \times 3600 \log(20 \times 3600) - 20 \times 3600 \log(0.08110897) - 20 \times 3600 = 914139.6,$$

dove $0.08110897 = \sum_i y_{Ai}^{-1} + \sum_i y_{Bi}^{-1}$. Nell'ipotesi H_1, la log-verosimiglianza realativa all'intera osservazione non è altro che la somma delle log-verosimiglianze associate alle 10 misure rilevate su ciascuno dei due tipi di automobile: $l(\theta_A, \theta_B) = l(\theta_A) + l(\theta_B)$. Massimizzando separatamente si ottiene

$$\begin{aligned} l(\hat{\theta}_A) + l(\hat{\theta}_B) &= 10 \times 3600 \log(10 \times 3600) - 10 \times 3600 \log(0.04075469) - 10 \times 3600 + \\ &\quad +10 \times 3600 \log(10 \times 3600) - 10 \times 3600 \log(0.04035429) - 10 \times 3600 \\ &= 456892.5 + 457247.9 = 914140.4, \end{aligned}$$

dove $0.04075469 = \sum_i y_{Ai}^{-1}$ e $0.04035429 = \sum_i y_{Bi}^{-1}$. Pertanto $W = 2(914140.4 - 914139.6) = 1.6$. Sotto H_0, W ha distribuzione limite χ_1^2. Facendo riferimento alle tavole della distribuzione χ_1^2, si ricava che il livello di significatività osservato (approssimato) del test è pari a 0.2. Ciò porta ad accettare l'ipotesi nulla.

20. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile X con distribuzione esponenziale di parametro $\lambda > 0$ ignoto, avente densità $f_X(x; \lambda) = \lambda e^{-\lambda x}$, per $x > 0$. Sia $F_X(x; \lambda)$ la funzione di ripartizione di X e, per $p \in (0, 1)$ fissato, si indichi con ξ_p il quantile di ordine p della distribuzione di X, cioè quel punto del supporto di X tale che $F_X(\xi_p; \lambda) = p$.

(a) Si ottenga ξ_p in funzione di λ e si ricavi lo stimatore di massima verosimiglianza $\hat{\xi}_p$.

(b) Si ricavi la distribuzione esatta di $\hat{\xi}_p$.

(c) Si ottenga un intervallo di confidenza per ξ_p, di livello (esatto) 0.95.

Si consideri la funzione di stima $q(t;\boldsymbol{x}) = \sum_{i=1}^n g(t;x_i)$ con

$$g(t;x_i) = \begin{cases} -1 & \text{se } x_i - t \le 0 \\ p/(1-p) & \text{se } x_i - t > 0 \end{cases}$$

(d) Si mostri che il funzionale statistico $T(\cdot)$ associato alla funzione $q(t;\boldsymbol{x})$ definisce ξ_p non solo per il modello esponenziale ma, più in generale, per la classe di distribuzioni continue aventi funzione di densità $f(x)$ strettamente positiva in un intorno del proprio quantile di ordine p.

Soluzione

(a) Dalla relazione $F_X(\xi_p;\lambda) = 1 - e^{-\lambda\xi_p} = p$ si ricava $\xi_p = -(1/\lambda)\log(1-p)$. Dato che lo stimatore di massima verosimiglianza per λ è $\hat{\lambda} = \bar{x}^{-1}$, con $\bar{x} = (1/n)\sum_i x_i$, si ha

$$\hat{\xi}_p = -\bar{x}\log(1-p),$$

per la proprietà di equivarianza (dello stimatore di massima verosimiglianza).

(b) Dato che $\sum_{i=1}^n X_i \sim Gamma(n,\lambda)$, si ha che

$$\hat{\xi}_p \sim Gamma(n, \lambda/\kappa)$$

con $\kappa = -(1/n)\log(1-p)$. Ovvero, $\hat{\xi}_p \sim Gamma(n, n/\xi_p)$.

(c) Dal risultato precedente, $\hat{\xi}_p/\xi_p \sim Gamma(n,n)$. Quindi un intervallo di confidenza di livello (esatto) 0.95 ha per estremi (inferiore e superiore) $\hat{\xi}_p/b$ e $\hat{\xi}_p/a$, dove $\Pr\{T < a\} = 0.025$ e $\Pr\{T > b\} = 0.025$ se $T \sim Gamma(n,n)$.

(d) In base alle ipotesi fatte, il quantile di ordine p della distribuzione caratterizzata dalla densità $f(x)$ è univocamente definito e risulta $\xi_p = F^{-1}(p)$. Qui, con $F(\cdot)$ indichiamo la funzione di ripartizione corrispondente a $f(x)$. Bisogna, allora, calcolare il funzionale $T(F)$ definito implicitamente dall'equazione

$$\int q(t;\boldsymbol{x})f(x)dx = 0.$$

Poiché

$$-\int_{-\infty}^{t} f(x)dx + \frac{p}{1-p}\int_{t}^{+\infty} f(x)dx = \frac{-F(t)+p}{1-p},$$

uguagliando a zero si ottiene $F(t) = p$, da cui $t = \xi_p$.

21. Dovendo valutare la precisione di uno strumento per la misura della radiazione elettromagnetica, si effettuano n misurazioni $x_1,\ldots,x_n$ su un campo elettromagnetico di intensità nota pari a μ microtesla. Si suppone che le misurazioni siano determinazioni indipendenti e identicamente distribuite di una variabile casuale X con distribuzione $N(\mu,\sigma^2)$. A priori, inoltre, si assume che $\tau = 1/\sigma^2$ abbia distribuzione appartenente alla famiglia $Gamma(\alpha,\beta)$[10].

[10]La funzione di densità di una v.c. Y con distribuzione $Gamma(\alpha,\beta)$ è $f(y;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\beta y}$, con media e varianza pari, rispettivamente, a α/β e α/β^2

(a) Si stabilisca se la famiglia $Gamma(\alpha, \beta)$ è coniugata naturale alla verosimiglianza.

Con $n = 10$ e $\mu = 10$, l'osservazione campionaria è tale che $\sum_i x_i = 97.96$ e $\sum_i x_i^2 = 1056.16$. Si sceglie per τ la distribuzione a priori con media 0.15 e varianza 0.4.

(b) Si fornisca una stima bayesiana di τ.

(c) Si fornisca una stima bayesiana di σ^2.

(d) Si dica quale tra i seguenti è l'intervallo di credibilità HPD (alta densità a posteriori) al 99% per τ, giustificando la risposta:

$$[0.0158, 0.2425] \;\; ; \;\; [0.0278, 0.1792] \;\; ; \;\; [0, 0.2425].$$

(e) Si stabilisca se, al crescere di n, la stima bayesiana di σ^2 data al punto (c) e la stima di massima verosimiglianza tendono a coincidere.

Soluzione

(a) La funzione di verosimiglianza (in termini del parametro τ) è

$$L(\mathbf{x}|\tau) \propto f(\mathbf{x}|\tau) = (2\pi)^{-n/2}\tau^{n/2} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}(x_i - \mu)^2\right).$$

Per la densità a posteriori si ha, quindi,

$$\begin{aligned}\pi(\tau|\mathbf{x}) &\propto \pi(\tau)L(\mathbf{x}|\tau) \propto \tau^{\alpha-1}\exp(-\beta\tau)\,\tau^{n/2}\exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\\ &\propto \tau^{n/2+\alpha-1}\exp\left(-\tau\left(\beta + \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right).\end{aligned}$$

Pertanto, $\tau|\mathbf{x}$ ha distribuzione $Gamma(\alpha^*, \beta^*)$ ove $\alpha^* = \alpha + n/2$ e $\beta^* = \beta + \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2$. La famiglia considerata è quindi coniugata naturale alla verosimiglianza.

(b) Se a priori deve essere $E(\tau) = 0.15$ e $var(\tau) = 0.4$, essendo la media di una $Gamma(\alpha, \beta)$ pari a α/β e la varianza pari a α/β^2, si ha

$$\beta = E(\tau)/var(\tau) = 0.375$$

$$\alpha = E(\tau)\beta = 0.05625.$$

Uno stimatore per τ è la media della distribuzione a posteriori α^*/β^* o la moda a posteriori $(\alpha^* - 1)/\beta^*$. Risulta $\alpha^* = 0.05625 + n/2 = 5.05625$ e $\beta^* = \beta + \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2 = \beta + \frac{1}{2}[\sum_{i=1}^{n} x_i^2 + n\mu^2 - 2\mu\sum_{i=1}^{n} x_i] = 0.375 + \frac{1}{2}[1056.16 + 1000 - 1959.2] = 48.855$. Quindi, utilizzando la media a posteriori, la stima di τ è $\alpha^*/\beta^* = 5.05625/48.855 = 0.1035$.

(c) Si ha che

$$\begin{aligned}E(\sigma^2|\mathbf{x}) &= \int_0^{+\infty}\tau^{-1}\pi(\tau|\mathbf{x})d\tau = \frac{(\beta^*)^{\alpha^*}}{\Gamma(\alpha^*)}\int_0^{+\infty}\tau^{\alpha^*-2}e^{-\beta^*\tau}d\tau = \frac{(\beta^*)^{\alpha^*}}{\Gamma(\alpha^*)}\frac{\Gamma(\alpha^*-1)}{(\beta^*)^{\alpha^*-1}}\\ &= \frac{\beta^*}{\alpha^*-1}.\end{aligned}$$

Quindi una stima bayesiana di σ^2 è $48.855/(5.05625 - 1) = 12.04438$.

(d) Data la forma della distribuzione gamma, l'intervallo HPD deve essere asimmetrico rispetto alla media a posteriori. Ciò esclude il secondo tra gli intervalli proposti. D'altro canto, l'HPD non può avere estremo inferiore 0, dato che il limite della densità è 0 in 0: rimane il primo intervallo tra quelli proposti.

(e) Adottiamo come stimatore bayesiano la media a posteriori. Sicché, detto $\hat{\sigma}^2 = \sum_i (x_i - \mu)^2/n$, si ha

$$E(\sigma^2|\mathbf{x}) = \frac{\beta + n\hat{\sigma}^2/2}{\alpha + n/2 - 1} = \frac{2\beta/n + \hat{\sigma}^2}{2\alpha/n + 1 - 2/n} \to_{n\to\infty} \hat{\sigma}^2.$$

22. Una particolare varietà di margherita presenta petali di due colori diversi, bianco e giallo. Siano X e Y le variabili che descrivono, rispettivamente, le lunghezze medie dei petali bianchi e gialli di un fiore e sia $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ un campione di misure rilevate su n margherite di quella varietà. Si supponga di poter assumere l'indipendenza tra X e Y e sia ρ la probabilità che i petali gialli siano, in media, più lunghi di quelli bianchi, cioè $\rho = \Pr\{Y > X\}$.

(a) Si mostri che $t(\boldsymbol{x}, \boldsymbol{y}) = (1/n) \sum_{i=1}^n I_{(0,+\infty)}(y_i - x_i)$ è uno stimatore consistente per ρ.

(b) Si costruisca un intervallo di confidenza per ρ, di livello approssimato γ, basato su t.

Si supponga ora che X e Y seguano leggi esponenziali, di parametri λ e θ rispettivamente. In particolare, quindi, $f(x;\lambda) = \lambda e^{-\lambda x}$, $x > 0$, $\lambda > 0$. Si supponga inoltre che, con $n = 100$, l'osservazione campionaria sia tale che $\sum_i x_i = 1873.3$ (millimetri) e $\sum_i y_i = 1987.7$.

(c) Si ricavi ρ in funzione di λ e θ e se ne fornisca una stima.

(d) Si risolva il problema di verifica d'ipotesi $H_0 : \lambda = \theta$ contro $H_1 : \lambda \neq \theta$.

Soluzione

(a) Si osservi che $I_{(0,+\infty)}(Y - X)$ ha distribuzione binomiale elementare $Bi(1, \rho)$. Per la legge dei grandi numeri, t converge in probabilità a $E[I_{(0,+\infty)}(Y - X)] = \rho$. Quindi t è stimatore consistente per ρ.

(b) Poiché $var[I_{(0,+\infty)}(Y - X)] = \rho(1 - \rho)$, per il teorema del limite centrale

$$\frac{\sqrt{n}(t - \rho)}{\sqrt{\rho(1 - \rho)}} \mathrel{\dot\sim} N(0, 1),$$

e un intervallo di confidenza, di livello approssimato γ, per ρ è dato da

$$t - \frac{\sqrt{t(1 - t)}}{\sqrt{n}} z_\gamma \leq \rho \leq t + \frac{\sqrt{t(1 - t)}}{\sqrt{n}} z_\gamma,$$

con z_γ tale che $\Phi(-z_\gamma) = (1 - \gamma)/2$ e $\Phi(\cdot)$ funzione di ripartizione della normale standard.

(c) Quando X e Y sono variabili casuali indipendenti, con distribuzione esponenziale di parametro λ e θ, rispettivamente, la funzione di verosimiglianza risulta essere

$$L(\lambda, \theta) = \prod_{i=1}^n f(x_i; \lambda) f(y_i; \theta) = \prod_{i=1}^n \lambda e^{-\lambda x_i} \theta e^{-\theta y_i},$$

cosicché

$$l(\lambda,\theta)=\log(L(\lambda,\theta))=n\log\lambda-\lambda\sum_i x_i+n\log\theta-\theta\sum_i y_i.$$

Ne consegue, evidentemente, che $\hat\lambda=n/\sum_i x_i$ e $\hat\theta=n/\sum_i y_i$. D'altro canto,

$$\begin{aligned}\rho=\Pr\{Y>X\} &= E[I_{(0,+\infty)}(Y-X)]=E[E[I_{(0,+\infty)}(Y-X)|X=x]]\\ &= E[\Pr\{Y>x\}]=E[e^{-\theta X}]=\int_0^{+\infty}e^{-\theta x}\lambda e^{-\lambda x}dx\\ &= \int_0^{+\infty}\lambda e^{-(\lambda+\theta)x}dx=\frac{\lambda}{\lambda+\theta}\int_0^{+\infty}(\lambda+\theta)e^{-(\lambda+\theta)x}dx\\ &= \frac{\lambda}{\lambda+\theta}.\end{aligned}$$

Quindi uno stimatore consistente per ρ è $\hat\rho=\hat\lambda/(\hat\lambda+\hat\theta)$. Coi dati forniti, la stima è $\hat\rho=(100/1873.3)/(100/1873.3+100/1987.7)=0.5148$.

(d) La statistica test adeguata è la statistica test log-rapporto di verosimiglianza profilo per ipotesi composite

$$W_P^{H_0}=2\left[\sup l(\lambda,\theta)-\sup_{H_0} l(\lambda,\theta)\right].$$

Essa ha distribuzione asintotica χ^2_1 sotto H_0. Ora,

$$\sup l(\lambda,\theta)=l(\hat\lambda,\hat\theta)=2n\log n-n\log(\sum_i x_i)-n\log(\sum_i y_i)-2n.$$

D'altra parte, sotto H_0, le $2n$ osservazioni possono ritenersi come costituenti un unico campione casuale semplice da una esponenziale di parametro $\lambda(=\theta)$. Quindi, sotto H_0, $\hat\lambda=\hat\theta=2n/(\sum_i x_i+\sum_i y_i)$ e

$$\sup_{H_0} l(\lambda,\theta)=2n\log 2n-2n\log(\sum_i x_i+\sum_i y_i)-2n.$$

Pertanto, risulta

$$W_P^{H_0}=2n\left[2\log(\sum_i x_i+\sum_i y_i)-\log(\sum_i x_i)-\log(\sum_i y_i)-2\log 2\right].$$

Con i dati forniti si ha $W_P^{H_0}=0.17566$. Poiché tale valore è inferiore a 3.84 (percentile di ordine 0.95 della distribuzione χ^2_1), l'ipotesi nulla non può essere rifiutata, al livello approssimato del 5%.

23. Sia $x_1,x_2,\dots,x_n$ un campione casuale semplice da una variabile X per la quale si considera la classe parametrica $\mathcal{F}=\{f(x;\theta),\ \theta>0\}$, con

$$f(x;\theta)=x^{-3/2}\theta^{1/2}(2\pi)^{-1/2}\exp\{-(\theta/2)(x-2+x^{-1})\},\qquad x>0.$$

(a) Si individui una statistica sufficiente minimale.

(b) Si ottenga la funzione di stima che definisce lo stimatore di massima verosimiglianza $\hat\theta$ per θ, stabilendo se esso è robusto al modello parametrico considerato.

(c) Si stabilisca se è possibile reperire uno stimatore ottimo per $1/\theta$.

(d) Si fornisca la regione critica più potente, di livello approssimato 0.05, per $H_0 : \theta \leq 1$ contro $H_1 : \theta > 1$.

Soluzione

(a) Si ha

$$\begin{aligned} f(x;\theta) &= x^{-3/2}\theta^{1/2}(2\pi)^{-1/2}\exp\{-(\theta/2)(x-2+x^{-1})\} \\ &= x^{-3/2}(2\pi)^{-1/2}\exp\{-(\theta/2)(x-2+x^{-1})+(1/2)\log\theta\}. \end{aligned}$$

Pertanto, la classe $\mathcal{F}$ costituisce una famiglia esponenziale regolare monoparametrica. Anche il modello per il campione casuale semplice è famiglia esponenziale regolare, con statistica canonica

$$T = \sum_{i=1}^{n}(x_i - 2 + x_i^{-1}) = \sum_{i=1}^{n}\frac{(x_i-1)^2}{x_i}.$$

T è statistica sufficiente minimale e completa.

(b) Per la funzione di verosimiglianza vale l'espressione

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) \propto \exp\left\{-(\theta/2)T + (n/2)\log\theta\right\},$$

cosicché

$$l(\theta) = -(\theta/2)T + (n/2)\log\theta$$

e

$$l_*(\theta) = -T/2 + n/(2\theta) = \frac{1}{2}\sum_{i=1}^{n}\left[\frac{1}{\theta} - \frac{(x_i-1)^2}{x_i}\right].$$

Essendo la funzione $g(x;\theta) = \frac{1}{\theta} - \frac{(x-1)^2}{x}$ non limitata (in x), lo stimatore di massima verosimiglianza $\hat{\theta}$ non è robusto al modello considerato.

(c) Lo stimatore ottimo (cioè a varianza minima tra i non distorti) per $1/\theta$ deve essere necessariamente funzione della statistica sufficiente minimale completa T. Ora, dalla proprietà di non distorsione dello score di verosimiglianza, si ha $-E(T) + n/\theta = 0$, da cui si ricava che T/n è stimatore ottimo per $1/\theta$.

(d) La regione critica cercata è basata sulla statistica T. Inoltre, $\hat{\theta} = n/T$ è stimatore di massima verosimiglianza per θ. Dato che $l_{**}(\theta) = -n/(2\theta^2)$, si ha $i(\theta) = n/(2\theta^2)$ e

$$\sqrt{n}\left(\frac{\hat{\theta}-1}{\sqrt{2}}\right) \sim N(0,1)$$

per n grande e sotto H_0. Quindi, H_0 è rifiutata (ad un livello approssimato del 5%) se $\sqrt{n}(\hat{\theta}-1) > k\sqrt{2}$, con k tale che $\Phi(k) = 0.95$, dove $\Phi(\cdot)$ indica la funzione di ripartizione della normale standard.

24. In un esperimento viene misurata la concentrazione (espressa in grammi per decilitro) di tiotimolina[11] nel sangue di $n = 10$ soggetti sani e $m = 5$ soggetti colpiti dal virus di Hueste[12]. Siano X e Y le variabili che descrivono i livelli di tiotimolina, rispettivamente, nei soggetti sani e contagiati e $x_1, x_2, \ldots, x_n$ e $y_1, y_2, \ldots, y_m$ i valori rilevati nei due gruppi. Si assume che le variabili (indipendenti) X e Y siano normali con medie μ_x e μ_y, rispettivamente, e varianza nota $\sigma^2 = 2$. Si assume, inoltre, che, a priori, μ_x e μ_y, siano indipendenti con identica distribuzione normale di media $\mu_0 = 16.5$ e varianza $\sigma_0^2 = 1$. L'osservazione campionaria è tale che $\bar{x} = 15.9$ e $\bar{y} = 18.4$ (sono le medie campionarie).

(a) Si determini la distribuzione a posteriori di (μ_x, μ_y).

(b) Si individui una regione di credibilità HPD, di livello 0.9, per (μ_x, μ_y).

(c) Si verifichi l'ipotesi $\mu_y > \mu_x$.

(d) Si fornisca una stima bayesiana della probabilità che, per un individuo contagiato, la concentrazione di tiotimolina nel sangue risulti minore di 17 g/dl.

Soluzione

(a) Posto $\boldsymbol{\mu} = (\mu_x, \mu_y)$, per la funzione di verosimiglianza vale la relazione

$$\begin{aligned} L(\boldsymbol{\mu}; \boldsymbol{x}, \boldsymbol{y}) &\propto \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{m}(y_i - \mu_y)^2\right)\right\} \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{j=1}^{n}(x_j - \mu_x)^2\right)\right\} \\ &= L_y(\mu_y; \boldsymbol{y}) L_x(\mu_x; \boldsymbol{x}), \end{aligned}$$

dove L_y e L_x sono le verosimiglianze (gaussiane) generate dai due distinti campioni. Pertanto, la distribuzione a posteriori è

$$\pi(\boldsymbol{\mu}|\boldsymbol{x}, \boldsymbol{y}) \propto L(\boldsymbol{\mu}; \boldsymbol{x}, \boldsymbol{y})\pi(\mu_x)\pi(\mu_y) = L_y(\mu_y; \boldsymbol{y})\pi(\mu_y)L_x(\mu_x; \boldsymbol{x})\pi(\mu_x),$$

cioè il prodotto delle distribuzioni a posteriori di μ_y e μ_x, che si ottengono separatamente dai due campioni gaussiani $y_1, \ldots, y_m$ e $x_1, \ldots, x_n$. Per tali distribuzioni a posteriori valgono le formule note. In particolare, per μ_y si ottiene, a posteriori, una distribuzione normale con varianza

$$\sigma_y^{*2} = (m/\sigma^2 + 1/\sigma_0^2)^{-1} = (5/2 + 1/1)^{-1} = 0.2857$$

e media

$$\mu_y^* = (m\bar{y}/\sigma^2 + \mu_0/\sigma_0^2)/(m/\sigma^2 + 1/\sigma_0^2) = (18.4 \cdot 5/2 + 16.5/1)/(5/2 + 1/1) = 17.86.$$

In definitiva, la distribuzione a posteriori è una normale bivariata con vettore delle medie $\boldsymbol{\mu}^* = (\mu_x^*, \mu_y*) = (16, 17.86)$ e matrice di varianza e covarianza

$$\Sigma^* = \begin{bmatrix} \sigma_x^{2*} & 0 \\ 0 & \sigma_y^{2*} \end{bmatrix} = \begin{bmatrix} 0.1667 & 0 \\ 0 & 0.2857 \end{bmatrix}.$$

[11] Asimov, I. (1953) Le applicazioni micropsichiatriche della tiotimolina (The Mycropsychiatric Applications of Thiotimoline, Astounding)

[12] Maine, C.E. (1962) The Darkest of Nights

(b) La regione di credibilità HPD è costituita da tutti i punti $\boldsymbol{\mu} = (\mu_x, \mu_y)$ tali che

$$k < \pi(\boldsymbol{\mu}|\boldsymbol{x}, \boldsymbol{y}) = (2\pi)^{-(n+m)/2}|\Sigma^*|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}^*)^T\Sigma^{*-1}(\boldsymbol{\mu} - \boldsymbol{\mu}^*)\right\},$$

ovvero, tali che

$$(\boldsymbol{\mu} - \boldsymbol{\mu}^*)^T\Sigma^{*-1}(\boldsymbol{\mu} - \boldsymbol{\mu}^*) < k'.$$

Dato che Σ^* è diagonale, la regione HPD cercata è costituita da tutti i punti $\boldsymbol{\mu} = (\mu_x, \mu_y)$ tali che

$$\frac{(\mu_y - \mu_y^*)^2}{\sigma_y^{2*}} + \frac{(\mu_x - \mu_x^*)^2}{\sigma_x^{2*}} < k'.$$

La quantità a sinistra è (a posteriori) somma dei quadrati di due variabili casuali normali standardizzate indipendenti ed ha perciò legge χ^2_2. La regione HPD di livello 0.9 cercata è perciò determinata ponendo nell'espressione di sopra $k' = 4.61$ (che è il percentile di ordine 0.9 della distribuzione χ^2_2). Si tratta di un'ellisse.

(c) Essendo a posteriori μ_y e μ_x indipendenti, si ha $(\mu_y - \mu_x|\boldsymbol{x}, \boldsymbol{y}) \sim N(\mu_y^* - \mu_x^*, \sigma_y^{*2} + \sigma_x^{*2})$ e quindi

$$\begin{aligned}\Pr\{\mu_y > \mu_x|\boldsymbol{x}, \boldsymbol{y}\} &= \Pr\{\mu_y - \mu_x > 0|\boldsymbol{x}, \boldsymbol{y}\} \\ &= 1 - \Phi\left(\frac{0 - (\mu_y^* - \mu_x^*)}{\sqrt{\sigma_x^{*2} + \sigma_y^{*2}}}\right) = \\ &= 1 - \Phi\left(\frac{0 - (17.86 - 16)}{\sqrt{0.2857 + 0.1667}}\right) = 0.9972\end{aligned}$$

L'ipotesi è dunque accettata.

(d) Utilizzando il risultato di cui alla nota 7, pag. xxviii, $Y \sim N(\mu_y^*, \sigma^2 + \sigma_y^{*2})$, cioè $Y \sim N(17.86, 2.2857)$. Quindi la stima cercata è pari a $\Pr\{Y < 17\} = \Phi\left(\frac{17-17.86}{\sqrt{2.2857}}\right) = 0.2847$. In alternativa si può fornire una stima ragionevole (stima *plug-in*) data da $\Phi\left(\frac{17-\mu_y^*}{\sqrt{2}}\right) = 0.2715$. Essa si ricava considerando che, condizionatamente al valore di μ_y, $\Pr\{Y < 17\} = \Phi\left(\frac{17-\mu_y}{\sqrt{2}}\right)$. In tale espressione si sostituisce quindi a μ_y la sua stima bayesiana μ_y^*.

25. Sia Y una variabile casuale discreta con supporto $\{1, 2, 3\}$ e distribuzione appartenente alla famiglia caratterizzata dalla legge $p(y; \theta)$, con spazio parametrico $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, specificata in tabella.

	θ_1	θ_2	θ_3	θ_4
y=1	0.4	0.2	0.2	0.6
y=2	0.1	0.1	0.4	0.3
y=3	0.5	0.7	0.4	0.1

Si supponga di disporre di due realizzazioni indipendenti, y_1 e y_2, di Y.

(a) Si fornisca lo stimatore di massima verosimiglianza per θ.

(b) Si costruisca il test più potente, di livello $\alpha = 0.09$, per il problema di verifica d'ipotesi $H_0 : \theta = \theta_2$ contro $H_1 : \theta = \theta_4$.

(c) Si calcoli la potenza del test di cui al punto precedente.

(d) Si proponga un test, di livello $\alpha = 0.09$, per risolvere il problema di verifica d'ipotesi $H_0 : \theta = \theta_2$, $H_1 : \theta \neq \theta_2$.

Soluzione

(a) La funzione di verosimiglianza può essere rappresentata dalle quattro tabelle che seguono

$L(\theta_1; y_1, y_2) = p(y_1; \theta_1)p(y_2; \theta_1)$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 1$	0.16	0.04	0.20
$y_1 = 2$	0.04	0.01	0.05
$y_1 = 3$	0.20	0.05	0.25

$L(\theta_2; y_1, y_2) = p(y_1; \theta_2)p(y_2; \theta_2)$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 1$	0.04	0.02	0.14
$y_1 = 2$	0.02	0.01	0.07
$y_1 = 3$	0.14	0.07	0.49

$L(\theta_3; y_1, y_2) = p(y_1; \theta_3)p(y_2; \theta_3)$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 1$	0.04	0.08	0.08
$y_1 = 2$	0.08	0.16	0.16
$y_1 = 3$	0.08	0.16	0.16

$L(\theta_4; y_1, y_2) = p(y_1; \theta_4)p(y_2; \theta_4)$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 1$	0.36	0.18	0.06
$y_1 = 2$	0.18	0.09	0.03
$y_1 = 3$	0.06	0.03	0.01

Si ha, dunque,

$$\hat{\theta} = \begin{cases} \theta_1 & \text{se } (y_1, y_2) \in \{(1,3), (3,1)\} \\ \theta_2 & \text{se } (y_1, y_2) \in \{(3,3)\} \\ \theta_3 & \text{se } (y_1, y_2) \in \{(2,2), (2,3), (3,2)\} \\ \theta_4 & \text{se } (y_1, y_2) \in \{(1,1), (1,2), (2,1)\} \end{cases}$$

(b) Il test più potente per il sistema d'ipotesi considerato è il test del rapporto di verosimiglianza $\lambda = L(\theta_4; y_1, y_2)/L(\theta_2; y_1, y_2)$. La statistica test assume i seguenti valori

$L(\theta_4; y_1, y_2)/L(\theta_2; y_1, y_2)$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 1$	9	9	0.4286
$y_1 = 2$	9	9	0.4286
$y_1 = 3$	0.4286	0.4286	0.0204

Ne segue che la regione critica più potente di livello 0.09 cercata è $\mathcal{R} = \{(y_1, y_2) : \lambda = 9\}$, ossia $\mathcal{R} = \{(1,1), (1,2), (2,1), (2,2)\}$, essendo, sotto H_0, $\Pr\{\lambda = 9\} = 0.09$.

(c) Sotto H_1, cioè per $\theta = \theta_4$, $\Pr\{(y_1, y_2) \in \mathcal{R}\} = 0.81$. Questa è la potenza del test.

(d) Un test adeguato è il test del rapporto di verosimiglianza

$$\lambda^* = \frac{L(\theta_2; y_1, y_2)}{\max_\theta L(\theta; y_1, y_2)}.$$

La tabella che riporta i valori massimi della verosimiglianza, corrispondenti ad ogni possibile realizzazione campionaria, è la seguente:

$\max_\theta L(\theta; y_1, y_2)$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 1$	0.36	0.18	0.20
$y_1 = 2$	0.18	0.16	0.16
$y_1 = 3$	0.20	0.16	0.49

Quindi, i valori assunti dalla statistica test λ^* sono

$L(\theta_2; y_1, y_2)/[\max_\theta L(\theta; y_1, y_2)]$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 1$	0.1111	0.1111	0.7
$y_1 = 2$	0.1111	0.0625	0.4375
$y_1 = 3$	0.7	0.4375	1

Tenuto conto della distribuzione della coppia (y_1, y_2) sotto H_0, ne segue che, al livello 0.09, l'ipotesi nulla è rifiutata se $\lambda^* \leq 0.1111$, ossia se $(y_1, y_2) \in \mathcal{R} = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$.

26. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile X con distribuzione beta di parametri (α, β), con $\alpha = \theta > 0$, $\beta = \theta^2$ e funzione di densità, quindi,

$$f(x; \theta) = \frac{\Gamma(\theta + \theta^2)}{\Gamma(\theta)\Gamma(\theta^2)} x^{\theta-1}(1 - x)^{\theta^2-1}, \qquad x \in (0, 1).$$

(a) Si fornisca una statistica sufficiente minimale per l'inferenza su θ.

(b) Si ottenga uno stimatore $\hat{\theta}$ per θ, basato sul metodo dei momenti.

(c) Si stabilisca se $\hat{\theta}$ è non distorto e consistente.

(d) Si ottenga un'approssimazione per la distribuzione di $\hat{\theta}$.

