Seconda prova in itinere di

Calcolo delle probabilità e Statistica (12 crediti)

10 gennaio 2007

Esercizio 1.

Un concessionario automobilistico ha in dotazione 4 autoveicoli nuovi e 2 usati, che divide equamente tra due vetrine, Γ e Ω. Per massimizzare l'impatto sulla clientela decide, ogni giorno, di scegliere a caso un'automobile da ogni vetrina e scambiarla.

a. Si può usare una *catena di Markov* per modelizzare lo stato delle auto esposte in ciascuna vetrina? Nel caso determinare la corrispondente matrice di transione e discuterne la natura classificandone gli stati.

b. Qual è la probabilità che, trascorso un anno e supponendo che non sia stata effettuata alcuna vendita, la vetrina Γ abbia in esposizione 3 veicoli nuovi?

Soluzione. Indichiamo con X_n il numero di macchine usate presenti nella prima vetrina il giorno n; questo numero specifica lo stato delle vetrine nel senso che, dato X_n, possiamo dire quante macchine di ogni tipo sono presenti in ciascuna vetrina. X_n assume valori in $E = \{0, 1, 2\}$ e inoltre la transizione da X_n a X_{n+1} dipende solo dallo stato X_n e non dalla storia precedente. Si tratta quindi di una catena di Markov. Osserviamo che anche Y_n il numero di macchine nuove presenti nella prima vetrina può essere usato; lo spazio degli stati diventa allora $\{1, 2, 3\}$.

Supponiamo $X_n = 2$; allora $X_{n+1} = 2$ se e solo se viene scelta per lo scambio l'auto nuova, quindi con probabilità $\frac{1}{3}$; altrimenti si ha $\Pr(X_{n+1} = 1 \mid X_n = 2) = \frac{2}{3}$.

Supponiamo $X_n = 0$: allora X_{n+1} rimane costante con probabilità $\frac{1}{3}$, aumenta di una unità con probabilità $\frac{2}{3}$.

Rimane il caso $X_n = 1$: al tempo $n + 1$, il numero di auto usate diminuisce se dalla prima vetrina scelgo quella usata e dalla seconda ne scelgo una nuova, quindi con probabilità $\frac{2}{9} = \frac{1}{3} * \frac{2}{3}$; aumenta con la stessa probabilità e rimane costante con probabilità $\frac{5}{9} = 1 - 2 * \frac{2}{9}$. Si ottiene quindi la seguente matrice di transizione per il processo

$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{2}{9} & \frac{5}{9} & \frac{2}{9} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Il processo di Markov $\{X_n\}$ è ergodico, quindi la distribuzione limite non dipende dalla distribuzione iniziale. Approssimativamente, per grandi tempi, la distribuzione coincide con quella limite, quindi per determinare $\Pr(X_n = 0)$, $n = 365$, calcoliamo la misura invariante π e approssimiamo il valore cercato con $\pi(0)$.

La misura invariante verifica $\pi T = \pi$, e si ottiene $\pi = (\frac{1}{5}, \frac{3}{5}, \frac{1}{5})$, quindi la probabilità cercata è approssimativamente $\frac{1}{5}$.

Esercizio 2.

La produzione di una certa lega prevede il 22.2% di rame. Un campione di 10 analisi del prodotto ha mostrato un contenuto medio di rame del 23.5% ed una deviazione standard campionaria dello 1.24%. Costruire un test di ipotesi bilatero, al livello di significatività dello 0.05, per verificare l'ipotesi che la lega ha mantenuto le proporzioni richieste. (Suggerimento: ricordatevi di scrivere tutti i valori in termini assoluti e non percentuali).

Soluzione. Posto $\mu = 0.222$, $n = 10$, $\bar{x} = 0.235$ e $s^2 = 0.0124$, si tratta di verificare se $\bar{x}$ appartiene all'intervallo $\mu \pm t_{0,975}(n-1)\frac{s}{\sqrt{n}}$:

$$0.235 \in (0.222 - 0.0124 * 2.2622/3.1623,\ 0.222 + 0.0124 * 2.2622/3.1623) = (0.213,\ 0.231)$$

quindi devo rifiutare l'ipotesi nulla.

Esercizio 3.

Sia $\{X_n\}$ una catena di Markov, omogenea nel tempo e nello spazio, su uno spazio finito degli stati E.

a. Definire il periodo di uno stato $a \in E$

b. Mostrare che, se la catena è irriducibile, tutti gli stati hanno lo stesso periodo.

c. Costruire un esempio di catena avente periodo 2

Soluzione. Indichiamo con $O(a)$ l'orbita di a, ossia l'insieme dei tempi n per cui $\Pr(X_n = a \mid X_0 = a) = p^{(n)}(a,a) > 0$. Si definisce periodo dello stato a il M.C.D. dell'orbita $O(a)$.

Supponiamo che a e b siano comunicanti tra loro; dimostriamo che hanno lo stesso periodo. Dato che in una catena irriducibile tutti gli stati comunicano tra loro, questo dimostra l'affermazione.

