Inferenza statistica – Scritto 1 – 13 giugno 2025

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 * e *

Parte A – Quesiti sulla seconda parte del programma

1. Sia $X_1, \dots, X_n$ un campione iid da $\mathrm{N}(\theta, \sigma^2)$ (con σ^2 **incognito**). Si consideri il sistema di ipotesi $H_0 : \theta = 3$ vs $H_1 : \theta < 3$. Scrivere l'espressione della statistica test $W(\mathbf{X}_n)$ per il confronto di tali ipotesi e la regola di rifiuto del test di ampiezza α.

> **Risp.**
> Problema N2: test su valore atteso di modello normale con varianza incognita. $W(\mathbf{X}_n) = \frac{\sqrt{n}}{S_n}(\overline{X}_n - \theta_0)$ che, sotto H_0, ha distribuzione t_{n-1}. Si rifiuta quindi l'ipotesi nulla con ampiezza α se $w.oss < t_{n-1,\alpha}$.

2. Con riferimento al precedente esercizio, si consideri un campione osservato di dimensione $n = 16$ con $\overline{x}_n = 2.5$ e $S_n^2 = 4$. Calcolare il valore osservato `w.oss` che si ottiene con i dati a disposizione e stabilire se, con i dati disponibili, l'ipotesi nulla viene accettata o meno in un test di ampiezza $\alpha = 0.1$. Fornire la rappresentazione grafica [IMP.: assegnare le etichette agli assi cartesiani].

> **Risp.**
> In questo caso: $w.oss = \frac{4}{2}(2.5 - 3) = -1$ e $t_{n-1,\alpha} = t_{15,0.1} = -1.341$. Si accetta quindi H_0. Per la rappresentazione grafica usare la densità della v.a. t_{15} e procedere come di consueto (consulta le dispense).

3. Con riferimento ai due quesiti precedenti, scrivere la formula del p-value $p_{\theta_0}(\mathbf{x}_n)$ che si ottiene con i dati disponibili. Rappresentare $p_{\theta_0}(\mathbf{x}_n)$ in un ulteriore grafico e dire quali conclusioni si traggono in merito all'accettazione/rifiuto dell'ipotesi nulla in un test di ampiezza $\alpha = 0.1$? [IMP.: non serve il calcolo numerico del p-value].

> **Risp.**
> Per questo sistema di ipotesi si ha che $p_{\theta_0}(\mathbf{x}_n) = \mathbb{P}_{\theta_0}[W(\mathbf{X}_n) < w.oss] = \mathbb{F}(-1) > 0.1$, dove $\mathbb{F}(\cdot)$ indica la funzione di ripartizione della v.a. t_{15}. Per la rappresentazione grafica vedi le dispense.

4. Sia $X_1, \ldots, X_n$ un campione iid da $\mathrm{N}(\mu, \theta)$ (con μ **incognito**). Considerare il sistema di ipotesi $H_0 : \theta = \frac{3}{2}$ vs $H_1 : \theta > \frac{3}{2}$. Scrivere l'espressione della statistica test $W(\mathbf{X}_n)$ per il confronto di tali ipotesi e la regola di rifiuto del test UMP di ampiezza α. [IMP: definire tutte le quantità che compaiono nelle formule usate].

> **Risp.**
> Problema N4 (test su varianza modello normale con valore atteso incognito). La statistica test è $W(\mathbf{X}_n) = \frac{(n-1)S_n^2}{\theta_0}$ che, sotto H_0, ha distribuzione χ^2_{n-1}. Il test UMP di ampiezza α rifiuta H_0 se $w.oss > q_{1-\alpha}^{\chi^2_{n-1}}$.

5. Con riferimento al precedente esercizio, si consideri un campione osservato di dimensione $n = 10$ con $S_n^2 = 3$. Calcolare il valore osservato `w.oss` che si ottiene con i dati a disposizione e stabilire se l'ipotesi nulla viene accettata o meno in un test di ampiezza $\alpha = 0.05$

> **Risp.**
> In questo caso $w.oss = \frac{9 \cdot 3}{3/2} = 18$ e $q_{1-\alpha}^{\chi^2_{n-1}} = q_{0.95}^{\chi^2_9} = 16.92$. Si rifiuta quindi H_0.

6. Con riferimento al precedente esercizio, calcolare l'intervallo di confidenza a code uguali di livello $1 - \alpha = 0.9$ per θ [ATTENZIONE: le tavole allegato del χ^2 forniscono i **valori critici** e non i quantili].

> **Risp.**
> $C_{1-\alpha} = \left[\frac{(n-1)S_n^2}{q_{0.975}}, \frac{(n-1)S_n^2}{q_{0.0.025}}, \right] = \left[\frac{27}{16.92}, \frac{27}{3.33}\right]$. I quantili sono del χ^2_9.

Parte B – Problema. Sia $\mathbf{X}_n = (X_1, \dots X_n)$ un campione iid da una popolazione X con funzione di densità $f_X(x;\theta) = \frac{\theta^3}{2x^4}\, e^{-\frac{\theta}{x}}\, \mathbb{I}_{\mathbb{R}^+}(x),\ \theta > 0$. Per la v.a. X sappiamo che $\mathbb{E}_\theta[X] = \frac{\theta}{2}$ e $\mathbb{V}_\theta[X] = \frac{\theta^2}{4}$.

