Prova d'esame

Calcolo delle probabilità e Statistica (12 crediti)

8 Gennaio 2013

Esercizio 1. Si definiscono due variabili aleatorie X e Y nel seguente modo. Si effettuano ripetuti lanci (indipendenti) di una moneta che dà testa con probabilità p e croce con probabilità $q = 1 - p$. Si termina nel momento in cui sono usciti tutti e due gli esiti: poniamo X il numero di teste uscite e Y il numero di croci uscite. Determinare la legge di X e di Y e la loro media. Sono indipendenti?

Soluzione. Indichiamo con E_k l'esito del lancio k della moneta. La famiglia $\{E_k\}$ è formata da variabili aleatorie indipendenti e equidistribuite, con legge $B(p)$. Consideriamo la distribuzione di X: si ottiene

$$\mathbb{P}(X = 1) = \mathbb{P}(X = 1 \mid E_1 = T)\mathbb{P}(E_1 = T) + \mathbb{P}(X = 1 \mid E_1 = C)\mathbb{P}(E_1 = C)$$

Si nota che se al primo lancio esce C, allora per forza $X = 1$; invece, se al primo lancio esce T, per avere $X = 1$ deve essere $E_2 = C$ (esce solo una T), quindi

$$\mathbb{P}(X = 1) = \mathbb{P}(E_2 = C)\mathbb{P}(E_1 = T) + \mathbb{P}(E_1 = C) = q(1 + p)$$

Per avere $X \geq 2$ deve essere necessariamente che il primo lancio sia T, quindi

$$\begin{aligned}\mathbb{P}(X = k) &= Pr(X = k \mid E_1 = T)\mathbb{P}(E_1 = T) = \mathbb{P}(\text{"}k \text{ teste e una croce consecutive"} \mid E_1 = T)\mathbb{P}(E_1 = T)\\ &= \mathbb{P}(\text{"}k-1 \text{ teste e una croce consecutive"})\mathbb{P}(E_1 = T) = p^k q\end{aligned}$$

Analogamente, risulta $\mathbb{P}(Y = 1) = p(1+q)$, $\mathbb{P}(Y = k) = q^k p$. Si osserva che X e Y non sono indipendenti, essendo, ad esempio $\mathbb{P}(X = 2, Y = 2) = 0 \neq \mathbb{P}(X = 2)\mathbb{P}(Y = 2)$.

Determiniamo infine la media di X e Y. Risulta

$$\mathbb{E}[X] = q(1 + p) + \sum_{k=2}^{\infty} kp^k q = q + \sum_{k=1}^{\infty} kp^k q = q + pq \sum_{k=1}^{\infty} kp^{k-1} = q + pq \frac{1}{(1-p)^2} = q + \frac{p}{q};$$

per simmetria, risulta $\mathbb{E}[Y] = p + \frac{q}{p}$.

Esercizio 2. Un rilevatore di particelle ha sensibilità pari a $p = 0.3$, ossia rileva una particella che lo attraversa con probabilità p. Sia X il numero di particelle che attraversano il rilevatore nel periodo di tempo in cui vengono contate 100 particelle.

(a) Trovare la legge di X e la sua media.

(b) Usando una opportuna approssimazione trovare la probabilità degli eventi $\{X \leq 280\}$, $\{X \leq 320\}$, $\{280 < X \leq 320\}$.

Soluzione. Sia X_1 il numero di particelle che attraversano il rilevatore per avere la prima rilevazione; allora X_1 ha legge di tipo geometrico,

$$\mathbb{P}(X_1 = j) = q^{j-1}p$$

essendo $q = 1 - p$. Indichiamo inoltre con $X_2, \ldots, X_{100}$ le particelle che attraversano il rilevatore dopo il primo per ottenere il secondo successo, dopo il secondo per ottenere il terzo, etc... La famiglia $\{X_k\}$ sono variabili aleatorie indipendenti e equidistribuite; in particolare, si ha $\mathbb{E}[X_1] = 1/p = 10/3$ e $\mathrm{Var}(X) = q/p^2 = 70/9$.

Si ha $X = \sum_{k=1}^{100} X_k$. In particolare, la media di X vale $333.\bar{3}$ e la sua varianza vale $7000/9$.

