Soluzioni
30/5/2003

Esercizio 1. $\theta = (\mu, \sigma^2)$, $\Theta_0 = \mathbb{R}\mathrm{x}\{1\}$, $\Theta = \mathbb{R}\mathrm{x}[1, \infty)$

$$\lambda(x) = \frac{\sup_{\Theta_0} L(\theta)}{\sup_{\Theta} L(\theta)}.$$

Le stime di massima verosimiglianza saranno

$$\hat{\theta} = \bar{x}, \quad s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}.$$

e inoltre

$$\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum (x_i - \bar{x})^2 = \frac{n}{2\sigma^2} \left(-1 + \frac{s^2}{\sigma^2} \right),$$

quindi L è crescente per $\sigma^2 > s^2$, decrescente se $\sigma^2 < s^2$ e ha un massimo per $\sigma^2 = s^2$. Da cui

$$\lambda(x) = \begin{cases} \frac{L(\bar{x}),1}{L(\bar{x}),1} \text{ se } s^2 < 1 \\ \frac{L(\bar{x},1)}{L(\bar{x}),s^2} \text{ se } s^2 \geq 1 \end{cases}$$

per cui

$$\lambda(x) = (s^2 e^{-s^2+1})^{\frac{n}{2}} < k \qquad \text{per } s^2 \geq 1.$$

E' facile verificare che $\lambda(x)$ è una funzione non crescente in s^2, e quindi la zona di rifiuto è data da

$$R = \{(x_1 \dots x_n) : s^2 > k\} = \{(x_1 \dots x_n) : ns^2/\sigma^2 > k'\}$$

Per trovare k' si ricorda che

$$\frac{ns^2}{\sigma^2} \sim \chi^2_{n-1}$$

che sotto H_0 diventa

$$ns^2 \sim \chi^2_{n-1}.$$

Il valore di k' è dato dalla condizione:

$$\alpha = P\left(\frac{ns^2}{\sigma^2} > k' | \sigma^2 = 1\right) = P(ns^2 > k') = P(ns^2 > \chi^2_{n-1,1-\alpha}),$$

allora

$$k' = \chi^2_{n-1,1-\alpha}.$$

Il test è UMP in quanto la zona di rifiuto non dipende dal valore di σ^2.

Esercizio 2. Sia

$$A = \left\{ (x_1 \dots x_n) : \frac{ns^2}{\sigma^2} \leq \chi^2_{n-1,1-\alpha} \right\}$$

la zona di accettazione del precedente test di ampiezza α. Utilizzando il criterio di inversione della zona di accettazione si ottiene l'intervallo di confidenza unilaterale di livello $1-\alpha$:

$$IC = \left\{ \sigma^2 : \sigma^2 \geq \frac{ns^2}{\chi^2_{n-1,1-\alpha}} \right\}.$$

Esercizio 3. Per utilizzare risultati noti sugli stimatori dei minimi quadrati per modelli lineari occorre linearizzare il modello $Y = \alpha X^\beta \epsilon$. Questo è possibile utilizzando la funzione logaritmo, ovvero:

$$\log Y = \log \alpha + \beta \log X + \log \epsilon.$$

A questo punto possiamo scrivere il modello come

$$Z = \delta + \beta W + \eta$$

da cui segue che

$$\hat{\beta} = \frac{s_{zw}}{s_w^2} \text{ e } \hat{\delta} = \bar{Z} - \hat{\beta}\bar{X}$$

Questi stimatori sono non distorti per la teoria nota sui modelli lineari del tipo:

$$y = \alpha + \beta X + \epsilon$$

Esercizio 4. Poniamo:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}, \quad \hat{\beta} = \frac{s_{xy}}{s_x^2}, \quad \hat{y}_i = \hat{\alpha} + \hat{\beta}x_i, \quad e_i = y_i - \hat{y}_i.$$

Allora:

$$\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = \sum_{i=1}^n e_i(\hat{y}_i - \bar{y}) = \sum_{i=1}^n e_i \hat{y}_i - \bar{y} \sum_{i=1}^n e_i.$$

Ma:

$$\sum_{i=1}^n e_i = \sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n y_i - \sum_{i=1}^n (\hat{\alpha} + \hat{\beta}x_i) = n\bar{y} - n\hat{\alpha} - n\hat{\beta}\bar{x} = 0$$