1. Determinare la statistica sufficiente minimale con il criterio di fattorizzazione[1] e lo stimatore di massima verosimiglianza di θ.

> **Risp.**
> $L(\theta) = c\dot{\theta}^{3n} e^{-t\theta}$, con $t = \sum_{i=1}^{n} \frac{1}{x_i}$, che è quindi statistica sufficiente minimnale per il modello (criterio di fattorizzazione). Si verifica facilmente che $\widehat{\theta}_{mv} = \frac{3n}{t}$.

2. Determinare l'informazione attesa di Fisher $I_n(\theta)$ e la distribuzione asintotica di $\widehat{\theta}_{mv}$. [IMP: giustificare la risposta].

> **Risp.**
> Si verifica facilmente che $\ell''(\theta) = -\frac{3n}{\theta^2}$ e quindi, senza necessità di valore atteso, si ha che $I_n(\theta) = \frac{3n}{\theta^2}$. Poichè il modello è REGOLARE (giustificazione), per le propr. degli SMV si ha che $\widehat{\theta}_{mv} \dot{\sim} \mathrm{N}(\theta, \theta^2/3n)$.

3. Determinare l'espressione di $\widetilde{C}_{1-\alpha}$, intervallo di confidenza asintotico di livello $1-\alpha$. Calcolare il valore di $\widehat{\theta}_{mv}$ e gli estremi dell'intervallo $\widetilde{C}_{1-\alpha}$ con $1-\alpha = 0.95$ per un campione osservato tale che $n = 20$ e $t = 20$.

> **Risp.**
> Si ha che $\widehat{\theta}_{mv} = 3$ e quindi $\widetilde{C} = \widehat{\theta}_{mv} \pm z_{1-\frac{\alpha}{2}} \sqrt{I_n(\widehat{\theta}_{mv})^{-1}} = \widehat{\theta}_{mv} \pm z_{1-\frac{\alpha}{2}} \frac{\widehat{\theta}_{mv}}{\sqrt{3n}} = 3 \pm 1.96 \frac{3}{\sqrt{60}}$.

[1] Ricorda che $\sum \frac{1}{x_i} \neq \frac{1}{\sum x_i}$.

4. Applicando l'opportuna definizione, verificare che il modello ha rapporto delle verosimiglianze crescente in $\widehat{\theta}_{mv}$ [IMP: motivare la risposta].

> **Risp.**
> Prendiamo $\theta_2 > \theta_1$. Si ha che $\lambda_{21} = \left(\frac{\theta_2}{\theta_1}\right)^{3n} \exp\{-t(\theta_2 - \theta_1)\} = \left(\frac{\theta_2}{\theta_1}\right)^{3n} \exp\{-\frac{3n}{\widehat{\theta}_{mv}}(\theta_2 - \theta_1)\}$ che è funzione crescente di $\widehat{\theta}_{mv}$ dal momento che $\theta_2 - \theta_1 > 0$.

5. Considerare il sistema di ipotesi $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$. Scrivere l'espressione della statistica test di Wald $\widetilde{W}_0(\mathbf{X}_n)$ e la regola di rifiuto di H_0 per un test di ampiezza α.

> **Risp.**
> $\widetilde{W}_0 = \frac{\widehat{\theta}_{mv} - \theta_0}{sd(\theta_0)}$, con $sd(\theta_0) = \sqrt{I_n(\theta_0)^{-1}}$. Si rifiuta H_0 se $|w.oss| > z_{1-\frac{\alpha}{2}}$. In questo caso $w.oss = \sqrt{3n}\frac{3n/t - \theta_0}{\theta_0}$.

6. Supporre ora che $\theta_0 = 2$. Calcolare il valore osservato di $\widetilde{W}_0(\mathbf{X}_n)$ assumendo che $n = 20$ e $t = 20$ e $\alpha = 0.05$ e stabilire se si accetta o si rifiuta l'ipotesi nulla. [IMP: NON serve la calcolatrice per motivare la risposta].

> **Risp.**
> $w.oss = \frac{3-2}{2}\sqrt{60} = \sqrt{15} > 1.96 = z_{0.975}$. Si rifiuta quindi H_0.

7. Considerare $n = 1$. Determinare l'espressione di $\lambda_{10}(X)$ supponendo che $\theta_0 = 1$ e $\theta_1 = 2$ e determinare l'insieme dei valori di x tali che $\lambda_{10}(x) > k$ per un generico $k > 8$.

Risp.
Per $n = 1$ si ha che $\lambda_{10} = 8e^{-\frac{1}{x}}$.
$\lambda_{10} > k \iff -\frac{1}{x} > \ln\frac{k}{8} \iff x < -\frac{1}{\ln k8} = k'$. Poichè per $k > 8$ si ha che $k' < 0$ e poichè $x \geq 0$, l'insieme cercato è l'insieme vuoto.

