Calcolo delle Probabilità 1 2004/05 - Laurea triennale
Foglio esercizi n.10 — Soluzioni

Esercizio 1. Abbiamo n variabili aleatorie $X_1, \dots, X_n$ indipendenti ed equidistibuite. Ciascuna X_i è una prova di Bernoulli di parametro p, quindi $E(X_i) = p$ e $Var(X_i) = p(1-p)$. Se indichiamo la media aritmetica delle X_i con $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^n X_i$, allora la disuguaglianza di Chebyshev diventa

$$P\left(\left\{\left|\overline{Y}_n - p\ \right| > 0,1\right\}\right) \le \frac{1}{n}\frac{p(1-p)}{(0,1)^2}$$

o equivalentemente

$$P\left(\left\{\left|\overline{Y}_n - p\ \right| \le 0,1\right\}\right) \ge 1 - \frac{1}{n}\frac{p(1-p)}{(0,1)^2}$$

Visto che p è incognito, è utile ricordare la disuguaglianza

$$1 - \frac{1}{n}\frac{p(1-p)}{(0,1)^2} \ge 1 - \frac{1}{4(0,1)^2 n}.$$

Affinché $P\left(\left\{\left|\overline{Y}_n - p\ \right| \le 0,1\right\}\right)$ sia maggiore di $0,99$ è sufficiente prendere n tale che

$$1 - \frac{1}{4(0,1)^2 n} \ge 0,99 \quad \Leftrightarrow \quad \frac{1}{4(0,1)^2 n} \le 0,01 \quad \Leftrightarrow \quad n \ge \frac{1}{4(0,1)^2 0,01} = 2500.$$

Esercizio 2. Il valore atteso di X_j è dato da

$$E(X_j) = \sum_{k=1}^{4} x_k P(X_j = x_k) = 0,15 + 0,25 + 2p.$$

a) Imponendo $E(X_j) = 1$ si trova $p = 0,3$
b) Calcoliamo innanzitutto

$$E(X_j^2) = \sum_{k=1}^{4} x_k^2 P(X_j = x_k) = 0,075 + 0,25 + 1,2 = 1,525$$

Quindi la varianza di X_j è data da $Var(X_j) = E(X_j^2) - 1 = 0,525 = \frac{21}{40}$.
c) La disuguaglianza di Chebyshev per Y diventa

$$P\left(\left\{|Y - 1\ | \le \frac{1}{5}\sqrt{\frac{21}{40}}\right\}\right) \ge 1 - \frac{1}{100}\frac{\frac{21}{40}}{\frac{1}{25}\frac{21}{40}} = 1 - \frac{1}{4} = \frac{3}{4}$$

d) Dal Teorema Centrale del Limite si deduce che

$$P\left(\{|Y_n - \mu\ | \le \epsilon\}\right) \simeq 2\Phi\left(\sqrt{\frac{n}{\sigma^2}}\epsilon\right) - 1$$

dove Y_n è la media aritmetica di n variabili aleatorie indipendenti e identicamente distribuite con valore atteso μ e varianza σ^2. Quindi nel nostro caso

$$P\left(\left\{|Y-1\,|\leq \frac{1}{5}\sqrt{\frac{21}{40}}\right\}\right)\simeq 2\Phi\left(\frac{10}{\sqrt{\frac{21}{40}}}\frac{1}{5}\sqrt{\frac{21}{40}}\right)-1=2\Phi(2)-1\simeq 0,9544.$$

Esercizio 3. La disuguaglianza di Chebyshev per la media aritmetica $\frac{S_n}{n}$ di n variabili aleatorie indipendenti e identicamente distribuite con valore atteso μ e varianza σ^2, dice che

$$P\left(\left\{\left|\frac{S_n}{n}-\mu\right|>\epsilon\right\}\right)\leq \frac{1}{n}\frac{\sigma^2}{\epsilon^2}$$