Soluzione

(a) Dati due punti $\boldsymbol{x_a}$ e $\boldsymbol{x_b}$ dello spazio campionario, per il rapporto tra le verosimiglianze si ha

$$\frac{L(\boldsymbol{x_a}; \theta)}{L(\boldsymbol{x_b}; \theta)} = \frac{\prod_{i=1}^n x_{ai}^{\theta-1}(1 - x_{ai})^{\theta^2-1}}{\prod_{i=1}^n x_{bi}^{\theta-1}(1 - x_{bi})^{\theta^2-1}} = \frac{\left(\prod_{i=1}^n x_{ai}\right)^{\theta-1} \left(\prod_{i=1}^n (1 - x_{ai})\right)^{\theta^2-1}}{\left(\prod_{i=1}^n x_{bi}\right)^{\theta-1} \left(\prod_{i=1}^n (1 - x_{bi})\right)^{\theta^2-1}}.$$

Tale rapporto non dipende da θ se e solo se

$$\prod_{i=1}^n x_{ai} = \prod_{i=1}^n x_{bi} \quad \text{e} \quad \prod_{i=1}^n (1 - x_{ai}) = \prod_{i=1}^n (1 - x_{bi}).$$

Quindi una statistica sufficiente minimale per l'inferenza su θ è la coppia $(\prod_{i=1}^n x_i, \prod_{i=1}^n (1 - x_i))$.

(b) Una variabile casuale beta di parametri (α, β) ha media $\alpha/(\alpha+\beta)$. Nel nostro caso, quindi, $E(X) = 1/(1+\theta)$. Ponendo $\bar{x} = 1/(1+\theta)$ e risolvendo in θ si ottiene lo stimatore basato sul metodo dei momenti $\hat{\theta} = (1/\bar{x}) - 1$.

(c) Dalla diseguaglianza di Jensen, $E(\hat{\theta}) \neq (1/E(\bar{x})) - 1 = \theta$, essendo $E(\bar{x}) = E(X) = 1/(1+\theta)$. Quindi $\hat{\theta}$ è distorto. D'altro canto, dato che la media campionaria è stimatore consistente per $1/(1+\theta)$ e che $g(z) = 1/z - 1$ è funzione continua in z, $\hat{\theta}$ risulta stimatore consistente per θ.

(d) Per il Teorema del Limite Centrale, $\bar{x}$ ha distribuzione approssimabile con quella normale di media $E(X)$ e varianza $var(X)/n$. Nel nostro caso,

$$var(X) = \frac{\theta^3}{(\theta+\theta^2)^2(\theta+\theta^2+1)} = \frac{\theta}{(1+\theta)^4 - \theta(1+\theta)^2}.$$

D'altro canto, la funzione $g(z)$ ha derivata $-1/z^2$. Applicando quindi il metodo delta, risulta che la distribuzione di $\hat{\theta}$ è approssimabile con la distribuzione normale di media $g(E(X)) = \theta$ e varianza $(E(X)^{-4} var(X)/n) = \frac{\theta(1+\theta)^2}{n[(1+\theta)^2-\theta]}$.

27. In una riserva naturale vivono tre mandrie di bisonti, chiamate *Sand Creek*, *Powder Creek* e *Bighorn*. È noto che i bisonti dal manto chiaro sono 1 su 10 nella mandria *Sand Creek*, 1 su 20 nella mandria *Powder Creek* e 1 su 100 in quella *Bighorn*.
In un covo di bracconieri la guardia forestale ritrova 60 pelli che, si scopre, provengono da un'unica battuta di caccia e, quindi, da un'unica mandria che è importante individuare. Gli agenti ritengono che, se è $2p$ la probabilità che i capi uccisi provengano dalla mandria *Sand Creek*, è p sia la probabilità che provengano dalla *Powder Creek* che quella che provengano dalla *Bighorn*. Inoltre, l'esame delle 60 pelli porta a rilevare che due dei bisonti uccisi erano bisonti dal manto chiaro.
Sulla base di queste informazioni si vuole impostare in termini bayesiani il problema dell'individuazione della mandria da cui provengono i capi uccisi, tra le tre alternative A: *Sand Creek*, B: *Powder Creek*, C: *Bighorn*.

(a) Si individui una distribuzione a priori che tenga conto delle informazioni di cui si dispone.

(b) Tenuto conto di quanto osservato, si individui la funzione di verosimiglianza e quindi la distribuzione a posteriori.

(c) Si dia una stima puntuale bayesiana.

(d) Si verifichi l'ipotesi: 'i capi uccisi provengono dalla mandria *Sand Creek*'.

Soluzione

(a) I possibili stati di natura sono le tre mandrie A, B, C. Assegnare una distribuzione a priori significa perciò assegnare una probabilità alla partizione $\{A, B, C\}$. Usiamo la notazione $P(E)$ per indicare la probabilità di un evento E; cioè, $P(E) = \Pr\{E\}$. Le indicazioni fornite prevedono che $P(A) = 2p$, $P(B) = P(C) = p$. Dovendo essere $P(A) + P(B) + P(C) = 1$, risulta $p = 1/4$ e quindi $P(A) = 0.5$, $P(B) = 0.25$ e $P(C) = 0.25$.

(b) Si sono osservati due capi dal manto chiaro su 60. Condizionatamente alla mandria di provenienza, il numero di capi dal manto chiaro su 60 capi selezionati casualmente è descritto da una variabile casuale binomiale di parametri 60 e $\pi_A = 0.1$, $\pi_B = 0.5$ o $\pi_C = 0.1$, rispettivamente per le mandrie A, B, C. Si ha cioè

$$\begin{aligned} P(E|A) &= \binom{60}{2}\pi_A^2(1-\pi_A)^{58} \propto 0.1^2 0.9^{58} = 2.218531 \times 10^{-5} \\ P(E|B) &= \binom{60}{2}\pi_B^2(1-\pi_B)^{58} \propto 0.05^2 0.95^{58} = 0.0001276172 \\ P(E|C) &= \binom{60}{2}\pi_C^2(1-\pi_C)^{58} \propto 0.01^2 0.99^{58} = 5.582661 \times 10^{-5}, \end{aligned}$$

dove E indica l'evento osservato. La distribuzione a posteriori risulta dunque da

$$\begin{aligned} P(A|E) &\propto q_A = P(A)P(E|A) = 0.5 \times 2.218531 \times 10^{-5} = 1.109266 \times 10^{-5} \\ P(B|E) &\propto q_B = P(B)P(E|B) = 0.25 \times 0.0001276172 = 3.190429 \times 10^{-5} \\ P(C|E) &\propto q_C = P(C)P(E|C) = 0.25 \times 5.582661 \times 10^{-5} = 1.395665 \times 10^{-5}, \end{aligned}$$

e, moltiplicando per $(q_A + q_B + q_C)^{-1} = 17558.15$,

$$P(A|E) = 0.1947665 \quad P(B|E) = 0.5601804 \quad P(C|E) = 0.2450530.$$

(c) Essendo gli stati di natura non ordinati e non numerici, l'unica sintesi della distribuzione a posteriori utilizzabile come stima puntuale bayesiana è la moda, quindi B.

(d) Ci si può riferire al sistema d'ipotesi $H_0 = A$ contro $H_1 = B \cup C$. Risulta

$$\frac{P(A|E)}{P(B \cup C|E)} = \frac{P(A|E)}{P(B|E) + P(C|E)} = \frac{0.1947665}{1 - 0.1947665} = 0.2418759$$

e si decide quindi contro l'ipotesi H_0.

28. Sia $x_1, x_2, \ldots, x_n$, con $n > 2$, un campione casuale semplice da una variabile X di Bernoulli di parametro $\theta \in (0, 1)$.

(a) Si ottenga la regione critica più potente, di livello α, per il sistema d'ipotesi $H_0 : \theta = \theta_0$ contro $H_1 : \theta > \theta_0$.

(b) Si ottenga la distribuzione della statistica $T = x_1(1 - x_2)$, mostrando, in particolare, che T è stimatore non distorto per la varianza (diciamo γ) di X.

(c) Si ottenga, utilizzando il risultato di Rao-Blackwell, uno stimatore non distorto per γ (basato sull'intera osservazione $x_1, x_2, \ldots, x_n$) con varianza non superiore a quella di T.

Soluzione Sia $S = \sum_{i=1}^n x_i$. Allora S ha distribuzione binomiale di parametri n e θ.

(a) Per il sistema d'ipotesi

$$H_0 : \theta = \theta_0 \quad \text{contro} \quad H_1 : \theta = \theta_1$$

con $\theta_1 > \theta_0$, il test più potente di livello α è il test del rapporto di verosimiglianza $\lambda = L(\theta_0)/L(\theta_1)$, che rifiuta H_0 per valori piccoli. Ora, $L(\theta) \propto \theta^s(1 - \theta)^{n-s}$, quindi

$$\frac{L(\theta_0)}{L(\theta_1)} = \frac{\theta_0^s(1-\theta_0)^{n-s}}{\theta_1^s(1-\theta_1)^{n-s}} = \left(\frac{1-\theta_0}{1-\theta_1}\right)^n \left[\frac{\theta_0(1-\theta_1)}{\theta_1(1-\theta_0)}\right]^s.$$

Poiché $\theta_1(1-\theta_0) > \theta_0(1-\theta_1)$, λ è funzione monotona decrescente di s. Quindi rifiutare H_0 per valori piccoli di λ equivale a rifiutare per valori grandi di s. Questo a prescindere dal particolare valore θ_1 fissato dall'ipotesi alternativa. Pertanto, il test più potente di livello α cercato ha regione critica $\{s : s > c_\alpha\}$, con c_α numero naturale scelto in modo che valga, approssimativamente,

$$\sum_{s=0}^{c_\alpha} \binom{n}{s} \theta_0^s (1-\theta_0)^{n-s} = 1-\alpha.$$

(b) La statistica T vale 1 se $x_1 = 1$ e $x_2 = 0$. Vale 0 altrimenti. Quindi

$$\Pr\{T=1\} = \Pr\{X_1 = 1, X_2 = 0\} = \theta(1-\theta) = \gamma.$$

Pertanto, T è una variabile casuale di Bernoulli di parametro γ ed ha, evidentemente, media γ.

(c) S è statistica sufficiente (minimale) per θ. Bisogna quindi ottenere

$$Z = E[T|S=s] = \Pr\{T=1|S=s\}.$$

Ora, $Z = 0$ se $S = 0$. Per $S > 0$ si ha

$$\begin{aligned} Z &= \Pr\{T=1|S=s\} = \Pr\{X_1=1, X_2=0|S=s\} \\ &= \frac{\Pr\{S=s, X_1=1, X_2=0\}}{\Pr\{S=s\}} = \frac{\Pr\{S=s|X_1=1, X_2=0\}\theta(1-\theta)}{\Pr\{S=s\}} \\ &= \frac{\Pr\{\sum_{i=3}^n X_i = s-1\}\theta(1-\theta)}{\Pr\{S=s\}} = \frac{\binom{n-2}{s-1}\theta^{s-1}(1-\theta)^{n-s-1}\theta(1-\theta)}{\binom{n}{s}\theta^s(1-\theta)^{n-s}} \\ &= \frac{\binom{n-2}{s-1}}{\binom{n}{s}} = \frac{s(n-s)}{n(n-1)}. \end{aligned}$$

In definitiva, lo stimatore cercato è

$$\frac{S(n-S)}{n(n-1)}.$$

29. In uno studio sull'affidabilità di un certo tipo di componenti elettronici, vengono rilevate le durate (nelle condizioni d'uso) di $n = 25$ componenti. Siano $y_1, y_2, \ldots, y_n$ i valori osservati e sia Y la variabile casuale che descrive la durata del generico componente. Si ritiene che per il logaritmo della durata, cioè per $X = \log(Y)$, sia adeguato un modello normale con parametri μ e σ^2. Sotto tale assunto, la durata media del generico componente (ossia la media di Y), diciamo τ, è esprimibile in funzione di μ e σ^2 e risulta $\tau = \exp(\mu + \sigma^2/2)$. Si indichino con $x_1, x_2, \ldots, x_n$ i logaritmi dei valori osservati.

(a) Si ottenga lo stimatore di massima verosimiglianza per τ, fornendone la distribuzione approssimata.

(b) Sapendo che l'osservazione campionaria è tale per cui $\sum_{i=1}^{25} x_i = 150.6$ e $\sum_{i=1}^{25} x_i^2 = 909.2$, si ottenga, sulla base del risultato di cui al punto (a), un intervallo di confidenza per τ di livello approssimato pari a 0.95.

Si supponga ora di voler stimare la durata mediana ρ. Nell'ipotesi parametrica formulata, risulta essere $\rho = \exp(\mu)$. Si supponga, però, che la vera legge di X sia un elemento della classe delle distribuzioni di Laplace con parametro di posizione θ e parametro di scala pari a 1 ($f_\theta(x) = (1/2)\exp(-|x-\theta|)$, $x \in \Re$, $\theta \in \Re$).

(c) Si stabilisca se, in tale situazione, lo stimatore di massima verosimiglianza per ρ risulta stimatore consistente per la durata mediana, nonostante l'errore nella specificazione del modello.

Soluzione Al campione casuale semplice $x_1, \dots, x_n$ da una normale di parametri μ e σ^2 corrispondono le stime di massima verosimiglianza $\hat{\mu} = \bar{x}$ e $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$, con $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$.

(a) Il parametro di interesse è legato ai parametri μ e σ^2 dalla relazione

$$\tau = g(\mu, \sigma^2) = \exp(\mu + \sigma^2/2).$$

Possiamo considerare la riparametrizzazione $(\mu, \sigma^2) \to (\tau, \sigma^2)$ definita da

$$\begin{aligned} \tau &= \exp(\mu + \sigma^2/2) \\ \sigma^2 &= \sigma^2, \end{aligned}$$

con trasformazione inversa

$$\begin{aligned} \mu &= \log(\tau) - \sigma^2/2 \\ \sigma^2 &= \sigma^2. \end{aligned}$$

Per la proprietà di equivarianza, la stima di massima verosimiglianza è pertanto

$$\hat{\tau} = \exp(\hat{\mu} + \hat{\sigma}^2/2).$$

Dato che $(\hat{\mu}, \hat{\sigma}^2)$ è asintoticamente normale,

$$((\hat{\mu}, \hat{\sigma}^2)^T - (\mu, \sigma^2)^T) \dot{\sim} N_2(0, I(\mu, \sigma^2)^{-1}),$$

con

$$I_{\mu,\sigma^2}(\mu, \sigma^2) = \begin{bmatrix} n/\sigma^2 & 0 \\ 0 & n/(2\sigma^4) \end{bmatrix}$$

(Azzalini p.82) e che $\hat{\tau}$ è funzione regolare di $(\hat{\mu}, \hat{\sigma}^2)$, si può usare il metodo delta per ottenere la distribuzione asintotica di $\hat{\tau}$. Posto

$$\Delta = \begin{pmatrix} \tau'_\mu \\ \tau'_{\sigma^2} \end{pmatrix} = \begin{pmatrix} \partial g(\mu, \sigma^2)/\partial\mu \\ \partial g(\mu, \sigma^2)/\partial\sigma^2 \end{pmatrix} = \begin{pmatrix} \exp(\mu + \sigma^2/2) \\ (1/2)\exp(\mu + \sigma^2/2) \end{pmatrix},$$

si ha allora, approssimativamente,

$$(\hat{\tau} - \tau) \sim N(0, \Delta^T I(\mu, \sigma^2)^{-1}\Delta).$$

Quindi $\hat{\tau}$ ha distribuzione approssimabile con quella normale di media τ e varianza

$$var(\hat{\tau}) = (\sigma^2/n)\exp(2\mu + \sigma^2) + (\sigma^4/(2n))\exp(2\mu + \sigma^2) = (\tau^2\sigma^2/n)(1 + \sigma^2/2).$$

(b) Sulla base dell'osservazione campionaria, si ha $\hat{\mu} = 6.024$ e $\hat{\sigma}^2 = 0.0794$. Quindi risulta $\hat{\tau} = 429.96$ e $v\hat{a}r(\hat{\tau}) = (\hat{\tau}^2\hat{\sigma}^2/n)(1 + \hat{\sigma}^2/2) = 610.658$. L'intervallo di confidenza al 95% per τ ha estremi

$$\hat{\tau} \pm 1.96\sqrt{v\hat{a}r(\hat{\tau})}.$$

Sostituendo i valori si ottiene l'intervallo $[381.5, 478.4]$.

(c) Nell'ipotesi parametrica considerata, lo stimatore di massima verosimiglianza per ρ è $\hat{\rho} = \exp(\hat{\mu}) = \exp(\bar{x})$. Nella classe di distribuzioni di Laplace, θ rappresenta la media e la mediana. Inoltre,
$$\Pr\{X < \theta\} = 0.5 = \Pr\{\exp(X) < \exp(\theta)\}.$$
Quindi $\exp(\theta)$ è la durata mediana quando X ha legge che è un elemento della classe di Laplace. Allora $\hat{\rho}$ è stimatore consistente nonostante l'errata specificazione, visto che $\bar{x}$ è stima cosistente per la media θ e $\exp(\cdot)$ è funzione continua.

30. Nelle acque antistanti un tratto di costa hawaiana capita con una certa frequenza che i sub entrino in contatto visivo con gli squali. Nella tabella sottostante sono riportati, per cinque anni, il numero di avvistamenti di squali in quelle acque e quante volte ad un avvistamento è seguito un attacco da parte di uno squalo.

	2002	2003	2004	2005	2006
numero di avvistamenti	8	9	9	6	10
numero di attacchi	5	3	5	3	5

Si supponga che la probabilità p che un avvistamento sia seguito da un attacco sia costante nel tempo e che gli avvistamenti siano indipendenti tra loro.

(a) Assumendo per p una distribuzione a priori uniforme, se ne trovi la distribuzione a posteriori sulla base dell'osservazione campionaria.

(b) Si dia una valutazione puntuale bayesiana di p. Si stabilisca, inoltre, qual è la forma degli intervalli di credibilità HPD che si ottengono per p.

(c) In base ad uno studio condotto a livello mondiale, si ritiene che la probabilità che l'"incontro" tra un sub e uno squalo dia luogo ad un attacco da parte del predatore sia pari a 0.45. Si verifichi l'ipotesi $H_0 : p = 0.45$ (contro $H_1 : p \neq 0.45$), assumendo un'a priori che assegni probabilità 0.5 a $p = 0.45$ e sia uniformemente distribuita altrove.

Soluzione

(a) Indicando con n_i e x_i, rispettivamente, il numero di avvistamenti e il numero di attacchi nell'anno i-esimo, si ha $X_i|p \sim Binom(n_i, p)$. Si ha, allora,
$$\pi(p|\boldsymbol{x}) \propto L(p; \boldsymbol{x})\pi(p) \propto \left[\prod_{i=1}^{5} p^{x_i}(1-p)^{n_i - x_i}\right] I_{(0,1)}(p).$$
Quindi,
$$\pi(p|\boldsymbol{x}) \propto p^s(1-p)^{n-s} I_{(0,1)}(p),$$
dove $n = \sum_{i=1}^5 n_i$ e $s = \sum_{i=1}^5 x_i$. Sicché, la distribuzione a posteriori è Beta$(1+s, 1+n-s)$, con $n = 42$ e $s = 21$.

(b) Una stima puntuale bayesiana per p è la media a posteriori, ovvero
$$(1+s)/(n+2) = 22/44 = 0.5\ .$$

Inoltre, poiché la funzione $\pi(p|\boldsymbol{x}) \propto p^{21}(1-p)^{21}$ è unimodale ($\arg\max_p p(1-p) = 0.5$) e simmetrica attorno a $p = 0.5$, gli intervalli di credibilità HPD sono necessariamente del tipo $(0.5 - c, 0.5 + c)$, con $c \in (0, 0.5)$ reale opportuno, tale da garantire il livello di credibilità richiesto.

(c) Posto $p_0 = 0.45$, si ha

$$\pi(p_0|\boldsymbol{x}) = \frac{\pi(p_0)p_0^s(1-p_0)^{n-s}}{\pi(p_0)p_0^s(1-p_0)^{n-s} + \int_{\{p \neq p_0\}} \pi(p)p^s(1-p)^{n-s}dp}$$

e,

$$\frac{\pi(p_0|\boldsymbol{x})}{1-\pi(p_0|\boldsymbol{x})} = \frac{0.5p_0^s(1-p_0)^{n-s}}{\int_0^1 0.5p^s(1-p)^{n-s}dp} = \frac{p_0^s(1-p_0)^{n-s}\Gamma(n+2)}{\Gamma(s+1)\Gamma(n-s+1)} = \frac{p_0^s(1-p_0)^{n-s}(n+1)!}{s!(n-s)!}.$$

Risulta $p_0^s(1-p_0)^{n-s} = 1.8411080E - 013$ e $\frac{(n+1)!}{s!(n-s)!} = 2.3145089E + 013$. Pertanto,

$$\frac{\pi(p_0|\boldsymbol{x})}{1-\pi(p_0|\boldsymbol{x})} = 4.26$$

e l'ipotesi H_0 è accettata.

31. Si suppoga di disporre di un'unica osservazione x dalla variabile casuale X, discreta, la cui distribuzione appartiene alla famiglia caratterizzata dalla legge $f(x;\theta)$, con spazio parametrico $\Theta = \{\theta_1, \theta_2, \theta_3\}$, specificata in tabella.

x	1	2	3	5	6	7
$f(x;\theta_1)$	0.02	0.03	0.04	0.02	0.03	0.86
$f(x;\theta_2)$	0.07	0.08	0.02	0.05	0.1	0.68
$f(x;\theta_3)$	0.2	0.05	0.03	0.15	0.54	0.03

(a) Si ottenga lo stimatore di massima verosimiglianza per θ e se ne fornisca la distribuzione.

(b) Si costruisca il test più potente, di livello $\alpha = 0.05$, per il problema di verifica d'ipotesi $H_0 : \theta = \theta_1$ contro $H_1 : \theta = \theta_2$.

(c) Si calcoli la potenza del test di cui al punto precedente.

(d) Si proponga un test, di livello $\alpha = 0.05$, per risolvere il problema di verifica d'ipotesi $H_0 : \theta = \theta_1$ contro $H_1 : \theta \neq \theta_1$.

Soluzione

(a) Si ha

$$\hat{\theta} = \begin{cases} \theta_1 & x = 3, 7 \\ \theta_2 & x = 2 \\ \theta_3 & x = 1, 5, 6 \end{cases}$$

La distribuzione dello stimatore è data da

$$\Pr\{\hat{\theta} = \theta_i; \theta_j\} = \begin{cases} \Pr\{X \in \{3,7\}; \theta_j\} & i = 1 \\ \Pr\{X = 2; \theta_j\} & i = 2 \\ \Pr\{X \in \{1,5,6\}; \theta_j\} & i = 3 \end{cases}$$

per $j = 1, 2, 3$, ed è riportata nella tabella seguente:

	$\hat{\theta} = \theta_1$	$\hat{\theta} = \theta_2$	$\hat{\theta} = \theta_3$
θ_1	0.90	0.03	0.07
θ_2	0.70	0.08	0.22
θ_3	0.06	0.05	0.89

(b) In base al lemma di Neymann-Pearson il test più potente ha regione di rifiuto di forma $f(x;\theta_2) > k_\alpha f(x;\theta_1)$, con k_α valore di soglia opportuno. Si calcola dunque il rapporto $f(x;\theta_2)/f(x;\theta_1)$ ottenendo

x	1	2	3	5	6	7
$f(x;\theta_1)$	0.02	0.03	0.04	0.02	0.03	0.86
$f(x;\theta_2)$	0.07	0.08	0.02	0.05	0.1	0.68
$f(x;\theta_2)/f(x;\theta_1)$	3.5	2.67	0.5	2.5	3.33	0.79

La soglia k_α deve essere tale per cui la probabilità che il rapporto sia maggiore di k_α nell'ipotesi H_0 sia (eventualmente approssimativamente) 0.05 . È immediato verificare che se $k_\alpha = 2.67$ tale condizione è verificata.

(c) Usando la tabella sopra riportata, la probabilità che il rapporto $f(x;\theta_2)/f(x;\theta_1)$ sia maggiore di k_α (con $k_\alpha = 2.67$) nell'ipotesi H_1 è pari a $\Pr\{X \in \{6,1\}; \theta_2\} = 0.17$.

(d) Un test appropriato è il test del rapporto di verosimiglianza.

x	1	2	3	5	6	7
$f(x;\theta_1)$	0.02	0.03	0.04	0.02	0.03	0.86
$f(x;\theta_2)$	0.07	0.08	0.02	0.05	0.1	0.68
$f(x;\theta_3)$	0.2	0.05	0.03	0.15	0.54	0.03
$f(x;\hat{\theta})/f(x;\theta_1)$	10	2.67	1	7.5	18	1

Si rifiuta per valori grandi di $f(x;\hat{\theta})/f(x;\theta_1)$. Quindi, per $\alpha = 0.05$, il test rifiuta H_0 se $f(x;\hat{\theta})/f(x;\theta_1) > 7.5$.

32. Si ritiene che la durata di vita (espressa in minuti) di certi insetti esposti ad un insetticida sia ben descritta da una variabile casuale X con funzione di densità $f(x;\theta) = \theta(1+x)^{-(1+\theta)}$, per $x > 0$, con $\theta > 0$ parametro ignoto. Si supponga di disporre di un campione casuale semplice $x_1, x_2, \ldots, x_n$ relativo alle durate di vita rilevate, in un esperimento, su n insetti.

(a) Si consideri la funzione di stima $q(\theta;\boldsymbol{x}) = \sum_{i=1}^n \{g(\theta;x_i) - a(\theta)\}$ con

$$g(\theta;x_i) = \begin{cases} 2e^{-1/\theta} & \text{se } x_i > 1 \\ (1+x_i)e^{-1/\theta} & \text{se } 0 < x_i \le 1 \\ e^{-1/\theta} & \text{se } x_i \le 0 \end{cases}$$

dove $a(\theta)$ è un'opportuna funzione di θ. Si ottenga l'espressione per la funzione $a(\cdot)$ che rende la funzione di stima $q(\cdot;\cdot)$ non distorta al modello $\mathcal{F} = \{f(x;\theta),\ \theta > 0\}$.

(b) Si confrontino, in termini di robustezza al modello $\mathcal{F}$, lo stimatore di massima verosimiglianza $\hat{\theta}$ per θ e lo stimatore definito dalla funzione di stima non distorta $q(\theta; \boldsymbol{x})$ individuata al punto (a).

(c) Si stabilisca se esiste (e si indichi, eventualmente, quale è) una funzione del parametro θ per la quale è possibile reperire lo stimatore ottimo (secondo la teoria classica).

(d) Supponendo $n = 300$ e assumendo che l'osservazione campionaria sia tale che $\sum_{i=1}^{300} \log(1 + x_i) = 259.73$, si ottenga la stima di massima verosimiglianza per la probabilità ρ che la durata di vita di un insetto esposto all'insetticida sia superiore a 10 minuti.

(e) Sempre con $n = 300$, come cambia la stima di ρ se si suppone che l'informazione campionaria sia: "solo 7 insetti sopravvivono dopo 20 minuti di esposizione" ?

Soluzione

(a) Deve risultare $E_\theta[q(\theta; \boldsymbol{x})] = 0$, $\forall\theta$. Ciò equivale a chiedere che sia $E_\theta[g(\theta; x)] = a(\theta)$, $\forall\theta$. Quindi, deve essere

$$\begin{aligned}
a(\theta) &= \int_0^1 (1+x)e^{-1/\theta}\theta(1+x)^{-(1+\theta)}dx + \int_1^{+\infty} 2e^{-1/\theta}\theta(1+x)^{-(1+\theta)}dx \\
&= \theta e^{-1/\theta}\int_1^2 t^{-\theta}dt + 2\theta e^{-1/\theta}\int_2^{+\infty} t^{-(1+\theta)}dt \\
&= \theta e^{-1/\theta}\left[\frac{t^{-\theta+1}}{-\theta+1}\right]_1^2 + 2\theta e^{-1/\theta}\left[\frac{t^{-\theta}}{-\theta}\right]_2^{+\infty} \\
&= \frac{\theta e^{-1/\theta}}{1-\theta}(2^{1-\theta}-1) + 2^{1-\theta}e^{-1/\theta} \\
&= \frac{e^{-1/\theta}}{1-\theta}(2^{1-\theta}-\theta).
\end{aligned}$$

avendo operato la sostituzione $t = x + 1$.

(b) Essendo la funzione di stima $q(\theta; \boldsymbol{x})$ limitata (in $\boldsymbol{x}$), lo stimatore da essa definito è robusto al modello considerato. D'altro canto, si ha

$$L(\theta) = \prod_{i=1}^n \theta(1+x_i)^{-(1+\theta)} = \theta^n \prod_{i=1}^n (1+x_i)^{-(1+\theta)}$$

e

$$l(\theta) = n\log\theta - (1+\theta)\sum_{i=1}^n \log(1+x_i).$$

Pertanto, lo score di verosimiglianza vale $l_*(\theta) = n/\theta - \sum_{i=1}^n \log(1+x_i)$ ed è funzione non limitata. Ne segue che lo stimatore $\hat{\theta}$ non è robusto a $\mathcal{F}$.

(c) Si può scrivere

$$f(x;\theta) = \exp\{\log\theta - (1+\theta)\log(1+x)\}.$$

Quindi $f(x;\theta)$ costituisce una famiglia esponenziale regolare monoparametrica. Sotto campionamento casuale semplice, con dimensione campionaria n, la statistica canonica è $T = \sum_{i=1}^n \log(1+x_i)$, che è statistica sufficiente minimale completa. Inoltre,

$$E(T) = n\frac{\frac{d}{d\theta}(-\log\theta)}{\frac{d}{d\theta}(-1-\theta)} = n/\theta.$$

Quindi lo stimatore a varianza minima tra i non distorti esiste per il parametro $\tau = 1/\theta$.

(d) Si ha

$$\Pr\{X > k\} = \int_k^{+\infty} \theta(1+x)^{-(1+\theta)}dx = \int_{k+1}^{+\infty} t^{-(1+\theta)}dt = (k+1)^{-\theta}.$$

Essendo $\hat{\theta} = n/\sum_{i=1}^n \log(1+x_i) = 300/259.73 = 1.155$, risulta che la stima di massima verosimiglianza per $\rho = 11^{-\theta}$ è $\hat{\rho} = 11^{-\hat{\theta}} = 0.0626$.

(e) La variabile casuale, diciamo M, di cui è realizzazione il numero di insetti che sopravvivono dopo 20 minuti di esposizione all'insetticida, ha distribuzione binomiale di parametri $n = 300$ e $p = \Pr\{X > 20\} = 21^{-\theta}$. Quindi, la verosimiglianza associata alla sola informazione relativa al numero m di insetti che sopravvivono dopo 20 minuti di esposizione è

$$L(\theta) \propto (21^{-\theta})^m(1-21^{-\theta})^{n-m}.$$

È noto che, per il modello binomiale, la stima di massima verosimiglianza per p è la frazione di "successi" osservati, ossia $\hat{p} = m/n$. Allora, la stima di massima verosimiglianza per θ è quel valore per il quale si verifica l'uguaglianza $21^{-\theta} = m/n$. Con $n = 300$ e $m = 7$, risulta $\hat{\theta} = -\frac{\log(m/n)}{\log 21} = 1.2343$. Di coseguenza, $\hat{\rho} = 11^{-\hat{\theta}} = 0.0518$

33. Un produttore di apparecchi elettromedicali è interessato a valutare il livello delle emissioni elettromagnetiche dei suoi prodotti, che sono assemblati utilizzando materiali provenienti da due fornitori distinti, diciamo A e B. In un esperimento, su un campione di n apparecchi prodotti viene misurato il livello delle emissioni in condizioni di uso standard. Detta Y la variabile che descrive il livello delle emissioni, si assume che essa abbia distribuzione esponenziale di parametro (reciproco della media) θ o $k\theta$ (con $\theta > 0$ ignoto e $k > 0$ costante nota) a seconda che l'apparecchio in prova risulti dall'assemblaggio di materiali del fornitore A o B, rispettivamente. L'osservazione campionaria è costituita dalle coppie (d_i, y_i), $i = 1, \ldots, n$, dove $d_i = 1$ se l'apparecchio i-esimo è assemblato con materiali forniti da B. Nel caso contrario, $d_i = 0$.

(a) Si scriva la funzione di verosimiglianza per θ e si individui la famiglia di a priori coniugata naturale alla verosimiglianza.

Si assuma che l'osservazione sia $\{(1,4),(0,2),(0,6),(1,9)(1,10)\}$, che sia $k = 1/3$ e si scelga di utilizzare, come a priori, l'elemento della famiglia coniugata con media 0.1 e varianza 0.3.

