Inferenza Statistica
II prova scritta a.a. 2012-2013 – 9 luglio 2013

SOLUZIONI

ATTENZIONE: I FOGLI A QUADRETTI CONSEGNATI
NON VERRANNO CONSIDERATI

• • •

Cognome e Nome: ______________________________ **Canale** F. De Santis ☐ L. Tardella ☐

Esercizio 1. Sia $X_1, \ldots, X_n$ un campione casuale proveniente da una popolazione tale che

$$E[X] = \frac{1}{3}\theta, \qquad V[X] = \frac{4}{45}\theta^2,$$

1. - Trovare lo stimatore dei momenti di θ, $\widehat{\theta}_M$;
   - determinare l'errore quadratico medio di $\widehat{\theta}_M$;
   - studiare la consistenza di $\widehat{\theta}_M$.
2. Determinare la distribuzione asintotica dello stimatore $\widehat{\theta}_M$ e l'intervallo di confidenza asintotico per θ.
3. 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 $\sum_{i=1}^n x_i = 1/3$.

Svolgimento:

1. - Ponendo $\overline{X}_n = \frac{1}{3}\theta$ si ottiene $\widehat{\theta}_M = 3\overline{X}_n$.
   - Si ha $E[\widehat{\theta}_M] = 3\frac{1}{3}\theta = \theta$, quindi $\widehat{\theta}_M$ è non distorto. Pertanto il suo MSE coincide con la varianza:
   $$MSE(\widehat{\theta}_M) = V(\widehat{\theta}_M) = V(3\overline{X}_n) = 9\frac{V(X)}{n} = \frac{4}{5n}\theta^2.$$
   Poiché $MSE(\widehat{\theta}_M) \to 0$ per $n \to \infty$, lo stimatore $\widehat{\theta}_M$ è consistente.
2. La distribuzione asintotica è
   $$\widehat{\theta}_M \approx N\left(\theta, \frac{4}{5n}\widehat{\theta}_M^2\right)$$
   e l'intervallo di confidenza approssimato:
   $$\widehat{\theta}_M \pm z_{\alpha/2}\sqrt{\frac{4}{5n}\widehat{\theta}_M^2} = 3\overline{X}_n \pm z_{\alpha/2}\sqrt{\frac{4}{5n}9\overline{X}_n^2}.$$

3. La stima dei momenti è

$$\widehat{\theta}_M = 3\overline{x}_n = 3 \cdot \frac{1}{3} \cdot \frac{1}{36} = \frac{1}{36} = 0.0278$$

e l'intervallo di confidenza asintotico

$$\widehat{\theta}_M \pm z_{\alpha/2}\sqrt{\frac{4}{5n}\widehat{\theta}_M^2} = \frac{1}{36} \pm 1.96 \cdot \frac{1}{36}\sqrt{\frac{4}{5 \cdot 36}} = (0.0197; 0.0359)$$

Esercizio 2. Sia $X_1 \dots X_n$ un campione casuale con valore atteso $E[X] = \mu$ e $V[X] = \sigma^2$. Si consideri la statistica

$$T(\mathbf{X}_n) = aS_n^2 \qquad a > 0,$$

dove S_n^2 indica la varianza campionaria corretta.

1. Mostrare che
$$MSE(T) = a^2 V(S_n^2) + (a-1)^2\sigma^4.$$
Mostrare inoltre che, assumendo $X_i \sim N(\mu, \sigma^2)$, si ottiene
$$MSE(T) = \frac{\sigma^4}{n-1}[2a^2 + (n-1)(a-1)^2].$$

2. Verificare che, al variare di a in $\mathbb{R}^+$, il valore minimo di $MSE(T)$ si ottiene per
$$a = (n-1)/(n+1).$$
[**Sugg**.: studiare $MSE(T)$ come funzione di a].

3. Sia
$$T^* = \frac{n-1}{n+1}S_n^2.$$
Determinare la probabilità di osservare un valore di T^* superiore a 2, assumendo $n = 10$ e $\sigma^2 = 4$.

