Metodi Matematici e Statistici
Prova scritta – 14/2/2008

Esercizio 1. Sia X una v.a. e siano $X_1, X_2, ..., X_n$ delle v.a. indipendenti e distribuite come X. Sia $S_n = X_1 + X_2 + ... + X_n$.

i) Se $X \sim Exp(0.5)$, calcolare media e varianza di S_3.

ii) Se $X \sim \mathcal{P}(2)$, calcolare la funzione generatrice di S_5.

iii) Se $X \sim B(10, 0.2)$, che variabile è S_2?

iv) Se $X \sim N(2, 3)$, calcolare $P(X_1 > 1, X_2 < 1)$.

v) Cosa si può dire di $P(S_n \le 2n)$, nei casi i, ii, iii, iv?

Esercizio 2. Indichiamo con N il numero di litri d'acqua venduti da un supermercato in un giorno, numero aleatorio, di solito molto elevato. Ignorando approssimativamente per semplicità il fatto che N sarebbe un numero intero, lo consideriamo come variabile continua e supponiamo per semplicità che sia una variabile gaussiana. Indichiamo la sua media con μ e la deviazione con σ.

i) Per una settimana vengono registrate le vendite d'acqua ottenendo i seguenti valori di N: 236, 312, 289, 350, 302, 245. Il gestore capisce che la stima di μ ottenuta da questi pochi numeri è povera. Quante altre osservazioni deve fare per stimare μ al 95% con una precisione di 24 litri?

ii) Decide comunque di registrare le vendite per 4 settimane e trova un valore medio sperimentale di 285 litri ed una deviazione standard sperimentale di 38 litri. Se vuole accontentare tutta la clientela il 99% dei giorni, quanti litri deve tenere pronti ogni giorno?

iii) Con la scelta fatta, accade che 3 volte su cento l'acqua non basta. A cosa può attribuire quest'errore di previsione?

iv) Dopo un anno ripete i controlli e trova, in 4 settimane, un valore medio sperimentale pari a 307 litri. La variazione rispetto alla stima precedente è attribuibile al caso oppure ad un aumento dei consumi di acqua? Rispondere al 95%. Trovare anche con quale significatività arriverebbe alla conclusione che una variazione casuale.

1 Soluzioni

Esercizio 1. i)

$$E[S_3] = 3E[X] = 6$$
$$Var[S_3] = 3Var[X] = \frac{3}{0.5^2} = 12.$$

ii)

$$\phi_{S_5}(t) = E\left[e^{tS_5}\right] = E\left[e^{tX_1} \cdots e^{tX_5}\right] = E\left[e^{tX}\right]^5$$
$$= \left(e^{-2}e^{2e^t}\right)^5 = e^{-10}e^{10e^t}.$$

iii) In vari modi (es. dalla generatrice) si riconosce che è una $B(20, 0.2)$.

iv)

$$
\begin{aligned}
P(X_1 > 1, X_2 < 1) &= P(X_1 > 1)\, P(X_2 < 1) \\
&= \Phi\left(\frac{1-2}{\sqrt{3}}\right)\left(1 - \Phi\left(\frac{1-2}{\sqrt{3}}\right)\right) = ...
\end{aligned}
$$

v) In generale, per il TLC, si può dire che

$$P(S_n \leq 2n) = P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq \frac{2n - n\mu}{\sqrt{n}\sigma}\right)$$

dove $\mu = 2$ in tutti i casi, quindi

$$= P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq 0\right) \sim \Phi(0) = 0.5.$$

Nel caso iv il risultato è esatto e non solo approssimato. Nei casi ii e iii si potrebbe usare il fatto che si conoscono le distribuzioni di S_n per svolgere un calcolo esatto.

Esercizio 2. i)

$$
\begin{aligned}
\overline{x} &= \frac{236 + 312 + 289 + 350 + 302 + 245}{6} = 289 \\
S^2 &= \frac{1}{5}\sum_{k=1}^{6}(x_k - 289)^2 = 1832.\,8
\end{aligned}
$$

$$\frac{\sqrt{1832.\,8}q_{0.975}}{\sqrt{n}} \leq 24$$

quindi $n \geq \frac{1832.\,8 \cdot 1.96^2}{24^2} = 12.\,224$, ovvero $n = 13$.

ii) Dobbiamo trovare il numero λ di litri tale che

$$P(N > \lambda) \leq 0.01.$$

Prendendo per buoni media e deviazione sperimentali, vale

$$P(N > \lambda) = 1 - \Phi\left(\frac{\lambda - 285}{38}\right)$$

quindi $\Phi\left(\frac{\lambda-285}{38}\right) = 0.99$, $\frac{\lambda-285}{38} = q_{0.99}$,

$$\lambda = 285 + 38 \cdot q_{0.99} = ...$$

Se invece si accetta solo 38 come deviazione vera e si considera l'incertezza nella stima di μ, si può affermare che al 95% vale (essendo $\frac{38 \cdot 1.96}{\sqrt{24}} = 15.\,203$)

$$269.\,8 \leq \mu \leq 300.\,2.$$

Quindi al 95% vale

$$1-\Phi\left(\frac{\lambda-269.8}{38}\right) \leq P\left(N>\lambda\right) \leq 1-\Phi\left(\frac{\lambda-300.2}{38}\right).$$

Dovendo essere $P\left(N>\lambda\right) \leq 0.01$, il valore più cautelativo di λ è quello che risolve

$$1-\Phi\left(\frac{\lambda-300.2}{38}\right)=0.01$$

ovvero $\Phi\left(\frac{\lambda-300.2}{38}\right)=0.99$,

$$\lambda=300.2+38\cdot q_{0.99}=...$$

iii) Se ha calcolato λ sulla base della media sperimentale 285, l'errore può essere dovuto al fatto che andava usata la stima più cautelativa $\lambda=300.2+38\cdot q_{0.99}$. Se invece ha già usato questa, l'errore è dovuto al caso, è avvenuto cioè qualcosa che si prevedeva accadesse con probabilità trascurabile. Ovviemente un'altra fonte di errore può essere l'uso dell'ipotesi gaussiana.

iv) Basta eseguire un test (che risulta significativo) e poi calcolare il valore p:

$$\begin{aligned} p &= 2-2\Phi\left(|z|\right) \\ z &= \frac{307-285}{38}\sqrt{24}=2.836. \end{aligned}$$

Con significatività $\alpha<p$ avremmo dovuto concludere che non c'era evidenza di un incremento dell'acquisto di acqua.