(b) Si fornisca una stima bayesiana del livello medio delle emissioni di un apparecchio assemblato con materiali forniti da A.

(c) Si fornisca una stima bayesiana della probabilità che il livello delle emissioni di un apparecchio assemblato con materiali forniti da B sia maggiore di 14.

(d) Utilizzando la tabella che segue, riportante alcuni valori della funzione di ripartizione a posteriori, si fornisca un intervallo di credibilità al 90% per θ.

θ	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19
$F(\theta\|\mathbf{d},\mathbf{y})$	0.023	0.032	0.044	0.058	0.075	0.093	0.114	0.136	0.161	0.187

θ	0.50	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59
$F(\theta\|\mathbf{d},\mathbf{y})$	0.898	0.907	0.915	0.923	0.930	0.936	0.942	0.947	0.952	0.957

Soluzione

(a) Il contributo alla verosimiglianza dell'i-ma osservazione è

$$L_i(\theta) \propto \begin{cases} \theta e^{-y_i\theta} & \text{se } d_i = 0, \\ (k\theta)e^{-(k\theta)y_i} & \text{se } d_i = 1 \end{cases}$$

cioè

$$L_i(\theta) \propto \theta k^{d_i} e^{-\theta k^{d_i} y_i}.$$

Pertanto, per la funzione di verosimiglianza si ha

$$L(\theta) \propto \prod_{i=1}^{n} \left(\theta k^{d_i} \exp\left\{-\theta k^{d_i} y_i\right\}\right) \propto \theta^n \exp\left\{-\theta \sum_{i=1}^{n} k^{d_i} y_i\right\}.$$

È evidente allora che la famiglia coniugata alla verosimiglianza è la famiglia di distribuzioni gamma. Infatti, presa la densità di una la $Gamma(a, b)$ come a priori, si ottiene come densità a posteriori

$$\pi(\theta|\mathbf{d}, \mathbf{y}) \propto \theta^{a+n-1} \exp\left\{-\theta(b + \sum_i y_i k^{d_i})\right\}.$$

La distribuzione a posteriori è, quindi, una $Gamma(a' = a + n, b' = b + \sum_i y_i k^{d_i})$.

(b) Il livello medio delle emissioni di un apparecchio assemblato con materiali forniti da A è pari a $1/\theta$. La media a posteriori di $1/\theta$ è

$$\begin{aligned} E(1/\theta|\mathbf{d}, \mathbf{y}) &= \int_0^{+\infty} \frac{b'^{a'}}{\Gamma(a')}(1/\theta)\theta^{a'-1} \exp\{-\theta b'\}\, d\theta \\ &= \int_0^{+\infty} \frac{b'^{a'}}{\Gamma(a')}\theta^{a'-2} \exp\{-\theta b'\}\, d\theta \\ &= \frac{b'^{a'}}{\Gamma(a')}\frac{\Gamma(a'-1)}{(b')^{a'-1}} \\ &= b'/(a'-1). \end{aligned}$$

I parametri della distribuzione a priori si ottengono impostando il sistema

$$\begin{cases} a/b = 0.1 \\ a/b^2 = 0.3 \end{cases} \Rightarrow \begin{cases} a = 0.03 \\ b = 0.33 \end{cases}$$

La statistica sufficiente $S = \sum_i y_i k^{d_i}$ è pari a 15.66. I parametri dell'a posteriori sono quindi $a' = 5.03$ e $b' = 15.99$. La media a posteriori di $1/\theta$ risulta essere 3.97.

(c) La probabilità che il livello delle emissioni di un apparecchio assemblato con materiali forniti da B superi una soglia $y^* > 0$ è $e^{-k\theta y^*}$. Una stima bayesiana di tale probabilità à data da

$$\begin{aligned} E(e^{-k\theta y^*}|\mathbf{d}, \mathbf{y}) &= \int_0^{+\infty} \frac{b'^{a'}}{\Gamma(a')} e^{-k\theta y^*} \theta^{a'-1} \exp\{-\theta b'\}\, d\theta \\ &= \frac{b'^{a'}}{\Gamma(a')} \int_0^{+\infty} \theta^{a'-1} \exp\{-\theta(b' + ky^*)\}\, d\theta \\ &= \frac{b'^{a'}}{\Gamma(a')}\frac{\Gamma(a')}{(b' + ky^*)^{a'}} \\ &= \left(\frac{b'}{b' + ky^*}\right)^{a'}. \end{aligned}$$

Per $y^* = 14$ si ottiene il valore 0.28.

(d) Per ispezione della tabella si ricava facilmente che un intervallo di credibilità al 90% è dato da $[0.10, 0.53]$.

34. Uno studio sulla probabilità del sesso alla nascita dei figli secondogeniti coinvolge un campione di n famiglie con soli due figli. I ricercatori suppongono sia ragionevole assumere l'equiprobabiltà dei sessi alla nascita per il primogenito. Inoltre, assumono che la probabilità che un secondogenito maschio segua un primogenito dello stesso sesso sia $p_m = (1+\mu)/2$ e che la probabilità che un secondogenito femmina segua un primogenito dello stesso sesso sia $p_f = (1+\varphi)/2$, con μ e φ parametri ignoti, entrambi interni all'intervallo (-1,1). Siano x_{mm}, x_{mf}, x_{fm}, x_{ff}, rispettivamente, il numero di famiglie nel campione con due figli maschi, il numero di famiglie con il primogenito maschio e il secondogenito femmina, il numero di famiglie con il primogenito femmina e il secondogenito maschio e il numero di famiglie con due figlie femmine.

(a) In base al modello ipotizzato, si scriva la funzione di verosimiglianza per (μ, φ).

(b) Si fornisca una regione di confidenza per (μ, φ), di livello approssimato 0.95.

(c) Con $n = 3500$, assumendo di di aver osservato $x_{mm} = 900$, $x_{mf} = 700$ e $x_{fm} = 1100$, si verifichi l'ipotesi $H_0 : p_m = p_f$ ad un livello di significatività (approssimato) del 5%.

Soluzione

(a) L'osservazione $(x_{mm},\ x_{mf},\ x_{fm},\ x_{ff})$ è realizzazione di una variabile multinomiale di indice n, a 4 celle, con vettore delle probabilità, che indichiamo con

$$\pi = (\pi_{mm}, \pi_{mf}, \pi_{fm}, \pi_{mm}),$$

tale che

$$\pi_{mm} = \frac{1}{2}\left(\frac{1+\mu}{2}\right) = \frac{1+\mu}{4} \qquad \pi_{mf} = \frac{1}{2}\left(1 - \frac{1+\mu}{2}\right) = \frac{1-\mu}{4}$$
$$\pi_{fm} = \frac{1}{2}\left(1 - \frac{1+\varphi}{2}\right) = \frac{1-\varphi}{4} \qquad \pi_{ff} = \frac{1}{2}\left(\frac{1+\varphi}{2}\right) = \frac{1+\varphi}{4}.$$

Risulta pertanto

$$L(\mu, \varphi) \propto (1+\mu)^{x_{mm}}(1-\mu)^{x_{mf}}(1-\varphi)^{x_{fm}}(1+\varphi)^{x_{ff}}.$$

(b) La funzione di log-verosimglianza risulta essere

$$l(\mu, \varphi) = x_{mm}\log(1+\mu) + x_{mf}\log(1-\mu) + x_{fm}\log(1-\varphi) + x_{ff}\log(1+\varphi).$$

Derivando si ottengono gli elementi dello score:

$$l_\mu(\mu, \varphi) = \frac{x_{mm}}{1+\mu} - \frac{x_{mf}}{1-\mu} \qquad l_\varphi(\mu, \varphi) = \frac{x_{ff}}{1+\varphi} - \frac{x_{fm}}{1-\varphi}$$

e uguagliando a zero si ricavano gli stimatori di massima verosimiglianza

$$\hat{\mu} = \frac{x_{mm} - x_{mf}}{x_{mm} + x_{mf}} \qquad \hat{\varphi} = \frac{x_{ff} - x_{fm}}{x_{ff} + x_{fm}}.$$

La regione di confidenza cercata è allora l'insieme

$$\{(\mu, \varphi) :\ 2[l(\hat{\mu}, \hat{\varphi}) - l(\mu, \varphi)] \le c\},$$

con c tale che $\Pr\{\chi^2_2 > c\} = 0.05$.

(c) Sotto l'ipotesi nulla si ha $\mu = \varphi$ e quindi

$$\begin{aligned} l_{H_0}(\mu) = l(\mu, \mu) &= x_{mm}\log(1+\mu) + x_{mf}\log(1-\mu) + x_{fm}\log(1-\mu) + x_{ff}\log(1+\mu) \\ &= (x_{mm} + x_{ff})\log(1+\mu) + (x_{mf} + x_{fm})\log(1-\mu). \end{aligned}$$

Derivando rispetto a μ e risolvendo l'equazione di verosimiglianza si ottiene la stima vincolata

$$\hat{\mu}_{H_0} = \frac{(x_{mm} + x_{ff}) - (x_{mf} + x_{fm})}{x_{mm} + x_{ff} + x_{mf} + x_{fm}} = \frac{(x_{mm} + x_{ff}) - (x_{mf} + x_{fm})}{n}.$$

Con i dati a disposizione, risulta allora

$$\hat{\mu} = \frac{200}{1600} = 0.125 \qquad \hat{\varphi} = \frac{-300}{1900} = -0.1579 \qquad \hat{\mu}_{H_0} = \frac{-100}{3500} = -0.02857,$$

e $l(\hat{\mu}, \hat{\varphi}) = 36.316$, $l_{H_0}(\hat{\mu}) = 1.4287$. La statistica test adeguata per risolvere il problema di verifica d'ipotesi in questione è la statistica test del rapporto di verosimiglianza $W_P^{H_0} = 2\{l(\hat{\mu}, \hat{\varphi}) - l_{H_0}(\hat{\mu})\}$, che ha distribuzione asintotica χ^2_1 sotto H_0 (l'ipotesi nulla esprime un vincolo su μ e φ). Poiché in questo caso $W_P^{H_0} = 69.7 > 3.84$, l'ipotesi nulla è rifiutata.

35. Sia $x_1, x_2, \dots, x_n$ $(n > 2)$ un campione casuale semplice da una variabile casuale X con distribuzione uniforme sull'intervallo con estremi θ e 5θ, con $\theta > 0$.

(a) Si ottenga lo stimatore di massima verosimiglianza $\hat{\theta}$ stabilendo se è non distorto.

(b) Si mostri che $\hat{\theta}$ è consistente.

(c) Si ricavi la funzione d'influenza, al modello parametrico considerato, dello stimatore $\tilde{\theta}$ ottenuto col metodo dei momenti.

(d) Si proponga un intervallo di confidenza per θ di livello approssimato 0.95.

(e) È possibile ottenere un intervallo di confidenza per θ, di livello 1, che non sia quello banale costituito dall'insieme dei reali positivi?

Soluzione

(a) Si ha $f(x;\theta) = [1/(4\theta)]I_{[\theta,\ 5\theta]}(x)$. Quindi,

$$L(\theta; \mathbf{x}) = \frac{1}{(4\theta)^n}\prod_{i=1}^{n} I_{[\theta,\ 5\theta]}(x_i) = \frac{1}{(4\theta)^n} I_{[\theta,\ 5\theta]}(x_{(1)}) I_{[\theta,\ 5\theta]}(x_{(n)}) = \frac{1}{(4\theta)^n} I_{\left[\frac{x_{(n)}}{5},\ x_{(1)}\right]}(\theta).$$

Dato che $1/(4\theta)^n$ è funzione decrescente di θ, lo stimatore di massima verosimiglianza è $\hat{\theta} = x_{(n)}/5$. Inoltre, risulta $x_{(n)}/5 < \theta$ con probabilità 1. Quindi $\hat{\theta}$ è distorto.

(b) Sia $\epsilon > 0$ una costante arbitraria. Allora,

$$\begin{aligned}\Pr\left\{\frac{X_{(n)}}{5} > \theta - \epsilon\right\} &= \Pr\{X_{(n)} > 5(\theta-\epsilon)\} = 1 - \Pr\{X_{(n)} \leq 5(\theta-\epsilon)\} \\ &= 1 - [F_X(5(\theta-\epsilon))]^n = 1 - \left[\frac{5(\theta-\epsilon)-\theta}{4\theta}\right]^n \to 1\end{aligned}$$

quando $n \to +\infty$, con $F_X(\cdot)$ che indica la funzione di ripartizione di X. Quindi $\hat{\theta}$ è stimatore consistente.

(c) Si ha $E(X) = (\theta + 5\theta)/2 = 3\theta$ e $var(X) = (5\theta - \theta)^2/12 = 4\theta^2/3$. Lo stimatore basato sul metodo dei momenti si ottiene ponendo $(1/n)\sum_i x_i = 3\theta$ e risolvendo in θ. Risulta $\tilde{\theta} = \bar{x}/3$, con $\bar{x}$ media campionaria. La funzione di stima che definisce tale stimatore è $\sum_i (x_i - 3\theta)$ e, al modello parametrico considerato, la funzione d'influenza associata è

$$IF(x; \tilde{\theta}, F_X) = \frac{x - 3\theta}{-E_{F_X}[d(x-3\theta)/d\theta]} = x/3 - \theta.$$

(d) Dal Teorema del Limite Centrale, $\bar{x}\dot{\sim}N(3\theta, 4\theta^2/(3n))$. Quindi $\tilde{\theta}\dot{\sim}N(\theta, 4\theta^2/(27n))$ e l'intervallo con estremi $\tilde{\theta} \pm 1.96\sqrt{4\tilde{\theta}^2/(27n)}$ è un intervallo di confidenza di livello approssimato 0.95 per θ.

(e) Evidentemente, $\theta \leq X_{(1)}$ e $X_{(n)}/5 \leq \theta$, ossia $X_{(n)}/5 \leq \theta \leq X_{(1)}$, con probabilità 1. Quindi l'intervallo $[x_{(n)}/5,\ x_{(1)}]$ è un intervallo di confidenza per θ di livello 1.

36. In un esperimento biologico viene misurata la lunghezza di certi microrganismi. I valori (in μm) rilevati su 10 unità sono $0.22, 0.2, 0.18, 0.16, 0.19, 0.31, 0.2, 0.046, 0.23$, e 0.27 (con somma 2.006). Si assume che essi siano realizzazioni indipendenti ed identicamente distribuite di una variabile $N(\theta, \sigma^2)$, con varianza $\sigma^2 = 0.3$ nota. Il parametro θ rappresenta dunque la lunghezza media di questa specie di microrganismi.

Si consideri come distribuzione a priori (impropria) per θ la funzione $\pi(\theta) \propto I_{[0,+\infty)}(\theta)$.

(a) Si mostri che la distribuzione a posteriori che si ottiene è valida.

(b) Si fornisca una stima puntuale bayesiana di θ.

(c) Si stabilisca se l'intervallo [0, 0.53] è un intervallo di credibilità HPD per θ.

(d) Si ricavi la funzione di ripartizione a posteriori e si verifichi l'ipotesi $\theta > 0.22$. [13]

[13] Indicando con $\Phi(\cdot)$ la funzione di ripartizione della normale standardizzata, si ha $\Phi(-1.1581) = 0.1234$ e $\Phi(0.112006) = 0.5448$.

Soluzione

(a) Si ha

$$\begin{aligned}\pi(\theta|\boldsymbol{x}) &\propto \exp\left\{-\frac{1}{2\sigma^2}\sum_i (x_i-\theta)^2\right\} I_{[0,+\infty)}(\theta)\\ &\propto \exp\left\{-\frac{n}{2\sigma^2}(\theta^2-2\theta\bar{x})\right\} I_{[0,+\infty)}(\theta)\\ &\propto \exp\left\{-\frac{n}{2\sigma^2}(\theta-\bar{x})^2\right\} I_{[0,+\infty)}(\theta)\end{aligned}$$

Si riconosce dunque il nucleo di una $N(\bar{x},\sigma^2/n)$, moltiplicato per la funzione indicatrice. Pertanto,

$$\int_0^{+\infty} \exp\left\{-\frac{n}{2\sigma^2}(\theta-\bar{x})^2\right\} d\theta = \sqrt{\frac{2\pi\sigma^2}{n}}[1-\Phi(-\bar{x}\sqrt{n}/\sigma)] < +\infty.$$

L'espressione della densità a posteriori è dunque

$$\pi(\theta|\boldsymbol{x}) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma[1-\Phi(-\bar{x}\sqrt{n}/\sigma)]} \exp\left\{-\frac{n}{2\sigma^2}(\theta-\bar{x})^2\right\} I_{[0,+\infty)}(\theta),$$

con $\sigma^2 = 0.3$ e $\bar{x} = 0.2006$.

(b) La stima bayesiana di θ più facilmente reperibile in questo caso è la moda della distribuzione a posteriori. Data la forma della densità a posteriori (normale -con media $\bar{x}$- troncata in zero), la moda risulta essere necessariamente

$$M_o = \max\{0,\bar{x}\} = \bar{x} = 0.2006.$$

Si può ottenere anche la mediana a posteriori osservando che la funzione di ripartizione di $\theta|\boldsymbol{x}$ è

$$\begin{aligned}F(\theta|\boldsymbol{x}) &= \int_0^\theta \pi(t|\boldsymbol{x})dt\\ &= \frac{1}{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}\int_0^\theta \frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{n}{2\sigma^2}(t-\bar{x})^2\right\}dt\\ &= \frac{1}{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}\left[\Phi\left(\sqrt{n}\frac{\theta-\bar{x}}{\sigma}\right)-\Phi\left(\sqrt{n}\frac{-\bar{x}}{\sigma}\right)\right]\end{aligned}$$

per $\theta > 0$ e 0 altrove. Quindi, risolvendo l'equazione $F(\theta|\boldsymbol{x}) = 1/2$, cioè

$$\begin{aligned}\frac{1}{2} &= \frac{1}{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}\left[\Phi\left(\sqrt{n}\frac{\theta-\bar{x}}{\sigma}\right)-\Phi\left(\sqrt{n}\frac{-\bar{x}}{\sigma}\right)\right]\\ \frac{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}{2} &= \Phi\left(\sqrt{n}\frac{\theta-\bar{x}}{\sigma}\right)-\Phi\left(\sqrt{n}\frac{-\bar{x}}{\sigma}\right)\\ \frac{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}{2}+\Phi\left(\sqrt{n}\frac{-\bar{x}}{\sigma}\right) &= \Phi\left(\sqrt{n}\frac{\theta-\bar{x}}{\sigma}\right)\\ \Phi^{-1}\left(\frac{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}{2}+\Phi\left(\sqrt{n}\frac{-\bar{x}}{\sigma}\right)\right) &= \sqrt{n}\frac{\theta-\bar{x}}{\sigma}\end{aligned}$$

si ottiene

$$M_e = \bar{x} + \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(\frac{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}{2} + \Phi\left(\sqrt{n}\frac{-\bar{x}}{\sigma}\right)\right)$$

(c) Tenendo conto della forma della densità a posteriori, dato che $2\bar{x} < 0.53$, l'intervallo $[0, 0.53]$ è un intervallo di credibilità HPD.

(d) Poiché

$$F(\theta|\boldsymbol{x}) = \frac{1}{1-\Phi(-\bar{x}\sqrt{n}/\sigma)}\left[\Phi\left(\sqrt{n}\frac{\theta-\bar{x}}{\sigma}\right) - \Phi\left(\sqrt{n}\frac{-\bar{x}}{\sigma}\right)\right]$$

per $\theta > 0$, la probabilità a posteriori associata all'ipotesi considerata è

$$1 - F(0.22|\boldsymbol{x}) = 1 - \frac{1}{1-\Phi(-1.158165)}[\Phi(0.112006) - \Phi(-1.158165)] = 0.5192.$$

L'ipotesi è dunque accettata.

37. Si suppoga di disporre di un'unica osservazione x dalla variabile casuale X, discreta, la cui distribuzione appartiene alla famiglia specificata nella tabella sottostante, caratterizzata dalla legge $p(x;\theta)$, con spazio parametrico $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$.

x	1	2	3	4	5	6
$p(x;\theta_1)$	0.02	0.07	0.04	0.10	0.56	0.21
$p(x;\theta_2)$	0.09	0.03	0.18	0.45	0.16	0.09
$p(x;\theta_3)$	0.04	0.13	0.08	0.20	0.16	0.39
$p(x;\theta_4)$	0.02	0.14	0.04	0.10	0.28	0.42
$p(x;\theta_5)$	0.07	0.08	0.14	0.35	0.12	0.24

(a) Si stabilisca se la statistica

$$T(x) = \begin{cases} 10 & \text{se } x \in \{1,3\} \\ 0 & \text{se } x \in \{2,4\} \\ -10 & \text{se } x \in \{5,6\} \end{cases}$$

è statistica sufficiente minimale per l'inferenza su θ.

(b) Si costruisca il test più potente, di livello $\alpha = 0.12$, per il problema di verifica d'ipotesi $H_0 : \theta = \theta_2$ contro $H_1 : \theta = \theta_5$.

(c) Si calcoli la potenza del test di cui al punto precedente.

(d) Si stabilisca se il test di cui al punto (b) è il più potente, tra quelli di livello $\alpha = 0.12$, per il problema di verifica d'ipotesi $H_0 : \theta = \theta_2$ contro $H_1 : \theta \in \{\theta_3, \theta_4, \theta_5\}$.

Soluzione

(a) La partizione di verosimiglianza ha orbite $\{1,3,4\}$, $\{2,6\}$ e $\{5\}$. La partizione indotta dalla statistica $T(x)$ non coincide con la partizione di verosimiglianza, quindi $T(x)$ non è statistica sufficiente minimale.

(b) Il test più potente è (lemma di Neyman-Pearson) il test del rapporto di verosimiglianza

$$\lambda(x) = \frac{p(x;\theta_5)}{p(x;\theta_2)}$$

che rifiuta l'ipotesi nulla per valori grandi. In questo caso si ha

x	1	2	3	4	5	6
$\lambda(x)$	0.77777	2.66666	0.77777	0.77777	0.75000	2.66666

Poiché $\Pr\{\lambda(X) \geq 2.66666 \mid H_0\} = 0.12$, il test ottimo di livello 0.12 ha regione critica $R = \{2, 6\}$.

(c) Risulta $\Pr\{x \in R \mid H_1\} = 0.08 + 0.24 = 0.32$.

(d) Consideriamo la particolare ipotesi alternativa $\theta = \theta_3$ e quindi il sistema d'ipotesi $H_0 : \theta = \theta_2$ contro $H_1 : \theta = \theta_3$. Si ha

x	1	2	3	4	5	6
$\lambda_3(x) = \frac{p(x;\theta_3)}{p(x;\theta_2)}$	0.44444	4.33333	0.44444	0.44444	1.00000	4.33333

e $\Pr\{\lambda_3(X) \geq 4.33333 \mid H_0\} = 0.12$. Qiundi, anche in questo caso, il test ottimo di livello più $\alpha = 0.12$ ha regione critica $R = \{2, 6\}$. Se come alternativa si fissa $\theta = \theta_4$, risulta

x	1	2	3	4	5	6
$\lambda_4(x) = \frac{p(x;\theta_4)}{p(x;\theta_2)}$	0.22222	4.66666	0.22222	0.22222	1.75000	4.66666

e $\Pr\{\lambda_4(X) \geq 4.66666 \mid H_0\} = \Pr\{X \in R \mid H_0\} = 0.12$. Evidentemente, allora, la regione critica R, individuata al punto (b), è quella ottima, di livello 0.12, per il problema $H_0 : \theta = \theta_2$ contro $H_1 : \theta \in \{\theta_3, \theta_4, \theta_5\}$.

38. Di recente è nata in Facoltà una nuova famiglia parametrica, nota ad alcuni come U_p *di Ruzza*. Una variabile X con distribuzione U_p di parametri μ, δ e α ha densità

$$f(x;\mu,\delta,\alpha) = \frac{\alpha+1}{2\delta^{\alpha+1}}|x-\mu|^{\alpha} I_{[\mu-\delta,\ \mu+\delta]}(x),$$

con $\mu \in \Re$, $\delta > 0$, $\alpha \geq 0$. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da X. Si suppongano, inizialmente, $\delta = 1$ e $\alpha = 2$ noti.

(a) Si stabilisca se le funzioni di stima $q_1(\mu;\boldsymbol{x}) = \sum_{i=1}^{n} I_{(-\infty,0]}(x_i - \mu) - n/2$ e $q_2(\mu;\boldsymbol{x}) = \sum_{i=1}^{n} \Phi(x_i - \mu) - n/2$ sono non distorte al modello $\mathcal{F} = \{f(x;\mu,1,2),\ \mu \in \Re\}$. Qui $\Phi(\cdot)$ indica la funzione di ripartizione della normale standardizzata.

(b) Supponendo di disporre di un valore iniziale $\tilde{\mu}_0$, si ottenga l'approssimazione dello stimatore $\tilde{\mu}$, definito dalla funzione di stima $q_2(\mu;\boldsymbol{x})$, fornita dal primo passo dell'algoritmo di Newton-Raphson.

Si suppongano ora $\mu = 1$ noto e δ e α ignoti.

(c) Si ottengano gli stimatori di massima verosimiglianza $\hat{\delta}$ e $\hat{\alpha}$.

(d) E' possibile affermare che $\hat{\delta}$ è non distorto ?

Soluzione

(a) Si tratta di stabilire se le variabili $[I_{(-\infty,0]}(X-\mu)-1/2]$ e $[\Phi(X-\mu)-1/2]$ hanno media nulla al modello $\mathcal{F}$. Si osservi che la funzione $f(x;\mu,1,2)$ è simmetrica attorno a μ che risulta, dunque, essere media e mediana della distribuzione di X. Di conseguenza,

$$E_\mu[I_{(-\infty,0]}(X-\mu)-1/2] = \Pr_\mu\{X \leq \mu\} - 1/2 = 0.$$

D'altra parte,

$$\begin{aligned} E_\mu[\Phi(X-\mu)-1/2] &= \int [\Phi(x-\mu)-1/2] f(x;\mu,1,2)dx \\ &= \frac{3}{2}\int_{\mu-1}^{\mu+1} [\Phi(x-\mu)-1/2](x-\mu)^2 dx \\ &= \frac{3}{2}\int_{-1}^{+1} [\Phi(z)-1/2] z^2 dz = 0, \end{aligned}$$

essendo la funzione z^2 simmetrica attorno allo zero e la funzione $\Phi(z)-1/2$ dispari. Quindi, entrambe le funzioni di stima considerate sono non distorte ad $\mathcal{F}$.

(b) Il primo passo dell'algoritmo di Newton Raphson fornisce l'approssimazione

$$\tilde{\mu}_1 = \tilde{\mu}_0 - \frac{q_2(\tilde{\mu}_0;\boldsymbol{x})}{\partial q_2(\mu;\boldsymbol{x})/\partial\mu|_{\mu=\tilde{\mu}_0}} = \tilde{\mu}_0 + \frac{\sum_i \Phi(x_i-\tilde{\mu}_0) - n/2}{\sum_i \phi(x_i-\tilde{\mu}_0)},$$

dove $\phi(\cdot)$ indica la funzione di densità della normale standardizzata.

(c) La funzione di verosimiglianza ha espressione

$$\begin{aligned} L(\delta,\alpha) &\propto \left(\frac{\alpha+1}{\delta^{\alpha+1}}\right)^n \prod_{i=1}^n |x_i-1|^\alpha I_{[1-\delta,\ 1+\delta]}(x_i) \\ &= \left(\frac{\alpha+1}{\delta^{\alpha+1}}\right)^n \prod_{i=1}^n |x_i-1|^\alpha I_{[|x_i-1|,\ +\infty)}(\delta) \\ &= \left[\left(\frac{\alpha+1}{\delta^{\alpha+1}}\right)^n \prod_{i=1}^n |x_i-1|^\alpha\right] I_{[max_i|x_i-1|,\ +\infty)}(\delta). \end{aligned}$$

Per ogni valore di $\alpha \geq 0$ fissato, L e funzione strettamente decrescente in δ, per $\delta \geq max_i|x_i-1|$. Quindi, fissato α, L è massima quando $\delta = max_i|x_i-1|$. Questo valore non dipende da α, quindi $\hat{\delta} = max_i|x_i-1|$. Inoltre,

$$L(\hat{\delta},\alpha) = \left(\frac{\alpha+1}{\hat{\delta}^{\alpha+1}}\right)^n \prod_{i=1}^n |x_i-1|^\alpha$$

e

$$l(\hat{\delta},\alpha) = \log(L(\hat{\delta},\alpha)) = n\log(\alpha+1) - n(\alpha+1)\log(\hat{\delta}) + \alpha\sum_{i=1}^n \log|x_i-1|.$$

Derivando rispetto a α e uguagliando a zero si ottiene

$$l_*(\hat{\delta},\alpha) = n/(\alpha+1) - n\log(\hat{\delta}) + \sum_{i=1}^n \log|x_i-1| = 0,$$

da cui

$$\hat{\alpha} = \frac{n}{n\log(\hat{\delta}) - \sum_{i=1}^{n} \log|x_i - 1|} - 1$$

(la derivata seconda della funzione $l(\hat{\delta}, \alpha)$ è sempre < 0).

(d) Essendo $\hat{\delta} = max_i |x_i - 1| < \delta$ con probabilità 1, lo stimatore non può che essere distorto.

39. Sia θ la probabilità che uno studente, iscritto ad un certo corso di laurea, superi un determinato esame al primo appello. Le convinzioni del docente su θ sono, a priori, formalizzabili mediante una variabile casuale Beta(a, b)[14] con $a = b = 1$. Siano $n \geq 1$ il numero di studenti che si presentano all'appello e X la variabile casuale che descrive il numero di studenti che superano l'esame. Sia x la realizzazione di X e si assumano indipendenti gli esiti dell'esame per i diversi studenti.

(a) Si scriva la verosimiglianza e si dica se la distribuzione a priori è un elemento di una famiglia coniugata naturale.

(b) Posto $x = n$, si trovi un intervallo di credibilità HPD con probabilità 0.95 per θ.

(c) Posto $n = 64$ e $x = 3$, si fornisca una stima puntuale per θ.

(d) Si fornisca una stima puntuale per θ nell'ipotesi che $n = 9$, $x = 3$ e che le convinzioni del docente siano formalizzate mediante una variabile discreta con supporto $\{0, 1/3, 2/3, 1\}$ e legge $(1/8, 3/8, 3/8, 1/8)$.

Soluzione

(a) La variabile casuale X può essere vista come una variabile che descrive il numero di successi in n prove indipendenti di Bernoulli, con probabilità di successo θ. La verosimiglianza è dunque

$$L(\theta; x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \propto \theta^x (1-\theta)^{n-x}.$$

Per determinare se la distribuzione a priori è coniugata occorre calcolare la distribuzione a posteriori e verificare se questa è della stessa famiglia della a priori. In questo caso si ha

$$\begin{aligned}\pi(\theta|x) \propto & L(\theta; x)\pi(\theta) \\ \propto & \binom{n}{x} \theta^x (1-\theta)^{n-x} \theta^{a-1} (1-\theta)^{b-1} \\ \propto & \theta^{x+a-1} (1-\theta)^{n-x+b-1}\end{aligned}$$

cioé la distribuzione a posteriori è ancora una beta. Si tratta quindi di famiglia coniugata.

(b) Se $x = n$ e $a = b = 1$, la distribuzione a posteriori è

$$\pi(\theta|x) = k\theta^n$$

dove la costante di normalizzazione k è

$$k = \left(\int_0^1 \theta^n d\theta \right)^{-1} = n + 1$$

[14] Se $Y \sim \mathrm{Beta}(a, b)$, allora $f(y; a, b) \propto y^{a-1}(1-y)^{b-1}$, per $y \in (0, 1)$, $a > 0$, $b > 0$. Inoltre, $E(Y) = a/(a + b)$.

Dato che la distribuzione a posteriori è definita in $[0,1]$ e ivi monotona crescente, l'intervallo di credibilità HPD è $[c,1]$ dove l'estremo c è determinato dall'equazione

$$0.95 = \int_c^1 \pi(\theta|x)d\theta \quad \text{o, equivalentemente,} \quad 0.05 = \int_0^c \pi(\theta|x)d\theta = c^{n+1}.$$

Quindi $c = 0.05^{1/(n+1)}$.