Posso ottenere direttamente la distribuzione di X osservando che $X = n$ vuol dire avere 99 successi su $n - 1$ tentativi e un successo al tentativo n, quindi per ogni $n \geq 100$ si ottiene

$$\mathbb{P}(X = n) = \binom{n-1}{99} p^{100}(1-p)^{n-100}$$

Usiamo l'approssimazione normale data dal teorema centrale; da

$$\frac{X - \mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}} \sim Z$$

si ottiene $\mathbb{P}(X \leq 270) \simeq \mathbb{P}(Z < -2.27) = 0.011$, $\mathbb{P}(X \leq 330) \simeq \mathbb{P}(Z < -0.12) = 0.452$, $\mathbb{P}(270 < X < 330) = 0.441$

Esercizio 3. Una successione di variabili aleatorie indipendenti ed equidistribuite $\{X_j,\ j = 1, \dots, n\}$ ha legge di Poisson di parametro λ.

(a) Determinare la media e la varianza di $S_n = X_1 + \dots X_n$;

(b) Sia $\bar{X} = \frac{1}{n} S_n$ la media aritmetica delle $\{X_j\}$. Trovate la distribuzione di $\bar{X}$, la sua media e la sua varianza. Dimostrate che $\bar{X}$ converge in probabilità e determinate la legge limite.

Soluzione. La distribuzione di S_n è data dalla somma di n distribuzioni di Poisson indipendenti, quindi ha legge di Poisson di parametro $n\lambda$; in particolare, $\mathbb{E}[S_n] = n\lambda$, $\mathrm{Var}(S_n) = n\lambda$.

Consideriamo ora $\bar{X}$. Dato che S_n ha valori interi, $\bar{X}$ ha valori nelle frazioni $\frac{k}{n}$ per $k \geq 0$; si ha

$$\mathbb{P}(\bar{X} = \frac{k}{n}) = \mathbb{P}(S_n = k) = e^{-n\lambda} \frac{(n\lambda)^k}{k!}.$$

Media e varianza di $\bar{X}$ valgono, rispettivamente, λ e $\frac{\lambda}{n}$. Posso applicare la legge dei grandi numeri per discutere la convergenza della successione $\bar{X}_n$ per $n \to \infty$; infatti le variabili X_k sono indipendenti, equidistribuite e hanno varianza finita. Risulta allora $\bar{X}_n \to \lambda$ per $n \to \infty$, ossia $\bar{X}$ approssima la variabile aleatoria degenere che vale λ con probabilità 1.

Esercizio 4. Si consideri la catena di Markov $\{X_n\}$ su $E = \{1, \dots, 5\}$ avente matrice di transizione

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

(a) Determinare l'orbita e il periodo dello stato 2. Corrisponde al periodo della catena? Perché?

(b) Determinare, se esiste, la distribuzione invariante del sistema. Quanto vale il limite $\lim_{n\to\infty} p_{2,j}^{(n)}$?

Soluzione. L'orbita dello stato 2 risulta $\{5, 6, 7, \dots\}$ e contiene tutti gli interi ≥ 5; in particolare il periodo vale 1. Inoltre, la catena è irriducibile quindi tutti gli stati hanno lo stesso periodo.

La distribuzione invariante esiste unica essendo il sistema irriducibile; inoltre, il sistema è regolare, quindi la distribuzione invariante è ergodica. Si ottiene

$$\pi = \left(\frac{1}{9}, \frac{1}{9}, \frac{4}{9}, \frac{2}{9}, \frac{1}{9} \right)$$

quindi $\lim_{n\to\infty} p_{2,j}^{(n)} = \pi(j)$ per $j \in E$.

Esercizio 5. Su un campione di 50 ragazzi, il 25% ha dichiarato di interessarsi di astronomia. Determinare un intervallo di confidenza al livello 95% per la percentuale di ragazzi che si interessano di astronomia. Determinare inoltre quante persone si dovrebbero intervistare perché questo intervallo di confidenza abbia ampiezza inferiore a 2%.

Soluzione. L'intervallo di confidenza sarà della forma

$$\hat{p} \in \left(\bar{p} \pm \phi_{0.975} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \right)$$

da cui si ottiene

$$\hat{p} \in \left(0.25 \pm 1.96 \frac{\sqrt{3}/4}{\sqrt{50}} \right) = (0.2 \pm 0.12) = \big(0.13,\ 0.37\big)$$

Perché l'intervallo di confidenza abbia ampiezza minore di 0.02, supponendo valida la proporzione, deve essere

$$1.96 \frac{\sqrt{3}/4}{\sqrt{n}} = 0.01$$

da cui si ottiene $n = (1.96 * \sqrt{3/16} * 100)^2 = 7203$, ossia si devono intervistare 7203 persone.