Corsi di Inferenza Statistica – Secondo Esonero. 31 Maggio 2010

Tutti i Canali.

ATTENZIONE: I FOGLI A QUADRETTI CONSEGNATI
NON VERRANNO CONSIDERATI

Cognome e Nome: ______________________________ **Canale:** ____________

Esercizio 1. Sia $X_1, \ldots, X_n$ un campione casuale proveniente dalla popolazione con funzione di densità di probabilità

$$f_X(x;\theta) = \frac{3}{\theta^3}x^2, \qquad 0 < x < \theta, \qquad \theta > 0.$$

1. Verificare che
$$\mathbb{E}[X] = \frac{3}{4}\theta, \qquad \mathbb{V}[X] = \frac{3}{80}\theta^2$$
e che lo stimatore dei momenti di θ è
$$\widehat{\theta}_M = \frac{4}{3}\,\overline{X}_n$$

   Svolgimento.

$$\mathbb{E}[X] = \int_0^\theta x\frac{3}{\theta^3}x^2dx = \frac{3}{\theta^3}\left[\frac{x^4}{4}\right]_0^\theta = \frac{3}{4}\theta$$

$$\mathbb{E}[X^2] = \int_0^\theta x^2\frac{3}{\theta^3}x^2dx = \frac{3}{\theta^3}\left[\frac{x^5}{5}\right]_0^\theta = \frac{3}{5}\theta^2$$

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{3}{5}\theta^2 - \left(\frac{3}{4}\theta\right)^2 = \frac{3}{80}\theta^2$$

   Per determinare lo stimatore dei momenti, si eguaglia la media campionaria al valore atteso:

$$\overline{X}_n = \mathbb{E}[X] \Longleftrightarrow \overline{X}_n = \frac{3}{4}\theta \Longleftrightarrow \widehat{\theta}_M = \frac{4}{3}\overline{X}_n$$

2. Determinare l'errore quadratico medio dello stimatore dei momenti e studiarne la consistenza.

   Svolgimento.

   Si verifica facilmente che lo stimatore dei momenti è non distorto, quindi l'errore quadratico medio è uguale alla varianza:

$$MSE(\widehat{\theta}_M) = \mathbb{V}(\widehat{\theta}_M) = \mathbb{V}(\frac{4}{3}\overline{X}_n) = \frac{16}{9}\frac{\mathbb{V}(X)}{n} = \frac{16}{9}\frac{1}{n}\frac{3}{80}\theta^2 = \frac{\theta^2}{15n}$$

   Poiché l'errore quadratico medio tende a 0 per $n \to \infty$ possiamo concludere che lo stimatore dei momenti è consistente.

3. Determinare la distribuzione asintotica dello stimatore $\widehat{\theta}_M$ e l'intervallo di confidenza asintotico per θ, assumendo come stima di $\mathbb{V}[X]$ la quantità ottenuta sostituendo a θ lo stimatore dei momenti.

   Svolgimento.

   Ricordando che $\overline{X}_n \approx N\left(\mathbb{E}[X], \frac{\mathbb{V}[X]}{n}\right)$ (Teorema del Limite Centrale), si ricava facilmente la distribuzione asintotica dello stimatore dei momenti:

$$\widehat{\theta}_M = \frac{4}{3}\overline{X}_n \approx N\left(\frac{4}{3}\mathbb{E}[X], \frac{16}{9}\frac{\mathbb{V}[X]}{n}\right) \approx N\left(\theta, \frac{\theta^2}{15n}\right)$$

   Assumendo come stima della varianza $\frac{\widehat{\theta}_M^2}{15n}$, possiamo trovare l'intervallo di confidenza asintotico per θ:

$$\tilde{IC}_{1-\alpha}(\theta) = \widehat{\theta}_M \pm z_{\frac{1-\alpha}{2}}\sqrt{\frac{\widehat{\theta}_M^2}{15n}} = \frac{4}{3}\overline{X}_n \pm z_{\frac{1-\alpha}{2}}\sqrt{\frac{(\frac{4}{3}\overline{X}_n)^2}{15n}}$$

4. Determinare la stima dei momenti e l'intervallo di confidenza asintotico di livello $1-\alpha = 0.95$ supponendo di avere osservato un campione di dimensione $n = 36$ per il quale $\sum_{i=1}^n x_i = \frac{3}{2}$.

   Svolgimento.

   La stima dei momenti è

$$\widehat{\theta}_M = \frac{4}{3}\overline{x}_n = \frac{4}{3}\frac{\sum_{i=1}^n x_i}{n} = \frac{4}{3}\frac{3}{2}\frac{1}{36} = \frac{1}{18} = 0.056$$

   e l'intervallo di confidenza asintotico a livello 0.95:

$$\tilde{IC}_{1-\alpha}(\theta) = \frac{1}{18} \pm 1.96\sqrt{\frac{(1/18)^2}{15 \cdot 36}} = [0.051; 0.06]$$

Esercizio 2. Sia $X_1, \dots, X_n$ un campione casuale proveniente dalla popolazione con funzione di densità di probabilità

$$f_X(x;\theta) = \theta^2 x e^{-\theta x}, \qquad x > 0, \qquad \theta > 0,$$

con

$$\mathbb{E}[X] = \frac{2}{\theta} \qquad \mathbb{V}[X] = \frac{2}{\theta^2}.$$

Si consideri il sistema di ipotesi

$$H_0: \quad \theta = \theta_0, \qquad H_1: \quad \theta = \theta_1, \qquad\qquad (\theta_0 > \theta_1).$$

1. Verificare che la regione di accettazione del test di Neyman-Pearson risulta essere l'insieme

$$A = \{\mathbf{x} \in \mathcal{X}^n : \quad \overline{x}_n < k\}, \qquad k > 0$$

Svolgimento.