(c) Come stima puntuale bayesiana prendiamo la media a posteriori. Essendo la distribuzione a posteriori una Beta$(x+1, n-x+1)$ la stima puntuale è $(x+1)/(n+2)$ che, nel caso $x = 3$ e $n = 64$, è pari a $4/66 = 0.06$.

(d) Il calcolo della distribuzione a posteriori è riassunto nella tabella

θ	$\pi(\theta)$	$L(\theta;x) \propto \theta^3(1-\theta)^6$	$L(\theta;x)\pi(\theta)$	$\pi(\theta\|x)$
0	0.13	0	0	0
0.33	0.38	0.00325	0.00121875	0.89
0.67	0.38	0.00041	0.00015375	0.11
1	0.13	0	0	0

Come stima puntuale bayesiana possiamo prendere la media della distribuzione a posteriori:

$$E(\theta|x) = 0.89\frac{1}{3} + 0.11\frac{2}{3} = 0.37.$$

40. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile X avente funzione di densità

$$f(x;\lambda,\theta) = \frac{x^{\theta-1}}{k(\theta)\lambda^\theta}\exp(-x/\lambda), \quad x > 0,$$

con $\theta > 0$ e $\lambda > 0$ parametri e $k(\theta)$ costante di normalizzazione che dipende solo da θ. La variabile casuale X ha media $\theta\lambda$ e varianza $\theta\lambda^2$.

(a) Supponendo θ noto, si ottenga lo stimatore per λ basato sul metodo dei momenti, mostrando che si tratta di uno stimatore non distorto.

(b) Si stabilisca se lo stimatore di cui al punto (a) è efficiente.

(c) Si supponga λ noto e θ ignoto. Si mostri che la statistica $S = \sum_{i=1}^n \log x_i$ è sufficiente per θ.

(d) Si stabilisca se esiste il test ottimo, di livello α fissato, per la verifica dell'ipotesi $H_0 : \theta = \theta_0$ contro $H_1 : \theta > \theta_0$, quando λ è noto.

Soluzione

(a) Poiché $E_\lambda(X) = \theta\lambda$, lo stimatore basato sul metodo dei momenti si ottiene ponendo $\bar{X} = \theta\lambda$, da cui, risolvendo, $\tilde{\lambda} = \bar{X}/\theta$, con $\bar{X}$ media campionaria. Si tratta di uno stimatore non distorto in quanto

$$E_\lambda(\tilde{\lambda}) = E_\lambda(\bar{X})/\theta = \lambda.$$

(b) Possiamo scrivere

$$L(\lambda,\theta) = \frac{\prod_{i=1}^n x_i^{\theta-1} e^{-x_i/\lambda}}{[k(\theta)\lambda^\theta]^n}$$

e

$$l(\lambda,\theta) = (\theta-1)\sum_{i=1}^n \log x_i - (1/\lambda)\sum_{i=1}^n x_i - n\log k(\theta) - n\theta\log\lambda.$$

Con θ noto,

$$l(\lambda) = -(1/\lambda)\sum_{i=1}^n x_i - n\theta\log\lambda \quad \text{e} \quad l_*(\lambda) = (1/\lambda^2)\sum_{i=1}^n x_i - n\theta/\lambda.$$

Pertanto,

$$l_{**}(\lambda) = -2n\bar{x}/\lambda^3 + n\theta/\lambda^2 \quad \text{e} \quad E_\lambda(-l_{**}) = n\theta/\lambda^2.$$

Ne segue che il limite inferiore di Cramér-Rao è $\lambda^2/(n\theta)$. D'altro canto,

$$var(\tilde{\lambda}) = var(\bar{X}/\theta) = (1/\theta^2)(\theta\lambda^2/n) = \lambda^2/(n\theta)$$

e $\tilde{\lambda}$ è dunque stimatore efficiente.

(c) Con λ noto e θ ignoto, si ha

$$l(\theta) = (\theta-1)\sum_{i=1}^n \log x_i - n\log k(\theta) - n\theta\log\lambda.$$

Ciò mostra, come conseguenza diretta del criterio di fattorizzazione di Neyman-Fisher, che $S = \sum_{i=1}^n \log x_i$ è sufficiente per θ.

(d) Sia $\theta_1 > \theta_0$ un valore fissato. Scriviamo il rapporto di verosimiglianza (con λ fissato e noto).

$$t^*(\theta_0,\theta_1) = \frac{L(\theta_1)}{L(\theta_0)} = \frac{[k(\theta_0)\lambda^{\theta_0}]^n \prod_{i=1}^n x_i^{\theta_1-1}}{[k(\theta_1)\lambda^{\theta_1}]^n \prod_{i=1}^n x_i^{\theta_0-1}}.$$

Passando al logaritmo,

$$\begin{aligned}\log t^* &= n\log k(\theta_0) + (\theta_1-1)\sum_{i=1}^n \log x_i - n\log k(\theta_1) - (\theta_0-1)\sum_{i=1}^n \log x_i \\ &= n[\log k(\theta_0) - \log k(\theta_1)] + (\theta_1-\theta_0)\sum_{i=1}^n \log x_i.\end{aligned}$$

Pertanto, il rapporto di verosimiglianza è funzione monotona crescente della statistica S (qualunque sia il valore θ_1 considerato) e quindi il test uniformemente più potente per il sistema d'ipotesi considerato rifiuta H_0 quando $S > c_\alpha$, dove il valore di soglia c_α è opportunamente scelto per conseguire il livello di significatività α desiderato.

41. Sia X la variabile aleatoria che descrive la durata di vita, in determinate condizioni, di alcuni microrganismi. È noto che per X è adeguata la scelta di un modello esponenziale $\mathcal{F} = \{f(x;\lambda),\ \lambda > 0\}$, con λ parametro ignoto e $f(x;\lambda) = \lambda\exp(-\lambda x),\ \ x > 0$.

(a) Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da X. Si mostri che la funzione di stima $q(\lambda; \boldsymbol{x}) = \sum_{i=1}^n g(\lambda; x_i)$, con

$$g(\lambda; x_i) = \begin{cases} -b & \text{se} \quad \lambda x_i > a + b \\ a - \lambda x_i & \text{se } 0 < \lambda x_i \leq a + b, \end{cases}$$

$a = 0.9378$ e $b = 1.84$ è non distorta al modello $\mathcal{F}$.

(b) L'osservazione campionaria, relativa a 10 microrganismi, è riportata, insieme ad altre informazioni, nella tabella che segue: Sulla base di tali informazioni, si calcoli la stima della durata

durate osservate (in ore) x_i	durate cumulate	$(a+b)/x_i$
0.56	0.56	4.9603571
1.79	2.35	1.5518436
2.16	4.51	1.2860185
3.46	7.97	0.80283237
3.67	11.64	0.75689373
3.81	15.45	0.72908136
4.52	19.97	0.61455752
5.53	25.50	0.50231465
6.57	32.07	0.42280061
80.23	112.30	0.034622959

di vita media dei microrganismi prodotta dalla funzione di stima $q(\cdot;\cdot)$.

(c) Si confronti la stima ottenuta al punto (b) con la corrispondente stima di massima verosimiglianza, commentando il risultato.

Soluzione

(a) Basta mostrare che $E_\lambda\{g(\lambda; X)\} = 0$. Tenendo presente che λX ha distribuzione esponenziale di media 1 (sotto $\mathcal{F}$), si ha

$$E_\lambda\{g(\lambda; X)\} = \int_0^{a+b} (a-t)e^{-t}dt \; - be^{-(a+b)} = a\{1-e^{-(a+b)}\} - \int_0^{(a+b)} te^{-t}dt \; - be^{-(a+b)}.$$

Integrando per parti,

$$\begin{aligned} E_\lambda\{g(\lambda; X)\} &= a\{1 - e^{-(a+b)}\} - [-(t+1)e^{-t}]_0^{a+b} - be^{-(a+b)} \\ &= a - ae^{-(a+b)} + (a+b+1)e^{-(a+b)} - 1 - be^{-(a+b)} \\ &= a - 1 + e^{-(a+b)} \\ &= 0 \end{aligned}$$

con i valori fissati per a e b (e tenendo conto delle approssimazioni numeriche).

(b) La stima di λ prodotta dalla funzione di stima $q(\lambda; \boldsymbol{x})$ si ottiene individuando la radice dell'equazione $\sum_{i=1}^n g(\lambda; x_i) = 0$. Ora, si può scrivere

$$\sum_{i=1}^n g(\lambda; x_i) = \sum_{x_i:\ \lambda x_i \leq a+b} (a - \lambda x_i) \; - \; b\eta(\lambda),$$

dove $\eta(\lambda)$ indica il numero (che è funzione di λ) di osservazioni x_i per cui $\lambda x_i > a + b$. Ovviamente, $\eta(\lambda) \in \{0, 1, 2, \ldots, n\}$. Quindi,

$$\begin{aligned}\sum_{i=1}^{n} g(\lambda; x_i) &= \{n - \eta(\lambda)\}a \quad - \lambda \sum_{x_i:\ \lambda x_i \leq a+b} x_i \quad - b\eta(\lambda) \\ &= na - (a+b)\eta(\lambda) - \lambda \sum_{x_i:\ \lambda x_i \leq a+b} x_i.\end{aligned}$$

Sulla base dell'osservazione campionaria (relativa a $n = 10$ microrganismi) e alle altre informazioni riportate in tabella, i valori di λ per cui nessuna durata osservata x_i è tale che $\lambda x_i > a + b$ (o equivalentemente $x_i > (a + b)/\lambda$) sono quelli minori di 0.034622959. I valori di λ compresi tra 0.034622959 e 0.42280061 sono quelli per cui solo la più grande durata osservata $x_{(10)}$ è tale che $\lambda x_{(10)} > a + b$. Analogamente, i valori di λ compresi tra 0.42280061 e 0.50231465 sono tali per cui solo 2 (le 2 più grandi) durate osservate sono maggiori di $(a + b)/\lambda$. E così via. Quindi:

- per $\lambda < 0.034622959$, $\eta(\lambda) = 0$ e $\sum_{i=1}^{n} g(\lambda; x_i) = 10a - 112.30\lambda$; uguagliando a zero si ottiene la radice $\tilde{\lambda} = 0.0835$, che però è fuori dal range dei valori considerati
- per $0.034622959 < \lambda < 0.42280061$, $\eta(\lambda) = 1$ e $\sum_{i=1}^{n} g(\lambda; x_i) = 10a - (a + b) - 32.07\lambda \quad \to \tilde{\lambda} = 0.2058$ (radice interna al range dei valori considerati)
- per $0.42280061 < \lambda < 0.50231465$, $\eta(\lambda) = 2$ e $\sum_{i=1}^{n} g(\lambda; x_i) = 10a - 2(a + b) - 25.50\lambda \quad \to \tilde{\lambda} = 0.1499$ (radice esterna al range dei valori considerati).
- in tutti gli altri casi si ottengono radici sempre esterne ai range dei valori di volta in volta considerati.

Pertanto, si ha $\tilde{\lambda} = 0.2058$ e la stima della durata di vita media prodotta da $q(\lambda; \boldsymbol{x})$ risulta essere

$$\tilde{\mu} = 1/0.2058 = 4.86.$$

(c) La stima di massima verosimiglianza della durata di vita media è $\hat{\mu} = \bar{x} = 112.30/10 = 11.23$. A differenza di $\tilde{\mu}$, che è stimatore robusto al modello $\mathcal{F}$ perchè ottenuto da una funzione di stima limitata, lo stimatore di massima verosimiglianza risente, evidentemente, della presenza nel campione di un dato "anomalo" (80.23).

42. Si consideri il campione

$$3.67 \quad 1.86 \quad 1.96 \quad 1.21 \quad 33.1 \quad 2.58 \quad 0.769 \quad 3.90$$

da una variabile casuale Y con distribuzione log-normale[15] di parametri $(\mu, 1)$.

(a) Si mostri che la famiglia di distribuzioni normali è coniugata naturale alla verosimiglianza per μ.

Si scelga per μ, nella famiglia coniugata naturale, la distribuzione a priori con media 2 e varianza 2.

(b) Si fornisca una stima bayesiana per $\gamma = E(Y)$.

(c) Si verifichi l'ipotesi $H_0 : \gamma > 7.39$.

[15] Se $X \sim N(\mu, \sigma^2)$, allora $Z = \exp(X)$ ha distribuzione log-normale di parametri μ e σ^2, con funzione di densità $f(z; \mu, \sigma^2) = \frac{1}{z\sigma\sqrt{2\pi}} \exp\{-\frac{1}{2\sigma^2}(\log(z) - \mu)^2\}$, per $z > 0$, e media $\exp(\mu + \sigma^2/2)$.

(d) Si ottenga l'intervallo di credibilità HPD al 95% per μ.

(e) Si ottenga un intervallo di credibilità al 95% per la mediana di Y.

Soluzione

(a) La funzione di verosimiglianza è invariante rispetto a trasformazioni biunivoche dei dati. Il campione trasformato $x_i = \log(y_i)$, $i = 1, \ldots, n$, proviene da una distribuzione $N(\mu, 1)$. Pertanto, se a priori $\mu \sim N(\mu_0, \sigma_0^2)$, a posteriori $\mu|\mathbf{y} \sim N\left(\frac{n\sigma_0^2\bar{x}+1}{\mu_0(n\sigma_0^2+1)}, \frac{\sigma_0^2}{n\sigma_0^2+1}\right)$. Scegliendo, in particolare, $\mu_0 = 2$ e $\sigma_0^2 = 2$, si trova come distribuzione a posteriori per μ una $N(1.1, 0.118)$.

(b) Indichiamo con μ^* e σ^{2*} i parametri della distribuzione a posteriori. Si ha $\gamma = \exp(\mu + 1/2) = \exp(1/2)\exp(\mu)$. Quindi $E(\gamma|\mathbf{y}) = \exp(1/2)E(\exp(\mu)|\mathbf{y})$. Avendo μ distribuzione a posteriori $N(\mu^*, \sigma^{2*})$, $\exp(\mu)$ ha distribuzione a posteriori lognormale di parametri μ^* e σ^{2*}, con media $E(\exp(\mu)|\mathbf{y}) = \exp(\mu^* + \sigma^{2*}/2) = 3.19$. Essendo $\exp(1/2) = 1.65$, si ha che la stima bayesiana per γ è $E(\gamma|\mathbf{y}) = 5.2635$.

(c) L'ipotesi $H_0 : E(Y) = \gamma > 7.39$ equivale all'ipotesi $\mu > \log 7.39 - 1/2 = 1.5$. Quindi, utilizzando la distribuzione a posteriori di μ, si ottiene

$$\Pr\{H_0|\mathbf{y}\} = \Pr\{\mu > 1.5\} = 1 - \Phi\left(\frac{1.5 - 1.1}{\sqrt{0.118}}\right) = 0.1221219.$$

E ciò porta a rifiutare H_0.

(d) L'intervallo di credibilità HPD per μ ha estremi $\mu^* \pm \Phi^{-1}(0.975)\sqrt{\sigma^{2*}}$. Risulta quindi essere

$$[0.426, 1.773].$$

(e) La mediana di Y è e^{μ}. Se $[\mu_1, \mu_2]$ è un intervallo di credibilità al 95% per μ, allora $[e^{\mu_1}, e^{\mu_2}]$ è un intervallo di credibilità al 95% per la mediana di Y. L'intervallo cercato risulta quindi essere

$$[e^{0.426}, e^{1.773}] = [1.53, 5.89].$$

43. Siano $x_1, \ldots, x_n$ e $y_1, \ldots, y_n$ il numero di persone in fila agli sportelli di due uffici postali, in due quartieri di Padova, alle ore 12.00 di n giorni lavorativi. Si suppone che $x_1, \ldots, x_n$ e $y_1, \ldots, y_n$ siano campioni casuali semplici da variabili X e Y, entrambe con distribuzione di Poisson, di media, rispettivamente, $\exp(\delta + \lambda)$ e $\exp(\lambda)$, con $\delta \in \Re$, $\lambda \in \Re$. Si assume, inoltre, l'indipendenza tra i due campioni.

(a) Si stabilisca se il modello risultante costituisce una famiglia esponenziale regolare.

(b) Si ottenga lo stimatore di massima verosimiglianza $(\hat{\delta}, \hat{\lambda})$, stabilendo se è non distorto.

(c) Si ricavi la funzione di log-verosimiglianza profilo per δ.

(d) Con $n = 100$, $\sum_{i=1}^n x_i = 480$ e $\sum_{i=1}^n y_i = 520$, si verifichi l'ipotesi $H_0 : \delta = 0$ contro $H_1 : \delta \neq 0$, ad un livello approssimato del 5%.

(e) Si fornisca un'approssimazione per la distribuzione della statistica $t = \sum_{i=1}^n x_i / \sum_{i=1}^n y_i$.

Soluzione

(a) Per la funzione di verosimiglianza vale la relazione

$$L(\delta,\lambda)=\prod_{i=1}^{n}\frac{e^{-e^{\delta+\lambda}}e^{(\delta+\lambda)x_i}}{x_i!}\frac{e^{-e^{\lambda}}e^{\lambda y_i}}{y_i!}\propto e^{-n(e^{\lambda+\delta}+e^{\lambda})}e^{(\lambda+\delta)\sum_i x_i+\lambda\sum_i y_i}.$$

Il modello costituisce quindi una famiglia esponenziale di ordine 2, con parametro canonico (δ,λ). Lo spazio parametrico è $\Re\times\Re$. Quindi si tratta di famiglia esponenziale regolare.

(b) Si ha

$$l(\delta,\lambda)=\log(L(\delta,\lambda))=-n(e^{\lambda+\delta}+e^{\lambda})+(\lambda+\delta)\sum_i x_i+\lambda\sum_i y_i$$

e, derivando,

$$l_\delta(\delta,\lambda)=-ne^{\lambda+\delta}+n\bar{x}\quad \text{e}\quad l_\lambda(\delta,\lambda)=-n(e^{\lambda+\delta}+e^{\lambda})+n(\bar{x}+\bar{y}).$$

Dalla prima relazione, ugugliando a zero, si ottiene $e^{\lambda+\delta}=\bar{x}$ che, inserita nella seconda relazione, porta ad ottenere

$$-n(\bar{x}+e^{\lambda})+n(\bar{x}+\bar{y})=0$$

da cui $\hat{\lambda}=\log(\bar{y})$. Dato che $\delta+\lambda=\log(\bar{x})$, si ha $\hat{\delta}=\log(\bar{x})-\log(\bar{y})$.
Sappiamo che $E(\bar{y})=E(Y)=e^{\lambda}$. Quindi, per la diseguaglianza di Jensen, $E(\hat{\lambda})=E(\log(\bar{y}))\neq\log(E(\bar{y}))=\lambda$, e lo stimatore di massima verosimiglianza è distorto.

(c) Da $l_\lambda(\delta,\lambda)=-n(e^{\lambda+\delta}+e^{\lambda})+n(\bar{x}+\bar{y})$, ponendo uguale a zero e risolvendo in λ con δ fissato, si ottiene

$$\hat{\lambda}_\delta=\log(\bar{x}+\bar{y})-\log(1+e^{\delta}).$$

Quindi,

$$l_P(\delta)=l(\delta,\hat{\lambda}_\delta)=-ne^{\hat{\lambda}_\delta}(1+e^{\delta})+n(\delta+\hat{\lambda}_\delta)\bar{x}+n\hat{\lambda}_\delta\bar{y}.$$

(d) Con i dati di cui si dispone, $\bar{x}=4.8$ e $\bar{y}=5.2$. Ne segue che $\hat{\delta}=\log(4.8)-\log(5.2)=-0.08$, $\hat{\lambda}=\log(5.2)=1.648$, $\hat{\lambda}_{\delta=0}=\log(10)-\log(2)=1.61$, $l_P(0)=609.43$ e $l_P(\hat{\delta})=610.238$. Quindi $W_P(0)=2(l_P(\hat{\delta})-l_P(0))=1.616$ e l'ipotesi nulla non può essere rifiutata ad un livello (approssimato) del 5%.

(e) Evidentemente, $t=\exp(\hat{\delta})$. Inoltre, in quanto stimatore di massima verosimiglianza, per $\hat{\delta}$ vale l'approssimazione $\hat{\delta}\dot{\sim}N(\delta,1/(ne^{\lambda+\delta}))$. Quindi, utilizzando il metodo delta, si ha $t\dot{\sim}N(e^{\delta},e^{\delta}/ne^{\lambda})$.

44. Si assume che la durata (in mesi) di certe lampadine sia descritta da una variabile casuale Y con funzione di densità $f(y;\gamma,\beta)=\gamma\beta^{-\gamma}y^{\gamma-1}\exp\{-(y/\beta)^{\gamma}\}$ e funzione di ripartizione $F(y;\gamma,\beta)=1-\exp\{-(y/\beta)^{\gamma}\}$, per $y>0$, con $\gamma>0$ e $\beta>0$ parametri ignoti. In una prova di affidabilità, un campione di n lampadine viene testato e, alla fine della prova, si conosce il numero n_1 di lampadine la cui durata è risultata compresa tra 1 e 2 mesi e il numero n_2 di lampadine la cui durata è risultata superiore a 2 mesi.

(a) Si scriva la funzione di verosimiglianza per (γ,β), individuando una statistica sufficiente minimale.

(b) Si ottenga lo stimatore di massima verosimiglianza $(\hat{\gamma}, \hat{\beta})$.

(c) Si indichi come ottenere una regione di confidenza per (γ, β), di livello approssimato 0.95.

Soluzione

(a) In base al modello ipotizzato per la variabile Y, si ha

$$\Pr\{Y > 2\} = 1 - \Pr\{Y \leq 2\} = 1 - F(2; \gamma, \beta) = e^{-(2/\beta)^\gamma},$$

$$\Pr\{1 < Y \leq 2\} = F(2; \gamma, \beta) - F(1; \gamma, \beta) = e^{-(1/\beta)^\gamma} - e^{-(2/\beta)^\gamma},$$

$$\Pr\{Y \leq 1\} = F(1; \gamma, \beta) = 1 - e^{-(1/\beta)^\gamma}.$$

La funzione di verosimiglianza è quella relativa a un campione di n realizzazioni indipendenti di una variabile casuale multinomiale a tre celle con parametri $\pi_0 = \Pr\{Y \leq 1\}$, $\pi_1 = \Pr\{1 < Y \leq 2\}$ e $\pi_2 = \Pr\{Y > 2\}$. Si ha, pertanto,

$$L(\gamma, \beta; n_1, n_2) \propto \left[1 - e^{-(1/\beta)^\gamma}\right]^{n-n_1-n_2} \left[e^{-(1/\beta)^\gamma} - e^{-(2/\beta)^\gamma}\right]^{n_1} \left[e^{-(2/\beta)^\gamma}\right]^{n_2}.$$

La coppia (n_1, n_2) è, evidentemente, statistica sufficiente minimale.

(b) Posto $n_0 = n - n_1 - n_2$, è noto che le stime di massima verosimiglianza per i parametri π_0 e π_2 sono, rispettivamente, n_0/n e n_2/n. Allora, le stime di massima verosimiglianza per γ e β si possono ottenere risolvendo il sistema di equazioni

$$1 - e^{-(1/\beta)^\gamma} = n_0/n$$

$$e^{-(2/\beta)^\gamma} = n_2/n.$$

Dalla prima equazione si ottiene $-\gamma \log(\beta) = \log[-\log(1 - n_0/n)]$. Quindi,

$$\hat{\beta}_\gamma = e^{-\log[-\log(1-n_0/n)]/\gamma}.$$

Dalla seconda equazione si ottiene $\gamma \log(2) - \gamma \log(\beta) = \log[-\log(n_2/n)]$. Sostituendo a β la sua stima vincolata, si ricava (dopo alcuni passaggi)

$$\hat{\gamma} = \frac{\log[-\log(n_2/n)] - \log[-\log(1 - n_0/n)]}{\log(2)}.$$

Infine,

$$\hat{\beta} = \exp\left\{-\frac{\log[-\log(1 - n_0/n)]}{\hat{\gamma}}\right\}.$$

(c) Utilizzando la statistica $W(\gamma, \beta) = 2[l(\hat{\gamma}, \hat{\beta}) - l(\gamma, \beta)]$, una regione di confidenza per (γ, β), di livello approssimato 0.95, è data da

$$\mathcal{R} = \{(\gamma, \beta) : \ W(\gamma, \beta) < c_\alpha\},$$

con c_α 95-esimo percentile della distribuzione χ^2_2.

45. Nel "gioco delle tre carte", il proponente la scommessa (banco) mostra tre carte scoperte (il fante di quadri $J\diamondsuit$, la donna di picche $Q\spadesuit$ e il re di cuori $K\heartsuit$) per poi disporle, coperte, sul tavolo. Lo scommettitore punta una somma per scoprire una carta e riceve la somma raddoppiata se la carta scoperta è la donna di picche.

In $n = 10$ mani del gioco, cui partecipano tre soggetti diversi (A, B e C), si registrano i seguenti risultati:

scommettitore	B	B	C	A	A	B	B	C	B	A
carta scoperta	$Q\spadesuit$	$Q\spadesuit$	$Q\spadesuit$	$J\diamondsuit$	$K\heartsuit$	$Q\spadesuit$	$K\heartsuit$	$J\diamondsuit$	$Q\spadesuit$	$K\heartsuit$

Si adotti l'ipotesi semplificatrice secondo cui la probabilità di vincere è la stessa per tutte le mani e i risultati delle diverse mani sono indipendenti.

(a) Si specifichi un modello (bayesiano) per l'inferenza sulla probabilità di vittoria θ, individuando la famiglia di *a priori* coniugata naturale alla verosimiglianza.

(b) Usando l'*a priori* di media $1/2$ e varianza $1/12$ (della classe individuata), si calcoli una stima bayesiana puntuale per θ.

(c) Usando l'*a priori* di cui al punto precedente, si dica se in base a un test bayesiano si accetta o rifiuta l'ipotesi secondo cui la probabilità di vincere è maggiore di 0.5 .

Si potrebbe sospettare che il soggetto B non sia un concorrente genuino ma un complice del biscazziere e che quindi il banco lo favorisca per invogliare altri partecipanti. Si considerino allora le sole mani che coinvolgono il giocatore B, e sia τ la probabilità di vittoria per B.

(d) Assumendo anche per τ un'*a priori* (nella famiglia coniugata) avente media $1/2$ e varianza $1/12$, si calcoli la probabilità *a posteriori* dell'evento $\{\tau > 0.75\}$.

Soluzione

(a) La variabile casuale X che descrive il "numero di successi in n prove indipendenti" a parità di condizioni, ha distribuzione binomiale con probabilità di successo θ. La funzione di verosimiglianza è, dunque,

$$L(\theta; x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \propto \theta^x (1-\theta)^{n-x}.$$

Per determinare se una famiglia di distribuzioni è coniugata naturale alla verosimiglianza occorre scegliere un elemento generico della famiglia come distribuzione *a priori*, calcolare la distribuzione a posteriori e verificare se questa è ancora elemento della stessa famiglia. Nel caso specifico, scegliendo come *a priori* una distribuzione Beta(a, b), si ha

$$\begin{aligned}
\pi(\theta|x) &\propto L(\theta; x)\pi(\theta) \\
&\propto \binom{n}{x} \theta^x (1-\theta)^{n-x} \theta^{a-1} (1-\theta)^{b-1} \\
&\propto \theta^{x+a-1} (1-\theta)^{n-x+b-1}.
\end{aligned}$$

Quindi, la distribuzione *a posteriori* è ancora una beta. La famiglia di distribuzioni beta è coniugata naturale alla verosimigliana per θ.

(b) Come stimatore puntuale bayesiano prendiamo la media della distribuzione *a posteriori*. Essendo la distribuzione *a priori* una uniforme su (0,1), la distribuzione *a posteriori* risulta una Beta$(x + 1, n - x + 1)$. La stima puntuale è, dunque, $(x + 1)/(n + 2) = 6/12 = 0.5$

(c) La distribuzione *a posteriori* è
$$\pi(\theta|x) \propto \theta^5(1-\theta)^5,$$
ed è simmetrica rispetto a 0.5. Quindi $\Pr\{\theta > 0.5|x\} = \Pr\{\theta < 0.5|x\}$ e il test non permette di decidere.

(d) Se ci si limita a 5 prove e 4 successi, la distribuzione *a posteriori* è
$$\pi(\tau|x) = 30\tau^4(1-\tau)^1 = 30(\tau^4 - \tau^5).$$
Quindi,
$$\Pr\{\tau > 3/4|x\} = 1 - 30\int_0^{3/4} \tau^4(1-\tau)d\tau = 0.466 \ .$$

46. La tabella che segue riporta il numero di tentati suicidi e di suicidi accertati in Italia per $n = 5$ anni (dal 2000 al 2004, fonte: ISTAT).

	2004	2003	2002	2001	2000
suicidi accertati	3265	3361	2966	2992	3096
tentati suicidi	3481	3412	2949	2918	3352
totale tentativi	6746	6773	5915	5811	6448

Siano Y_i e X_i, per $i = 1, 2, \ldots, n$, le variabili casuali che descrivono, rispettivamente, il totale annuo di tentativi di suicidio e il numero annuo di suicidi accertati. Si supponga che la generica Y_i abbia distribuzione di Poisson di media λ e che le coppie $(Y_1, X_1), (Y_2, X_2), \ldots, (Y_n, X_n)$ siano indipendenti. Si assuma, infine, che la probabilità p che un tentativo di suicidio si concluda tragicamente sia costante tra gli individui e nel tempo.

(a) Si scriva la funzione di verosimiglianza per il parametro $\theta = (\lambda, p)$, ottenendo la stima $\hat{\theta}$ corrispondente.

(b) Si mostri che
$$I(\theta) = \begin{pmatrix} \frac{n}{\lambda} & 0 \\ 0 & \frac{n\lambda}{p(1-p)} \end{pmatrix}$$
è la matrice d'informazione attesa e si ottenga la stima della matrice di varianze e covarianze di $\hat{\theta}$.

(c) Siano $\bar{Y} = (1/n)\sum_{i=1}^n Y_i$ e $\bar{X} = (1/n)\sum_{i=1}^n X_i$. Si consideri la regione
$$R_c = \left\{(\lambda, p) : \frac{(\bar{Y} - \lambda)^2}{\frac{\lambda}{n}} + \frac{(\bar{X}/\bar{Y} - p)^2}{\frac{p(1-p)}{n\lambda}} < c\right\},$$
stabilendo per quale valore di c essa è una regione di confidenza per θ, di livello approssimato 0.95.

(d) Si ottenga la funzione di verosimiglianza profilo per il parametro $\tau = \lambda p$, che rappresenta il numero medio annuo di suicidi.

Soluzione

(a) Siano x_i e y_i le determinazioni delle variabili X_i e Y_i, rispettivamente, per $i = 1, \ldots, n$. Quindi x_i rappresenta il numero di suicidi "riusciti" nell'anno i, mentre y_i rappresenta il numero di suicidi tentati nello stesso anno. Dalle ipotesi formulate si deduce che la variabile X_i ha distribuzione, condizionata all'evento $Y_i = y_i$, che è binomiale di parametri y_i e p. Ne segue che la funzione di verosimiglianza relativa all'osservazione campionaria $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ ha espressione

$$L(\theta) = \prod_i \frac{\lambda^{y_i}}{y_i!} e^{-\lambda} \binom{y_i}{x_i} p^{x_i} (1-p)^{y_i - x_i}.$$

Quindi,

$$L(\theta) \propto \lambda^{\sum_i y_i} e^{-n\lambda} p^{\sum_i x_i} (1-p)^{\sum_i y_i - \sum_i x_i}.$$

Siano $\bar{x} = \frac{1}{n} \sum_i x_i$ e $\bar{y} = \frac{1}{n} \sum_i y_i$. Si ha

$$L(\theta) \propto \lambda^{n\bar{y}} e^{-n\lambda} p^{n\bar{x}} (1-p)^{n\bar{y} - n\bar{x}}$$

e per la log-verosimiglianza si ha l'espressione

$$l(\theta) = n\bar{y} \log \lambda - n\lambda + n\bar{x} \log p + (n\bar{y} - n\bar{x}) \log(1-p).$$

Derivando rispetto a λ, si ottiene

$$\frac{\partial l(\theta)}{\partial \lambda} = \frac{n\bar{y}}{\lambda} - n,$$

da cui, eguagliando a zero, si ricava lo stimatore $\hat{\lambda} = \bar{Y}$ e quindi la stima $\hat{\lambda} = 6338.6$. Allo stesso modo, derivando la log-verosimiglianza rispetto a p, si ottiene

$$\frac{\partial l(\theta)}{\partial p} = \frac{n\bar{x}}{p} - \frac{n\bar{y} - n\bar{x}}{1-p} = n \frac{\bar{x} - \bar{y}p}{p(1-p)}$$

da cui lo stimatore $\hat{p} = \bar{X}/\bar{Y}$ e quindi la stima $\hat{p} = 3136/6338.6 = 0.4947$.

