Statistica I. Ingegneria Gestionale. Scritto del 31/01/2012

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

Esercizio 1. Si vuole misurare il livello di traffico su un tratto di strada ed i cambiamenti che possono essere indotti da un nuovo raccordo.

i) (3 pt) Intanto si vuole stimare il traffico medio orario nelle ore di punta. Si deve decidere quante ore (di punta) di osservazione servono per stimare il traffico medio in modo che l'errore relativo sia al massimo del 10% e che ci si possa fidare del risultato al 95%. Che formula si deve usare e come si può fare per rimediare alla mancanza di certi dati? Impostare in modo esaustivo il problema, senza usare formulette preconfezionate.

ii) (3 pt) Raccolti i dati per 20 ore, cioè dei numeri $x_1, ..., x_{20}$, si trova $\overline{x} = 947.3$, $s = 85.6$. Che probabilità c'è in una generica ora passino più di 1100 macchine?

..., ..., ..., ..., 0.037

iii) (4 pt) Considerando anche l'incertezza nella stima della media al 95%, qual'è il numero massimo di macchine che ci si deve aspettare, in una generica ora, 9 volte su 10?

..., ..., ..., ..., 1094.4.

iv) (4 pt) In un periodo di punta si considerano 40 ore indipendenti ed il totale dei passaggi in esse, detto N. Calcolare, la probabilità che N sia maggiore di 40000.

..., ..., ..., ..., 0.00005.

v) (4 pt) Viene costruito e messo in funzione il nuovo raccordo, che avrebbe lo scopo di diminuire il traffico sul tratto di strada considerato fino ad ora. Si decide di adottare un test unilaterale al 90% per capire se il nuovo raccordo ha effetto. Il test sarà basato sull'osservazione di altre 20 ore. Se il numero medio di passaggi diminuirà di 80 unità, che probabilità ci sarà di non accorgersene con questo test? Scrivere matematicamente la probabilità richiesta e, partendo da lì, ricavare il risultato, usando solamente le formule riportate sotto.

..., ..., ..., ..., 0.002.

vi) (4 pt) Le nuove 20 osservazioni hanno dato $\overline{x} = 805$. Supponiamo che si decida di usare una v.a. esponenziale per descrivere il numero di passaggi all'ora. Qual'è il numero massimo di macchine che ci si deve aspettare, in una generica ora, 9 volte su 10? Si commenti il risultato dal punto di vista del realismo applicativo.

..., ..., ..., ..., 1853.6.

Esercizio 2. i) (3 pt) Calcolare $P(|X| > 1)$ se $X \sim N(0, 2)$.

..., ..., ..., ..., 0.479 6.

ii) (2 pt) Se Y vale ± 1 con egual probabilità, calcolare la sua varianza.

$$..., \quad ..., \quad ..., \quad ..., \quad = 0.477.$$

ii) (3 pt) Se X ed Y sono indipendenti, calcolare $P\left(\left(\frac{X}{\sqrt{2}} - Y\right)^2 < 1\right)$.

Quantili utili per questo compito:

$x =$	0.707	1.28	1.784	1.96	2	2.8996	3.297	3.894
$\Phi(x) =$	0.7602	0.9	0.963	0.975	0.977	0.998	0.9995	0.99995

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) Vedere un compito precedente.

ii) Descritto il traffico orario con una v.a. gaussiana X di media 947.3 e deviazione 85.6, vale

$$P(X > 1100) = 1 - \Phi\left(\frac{1100 - 947.3}{85.6}\right) = 1 - \Phi(1.784)$$
$$= 1 - 0.963 = 0.037.$$

iii) Per la stima della media al 95% abbiamo ($\alpha = 0.05$, $\alpha/2 = 0.025$, $1 - \alpha/2 = 0.975$, $q_{1-\alpha/2} = 1.96$)

$$\mu = 947.3 \pm \frac{85.6 \cdot 1.96}{\sqrt{20}} = 947.3 \pm 37.516$$

quindi la media potrebbe anche essere

$$947.3 + 37.516 = 984.82.$$

In tal caso il numero massimo 9 volte su 10 sarebbe

$$984.82 + 85.6 \cdot q_{0.9} = 984.82 + 85.6 \cdot 1.28 = 1094.4.$$

iv) Avendo fin qui supposto che il numero di passaggi fosse gaussiano, possiamo affermare che N è gaussiana di media $40 \cdot 947.3$ e varianza $40 \cdot 85.6^2$. Quindi

