Statistica per la ricerca sperimentale e tecnologica
Corso di Laurea in Informatica, Università di Roma "Tor Vergata"
Anno accademico: `2006-2007`
Titolare del corso: `Claudio Macci`
Simulazione 2

Esercizio 1. Un'urna contiene 5 palline numerate da 1 a 5. Si estraggono in blocco 2 palline a caso e consideriamo le seguenti variabili aleatorie: X indica il numero di palline con numero *pari* estratte; Y indica il massimo tra i due numeri estratti; Z indica il minimo tra i due numeri estratti.
`D1`) Trovare la densità discreta di X.
`D2`) Trovare la densità congiunta di (Y, Z).

Esercizio 2. Si lanci un dado equo: se esce un numero pari si lancia una moneta equa, se esce un numero dispari si lancia una moneta la cui probabilità che esca testa è $\frac{1}{3}$.
`D3`) Calcolare la probabilità che esca testa.
`D4`) Calcolare la probabilità sia uscito un numero pari nel lancio del dado sapendo che è uscita testa.

Esercizio 3. Un'urna contiene 100 palline numerate da 1 a 100. Si estraggono a caso 50 palline, una alla volta e con reinserimento. Sia X la variabile aleatoria che conta il numero di volte che esce il numero 53.
`D5`) Calcolare $p_X(0)$.
`D6`) Calcolare $p_X(10)$ sfruttando l'approssimazione di Poisson della binomiale.

Esercizio 4. Sia X una variabile aleatoria con densità $f_X(t) = \frac{1}{t \log 2}$ per $t \in [1, 2]$ e $f_X(t) = 0$ altrimenti. Poi sia (X_n) una successione di variabili aleatorie indipendenti e tutte con la stessa distribuzione di X. Infine poniamo $\overline{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$.
`D7`) Calcolare $P(X > 3/2)$.
`D8`) Trovare il valore di m per cui si ha $\lim_{n \to \infty} P(|\overline{X}_n - m| \geq \varepsilon) = 0$ per ogni $\varepsilon > 0$.

Esercizio 5. Siano X una variabile aleatoria con densità $f_X(t) = te^{-t}$ per $t \geq 0$ e $f_X(t) = 0$ per $t < 0$. Sia $Y = [X]$ dove $[x] = \max\{k \in \mathbb{Z} : k \leq x\}$ è la *parte intera* di x.
`D9`) Calcolare $p_Y(0)$.
`D10`) Calcolare $\mathbb{E}[X]$. *Suggerimento*: è utile osservare che X ha distribuzione Gamma con certi parametri e ricordare la speranza matematica di una variabile aleatoria con distribuzione Gamma.

Esercizio 6. Sia X una variabile aleatoria normale con media $\mu = 2$ e varianza $\sigma^2 = 16$.
`D11`) Calcolare $P(X > 0)$.
Ora supponiamo che μ sia incognito e consideriamo un campione casuale di $n = 100$ osservazioni con la stessa distribuzione di X. Il valore della media campionaria è $\overline{x}_n = 2$.
`D12`) Trovare un intervallo di confidenza per μ al livello $1 - \alpha = 0.95$.

Cenno alle soluzioni (*Ogni segnalazione di errori o sviste (sempre possibili) è gradita*)

Esercizio 1.
D1) La variabile aleatoria X ha distribuzione ipergeometrica; in dettaglio si ha $p_X(k) = \frac{\binom{2}{k}\binom{3}{2-k}}{\binom{5}{2}}$ per $k \in \{0,1,2\}$, da cui $p_X(0) = \frac{3}{10}$, $p_X(1) = \frac{6}{10}$ e $p_X(2) = \frac{1}{10}$.
D2) I sottoinsiemi di 2 elementi di $\{1,2,3,4,5\}$ sono $\binom{5}{2} = 10$ e ciascuno ha la stessa probabilità di essere estratto. Quindi la densità congiunta di (Y,Z) è la seguente: $p_{(Y,Z)}(2,1) = p_{(Y,Z)}(3,1) = p_{(Y,Z)}(4,1) = p_{(Y,Z)}(5,1) = p_{(Y,Z)}(3,2) = p_{(Y,Z)}(4,2) = p_{(Y,Z)}(5,2) = p_{(Y,Z)}(4,3) = p_{(Y,Z)}(5,3) = p_{(Y,Z)}(5,4) = \frac{1}{10}$.

