

Tema di Statistica Matematica

Sessione invernale - Secondo appello

22 febbraio 2022

1) In uno studio sull'affidabilità di un certo tipo di componenti elettronici, vengono rilevate le durate di $n = 50$ componenti. Da precedenti studi, inoltre, si ha motivo di ritenere che la legge di distribuzione del logaritmo della durata sia adeguatamente descritta da un modello Normale di parametri μ e σ^2 , cioè $X = \ln(Y) \sim N(\mu, \sigma^2)$.

Sotto tale assunzione, la durata media del generico componente, $\tau = \mathbb{E}(Y)$, è esprimibile in funzione di μ e σ^2 e risulta $\tau = \exp\{\mu + \frac{\sigma^2}{2}\}$. Indichiamo con x_1, x_2, \dots, x_{50} i valori dei logaritmi delle durate osservate y_1, y_2, \dots, y_{50} .

- a) Trovare lo stimatore di massima verosimiglianza per τ e la sua distribuzione approssimata.
- b) Sapendo che $\sum_{i=1}^{50} x_i = 301.2$ e $\sum_{i=1}^{50} x_i^2 = 1818.4$ e sulla base dei risultati ottenuti al punto a), calcolare un intervallo di confidenza per τ di livello approssimato pari a 0.95.

2) La variabile aleatoria X rappresenta la frazione di memoria principale allocabile di un server che viene richiesta da un job qualsiasi. Si assuma che X segua una distribuzione avente densità

$$f(x; \theta) = \frac{1}{\theta} x^{\frac{1}{\theta}-1} \mathbb{I}_{(0,1)}(x), \quad \theta > 0. \quad (1)$$

Un valore basso di θ implica la preponderanza di "grossi" job; invece, se $\theta = 1$ la distribuzione delle richieste di memoria è uniforme.

Sia (X_1, X_2, \dots, X_n) un campione casuale estratto dalla distribuzione avente densità $f(x; \theta)$ data in (1) e si considerino le seguenti ipotesi: $H_0 : \theta = 2$ vs. $H_1 : \theta = 0.2$.

- a) La famiglia di distribuzioni in questione è regolare?
- b) Trovare una statistica che sia sufficiente per il parametro θ .
- c) Determinare la densità di $-\sum_{i=1}^n \ln(X_i)$ sotto H_0 e sotto H_1 . [NB: $\ln(X_i)$ indica il logaritmo naturale di X_i].
- d) Costruire il test più potente (MP) di livello α per il dato sistema di ipotesi. Il test più potente trovato è anche uniformemente più potente (per alternative unilaterali)?
- e) Sia $\alpha = 0.05$, $n = 10$ e $\prod_{i=1}^{10} x_i = 0.00012$. Alla luce di questi dati, accettate o meno l'ipotesi H_0 ? E se α fosse uguale a 0.01? E se fosse $\alpha = 0.10$?

3) Sia (X_1, X_2, \dots, X_n) un campione casuale da una distribuzione avente funzione di densità $f(x; \theta)$. Dimostrare che massimizzare rispetto a θ la funzione di verosimiglianza $L(\theta|\mathbf{x})$ è equivalente a massimizzare rispetto a θ la funzione di log-verosimiglianza $\ell(\theta|\mathbf{x})$.

4) Alcuni ricercatori sono interessati ad acquisire informazioni sulla durata minima del periodo di incubazione di una certa malattia. Un campione di n cavie viene quindi esposto all'agente infettivo e per ogni cavia viene rilevato il tempo X trascorso fino alla comparsa dei primi sintomi della malattia. Si ritiene ragionevole ipotizzare che la variabile casuale X segua un modello parametrico di Pareto la cui densità è data da

$$f(x; \lambda, \theta) = \theta \lambda^\theta x^{-(\theta+1)} \mathbb{I}_{[\lambda, +\infty)}(x),$$

in cui $\lambda > 0$ rappresenta la quantità di interesse mentre il parametro $\theta > 0$ è supposto noto, $\theta = 2$.

- a) La famiglia di distribuzioni associata alla precedente densità appartiene a famiglia esponenziale?
- b) Trovare una statistica sufficiente e minimale (che indichiamo con S_n) per λ .
- c) Trovare lo stimatore di massima verosimiglianza per λ , stabilendo se esso è non distorto.
- d) Calcolare la funzione di ripartizione della statistica sufficiente individuata al punto b) e stabilire se esiste uno stimatore non distorto per il parametro λ nella classe degli stimatori del tipo $c S_n$ con c costante opportuna da calcolare.

5) Sapendo che $n = 36$, $\bar{x}_n = 15$, $\bar{Y}_n = 8$, $s_x^2 = 9$, $s_Y^2 = 4$ e $\text{Corr}(x, Y) = \frac{1}{3}$, trovare la stima a minimi quadrati della retta di regressione di Y su x e valutarne il grado di bontà di adattamento ai dati osservati riassunti nelle misure di sintesi poc'anzi proposte. Commentare il risultato.