Svolgimento:

1. Calcoliamo l'errore quadratico medio:
$$MSE(T) = V(T)+B^2(T) = V(aS_n^2)+[E(aS_n^2)-\sigma^2]^2 = a^2V(S_n^2)+[a\sigma^2-\sigma^2]^2 = a^2V(S_n^2)+(a-1)^2\sigma^4$$
Se $X_i \sim N(\mu, \sigma^2)$ si ha che $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$ e quindi
$$V(S_n^2) = V\left(\frac{\sigma^2}{(n-1)}\frac{(n-1)S_n^2}{\sigma^2}\right) = \frac{\sigma^4}{(n-1)^2}V\left(\frac{(n-1)S_n^2}{\sigma^2}\right) = \frac{\sigma^4}{(n-1)^2}2(n-1) = \frac{\sigma^4}{(n-1)}2$$
Si ha quindi:
$$MSE(T) = 2a^2\frac{\sigma^4}{(n-1)} + (a-1)^2\sigma^4 = \frac{\sigma^4}{n-1}[2a^2 + (n-1)(a-1)^2].$$

2. Studiamo $MSE(T)$ in funzione di a. La derivata prima rispetto ad a è
$$\frac{d}{da}MSE(T) = \frac{\sigma^4}{n-1}\left(4a + 2(n-1)(a-1)\right).$$
Questa funzione si annulla per $a = \frac{n-1}{n+1}$. Inoltre la derivata seconda rispetto ad a
$$\frac{d^2}{da^2}MSE(T) = \frac{\sigma^4}{n-1}\left(4 + 2(n-1)\right)$$
è sempre positiva, per cui $a = \frac{n-1}{n+1}$ è un punto di minimo.

3. Calcoliamo

$$P\left(\frac{n-1}{n+1}S_n^2 > 2\right) = P\left(\frac{\sigma^2}{n+1}\frac{n-1}{\sigma^2}S_n^2 > 2\right) = P\left(\chi^2_{n-1} > 2\frac{n+1}{\sigma^2}\right) = P\left(\chi^2_{n-1} > 5.5\right) = 0.789$$

(**Nota:** dalle tavole della chi quadrato il valore critico piú vicino a 5.5 è 5.07 che corrisponde a una probabilità pari a 0.75)

Esercizio 3. Sia $X_1, \dots, X_n$ un campione casuale proveniente dalla popolazione con funzione di densità di probabilità

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

con

$$E[X] = \frac{1}{\theta} \qquad\qquad V[X] = \frac{1}{\theta^2}.$$

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

$$A = \{\mathbf{x}_n : \quad \overline{x}_n < k\}, \qquad k > 0.$$

2. Utilizzando la distribuzione asintotica di $\overline{X}_n$, verificare che il valore di k per il quale la probabilità di errore di I specie del test di cui al punto 1) pari a α risulta essere

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

3. - Stabilire se, per un campione in cui la media campionaria pari a 1.5, l'ipotesi nulla viene accettata o rifiutata, per $\alpha = 0.05$.

   - Determinare la potenza del test considerato, assumendo $\alpha = 0.05$, $\theta_0 = 2$ e $\theta_1 = 1$ $n = 25$.

Svolgimento:

1. La funzione di verosimiglianza è:

$$L(\theta) = \theta^n e^{-\theta \sum_{i=1}^n x_i}.$$

   Il rapporto tra le verosimiglianze è:

$$\frac{L(\theta_0)}{L(\theta_1)} = \left(\frac{\theta_0}{\theta_1}\right)^n e^{-(\theta_0 - \theta_1)\sum_{i=1}^n x_i}$$

   che è funzione decrescente di $\sum_{i=1}^n x_i$ e quindi anche di $\overline{x}_n$, da cui segue che la forma della regione di accettazione è

$$A = \{\mathbf{x}_n : \quad \overline{x}_n < k\}, \qquad k > 0.$$

2. Per determinare k_α, fissiamo

$$\alpha = P(R|\theta = \theta_0) = P(\overline{X}_n > k_\alpha|\theta = \theta_0) = P\left(\frac{\overline{X}_n - \frac{1}{\theta_0}}{\sqrt{\frac{1}{n\theta_0^2}}} > \frac{k_\alpha - \frac{1}{\theta_0}}{\sqrt{\frac{1}{n\theta_0^2}}}\right) = P\left(Z > \frac{k_\alpha - \frac{1}{\theta_0}}{\sqrt{\frac{1}{n\theta_0^2}}}\right)$$

   da cui segue che

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

3. - Se $\alpha = 0.05$, $\theta_0 = 2$, $n = 25$ si ha $k_\alpha = \frac{1}{2} + 1.65\sqrt{\frac{1}{25\cdot 4}} = 0.665$. Poiché $\overline{x}_n = 1.5 > k_\alpha = 0.665$ si rifiuta l'ipotesi nulla.

   - La potenza del test, assumendo $\alpha = 0.05$, $\theta_0 = 2$ e $\theta_1 = 1$, $n = 25$ è pari a

$$\alpha = P(R|\theta = \theta_1) = P(\overline{X}_n > k_\alpha|\theta = \theta_1) = P\left(Z > \frac{0.665 - 1}{\sqrt{\frac{1}{25}}}\right) = 1 - \Phi(-1.675) = 0.953.$$