Statistica I. Ingegneria Gestionale. Scritto del 17/07/2012

Cerchiare, su questo foglio, le risposte corrette e risolvere per esteso gli esercizi sui fogli assegnati.

Esercizio 1. Un operatore finanziario osserva i dati in rete relativi al prezzo giornaliero di un certo prodotto finanziario.

i) (3 pt) Visualizza un istogramma dei valori relativi a 25 giorni e decide inizialmente di usare una v.a. esponenziale di parametro λ per il prezzo giornaliero. Che probabilità c'è il prezzo dei giorni futuri superi 2μ, dove μ è la media?

0.035, 0.135, 0.235, 0.335, 0.435.

ii) (3 pt) Con più dati a disposizione decide di usare una gaussiana per il prezzo giornaliero. Dai dati stima media e deviazione pari a 27.5 e 12.2. Rispondere ora alla domanda precedente.

0.412, 0.312, 0.212, 0.112, 0.012.

iii) (4 pt) Vuole trovare un intervallo $[\mu - \delta, \mu + \delta]$ entro cui stia il prezzo giornaliero, con probabilità 0.9. Quanto vale δ? Aiutarsi con un grafico.

16.069, 18.069, 20.069, 22.069, 24.069.

iv) (4 pt) Un paio di mesi dopo decide di controllare se il prezzo sia rimasto statisticamente invariato. Utilizza i valori degli ultimi 30 giorni, che hanno una media pari a 32.5. Fino a quale livello di significatività si arriva a concludere che le cose sono cambiate? Non usare formule preconfezionate; spiegare bene cosa volete calcolare e che procedimento usate.

0.0148, 0.0248, 0.0348, 0.0448, 0.0548.

v) (4 pt) L'operatore chiama di classe A ogni giornata in cui il prezzo è superiore a 35, di classe B quando è inferiore. Supponiamo che l'immediato futuro sia simile agli ultimi 30 giorni. Calcolare la probabilità che ci siano 3 giornate A su 5, nelle prossime 5. Introdurre opportune v.a. aleatorie e formulare con esse il problema.

0.248, 0.348, 0.448, 0.548, 0.648.

vi) (3 pt) Calcolare la probabilità che, tra le prossime 30 giornate, ce ne siano più di 10 di classe A.

0.4287, 0.5287, 0.6287, 0.7287, 0.8287.

Esercizio 2. i) (3 pt) Sia X una v.a. di Poisson di parametro λ. Siano X_1 e X_2 due copie indipendenti di X. Trovare λ per cui $E\left[(X_1 - X_2)^2\right] = 1$.

0.2, 0.3, 0.4, 0.5, 0.6.

ii) (3 pt) Assumendo $\lambda = 2$, calcolare $P(X_1 + X_2 \geq 2)$.

iii) (3 pt) Assumendo $\lambda = 2$, calcolare $E[g(X)]$ dove $g(x) = x$ per $x \leq 2$, $g(x) = 2$ per $x \geq 2$.

$$1.2587, \qquad 1.4587, \qquad 1.6587, \qquad 1.8587, \qquad 2.0587.$$

Quantili utili per questo compito:

$x =$	-2.357	-0.949	0.205	1.645	2.245	2.254
$\Phi(x) =$	0.0092	0.1713	0.5812	0.95	0.9876	0.9879

Formule e teoremi potenzialmente utili. 1) Se X è $B(n,p)$, allora $P(N = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $E[X] = np$, $Var[X] = np(1-p)$; se X è $\mathcal{P}(\lambda)$, allora $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $E[X] = Var[X] = \lambda$; se X è esponenziale di parametro λ, $F(x) = 1 - e^{-\lambda x}$, $E[X] = \sigma[X] = \frac{1}{\lambda}$. 2) Se $X_1, ..., X_n$ sono $N(\mu, \sigma^2)$, allora $P\left(\left|\overline{X} - \mu\right| \leq \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1 - \alpha$. 3) Se $X_1, ..., X_n$ sono $N(\mu, \sigma^2)$, allora $Z = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n}$ è una $N(d, 1)$ con $d = \frac{\mu - \mu_0}{\sigma}\sqrt{n}$.

1 Soluzioni

Esercizio 1. i) Detto X il prezzo giornaliero, ricordando che $\mu = \frac{1}{\lambda}$,

$$P(X > 2\mu) = e^{-\lambda 2\mu} = e^{-2} = 0.135.$$

ii)

$$\begin{aligned} P(X > 2\mu) &= 1 - \Phi\left(\frac{2\mu - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\mu}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{27.5}{12.2}\right) = 1 - \Phi(2.254) \\ &= 1 - 0.9879 = 0.012. \end{aligned}$$

iii) Una risoluzione grafica semplifica molto i passaggi. Dal grafico si vede che il punto $\mu+\delta$ ha alla sua sinistra area pari a 0.95, quindi soddisfa l'equazione

