Inferenza statistica – Scritto 1, Parte A
10 giugno 2021

Cognome, nome e n. di matricola: SOLUZIONI

A – Quesiti [dove necessario giustificare adeguatamente le risposte]

1. Sia $X_1, \ldots, X_n$ un campione causale da una popolazione $X \sim \text{Ber}(\theta)$ Per un campione di dimensione $n = 30$ si ha che $\sum_{i=1}^{n} x_i = 20$. Calcolare l'intervallo di confidenza asintotico per θ di livello 0.80.

   Risp.
   $\bar{X}_n | \theta \overset{\sim}{\sim} N(\theta, \theta(1-\theta)/n) \Rightarrow \tilde{C} = \bar{x}_n \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{x}_n(1-\bar{x}_n)}{n}}$
   $\bar{x}_n = 20/30 = 2/3 \quad z_{1-\alpha/2} = z_{0.9} = 1.28$
   $\Rightarrow \tilde{C} = 2/3 \pm 1.28 \sqrt{\frac{2/3 \times 1/3}{30}} = [0.56, 0.78]$

2. Sia $X_1, \ldots, X_n \sim \text{Pois}(\theta)$ i.i.d..

   (a) Calcolare l'intervallo di confidenza di livello $1 - \alpha = 0.90$ per $g(\theta) = e^{-\theta}$ sapendo che $\sum_{i=1}^{n} x_i = 36$, $n = 18$.

   (b) Stabilire se, sulla base di questo intervallo, è plausibile supporre che il vero valore di $g(\theta)$ sia pari a 0.2 (rispondere e motivare).

   $\begin{cases} \alpha = 0.1 \Rightarrow \\ z_{1-\alpha/2} = z_{0.95} = 1.645 \end{cases}$

   Risp. (a)
   METODO $\Delta \Rightarrow \tilde{C} = e^{-\bar{x}_n} \pm z_{1-\frac{\alpha}{2}} e^{-\bar{x}_n} \sqrt{\bar{x}_n/n}$
   $\bar{x}_n = 36/18 = 2$
   $\tilde{C} = e^{-2} \pm 1.645\, e^{-2} \sqrt{2/18} = [0.06, 0.21]$

   Risp. (b) $0.2 \in [0.06\ \ 0.21] \Rightarrow$ SI

3. Sia $X_1, \ldots, X_n \sim N(\theta, 1/2)$ i.i.d.. Scrivere l'espressione generica della statistica test di Karlin-Rubin (UMP) per il confronto delle ipotesi $H_0 : \theta = 0.6$ vs. $H_1 : \theta < 0.6$. Sapendo che $\sum_{i=1}^{n} x_i = 12$, $n = 30$, determinare il valore numerico della statistica test, scrivere la regola di rifiuto del test di ampiezza $\alpha = 0.05$ (in funzione dell'opportuno percentile) e stabilire se si accetta o rifiuta H_0.

   Risp.
   W= $\sqrt{n}(\bar{x}_n - \theta_0)/\sigma$
   valore numerico W= $\sqrt{30}(12/30 - 0.6)/\sqrt{0.5}$
   Regola generale per questo test: RIF Ho se $W_{oss} < z_\alpha$
   In questo caso specifico: $w_{oss} = -1.55 \quad z_\alpha = -1.64 \Rightarrow$ ACC Ho

4. Sia $X_1, \ldots, X_n \sim N(\theta, 1)$. Si consideri il campione osservato con: $\bar{x} = 4$, $\sum(x_i - \bar{x})^2 = 64$, $n = 25$. Scrivere in funzione di $\Phi(\cdot)$ o di `pnorm`$(\cdot)$ l'espressione del p-value per la verifica delle ipotesi $H_0 : \theta = 3.5$ vs. $H_1 : \theta \neq 3.5$. Calcolare il p-value e stabilire se si accetta o rifiuta H_0 in un test di ampiezza $\alpha = 0.01$.

   Risp.
   espressione w.oss: $w_{oss} = \sqrt{n}(\bar{x}_n - \theta_0)/\sigma = \sqrt{64}\,(4 - 3.5) = 2.5$
   valore di w.oss: 2.5
   espressione generica p-value: $2\Phi(-|w_{oss}|) =$ ~~0.012~~
   valore p-value: $0.012 = 2 * \Phi(-2.5)$
   decisione: $0.012 > 0.01 \Rightarrow$ ACCETTO Ho

5. Sia (X_1, X_2, X_3) un campione causale di dimensione $n = 3$ da una popolazione di Poisson di parametro θ. Considerare il sistema di ipotesi $H_0 : \theta = 2$ vs. $\theta = 1$. Si consideri la regione di rifiuto dell'ipotesi nulla del test definita da R $= \{\mathbf{x}_n : \sum_0^3 x_i > 2\}$. Calcolare la probabilità di errore di II specie associata alla regione R.