Applicando il Lemma di Neyman-Pearson, la regione di accettazione sarà del tipo

$$A = \{\mathbf{x} \in \mathcal{X}^n : \quad \frac{L(\theta_0;\mathbf{x})}{L(\theta_1;\mathbf{x})} \geq c\}.$$

Per il modello considerato la verosimiglianza risulta essere

$$L(\theta;\mathbf{x}) = \prod_{i=1}^{n} \left(\theta^2 x_i e^{-\theta x_i}\right) \propto \theta^{2n} exp\left\{-\theta \sum_{i=1}^{n} x_i\right\}$$

e quindi si ha:

$$\frac{L(\theta_0;\mathbf{x})}{L(\theta_1;\mathbf{x})} = \frac{\theta_0^{2n} exp\left\{-\theta_0 \sum_{i=1}^{n} x_i\right\}}{\theta_1^{2n} exp\left\{-\theta_1 \sum_{i=1}^{n} x_i\right\}} = \left(\frac{\theta_0}{\theta_1}\right)^{2n} exp\left\{-(\theta_0 - \theta_1)\sum_{i=1}^{n} x_i\right\}.$$

Dal momento che $-(\theta_0 - \theta_1) < 0$ il rapporto delle verosimiglianze è funzione decrescente di $\sum_{i=1}^{n} x_i$ (o equivalentemente di $\overline{x}_n$) e dunque possiamo concludere che la regione di accettazione è del tipo

$$A = \{\mathbf{x} \in \mathcal{X}^n : \quad \sum_{i=1}^{n} x_i < k'\} = \{\mathbf{x} \in \mathcal{X}^n : \quad \overline{x}_n < k\}$$

2. Determinare la distribuzione asintotica di $\overline{X}_n$.

Svolgimento.

$$\overline{X}_n \approx N\left(\mathbb{E}[X], \frac{\mathbb{V}[X]}{n}\right) \approx N\left(\frac{2}{\theta}, \frac{2}{n \cdot \theta^2}\right)$$

3. Verificare che, utilizzando la distribuzione asintotica di $\overline{X}_n$, il valore di k per il quale la probabilità di errore di I specie del test (vedi il punto 1.) è pari ad α è

$$k_\alpha = \frac{2}{\theta_0} + \sqrt{\frac{2}{n\theta_0^2}} z_{1-\alpha}.$$

Determinare il valore di k_α per $\alpha = 0.05$, $\theta_0 = 2$ e $n = 25$.

Svolgimento.

Fissando α e utilizzando la regione di accettazione di cui al punto 1 si ha

$$\alpha = P(\mathbf{X} \in R = A^c | H_0) = P(\overline{X}_n > k | \theta = \theta_0).$$

Sotto H_0 vale la distribuzione asintotica di cui al punto 2, e standardizzando si ha:

$$\alpha = P\left(\frac{\overline{X}_n - \frac{2}{\theta_0}}{\sqrt{\frac{2}{n\theta_0^2}}} > \frac{k - \frac{2}{\theta_0}}{\sqrt{\frac{2}{n\theta_0^2}}}\right) = 1 - P\left(Z < \frac{k - \frac{2}{\theta_0}}{\sqrt{\frac{2}{n\theta_0^2}}}\right) \Longleftrightarrow 1 - \alpha = \Phi\left(\frac{k - \frac{2}{\theta_0}}{\sqrt{\frac{2}{n\theta_0^2}}}\right)$$

Applicando la funzione inversa alla funzione di ripartizione si ottiene:

$$\frac{k - \frac{2}{\theta_0}}{\sqrt{\frac{2}{n\theta_0^2}}} = z_{1-\alpha}$$

da cui si ricava

$$k_\alpha = \frac{2}{\theta_0} + z_{1-\alpha}\sqrt{\frac{2}{n\theta_0^2}}.$$

Per $\alpha = 0.05$, $\theta_0 = 2$ e $n = 25$, si ha

$$k_\alpha = 1 + 1.64\sqrt{\frac{2}{25 \cdot 4}} = 1.23.$$

4. Determinare la potenza del test considerato per $\theta_1 = 1$.

Svolgimento.

La potenza, in corrispondenza dell'ipotesi alternativa semplice $\theta_1 = 1$, risulta essere

$$1-\beta = P(\mathbf{X} \in R = A^c | H_1) = P(\overline{X}_n > k_\alpha | \theta = \theta_1) = 1-\Phi\left(\frac{1.23 - \frac{2}{1}}{\sqrt{\frac{2}{25}}}\right) = 1-\Phi(-2.72) = 0.997$$

5. Stabilire se, per un campione osservato di dimensione $n = 25$ e con media campionaria pari a 1, l'ipotesi nulla viene accettata o rifiutata.

Svolgimento.

Poiché il valore osservato della media campionaria $\overline{x}_{25} = 1 < k_\alpha = 1.23$ cade nella regione di accettazione, non si può rifiutare l'ipotesi nulla a livello $\alpha = 0.05$.