$$P(X \leq \mu + \delta) = 0.95.$$

Coi soliti passaggi si arriva a

$$\mu + \delta = \mu + \sigma q_{0.95} = \mu + 12.2 \cdot 1.645$$

da cui

$$\delta = 12.2 \cdot 1.645 = 20.069.$$

iv) Il valore estremo della significatività è il p-value: per $\alpha > p$ il test risulta significativo, cioè si afferma che le cose sono cambiate. Dobbiamo quindi calcolare il p-value del test. Il test in questione è quello bilaterale, visto che non c'è alcuna indicazione a priori sulla direzione. Non usiamo la formula finale del p-value ma partiamo dalla sua definizione e, nello specifico di questo esercizio, dalla sua definizione in quanto valore estremo della significatività per cui il test risulta significativo. Quando $|z| > q_{1-\frac{\alpha}{2}}$ ($z = \frac{\overline{x}-\mu_0}{\sigma}\sqrt{n}$), il test è significativo, quindi il valore limite p si trova risolvendo

$$|z| = q_{1-\frac{p}{2}}$$

ovvero $1 - \frac{p}{2} = \Phi(|z|)$,

$$p = 2 - 2\Phi(|z|).$$

Sostituendo,

$$\begin{aligned} p &= 2 - 2\Phi\left(\frac{32.5 - 27.5}{12.2}\sqrt{30}\right) = 2 - 2 \cdot \Phi(2.245) \\ &= 2 - 2 \cdot 0.9876 = 0.0248 \end{aligned}$$

v) Chiamiamo $Y_1, ..., Y_5$ delle v.a. di Bernoulli che valgono 1 se il relativo giorno è di classe A. Sono $B(1, p)$ dove

$$\begin{aligned} p &= P(X > 35) = 1 - \Phi\left(\frac{35 - 32.5}{12.2}\right) = 1 - \Phi(0.205) \\ &= 1 - 0.5812 = 0.4188. \end{aligned}$$

La loro somma $S = Y_1 + ... + Y_5$ è una $B(5, 0.4188)$ e rappresenta il numero di giorni di classe A. Dobbiamo pertanto calcolare

$$P(S = 3) = \binom{5}{3} 0.4188^3 (1 - 0.4188)^2 = 0.248.$$

vi) Con notazioni simili,

$$\begin{aligned} P(S > 10) &= P(Y_1 + ... + Y_{30} > 10) \\ &= P\left(\frac{Y_1 + ... + Y_{30} - 30 \cdot p}{\sqrt{30 \cdot p \cdot (1-p)}} > \frac{10 - 30 \cdot p}{\sqrt{30 \cdot p \cdot (1-p)}}\right) \end{aligned}$$

che per il TLC è circa uguale a

$$\begin{aligned} 1 - \Phi\left(\frac{10 - 30 \cdot 0.4188}{\sqrt{30 \cdot 0.4188 \cdot (1 - 0.4188)}}\right) &= 1 - \Phi(-0.949) \\ &= 1 - 0.1713 = 0.8287. \end{aligned}$$

Esercizio 2. i)

$$\begin{aligned} E\left[(X_1 - X_2)^2\right] &= E\left[X_1^2 + X_2^2 - 2X_1X_2\right] \\ &= E\left[X_1^2\right] + E\left[X_2^2\right] - 2E[X_1]E[X_2] \\ &= 2E\left[X_1^2\right] - 2E[X_1]^2 = 2E\left[X_1^2\right] - 2\lambda^2 \end{aligned}$$

$$E\left[X_1^2\right] = Var[X_1] + E[X_1]^2 = \lambda + \lambda^2$$

quindi

$$1 = E\left[(X_1 - X_2)^2\right] = 2\left(\lambda + \lambda^2\right) - 2\lambda^2 = 2\lambda$$

da cui $\lambda = 0.5$.

ii)

$$P(X_1 + X_2 \geq 2) = 1 - P(X_1 + X_2 \leq 1)$$

$$\begin{aligned} P(X_1 + X_2 \leq 1) &= P(X_1 = 0, X_2 = 0) + P(X_1 = 0, X_2 = 1) \\ &+ P(X_1 = 1, X_2 = 0) \\ &= e^{-\lambda}e^{-\lambda} + 2e^{-\lambda}e^{-\lambda}\lambda = e^{-2\lambda} + 2e^{-2\lambda}\lambda \end{aligned}$$

da cui

$$P(X_1 + X_2 \geq 2) = 1 - e^{-2\lambda} - 2e^{-2\lambda}\lambda.$$

Si potrebbe usare il fatto che $Y = X_1 + X_2$ è una Poisson di parametro 2λ per semplificare:

$$P(Y \geq 2) = 1 - P(Y \leq 1) = 1 - e^{-2\lambda} - e^{-2\lambda}2\lambda.$$

A questo punto si sostituisce:

$$P\left(Y \geq 2\right) = 1 - e^{-4} - e^{-4}4 = 0.908.$$

iii) La v.a. $Y = g\left(X\right)$ assume i valori 0,1,2, con probabilità:

$$\begin{aligned}
&P\left(Y = 0\right) = P\left(X = 0\right) = e^{-\lambda} \\
&P\left(Y = 1\right) = P\left(X = 1\right) = e^{-\lambda}\lambda \\
&P\left(Y = 2\right) = P\left(X \geq 2\right) = 1 - P\left(X < 2\right) = 1 - e^{-\lambda} - e^{-\lambda}\lambda.
\end{aligned}$$

Pertanto,

$$\begin{aligned}
E\left[g\left(X\right)\right] &= e^{-\lambda}\lambda + 2\left(1 - e^{-\lambda} - e^{-\lambda}\lambda\right) \\
&= 2 - 2e^{-2} - e^{-2}2 = 1.4587.
\end{aligned}$$