Inferenza statistica – Esonero 1, Parte A1 – 6 aprile 2024

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 8 e 9

A – Quesiti [dove necessario giustificare adeguatamente le risposte]

1. Sia $X|\theta \sim \text{Pois}\left(\frac{\theta}{2}\right)$. Determinare l'espressione di $\tau(\theta) = \mathbb{P}_\theta(X = 0)$ e stabilire per quali valori di θ tale probabilità risulta maggiore o uguale a $\frac{1}{2}$.

> **Soluzione**.
>
> - $f_X(x;\theta) = \frac{1}{x!}e^{-\frac{\theta}{2}}\left(\frac{\theta}{2}\right)^x$ e quindi $\mathbb{P}_\theta(X = 0) = e^{-\frac{\theta}{2}}$.
> - $\mathbb{P}_\theta(X = 0) = e^{-\frac{\theta}{2}} \geq \frac{1}{2} \iff \frac{\theta}{2} \leq \ln 2 \iff \theta \leq 2\ln 2$.

2. Sia (X_1, X_2, X_3) un campione casuale con ciascuna X_i avente distribuzione indicata nel precedente esercizio. Fornire le stime di massima verosimiglianza di θ e di $\tau(\theta)$ che si ottengono per il campione osservato $(x_1 = 2, x_2 = 1, x_3 = 6)$.

> **Soluzione**.
>
> - Se si pone $\lambda = \frac{\theta}{2}$, poichè $\widehat{\lambda}_{mv} = \overline{x}_n$, per invarianza della smv si ha che $\widehat{\theta}_{mv} = 2\overline{x}_n = 6$.
> - $\widehat{\tau} = e^{-\frac{\widehat{\theta}_{mv}}{2}} = e^{-\overline{x}_n} = e^{-3}$.

3. Sia X una variabile aleatoria con funzione di densità $f_X(x) = \frac{x}{a} \cdot \mathbb{I}_{[0,1]}(x)$, $a > 0$. Determinare il valore di a affinchè $f_X(x)$ sia una funzione di densità in $[0, 1]$.

> **Soluzione**.
> $\int_0^1 \frac{x}{a}dx = 1 \iff a = \int_0^1 xdx \iff a = \frac{x^2}{2}\Big|_0^1 \iff a = \frac{1}{2}$. Pertanto $f_X(x) = 2xI_{[0,1]}(x)$.

4. Con riferimento al precedente esercizio, determinare l'espressione della famiglia di scala $g_s(\cdot;\theta)$ che si ottiene utilizzando $f_X(x)$ come funzione di densità standard e $\theta > 0$ come parametro di scala.

> **Soluzione**.
> $g_s(x;\theta) = \frac{1}{\theta}f_X\left(\frac{x}{\theta}\right) = \frac{2}{\theta}\frac{x}{\theta}I_{[0,1]}\left(\frac{x}{\theta}\right) = \frac{2x}{\theta^2}I_{[0,\theta]}(x)$.

5. Con riferimento ai precedenti esercizi, determinare il valore atteso della v.a. X e della v.a. Y con funzione di densità $g_s(y;\theta)$.

> **Soluzione**.
>
> - $\mathbb{E}(X) = \int_0^1 2x^2dx = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3}$.
> - $Y \sim g_s(\cdot;\theta) \iff Y = \theta X$. Pertanto $\mathbb{E}_\theta(Y) = \frac{2}{3}\theta$.

6. Sia $X|\theta \sim \text{Unif}[\theta, 1]$, $\theta < 1$. Determinare l'espressione di $f_X(x;\theta)$, $\mathbb{E}_\theta[X]$ e $\mathbb{V}_\theta[X]$ (funzione di densità, valore atteso e varianza di X).

Soluzione. Ricordare che: $Y \sim \text{Unif}[a,b] \Longrightarrow f_X(x;\theta) = \frac{1}{b-a} I_{[a,b]}(x), \mathbb{E}(Y) = \frac{a+b}{2}$ e $\mathbb{V}(Y) = \frac{(b-a)^2}{12}$.
In questo caso, quindi, $f_X(x;\theta) = \frac{1}{1-\theta} I_{[\theta,1]}(x)$, $\mathbb{E}_\theta(X) = \frac{\theta+1}{2}$ e $\mathbb{V}_\theta(X) = \frac{(1-\theta)^2}{12}$.
Oppure si rifanno tutti i calcoli ...