(b) Calcoliamo le derivate seconde per ottenere l'informazione di Fisher. È chiaro che la derivata mista è nulla, mentre

$$\frac{\partial^2 l(\theta)}{\partial \lambda^2} = -\frac{n\bar{y}}{\lambda^2},$$

$$\frac{\partial^2 l(\theta)}{\partial p^2} = -\frac{n\bar{x}}{p^2} - \frac{n(\bar{y} - \bar{x})}{(1-p)^2}.$$

Di conseguenza si ha

$$-E\left(\frac{\partial^2 l(\theta)}{\partial \lambda^2}\right) = \frac{nE(\bar{y})}{\lambda^2} = \frac{n}{\lambda}$$

$$-E\left(\frac{\partial^2 l(\theta)}{\partial p^2}\right) = \frac{nE(\bar{x})}{p^2} + \frac{nE(\bar{y} - \bar{x})}{(1-p)^2} = \frac{n\lambda}{p} + \frac{n\lambda}{1-p} = \frac{n\lambda}{p(1-p)},$$

tenendo presente che $E[X_i] = E\{E[X_i|Y_i]\} = E[pY_i] = p\lambda$. Inoltre, la stima della matrice di varianze e covarianze è

$$I^{-1}(\hat{\theta}) = \begin{pmatrix} \frac{\hat{\lambda}}{n} & 0 \\ 0 & \frac{\hat{p}(1-\hat{p})}{n\ \hat{\lambda}} \end{pmatrix} = \begin{pmatrix} 1267.72 & 0 \\ 0 & 7.887306 \times 10^{-6} \end{pmatrix}.$$

(c) La quantità
$$\frac{(\bar{X}-\lambda)^2}{\frac{\lambda}{n}}+\frac{(\bar{X}/\bar{Y}-p)^2}{\frac{p(1-p)}{n\lambda}}$$
ha distribuzione asintotica χ^2_2, poiché è somma di due variabili con distribuzione asintotica normale standard e indipendenti. Quindi c deve essere il percentile di ordine $1-\alpha$ di una variabile casuale χ^2 con 2 gradi di libertà.

(d) Si riparametrizzi ponendo $\tau=\lambda p$ e $p=p$. Per il nuovo parametro (τ,p) la verosimiglianza diviene
$$L(\tau,p)\propto \tau^{n\bar{y}}e^{-n\tau/p}p^{-n\bar{y}}p^{n\bar{x}}(1-p)^{n\bar{y}-n\bar{x}}.$$
Quindi
$$l(\tau,p)=n\bar{y}\log\tau-n\tau/p+(n\bar{x}-n\bar{y})\log p+(n\bar{y}-n\bar{x})\log(1-p).$$
La verosimiglianza profilo per τ si ottiene massimizzando rispetto a p, per τ fissato. Ponendo $\partial l(\tau,p)/\partial p=0$ si ottiene
$$\frac{n\tau}{p^2}+\frac{n\bar{x}-n\bar{y}}{p}-\frac{n\bar{y}-n\bar{x}}{1-p}=0,$$
da cui
$$p_\tau=\frac{\tau}{\tau+\bar{y}-\bar{x}}.$$
Quindi la verosimiglianza profilo è
$$L_P(\tau)\propto \tau^{n\bar{y}}e^{-n(\tau+\bar{y}-\bar{x})}\left(\frac{\bar{y}-\bar{x}}{\tau}\right)^{n\bar{y}-n\bar{x}}.$$

47. Sia $x_1,x_2,\ldots,x_n$ un campione casuale semplice da una variabile X con funzione di densità $f(x;\theta,\gamma)=(\gamma/2)\exp(-\gamma|x-\theta|)$, per $x\in\Re$, con $\theta\in\Re$ e $\gamma>0$ parametri ignoti.

(a) Si stabilisca se la classe parametrica il cui generico elemento è $f(x;\theta,\gamma)$ costituisce una famiglia esponenziale.

(b) Si ottenga lo stimatore di massima verosimiglianza per (θ,γ).

Si supponga $\gamma=1$ noto e θ ignoto. Sia $\mathcal{F}=\{f(x;\theta),\ \theta\in\Re\}$, con $f(x;\theta)=(1/2)\exp(-|x-\theta|)$.

(c) Si mostri che la funzione di stima $\frac{1}{n}\sum_{i=1}^n \mathrm{sgn}(x_i-\theta)$, in cui $\mathrm{sgn}(\cdot)$ indica la funzione segno, definisce, come stimatore per θ, la mediana campionaria.

(d) Si confrontino, in termini di robustezza al modello $\mathcal{F}$, lo stimatore per θ di cui al punto (c) e lo stimatore basato sul metodo dei momenti.

Soluzione

(a) Dato che
$$|x-\theta|=\begin{cases} x-\theta & \text{se } x-\theta\geq 0\\ \theta-x & \text{se } x-\theta<0\end{cases}$$
la funzione di densità $f(x;\theta,\gamma)$ non può essere scritta nella forma caratterizzante un elemento di una famiglia esponenziale. Quindi, la classe parametrica considerata non costituisce una famiglia esponenziale.

(b) Si ha

$$L(\theta, \gamma) = \prod_{i=1}^{n} (\gamma/2) \exp(-\gamma |x_i - \theta|) = (\gamma/2) \exp\left(-\gamma \sum_{i=1}^{n} |x_i - \theta|\right).$$

Quindi,

$$l(\theta, \gamma) = \log(L(\theta, \gamma)) = n \log(\gamma) - \gamma \sum_{i=1}^{n} |x_i - \theta|.$$

Per γ fissato, $l(\theta, \gamma)$ è massima quando è minima la quantità $\sum_i |x_i - \theta|$. Ciò si verifica quando θ è la mediana campionaria (nota proprietà della mediana campionaria). Quindi $\hat{\theta}_\gamma = Me$. Evidentemente, il valore che massimizza $l(\theta, \gamma)$, per γ fissato, non dipende da γ. Quindi la mediana campionaria è lo stimatore di massima verosimiglianza per θ, $\hat{\theta} = Me$. Derivando poi $l(Me, \gamma)$ rispetto a γ e eguagliando a zero si ottiene

$$\hat{\gamma} = \frac{n}{\sum_i |x_i - Me|} \; .$$

(c) Sia

$$g(x_i; \theta) = sgn(x_i - \theta) = \begin{cases} 1 & \text{se } \; x_i - \theta > 0 \\ 0 & \text{se } \; x_i - \theta = 0 \\ -1 & \text{se } \; x_i - \theta < 0 \end{cases}$$

La funzione $g(z) = sgn(z)$ è funzione dispari; inoltre, la funzione $f(z) = (1/2) \exp(-|z|)$ è simmetrica attorno all'origine. Ne segue che la funzione di stima $q(\boldsymbol{x}; \theta) = (1/n) \sum_i g(x_i; \theta)$ è non distorta al modello $\mathcal{F}$. Inoltre, indicando con $x_{(i)}$ l'i-esimo elemento della statistica d'ordine per un campione di dimensione n da X, si ha che:

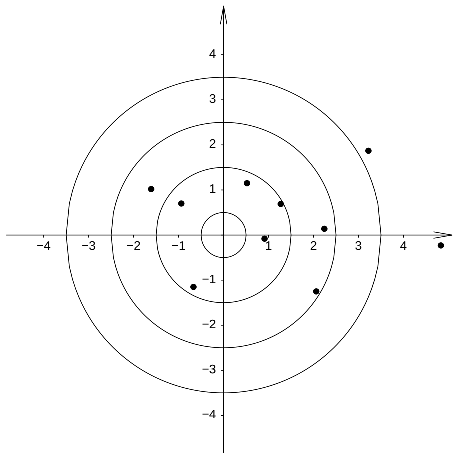
- se n è dispari, allora $q(\boldsymbol{x}; \theta) = 0$ se $\theta = x_{(\frac{n+1}{2})}$;
- se n è pari, allora $q(\boldsymbol{x}; \theta) = 0$ per ogni valore di θ tale che $x_{(\frac{n}{2})} < \theta < x_{(\frac{n}{2}+1)}$; in particolare, per

$$\theta = \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}.$$

Quindi, la funzione di stima $q(\boldsymbol{x}; \theta)$ definisce, come stimatore per θ la mediana campionaria.

(d) La funzione $g(x; \theta)$ è limitata in x. Ne segue che la mediana campionaria è stimatore robusto al modello $\mathcal{F}$. Lo stimatore per θ basato sul metodo dei momenti è la media campionaria (θ è la media di X), definita dalla funzione di stima $(1/n) \sum_i (x_i - \theta)$. Si tratta, evidentemente, di uno stimatore non robusto a $\mathcal{F}$.

48. Si sospetta che un fucile sia difettoso nel senso che i proiettili da esso sparati deviano verso destra rispetto al punto mirato. Al fine di verificare tale ipotesi, si effettuano $n = 10$ prove di tiro contro un bersaglio. I tiri sono effettuati da una macchina di precisione. Fissando l'origine di un piano cartesiano in corrispondenza del centro del bersaglio e gli assi in corrispondenza delle direzioni sinistra-destra e basso-alto (si veda la figura), si rilevano i punti colpiti come coppie di coordinate (x_i, y_i). I valori sono riportati nella tabella che segue.



i	x_i	y_i
1	0.91	−0.08
2	−0.67	−1.15
3	2.24	0.14
4	−1.61	1.02
5	3.22	1.87
6	−0.94	0.70
7	2.06	−1.25
8	1.27	0.69
9	0.52	1.15
10	4.83	−0.23
somma	11.83	2.86

Si assume che le coordinate (X, Y) del punto colpito da un generico sparo seguano leggi gaussiane indipendenti con medie, rispettivamente, μ_x e μ_y e varianze note σ_x^2 e σ_y^2. Inoltre, gli esiti dei vari tiri si assumono indipendenti.

(a) Scegliendo per μ_x e μ_y due distribuzioni a priori indipendenti normali di media 0 e varianza unitaria, si fornisca la distribuzione a posteriori per la coppia (μ_x, μ_y).

(b) Si fornisca uno stimatore puntuale bayesiano per la quantità $\theta = \mu_x^2 + \mu_y^2$, che rappresenta un indicatore globale del livello di precisione del fucile.

(c) Si ponga $\sigma_x^2 = 2$. Si può affermare, in un'ottica bayesiana, che i dati supportano l'ipotesi "il fucile devia a destra"?

Soluzione

(a) A priori, si assume $\mu_x \sim N(0,1)$ e $\mu_y \sim N(0,1)$ indipendenti. La verosimiglianza per (μ_x, μ_y) è

$$L(\mu_x, \mu_y) \propto \prod_{i=1}^{n} \phi\left(\frac{x_i - \mu_x}{\sigma_x}\right) \phi\left(\frac{y_i - \mu_y}{\sigma_y}\right),$$

dove $\phi(\cdot)$ è la funzione di densità di una normale standard. Se con π_x e π_y indichiamo le distribuzioni a priori per μ_x e μ_y, rispettivamente, si ha che, a posteriori,

$$\pi(\mu_x, \mu_y | (\boldsymbol{x}, \boldsymbol{y})) \propto \pi_x(\mu_x)\pi_y(\mu_y) \prod_{i=1}^{n} \phi\left(\frac{x_i - \mu_x}{\sigma_x}\right) \phi\left(\frac{y_i - \mu_y}{\sigma_y}\right).$$

Quindi,

$$\pi(\mu_x, \mu_y | (\boldsymbol{x}, \boldsymbol{y})) \propto \left(\pi_x(\mu_x) \prod_{i=1}^{n} \phi\left(\frac{x_i - \mu_x}{\sigma_x}\right)\right) \left(\pi_y(\mu_y) \prod_{i=1}^{n} \phi\left(\frac{y_i - \mu_y}{\sigma_y}\right)\right),$$

ossia la distribuzione a posteriori per (μ_x, μ_y) è il prodotto delle distribuzioni a posteriori per μ_x e μ_y. Queste ultime sono distribuzioni normali; in particolare,

$$\pi(\mu_x|\boldsymbol{x}) = N\left(\frac{\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}}, \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}}\right)$$

dove $\sigma_0 = 1$ e $\mu_0 = 0$. Pertanto, la distribuzione a posteriori per la coppia (μ_x, μ_y) è una normale bivariata a componenti indipendenti

$$N\left(\begin{bmatrix}\frac{\frac{n\bar{x}}{\sigma_x^2}}{\frac{n}{\sigma_x^2}+1}\\ \frac{\frac{n\bar{y}}{\sigma_x^2}}{\frac{n}{\sigma_y^2}+1}\end{bmatrix}, \begin{bmatrix}\frac{1}{\frac{n}{\sigma_x^2}+1} & 0\\ 0 & \frac{1}{\frac{n}{\sigma_y^2}+1}\end{bmatrix}\right). \tag{2}$$

(b) Come è noto, uno stimatore puntuale bayesiano è la media a posteriori. Si ha allora

$$E(\theta|(\boldsymbol{x},\boldsymbol{y})) = E(\mu_x^2+\mu_y^2|(\boldsymbol{x},\boldsymbol{y})) = E(\mu_x^2|(\boldsymbol{x},\boldsymbol{y})) + E(\mu_y^2|(\boldsymbol{x},\boldsymbol{y})). \tag{3}$$

Inoltre,

$$E(\mu_x^2|(\boldsymbol{x},\boldsymbol{y})) = var(\mu_x|(\boldsymbol{x},\boldsymbol{y})) + (E(\mu_x|(\boldsymbol{x},\boldsymbol{y})))^2 = \frac{1}{\frac{n}{\sigma_y^2}+1} + \left(\frac{\frac{n\bar{x}}{\sigma_x^2}}{\frac{n}{\sigma_x^2}+1}\right)^2. \tag{4}$$

Sostituendo la (4) e l'analoga formula relativa a μ_y^2 nella (3) si ottiene infine

$$E(\theta|(\boldsymbol{x},\boldsymbol{y})) = \frac{1}{\frac{n}{\sigma_y^2}+1} + \left(\frac{\frac{n\bar{x}}{\sigma_x^2}}{\frac{n}{\sigma_x^2}+1}\right)^2 + \frac{1}{\frac{n}{\sigma_y^2}+1} + \left(\frac{\frac{n\bar{y}}{\sigma_y^2}}{\frac{n}{\sigma_y^2}+1}\right)^2.$$

(c) Per verificare l'ipotesi "il fucile devia a destra" dobbiamo calcolare la probabilità a posteriori di $H_0 : \mu_x > 0$ (notiamo che μ_y è ininfluente). Usando la formula (2) per la distribuzione a posteriori e i dati di cui disponiamo, otteniamo

$$\mu_x|(\boldsymbol{x},\boldsymbol{y}) \sim N(0.985, 0.1666)$$

e quindi

$$\Pr\{H_0|\boldsymbol{x},\boldsymbol{y}\} = \Pr\{\mu_x > 0|(\boldsymbol{x},\boldsymbol{y})\} = 1 - \Phi\left(-0.985/\sqrt{0.1666}\right) \cong 1.$$

Ciò porta a decidere in favore di H_0.

49. Alcuni ricercatori sono interessati ad acquisire informazioni sulla durata minima del periodo di incubazione di una certa malattia. Un campione di n cavie viene quindi esposto ad un agente infettivo e per ogni cavia viene rilevato il tempo X trascorso fino alla comparsa dei primi sintomi della malattia. Si ritiene ragionevole per la variabile X un modello parametrico di Pareto

$$f(x;\lambda,\theta) = \theta\lambda^\theta x^{-(\theta+1)} \qquad x \geq \lambda, \quad \lambda > 0, \quad \theta > 0,$$

in cui λ rappresenta la quantità d'interesse e θ è supposto noto, $\theta = 2$.

(a) Si ottenga lo stimatore di massima verosimiglianza per λ, stabilendo se è non distorto.

(b) Si mostri che la statistica $S = \min\{X_1, X_2, \ldots, X_n\}$ è sufficiente. È anche minimale ?

(c) Si calcoli la funzione di ripartizione per la variabile casuale S e si stabilisca se esiste uno stimatore non distorto per λ nella classe degli stimatori del tipo cS, con c costante opportuna. [16]

Soluzione

(a) Sia $x_{(1)}$ il più piccolo valore nel campione. Si ha

$$L(\lambda) \propto \lambda^{2n}\left[\prod_{i=1}^{n} x_i^{-3}\right] I_{[\lambda,+\infty)}(x_{(1)}) \propto \lambda^{2n} I_{(0,x_{(1)}]}(\lambda).$$

Si tratta di una funzione strettamente crescente in λ sull'intervallo $(0, x_{(1)}]$. Quindi $\hat{\lambda} = S$. Poiché $\hat{\lambda} > \lambda$ con probabilità 1, lo stimatore di massima verosimiglianza è distorto.

(b) Come conseguenza diretta del criterio di fattorizzazione di Neyman-Fisher, S è statistica sufficiente. Inoltre, la partizione dello spazio campionario indotta da S coincide con quella di verosimiglianza:

$$L(\lambda; \boldsymbol{x_0}) \propto L(\lambda; \boldsymbol{x_1}) \iff S(\boldsymbol{x_0}) = S(\boldsymbol{x_1}).$$

Quindi S è statistica sufficiente minimale.

(c) Calcoliamo la funzione di ripartizione della variabile casuale S. Si ha

$$F_S(s;\lambda) = Pr\{S \le s\} = Pr\{X_{(1)} \le s\} = 1 - Pr\{X_{(1)} > s\} = 1 - [1 - F(s;\lambda,2)]^n,$$

dove $F(x;\lambda,\theta)$ indica la funzione di ripartizione di una variabile casuale di Pareto di parametri λ e θ. Ora,

$$F(x;\lambda,\theta) = \int_{\lambda}^{x} \theta\lambda^{\theta} u^{-(\theta+1)} du = 1 - (\lambda/x)^{\theta}, \quad \text{per} \quad x \ge \lambda,$$

quindi

$$F_S(s;\lambda) = 1 - (\lambda/s)^{2n}, \quad \text{per} \quad s \ge \lambda.$$

Dunque S ha distribuzione di Pareto di parametri λ e $2n$. Ne segue che

$$E(cS) = cE(S) = 2n\lambda c/(2n-1),$$

e lo stimatore cS è non distorto se $c = (2n-1)/(2n)$.

50. Si consideri per una variabile X il modello di Pareto, con funzione di densità $f(x;\lambda,\theta)$ data nell'esercizio precedente. Si supponga $\lambda = 2$ noto e $\theta > 1$ ignoto, cosicché $\mathcal{F} = \{f(x;2,\theta), \theta > 1\}$. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da X.

(a) Si ottenga lo stimatore per θ basato sul metodo dei momenti, individuando il funzionale statistico che lo definisce.

[16] Suggerimento: una variabile casuale di Pareto di parametri $\lambda > 0$, $\theta > 1$, ha media $\theta\lambda/(\theta - 1)$.

Si consideri la funzione di stima $q(\theta; \boldsymbol{x}) = \sum_{i=1}^n \{g(\theta; x_i) - b(\theta, k)\}$ con

$$g(\theta; x_i) = \begin{cases} k(\theta-1) & \text{se } x_i > k \\ (\theta-1)x_i & \text{se } x_i \leq k \end{cases}$$

dove $k > 2$ è una costante (finita) fissata e $b(\theta, k)$ è una opportuna funzione di θ e k.

(b) Si trovi l'espressione per la quantità $b(\cdot, \cdot)$ che rende la funzione di stima $q(\cdot; \cdot)$ non distorta al modello $\mathcal{F}$.

(c) Si confrontino, in termini di robustezza al modello $\mathcal{F}$, lo stimatore di massima verosimiglianza per θ, lo stimatore definito dalla funzione di stima non distorta $q(\theta; \boldsymbol{x})$ individuata al punto (b) e lo stimatore basato sul metodo dei momenti.

Soluzione

(a) Eguagliando momento campionario e momento teorico si ottiene

$$\frac{1}{n}\sum_{i=1}^n x_i = 2\theta/(\theta-1),$$

da cui $\tilde{\theta} = \bar{x}/(\bar{x}-2)$, con $\bar{x}$ media campionaria. Evidentemente, l'equazione di stima che definisce $\tilde{\theta}$ è

$$\frac{1}{n}\sum_{i=1}^n \left[x_i - \frac{2\theta}{\theta-1}\right] = 0.$$

Pertanto, il funzionale statistico associato a $\tilde{\theta}$ è il funzionale $T(F)$ definito implicitamente dall'equazione

$$\int \left[x - \frac{2T}{T-1}\right] dF(x) = 0.$$

(b) Calcoliamo la media di $g(\theta; X)$. Si ha

$$\begin{aligned} E_\theta\{g(\theta; X)\} &= \int_2^k (\theta-1)x\ \theta 2^\theta x^{-(\theta+1)} dx + \ k(\theta-1)\int_k^{+\infty} \theta 2^\theta x^{-(\theta+1)} dx \\ &= 2\theta \int_2^k (\theta-1)2^{\theta-1} x^{-[(\theta-1)+1]} dx + k(\theta-1)[1 - F(k; 2, \theta)] \\ &= 2\theta F(k; 2, \theta-1) + k(\theta-1)[1 - F(k; 2, \theta)] \\ &= 2\theta[1 - (2/k)^{\theta-1}] + k(\theta-1)(2/k)^\theta \\ &= 2\theta - k(2/k)^\theta. \end{aligned}$$

Quindi, perché la funzione di stima $q(\cdot; \cdot)$ sia non distorta al modello $\mathcal{F}$ deve essere $b(\theta, k) = 2\theta - k(2/k)^\theta$.

(c) Ricaviamo la funzione di stima che definisce lo stimatore di massima verosimiglianza. Si ha

$$L(\theta) \propto \theta^n 2^{n\theta} \prod_{i=1}^n x_i^{-(\theta+1)}.$$

Qundi, la funzione di log-verosimiglianza ha espressione

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

e

$$l_*(\theta) = \frac{dl(\theta)}{d\theta} = n/\theta + n \log 2 - \sum_{i=1}^{n} \log x_i.$$

Pertanto lo score di verosimiglianza non è funzione limitata, e lo stesso vale per la funzione di stima che definisce lo stimatore basato sul metodo dei momenti. I due stimatori associati (max-ver. e met. momenti) sono dunque non robusti a $\mathcal{F}$. La funzione di stima $q(\theta; \boldsymbol{x})$ è invece limitata in $\boldsymbol{x}$, per ogni elemento di $\mathcal{F}$.

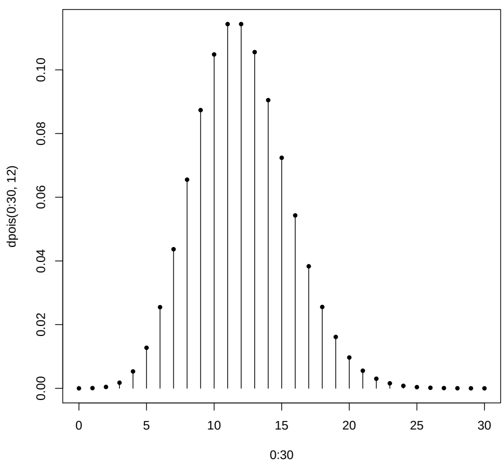
51. Siano N e M, rispettivamente, il numero di falli effettivamente commessi e il numero di falli fischiati dall'arbitro durante la finale dei campionati mondiali di calcio. Si supponga sia valida l'ipotesi semplificatrice secondo la quale l'arbitro fischia un fallo effettivamente commesso con probabilità θ (mentre è nulla la probabilità che venga fischiato un fallo inesistente), che tale probabilità sia costante nel tempo, e che la decisione di fischiare un fallo in una particolare circostanza sia indipendente dalle scelte operate in altri momenti. Sulla base del numero M di falli fischiati dall'arbitro, si vuole valutare, in termini bayesiani, quanti falli siano stati commessi nella partita. Si suppone che N abbia, a priori, distribuzione di Poisson di parametro $\lambda > 0$.

(a) Stante le ipotesi formulate, si dica qual è la verosimiglianza associata all'osservazione.

(b) Si calcoli la $\Pr\{N = 5 | M = 10\}$.

(c) Si individui la distribuzione a posteriori per N.

(d) Con $M = 20$, $\lambda = 30$ e $\theta = 0.6$, si individui l'intervallo di credibilità HPD al 60% (approssimativamente) per il numero di falli commessi. (Si utilizzino il grafico e i valori della funzione di probabilità della variabile casuale di Poisson di parametro 12, $p(x) = \Pr\{X = x\}$, di seguito forniti.)



x	$p(x)$	x	$p(x)$	x	$p(x)$
0	0.0000	11	0.1144	22	0.0030
1	0.0001	12	0.1144	23	0.0016
2	0.0004	13	0.1056	24	0.0008
3	0.0018	14	0.0905	25	0.0004
4	0.0053	15	0.0724	26	0.0002
5	0.0127	16	0.0543	27	0.0001
6	0.0255	17	0.0383	28	0.0000
7	0.0437	18	0.0255	29	0.0000
8	0.0655	19	0.0161	30	0.0000
9	0.0874	20	0.0097		
10	0.1048	21	0.0055		

Soluzione

(a) L'osservazione campionaria è costituita dal numero di falli fischiati dall'arbitro. Stante le ipotesi formulate, il numero di falli fischiati, condizionatamente al fatto siano stati commessi N falli, ha distribuzione binomiale di parametri N e θ. Quindi, la funzione di verosimiglianza è
$$\Pr\{M = m|N\} = \binom{N}{m}\theta^m(1-\theta)^{N-m} \propto \frac{n!}{(N-m)!}(1-\theta)^N,$$
per $m = 0, 1, 2, \ldots, N$

(b) $\Pr\{N = 5|M = 10\} = 0$. Se infatti sono stati fischiati 10 falli e si è detto che si ha probabilità 0 che l'arbitro fischi un fallo inesistente, deve risultare $N \geq M$. Formalmente,
$$\Pr\{N = 5|M = 10\} = \frac{\Pr\{N = 5\}\Pr\{M = 10|N = 5\}}{\Pr\{M = 10\}}$$
dove $\Pr\{M = 10|N = 5\}$, calcolata tenendo conto che $M|N = 5$ è binomiale di parametri 5 e θ, è nulla.

(c) La distribuzione a priori per N è una Poisson(λ). A posteriori si ha, allora,
$$\begin{aligned}\Pr\{N = n|M = m\} &\propto \frac{n!}{(n-m)!}(1-\theta)^n\frac{\lambda^n}{n!}e^{-\lambda}I(m \leq n)\\ &\propto \frac{1}{(n-m)!}(\lambda(1-\theta))^n I(m \leq n)\\ &\propto \frac{1}{(n-m)!}(\lambda(1-\theta))^{n-m}e^{-\lambda(1-\theta)}I(m \leq n)\end{aligned}$$
Cioè, a posteriori, $N - m \sim$ Poisson($\lambda(1-\theta)$).

(d) Se $p(\cdot)$ denota la funzione di probabilità della variabile casuale di Poisson di parametro 12, la distribuzione a posteriori di N è
$$\Pr\{N = n|M = 20\} = p(n-20)$$
per $n \geq m$ e 0 altrimenti. Per ispezione del grafico e della tabella forniti, si trova allora che l'intervallo di credibilità HPD al 60% (approssimativamente) per N ha estremi 29 e 34. In effetti, $\Pr\{29 \leq N \leq 34|M = 20\} = 0.6151$.

52. Sia $x_1, x_2, \ldots, x_n$, un campione casuale semplice da una variabile X con funzione di densità $f(x;\theta) = \theta(1+x)^{-\theta-1}$, con $x > 0$ e $\theta > 0$ parametro ignoto.

(a) Si ottenga lo stimatore di massima verosimiglianza, $\hat{\gamma}$, per $\gamma = 1/\theta$.

(b) Si ricavi la distribuzione di $\log(1+X)$ e si mostri che $\hat{\gamma}$ è stimatore non distorto per γ.

(c) Si stabilisca se $\hat{\gamma}$ è stimatore efficiente, nel senso che la sua varianza raggiunge il limite inferiore di Cramér-Rao.

(d) Si costruisca un intervallo di confidenza per γ, con livello di copertura esatto pari a 0.95.

Soluzione

(a) La funzione di verosimiglianza è

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \theta^n \prod_{i=1}^{n} (1+x_i)^{-\theta-1}.$$

Quindi la log-verosimiglianza è $l(\theta) = n\log(\theta) - (\theta+1)\sum_{i=1}^{n}\log(1+x_i)$ e lo *score* è

$$l_*(\theta) = (n/\theta) - \sum_{i=1}^{n} \log(1+x_i).$$

Inoltre, si ha $\partial^2 l(\theta)/\partial\theta^2 = -n/\theta^2$. Pertanto, lo stimatore di massima verosimiglianza è $\hat{\theta} = n/[\sum_{i=1}^{n}\log(1+x_i)]$ e, per la proprietà di equivarianza,

$$\hat{\gamma} = (1/n)\sum_{i=1}^{n}\log(1+x_i).$$

(b) Posto $Y = \log(1+X)$, si ha $X = e^Y - 1$ e $\frac{d}{dy}(e^y - 1) = e^y$. Quindi

$$f_Y(y) = f_X(e^y - 1)e^y = \theta(1+e^y-1)^{-\theta-1}e^y = \theta e^{-\theta y}.$$

Pertanto, Y ha distribuzione esponenziale di parametro θ. Ne segue che $E(\hat{\gamma}) = E(Y) = \gamma$.

(c) Si ha che $var(\hat{\gamma}) = (n/n^2)var(Y) = (\gamma^2/n)$, essendo γ^2 la varianza di Y. D'altro canto,

$$l_*(\gamma) = \left[n\gamma - \sum_{i=1}^{n}\log(1+x_i)\right](-1/\gamma^2) = -(n/\gamma) + (1/\gamma^2)\sum_{i=1}^{n}\log(1+x_i).$$

Allora,

$$l_{**}(\gamma) = \frac{dl_*(\gamma)}{d\gamma} = (n/\gamma^2) - (2/\gamma^3)\sum_{i=1}^{n}\log(1+x_i)$$

e

$$i(\gamma) = E(-l_{**}(\gamma)) = -(n/\gamma^2) + (2n/\gamma^2) = n/\gamma^2.$$

Ne segue, dunque, che $\hat{\gamma}$ è stimatore efficiente essendo $var(\hat{\gamma}) = i^{-1}(\gamma)$.

(d) Dato che $n\hat{\gamma} = \sum_{i=1}^{n}\log(1+x_i) \sim Gamma(n, 1/\gamma)$, risulta che $n\hat{\gamma}/\gamma \sim Gamma(n,1)$ è una quantità pivotale. Quindi un intervallo di confidenza esatto per γ si può ottenere come

$$bn\hat{\gamma} \le \gamma \le an\hat{\gamma}$$

dove a e b sono tali che $\Pr\{H < a\} = \Pr\{H > b\} = 0.025$, se $H \sim Gamma(n,1)$.

53. Sia $y_1, y_2, \ldots, y_n$ un campione casuale semplice da una variabile Y con funzione di densità

$$f(y;\lambda,\theta) = \begin{cases} \lambda\exp[-\lambda(y-\theta)] & \text{se } y \ge \theta \\ 0 & \text{altrimenti,} \end{cases}$$

dove $\lambda > 0$ e $\theta \in \Re$ sono parametri ignoti.

(a) Si ottenga lo stimatore di massima verosimiglianza per la coppia (λ, θ).

Si supponga $\theta = 1$ noto e si consideri il modello $\mathcal{F}$=$\{f(y; \lambda, 1), \quad \lambda > 0\}$ per Y.

