Soluzioni

Calcolo delle probabilità e Statistica (12 crediti)

3 luglio 2007

---

1. Probabilità almeno 3 periodi di magra

$$\Pr(X \geq 3) = 1 - \Pr(X=0) - \Pr(X=1) - \Pr(X=2) = 1 - e^{-\lambda}(1 + \lambda + \frac{1}{2}\lambda^2) = 1 - 5e^{-2} \simeq 0,32$$

Sia Y il numero di periodi di secca; allora condizionato rispetto a $X = x$, Y ha distribuzione binomiale di parametri x e $p = \frac{1}{2}$; per trovare la distribuzione di Y procediamo in questo modo.

$$\Pr(Y=0) = \sum_{x=0}^{\infty} \Pr(Y=0, X=x) = \sum_{x=0}^{\infty} \Pr(Y=0 \mid X=x) \Pr(X=x) = \sum_{x=0}^{\infty} (\frac{1}{2})^x e^{-\lambda} \frac{1}{x!} \lambda^x$$
$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{1}{x!} (\frac{\lambda}{2})^x = e^{-\lambda} e^{\lambda/2} = e^{-\lambda/2}$$

$$\Pr(Y=1) = \sum_{x=0}^{\infty} \Pr(Y=1, X=x) = \sum_{x=0}^{\infty} \Pr(Y=1 \mid X=x) \Pr(X=x) = \sum_{x=1}^{\infty} x (\frac{1}{2})^x e^{-\lambda} \frac{1}{x!} \lambda^x$$
$$= e^{-\lambda} (\frac{\lambda}{2}) \sum_{x=1}^{\infty} \frac{1}{(x-1)!} (\frac{\lambda}{2})^{x-1} = (\frac{\lambda}{2}) e^{-\lambda} e^{\lambda/2} = (\frac{\lambda}{2}) e^{-\lambda/2}$$

$$\Pr(Y=2) = \sum_{x=0}^{\infty} \Pr(Y=2, X=x) = \sum_{x=2}^{\infty} \Pr(Y=2 \mid X=x) \Pr(X=x) = \sum_{x=2}^{\infty} \binom{x}{2} (\frac{1}{2})^x e^{-\lambda} \frac{1}{x!} \lambda^x$$
$$= e^{-\lambda} \frac{1}{2!} (\frac{\lambda}{2})^2 \sum_{x=1}^{\infty} \frac{1}{(x-2)!} (\frac{\lambda}{2})^{x-2} = \frac{1}{2!} (\frac{\lambda}{2})^2 e^{-\lambda} e^{\lambda/2} = \frac{1}{2!} (\frac{\lambda}{2})^2 e^{-\lambda/2}$$

In generale si dimostra che Y ha distribuzione di Poisson di parametro $(\lambda/2)$, quindi il numero medio di periodi di secca ogni anno è 1.

2. Sia Y la lunghezza di AP: allora Y ha distribuzione uniforme in $(0, \ell)$. Il triangolo APQ è ancora rettangolo isoscele avente ipotenusa $AP = Y$ e cateti $AQ = PQ = Y/\sqrt{2}$ e la sua area è pari a $X = \frac{1}{4}Y^2$; la distribuzione di X è allora compresa tra 0 e $\frac{1}{4}\ell^2$ ed ha funzione di ripartizione, per $0 < t < \ell^2/4$

$$\Pr(X \leq t) = \Pr(Y^2 < 4t) = \Pr(0 < Y < 2\sqrt{t}) = \frac{2}{\ell}\sqrt{t}.$$

La sua media si può calcolare direttamente dalla formula

$$\mathbb{E}[X] = \mathbb{E}[\frac{1}{4}Y^2] = \int_0^{\ell} \frac{1}{4} y^2 \, \frac{\mathrm{d}y}{\ell} = \frac{1}{12}\ell^2.$$

3. La catena è irriducibile, a stati finiti e aperiodica (ad esempio, l'orbita di A contiene 3 e 4, quindi il suo MCD è 1), quindi è ergodica ed ammette un'unica misura invariante. Scriviamo la matrice di transizione

$$P = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

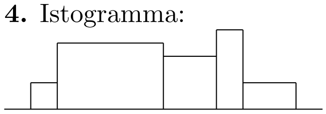
Allora $\Pr(X_1 = A \mid \pi_0) = \frac{1}{4}$, la probabilità di arrivare in D al primo passo è $\frac{1}{4}(\frac{1}{3} + \frac{1}{2} + 1) = \frac{11}{24}$, la probabilità di arrivare in D al secondo passo è $\frac{1}{4}(\frac{1}{2} + \frac{1}{2} + \frac{1}{3}) = \frac{1}{3}$ e quella di arrivare a D al terzo passo è $\frac{1}{4}(\frac{5}{18} + \frac{1}{6} + \frac{1}{3} + \frac{1}{2}) = \frac{23}{72} \simeq 0,319$. La distribuzione invariante della catena si ottiene del sistema

$$\begin{cases} d = a \\ \frac{1}{3}a = b \\ \frac{1}{3}a + \frac{1}{2}b = c \\ a + b + c + d = 1 \end{cases}$$