Da cui, sostituendo,

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = \sum_{i=1}^{n} e_i\hat{y}_i = \hat{\alpha}\sum_{i=1}^{n} e_i + \hat{\beta}\sum_{i=1}^{n} e_i x_i = \hat{\beta}\sum_{i=1}^{n} e_i x_i,$$

dobbiamo solo dimostrare che $\sum e_i x_i = 0$, infatti:

$$\sum_{i=1}^{n} e_i x_i = \sum_{i=1}^{n} y_i x_i - \hat{a}\sum_{i=1}^{n} x_i - \hat{\beta}\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y} + n\hat{\beta}\bar{x}^2 - \hat{\beta}\sum_{i=1}^{n} x_i^2.$$

Possiamo inoltre riscrivere la covarianza e la varianza campionaria in questo modo:

$$ns_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y},$$

$$ns_x^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2,$$

quindi

$$\sum_{i=1}^{n} e_i x_i = ns_{xy} - \hat{\beta}s_x^2 = 0. \qquad \square$$

Esercizio 5. Definiamo la quantità

$$Y = \sum_{i=1}^{n} X_i,$$

ed osserviamo che $Y \sim Gamma(2n, 1/\theta)$, ovvero:

$$f_Y(y;\theta) = \frac{1}{\Gamma(2n)}\frac{1}{\theta^{2n}} y^{2n-1} e^{-\frac{y}{\theta}} I_{(0,\infty)}(y),$$

pongo

$$Z = \frac{2Y}{\theta} \Rightarrow Y = \frac{\theta Z}{2},$$

allora:

$$f_Z(z;\theta) = f_Y(y(z);\theta)\left|\frac{d}{dz}y(z)\right| = \frac{1}{\Gamma(2n)}\frac{1}{2^{2n}} z^{2n-1} e^{-\frac{z}{2}} I_{(0,\infty)}(z).$$

Quindi $Z \sim Gamma(2n, 1/2) = \chi^2_{4n}$. Poiché Z ha distribuzione che non dipende da θ, essa rappresenta una quantità pivotale per θ. Per quanto riguarda gli intervalli di confidenza, è sufficiente utilizzare i quantili della distribuzione χ^2_{4n}, in particolare si ha:

$$Prob(Z \le x_\alpha) = 1 - \alpha$$

da cui segue che un intervallo di confidenza di livello $1-\alpha$ è dato da:

$$\left[\frac{2\sum X_i}{x_\alpha}, \infty\right).$$

Analogamente l'altro intervallo di confidenza unilaterale è dato da

$$\left[0, \frac{2\sum X_i}{x_{1-\alpha}}\right),$$

ed infine l'intervallo di confidenza bilaterale:

$$\left[\frac{2\sum X_i}{x_{\alpha/2}}, \frac{2\sum X_i}{x_{1-\alpha/2}}\right).$$

Esercizio 6. Come è noto la fnzione di densità di Y è data da:

$$f_Y(y;\theta) = \frac{ny^{n-1}}{\theta^n} I_{(0,\theta)}(y).$$

A partire da tale densità possiamo calccolare la distribuzione di Z, infatti

$$Y = \theta e^{-\frac{z}{2n}},$$

e quindi:

$$f_Z(z;\theta) = f_Y(y(z);\theta)\left|\frac{d}{dz}y(z)\right| = -\frac{1}{2}e^{-\frac{z}{2}} I_{(0,\infty)}(z).$$

Ovvero $Z \sim Exp(1/2) = Gamma(1, 1/2) = \chi^2_2$ che non dipende da θ per cui Z è una quantità pivotale. Gli intervalli di confidenza si possono determinare utilizzando i quantili della distribuzione χ^2 con due gradi di libertà. In particolare utilizzando le relazioni:

$$Prob(Z \le x_\alpha) = 1 - \alpha,$$

$$Prob(Z > x_{1-\alpha}) = 1 - \alpha$$

e

$$Prob(x_{1-\alpha} < Z \le x_{\alpha/2}) = 1 - \alpha,$$

si hanno i corrispondenti intervalli di confidenza:

$$\left(0, Ye^{-\frac{x_\alpha}{2n}}\right],$$

$$\left(Ye^{-\frac{x_{1-\alpha}}{2n}}, \infty\right]$$

e

$$\left(Ye^{-\frac{x_{1-\alpha/2}}{2n}}, Ye^{-\frac{x_{\alpha/2}}{2n}}\right].$$