Inferenza statistica – Parte A – 6 giugno 2023 SOLUZIONI A

Cognome, nome e n. di matricola: ______

[SI CORREGGE SOLO QUANTO RIPORTATO IN QUESTI FOGLI]
Gli studenti che hanno diritto alla riduzione del compito possono saltare i quesiti 2, 5 e 9

A – Quesiti [dove necessario giustificare adeguatamente le risposte]

1. Sia X una v.a. con funzione di densità $f_X(x;\theta) = \frac{1}{\theta-1}\exp\left\{-\frac{x}{\theta-1}\right\}$, $x \geq 0$, $\theta > 1$, con $\mathbb{E}[X] = \theta - 1$. Determinare lo stimatore di massima verosimiglianza di θ e stabilire, attraverso i teoremi di Rao-Blackwell e Lehmann-Scheffè, se è UMVUE.

Risp. $X \sim Esp(\theta-1) \Rightarrow X \sim Esp(\lambda) \quad \lambda = \theta - 1$
$\Rightarrow$ UMVUE di λ è $\bar{X}_n$ e di θ è $\bar{X}_n + 1$
in quanto: fam espon UNIPARAM, $\mathbb{E}(\bar{X}_n + 1) = \theta$,
$\bar{X}_n$ è SUFF e COMPLETA

$\boxed{\hat{\theta}_{MV} = \bar{X}_n + 1}$

2. Con riferimento al precedente esercizio, determinare il limite inferiore Cramer-Rao per gli stimatori non distorti di θ.

Risp. $\mathbb{V}[\bar{X}_n + 1] = \mathbb{V}(x) = (\theta-1)^2 = cr(\theta)$ in quanto fam espon^n unipar^n $\Rightarrow cr(\theta) = I_n^{-1}(\theta) = \mathbb{V}(\text{UMVUE})$

$\boxed{cr(\theta) = (\theta-1)^2/n}$

3. Sia $X_1, \ldots, X_n$ un campione casuale da $f_X(x;\theta) = \theta\exp\{-\theta x\}$, $x \geq 0$, $\theta > 0$. Determinare la varianza asintotica dello stimatore di massima verosimiglianza di $\tau = \ln\theta$.

Risp. $g(\theta) = \ln\theta \Rightarrow g'(\theta) = 1/\theta$
$L(\theta) = \theta^n e^{-\Sigma x_i \theta} \Rightarrow \ell(\theta) = n\ln\theta - (\Sigma x_i)\theta$
$\Rightarrow \ell'(\theta) = n/\theta - \Sigma x_i \Rightarrow \ell''(\theta) = -n/\theta^2 \cdot \hat{\theta}_{MV} = 1/\bar{x}_n$
$\Rightarrow I_n^{-1}(\theta) = \theta^2/n$
$\Rightarrow \boxed{\mathbb{V}[\hat{\tau}] \simeq g'(\theta)^2 \cdot I_n^{-1}(\theta) = \frac{1}{\theta^2}\cdot\frac{\theta^2}{n} = 1/n}$

4. Sia $X_1, \ldots, X_n$ un campione casuale da $N(\mu, \sigma^2)$. Supponendo che $n = 11$, $S_0^2 = 4.5$, $S_n^2 = 4.9$ e $\bar{x}_n = 1.3$, determinare l'intervallo di confidenza per σ^2 di livello $1 - \alpha = 0.90$.

Risp.
$$C = \left[\frac{(n-1)S_n^2}{\chi^2_{n-1;1-\frac{\alpha}{2}}}, \frac{(n-1)S_n^2}{\chi^2_{n-1;\frac{\alpha}{2}}}\right] = \left[\frac{10 * 4.9}{18.31}, \frac{10*4.9}{3.94}\right] = [2.67, 12.44]$$

Con S_0^2
$$\left[\frac{11*4.5}{19.68}, \frac{11*4.5}{4.57}\right] = [2.52, 10.83]$$

5. Con riferimento all'esercizio precedente, determinare l'intervallo di confidenza per μ di livello $1 - \alpha = 0.90$.