Sia d il periodo di a e $q \neq d$ il periodo di b; questo significa che $p^{(nd)}(a,a) > 0$ per ogni $n \in \mathbb{N}$. Sia α il minimo numero di passi necessari per passare da $a \to b$ e sia β il minimo numero di passi per andare da $b \to a$; allora $\alpha + \beta$ è multiplo di d, in quanto appartiene all'orbita di a. D'altronde, anche $\alpha + nq + \beta$ appartiene all'orbita di a per ogni $n \in \mathbb{N}$, quindi d deve dividere nq per ogni $n \in \mathbb{N}$, e quindi d deve dividere q. Lo stesso ragionamento implica però che q deve dividere p, e quindi necessariamente $d = q$.

La più semplice catena di periodo 2 è espressa dalla matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Esercizio 4.

In un esperimento aleatorio relativo al lancio di una moneta si ottiene il seguente risultato (T, C, C, C). A priori ci attendevamo che la moneta non fosse truccata e che la probabilità di test fosse approssimativamente 0.5. Detta X la variabile relativa al numero delle teste e p la probabilità che esca testa in un singolo lancio:

a. scrivere la distribuzione a priori di X dato p;

b. scegliamo come distribuzione a priori per p una distribuzione del tipo Beta(3,3): $h(p) = Cp^2(1-p)^2\mathbf{1}_{(0,1)}(p)$. Verificare che la media a priori è pari a 0.5. Determinare la distribuzione a posteriori $h(p \mid (T, C, C, C))$. Determinare lo stimatore bayesiano del parametro p date le osservazioni.

Ricordiamo che
$$\int_0^1 x^{n-1}(1-x)^{m-1}\,\mathrm{d}x = \frac{(n-1)!\,(m-1)!}{(n+m-1)!}.$$

Soluzione. X ha distribuzione binomiale di parametri $n = 4$ e p: $\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Osserviamo che la media a priori per h è pari al rapporto

$$\frac{\int_0^1 x^3(1-x)^2 \,\mathrm{d}x}{\int_0^1 x^2(1-x)^2 \,\mathrm{d}x} = \frac{3!\,2!/6!}{2!\,2!/5!} = \frac{1}{2}.$$

Poniamo $T = 1$, $C = 0$. Indichiamo con $f(x \mid p) = p^x(1-p)^{1-x}$ la distribuzione di un lancio; si ha $f((T, C, C, C) \mid p) = p(1-p)^3$, da cui

$$h(p \mid (T, C, C, C)) = kp^3(1-p)^5, \qquad k = \frac{9!}{3!\,5!} = 504;$$

lo stimatore bayesiano di p è pari a

$$\int_0^1 kp^4(1-p)^5 \,\mathrm{d}x = \frac{9!}{3!\,5!} \frac{4!\,5!}{10!} = \frac{4}{10} = 0,4$$

Esercizio 5.

Una catena di Markov X sullo spazio degli stati $E = \{0, 1, 2, 3, 4\}$ è definita dalla seguente matrice di transizione

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/5 & 2/5 & 1/5 & 1/5 & 0 \\ 0 & 1/5 & 2/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

a. Determinare le classi irriducibili degli stati. Determinare quali stati sono transitori, ricorrenti oppure assorbenti.

b. Determinare, se esistono, la/le misure invarianti per il sistema. Quante sono?

c. Sia $\pi_0 = (0, 0, 1/2, 1/2, 0)$ la distribuzione iniziale. Qual è la probabilità che il sistema si trovi nello stato 0 al passo 1? al passo 2? Determinare la probabilità che il sistema raggiunga, prima o poi, lo stato 0.

Soluzione. Vi sono 3 classi irriducibili: $\{0\}$ e $\{4\}$ formano due classi assorbenti, $\{1, 2, 3\}$ formano una classe di stati transitori. Esistono infinite misure invarianti del tipo $(a, 0, 0, 0, 1-a)$ per ogni $a \in [0, 1]$.

Abbiamo $\Pr(X_1 = 0 \mid \pi_0) = \frac{1}{2} * 0 + \frac{1}{2} * \frac{1}{5} = \frac{1}{10}$ e $\Pr(X_2 = 0 \mid \pi_0) = 4 * \frac{1}{5} * \frac{1}{5} + \frac{1}{5} = \frac{9}{25}$. Indichiamo con P_1, P_2, P_3 la probabilità di finire in 0 partendo dagli stati 1, 2, 3 rispettivamente. Scriviamo il sistema

$$\begin{cases} P_1 = \frac{1}{5} + \frac{2}{5}P_1 + \frac{1}{5}P_2 + \frac{1}{5}P_3 \\ P_2 = \qquad \frac{1}{5}P_1 + \frac{2}{5}P_2 + \frac{1}{5}P_3 \\ P_3 = \frac{1}{5} + \frac{1}{5}P_1 + \frac{1}{5}P_2 + \frac{1}{5}P_3 \end{cases}$$

che ha soluzione

$$P_1 = \frac{5}{8}, \qquad P_2 = \frac{3}{8}, \qquad P_3 = \frac{1}{2}$$

da cui la probabilità di essere assorbito in 0 è $\frac{1}{2} * \frac{3}{8} + \frac{1}{2} * \frac{1}{2} = \frac{7}{16}$.