(b) Si ottenga la funzione d'influenza per lo stimatore di massima verosimiglianza $\hat{\lambda}$, calcolata al generico elemento di $\mathcal{F}$.

(c) Si consideri la funzione di stima $q(\lambda; \boldsymbol{y}) = \sum_{i=1}^{n} g(\lambda; y_i)$, con

$$g(\lambda; y) = (y-1)I_{(0,k]}(y-1) + kI_{(k,+\infty)}(y-1) + b(\lambda, k),$$

dove k è una costante positiva fissata. Si trovi l'espressione per la quantità $b(\cdot, \cdot)$ che rende la funzione di stima q non distorta al modello $\mathcal{F}$ [17].

(d) Si confrontino, in termini di robustezza al modello $\mathcal{F}$, lo stimatore di massima verosimiglianza $\hat{\lambda}$ e lo stimatore definito dalla funzione di stima q di cui al punto (c).

Soluzione

(a) In questo caso,

$$\begin{aligned} L(\lambda, \theta) &= \prod_{i=1}^{n} \left[\lambda \exp(-\lambda(y_i - \theta)) I_{[\theta,+\infty)}(y_i)\right] \\ &= \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} (y_i - \theta)\right) I_{(-\infty, y_{(1)}]}(\theta), \end{aligned}$$

dove con $y_{(1)}$ si indica il più piccolo valore osservato. Quindi la funzione di verosimiglianza è non nulla solo per $\theta \le y_{(1)}$. Pertanto, il punto di massimo va cercato per $\theta \le y_{(1)}$. Ora, per ogni λ fissato e per $\theta \le y_{(1)}$, la funzione

$$\lambda^n \exp\left(-\lambda \sum_{i=1}^{n} (y_i - \theta)\right)$$

è funzione monotona crescente di θ. Quindi, il punto di massimo assoluto di $L(\lambda, \theta)$ deve necessariamente essere per $\theta = y_{(1)}$. Quindi $\hat{\theta} = y_{(1)}$. Inoltre,

$$l(\lambda, y_{(1)}) = \log L(\lambda, y_{(1)}) = n \log(\lambda) - \lambda \sum_{i=1}^{n} (y_i - y_{(1)})$$

e

$$\frac{\partial l(\lambda, y_{(1)})}{\partial \lambda} = (n/\lambda) - \sum_{i=1}^{n} (y_i - y_{(1)}).$$

Ne segue dunque che

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} (y_i - y_{(1)})}.$$

[17] Suggerimento: $\int_{a_1}^{a_2} t \exp(-t)\, dt = [-(t+1)\exp(-t)]_{a_1}^{a_2}$.

(b) Con $\theta = 1$ noto, lo *score* di verosimiglianza è

$$l_*(\lambda) = n/\lambda - \sum_{i=1}^{n}(y_i - 1) = \sum_{i=1}^{n}[(1/\lambda) - y_i + 1)] = \sum_{i=1}^{n} g^*(\lambda; y_i).$$

Inoltre, $\frac{dg^*(\lambda;y)}{d\lambda} = -1/\lambda^2$ e $-E_{f(y;\lambda,1)}(-1/\lambda^2) = 1/\lambda^2$. Quindi, al generico elemento di $\mathcal{F}$, la funzione d'influenza per lo stimatore di massima verosimiglianza $\hat{\lambda}$ è

$$IF(x;\lambda) = \lambda^2[(1/\lambda) - x + 1] = \lambda - \lambda^2 x + \lambda^2.$$

(c) Perché la funzione di stima q sia non distorta ad $\mathcal{F}$ deve essere $\int g(\lambda;y)f(y;\lambda,1)dy = 0$ per ogni $\lambda > 0$. Deve pertanto risultare

$$\begin{aligned}
-b(\lambda,k) &= \int_1^{k+1}(y-1)\lambda e^{-\lambda(y-1)}dy + k\int_{k+1}^{+\infty}\lambda e^{-\lambda(y-1)}dy \\
&= \int_0^k t\lambda e^{-\lambda t}dt + k\int_k^{+\infty}\lambda e^{-\lambda t}dt \\
&= (1/\lambda)\int_0^{\lambda k} u e^{-u}du + ke^{-\lambda k} \\
&= (1/\lambda)[-(\lambda k+1)e^{-\lambda k} - 1] + ke^{-\lambda k} \\
&= (1/\lambda)(1 - e^{-\lambda k}).
\end{aligned}$$

(d) La funzione di stima $q(\lambda; \boldsymbol{y})$, contrariamente a quanto accade per lo *score* di verosimiglianza, è funzione limitata (in $\boldsymbol{y}$). Quindi, lo stimatore da essa definito, a differenza di $\hat{\lambda}$, è robusto a $\mathcal{F}$.

54. Il club degli amici dei roditori, che raccoglie possessori di roditori domestici, è interessato a ottenere informazioni sul livello d'istruzione dei suoi soci. In particolare, il club è interessato a raccogliere informazioni sulle proporzioni $(\theta_1, \theta_2, \theta_3)$ di soci con livello d'istruzione, rispettivamente, basso, medio e alto. Un'indagine, condotta su un campione (casuale semplice) di soci, fornisce come risultato il numero di intervistati (x_1, x_2, x_3) che dichiarano, rispettivamente, un titolo di studio di livello basso, medio e alto.

(a) Si mostri che la famiglia di distribuzioni di Dirichlet$(\alpha_1, \alpha_2, \alpha_3)$ [18]

$$\pi(\theta_1,\theta_2,\theta_3) = \begin{cases} \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_1+\alpha_2+\alpha_3)}\theta_1^{\alpha_1}\theta_2^{\alpha_2}\theta_3^{\alpha_3} & \text{se } \theta_1+\theta_2+\theta_3 = 1 \\ 0 & \text{altrimenti,} \end{cases}$$

con $\alpha_1, \alpha_2, \alpha_3$ reali positivi, è coniugata naturale alla verosimiglianza per $(\theta_1, \theta_2, \theta_3)$.

Si supponga di osservare la terna (51,30,19) e di voler usare come a priori l'elemento della famiglia Dirichlet per cui $\sum_{j=1}^{3}\alpha_j = 10$ e la media a priori di $(\theta_1, \theta_2, \theta_3)$ è pari a $(0.4, 0.4, 0.2)$.

(b) Si fornisca una stima puntuale bayesiana per la proporzione di soci che posseggono un livello d'istruzione medio o basso.

[18] Se (Y_1, Y_2, Y_3) ha distribuzione di Dirichlet di parametri $\alpha_1, \alpha_2, \alpha_3$, allora la variabile marginale Y_i ha distribuzione Beta$(\alpha_i, \sum_{j=1}^{3}\alpha_j - \alpha_i)$.

(c) Si stabilisca la forma degli intervalli di credibilità HPD per la proporzione di soci con livello d'istruzione basso.

(d) Cosa si può dire relativamente all'ipotesi H_0: *la maggior parte dei soci possiede un livello d'istruzione basso*?

Soluzione

(a) La funzione di verosimiglianza relativa all'osservazione (x_1, x_2, x_3) è quella multinomiale:

$$L(x_1, x_2, x_3|\theta_1, \theta_2, \theta_3) = \frac{n!}{x_1!x_2!x_3!}\theta_1^{x_1}\theta_2^{x_2}\theta_3^{x_3}.$$

Pertanto, per la densità a posteriori vale la relazione

$$\begin{aligned}\pi(\theta_1, \theta_2, \theta_3|\mathbf{x}) &\propto \pi(\theta_1, \theta_2, \theta_3)L(x_1, x_2, x_3|\theta_1, \theta_2, \theta_3)\\ &\propto \theta_1^{\alpha_1}\theta_2^{\alpha_2}\theta_3^{\alpha_3}I(\theta_1+\theta_2+\theta_3=1)\theta_1^{x_1}\theta_2^{x_2}\theta_3^{x_3}\\ &\propto \theta_1^{\alpha_1+x_1}\theta_2^{\alpha_2+x_2}\theta_3^{\alpha_3+x_3}I(\theta_1+\theta_2+\theta_3=1)\end{aligned}$$

La distribuzione a posteriori è dunque una Dirichlet$(\alpha_1+x_1, \alpha_2+x_2, \alpha_3+x_3)$: la famiglia di distribuzioni di Dirichlet è coniugata naturale alla verosimiglianza multinomiale.

(b) Dato che $\theta_i \sim$ Beta$(\alpha_i, \sum\alpha_j - \alpha_i)$, si ha

$$E(\theta_1, \theta_2, \theta_3) = \left(\frac{\alpha_1}{\sum_i\alpha_i}, \frac{\alpha_2}{\sum_i\alpha_i}, \frac{\alpha_3}{\sum_i\alpha_i}\right).$$

Quindi

$$\begin{aligned}\alpha_1 &= 0.40\sum_i\alpha_i = 4\\ \alpha_2 &= 0.40\sum_i\alpha_i = 4\\ \alpha_3 &= 0.20\sum_i\alpha_i = 2,\end{aligned}$$

stante il vincolo $\sum_i\alpha_i = 10$. Con l'a priori che ha questi valori per α_1, α_2, α_3 e il campione (51,30,19), la distribuzione a posteriori è Dirichlet$(4+51, 4+30, 2+19)$ e

$$E(\theta_1+\theta_2|\mathbf{x}) = E(\theta_1|\mathbf{x}) + E(\theta_2|\mathbf{x}) = \frac{55}{110} + \frac{34}{110} = \frac{89}{110}.$$

(c) La distribuzione a posteriori per θ_1 è

$$(\theta_1|\mathbf{x}) \sim \text{Beta}(55, 55).$$

Si tratta di una distribuzione simmetrica e unimodale. L'intervallo di credibilità HPD è quindi simmetrico rispetto alla media che è pari a 0.5.

(d) Si noti che $\theta_1 > \theta_2 + \theta_3$ se e solo se $\theta_1 > 0.5$. Essendo $(\theta_1|\mathbf{x}) \sim$ Beta$(55, 55)$, risulta

$$\Pr\{H_0|\mathbf{x}\} = 0.5.$$

Nulla si può dire, dunque, circa l'ipotesi H_0.

55. Si suppoga di disporre di un'unica osservazione y dalla variabile casuale Y, discreta, la cui distribuzione appartiene alla famiglia caratterizzata dalla legge $p(y;\theta)$, con spazio parametrico $\Theta = \{1,2,3,4,5,6\}$, specificata in tabella.

y	10	20	30	40	50	60
$p(y;\theta=1)$	0.5	0.2	0.1	0.1	0.1	0
$p(y;\theta=2)$	0.2	0.5	0.1	0.1	0.1	0
$p(y;\theta=3)$	0.1	0.2	0.5	0.1	0.1	0
$p(y;\theta=4)$	0.1	0.1	0.2	0.5	0.1	0
$p(y;\theta=5)$	0.1	0.1	0.1	0.2	0.5	0
$p(y;\theta=6)$	0	0.1	0.1	0.1	0.2	0.5

(a) Si ottenga lo stimatore di massima verosimiglianza $\hat{\theta}$, stabilendo se è non distorto. Si può affermare che $\hat{\theta}$ è statistica sufficiente minimale?

(b) Si proponga un test, di livello $\alpha = 0.3$, per risolvere il problema di verifica d'ipotesi $H_0 : \theta = 1$ contro $H_1 : \theta \neq 1$.

(c) Supponendo che l'osservazione sia $y = 20$, si calcoli una stima della varianza di $\hat{\theta}$.

(d) Supponendo che l'osservazione sia $y = 20$, si ottenga una regione di confidenza per θ di livello 0.7.

Soluzione

(a) Dall'ispezione della tabella fornita si evince immediatamente che $\hat{\theta} = 1$ se $y = 10$, $\hat{\theta} = 2$ se $y = 20$, $\hat{\theta} = 3$ se $y = 30$, $\hat{\theta} = 4$ se $y = 40$, $\hat{\theta} = 5$ se $y = 50$, $\hat{\theta} = 6$ se $y = 60$. In definitiva, quindi,

$$\hat{\theta} = y/10.$$

D'altro canto,

$$E_{\theta=1}(\hat{\theta}) = (10/10)\times 0.5 + (20/10)\times 0.2 + (30/10)\times 0.1 + (40/10)\times 0.1 + (50/10)\times 0.1 + 0 = 2.1.$$

Ciò basta per affermare che lo stimatore è distorto. Inoltre, è facile verificare che la partizione di verosimiglianza è costituita dagli elementi $\{10\}, \{20\}, \{30\}, \{40\}, \{50\}, \{60\}$. Quindi $\hat{\theta}$ è statistica sufficiente minimale.

(b) Una statistica test adeguata è la statistica test del rapporto di verosimiglianza $\lambda = \frac{L(\theta_0)}{L(\hat{\theta})} = \frac{p(y;\theta_0)}{p(y;\hat{\theta})}$. I valori che essa assume sono riportati nella tabella che segue.

	$y=10$	$y=20$	$y=30$	$y=40$	$y=50$	$y=60$
λ	1	0.4	0.2	0.2	0.2	0

L'ipotesi nulla è rifiutata per valori piccoli di λ e risulta che $\Pr\{\lambda < 0.4 | H_0\} = 0.3$. Ne segue che, al livello 0.3, l'ipotesi nulla è rifiutata se $\lambda < 0.4$, ossia se $y \notin \mathcal{A} = \{10, 20\}$.

(c) Si ha che

$$var_\theta(\hat{\theta}) = E_\theta(\hat{\theta}^2) - E^2_\theta(\hat{\theta})$$

e $v\hat{a}r(\hat{\theta}) = var_{\hat{\theta}}(\hat{\theta})$. D'altro canto, se $y = 20$, risulta $\hat{\theta} = 2$ e

$$E_{\theta=2}(\hat{\theta}) = 0.2 + 1 + 0.3 + 0.4 + 0.5 + 0 = 2.4,$$

$$E_{\theta=2}(\hat{\theta}^2) = 1 \times 0.2 + 4 \times 0.5 + 9 \times 0.1 + 16 \times 0.1 + 25 \times 0.1 + 0 = 7.2.$$

Pertanto, si ha $v\hat{a}r(\hat{\theta}) = 7.2 - 2.4^2 = 1.44$.

(d) La regione di confidenza può essere costruita invertendo la regione di accettazione, diciamo $\mathcal{A}_\theta$, del test del rapporto di verosimiglianza di livello $\alpha = 0.3$:

$$\hat{\Theta}(y) = \{\theta \in \Theta : y \in \mathcal{A}_\theta\}.$$

Al punto (b) si è stabilito che, quando $\theta = 1$, $\mathcal{A}_1 = \{10, 20\}$. In maniera analoga si ricava che $\mathcal{A}_2 = \{10, 20\}$, $\mathcal{A}_3 = \{20, 30\}$, $\mathcal{A}_4 = \{30, 40\}$, $\mathcal{A}_5 = \{40, 50\}$ e $\mathcal{A}_6 = \{50, 60\}$. Allora, l'osservazione $y = 20$ appartiene ad $\mathcal{A}_1$, $\mathcal{A}_2$ e $\mathcal{A}_3$. Pertanto, la regione di confidenza cercata è

$$\hat{\Theta}(y = 20) = \{1, 2, 3\}.$$

56. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile discreta X avente funzione di probabilità

$$f(x;\theta) = \frac{\theta^x a_x}{b(\theta)}, \qquad \text{per} \ \ x = 0, 1, 2, 3, \ldots,$$

con $\theta > 0$ parametro ignoto, $\{a_x\}$ insieme di costanti non negative e $b(\theta)$ funzione reale derivabile almeno due volte.

(a) Si stabilisca se lo stimatore di massima verosimiglianza $\hat{\theta}$ è (qualitativamente) robusto al modello $\mathcal{F}$ avente legge marginale $f(x;\theta)$.

(b) Posto $\tau = \theta \frac{b'(\theta)}{b(\theta)}$, con $b'(\theta) = db(\theta)/d\theta$, si ricavi lo stimatore ottimo (a varianza minima tra i non distorti) per τ. Si fornisca un'espressione per la varianza di tale stimatore.

Si consideri il caso particolare in cui $b(\theta) = -\log(1-\theta)$, $\theta \in (0, 1)$ e $a_x = 0$ se $x = 0$, $a_x = 1/x$ se $x > 0$. Si supponga che sia $n = 10$ e che l'osservazione campionaria sia tale che $\sum_{i=1}^{10} x_i = 35$.

(c) Si ottenga l'approssimazione della stima di massima verosimiglianza $\hat{\theta}$ fornita dal primo passo dell'algoritmo di Newton-Raphson con valore inizale $\hat{\theta}_0 = 0.5$.

(d) Si fornisca un intervallo di confidenza per $\rho = -\log(1-\theta)$, di livello approssimato 0.90.

Soluzione

(a) Poiché $f(x;\theta) = a_x \exp[x\log(\theta) - \log(b(\theta))]$, le funzioni di verosimiglianza e di log-verosimiglianza hanno espressione

$$L(\theta) = \exp[\log(\theta)\sum_{i=1}^n x_i - n\log(b(\theta))] \quad \text{e} \quad l(\theta) = \log(\theta)\sum_{i=1}^n x_i - n\log(b(\theta)).$$

Quindi,

$$l_*(\theta) = \frac{dl(\theta)}{d\theta} = \frac{1}{\theta}\sum_{i=1}^n x_i - n\frac{b'(\theta)}{b(\theta)} = \sum_{i=1}^n [x_i/\theta - b'(\theta)/b(\theta)].$$

Essendo la funzione di stima che definisce $\hat{\theta}$ non limitata, lo stimatore di massima verosimiglianza non è qualitativamente robusto al modello parametrico considerato.

(b) Il modello parametrico considerato costituisce una famiglia esponenziale di odine uno. Si tratta di una famiglia esponenziale regolare, in quanto il parametro canonico $\log(\theta)$ varia su tutto l'insieme dei reali. Ne segue che, sotto campionamento casuale semplice, la statistica canonica $T = \sum_{i=1}^n x_i$ è statistica sufficiente minimale completa. Dalla relazione $E_\theta[l_*(\theta)] = 0$ (prima identità di Bartlett), si ricava che $E_\theta(T/n) = \theta b'(\theta)/b(\theta) = \tau$. Inoltre, l'informazione osservata risulta essere

$$j(\theta) = \frac{\sum_{i=1}^n x_i}{\theta^2} + n\frac{b''(\theta)b(\theta) - b'(\theta)^2}{b(\theta)^2}$$

e, quindi, per l'informazione attesa vale la relazione

$$i(\theta) = n\left[\frac{b'(\theta)}{\theta b(\theta)} + \frac{b''(\theta)b(\theta) - b'(\theta)^2}{b(\theta)^2}\right] = ni_1(\theta).$$

Poiché $var_\theta(l_*(\theta)) = i(\theta)$, si ha $var(T) = \theta^2 i(\theta)$. In definitiva, essendo T/n stimatore non distorto per τ, funzione della statistica sufficiente minimale completa, esso è stimatore ottimo. Un'espressione per la varianza di T/n è $var(T/n) = \theta^2 i_1(\theta)/n$.

(c) Posto $b(\theta) = -\log(1-\theta)$, si ha

$$b'(\theta) = \frac{1}{1-\theta}, \quad b''(\theta) = \frac{1}{(1-\theta)^2},$$

$$l_*(\theta) = \frac{\sum_{i=1}^n x_i}{\theta} + \frac{n}{(1-\theta)\log(1-\theta)}, \quad j(\theta) = \frac{\sum_{i=1}^n x_i}{\theta^2} - \frac{n + n\log(1-\theta)}{(1-\theta)^2\log^2(1-\theta)}.$$

Quindi, con i dati forniti $\sum_{i=1}^n x_i = 35$ e $n = 10$, risulta $l_*(0.5) = 41.14$ e $j(0.5) = 114.453$. Ne segue che l'approssimazione della stima di massima verosimiglianza $\hat\theta$ fornita dal primo passo dell'algoritmo di Newton-Raphson è

$$\hat\theta_1 = \hat\theta_0 + \frac{l_*(\hat\theta_0)}{j(\hat\theta_0)} = 0.5 + 41.14/114.453 = 0.86.$$

(d) Usando $\hat\theta_1$ come approssimazione della stima di massima verosimiglianza, un intervallo di confidenza per θ, di livello approssimato 0.90, è dato da

$$(\hat\theta_1 - 1.64/\sqrt{j(\hat\theta_1)}, \quad \hat\theta_1 + 1.64/\sqrt{j(\hat\theta_1)}),$$

ovvero $(0.736, 0.984)$, dato che $j(0.86) = 174.836$. Poiché ρ è funzione monotona crescente di θ, l'intervallo cercato per ρ è $(-\log(1-0.736), \quad -\log(1-0.984))$.

57. Un aereo è disperso in una regione Θ che, ai fini delle successive operazioni di ricerca, è ragionevole suddividere in tre zone denominate $\theta_1, \theta_2, \theta_3$. Inizialmente si valuta pari a π_i la probabilità che l'aereo sia precipitato nella zona θ_i. Per ritrovare il velivolo vengono effettuate delle ricognizioni nelle varie zone, ed è noto che la probabilità di ritrovare l'aereo con una ricognizione nella zona θ_i, condizionatamente al fatto che l'aereo si trovi effettivamente nella zona θ_i, è p_i. Si assume inoltre che gli esiti delle ricognizioni (in una stessa zona o in zone diverse) siano indipendenti condizionatamente al luogo del disastro.

(a) Tre ricognizioni, ciascuna effettuata in una zona diversa, non danno luogo al ritrovamento. Sulla base di questa osservazione si individui la distribuzione a posteriori della "zona del disastro" impostando il problema in termini Bayesiani (specificando esplicitamente distribuzione a priori e verosimiglianza).

Si pongano $\pi_1 = \pi_2 = \pi_3$ e $p_1 = 1/6$, $p_2 = p_3 = 1/3$.

(b) Si dica qual è la zona del disastro più probabile sulla base dell'esito delle tre ricognizioni.

(c) Si proponga uno stimatore puntuale bayesiano per la zona del disastro.

(d) Si verifichi, sulla base dell'esito delle tre ricognizioni, l'ipotesi nulla "l'aereo è precipitato nella zona θ_1" contro l'ipotesi alternativa "l'aereo non è precipitato nella zona θ_1".

(e) Se dopo $3n$ ricognizioni, di cui n per ciascuna regione, non si è ancora ritrovato l'aereo, si commenti l'affermazione *ne sappiamo quanto prima* con riferimento a $n = 1$, $n = 5$, $n = 20$.

Soluzione

L'oggetto dell'inferenza è il luogo ove l'aereo è precipitato; $\theta_1, \theta_2, \theta_3$ sono i possibili stati di natura. La distribuzione *a priori* sugli stati di natura è $\pi_i = \Pr\{\theta_i\}$ $\;i = 1, 2, 3$. Chiaramente, $\pi_1 + \pi_2 + \pi_3 = 1$. L'esperimento sulla base del quale fare inferenza è quello che genera la sequenza degli esiti delle ricognizioni effettuate. Conviene indicare in generale con X_{ij} la variabile casuale che descrive l'"esito dell'j-ma ricognizione nell'area θ_i". Assumiamo che X_{ij} valga 1 se la ricognizione ha avuto successo (l'aereo è stato ritrovato), 0 altrimenti. Con questa notazione, allora, il contributo alla verosimiglianza di una osservazione x_{ij} è determinato dalle probabilità

$$\Pr\{X_{ij} = 1|\theta_h\} = \begin{cases} p_h & h = i \\ 0 & h \neq i \end{cases},$$

e

$$\Pr\{X_{ij} = 0|\theta_h\} = 1 - \Pr\{X_{ij} = 1|\theta_h\} = \begin{cases} 1 - p_h & h = i \\ 1 & h \neq i \end{cases}.$$

(a) Nel caso di tre ricognizioni con esito negativo l'osservazione è

$$E = \{(X_{11} = 0) \cap (X_{21} = 0) \cap (X_{31} = 0)\}.$$

La verosimiglianza è data da

$$\begin{aligned} \Pr\{(X_{11} = 0) \cap (X_{21} = 0) \cap (X_{31} = 0)|\theta_h\} &= \Pr\{X_{11} = 0|\theta_h\} \Pr\{X_{21} = 0|\theta_h\} \Pr\{X_{31} = 0|\theta_h\} \\ &= \begin{cases} 1 - p_1 & \text{se } h = 1 \\ 1 - p_2 & \text{se } h = 2 \\ 1 - p_3 & \text{se } h = 3. \end{cases} \end{aligned}$$

Di conseguenza la distribuzione *a posteriori* è

$$\Pr\{\theta_h|E\} \propto \Pr\{E|\theta_h\} \Pr\{\theta_h\} = (1 - p_h)\pi_h.$$

Il fattore di normalizzazione è $k = \sum_{i=1}^{3} (1 - p_i)\pi_i$ e dunque

$$\Pr\{\theta_h|E\} = (1 - p_h)\pi_h/k.$$

(b) Se supponiamo che $\pi_1 = \pi_2 = \pi_3$, cioè $\pi_i = 1/3\ \forall i$, $p_1 = 1/6$, $p_2 = p_3 = 1/3$, si ha

$$\Pr\{\theta_h|E\} \propto \Pr\{E|\theta_h) P(\theta_h\} = (1-p_h)\pi_h = \begin{cases} 0.28 & \text{se } h = 1 \\ 0.22 & \text{se } h = 2,3 \end{cases}$$

e, normalizzando ($k = 0.72$),

$$\Pr\{\theta_h|E\} = \begin{cases} 0.38 & \text{se } h = 1 \\ 0.31 & \text{se } h = 2,3 \end{cases}$$

La "zona del disastro" più probabile è quindi θ_1.

(c) Tra i più comuni stimatori puntuali bayesiani sono la media, la mediana e la moda della distribuzione a posteriori. Essendo gli stati di natura non numerici e non ordinati media e mediana non sono opzioni disponibili. La moda a posteriori è quindi la stima puntuale bayesiana più naturale. La moda della distirbuzione a posteriori è, in altre parole, la "zona del disastro" più probabile: quindi θ_1.

(d) Calcoliamo le probabilità a posteriori assegnate alle ipotesi

$$\begin{cases} H_0 : & \theta_1 \\ H_1 : & \theta_2 \cup \theta_3. \end{cases}$$

Si ha

$$\Pr\{H_0|E\} = \Pr\{\theta_1|E\} = 0.38 \quad \Pr\{H_1|E\} = 1 - \Pr\{H_0|E\} = 0.62.$$

Pertanto, $\frac{\Pr\{H_0|E\}}{\Pr\{H_1|E\}} = \frac{0.38}{0.62} = 0.613 < 1$ e l'ipotesi H_0 è rifiutata.

(e) *Ne sappiamo quanto prima* vorrebbe dire che l'osservazione campionaria non fornisce elementi in più per decidere quale sia la zona ove è precipitato l'aereo. Questo sarebbe vero se le osservazioni lasciassero invariata la distribuzione sugli stati di natura. Non è però così, come emerge dalla risposta alla domanda (b): la distribuzione *a posteriori* è diversa dalla distribuzione *a priori* e le informazioni raccolte ci fanno propendere per la zona θ_1. Nel caso $n = 5$ e $n = 20$ la cosa è ancora più evidente. Infatti, con $E = \{\cap_{j=1}^{5} \cap_{i=1}^{3} (X_{ij} = 0)\}$ la distribuzione a posteriori è

$$\Pr\{\theta_h|E\} \propto \Pr\{E|\theta_h\} \Pr\{\theta_h\} = (1-p_h)^5\pi_h = \begin{cases} 0.13 & \text{se } h = 1 \\ 0.04 & \text{se } h = 2,3 \end{cases}$$

e, normalizzando ($k = 0.21$),

$$\Pr\{\theta_h|E\} = \begin{cases} 0.62 & \text{se } h = 1 \\ 0.19 & \text{se } h = 2,3\ . \end{cases}$$

Con $E = \{\cap_{j=1}^{20} \cap_{i=1}^{3} (X_{ij} = 0)\}$,

$$\Pr\{\theta_h|E\} \propto \Pr\{E|\theta_h\} \Pr\{\theta_h\} = (1-p_h)^{20}\pi_h = \begin{cases} 0.0087 & \text{se } h = 1 \\ 0.0001 & \text{se } h = 2,3 \end{cases}$$

e, normalizzando,

$$\Pr\{\theta_h|E\} = \begin{cases} 0.978 & \text{se } h = 1 \\ 0.011 & \text{se } h = 2,3\ . \end{cases}$$

58. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile X. Si consideri per X il modello statistico $\mathcal{F} = \{f(x;\mu),\ \mu \in \Re\}$, con μ parametro ignoto e funzione di densità

$$f(x;\mu) = \exp\{(x-\mu) - \exp(x-\mu)\}, \quad x \in \Re.$$

La funzione di ripartizione relativa a $f(x;\mu)$ è $F(x;\mu) = 1 - \exp\{-\exp(x-\mu)\}$.

(a) Si stabilisca se lo stimatore di massima verosimiglianza per μ è robusto a $\mathcal{F}$ rispetto a piccole deviazioni e/o dati anomali.

Si consideri la funzione di stima $q(\mu;\boldsymbol{x}) = \sum_{i=1}^n \{g(\mu;x_i) + \exp[-\exp(k)] - 1\}$ con

$$g(\mu;x_i) = \begin{cases} \exp(k) & \text{se } x_i - \mu > k \\ \exp(x_i - \mu) & \text{se } x_i - \mu \le k \end{cases}$$

dove $k > 0$ è una costante (finita) fissata.

(b) Si mostri che la funzione di stima $q(\cdot;\cdot)$ è non distorta al modello $\mathcal{F}$ [19].

(c) Cosa si può dire circa le differenze, in termini di robustezza e efficienza al modello $\mathcal{F}$, tra lo stimatore di massima verosimiglianza per μ e lo stimatore definito dalla funzione di stima $q(\mu;\boldsymbol{x})$?

(d) Si fornisca un intervallo di confidenza, di livello approssimato 0.95, per la mediana (diciamo ρ) di X.

Soluzione

(a) Si ha

$$L(\mu) = \prod_{i=1}^n \exp\{(x_i-\mu) - \exp(x_i-\mu)\}, \ \text{ e } \ l(\mu) = \log(L(\mu)) = \sum_{i=1}^n \{(x_i-\mu) - \exp(x_i-\mu)\}.$$

Pertanto,

$$l_*(\mu) = \frac{dl(\mu)}{d\mu} = \sum_{i=1}^n g_*(\mu;x_i), \quad \text{con} \quad g_*(\mu;x_i) = \exp(x_i-\mu) - 1.$$

Poiché la funzione $[\exp(z) - 1]$ tende a -1 quando $z \to -\infty$ e a $+\infty$ quando $z \to +\infty$, lo score di verosimiglianza non è funzione limitata e quindi lo stimatore di massima verosimiglianza

$$\hat{\mu} = -\log \frac{n}{\sum_{i=1}^n \exp(x_i)}$$

non è robusto a $\mathcal{F}$.

(b) Calcoliamo la media di $g(\mu;X)$. Si ha

$$E_\mu\{g(\mu;X)\} = \int_{-\infty}^{k+\mu} \exp(x-\mu)\exp\{(x-\mu) - \exp(x-\mu)\}dx + \exp(k)\exp\{-\exp(k)\}.$$

[19] Suggerimento: $\int_{b1}^{b2} te^{-t}dt = [-(t+1)e^{-t}]_{b_1}^{b2}$

Posto $\exp(x-\mu)=t$,

$$\begin{aligned}\int_{-\infty}^{k+\mu}\exp(x-\mu)\exp\{(x-\mu)-\exp(x-\mu)\}dx &= \int_0^{\exp(k)} t\exp\{\log(t)-t\}(1/t)dt\\ &= \int_0^{\exp(k)} t\exp(-t)dt,\\ &= [-(t+1)\exp(-t)]_0^{\exp(k)}\\ &= -\{\exp(k)+1\}\exp\{-\exp(k)\}+1\end{aligned}$$

utilizzando il suggerimento. Quindi

$$E_\mu\{g(\mu;X)\}=1-\exp\{-\exp(k)\},$$

e la funzione di stima q è non distorta.