8. Con riferimento al punto precedente, calcolare $\mathbb{P}_\theta[\lambda_{10}(X) > k]$, per $k > 8$. Esprimere i risultati in termini della f.ne di ripartizione di v.a. gamma o gamma inversa.

Risp.
Per $k > 8$ si ha che, per ogni $\theta > 0$, $\mathbb{P}_\theta[\lambda_{10} > k] = \mathbb{P}(X < k') = \mathbb{F}(k') = 0$, dove $\mathbb{F}(\cdot)$ indica la funzione di ripartizione della v.a. gamma inversa di parametri $(3, \theta)$.

9. Con riferimento al punto precedente, determinare:

(a) l'espressione k_α per k per il quale il test ha probabilità di errore di primo tipo pari ad α (Esprimere i risultati in funzione dei quantili di v.a. gamma o gamma inversa.);

(b) la potenza del test che si ottiene.

Risp.
(a)-(b)
Per quanto visto sopra (risultato valido per ogni θ), se $\alpha > 0$ non è possibile determinare una soglia k_α tale che il test abbia ampiezza α e quindi neanche la potenza di tale test.

Parte C - Quesiti sulla prima parte del programma

1. Sia $\mathbf{X}_n = (X_1, \ldots X_n)$ un campione iid da una popolazione X con funzione di densità $f_X(x;\theta) = \frac{\theta^3}{2x^4}\, e^{-\frac{\theta}{x}}\, \mathbb{I}_{\mathbb{R}^+}(x),\ \theta > 0$. Stabilire se θ è un parametro di forma o di scala.

   Risp.
   Ponendo $\theta = 1$ in $f_X(x;\theta)$ e poi scrivendo l'espressione di $g_s(x;\theta)$ si verifica facilmente che θ è un parametro di scala.

2. Sia $X_1, \ldots, X_n$ un campione casuale da una popolazione bernoulliana di parametro $\theta \in [0,1]$. Deteminare lo stimatore di massima verosimiglianza di $g(\theta) = \frac{\theta}{1-\theta}$ e la sua distribuzione asintotica.

   Risp.
   Nel modello bernoulliano si ha che $\widehat{\theta}_{mv} = \overline{X}_n$. Osservando che $g'(\theta) = \frac{1}{(1-\theta)^2}$ e $I_n(\theta) = \frac{n}{\theta(1-\theta)}$ per il metodo delta si ha che $\widehat{g(\theta)} = g(\widehat{\theta}_{mv}) = \frac{\overline{X}_n}{1-\overline{X}_n}$ e $\widehat{g(\theta)} \dot{\sim} \mathrm{N}\left(\frac{\theta}{1-\theta}, \frac{\theta}{n(1-\theta)^3}\right)$.

3. Sia $X_1, \ldots, X_n$ un campione casuale da una popolazione Unif in $[-\theta, 2\theta]$. Determinare lo stimatore dei momenti $\widehat{\theta}_m$ di θ e verificare se risulta non distorto.

   Risp.
   $\mathbb{E}_\theta[X] = \frac{\theta}{2}$ e quindi $\widehat{\theta}_m = 2\overline{X}_n$. Inoltre $\mathbb{E}_\theta[2\overline{X}_n] = \theta$ per ogni θ. Lo stimatore è quindi non distorto.

4. Con riferimento al precedente esercizio, determinare la funzione $\mathrm{MSE}_\theta(\widehat{\theta}_m)$ e studiare la consistenza dello stimatore dei momenti di θ.

> **Risp.**
> $\mathrm{MSE}_\theta(2\overline{X}_n) = \mathbb{V}_\theta(2\overline{X}_n) = ... = \frac{3\theta^2}{n}$ che tende a zero per ogni θ. Lo stimatore è quindi consistente in senso forte e debole.

5. Sia $X_1, \ldots, X_n$ un campione casuale da una popolazione di Poisson di parametro θ. Determinare in funzione di θ la quantità $\mathbb{P}_\theta[X < 2]$ e stimarla supponendo che $n = 10$ e $\sum_{i=1}^n x_i = 20$.

> **Risp.**
> $\mathbb{P}(X < 2) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = e^{-\theta} + \theta e^{-\theta}$. Poichè $\widehat{\theta}_{mv} = 2$, la stima della quantità cercata è $e^{-\overline{X}_n}(1 + \overline{X}_n)$.

6. Sia $X_1, \ldots, X_n$ un campione casuale da $\mathrm{EN}(\theta)$. Dimostrare che lo stimatore UMVUE di θ non può raggiungere il limite inferiore di Cramer-Rao. [Sugg.: utilizzare la funzione `score`].

> **Risp.**
> Il modello è famiglia esponenziale uniparametrica. La statistica sufficiente minimale e completa è $T = \sum_{i=1}^n X_i$. Si verifica che lo stimatore UMVUE è $\frac{n-1}{n}\frac{1}{\overline{X}_n}$, mentre $S_n(\mathbf{X}_n, \theta) = ... = \frac{n}{\theta} - n\overline{X}_n$. Funzione score e UMVUE non sono quindi legati da una relazione lineare e quindi l'UMVUE non raggiunge il LICR.