

Calcolo delle Probabilità e Statistica

Test di verifica - 17/05/2004

Il candidato risolva due (e solo due) problemi tra i seguenti, motivando le proprie risposte.

1. Sia X una variabile aleatoria avente densità

$$f(x) = \begin{cases} \frac{1}{2x^2} & \text{per } |x| > 1, \\ 0 & \text{altrimenti} \end{cases}$$

- (a) Calcolare la funzione di ripartizione di X .
 - (b) Calcolare la densità di $Y = \log |X|$. Si tratta di una densità nota?
 - (c) Calcolare $\mathbf{E}[\sqrt{|X|Y}]$ e mostrare che la v.a. XY non ha speranza finita.
2. Domenica prossima in una data città si terrà il ballottaggio per l'elezione a sindaco tra i due candidati A e B . Dai sondaggi effettuati, risulta che il candidato A ha probabilità p di essere votato da un generico elettore (e per il candidato B la stessa probabilità vale $1 - p$ - si esclude cioè la possibilità di scheda bianca o nulla). Consideriamo la famiglia di v. a.

$$X_i = \begin{cases} 1 & \text{se l'elettore } i\text{-esimo vota per } A \\ 0 & \text{se l'elettore } i\text{-esimo vota per } B. \end{cases}$$

- (a) Qual è la legge di ciascuna X_i ?
 - (b) Supponiamo che gli elettori siano 1200 e che votino in modo indipendente. Se $p = 0.52$, qual è la probabilità che il candidato A sia eletto (cioè che ottenga più della metà dei voti)?
 - (c) Nell'ipotesi che gli elettori siano n e che la probabilità che A sia votato dal generico elettore sia p , calcolare, per n grande ($n \rightarrow \infty$) e in funzione di p la probabilità che A sia eletto.
3. 400 osservazioni indipendenti di una v.a. normale X di varianza $\sigma^2 = 1$ hanno fornito un intervallo di fiducia bilaterale per la media μ di X avente lunghezza 0.1.
- (a) Qual è il livello dell'intervallo?
 - (b) Qual è la lunghezza dell'intervallo di fiducia bilaterale calcolato a partire da 900 osservazioni della stessa variabile e del livello calcolato al punto (a) ?
 - (c) La lunghezza del secondo intervallo è minore o maggiore di quella del primo? Perché? Giustificare opportunamente la risposta.

Calcolo delle Probabilità e Statistica

Soluzioni del test di verifica - 17/05/2004

1. (a) Distinguiamo i tre casi: (i) $x < -1$, (ii) $-1 \leq x < 1$, (iii) $x \geq 1$. Per $x < -1$ si ha

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^x \frac{1}{2t^2}dt = -\frac{1}{2x}$$
 (si noti che $x < 0$, quindi $F(x) = -1/2x > 0$). Nel caso (ii), si ha:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^{-1} \frac{1}{2t^2}dt = \frac{1}{2}$$
 dato che $f(t) = 0$ per $-1 \leq t \leq 1$. Infine, nel caso (iii), avremo:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^{-1} \frac{1}{2t^2}dt + \int_1^x \frac{1}{2t^2}dt = 1 - \frac{1}{2x}.$$
 Per quanto riguarda quest'ultimo caso, si poteva anche ragionare nel modo seguente: dato che X è simmetrica, si avrà $F(x) = 1 - F(-x)$ e quindi per $x > 1$, si ha $F(x) = 1 - (-1/(-2x)) = 1 - 1/2x$.
- (b) Cominciamo con il calcolare la funzione di ripartizione di Y .

$$P(Y \leq y) = P(\log |X| \leq y) = P(|X| \leq e^y) = P(-e^y \leq X \leq e^y) = F(e^y) - F(-e^y).$$
 Bisogna dunque distinguere i due casi: (i) $y < 0$ e (ii) $y \geq 0$, in cui si ha rispettivamente $e^y < 1$ e $e^y \geq 1$. Nel caso (i) si ha $P(Y \leq y) = 0$, poiché $F(e^y) = F(-e^y) = 1/2$. Nel caso (ii) si ha invece $F(e^y) = 1 - \frac{1}{2e^y} = 1 - e^{-y}/2$ e $F(-e^y) = -\frac{1}{2(-e^y)} = e^{-y}/2$ da cui segue $P(Y \leq y) = 1 - e^{-y}$. Senza necessità di ulteriori calcoli, si riconosce la densità $\mathcal{E}(1)$ (esponenziale di parametro 1).
- (c) Si ha

$$\begin{aligned} \mathbf{E}[\sqrt{|X|}Y] &= \mathbf{E}[\sqrt{|X|} \log |X|] = \frac{1}{2} \int_{-\infty}^{-1} \frac{\sqrt{|x|} \log |x|}{x^2} dx + \frac{1}{2} \int_1^{+\infty} \frac{\sqrt{|x|} \log |x|}{x^2} dx \\ &= \int_1^{+\infty} \frac{\sqrt{x} \log x}{x^2} dx = \int_1^{+\infty} \frac{\log x}{x\sqrt{x}} dx. \end{aligned}$$