Risp.
$$C = \bar{x}_n \pm t_{n-1;1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}} \qquad \alpha = 0.1 \quad \alpha/2 = 0.05$$
$$= 1.3 \pm 1.81\frac{\sqrt{4.9}}{\sqrt{11}} = [0.09, 2.51]$$

$U[a,b] \Rightarrow \mathbb{E}(X) = \frac{a+b}{2}$
$V(X) = (b-a)^2/12$

6. Sia $X_1, \ldots, X_n$ un campione casuale dalla popolazione uniforme nell'intervallo $[-\theta, 2\theta]$, $\theta > 0$. Determinare l'intervallo di confidenza asintotico di livello $1-\alpha$ per θ basato sullo stimatore dei momenti.

Risp. $\mathbb{E}(X) = \frac{2\theta - \theta}{2} = \frac{\theta}{2} \Rightarrow \boxed{\hat{\theta}_M = 2\bar{X}_n}$

$\mathbb{E}(\hat{\theta}_n) = \theta \quad V(\hat{\theta}_M) = 4\,V(X)/n = \frac{\cancel{4}(3\theta)^2}{3\,\cancel{12}\,n} = \frac{3\theta^2}{n}$

$\Rightarrow$

$\tilde{C} = \hat{\theta}_M \pm z_{1-\frac{\alpha}{2}} \frac{\sqrt{3}\,\hat{\theta}_M}{\sqrt{n}} = 2\bar{X}_n \pm \frac{2\sqrt{3}\,\bar{X}_n z_{1-\frac{\alpha}{2}}}{\sqrt{n}}$

7. Sia $X_1, \ldots, X_n$ un campione casuale dalla popolazione $f_X(x;\theta) = \frac{\theta}{x^{\theta+1}}$, $x > 1$, $\theta > 0$. Determinare una statistica sufficiente T rispetto alla quale il modello ha rapporto delle verosimiglianze monotono e scrivere la generica regione di rifiuto del test di Karlin-Rubin per le ipotesi $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$.

Risp.

$L(\theta) = \theta^n / \left(\prod^n x_i\right)^{\theta} = \theta^n / T^{\theta} \qquad \boxed{T = \prod^n x_i}$

$(\theta_1, \theta_2) \quad \theta_1 < \theta_2$

$\lambda_{21} = (\theta_2/\theta_1)^n \left(1/T\right)^{\theta_2 - \theta_1} \quad \downarrow \text{ in } T \text{ e } \uparrow \text{ in } 1/T$

$\Rightarrow R = \{\underline{x}_n : \prod^n x_i < k\}$

8. Con riferimento al quesito precedente, determinare lo stimatore di massima verosimiglianza di θ e la sua distribuzione asintotica.

Risp. $\ell(\theta) = n \ln\theta - \theta \ln T \;;\; \ell'(\theta) = n/\theta - \ln T$

$\Rightarrow \hat{\theta}_{MV} = n/\ln T \;.\; \ell''(\theta) = -n/\theta^2 \Rightarrow I_n(\theta) = n/\theta^2$

$\Rightarrow \boxed{\hat{\theta}_{MV} \overset{a}{\sim} N\left(\theta, \theta^2/n\right)} \qquad \boxed{\hat{\theta}_{MV} = n/\ln T}$

9. Con riferimento ai quesiti precedenti, determinare l'espressione della statistica test $\widetilde{W}_0$ e la regola di rifiuto del test di ampiezza α.