7. Sia $X_1, \ldots, X_n$ un campione casuale da una popolazione con distribuzione $\text{Unif}[\theta, 1]$, $\theta < 1$ (come nel precedente esercizio). Determinare la funzione di verosimiglianza di θ, una statistica sufficiente minimale e la stima di massima verosimiglianza di θ.

Soluzione.

- $L(\theta) = \prod_{i=1}^n \frac{1}{1-\theta} I_{[\theta,1]}(x_i) = \frac{1}{(1-\theta)^n} I_{(-\infty, x_{(1)}]}(\theta)$.
- $L(\theta)$ è crescente in θ in $(-\infty, x_{(1)}]$ e nulla in $(x_{(1)}, 1)$; pertanto $\widehat{\theta}_{mv} = x_{(1)}$.
- Per il criterio di fattorizzazione $x_{(1)}$ è statistica suffficiente minimale (ha partizione coincidente con quella di verosimiglianza).

8. Con riferimento al precedente esercizio, determinare la funzione di verosimiglianza relativa di θ e l'espressione di $L_q(\mathbf{x}_n)$, insieme di verosimiglianza di livello $q \in (0,1)$ per θ. Fornire risultati analitici e rappresentazione grafica.

Soluzione.

- $\overline{L}(\theta) = \left[\frac{1-x_{(1)}}{1-\theta}\right]^n \; I_{(-\infty, x_{(1)}]}(\theta)$
- $\overline{L}(\theta) \geq q \Longleftrightarrow \begin{cases} \left[\frac{1-x_{(1)}}{1-\theta}\right]^n \geq q \\ \theta \leq x_{(1)} \end{cases} \Longleftrightarrow \theta \in [1 - q^{-1/n}(1 - x_{(1)}), x_{(1)}] = L_q(\mathbf{x}_n)$.
- Rappresentazione grafica: $\overline{L}(\theta)$ è positiva e crescente in $(-\infty, x_{(1)}]$ e vale zero per $\theta > x_{(1)}$. In $x_{(1)}$ la funzione $\overline{L}(\cdot)$ vale 1.

  Funzione R per rappresentazione grafica

```
L.rel=function(theta, xn){
  n=length(xn)
  x.1=min(xn)
  0*(theta>min(xn)) + (((1-x.1)/(1-theta))^n)*(theta<=x.1)
}
# ad esempio con i dati seguenti
dati=c(-2,2,3,-2,5)
curve(L.rel(x,dati),from=-4,to=1,xlab=expression(theta),ylab="verosim. relativa")
# per insieme di verosimiglianza Lq
q=0.7
abline(h=q)
n=length(dati)
# estermo inferiore insieme Lq
L.n=1-q^(-1/n)*(1-min(dati))
abline(v=L.n)
```

9. Siano X_1 e X_2 due v.a. con distribuzione $\text{Ber}(\theta)$ indipendenti tra loro. Determinare la distribuzione di probabilità della v.a. $U = (X_1 - X_2)$ [ovvero: riportare in una tabella i valori che U può assumere e con quali probabilità].

Soluzione. Poichè X_1 e X_2 possono entrambe assumere i valori $\{0, 1\}$, $Y = X_1 - X_2$ assume i valori $\{-1, 0, 1\}$. In particolare, tenendo conto della indipendenza e somiglianza di X_1 e X_2, si ha che

x_1	x_2	y	$\mathbb{P}_\theta(Y = y)$
0	0	0	$(1-\theta) \times (1-\theta)$
1	0	1	$\theta \times (1-\theta)$
0	1	-1	$(1-\theta) \times \theta$
1	1	0	$\theta \times \theta$

Riassumendo e ordinando

y	$\mathbb{P}_\theta(Y = y)$
-1	$\theta(1-\theta)$
0	$\theta^2 + (1-\theta)^2$
1	$\theta(1-\theta)$