(c) Si può notare che la funzione $g(\cdot;\cdot)$ è ottenuta troncando lo score di verosimiglianza (relativo alla generica osservazione). Pertanto, a differenza dello stimatore di massima verosimiglianza per μ, quello definito dalla funzione q è robusto a $\mathcal{F}$. L'efficienza di tale stimatore (rispetto a quello di massima verosimiglianza) aumenta la crescere della soglia k. Ciò, ovviamente, a scapito delle sue doti di robustezza.

(d) Ponendo $F(\rho;\mu)=1/2$ e risolvendo in ρ si ottiene che per la mediana di X vale l'espressione $\rho=\mu+\log(-\log(1/2))$. Quindi, un intervallo di confidenza per ρ si ricava facilmente da un intervallo per μ. In particolare, poiché $l_{**}(\mu)=\frac{d^2l(\mu)}{d\mu^2}=-\sum_{i=1}^n\exp(x_i-\mu)$, si ha che

$$j(\hat\mu)=-l_{**}(\hat\mu)=\sum_{i=1}^n\exp(x_i-\hat\mu)=\exp(-\hat\mu)\sum_{i=1}^n\exp(x_i)=n.$$

Dunque, un intervallo di confidenza per μ, di livello approssimato 0.95, ha estremi $\hat\mu\pm 1.96/\sqrt{n}$ e il corrispondente intervallo per ρ ha estremi $[\hat\mu+\log(-\log(1/2))]\pm1.96/\sqrt{n}$.

59. L'efficacia di un analgesico viene testata su un gruppo di m soggetti maschi e su un gruppo di n soggetti femmine, tutti sofferenti di emicrania. Ogni soggetto riceve il trattamento per un periodo fissato, al termine del quale dichiara se esso ha sortito un effetto positivo, determinando un miglioramento della qualità della propria vita. Siano γ e β le probabilità (ignote) che il trattamento determini un tale miglioramento su un maschio e una femmina, rispettivamente, e siano x e y il numero di maschi e femmine (tra i soggetti osservati) che si dichiarano effettivamente soddisfatti del trattamento. Obiettivo principale dello studio è stabilire se il trattamento è più efficace sui maschi che sulle femmine.
Si indichino con X e Y le variabili di cui sono realizzazioni x e y, e si pongano $S=X+Y$, $N=m+n$, $\delta=\gamma-\beta$.

(a) Si ricavi il test di livello approssimato $\alpha=0.05$, basato sulla funzione di verosimiglianza profilo per δ, che risolve il problema di verifica d'ipotesi $H_0:\delta=0$ contro $H_1:\delta>0$.

(b) Si mostri che, assumendo valida l'ipotesi nulla di cui al punto (a), la legge di $X|S=s$ è quella ipergeometrica di indici s e N e parametro m.

(c) Si spieghi perché è ragionevole risolvere il problema di verifica di ipotesi di cui al punto (a) mediante il test condizionato (al valore osservato s) che rifiuta l'ipotesi nulla se $x>k$, con k soglia da fissare opportunamente.

(d) Con $m = 24$, $n = 20$ e assumendo che l'osservazione campionaria sia tale che $s = 37$, si determini il livello di significatività (esatto) più piccolo effettivamente conseguibile con il test di cui al punto (c).

Soluzione

(a) Nel contesto considerato, X e Y sono due variabili casuali indipendenti con distribuzione binomiale di parametri (m, γ) e (n, β), rispettivamente. Pertanto, per la funzione di verosimiglianza vale la relazione

$$L(\gamma, \beta) \propto \gamma^x (1-\gamma)^{m-x} \beta^y (1-\beta)^{n-y}.$$

Passando alla parametrizzazione (δ, β), con $\delta = \gamma - \beta$, si ha

$$L(\delta, \beta) = (\delta+\beta)^x (1-\delta-\beta)^{m-x} \beta^y (1-\beta)^{n-y}$$

e

$$l(\delta, \beta) = \log(L(\delta, \beta)) = x\log(\delta+\beta) + (m-x)\log(1-\delta-\beta) + y\log(\beta) + (n-y)\log(1-\beta).$$

È chiaro che la stima di massima verosimiglianza per (γ, β) è $(\hat\gamma, \hat\beta)$, con $\hat\gamma = x/m$ e $\hat\beta = y/n$. Per la proprietà di equivarianza, allora, $\hat\delta = \hat\gamma - \hat\beta$. Inoltre,

$$\frac{\partial l(\delta, \beta)}{\partial \beta} = \frac{x}{\delta+\beta} - \frac{m-x}{1-\delta-\beta} + \frac{y}{\beta} - \frac{n-y}{1-\beta}$$

e, per $\delta = 0$,

$$\left.\frac{\partial l(\delta, \beta)}{\partial \beta}\right|_{\delta=0} = \frac{x+y}{\beta} - \frac{m+n-x-y}{1-\beta} = \frac{s}{\beta} - \frac{N-s}{\beta},$$

dove $s = x+y$. La stima di massima verosimiglianza vincolata $\hat\beta_{\delta=0}$ si ottiene dall'equazione

$$\frac{s}{\beta} - \frac{N-s}{\beta} = 0$$

e risulta essere $\hat\beta_{\delta=0} = s/N$. Quindi, la statistica test adeguata per il problema di verifica d'ipotesi trattato è

$$r_P(0) = sgn(\hat\delta - 0)\sqrt{W_P(0)},$$

con $W_P(0) = 2\{l(\hat\delta, \hat\beta) - l(0, \hat\beta_{\delta=0})\}$. Ad un livello di significatività approssimato 0.05, l'ipotesi nulla si rifiuta se $r_P(0) > 1.64$.

(b) Condizionatamente a $S = s$, X ha supporto dato dai numeri interi compresi tra $\max\{0, s-n\}$ e $\min\{s, m\}$ (estremi inclusi). Inoltre, sotto H_0, cioè quando $\gamma = \beta$, S ha distribuzione binomiale di parametri (N, γ) e,

$$\begin{aligned}
\Pr\{X = x | S = s\} &= \frac{\Pr\{(S = s) \cap (X = x)\}}{\Pr\{S = s\}} = \frac{\Pr\{X + Y = s | X = x\}\Pr\{X = x\}}{\Pr\{S = s\}} \\
&= \frac{\Pr\{Y = s - x\}\Pr\{X = x\}}{\Pr\{S = s\}} \\
&= \frac{\binom{n}{s-x}\gamma^{s-x}(1-\gamma)^{n-s-x}\binom{m}{x}\gamma^x(1-\gamma)^{m-x}}{\binom{N}{s}\gamma^s(1-\gamma)^{N-s}} \\
&= \frac{\binom{n}{s-x}\binom{m}{x}}{\binom{N}{s}}.
\end{aligned}$$

(c) Dato il numero complessivo s di soggetti che si dichiarano soddisfatti del trattamento, è chiaro che un numero x (relativamente) elevato di soggetti maschi potrebbe indicare come verosimile una maggiore efficacia del trattamento sui maschi piuttosto che sulle femmine. È quindi ragionevole pensare di risolvere il problema di verifica di ipotesi di cui al punto (a) mediante il test condizionato (al valore osservato s) che rifiuta l'ipotesi nulla se $x > k$, con k soglia da fissare opportunamente.

(d) Con i dati forniti, il più grande valore del supporto della legge di $X|(S = 37)$ è $\min\{24, 37\} = 24$. Il più piccolo livello di significatività effettivamente conseguibile con il test che rifiuta H_0 se $X > k$ è, allora, pari a

$$\Pr\{X = 24|S = 37\} = \frac{\binom{20}{13}}{\binom{44}{37}}.$$

60. Sia $x_1, x_2 \ldots, x_n$ un campione casuale semplice di numerosità $n = 10$ da una variabile X con distribuzione di Bernoulli di parametro θ. Si scelga per θ la seguente distribuzione a priori discreta.

θ	0.1	0.2	0.3	0.4	0.5	0.6
$\pi(\theta)$	0.05	0.38	0.05	0.05	0.39	0.08

Sia poi $(0, 0, 1, 0, 1, 1, 0, 0, 0, 0)$ l'osservazione campionaria.

(a) Si calcoli la distribuzione a posteriori e si fornisca una stima bayesiana di θ.

(b) Si individuino tutte le regioni di credibilità HPD per θ, indicandone i rispettivi livelli.

(c) Posto $\tau = n\ var(X|\theta)$, si verifichi l'ipotesi $\tau > 2$.

Soluzione

(a) Chiaramente, per la funzione di verosimiglianza vale la relazione

$$L(\theta) = L(\mathbf{x}|\theta) \propto \theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}.$$

Con l'osservazione campionaria di cui si dispone, gli elementi del problema possono, allora, riassumersi nella tabelle che segue, in cui la penultima colonna fornisce la distribuzione a posteriori $\pi(\theta|\mathbf{x}) \propto \pi(\theta)L(\theta)$.

$\pi(\theta)$	θ	$L(\theta)$	$L(\theta)\pi(\theta)$	$\pi(\theta\|\mathbf{x})$	τ
0.05	0.1	0.00048	0.00002	0.02	0.9
0.38	0.2	0.00168	0.00064	0.5	1.6
0.05	0.3	0.00222	0.00011	0.09	2.1
0.05	0.4	0.00179	0.00009	0.07	2.4
0.39	0.5	0.00098	0.00038	0.3	2.5
0.08	0.6	0.00035	0.00003	0.02	2.4

Una stima bayesiana di θ può essere la moda della distribuzione a posteriori, quindi 0.2.

(b) Per ispezione della tabella, si può concludere che le regioni di credibilità HPD per θ sono gli insiemi $\{0.2\}$, $\{0.2, 0.5\}$, $\{0.2, 0.3, 0.5\}$, $\{0.2, 0.3, 0.4, 0.5\}$ e $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$. Si tratta di regioni di livello 0.5, 0.8, 0.89, 0.96 e 1, rispettivamente.

(c) Si ha $\tau = n\ var(X|\theta) = 10\theta(1-\theta)$. I valori assunti da τ sono riportati nell'ultima colonna della tabella. Evidentemente,

$$\Pr\{\tau > 2|\mathbf{x}\} = \Pr\{\theta > 0.2|\mathbf{x}\} = 0.48.$$

Quindi, l'ipotesi formulata non può essere accettata.

61. Sia $x_1, \ldots, x_n$ un campione casuale semplice di dimensione n da una variabile continua X, avente funzione di densità

$$f(x;\theta) = \frac{\Gamma(2\theta)}{\Gamma^2(\theta)} x^{\theta-1}(1-x)^{\theta-1},$$

per $x \in (0,1)$ e con $\theta > 0$ parametro ignoto.

(a) Si individui una statistica sufficiente minimale per l'inferenza su θ.

(b) Posto $n = 1$, si fornisca la regione critica più potente, di livello (esatto) $\alpha = 0.05$, per il problema di verifica d'ipotesi $H_0 : \theta = 1$ contro $H_1 : \theta > 1$.

Con $n > 1$, si supponga di voler risolvere il problema di verifica d'ipotesi di cui al punto (b) utilizzando il test basato sulla statistica $t = \sum_{i=1}^n I_{(\frac{1}{2}-c,\ \frac{1}{2}+c)}(x_i)$, con $c \in [0.1, 0.2]$ costante fissata, che rifiuta H_0 se $t > k$, essendo k una soglia da fissare opportunamente.

(c) Con $n = 50$ e $c = 0.15$, si scelga k in modo che il test abbia livello approssimato 0.05.

(d) Si fornisca un valore approssimato per la potenza del test di cui al punto (c) in corrispondenza dell'alternativa $\theta = 2$.

Soluzione

(a) Si ha

$$L(\theta) \propto \left[\frac{\Gamma(2\theta)}{\Gamma^2(\theta)}\right]^n \prod_{i=1}^n [x_i(1-x_i)]^{\theta-1}.$$

Quindi, presi due punti $\boldsymbol{x_a}$ e $\boldsymbol{x_b}$ dello spazio campionario, il rapporto tra le verosimiglianze

$$\frac{L(\boldsymbol{x_a};\theta)}{L(\boldsymbol{x_b};\theta)} = \frac{\prod_{i=1}^n [x_{ai}(1-x_{ai})]^{\theta-1}}{\prod_{i=1}^n [x_{bi}(1-x_{bi})]^{\theta-1}} = \left[\frac{\prod_{i=1}^n x_{ai}(1-x_{ai})}{\prod_{i=1}^n x_{bi}(1-x_{bi})}\right]^{\theta-1}$$

non dipende da θ se e solo se $\boldsymbol{x_a}$ e $\boldsymbol{x_b}$ sono tali che

$$\prod_{i=1}^n x_{ai}(1-x_{ai}) = \prod_{i=1}^n x_{bi}(1-x_{bi}).$$

Ne segue che statistica sufficiente minimale per l'inferenza su θ è $\prod_{i=1}^n x_i(1-x_i)$ o, equivalentemente, $\sum_{i=1}^n [\log(x_i) + \log(1-x_i)]$.

(b) Con $n = 1$, consideriamo il sistema d'ipotesi $H_0^* : \theta = \theta_0$ contro $H_1^* : \theta = \theta_1$, con $\theta_1 > \theta_0$ valore fissato. Il test più potente per tale sistema d'ipotesi è il test del rapporto di verosimiglianza (Lemma di Neymann-Pearson), basato sulla statistica

$$\lambda(x_1) = \frac{L(\theta_1)}{L(\theta_0)} \propto \frac{[x_1(1-x_1)]^{\theta_1-1}}{[x_1(1-x_1)]^{\theta_0-1}} = [x_1(1-x_1)]^{\theta_1-\theta_0}.$$

Tale test rifiuta H_0^* per valori grandi di $\lambda(x_1)$. Dato che $\theta_1 > \theta_0$, $\lambda(x_1)$ è funzione monotona crescente di $x_1(1-x_1)$. Quindi, rifiutare per valori grandi di $\lambda(x_1)$ equivale a rifiutare per valori grandi di $x_1(1-x_1) = x_1 - x_1^2$, ossia per $x_1 \in (1/2-k, 1/2+k)$, con k soglia da fissare opportunamente. Poiché la regione $R = \{x_1 : 1/2-k < x_1 < 1/2+k\}$ non dipende dalla particolare alternativa θ_1 fissata, essa è la regione critica più potente per il sistema d'ipotesi $H_0 : \theta = \theta_0$ contro $H_1 : \theta > \theta_0$. Con $\theta_0 = 1$, sotto H_0, X ha distribuzione uniforme su $(0,1)$. Pertanto, il valore di k che permette di ottenere un livello di significatività esatto pari a 0.05 è 0.025.

(c) Sotto H_0, la variabile casuale $I_{(1/2-c,\ 1/2+c)}(X)$ ha distribuzione di Bernoulli di parametro $2c$. Quindi, sempre sotto H_0, $t = \sum_{i=1}^n I_{(1/2-c,\ 1/2+c)}(X_i)$ ha distribuzione binomiale di parametri $(n, 2c)$. Usando l'approssimazione normale alla binomiale, si conclude che $t \dot{\sim} N(2nc, 2nc(1-2c))$ sotto H_0. Con i valori forniti, $t \dot{\sim} N(15, 10.5)$. Quindi, poiché

$$\Pr_{H_0}\left\{\frac{t-15}{\sqrt{10.5}} > 1.64\right\} = \Pr_{H_0}\{t > 1.64\sqrt{10.5}+15\} = 0.05,$$

risulta $k = 1.64\sqrt{10.5}+15 = 20.3$.

(d) Sotto l'alternativa $\theta = 2$, la funzione di densità di X è $f(x;2) = 6x - 6x^2$. Quindi,

$$\Pr_{\theta=2}\{1/2-c < X < 1/2+c\} = \int_{1/2-c}^{1/2+c} (6u - 6u^2)du = 0.4365.$$

Pertanto, quando $\theta = 2$, $t \dot{\sim} N(0.4365n, 0.4365(1-0.4365)n)$, cioè $t \dot{\sim} N(21.8, 12.3)$, e

$$\Pr_{\theta=2}\{t > 20.3\} = \Pr_{\theta=2}\left\{\frac{t-21.8}{\sqrt{12.3}} > \frac{20.3-21.8}{\sqrt{12.3}}\right\} = \Pr_{\theta=2}\left\{\frac{t-21.8}{\sqrt{12.3}} > -0.4277\right\} = 0.665.$$

62. Sia $\mathbf{y} = (y_1, \ldots, y_n)$ realizzazione della variabile casuale multivariata $\mathbf{Y} = (Y_1, \ldots, Y_n)$ con componenti indipendenti e $Y_i \sim N(\theta x_i, 1 + x_i^2)$, dove $\theta \in \Re$ è un parametro ignoto e i valori x_i $(i = 1, \ldots, n)$ sono costanti note, non tutte nulle e tali che

$$\sum_{i=1}^n \frac{x_i^2}{1+x_i^2} = 1\,.$$

Si consideri il problema di verifica d'ipotesi

$$H_0 \,:\, \theta = 0 \qquad \text{contro} \qquad H_1 \,:\, \theta = 1\,.$$

Per $c \in \Re$, sia $R_c = \{\mathbf{y} \in \Re^n \,:\, t(\mathbf{y}) > c\}$, con

$$t(\mathbf{y}) = \sum_{i=1}^n \frac{x_i y_i}{1+x_i^2}\,.$$

(a) Si mostri che una regione critica del tipo R_c minimizza la probabilità di errore del secondo tipo per una fissata probabilità di errore del primo tipo.

(b) Si verifichi se il test con regione di rifiuto R_c è non distorto.

(c) Si mostri che $t(\mathbf{Y})$ è stimatore di massima verosimiglianza per θ. Si stabilisca se tale stimatore è non distorto e consistente.

Soluzione

(a) Per $Y_1, Y_2, \ldots, Y_n$ variabili casuali indipendenti con $Y_i \sim N(\mu_i, \sigma_i^2)$ e varianze note, la funzione di verosimiglianza è

$$\begin{aligned} L(\mu_1, \ldots, \mu_n) &\propto \exp\left\{-(1/2)\sum_{i=1}^{n}(1/\sigma_i^2)(y_i - \mu_i)^2\right\} \\ &\propto \exp\left\{-(1/2)\sum_{i=1}^{n}(\mu_i^2/\sigma_i^2)\right\}\exp\left\{(1/2)\sum_{i=1}^{n}(2y_i\mu_i/\sigma_i^2)\right\}. \end{aligned}$$

Nel caso specifico in cui $\sigma_i^2 = 1 + x_i^2$ e $\mu_i = \theta x_i$, si ha

$$\begin{aligned} L(\theta) &\propto \exp\left\{-(\theta^2/2)\sum_{i=1}^{n}[x_i^2/(1+x_i^2)]\right\}\exp\left\{\theta\sum_{i=1}^{n}[y_i x_i/(1+x_i^2)]\right\} \\ &= \exp(-\theta^2/2)\exp\left\{\theta\sum_{i=1}^{n}[y_i x_i/(1+x_i^2)]\right\}. \end{aligned}$$

Quindi la statistica test TRV (test del rapporto di verosimiglianza) è

$$L(1)/L(0) = \exp(-1/2)\exp\{t(\mathbf{y})\},$$

e rifiutare H_0 per grandi valori della statistica test TRV (test ottimo secondo il Lemma di Neyman-Pearson) equivale a rifiutare per grandi valori di $t(\mathbf{y})$.

(b) Dato che $t(\mathbf{Y})$ è combinazione lineare di variabili casuali normali, ha distribuzione normale. Sotto H_0, $Y_i \sim N(0, 1+x_i^2)$. Quindi, $x_iY_i/(1+x_i^2) \sim N(0, x_i^2/(1+x_i^2))$ e $t(\mathbf{Y}) \sim N(0,1)$. Sotto H_1, $Y_i \sim N(x_i, 1+x_i^2)$. Quindi, $x_iY_i/(1+x_i^2) \sim N(x_i^2/(1+x_i^2), x_i^2/(1+x_i^2))$ e $t(\mathbf{Y}) \sim N(1,1)$. Allora,

$$\alpha_c = Pr\{t(\mathbf{Y}) > c|H_0\} = 1 - \Phi(c)$$

e

$$Pr\{t(\mathbf{Y}) > c|H_1\} = 1 - \Phi(c-1).$$

Per ogni $c \in \Re$ fissato risulta

$$Pr\{t(\mathbf{Y}) > c|H_1\} = 1 - \Phi(c-1) > 1 - \Phi(c) = Pr\{t(\mathbf{Y}) > c|H_0\} = \alpha_c,$$

quindi il test è non distorto.

(c) In generale, $Y_i \sim N(\theta x_i, 1+x_i^2)$. Quindi, $x_iY_i/(1+x_i^2) \sim N(\theta x_i^2/(1+x_i^2), x_1^2/(1+x_i^2))$ e $t(\mathbf{Y}) \sim N(\theta, 1)$, qualunque sia la dimensione campionaria n. D'altro canto,

$$l(\theta) = \log(L(\theta)) = -\theta^2/2 + \theta\sum_{i=1}^{n}[y_i x_i/(1+x_i^2)],$$

$$l_*(\theta) = \frac{dl(\theta)}{d\theta} = -\theta + \sum_{i=1}^{n}[y_i x_i/(1+x_i^2)],$$

cosicché lo stimatore di massima verosimiglianza è $\hat{\theta} = t(\mathbf{Y})$. Pertanto, dato che $t(\mathbf{Y}) \sim N(\theta, 1)$, lo stimatore di massima verosimiglianza per θ è non distorto ma non è consistente.

63. Siano $x_1, \ldots, x_n$ determinazioni i.i.d. di una variabile casuale X con distribuzione uniforme sull'intervallo $(0, \theta)$.

(a) Si mostri che la famiglia di distribuzioni di Pareto, con parametri $\alpha > 0$, $\beta > 0$ e funzione di densità $f(y; \alpha, \beta) = \alpha \frac{\beta^\alpha}{y^{\alpha+1}}$ per $y \geq \beta$, è coniugata naturale alla verosimiglianza per θ.

(b) Supponendo di osservare il campione $5.54, 1.14, 1.89, 3.94, 1.53, 4.17, 3.09, 4.22, 2.64, 2.06$ e di scegliere, nella famiglia coniugata, la *a priori* di parametri $\alpha_0 = 2$ e $\beta_0 = 1$, si verifichi l'ipotesi $5.8 < \theta < 6$.

(c) Con l'osservazione campionaria e la scelta della distribuzione *a priori* di cui al punto (b), si fornisca, come valutazione bayesiana di θ, la mediana *a posteriori*.

Soluzione

(a) Si ha

$$L(\mathbf{x}; \theta) = \frac{1}{\theta^n} I_{(x_{(n)}, +\infty)}(\theta)$$

dove $I(\cdot)$ è la funzione indicatrice e $x_{(n)}$ è il massimo delle osservazioni x_i. La densità a priori è

$$\pi(\theta) = \alpha \frac{\beta^\alpha}{\theta^{\alpha+1}} I_{(\beta, +\infty)}(\theta).$$

Quindi, per la densità a posteriori vale la relazione

$$\begin{aligned} \pi(\theta|\mathbf{x}) &\propto L(\mathbf{x}; \theta)\pi(\theta) \\ &\propto \frac{1}{\theta^{\alpha+n+1}} I_{(x_{(n)}, +\infty)}(\theta) I_{(\beta, +\infty)}(\theta) \\ &\propto \frac{1}{\theta^{\alpha+n+1}} I_{(\max\{x_{(n)}, \beta\}, +\infty)}(\theta). \end{aligned}$$

Ne segue che $\theta|\mathbf{x}$ ha distribuzione di Pareto con parametri $\alpha' = \alpha + n$ e $\beta' = \max\{x_{(n)}, \beta\}$.

(b) Nelle ipotesi fatte e con i dati forniti, $\alpha' = 12$, $x_{(n)} = 5.54$ e $\beta' = 5.54$. Allora, per $t > 5.54$, si ha,

$$\begin{aligned} \Pr\{\theta > t|\mathbf{x}\} &= \int_t^{+\infty} \alpha' \beta'^{\alpha'} \theta^{-\alpha'-1} d\theta \\ &= \beta'^{\alpha'} \theta^{-\alpha'} \Big|_t^{+\infty} \\ &= (\beta'/t)^{\alpha'} = (5.54/t)^{12}. \end{aligned}$$

Quindi,

$$\Pr\{5.8 < \theta < 6|\mathbf{x}\} = \Pr\{\theta > 5.8|\mathbf{x}\} - \Pr\{\theta > 6|\mathbf{x}\} = 0.193.$$

Ciò porta a rifiutare l'ipotesi formulata.

(c) La mediana di una distribuzione di Pareto di parametri α e β è la soluzione dell'equazione in t

$$\frac{1}{2} = \int_t^{+\infty} \alpha\beta^\alpha \theta^{-\alpha-1} d\theta = \alpha\beta^\alpha \left. \frac{\theta^{-\alpha}}{-\alpha} \right|_t^{+\infty} = \beta^\alpha t^{-\alpha},$$

da cui $t = (2\beta^\alpha)^{1/\alpha} = 2^{1/\alpha}\beta$. Quindi, la mediana a posteriori è

$$Me(\theta|\mathbf{x}) = 2^{1/\alpha'}\beta' = 5.54 \sqrt[12]{2} = 5.72.$$

64. In un esperimento si spara con un fucile contro un bersaglio di forma circolare. Si suppone che, rispetto al riferimento ideale costituito da un sistema di assi cartesiani che hanno origine nel centro del bersaglio, le coordinate dei punti colpiti siano distribuite secondo una variabile casuale $\mathcal{N}(0, \sigma^2)$ e siano indipendenti tra di loro. Allo scopo di stimare σ^2 si sparano un certo numero di colpi, e si registra il punto esatto che viene colpito solo se la distanza dal centro è minore di c, raggio del bersaglio, cioè solo se il bersaglio è centrato. Altrimenti, si sa che il colpo è andato a vuoto. Si spara sino a che si osservano m (valore fissato) colpi nel bersaglio. Siano dunque (x_i, y_i) $i = 1, \ldots, m$ le coordinate dei punti che hanno colpito il bersaglio e sia n il numero totale di colpi sparati. [20]

(a) Si mostri che la probabilità di non centrare il bersaglio è pari a $e^{-c^2/2\sigma^2}$.

(b) Posto che per la funzione di verosimiglianza per σ^2 vale l'espressione

$$L(\sigma^2; \mathbf{x}, \mathbf{y}, n) \quad \propto \quad e^{-(n-m)c^2/2\sigma^2} \prod_{i=1}^{m} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} x_i^2\right\} \exp\left\{-\frac{1}{2\sigma^2} y_i^2\right\},$$

si individui una statistica sufficiente minimale.

(c) Si ottenga lo stimatore di massima verosimiglianza per σ^2.

(d) Si dica se lo stimatore di massima verosimiglianza di cui al punto precedente è corretto, sapendo che

$$E(X^2 + Y^2) = 2\sigma^2 - c^2 e^{-c^2/2\sigma^2}/(1 - e^{-c^2/2\sigma^2}),$$

con (X, Y) variabile casuale di cui sono determinazioni le osservazioni (x_i, y_i), $i = 1, \ldots, m$.

Soluzione

(a) Indichiamo con x' e y' le coordinate (eventualmente non osservate) del generico (tra gli n) colpo sparato. Essendo $(X'^2 + Y'^2)/\sigma^2$ distribuito secondo una varaibile casuale χ^2_2, $Z' = X'^2 + Y'^2$ ha distribuzione esponenziale di parametro $1/(2\sigma^2)$, con funzione di densità

$$f(z) = \frac{1}{2\sigma^2} e^{-z/2\sigma^2}.$$

Perciò, $\Pr(X'^2 + Y'^2 > c^2) = e^{-c^2/2\sigma^2}$.

(b) Il rapporto tra le verosimiglianze relative a due punti diversi dello spazio campionario è

$$\frac{L(\sigma^2; \mathbf{x}, \mathbf{y}, n)}{L(\sigma^2; \mathbf{x}^*, \mathbf{y}^*, n^*)} \quad \propto \quad \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}(x_i^2 + y_i^2) - \sum_{i=1}^{m}(x_i^{*2} + y*_i^{*2})\right]\right\} (e^{-c^2/2\sigma^2})^{n-n*}$$

sicché la statistica sufficiente minimale è $\sum_{i=1}^{m}(x_i^2 + y_i^2) + c^2 n$.

(c) Posto $S = \sum_{i=1}^{m}(x_i^2 + y_i^2)$, la funzione di verosimiglianza ha espressione

$$L(\sigma^2; S, n) \propto \frac{1}{(\sigma^2)^m} e^{-\frac{1}{2\sigma^2}S} e^{-c^2(n-m)/2\sigma^2}.$$

Si ottiene quindi la funzione di log-verosimiglianza

$$l(\sigma^2; S, n) = -m \log(\sigma^2) - \frac{1}{2\sigma^2}(S + c^2(n - m)),$$

[20] Si osservi che, in questo caso, n è determinazione di una variabile casuale (diciamo N).

la cui derivata

$$l_*(\sigma^2; S, n) = -m\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\left(S + c^2(n-m)\right)$$

è nulla in

$$\hat{\sigma}^2 = \frac{S + c^2(n-m)}{2m}.$$

In tale punto la derivata seconda è negativa: $\hat{\sigma}^2$ è quindi un massimo locale. La funzione è inoltre derivabile ed è facile vedere che il limite di L è 0, sia se $\sigma^2 \to 0$ sia che se $\sigma^2 \to \infty$. Quindi $\hat{\sigma}^2$ è punto di massimo assoluto.

(d) Indicando con $Z_i = X_i^2 + Y_i^2$, si ha

$$\begin{aligned} E(\hat{\sigma}^2) &= E\left(\frac{\sum_{i=1}^m Z_i + c^2(N-m)}{2m}\right) \\ &= \frac{E(Z_1)}{2} + \frac{c^2}{2m}E(N-m). \end{aligned}$$

$E(Z_1)$ è dato nel testo, mentre N è binomiale negativa, con media $m/(1-e^{-c^2/2\sigma^2})$. Si ha perciò

$$E(\hat{\sigma}^2) = \sigma^2 + \frac{1}{1-e^{-c^2/2\sigma^2}}\left(-\frac{c^2}{2}e^{-c^2/2\sigma^2} + \frac{c^2}{2}e^{-c^2/2\sigma^2}\right) = \sigma^2.$$

Lo stimatore è dunque non distorto.

65. Sia $x_1, x_2, \ldots, x_n$ un campione casuale semplice da una variabile casuale X con distribuzione normale di media μ e varianza σ^2. Sia $\tau = \mu + 2\sigma$ un quantile della distribuzione di X di interesse. Si supponga $n = 20$ e che l'osservazione campionaria sia tale che $\sum_{i=1}^{20} x_i = 84.2$ e $\sum_{i=1}^{20} x_i^2 = 375.4$.

(a) Si costruisca un intervallo di confidenza per τ di livello di copertura approssimato 0.95.

(b) Si verifichi, ad un livello di significatività (approssimato) del 5%, l'ipotesi $H_0 : \tau = 8$ contro $H_1 : \tau > 8$.

(c) Si calcoli la potenza del test di cui al punto (b) per $\tau = 9$.

(d) Si supponga $\sigma^2 = 1$ noto. Si costruisca un intervallo di confidenza per μ (di livello, eventualmente approssimato, 0.95) che sia valido anche nel caso in cui la vera distribuzione di X risultasse appartenere alla famiglia $\{f(x;\mu) = (1/2)\exp(-|x-\mu|),\ x \in \Re,\ \mu \in \Re\}$.