Risp. $\beta = P(\text{Acc } H_0 \mid \theta_1) = P\left[\sum X_i \le 2 \mid \theta = 1\right]$

Qui $\sum_0^3 X_i \mid \theta_1 \sim$ Poisson $(n\,\theta_1) =$ Poisson (3)

$\Rightarrow$

$\beta =$ sum [dpois (0:2, 3)] $= 0.677$

6. Sia $X_1, \ldots, X_n$ un campione causale da una popolazione con funzione di densità uniforme nell'intervallo $[0, \theta]$, $\theta > 0$. Determinare l'espressione dell'intervallo di confidenza **esatto** unilaterale di livello $1 - \alpha = 0.95$, supponendo che $n = 20$ e $x_{(n)} = 2$.

[Ricordare che $X_{(n)}/\theta \sim$ Beta(n,1) e prendere $q_2 = 1$].

Risp. $\frac{X_{(n)}}{\theta}$, Beta $(n, 1) \Rightarrow C = [X_{(n)}, X_{(n)}/q_\alpha]$

$q_\alpha =$ qbeta $(\alpha, n, 1) =$ qbeta $(0.05, 20, 1)$

$\Rightarrow$

$C = \left(2, \frac{2}{0.86}\right) = [2, 2.3]$

7. Con riferimento al precedente esercizio, determinare l'espressione dell'intervallo di confidenza **asintotico** di livello $1 - \alpha = 0.95$ basato sullo stimatore dei momenti di θ e i valori numerici di L e U, sapendo che $\overline{x}_n = 1.5$.

Risp. Poiché $2\overline{X}_n = \hat{\theta}_M \overset{\cdot}{\sim} N(\theta, \theta^2/3n) \Rightarrow$

$\tilde{C} = 2\overline{X}_n \pm z_{1-\frac{\alpha}{2}} (2\overline{X}_n)/\sqrt{3n}$ $\qquad \hat{\theta}_n = 2 \times 1.5 = 3$

$= (2 * 1.5) \pm 1.96\, (2 * 1.5) / \sqrt{3 * 20} = [2.24, 3.76]$

8. Modello normale $N(\mu\sigma^2)$ (entrambi i parametri incogniti). La statistica test per verifica di ipotesi nulla puntuale contro alternativa bilaterale su σ^2 ha, sotto ipotesi nulla, distribuzione:

☐ χ^2_n ☒ χ^2_{n-1} ☐ t_n ☐ t_{n-1} ☐ $N(0,1)$

9. Date le v.a. $U \sim$ N(0, 1), $V_q = \chi^2_q$ e $V_p = \chi^2_p$ indipendenti. Come si definiscono le v.a. T_q di Student e di Fisher $F_{q,p}$?

$T_q =$ $U / \sqrt{V_q/q}$

$F_{q,p} =$ $\dfrac{V_q/q}{V_p/p}$

10. Indicare l'espressione dell'intervallo di confidenza per $\psi = \sigma_a^2/\sigma_b^2$ nel caso di due campioni di dimensioni n_a e n_b da popolazioni normali indipendenti N(μ_a, σ_a^2) e N(μ_b, σ_b^2) con valori attesi noti (definire le quantità che si utilizzano).

VEDI DISPENSE

Inferenza statistica – Scritto 1, Parte B
10 giugno 2021

Cognome, nome e n. di matricola: ______________________

B – Problema. Sia $X_1, \ldots X_n$ un campione casuale in cui si assume per X_i la seguente funzione di densità di probabilità:

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

1. Verificare che le v.a. X_i hanno distribuzione Gamma e individuare i parametri di forma (di solito indicato con α) e scala (di solito indicato con β). Determinare quindi valore atteso e varianza delle v.a. X_i.

   Risp.

   $f_X(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\beta x} \quad x \geq 0 \quad \theta > 0$

   fmo

   $\alpha = 2 \quad \beta = \theta \quad \Rightarrow \quad \mathbb{E}_\theta(X) = 2\theta \quad \mathbb{V}_\theta(X) = 2\theta^2$

2. Determinare $\widehat{\theta}_M$, lo stimatore dei momenti del parametro θ e stabilire se è UMVUE di θ

   Risp.

   $\hat{\theta}_M : \quad \mathbb{E}_\theta(X) = \overline{X}_n \Rightarrow 2\theta = \overline{X}_n$

   $\Rightarrow \hat{\theta}_n = \overline{X}_n / 2$

   Poiché $L(\theta) \propto \frac{1}{\theta^{2n}} e^{-\sum x_i/\theta} \Rightarrow \sum x_i$ è SUFF

   (+ fam. ESPON)

   $\Rightarrow \hat{\theta}_n$ è $\left\{\begin{matrix}\text{ND} \\ \text{fne di STAT SUFF COMPLETA}\end{matrix}\right.$ $\Rightarrow \hat{\theta}_n \equiv \hat{\theta}_{UMVUE}$

3. Determinare l'informazione attesa e l'approssimazione normale della distribuzione campionaria di $\widehat{\theta}_M(\mathbf{X}_n)$. Fornire l'espressione di uno stimatore della varianza asintotica di $\widehat{\theta}_{mv}(\mathbf{X}_n)$. $(*) \Rightarrow \hat{\theta}_n \dot{\sim} N\left(\theta, \frac{\theta^2}{2n}\right)$

   Risp.

