**Statistica I. Ingegneria Gestionale. Scritto del 26/06/2012**

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

**Esercizio 1.** Si possiede un foglio Excel su cui sono state registrate le vendite giornaliere di piastrelle di un'azienda. Esaminiamo, a seconda del quesito, due scenari (domande i-iv; v-vi).

i) (3 pt) Si hanno i dati di 50 giorni, si decide di usare una gaussiana per descrivere il volume di vendite giornaliere, si calcolano media e deviazione standard campionarie, che risultano pari a 25, 12. Che precisione relativa abbiamo, al 90%?

$$0.311\,61, \qquad 0.111\,67, \qquad 0.213\,67, \qquad 0.011\,52, \qquad 0.414\,64.$$

ii) (4 pt) Che probabilità c'è di ricevere una richiesta giornaliera superiore a 40 piastrelle? Impostare il problema e corredare la risposta di una spiegazione grafica.

$$0.006, \qquad 0.206, \qquad 0.013, \qquad 0.106, \qquad 0.316.$$

iii) (3 pt) Se ogni piastrella comporta un guadagno di 5 euro, che guadagno minimo al 90% ci aspettiamo? Impostare il problema e corredare la risposta di una spiegazione grafica.

$$9.622, \qquad 8.622, \qquad 7.622, \qquad 6.622, \qquad 5.622.$$

iv) (4 pt) L'anno successivo pensiamo che le vendite possano essere cambiate. Da altre 50 registrazioni giornaliere si calcola una media pari a 21. 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.

$$1.4216, \qquad 1.5816, \qquad 1.7116, \qquad 1.8216, \qquad 1.9816.$$

v) (4 pt) Cambiamo scenario. Si osserva che i dati giornalieri sono frammentari: o non si è venduto nulla oppure uno stock da 50 piastrelle; per cui una descrizione con una gaussiana non è sensata. Supponiamo che, su 300 giorni, 210 volte non si sia venduto nulla. Supponiamo i giorni indipendenti. Raggruppiamo le vendite a gruppi di 30 giorni, per osservare valori meno discreti. Indichiamo con $X$ la vendita relativa a 90 giorni. Che distribuzione usereste per $X$ (perché) e come calcolereste la probabilità che $X$ superi 600?

$$0.1161, \qquad 0.2261, \qquad 0.3361, \qquad 0.4461, \qquad 0.5561.$$

vi) (4 pt) Nello scenario del punto (v), detto $Y$ il numero di piastrelle venduto in 5 giorni, che probabilità c'è che sia pari a 100 (esattamente)? (Si suggerisce di rapportare lo stock da 50 pezzi all'unità).

$$0.1087, \qquad 0.2087, \qquad 0.3087, \qquad 0.4087, \qquad 0.5087.$$

**Esercizio 2.** i) (3 pt) Sia $X$ una v.a. che assume i valori $-1, 0, 1$ con egual probabilità. Calcolare $Var[X+1]$.

1/3, 2/3, 1, 4/3, 5/3.

ii) (3 pt) Se $X_1, X_2$ sono due copie indipendenti di $X$, calcolare $P(X_1 + X_2 \leq 0)$.

1/6, 1/3, 1/2, 2/3, 5/6.

iii) (3 pt) Calcolare $E\left[\frac{X_1+1}{1+|X_2|}\right]$.

1/3, 2/3, 1, 4/3, 5/3.

**Quantili utili per questo compito:**

| | | | | | | |
|---|---|---|---|---|---|---|
| $x =$ | $-2.357$ | $1.031$ | $1.25$ | $1.195$ | $1.2815$ | $1.645$ |
| $\Phi(x) =$ | $0.0092$ | $0.8487$ | $0.894$ | $0.8839$ | $0.9$ | $0.95$ |