Soluzione

(a) Al campione casuale semplice $x_1, \ldots, x_n$ da una normale di parametri μ e σ^2 corrispondono le stime di massima verosimiglianza $\hat{\mu} = \bar{x}$ e $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^n x_i^2 - \bar{x}^2$, con $\bar{x} = (1/n)\sum_{i=1}^n x_i$. Il parametro di interesse è legato ai parametri μ e σ^2 dalla relazione $\tau = \mu + 2\sigma$. Per la proprietà di equivarianza (immaginando di riparametrizzare), la stima di massima verosimiglianza è

$$\hat{\tau} = \hat{\mu} + 2\hat{\sigma}.$$

Dato che $(\hat{\mu}, \hat{\sigma}^2)$ è asintoticamente normale,

$$((\hat{\mu}, \hat{\sigma}^2)^T - (\mu, \sigma^2)^T) \dot{\sim} N_2(0, I(\mu, \sigma^2)^{-1}),$$

con

$$I_{\mu,\sigma^2}(\mu, \sigma^2) = \begin{bmatrix} n/\sigma^2 & 0 \\ 0 & n/(2\sigma^4) \end{bmatrix}$$

(Azzalini, pagina 82) e che, in base ad uno sviluppo di Taylor della funzione $\sqrt{x}$, vale la relazione (approssimazione)

$$\sqrt{\hat{\sigma}^2} \doteq \sqrt{\sigma^2} + \frac{1}{2\sqrt{\sigma^2}}(\hat{\sigma^2} - \sigma^2),$$

si ha che $\hat{\tau}$ è approssimativamente normale con media τ e varianza $\sigma_\tau^2 = \frac{\sigma^2}{n} + 2^2(\frac{2\sigma^4}{n}\frac{1}{4\sigma^2}) = \frac{3\sigma^2}{n}$. Quindi, con i dati forniti, si ha $\hat{\mu} = 4.21$, $\hat{\sigma}^2 = 1.046$ e

$$(\hat{\tau} - \tau) \dot{\sim} N(0, 0.157).$$

Inoltre, la stima di massima verosimiglianza per τ è 6.25. Ne segue che l'intervallo di confidenza cercato per τ è

$$\{\tau : \hat{\tau} - 1.96\hat{\sigma}_\tau < \tau < \hat{\tau} + 1.96\hat{\sigma}_\tau\}$$

con $\hat{\tau} = 6.25$ e $\hat{\sigma}_\tau = \sqrt{0.157} = 0.396$. In definitiva, l'intervallo cercato è (5.47, 7.02).

(b) L'ipotesi nulla è rifiutata se

$$\frac{\hat{\tau} - \tau_0}{\hat{\sigma}_\tau} > 1.64,$$

con $\tau_0 = 8$. Poiché $\hat{\tau} - \tau_0 < 0$, l'ipotesi nulla non può essere rifiutata al livello approssimato del 5%.

(c) Per la fissata alternativa $\tau = 9$ si ha

$$\Pr_{\tau=9}\{(\hat{\tau} - \tau_0)/\hat{\sigma}_\tau > 1.64\} = \Pr_{\tau=9}\{(\hat{\tau} - 9)/\hat{\sigma}_\tau > 1.64 - 1/\hat{\sigma}_\tau\} = 1 - \Phi(-0.885) = 0.812.$$

(d) Il parametro μ rappresenta la media anche per la famiglia $\mathcal{F}$ data da

$$\{f(x; \mu) = (1/2)\exp(-|x - \mu|),\ x \in \Re,\ \mu \in \Re\}.$$

Essendo lo score di verosimiglianza (relativo la modello parametrico fissato) funzione di stima non distorta anche sotto $\mathcal{F}$, l'intervallo di confidenza per μ può essere ottenuto a partire dallo stimatore di massima verosimiglianza $\hat{\mu}$. In questa situazione, una stima consistente (anche sotto $\mathcal{F}$) per la varianza asintotica di $\hat{\mu} = \bar{x}$ è data da (stimatore *sandwich*)

$$\hat{\sigma}^2_{\hat{\mu}} = \frac{1}{j(\hat{\mu})^2}\sum_{i=1}^{20} l_*(\hat{\mu}; x_i)^2,$$

dove $l_*(\mu; x_i) = x_i - \mu$ e $j(\mu) = n$ sono lo score e l'informazione osservata al modello parametrico fissato. Ne segue che $\hat{\sigma}^2_{\hat{\mu}} = (1/n^2)\sum_i(x_i - \bar{x})^2 = \hat{\sigma}^2/n$, dove $\hat{\sigma}^2 = 1.046$ è la varianza campionaria. Pertanto, l'intervallo cercato è

$$\left\{\mu : \hat{\mu} - 1.96\frac{\hat{\sigma}}{n} < \mu < \hat{\mu} - 1.96\frac{\hat{\sigma}}{n}\right\},$$

cioè [4.11, 4.31].

66. Nell'ultimo appello d'esame, hanno consegnato l'elaborato 116 studenti, dei quali 58 iscritti ad un corso di laurea del vecchio ordinamento (v.o.). Il punteggio cumulato conseguito dagli elaborati degli studenti v.o. è risultato essere pari a 720.5, mentre il punteggio cumulato relativo alla totalità degli elaborati è risultato essere 1446.5. Si indichino con X e Y le variabili casuali che descrivono il punteggio conseguito dal generico studente v.o. e dal generico studente del nuovo ordinamento (n.o.), rispettivamente. Si assuma che sia $X \sim N(\mu_x, 24.6)$, $Y \sim N(\mu_y, 36.9)$ e che ci sia indipendenza tra gli esiti dei vari elaborati. Si consideri per (μ_x, μ_y) una distribuzione a priori normale bidimensionale, con marginali indipendenti entrambe $N(13, 25)$.

(a) Si ricavi la distribuzione a posteriori per (μ_x, μ_y).

(b) Posto $\delta = \mu_y - \mu_x$, si verifichi l'ipotesi $\delta > 0$.

(c) Si ricavi un intervallo di credibilità HPD al 95% per δ.

(d) Si dia una valutazione bayesiana della probabilità di successo (punteggio ≥ 18) alla prossima prova d'esame, per gli studenti v.o. e n.o..

Soluzione

(a) Dato che le variabili X e Y sono indipendenti e che le due distribuzioni a priori marginali sono indipendenti e coniugate naturali alle verosimiglianze associate ai campioni $x_1, \dots, x_n$ e $y_1, \dots, y_n$ (con $n = 58$), rispettivamente, la distribuzione a posteriori per (μ_x, μ_y) è normale bidimensionale, con marginali ancora indipendenti. In particolare, si ha $\mu_x|\mathbf{x} \sim N(\mu_x^*, \sigma_x^{*2})$ e $\mu_y|\mathbf{y} \sim N(\mu_y^*, \sigma_y^{*2})$, con

$$\mu_x^* = \frac{13 \times 24.6 + n\bar{x} \times 25}{24.6 + n \times 25} = 12.432 \qquad \sigma_x^{*2} = \frac{24.6 \times 25}{24.6 + n \times 25} = 0.417$$

e

$$\mu_y^* = \frac{13 \times 36.9 + n\bar{y} \times 25}{36.9 + n \times 25} = 12.53 \qquad \sigma_y^{*2} = \frac{36.9 \times 25}{36.9 + n \times 25} = 0.62,$$

essendo $n\bar{x} = 720.5$ e $n\bar{y} = 1446.5 - 720.5 = 726$.

(b) Evidentemente, $\delta|(\mathbf{x}, \mathbf{y}) \sim N(\mu_y^* - \mu_x^*,\ \sigma_y^{*2} + \sigma_x^{*2})$. Poiché $\mu_y^* - \mu_x^* > 0$, risulta $\Pr\{\delta > 0|\mathbf{x}, \mathbf{y}\} > 0.5$. Più precisamente,

$$\Pr\{\delta > 0|\mathbf{x}, \mathbf{y}\} = \Phi((\mu_y^* - \mu_x^*)/\sqrt{\sigma_y^{*2} + \sigma_x^{*2}}) = \Phi(0.098/1.0183) = 0.538.$$

L'ipotesi formulata è dunque accettata.

(c) L'intervallo HPD al 95% per δ ha per estremi i punti $0.098 - 1.96 \times 1.0183$ e $0.098 + 1.96 \times 1.0183$. Risulta quindi essere [-1.898, 2.094].

(d) Poiché $X|\mu_x \sim N(\mu_x, 24.6)$ e $\mu_x|(\mathbf{x}, \mathbf{y}) \sim N(\mu_x^*, \sigma_x^{*2})$, si ha che (nota 7, pag. xxviii)

$$X|(\mathbf{x}, \mathbf{y}) \sim N(\mu_x^*, 24.6 + \sigma_x^{*2}).$$

Analogamente, risulta $Y|(\mathbf{x}, \mathbf{y}) \sim N(\mu_y^*, 36.9 + \sigma_y^{*2})$. Quindi,

$$\Pr\{X|(\mathbf{x}, \mathbf{y}) \geq 18\} = 1 - \Phi((18 - 12.432)/\sqrt{24.6 + 0.417}) = 0.133$$

e

$$\Pr\{Y|(\mathbf{x}, \mathbf{y}) \geq 18\} = 1 - \Phi((18 - 12.53)/\sqrt{36.9 + 0.62}) = 0.186.$$

67. Sia $y_1, \dots, y_n$ un campione casuale da una variabile Y a valori in $\in \Re$ con funzione di densità $f(y;\gamma) = g(y-\gamma)$, dove

$$g(u) = \begin{cases} C(k)e^{-u^2/2} & \text{se } |u| \le k \\ C(k)e^{k^2/2-k|u|} & \text{se } |u| > k \end{cases}$$

con $k > 0$ costante fissata, $C(k)$ opportuna costante di normalizzazione e $\gamma \in \Re$ parametro ignoto.

(a) Si ricavi lo stimatore per γ basato sul metodo dei momenti.

(b) Si fornisca una statistica sufficiente per l'inferenza su γ (che non sia quella banale costituita dall'intera osservazione $y_1, \dots, y_n$).

(c) Si stabilisca se lo stimatore di massima verosimiglianza $\hat{\gamma}$ (definito come radice dell'equazione di verosimiglianza) è robusto al modello parametrico considerato.

(d) Quale è il funzionale statistico associato a $\hat{\gamma}$?

Soluzione

(a) Si osservi che la funzione $g(u)$ è simmetrica rispetto a 0. Ne segue che la funzione di densità $f(y;\gamma)$ è simmetrica rispetto a γ. Quindi γ è la media di Y. Pertanto, lo stimatore per γ basato sul metodo dei momenti è, semplicemente, la media campionaria $\bar{y}$.

(b) Poiché siamo sotto campionamento casuale semplice, una statistica sufficiente per γ, che ha una certa capacità di sintesi rispetto alla semplice osservazione $y_1, \dots, y_n$, è la statistica d'ordine $y_{(1)}, \dots, y_{(n)}$. In effetti, due punti dello spazio campionario, equivalenti in termini di statistica d'ordine, generano la stessa funzione di verosimiglianza per γ.

(c) Il contributo alla verosimiglianza della generica osservazione y_i è

$$\propto e^{-(y_i-\gamma)^2/2}, \quad \text{se} \quad -k \le y_i - \gamma \le k,$$

oppure

$$\propto e^{k^2/2-k(y_i-\gamma)}, \quad \text{se} \quad y_i - \gamma > k,$$

o

$$\propto e^{k^2/2+k(y_i-\gamma)}, \quad \text{se} \quad y_i - \gamma < -k.$$

Il contributo alla log-verosimiglianza può dunque essere $-(y_i-\gamma)^2/2$, o $k^2/2 - k(y_i-\gamma)$, o $k^2/2 + k(y_i - \gamma)$. Ne segue che lo score di verosimiglianza ha espressione $l_*(\gamma, \mathbf{y}) = \sum_{i=1}^n h(\gamma, y_i)$, con

$$h(\gamma, y_i) = \begin{cases} -k & \text{se } y_i - \gamma < -k \\ y_i - \gamma & \text{se } -k \le y_i - \gamma \le k \\ k & \text{se } y_i - \gamma > k. \end{cases}$$

Dato che la funzione h è limitata in y_i, lo stimatore di massima verosimiglianza è robusto al modello considerato.

(d) Il funzionale statistico associato a $\hat{\gamma}$ è $T(F)$ definito implicitamente dall'equazione

$$\int h(T, y) dF(y) = 0.$$

68. Sia $x_1, \ldots, x_n$ un campione casuale semplice da una variabile X continua, con densità $f(x;\theta) = \theta(x+1)^{-(\theta+1)}$, per $x \geq 0$ e $\theta > 0$ parametro ignoto. Si indichi con F la funzione di ripartizione di X.

(a) Si mostri che la variabile $Z = -2\log(1 - F(X))$ ha distribuzione χ^2_2.

(b) Si usi il risultato di cui al punto (a) per costruire un intervallo di confidenza per θ, di livello esatto 0.95.

(c) Si stabilisca se lo stimatore di massima verosimiglianza per θ è ottimo secondo la teoria classica.

(d) Si ottenga la regione critica più potente di livello $\alpha = 0.05$ per il problema di verifica d'ipotesi $H_0 : \theta = \theta_0$ contro $H_1 : \theta > \theta_0$.

Soluzione

(a) Si ha

$$\Pr\{Z \leq z\} = \Pr\{-2\log(1 - F(X)) \leq z\} = \Pr\{F(X) \leq 1 - e^{-z/2}\} = 1 - e^{-z/2},$$

essendo, com'è noto, $F(X) \sim U(0,1)$. Quindi Z ha distribuzione esponenziale di parametro $1/2$ (media 2), ovvero $Z \sim Gamma(1, 1/2)$ o, equivalentemente, $Z \sim \chi^2_2$.

(b) In base al risultato di cui al punto (a), si ha che $-2\sum_{i=1}^n \log(1-F(x_i)) \sim Gamma(n, 1/2) \equiv \chi^2_{2n}$. D'altra parte,

$$F(x) = \int_0^x \theta(t+1)^{-(\theta+1)}dt = \int_1^{x+1} u^{-(\theta+1)}du = 1 - (x+1)^{-\theta}.$$

Quindi

$$-2\sum_{i=1}^n \log(1 - F(x_i)) = -2\sum_{i=1}^n \log(x_i + 1)^{-\theta} = 2\theta \sum_{i=1}^n \log(x_i + 1) \sim \chi^2_{2n}.$$

Pertanto, un intervallo di confidenza di livello esatto 0.95 per θ è dato da

$$\frac{k_{0.025}}{2\sum_i \log(x_i + 1)} < \theta < \frac{k_{0.975}}{2\sum_i \log(x_i + 1)},$$

dove k_α indica il quantile di ordine α della distribuzione χ^2_{2n}.

(c) Si ha

$$L(\theta) = \prod_{i=1}^n \theta(x_i + 1)^{-(\theta+1)} = \theta^n \prod_{i=1}^n (x_i + 1)^{-(\theta+1)}.$$

Quindi, $l(\theta) = n\log\theta - (\theta + 1)\sum_{i=1}^n \log(x_i + 1)$ e $l_*(\theta) = n/\theta - \sum_{i=1}^n \log(x_i + 1)$. Ne segue che

$$\hat{\theta} = \frac{n}{\sum_{i=1}^n \log(x_i + 1)}.$$

Inoltre, dalla prima identità di Bartlett,

$$E_\theta[l_*(\theta)] = n/\theta - E_\theta\left[\sum_{i=1}^n \log(x_i + 1)\right] = 0,$$

risulta che $E_\theta[\sum_{i=1}^n \log(x_i+1)] = n/\theta$. Allora, utilizzando la diseguaglianza di Jensen,

$$E_\theta[\hat{\theta}] = E_\theta\left[\frac{n}{\sum_{i=1}^n \log(x_i+1)}\right] \neq \frac{n}{E_\theta[\sum_{i=1}^n \log(x_i+1)]} = \theta.$$

Pertanto, lo stimatore di massima verosimiglianza è distorto e non può quindi essere ottimo secondo la teoria classica.

(d) È facile verificare che la famiglia parametrica considerata costituisce una famiglia esponenziale regolare di ordine 1 con statistica canonica (quindi sufficiente, minimale e completa) $t(\mathbf{x}) = \sum_{i=1}^n \log(x_i+1)$. Pertanto, il test ottimo per il problema di verifica d'ipotesi considerato ha regione critica $\mathcal{R}$ basata su $t(\mathbf{x})$ o opportuna trasformata. Sappiamo che, sotto H_0, $2\theta_0 t(\mathbf{x})$ ha distribuzione χ^2_{2n}. Inoltre, $t(\mathbf{x}) = n/\hat{\theta}$ e valori grandi della stima di massima verosimiglianza sono, evidentemente, contrari all'ipotesi nulla. Ne segue che la regione critica più potente cercata è

$$\mathcal{R} = \{\mathbf{x} : 2\theta_0 t(\mathbf{x}) < k_{0.05}\}.$$

69. Un quiz televisivo prevede tre possibili livelli di vincita in denaro, oltre che, naturalmente, un eventuale gadget come premio di consolazione. Gli ideatori del programma ritengono che le probabilità che un concorrente arrivi alla vincita siano β, $\beta/2$ e $\beta/4$ (rispettivamente per i tre livelli) e in una serie pilota del programma, con n concorrenti partecipanti, si registrano x_1, x_2 e x_3 vincite di primo, di secondo e di terzo livello, rispettivamente. Si scelga per β una distribuzione a priori uniforme sull'intervallo $(0, \frac{4}{7})$.

(a) Si ricavi, a meno di una costante di normalizzazione, la densità della distribuzione a posteriori di β.

(b) Supponendo che sia $x_1 + x_2 + x_3 > 0$, si fornisca una stima puntuale bayesiana per β.

(c) Si stabilisca la forma degli intervalli di credibilità HPD per β nel caso in cui l'osservazione campionaria fosse tale che $x_1 + x_2 + x_3 = 0$.

Soluzione

(a) Per la distribuzione a priori si ha $\pi(\beta) \propto I_{(0,4/7)}(\beta)$. Per la funzione di verosimiglianza, associata all'osservazione $(x_1, x_2, x_3, n - x_1 - x_2 - x_3)$, vale la relazione

$$L(\beta) \propto \left(1 - \frac{7\beta}{4}\right)^{n-x_1-x_2-x_3} \beta^{x_1} \left(\frac{\beta}{2}\right)^{x_2} \left(\frac{\beta}{4}\right)^{x_3} \propto (4-7\beta)^{n-s}\beta^s,$$

dove si è posto $s = \sum_{i=1}^3 x_i$. Quindi, per la distribuzione a posteriori otteniamo l'espressione

$$\pi(\beta|\mathbf{x}) \propto (4-7\beta)^{n-s}\beta^s I_{(0,4/7)}(\beta).$$

(b) Sia $s > 0$. Poniamo $h(\beta) = (4-7\beta)^{n-s}\beta^s I_{(0,4/7)}(\beta)$. Derivando la funzione $h(\beta)$ sull'intervallo $(0, 4/7)$, si ha

$$\begin{aligned} h'(\beta) &= \frac{dh(\beta)}{d\beta} = -7(n-s)(4-7\beta)^{n-s-1}\beta^s + s\beta^{s-1}(4-7\beta)^{n-s} \\ &= (4-7\beta)^{n-s}\left[s\beta^{s-1} - \frac{7(n-s)\beta^s}{4-7\beta}\right] = (4-7\beta)^{n-s-1}(4s\beta^{s-1} - 7n\beta^s) \\ &= (4-7\beta)^{n-s-1}\beta^{s-1}(4s - 7n\beta). \end{aligned}$$

Tale funzione è zero quando $4s - 7n\beta = 0$, ossia quando $\beta = \beta_0 = \frac{4s}{7n}$. Inoltre, risulta $h'(\beta) > 0$ se $\beta < \beta_0$ e $h'(\beta) < 0$ se $\beta > \beta_0$. Quindi, la funzione $h(\beta)$, nell'intervallo $(0, 4/7)$, è cresente per $\beta < \beta_0$ ed è decrescente per $\beta > \beta_0$. Pertanto, β_0 è la moda della distribuzione a posteriori e costituisce una stima bayesiana di β.

(c) Se fosse $s = 0$, si avrebbe

$$\pi(\beta|\mathbf{x}) \propto (4 - 7\beta)^n I_{(0,4/7)}(\beta).$$

In questo caso, la funzione $h(\beta) = (4 - 7\beta)^n I_{(0,4/7)}(\beta)$ risulterebbe monotona decrescente nell'intervallo $(0, 4/7)$. Di conseguenza, gli intervalli di credibilità HPD sarebbero tutti intervalli di tipo $(0, c)$, con c valore da fissare opportunamente (in base al livello di credibilità richiesto).

70. Sia $y_1, \dots, y_n$ un campione casuale semplice di dimensione $n > 1$ da una variabile discreta Y, con supporto $\{0, 1, 2, 3, \dots\}$. Si assuma che Y abbia legge geometrica, con parametro $\theta \in (0, 1)$, funzione di probabilità $p(y; \theta) = \theta(1 - \theta)^y$ e media $(1 - \theta)/\theta$.

(a) Si utilizzi il risultato di Rao-Blackwell per ottenere uno stimatore non distorto per θ che abbia varianza non superiore a quella di $T = I_{\{0\}}(y_1)$.

(b) Si può affermare che quello individuato al punto (a) è lo stimatore ottimo per θ ?

(c) Si ottenga un intervallo di confidenza di livello approssimato 0.95 per la quantità $\tau = \Pr\{Y \leq 1\}$.

(d) Si stabilisca se lo stimatore di massima verosimiglianza $\hat{\theta}$, ottenuto sotto il modello parametrico considerato, risulta stimatore consistente per la $\Pr\{Y = 0\}$ anche se la vera legge di Y è un elemento della classe di distribuzioni di Poisson?

Soluzione

(a) La classe parametrica $\{p(y; \theta),\ \theta \in (0, 1)\}$ costituisce una famiglia esponenziale regolare monoparametrica. Sotto campionamento casuale semplice, la statistica canonica è $S = \sum_{i=1}^n y_i$, che è statistica sufficiente minimale e completa. Essendo $\theta = \Pr\{Y = 0\}$, T è, evidentemente, stimatore non distorto per θ. Bisogna, allora, calcolare

$$\begin{aligned}
E[T|S = s] &= \Pr\{T = 1|S = s\} = \Pr\{y_1 = 0|S = s\} \\
&= \frac{\Pr\{(y_1 = 0) \cap (S = s)\}}{\Pr\{S = s\}} = \frac{\Pr\{S = s|y_1 = 0\} \Pr\{y_1 = 0\}}{\Pr\{S = s\}} \\
&= \frac{\Pr\{\sum_{i=2}^n y_i = s\} \Pr\{y_1 = 0\}}{\Pr\{\sum_{i=n}^n y_i = s\}} = \frac{\binom{s+n-2}{n-2}\theta^{n-1}(1-\theta)^s\theta}{\binom{s+n-1}{n-1}\theta^n(1-\theta)^s} \\
&= \frac{\frac{(s+n-2)!}{(n-2)!s!}}{\frac{(s+n-1)!}{(n-1)!s!}} = \frac{n-1}{n+s-1},
\end{aligned}$$

tenendo conto che la somma di k variabili casuali geometriche di parametro θ indipendenti ha distribuzione binomiale negativa di parametri k e θ. Lo stimatore cercato è, dunque,

$$\frac{n-1}{n+S-1}.$$

(b) Dato che lo stimatore trovato al punto precedente è funzione della statistica sufficiente minimale completa S (ed è non distorto per costruzione), esso è lo stimatore ottimo per l'inferenza su θ.

(c) Si osservi che $\tau = \Pr\{Y \leq 1\} = \Pr\{Y = 0\} + \Pr\{Y = 1\} = \theta + \theta(1-\theta) = 2\theta - \theta^2$. Perciò, $\tau = g(\theta)$, con $g'(\theta) = dg(\theta)/d\theta = 2 - 2\theta > 0$ per ogni valore di $\theta \in (0, 1)$. Quindi, τ è funzione monotona crescente di θ.
D'altro canto, la funzione di verosimiglianza ha espressione

$$L(\theta) = \prod_{i=1}^{n} [\theta(1-\theta)^{y_i}] = \theta^n (1-\theta)^s.$$

Quindi,

$$l(\theta) = \log(L(\theta)) = n \log(\theta) + s \log(1-\theta),$$

e, derivando rispetto a θ,

$$l_*(\theta) = \frac{n}{\theta} - \frac{s}{1-\theta}.$$

Inoltre, per la derivata seconda della log-verosimiglianza si ha

$$l_{**}(\theta) = -\frac{n}{\theta^2} - \frac{s}{(1-\theta)^2}$$

e risulta

$$i(\theta) = -E[l_{**}(\theta)] = \frac{n}{\theta^2} + \frac{n}{\theta(1-\theta)} = \frac{n}{\theta^2(1-\theta)},$$

dato che $E(Y) = (1-\theta)/\theta$. Allora, per lo stimatore di massima verosimiglianza $\hat{\theta}$ vale l'approssimazione $(\hat{\theta} - \theta)\dot{\sim} N(0, \theta^2(1-\theta)/n)$, e un intervallo di confidenza per θ, di livello approssimato 0.95, ha per estremi i valori $\hat{\theta} - 1.96\sqrt{\hat{\theta}^2(1-\hat{\theta})/n}$ e $\hat{\theta} + 1.96\sqrt{\hat{\theta}^2(1-\hat{\theta})/n}$.
Di conseguenza, un intervallo di confidenza per τ, di livello approssimato 0.95, ha per estremi i valori $g\left(\hat{\theta} - 1.96\sqrt{\hat{\theta}^2(1-\hat{\theta})/n}\right)$ e $g\left(\hat{\theta} + 1.96\sqrt{\hat{\theta}^2(1-\hat{\theta})/n}\right)$.

(d) Sotto il modello parametrico considerato (distribuzione marginale geometrica), lo stimatore di massima verosimiglianza $\hat{\theta} = n/(n+S)$ è stimatore consistente per $\theta = \Pr\{Y = 0\}$. Se la vera legge di Y fosse un elemento della classe di distribuzioni di Poisson, diciamo di parametro λ, avremmo $\Pr\{Y = 0\} = e^{-\lambda}$. D'altro canto, $S/n = \sum_i y_i/n$ sarebbe (Legge dei Grandi Numeri) stimatore consistente per $E(Y) = \lambda$. Allora, essendo $\hat{\theta}$ funzione continua di S/n, si avrebbe, al crescere di n,

$$\hat{\theta} = \frac{1}{1 + \frac{S}{n}} \to \frac{1}{1+\lambda} \neq e^{-\lambda},$$

in probabilità. Quindi, $\hat{\theta}$ non risulterebbe stimatore consistente per $\Pr\{Y = 0\}$.

71. Sia $y_1, y_2, \ldots, y_n$ un campione casuale semplice da una variabile Y. Si consideri il modello statistico $\mathcal{F}$ che prevede che la legge di Y sia un elemento della classe parametrica normale $N(\lambda, \lambda)$, con funzione di densità

$$f(y; \lambda) = \frac{1}{\sqrt{2\pi\lambda}} \exp\left\{-\frac{1}{2\lambda}(y-\lambda)^2\right\},$$

per $y \in \Re$ e $\lambda > 0$ parametro ignoto.

(a) Si stabilisca se lo stimatore di massima verosimiglianza $\hat{\lambda}$ è non distorto.

(b) Si ottenga un'approssimazione per la distribuzione della statistica $T = (1/n)\sum_{i=1}^n y_i^2$.

(c) Si fornisca la regione critica più potente, di livello approssimato $\alpha = 0.05$, per il problema di verifica d'ipotesi $H_0 : \lambda = 1$ contro $H_1 : \lambda > 1$.

(d) Si supponga ora che legge di Y sia un elemento della classe di distribuzioni gamma $Ga(\lambda, 1)$, con parametro di forma λ e parametro di scala 1. Si può affermare che lo stimatore di massima verosimiglianza $\hat{\lambda}$ (ottenuto sotto $\mathcal{F}$) rimane, in questo caso, stimatore consistente per la media di Y?

Soluzione

(a) Per la funzione di verosimiglianza vale la relazione

$$L(\lambda) \propto \lambda^{-n/2} \exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^n (y_i - \lambda)^2\right\}.$$

Quindi,

$$l(\lambda) = \log(L(\lambda)) = -\frac{n}{2}\log(\lambda) - \frac{1}{2\lambda}\sum_{i=1}^n (y_i - \lambda)^2,$$

e

$$\begin{aligned} l_*(\lambda) = \frac{dl(\lambda)}{d\lambda} &= -\frac{n}{2\lambda} + \frac{2\lambda\sum_i (y_i - \lambda) + \sum_i (y_i - \lambda)^2}{2\lambda^2} \\ &= \frac{-n\lambda - n\lambda^2 + nT}{2\lambda^2} = -\frac{n}{2\lambda} - \frac{n}{2} + \frac{nT}{2\lambda^2}, \end{aligned}$$

con $T = (1/n)\sum_{i=1}^n y_i^2$. Uguagliando a zero, si ottiene un'equazione di secondo grado, $-n\lambda - n\lambda^2 + nT = 0$, che risolta fornisce due radici, di cui una negativa (quindi non considerabile dato che $\lambda > 0$), e una positiva, che rappresenta lo stimatore di massima verosimiglianza

$$\hat{\lambda} = \frac{-1 + \sqrt{1 + 4T}}{2}.$$

Ora, usando la diseguaglianza di Jensen e il fatto che $E(T) = E(Y^2) = \lambda + \lambda^2$, si ricava che

$$E(\hat{\lambda}) \neq \sqrt{\frac{1}{4} + E(T)} - \frac{1}{2} = \sqrt{\frac{1}{4} + \lambda + \lambda^2} - \frac{1}{2} = \sqrt{\left(\frac{1}{2} + \lambda\right)^2} - \frac{1}{2} = \lambda.$$

Pertanto, $\hat{\lambda}$ è distorto.

(b) La derivata seconda della log-verosimiglianza vale

$$l_{**}(\lambda) = \frac{n}{2\lambda^2} - \frac{nT}{\lambda^3},$$

cosicché l'informazione attesa ha espressione

$$i(\lambda) = E[-l_{**}(\lambda)] = -\frac{n}{2\lambda^2} + \frac{n(\lambda + \lambda^2)}{\lambda^3} = \frac{n(1 + 2\lambda)}{2\lambda^2}.$$

Ne segue che, per la distribuzione dello stimatore di massima verosimiglianza, vale l'approssimazione
$$\hat{\lambda} \dot{\sim} N\left(\lambda, \frac{2\lambda^2}{n(1+2\lambda)}\right).$$
D'altra parte, T è funzione di $\hat{\lambda}$, essendo $T = \left(\lambda + \frac{1}{2}\right)^2 - \frac{1}{4}$. Quindi, usando il metodo delta, si può concludere che T ha distribuzione approssimabile con quella normale di media $\lambda + \lambda^2$ e varianza
$$\frac{2\lambda^2}{n(1+2\lambda)}(2\lambda+1).$$
Si osservi che allo stesso risultato si arriva invocando il Teorema del Limite Centrale e utilizzando la seconda identità di Bartlett.

(c) La famiglia parametrica considerata (sottoclasse delle distribuzioni normali) è famiglia esponenziale regolare monoparametrica. Sotto campionamento casuale semplice, la statistica canonica è $\sum_{i=1}^n y_i^2$. Quindi il test uniformemente più potente per il problema di verifica d'ipotesi considerato è necessariamente basato su $\sum_{i=1}^n y_i^2$ o sua trasformata. Dato che lo stimatore di massima verosimiglianza $\hat{\lambda}$ è funzione di T (quindi di $\sum_{i=1}^n y_i^2$), la regione critica più potente è semplicemente quella fornita dal test di Wald, cioè
$$\mathcal{R} = \left\{ \mathbf{y} : \frac{\hat{\lambda} - \lambda_0}{\sqrt{\frac{2\lambda_0^2}{n(1+2\lambda_0)}}} > k \right\},$$
con $\lambda_0 = 1$ e $k = 1.64$.

(d) Se fosse $Y \sim Gamma(\lambda, 1)$, si avrebbe (ancora) $E(Y) = var(Y) = \lambda$. Quindi, si avrebbe $E(Y^2) = \lambda + \lambda^2$ e, per la Legge dei Grandi Numeri, T risulterebbe (ancora) stimatore consistente per $\lambda + \lambda^2$. D'altro canto, lo stimatore di massima verosimiglianza $\hat{\lambda}$, ottenuto sotto l'ipotesi di normalità, è funzione di T, ossia $\hat{\lambda} = \frac{-1+\sqrt{1+4T}}{2} = g(T)$. Poiché la funzione $g(\cdot)$ è continua, se fosse $Y \sim Gamma(\lambda, 1)$, lo stimatore $\hat{\lambda}$ sarebbe stimatore consistente per
$$g(\lambda + \lambda^2) = \frac{-1 + \sqrt{1 + 4(\lambda + \lambda^2)}}{2} = \lambda,$$
ossia per la media di Y.