Esercizio 6.
Un carattere è distribuito in una popolazione secondo la legge

$$f(x \mid \theta) = \begin{cases} \frac{\theta}{2} & x = \pm 1 \\ 1-\theta & x = 0. \end{cases}$$

Si effettua un campionamento di ampiezza $n = 10$, da cui si ottengono 3 valori $+1$, 4 valori -1 e i rimanenti 0. Determinare la stima di massima verosimiglianza di θ.

Soluzione. Indichiamo con $\bar{X} = (1, 1, 1, -1, -1, -1, -1, 0, 0, 0)$ il vettore delle osservazioni. Scriviamo la funzione di massima verosimiglianza

$$L(\theta \mid \bar{X}) = \left(\tfrac{\theta}{2}\right)^7 (1-\theta)^3$$

e cerchiamo il massimo di questa funzione che si ottiene per $\theta = \frac{7}{10}$.

Esercizio 4.
Nella tabella sono riportati gli esiti dei rilevamenti della pressione arteriosa massima in un gruppo di maschi quarantenni. I dati sono espressi in millimetri di mercurio (mm Hg) e arrotondati alle cinquine. Tracciare un istogramma dei dati, utilizzando al più 8 colonne. Determinare (usando al più una calcolatrice non programmabile) mediana, media e varianza campionaria.

valori	95	100	105	110	115	120	125	130	135
freq.ass.	1	1	2	3	5	6	10	15	21
valori	140	145	150	155	160	165	170	175	180
freq.ass.	19	14	13	8	5	4	3	3	2

Soluzione. Si ottengono i valori

$$\sum x_i = 18800, \qquad \sum x_i^2 = 2654100, \ qquadn = 135$$

da cui $m = x_{68} = 140$, $\bar{x} = 139.26$, ed infine

$$s^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum x_i^2 - \bar{x}\right) = 268.85$$

Esercizio 6.
Supponiamo di giocare con un dado onesto a 5 facce. Ad ogni istante lanciamo il dado ed indichiamo con X_n il punteggio massimo realizzato nei primi n lanci.

a. Si può usare una *catena di Markov* per modellizzare lo stato della quantità X_n? Nel caso determinare la corrispondente matrice di transizione, discuterne la natura classificandone gli stati e trovare eventuali misure invarianti e/o ergodiche.

Esercizio 7.
Supponiamo che Chiara e Riccardo giochino con un dado onesto a 5 facce puntando 1 euro a testa nella seguente scommessa:

· Chiara e Riccardo lanciano, a turno, il dado a 5 facce: Chiara vince se ottiene un punteggio strettamente più alto di quello ottenuto da Riccardo, altrimenti vince Riccardo.

Supponendo che i capitali iniziali dei due giocatori siano di 10 euro per Chiara e di 5 euro per Riccardo:

a. Modelizzare l'andamento delle scommesse tra Chiara e Riccardo facendo uso di un' opportuna *catena di Markov*

b. Determinare la probabilità che ha Riccardo di vincere l'intero capitale di Chiara

c. Calcolare la durata media del gioco.

Esercizio 8.

Di un cavo di gomma prodotto industrialmente sono stati esaminati tratti consecutivi lunghi ciascuno 100 m e si sono individuati quanti sono i difetti in ciascuno di essi. Si è ottenuta la seguente distribuzione

n.ro difetti riscontrati	0	1	2	3	4	5
n.ro tratti con tale n.ro di difetti	35	45	40	23	4	3

Supponiamo che il numero di difetti sia distribuito secondo una legge di Poisson di parametro λ. Calcolare lo stimatore di massima verosimiglianza per λ.

Soluzione. Lo stimatore di massima verosimiglianza è pari alla media campionaria $\bar{x} = 1.5$

Esercizio 9.

Il gruppo ciclistico Baldo, composto da 100 corridori, è impegnato in alcuni test di potenza/resistenza. In particolare i suoi atleti sono chiamati a percorrere la tappa Trento-Bondone usando rapporti tali da rimanere nell' intervallo di pedalata compreso tra 60 e 75 *rotazioni per minuto (rpm)* medie. Alla fine del test si raccolgono i seguenti dati:

RPM media	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
N.ro Atleti	3	1	1	10	2	6	15	20	7	9	13	5	6	1	1

In relazione ai dati riportati in tabella:

a. Ricavare un istogramma delle frequenze per le classi $60-64$, $65-66$, 67, $68-69$, $70-74$ relativamente al dato *rpm media*

b. Determinare media e varianza campionaria

c. Determinare intervalli di confidenza al livello del 95% e del 99% per la stima della *rpm media* ottenuta dal gruppo ciclistico nel test.