Nel nostro caso si ha

$$P\left(\left\{\left|\frac{S_n}{n}-\frac{3}{2}\right|>\frac{1}{10}\right\}\right)\leq \frac{1}{n}\frac{3}{4}100$$

e quindi n deve essere tale che

$$\frac{1}{n}\frac{3}{4}100\leq 10^{-3}\quad\Leftrightarrow\quad n\geq 75000.$$

Grazie al Teorema Centrale del Limite si ha

$$\lim_{n\to\infty}P\left(\left\{\frac{2S_n-3n}{\sqrt{n}}>0\right\}\right)=\lim_{n\to\infty}P\left(\left\{\frac{S_n-\frac{3}{2}n}{\sqrt{n\frac{3}{4}}}>0\right\}\right)=$$

$$=\lim_{n\to\infty}\left(1-P\left(\left\{\frac{S_n-\frac{3}{2}n}{\sqrt{n\frac{3}{4}}}\leq 0\right\}\right)\right)=1-\Phi(0)=\frac{1}{2}$$

Esercizio 4. Siano $X_1,\ldots,X_n$ variabili aleatorie indipendenti e identicamente distribuite con valore atteso $\mu=1$ e deviazione standard $\sigma=1/10$, e sia $S_n=\sum_{i=1}^n X_i$. La variabile aleatoria X_i misura il peso della i-esima auto mentre S_n misura il peso complessivo di n auto. Vogliamo determinare n tale che

$$P\left(\{S_n>200\}\right)\geq\frac{1}{10}$$

Dal Teorema Centrale del Limite si ha

$$P\left(\{S_n>200\}\right)=P\left(\left\{\frac{S_n-n}{\sqrt{\frac{n}{100}}}>\frac{200-n}{\sqrt{\frac{n}{100}}}\right\}\right)\simeq 1-\Phi\left(\frac{200-n}{\sqrt{\frac{n}{100}}}\right)\geq\frac{1}{10}$$

Quest'ultima disuguaglianza equivale a

$$\Phi\left(\frac{200-n}{\sqrt{\frac{n}{100}}}\right) \le \frac{1}{10} \quad \Leftrightarrow \quad \frac{200-n}{\sqrt{\frac{n}{100}}} \le 1,2816 \quad \Leftrightarrow \quad n+0,12816\sqrt{n}-200 \ge 0$$

Per maggiore chiarezza, conviene riscrivere quest'ultima disequazione sostituendo $\sqrt{n}$ con t. Se indichiamo con t_1 e t_2 le soluzioni dell'equazione $t^2 + 0,12816\,t - 200 = 0$, allora la disequazione $t^2 + 0,12816\,t - 200 \ge 0$ ha come soluzione $t \le t_1$ e $t \ge t_2$. Approssimativamente si ha $t_1 \simeq -\sqrt{200}$ e $t_2 \simeq \sqrt{200}$ (ma si osservi che t_2 è minore di $\sqrt{200}$), inoltre a noi interessano solo le soluzioni positive, ovvero $t \ge t_2 \simeq \sqrt{200}$. Quindi n deve essere tale che $\sqrt{n} \ge \sqrt{200}$, ovvero $n \ge 200$.

Esercizio 5. Siano $X_1, \ldots, X_{30}$ variabili aleatorie indipendenti e identicamente distribuite con valore atteso $\mu = 3000$ e deviazione standard $\sigma = 500$, e sia $S_{30} = \sum_{i=1}^{30} X_i$. La variabile aleatoria X_i conta il numero di viaggiatori nella i-esima giornata mentre S_{30} misura il numero complessivo di viaggiatori in 30 giorni. Vogliamo calcolare

$$P(\{S_{30} > 100\,000\}) = P\left(\left\{\frac{S_{30} - 90\,000}{\sqrt{30}\;500} > \frac{100\,000 - 90\,000}{\sqrt{30}\;500}\right\}\right) \simeq$$
$$\simeq 1 - \Phi\left(\frac{10\,000}{\sqrt{30}\;500}\right) = 1 - \Phi\left(\frac{20}{\sqrt{30}}\right) \simeq 0,0001$$

Esercizio 6. Consideriamo la ripetizione (infinite volte) di una prova di Bernoulli di parametro p e consideriamo la variabile aleatoria X che conta il numero di tentativi necessari per avere il primo successo. La distribuzione di probabilità di X è data da $P(X = k) = p\,(1-p)^{k-1}$ per k intero maggiore o uguale ad 1. Nel nostro caso la prova di Bernoulli consiste nel lancio di un dado e la prova ha successo quando esce un sei. Dall'ipotesi che il dado sia equilibrato si deduce $p = 1/6$.