   $L(\theta) = \theta^{-2n} e^{-T/\theta} \Rightarrow \ell(\theta) = -2n \ln\theta - T/\theta$

   $\Rightarrow \ell'(\theta) = -\frac{2n}{\theta} + \frac{T}{\theta^2} \qquad \ell''(\theta) = \frac{2n}{\theta^2} - \frac{2T}{\theta^3}$

   $\mathbb{E}(\ell'') = \frac{2n}{\theta^2} - \frac{4n}{\theta^2} = -\frac{2n}{\theta^2} \Rightarrow I_n = \frac{\theta^2}{2n} \; (*)$

4. Determinare l' espressione dell'intervallo di confidenza asintotico $\tilde{C}$ di livello $1-\alpha$ per θ basato su $\widehat{\theta}_M$.

   Risp.

   $\tilde{C} = \hat{\theta}_n \pm z_{1-\alpha/2} \, \hat{\theta}_n / \sqrt{2n} \quad$ con $\quad \hat{\theta}_n = \overline{X}_n / 2$

   $\tilde{\mathbb{V}}_\theta[\hat{\theta}_n] = I_n^{-1}(\theta) = (\overline{X}/2)^2 / 2n = \overline{X}_n / 8n$

5. Determinare l'espressione della lunghezza aleatoria di $\tilde{C}$ e il suo valore atteso.

   Risp.

   $\mathcal{L} = 2 z_{1-\alpha/2} \, \hat{\theta}_n / \sqrt{2n} \Rightarrow \mathbb{E}_\theta(\mathcal{L}) = 2 z_{1-\frac{\alpha}{2}} \theta / \sqrt{2n}$

OPPURE : $I_n(\theta) = cr(\theta)^{-1} = [\mathbb{V}(\hat{\theta}_{UMVUE})]^{-1} = \left(\frac{\theta^2}{2n}\right)^{-1} = \frac{2n}{\theta^2}$

$\hat{\theta}_n = \dfrac{\bar{X}_n}{2} = \dfrac{80/20}{2} = 2$

6. Supponendo di avere osservato un campione di dimensione $n = 20$ in cui $Y_n = \sum_{i=1}^{n} x_i = 80$, determinare i valori numerici degli estremi dell'intervallo $\tilde{C}$ di livello $1 - \alpha = 0.85$ per θ.

Risp. $\tilde{C} = 2 \pm 1.96\,(2)/\sqrt{40} = \boxed{[1.38,\ 2.62]}$

7. Determinare una statistica sufficiente rispetto alla quale il modello ha rapporto delle verosimiglianze monotono (non decrescente).

Risp. $L(\theta) = 1/\theta^{2n}\, e^{-T/\theta}$ $\quad \theta_1 < \theta_2$

$\lambda_{21} = \left(\dfrac{\theta_1}{\theta_2}\right)^{2n} \exp\left\{-T\underbrace{\left(\dfrac{1}{\theta_2} - \dfrac{1}{\theta_1}\right)}_{<0}\right\}$ = fne ↑ di T

$\boxed{T = \sum x_i}$

8. Si consideri il sistema di ipotesi $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta_0$. Scrivere la (generica) regione di rifiuto di H_0 del test UMP (di Karlin-Rubin) in funzione della statistica sufficiente e dello stimatore dei momenti di θ.

Risp. RIF H_0 se $\sum x_i > K$

ovvero se $\dfrac{\sum x_i}{2n} > K'$ $\quad K \in \mathbb{R}^+$

1.45

9. Si consideri il sistema di ipotesi $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \gtrless \theta_0$, con $\theta_0 = 2/5$. Supporre di avere osservato un campione di dimensione $n = 20$ in cui $Y_n = \sum_{i=1}^{n} x_i = 80$. Calcolare il valore osservato $\widetilde{w}_0$ della statistica test $\widetilde{W}_0$ (test di Wald). (usare $\hat{\theta}_n$)

Risp.

Espressione di $\widetilde{W}_0$ $\quad (\hat{\theta}_n - \theta_0)/\sqrt{I_n^{-1}(\theta_0)} = (\bar{X}_n/2 - \theta_0)/\sqrt{\theta_0^2/2n}$

$= \sqrt{2n}\,(\bar{x}_n/2 - \theta_0)/\theta_0$

Valore numerico di $\widetilde{w}_0$ $\quad \sqrt{40}\left(\dfrac{4}{2} - 2/5\right)/(2/5) = 25.29$

10. Stabilire se si accetta o meno l'ipotesi nulla del test di ampiezza $\alpha = 0.01$ per θ.

Risp. RIF H_0 se $W_{oss} > q_{1-\alpha}$

$25.40 > 2.33$ NON

$\Rightarrow$ ACC H_0