Con il cambiamento di variabile $y = \log x$ l'ultimo integrale si trasforma in

$$\int_0^{+\infty} ye^{-y/2} dy = 2 \int_0^{+\infty} \frac{y}{2} e^{-y/2} dy = 4.$$

Si può anche osservare che, dato che la funzione $\log z$ è invertibile, si ha $|X| = e^Y$ e quindi $E[\sqrt{|X|}Y] = E[Ye^{Y/2}]$, da cui si ricava

$$E[\sqrt{|X|}Y] = E[e^{Y/2}Y] = \int_0^{\infty} ye^{y/2}e^{-y}dy = \int_0^{+\infty} ye^{-y/2}dy$$

e si conclude come sopra. In modo analogo, si verifica che

$$E[|XY|] = E[|X|Y] = E[e^Y Y] = \int_0^{\infty} ye^ye^{-y}dy = \int_0^{+\infty} ydy = +\infty$$

e quindi XY non ha speranza matematica finita.

2. (a) Ciascuna X_i ha evidentemente legge $B(1, p)$ (bernoulliana di parametro p).
 (b) La probabilità richiesta è $P\left(\sum_{i=1}^n X_i > \frac{n}{2}\right)$. che si può scrivere equivalentemente nella forma

$$P\left(\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} > \frac{n((1/2) - p)}{\sqrt{np(1-p)}}\right) = P\left(\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} > \frac{\sqrt{n}((1/2) - p)}{\sqrt{p(1-p)}}\right)$$

Per $n = 1200$, $p = 0.52$, si può usare l'approssimazione normale (dato che sono soddisfatte entrambe le condizioni $np \geq 5$, $n(1-p) \geq 5$). Si ha pertanto che la quantità precedente è approssimativamente uguale a

$$1 - \Phi\left(\frac{\sqrt{1200}(0.5 - 0.52)}{\sqrt{0.52(1 - 0.52)}}\right) = 1 - \Phi(-1.39) = \Phi(1.39) = 0.9177.$$

- (c) Come nel punto precedente, possiamo dire che per n grande, la probabilità che A sia eletto è approssimativamente uguale a

$$1 - \Phi\left(\frac{\sqrt{n}((1/2) - p)}{\sqrt{p(1-p)}}\right) = \Phi\left(\frac{\sqrt{n}(p - 1/2)}{\sqrt{p(1-p)}}\right).$$

Per $n \rightarrow \infty$, si ha

$$\lim_{n \rightarrow \infty} \frac{\sqrt{n}(p - 1/2)}{\sqrt{p(1-p)}} = \begin{cases} -\infty & \text{se } p < 1/2 \\ 0 & \text{se } p = 1/2 \\ +\infty & \text{se } p > 1/2. \end{cases}$$

Quindi

$$\lim_{n \rightarrow \infty} \Phi\left(\frac{\sqrt{n}(p - 1/2)}{\sqrt{p(1-p)}}\right) = \begin{cases} 0 & \text{se } p < 1/2 \\ \frac{1}{2} & \text{se } p = 1/2 \\ 1 & \text{se } p > 1/2. \end{cases}$$

3. (a) Dalla formula dell'intervallo di fiducia bilaterale per la media di una normale con varianza nota $\left(\bar{X} - \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}\right)$ si ricava che la sua lunghezza è

$$2\frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}.$$

(dove ricordiamo che ϕ_α denota il quantile di ordine α della distribuzione $\mathcal{N}(0, 1)$). Sapendo che $n = 400$, $\sigma^2 = 1$ e la lunghezza è pari a 0.1, si ricava l'equazione

$$\phi_{1-\alpha/2} = 1,$$

ovvero $1 - \alpha/2 = \Phi(1) = 0.841$ e $\alpha = 0.318$. Quindi il livello dell'intervallo è $1 - \alpha = 0.682$.

- (b) Utilizzando il fatto che $\phi_{1-\alpha/2} = 1$ e $n = 900$, si trova che la lunghezza richiesta è

$$2\frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2} = \frac{2}{30} = \frac{1}{15} = 0.067.$$

- (c) La lunghezza trovata nel punto (b) risulta inferiore a quella del punto (a): maggiore è il numero di osservazioni, migliore è l'approssimazione che si trova.