Risp. $\tilde{W}_0 = \left(\hat{\theta}_{MV} - \theta_0\right) I_n^{-1/2}(\theta_0)$

$= \left(\hat{\theta}_{MV} - \theta_0\right) / \sqrt{\theta_0^2/n} = \frac{\sqrt{n}}{\theta_0}\left(\hat{\theta}_{MV} - \theta_0\right)$

$\Rightarrow$ (k.t tipo A)

$R = \{\underline{x}_n : \tilde{W}_0 > z_{1-\alpha}\}$

10. Sia $X_1, \ldots, X_n$ un campione casuale da $N(\mu, \sigma^2)$ con $\sigma = 8.9$ Si consideri il sistema di ipotesi $H_0 : \mu = 70$ vs. $H_1 : \mu > 70$. Per il campione osservato si ha: $n = 100$, $\bar{x}_n = 71.8$. Calcolare il p-value e stabilire se si accetta l'ipotesi nulla con $\alpha = 0.05$.

Risp. $p(\underline{x}_n) = 1 - \Phi(w^{oss})$ con $w^{oss} = \sqrt{n}(\bar{x}_n - \theta_0)/\sigma$

$w^{oss} = \sqrt{100}\,(71.8 - 70)/8.9 = 2.02$

$\Rightarrow$

$1 - \Phi(2.02) = 0.02 < 0.05 \Rightarrow$ RIF H_0

Inferenza statistica – Parte B – 6 giugno 2023

Cognome, nome e n. di matricola: SOLUZIONI B

[SI CORREGGE SOLO QUANTO RIPORTATO IN QUESTI FOGLI]
Gli studenti che hanno diritto alla riduzione del compito possono saltare i quesiti 6, 5 e 9

B – Problema.
Sia $\mathbf{X}_n = (X_1, \ldots X_n)$ un campione i.i.d. da una popolazione X con funzione di densità

$$f_X(x;\theta) = (\theta+1)x^{\theta}, \qquad x \in [0,1], \quad \theta > -1.$$

1. Determinare: (a) funzione di verosimiglianza, $L(\theta;\mathbf{x}_n)$, (b) statistica sufficiente $T(\mathbf{x}_n)$, (c) stima di massima verosimiglianza di θ, $\widehat{\theta}_{mv}(\mathbf{x}_n)$.

   Risp. (a) $L(\theta) = (\theta+1)^n \left(\prod^n x_i\right)^{\theta} \quad \theta > -1$; (b) $\boxed{T = \prod^n x_i}$

   (c) $\ell(\theta) = n \ln(\theta+1) + \theta \ln T$

   $\ell'(\theta) = n/(\theta+1) + \ln T \Rightarrow \boxed{\hat{\theta}_{MV} = -n/\ln T - 1}$

   $\ell''(\theta) = -n/(\theta+1)^2 < 0 \quad \forall \theta \Rightarrow$ OK

2. Determinare informazione attesa $I_n(\theta)$ e approssimazione normale per la distribuzione di $\widehat{\theta}_{mv}(\mathbf{X}_n)$.

   Risp. $I_n(\theta) = n/(\theta+1)^2$

   $\boxed{\hat{\theta}_{MV} \,\dot{\sim}\, N\left(\theta, (\theta+1)^2/n\right)}$

3. Utilizzando la distribuzione asintotica di $\widehat{\theta}_{mv}(\mathbf{X}_n)$, determinare la quantità pivotale asintotica $\widetilde{Q}(\mathbf{X}_n,\theta)$ e il corrispondente intervallo di confidenza $\widetilde{C}(\mathbf{X}_n)$ per θ di livello $1-\alpha$.

   Risp. $\tilde{Q} = \sqrt{n}\dfrac{(\hat{\theta}_{MV} - \theta)}{\theta+1}$

   $\downarrow$

   $\tilde{Q} = \sqrt{n}\,(\hat{\theta}_{MV} - \theta)/(\hat{\theta}_{MV}+1) \Rightarrow \boxed{\tilde{C} = \hat{\theta}_{MV} \pm z_{1-\frac{\alpha}{2}} \dfrac{\hat{\theta}_{MV}+1}{\sqrt{n}}}$

   con $\hat{\theta}_{MV} = -n/\ln T - 1$

4. Determinare lunghezza aleatoria $\mathcal{L}(\mathbf{X}_n)$ e lunghezza attesa $e_n(\theta)$ di $\widetilde{C}(\mathbf{X}_n)$. [Utilizzare la distribuzione asintotica di $\widehat{\theta}_{mv}(\mathbf{X}_n)$].