**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 $\lambda$, $F(x) = 1 - e^{-\lambda x}$, $E[X] = \sigma[X] = \frac{1}{\lambda}$. 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) La precisione relativa al 90% è data dalla formula ($q_{1-\frac{\alpha}{2}} = q_{1-\frac{0.1}{2}} = q_{0.95}$) $\frac{\sigma q_{0.95}}{\sqrt{n}|\mu|}$ che approssimiamo a

$$\frac{\sigma q_{0.95}}{\sqrt{n}\,|\mu|} \simeq \frac{12 \cdot q_{0.95}}{\sqrt{50} \cdot |25|} = \frac{12 \cdot 1.645}{\sqrt{50} \cdot 25} = 0.111\,67.$$

ii) Detta $V$ la v.a. "vendita giornaliera", dobbiamo calcolare

$$\begin{aligned} P(V > 40) &= 1 - \Phi\left(\frac{40-25}{12}\right) = 1 - \Phi(1.25) \\ &= 1 - 0.894 = 0.106. \end{aligned}$$

iii) Detta $V$ la v.a. "vendita giornaliera", e $G$ il "guadagno", $G = 5V$, dobbiamo cercare $\lambda$ tale che

$$0.1 = P(G < \lambda) = P(5V < \lambda) = P(V < \lambda/5).$$

Ragionando come al solito (lo studente deve riportare le giustificazioni) si trova

$$\lambda/5 = \mu - \sigma q_{0.9} = 25 - 12 \cdot 1.2815 = 9.622.$$

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,

$$\begin{aligned} p &= 2 - 2\Phi\left(\frac{21-25}{12}\sqrt{50}\right) = 2 - 2\Phi(-2.357) \\ &= 2 - 2 \cdot 0.0092 = 1.9816. \end{aligned}$$

v) La frequenza empirica di vendita di uno stock è $\widehat{p} = \frac{90}{300} = 0.3$. La vendita giornaliera $V$ si può scrivere come

$$V = 50 \cdot W$$

dove $W$ è una Bernoulli $B(1, 0.3)$. Dette $V_1, ..., V_{30}$ le vendite di 30 giorni, poniamo $X = V_1 + ... + V_{30}$. Per il teorema limite centrale,

$$\frac{V_1 + ... + V_{30} - 30 \cdot E[V]}{\sqrt{30}\sqrt{Var[V]}}$$

è approssimativamente gaussiana, quindi (anche se un po' intuitivamente) lo è anche $V_1 + ... + V_{30}$. Quindi, per $X$ usiamo come modello una

$$N\left(30 \cdot 0.3 \cdot 50, 30 \cdot 50^2 \cdot 0.3 \cdot 0.7\right) = N(450, 15750).$$

Quindi

$$\begin{aligned} P(X > 600) &= 1 - \Phi\left(\frac{600 - 450}{\sqrt{15750}}\right) = 1 - \Phi(1.195) \\ &= 1 - 0.8839 = 0.1161. \end{aligned}$$

vi) Pensando al numero 50 (lo stock) come un'unità, le vendite di 5 giorni sono delle Bernoulli $B(1, 0.3)$, che indichiamo con $W_1, ..., W_5$. Il numero di stock venduti è $S = W_1 + ... + W_5$, $B(5, 0.3)$. L'evento $Y = 100$ equivale a $S = 2$. Quindi dobbiamo calcolare

$$P(S = 2) = \binom{5}{2} 0.3^2 0.7^3 = 0.3087.$$

**Esercizio 2.** i)

$$Var[X + 1] = Var[X] = E\left[X^2\right]$$

perché la varianza è invariante per traslazioni e la v.a. $X$ ha media nulla. Poi

$$= 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}.$$

ii)

$$\begin{aligned} P(X_1 + X_2 > 0) &= P(X_1 = 1, X_2 = 1) + P(X_1 = 1, X_2 = 0) \\ &+ P(X_1 = 0, X_2 = 1) \\ &= \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{3} \end{aligned}$$

quindi $P(X_1 + X_2 \leq 0) = 2/3$.

iii)

$$\begin{aligned} E\left[\frac{X_1 + 1}{1 + |X_2|}\right] &= E[X_1 + 1] \cdot E\left[\frac{1}{1 + |X_2|}\right] = E\left[\frac{1}{1 + |X_2|}\right] \\ &= \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{3}. \end{aligned}$$