da cui

$$\pi = (\frac{6}{17}, \frac{2}{17}, \frac{3}{17}, \frac{6}{17}).$$

4. Istogramma:



La media aritmetica e la deviazione standard si calcolano prendendo per ogni classe il valore centrale

$$\begin{aligned} \mu &= \frac{1}{n} \sum n_i \bar{x}_i = \frac{1}{200}(10 * 0,5 + 100 * 3 + 40 * 6 + 30 * 7,5 + 20 * 9) = 4,75 \\ s^2 &= \frac{1}{n-1} \sum n_i (\bar{x}_i - \mu)^2 \\ &= \frac{1}{199}(10 * (-4,25)^2 + 100 * (-1,75)^2 + 40 * (1,25)^2 + 30 * (2,75)^2 + 20 * (4,25)^2) \simeq 5,716 \\ s &\simeq 2,39 \end{aligned}$$

Indicando con x_i i dati ordinati, essendo $n/2 = 100$, $n/4 = 50$ e $3 * n/4 = 150$, la mediana e i quartili risultano essere

$$m = \frac{x_{100} + x_{101}}{2}, \qquad q_1 = \frac{x_{50} + x_{51}}{2}, \qquad q_3 = \frac{x_{150} + x_{151}}{2}$$

I valori x_{50}, x_{51}, x_{100} e x_{101} cadono nella seconda classe, quindi in prima approssimazione possiamo prendere $q_1 = m = 3$; allo stesso modo, essendo x_{150} nella terza classe e x_{151} nella quarta classe, avremo $q_3 = \frac{6+7,5}{2} = 6,75$.

Con una migliore precisione osserviamo che il primo quartile si trova passato il 40% della seconda fascia

$$q_1 = 1 + 4 * \frac{40,5}{100} = 2,62$$

la mediana passato il 90% della seconda fascia

$$m = 1 + 4 * \frac{90,5}{100} = 4,62$$

ed infine il terzo quartile si trova a cavallo tra la terza e la quarta fascia, quindi $q_3 = 7$; la distanza interquartile è

$$IQR = q_3 - q_1 = 7 - 1,42 = 5,58$$

5. Per ragioni di simmetria vediamo che NA e NB sono equidistribuite, ma non sono indipendenti, in quanto $NA + NB = 4$. Osserviamo che la distribuzione congiunta si ottiene dall'osservazione

$$\Pr(NA = x, NB = y) = \begin{cases} 0, & x + y \neq 4 \\ \Pr(NA = x), & x + y = 4 \end{cases}$$

quindi conoscere la marginale mi dà anche l'informazione sulla congiunta. Si ottiene (scriviamo $40 = 4n$ con $n = 10$ in modo da fare subito anche il caso generale)

$$\Pr(NA = 0) = \frac{\binom{4}{0}\binom{4n-4}{2n}}{\binom{4n}{2n}}$$

perché nei casi favorevoli devo scegliere 0 assi su 4 e $20 = 2n$ carte tra le carte del mazzo che non sono assi, mentre i casi totali sono il numero di modi in cui distribuisco la metà delle carte, quindi

$$\Pr(NA = 0) = \frac{(4n-4)!(2n)!}{(2n-4)!(4n)!} = \frac{2n(2n-1)(2n-2)(2n-3)}{4n(4n-1)(4n-2)(4n-3)} = \frac{1}{2}\frac{(n-1)(2n-3)}{(4n-1)(4n-3)}$$

sostituendo il valore $n = 10$ si ottiene

$$\Pr(NA = 0) \simeq 0,053$$

e questo stesso vale per $\Pr(NA = 4) = \Pr(NB = 0) = \Pr(NA = 0)$; calcoliamo anche

$$\Pr(NA = 1) = \Pr(NA = 3) = \frac{\binom{4}{1}\binom{4n-4}{2n-1}}{\binom{4n}{2n}} = \frac{4n(n-1)}{(4n-1)(4n-3)}$$

da cui $\Pr(NA = 1) = \Pr(NA = 3) \simeq 0,249$; infine

$$\Pr(NA = 2) = \frac{\binom{4}{2}\binom{4n-4}{2n-2}}{\binom{4n}{2n}} = \frac{3n(2n-1)}{(4n-1)(4n-3)}$$

da cui $\Pr(NA = 2) \simeq 0,395$.

La media di NA è

$$\begin{aligned}\mathbb{E}[NA] &= \Pr(NA = 1) + 2\Pr(NA = 2) + 3\Pr(NA = 3) + 4\Pr(NA = 4) \\ &= 2\frac{(n-1)(2n-3)}{(4n-1)(4n-3)} + 2\frac{3n(2n-1)}{(4n-1)(4n-3)} + 4\frac{4n(n-1)}{(4n-1)(4n-3)} = 2.\end{aligned}$$

Quando il numero delle carte n tende all'infinito, restando fermi i 4 assi, allora NA tende alla variabile aleatoria con la seguente distribuzione

x	0	1	2	3	4
$p(x)$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{6}{16}$	$\frac{1}{4}$	$\frac{1}{16}$