   Risp. $\mathcal{L}(X_n) = 2 z_{1-\alpha/2} (\hat{\theta}_{MV}+1)/\sqrt{n}$

   $e_n(\theta) = 2 z_{1-\frac{\alpha}{2}} (\theta+1)/\sqrt{n}$

5. Supporre che $n = 10$, $\alpha = 0.05$ e $\ln[\prod_{i=1}^n x_i] = -5$. Determinare gli estremi di $\widetilde{C}(\mathbf{x}_n)$.

   $\hat{\theta}_{MV} = -\dfrac{10}{-5} - 1 = 1$

   Risp. $\tilde{C} = 1 \pm (1.96)\dfrac{2}{\sqrt{10}} = [-0.24, 2.24]$

6. Sia $\theta_d = 1$ e $\ell_0 = 2$. Determinare il minimo valore di n tale che $e_n(\theta_d) < \ell_0$.

Risp. $e_n(\theta_d) = 2 z_{1-\frac{\alpha}{2}} (\theta_d+1)/\sqrt{n} < \ell_0 \iff$

$n > \left\lceil \frac{4 z^2 (\theta_d+1)^2}{\ell_0^2} \right\rceil = \left\lceil \frac{4 * 1.96^2 * 4}{4} \right\rceil = 16$

7. Si consideri il sistema di ipotesi $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ (con $\theta_0 < \theta_1$). Determinare l'espressione del rapporto delle verosimiglianze $\lambda_{01}(\mathbf{x}_n)$ e la generica regione di rifiuto di H_0 in funzione della statistica sufficiente $T(\mathbf{x}_n)$ trovata nel punto 1.

Risp. $L(\theta) = (\theta+1)^n \times T^{\theta}$ $\quad T = \prod_{i}^{n} x_i$

$\lambda_{01} = \left(\frac{\theta_0+1}{\theta_1+1}\right)^n T^{(\theta_0-\theta_1)}$, decr risp T essendo $(\theta_0-\theta_1) < 0$ –

$\Rightarrow$

$\lambda_{01} < K \iff T > K \Rightarrow R = \{x_n : T = \prod_{i}^{n} x_i > K\}$

8. Determinare l'espressione della regione di rifiuto di H_0 in funzione di $\widehat{\theta}_{mv}$ e il valore soglia k_α del test di ampiezza α. [Utilizzare la distribuzione asintotica di $\widehat{\theta}_{mv}(\mathbf{X}_n)$].

usare $\hat{\theta}_{MV}/\theta_0 \overset{\sim}{\sim} N\left(\theta_0, \frac{(\theta_0+1)^2}{n}\right)$

Risp. $\hat{\theta}_{MV} = -n/\ln T - 1$ $\quad$ Rif $\iff \hat{\theta}_{MV} > K$

K_α : $P_{\theta_0}[\hat{\theta}_{MV} > K] = \alpha \iff P_{\theta_0}\left[Z > \frac{(K-\theta_0)\sqrt{n}}{\theta_0+1}\right] = \alpha$

$\iff$

$\frac{(K-\theta_0)\sqrt{n}}{\theta_0+1} = z_{1-\alpha} \iff \boxed{K_\alpha = \frac{\theta_0+1}{\sqrt{n}} z_{1-\alpha} + \theta_0}$

9. Determinare l'espressione di $1-\beta$, potenza del test di ampiezza α. [Utilizzare la distribuzione asintotica di $\widehat{\theta}_{mv}(\mathbf{X}_n)$].

usare $\hat{\theta}_{MV}/\theta_1 \overset{\sim}{\sim} N\left(\theta_1, \frac{(\theta_1+1)^2}{n}\right)$