$$P(N > 40000) = 1 - \Phi\left(\frac{40000 - 40 \cdot 947.3}{\sqrt{40} \cdot 85.6}\right) = 1 - \Phi(3.894)$$
$$= 1 - 0.99995 = 0.00005.$$

v) Le premesse ed i ragionamenti sono simili ad altri esercizi e non vengono qui ripetute. Non ci accorgiamo della diminuzione se $Z \geq -q_{0.9}$, dove però Z è una $N(d, 1)$ con $d = \frac{\mu - \mu_0}{\sigma}\sqrt{n}$ e qui vale

$$d = \frac{\mu - \mu_0}{\sigma}\sqrt{n} = \frac{-80}{85.6}\sqrt{20} = -4.1796.$$

Vale

$$P(Z \geq -q_{0.9}) = 1 - \Phi(-q_{0.9} - d) = 1 - \Phi(-1.28 + 4.1796)$$
$$= 1 - \Phi(2.8996) = 1 - 0.998 = 0.002.$$

vi) Ora X è esponenziale. La media $E[X] = \lambda^{-1}$ si può stimare con $\overline{x} = 805$, quindi $\lambda \sim \frac{1}{805} = 0.00124$. Cerchiamo ora il numero t tale che

$$P(X \leq t) = 0.9$$

cioè

$$1 - e^{-\lambda t} = 0.9$$

ovvero $e^{-\lambda t} = 0.1$, $\lambda t = -\log 0.1$,

$$t = -\frac{\log 0.1}{\lambda} = -805 \cdot \log 0.1 = 1853.6.$$

Il risultato è troppo pessimistico persino rispetto alla stima cautelativa fatta sopra. L'"errore" sta nel supporre X esponenziale. Se così fosse, la deviazione standard sarebbe uguale alla media, cioè enorme (da qui la stima così pessimistica), fatto in disaccordo con i numeri dati all'inizio.

Esercizio 2. i)

$$\begin{aligned} P(|X| > 1) &= 1 - P(|X| \leq 1) = 1 - P(-1 \leq X \leq 1) \\ &= 1 - \Phi\left(\frac{1-0}{\sqrt{2}}\right) + \Phi\left(\frac{-1-0}{\sqrt{2}}\right) \\ &= 2 - 2 \cdot \Phi(0.707) = 2 - 2 \cdot 0.7602 = 0.479\,6 \end{aligned}$$

ii) $E[Y] = 0$, Y^2 vale 1 con probabilità 1, quindi $E\left[Y^2\right] = 1$, da cui infine

$$Var[Y] = 1.$$

iii)

$$\begin{aligned} &P\left(\left(\frac{X}{\sqrt{2}} - Y\right)^2 < 1\right) \\ &= P\left(\left(\frac{X}{\sqrt{2}} - Y\right)^2 < 1 | Y = 1\right) P(Y=1) + P\left(\left(\frac{X}{\sqrt{2}} - Y\right)^2 < 1 | Y = -1\right) P(Y=-1) \\ &= \frac{1}{2} P\left(\left(\frac{X}{\sqrt{2}} - 1\right)^2 < 1\right) + \frac{1}{2} P\left(\left(\frac{X}{\sqrt{2}} + 1\right)^2 < 1\right) \\ &= \frac{1}{2} P\left(-1 < \frac{X}{\sqrt{2}} - 1 < 1\right) + \frac{1}{2} P\left(-1 < \frac{X}{\sqrt{2}} + 1 < 1\right) \\ &= \frac{1}{2} P\left(0 < \frac{X}{\sqrt{2}} < 2\right) + \frac{1}{2} P\left(-2 < \frac{X}{\sqrt{2}} < 0\right) \\ &= \frac{1}{2} P\left(-2 < \frac{X}{\sqrt{2}} < 2\right) = \frac{1}{2}(\Phi(2) - \Phi(-2)) = \frac{1}{2}(2\Phi(2) - 1) \\ &= 0.977 - 0.5 = 0.477. \end{aligned}$$