Esercizio 2. Sia T l'evento "esce testa" e D l'evento "esce dispari".
D3) Per la formula delle probabilità totali si ha $P(T) = P(T|D)P(D) + P(T|D^c)P(D^c) = \frac{1}{3}\frac{3}{6} + \frac{1}{2}\frac{3}{6} = \left[\frac{1}{3} + \frac{1}{2}\right]\frac{3}{6} = \frac{2+3}{6}\frac{1}{2} = \frac{5}{6}\frac{1}{2} = \frac{5}{12}$.
D4) Per la formula di Bayes, e sfruttando il valore di $P(T)$ calcolato prima, si ha $P(D^c|T) = \frac{P(T|D^c)P(D^c)}{P(T)} = \frac{\frac{1}{2}\frac{3}{6}}{\frac{5}{12}} = \frac{\frac{3}{12}}{\frac{5}{12}} = \frac{3}{5}$.

Esercizio 3. La variabile aleatoria X ha distribuzione ha distribuzione binomiale con parametri $n = 50$ (numero di estrazioni) e $p = \frac{1}{100}$ (probabilità di estrarre il numero 53 in ogni estrazione). Quindi $p_X(k) = \binom{50}{k}(\frac{1}{100})^k(1 - \frac{1}{100})^{50-k}$ per $k \in \{0,\dots,50\}$.
D5) Si ha $p_X(0) = (\frac{99}{100})^{50}$.
D6) Si ha $p_X(10) = \binom{50}{10}(\frac{1}{100})^{10}(\frac{99}{100})^{40}$ e, sfruttando l'approssimazione di Poisson della binomiale $p_X(k) \approx \frac{\lambda^k}{k!}e^{-\lambda}$ con $\lambda = np = 50\frac{1}{100} = \frac{1}{2}$, otteniamo il valore approssimato $p_X(10) \approx \frac{e^{-\frac{1}{2}}}{2^{10}10!}$.

Esercizio 4.
D7) Si ha $P(X > 3/2) = \int_{3/2}^{2} \frac{1}{t \log 2}dt = \left[\frac{\log t}{\log 2}\right]_{t=3/2}^{t=2} = \frac{\log 2 - \log 3/2}{\log 2} = 1 - \frac{\log 3/2}{\log 2}$.
D8) Il valore di m richiesto è $m = \mathbb{E}[X] = \int_{-\infty}^{\infty} t f_X(t)dt = \int_1^2 t\frac{1}{t \log 2}dt = \frac{1}{\log 2}$.

Esercizio 5.
D9) Si ha $p_Y(0) = P(0 \leq X < 1) = \int_0^1 te^{-t}dt = [-te^{-t}]_{t=0}^{t=1} - \int_0^1 -e^{-t}dt = -e^{-1} - [e^{-t}]_{t=0}^{t=1} = -e^{-1} - (e^{-1} - 1) = 1 - 2e^{-1}$.
D10) Osserviamo che $f_X(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)}t^{\alpha-1}e^{-\lambda t}$ per $t \geq 0$ con $\alpha = 2$ e $\lambda = 1$, e $f_X(t) = 0$ per $t < 0$. Quindi seguendo il suggerimento si ha $\mathbb{E}[X] = \frac{\alpha}{\lambda} = 2$.

Esercizio 6.
D11) La v.a. $Z_X = \frac{X-2}{\sqrt{16}}$ è la standardizzata di X e si ha $P(X > 0) = P(\frac{X-2}{\sqrt{16}} > \frac{0-2}{\sqrt{16}}) = P(Z_X > -2/4) = 1 - \Phi(-2/4) = 1 - (1 - \Phi(2/4)) = \Phi(2/4) = \Phi(0.5) = 0.69146$.
D12) L'intervallo di confidenza richiesto è $\left[\overline{x}_n - \phi_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \overline{x}_n + \phi_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right]$. Si ha $n = 100$, $\overline{x}_n = \overline{x}_{100} = 2$, $\sigma = \sqrt{16} = 4$; inoltre $1 - \frac{\alpha}{2} = 0.975$ segue da $1 - \alpha = 0.95$ e quindi $\phi_{1-\frac{\alpha}{2}} = 1.96$. In conclusione l'intervallo di confidenza richiesto è $[1.216, 2.784]$.

Commenti.
D1) Si ha $p_X(0) + p_X(1) + p_X(2) = \frac{3+6+1}{10} = 1$ in accordo con la teoria.
D9) La variabile aleatoria X ha la stessa distribuzione della variabile aleatoria T_2 che indica il tempo aleatorio del secondo evento di un processo di Poisson (N_t), la cui intensità è $\lambda = 1$ (è il secondo evento perché $\alpha = 2$). Quindi possiamo ottenere il risultato in maniera alternativa come segue: $p_Y(0) = P(0 \leq X < 1) = P(T_2 < 1) = P(N_1 \geq 2) = 1 - (P(N_1 = 0) + P(N_1 = 1)) = 1 - \frac{(1\cdot1)^0}{0!}e^{-(1\cdot1)} - \frac{(1\cdot1)^1}{1!}e^{-(1\cdot1)} = 1 - 2e^{-1}$.