Risp. $1-\beta = P_{\theta_1}[R] = P_{\theta_1}\left[\hat{\theta}_{MV} \geq \frac{(\theta_0+1) z_{1-\alpha}}{\sqrt{n}} + \theta_0\right]$

$= P_{\theta_1}\left[Z \geq \frac{\frac{\theta_0+1}{\sqrt{n}} z_{1-\alpha} + \theta_0 - \theta_1}{(\theta_1+1)/\sqrt{n}}\right]$

$= 1 - \Phi\left[\frac{\theta_0+1}{(\theta_1+1)} z_{1-\alpha} + \sqrt{n}(\theta_0-\theta_1)/(\theta_1+1)\right]$

10. Si consideri il sistema di ipotesi $H_0 : \theta = \theta_0$ vs. $H_1 : \theta < \theta_0$. Determinare la versione della statistica sufficiente rispetto alla quale è crescente il rapporto delle verosimiglianze e determinare la regola di rifiuto del test UMP di ampiezza α. [Utilizzare la distribuzione asintotica di $\widehat{\theta}_{mv}(\mathbf{X}_n)$].

Risp. $\theta_1 < \theta_2$ $\quad \lambda_{21} = \left(\frac{\theta_2+1}{\theta_1+1}\right)^n T^{(\theta_2-\theta_1)}$ $\quad T = \prod_{i}^{n} x_i$

$\Rightarrow$ (essendo $\theta_2-\theta_1>0$)

$\lambda_{21} \uparrow$ con $T = \prod_{i} x_i \Rightarrow (KR)$ $R = \{x_n : T < K\}$

Ma

$T < K \iff \hat{\theta}_{MV} < K \Rightarrow R_\alpha = \left\{x_n : \frac{(\hat{\theta}_{MV}-\theta_0)\sqrt{n}}{\theta_0+1} < z\right.$

Inferenza statistica – Parte C – 6 giugno 2023

Cognome, nome e n. di matricola: SOLUZIONI e

[SI CORREGGE SOLO QUANTO RIPORTATO IN QUESTI FOGLI]
Gli studenti che hanno diritto alla riduzione del compito possono saltare i quesiti 4, 7 e 8

A – Quesiti [dove necessario giustificare adeguatamente le risposte]

1. Sia $f_X(x;\theta) = \binom{4}{x}\theta^x(1-\theta)^{4-x}$, $x = 0, 1, 2, 3, 4$ $\theta \in [0,1]$. Determinare $\mathbb{E}_\theta[\overline{X}_n]$ e $\mathbb{V}_\theta[\overline{X}_n]$.

   Risp. $X \sim \text{Binom}(4,\theta) \Rightarrow \mathbb{E}(\bar{X}_n) = \mathbb{E}(X) = 4\theta$

   $\mathbb{V}(\bar{X}_n) = \frac{\mathbb{V}(X)}{n} = \frac{4\theta(1-\theta)}{n}$

2. $X_1, \ldots, X_{n+1}$ i.i.d. $N(\mu, \sigma^2)$. Considerare $\overline{X}_{n+1} = \frac{X_{n+1}+n\overline{X}_n}{n+1}$. Calcolare $\mathbb{E}[\overline{X}_{n+1}]$.

   Risp. $\mathbb{E}(\bar{X}_{n+1}) = \mathbb{E}(X) = \mu$

3. Con riferimento al precedente esercizio, calcolare $\mathbb{V}[\overline{X}_{n+1}]$.

   Risp. $\mathbb{V}(\bar{X}_{n+1}) = \frac{\mathbb{V}(X)}{n+1} = \frac{\sigma^2}{n+1}$

4. Con riferimento al problema precedente, stabilire se $\overline{X}_{n+1}$ è uno stimatore consistente di μ.

   Risp. $\text{MSE}(\bar{X}_{n+1}) = \frac{\sigma^2}{n+1} + 0 \mapsto 0 \quad \forall \mu \in \mathbb{R}$

   $\Rightarrow$ stim consist di μ

5. Sia $X_1, \ldots, X_n$ un campione casuale da $N(\mu, \sigma^2)$. Verificare che lo stimatore S_n^2 di σ^2 è consistente.