a) $P(X = 4) = \frac{1}{6}\left(\frac{5}{6}\right)^3$

b) $P(X \le 3) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{1}{6} + \frac{1}{6}\frac{5}{6} + \frac{1}{6}\left(\frac{5}{6}\right)^2$

Alternativamente $P(X \le 3) = 1 - P(X > 3) = 1 - \left(\frac{5}{6}\right)^3$

Nota bene: le due soluzioni sono uguali grazie al fatto che

$$\frac{1}{6} + \frac{1}{6}\frac{5}{6} + \frac{1}{6}\left(\frac{5}{6}\right)^2 = \frac{1}{6}\left(1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2\right) = \frac{1}{6}\left(\frac{1-\left(\frac{5}{6}\right)^3}{1-\frac{5}{6}}\right)$$

c) La probabilità che il sei esca dopo il sesto lancio equivale alla probabilità che il sei non esca nei primi sei lanci. Quindi $P(X > 6) = \left(\frac{5}{6}\right)^6$. In alternativa si può calcolare questa probabilità come segue

$$P(X > 6) = 1 - P(X \leq 6) = 1 - \sum_{k=1}^{6} P(X = k) = 1 - \sum_{k=1}^{6} \frac{1}{6} \left(\frac{5}{6}\right)^{k-1}$$

$$= 1 - \frac{1}{6} \sum_{k=1}^{6} \left(\frac{5}{6}\right)^{k-1} = 1 - \frac{1}{6} \frac{1 - \left(\frac{5}{6}\right)^6}{1 - \frac{5}{6}} = 1 - \left(1 - \left(\frac{5}{6}\right)^6\right) = \left(\frac{5}{6}\right)^6$$

Esercizio 7. Si tratta di 10 prove indipendenti, ciascuna delle quali può avere tre esiti: vittoria di Massimo con probabilità 1/3, vittoria di Paolo con probabilità 1/3, pareggio con probabilità 1/3.

a) $A = \{X_1 = 4\} \cap \{X_2 = 3\} = \{X_1 = 4\} \cap \{X_2 = 3\} \cap \{X_3 = 3\}$. Osserviamo che la probabilità di una determinata sequenza di 4 vittorie di Massimo, 3 di Paolo e 3 pareggi è data da $\left(\frac{1}{3}\right)^4 \left(\frac{1}{3}\right)^3 \left(\frac{1}{3}\right)^3 = \left(\frac{1}{3}\right)^{10}$ e che il numero di tali sequenze è pari a $\frac{10!}{4!\,3!\,3!}$. Quindi $P(A) = \frac{10!}{4!\,3!\,3!} \left(\frac{1}{3}\right)^{10}$

b) Si osservi che X_1 è una variabile aleatoria bernoulliana di parametri $p = 1/3$ e $n = 10$. Pertanto

$$P(X_1 = k) = \binom{10}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{10-k}$$

Calcoliamo infine la probabilità condizionata

$$P(X_2 = 3, X_3 = 3 | X_1 = 4) = \frac{P(X_2 = 3, X_3 = 3, X_1 = 4)}{P(X_1 = 4)} = \frac{\frac{10!}{4!\,3!\,3!} \left(\frac{1}{3}\right)^{10}}{\frac{10!}{4!\,6!} \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^6} = \frac{6!}{3!\,3!\,2^6}$$

c) Osserviamo che $X_1 + X_2 + X_3 = 10$ e quindi $Y = X_2 + X_3 = 10 - X_1$. Dato che $X_1 \sim B(10, 1/3)$ allora

$$P(Y = k) = \binom{10}{10-k} \left(\frac{1}{3}\right)^{10-k} \left(\frac{2}{3}\right)^k$$

d) $Cov(X_1, Y) = Cov(X_1, 10 - X_1) = -Cov(X_1, X_1) = -Var(X_1) = -\frac{20}{9}$ poiché la varianza di una variabile aleatoria bernoulliana $B(n, p)$ è data da $np(1 - p)$.