   Risp. $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1} \Rightarrow \mathbb{E}[S_n^2] = \sigma^2 \quad \forall \sigma^2$

   $\mathbb{V}[S_n^2] = \frac{2\sigma^4}{n-1}$

   $\Rightarrow \text{MSE}(S_n^2) = \mathbb{V}(S_n^2) = \frac{2\sigma^4}{n-1} \xrightarrow{n} 0 \quad \forall \sigma^2 > 0 \Rightarrow$ CONS

6. $X_1, \ldots, X_n$ i.i.d. Pois(θ) i.i.d. - Calcolare la stima di MV di $g(\theta) = \mathbb{P}(X = 0; \theta)$, per un campione osservato in cui $\sum_{i=1}^{n} x_i = 20$ e $n = 10$?

Risp.

$\mathbb{P}(X=0;\theta) = e^{-\theta} \Rightarrow$ (equiv 8MV)

$\widehat{g(\theta)} = g(\hat{\theta}_{MV}) = e^{-\bar{X}_n} \Rightarrow (\bar{x}_n = 2)$

$g(\hat{\theta}_{MV}(x_n)) = e^{-2}$

7. $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$. Determinare uno stimatore non distorto di $\psi = 3\mu - 2$ che sia funzione di $\overline{X}_n$.

Risp.

$\mathbb{E}(\bar{X}_n) = \mu \quad \forall\mu \Rightarrow \mathbb{E}(3\bar{X}_n - 2) = 3\mu - 2 \quad \forall\mu$

$\Rightarrow \hat{\psi} = 3\bar{X}_n - 2$ è ND per ψ.

8. Con riferimento al precedente esercizio, determinare la varianza dello stimatore di ψ.

Risp.

$\mathbb{V}[\hat{\psi}] = 9\,\mathbb{V}(\bar{X}_n) = 9\dfrac{\sigma^2}{n}$

9. Sia $X \sim$ Unif[0, 1]. Determinare $g_{ps}(x; \mu, \sigma)$ (funzione di densità della famiglia posizione-scala). Con riferimento al precedente esercizio, determinare valore atteso e varianza della v.a. Y con funzione di densità $g_{ps}(\cdot; \mu, \sigma)$.

Risp. $f_X(x) = 1 \cdot \mathbb{I}_{[0,1]}(x)$

$\Rightarrow$

$g_{ps}(x; \mu, \sigma) = \dfrac{1}{\sigma} f\left(\dfrac{x-\mu}{\sigma}\right) = \dfrac{1}{\sigma}\mathbb{I}_{[\mu, \mu+\sigma]}(x)$

f.ne di dens di $Y = \mu + \sigma X \Rightarrow \mathbb{E}[Y] = \mu + \sigma \cdot 1/2 = \mu + \sigma/2$

$\mathbb{V}[Y] = \sigma^2 \mathbb{V}(X) = \sigma^2/12$

10. Sia X una v.a. che assume i valori $(0, 1, 2)$ con probabilità $(1 - \frac{\theta}{2}, \frac{\theta}{4}, \frac{\theta}{4})$, $\theta \in [0, 1]$. Determinare $\mathbb{E}[X]$ e lo stimatore dei momenti di θ basato su un campione casuale.

Risp.

$\mathbb{E}(X) = 0(1 - \theta/2) + 1\,\theta/4 + 2\,\theta/4 = \dfrac{\theta}{2} + \dfrac{\theta}{4}$

$= \dfrac{3}{4}\theta$

$\boxed{\hat{\theta}_M = \dfrac{4}{3}\bar{X}_n}$