■ 15 febbraio 2016 es.1 (Lucy e Sally)

Lucy e Sally disputano il seguente gioco: Da un mazzo di 40 carte, Lucy ne estrae una, senza reintrodurla; poi Sally ne estrae due, poi Lucy tre, poi ancora Sally quattro, e così via finché ci sono carte nel mazzo.
a) Qual è la probabilità che sia Sally a estrarre l'asso di denari?
b) Qual è la probabilità che sia Sally a estrarre l'asso di denari e Lucy a estrarre l'asso di spade?

Soluzione

a) Dopo n estrazioni complessive delle due giocatrici le carte estratte sono in tutto $\frac{n(n+1)}{2}$; dopo 8 estrazioni sono 36; la nona estrazione, che tocca a Lucy, sarà di sole 4 carte che esauriranno il mazzo. Alla fine il numero di carte estratte da Lucy sarà 1+3+5+7+4=20, e altrettante ne avrà Sally. La probabilità che una determinata carta (per esempio l'asso di denari) sia alla fine in mano a Lucy o a Sally vale quindi $\frac{20}{40} = \frac{1}{2}$.

b) Siano A l'evento "Sally estrae l'asso di denari", B l'evento "Lucy estrae l'asso di spade"; cerchiamo $P(A \wedge B)$. Si ha:

$P(A \wedge B) = P(A) \cdot P(B|A)$

$P(B|A)$ si può calcolare nel modo seguente: se sappiamo che l'asso di denari è in mano a Sally, le 20 carte di Lucy provengono dalle 39 carte del mazzo privato dell'asso di denari; quindi la probabilità che tra esse ci sia una determinata carta, per esempio l'asso di spade, vale $\frac{20}{39}$. Perciò $P(A \wedge B) = P(A) \cdot P(B|A) = \frac{1}{2} \cdot \frac{20}{39} = \frac{10}{39}$.

Un altro modo per calcolare $P(A \wedge B)$ è il seguente. L'evento $A \wedge B$ si può descrivere come "Sally estrae l'asso di denari ma non l'asso di spade. Gli insiemi di 20 carte estratte dal mazzo, contenenti l'asso di denari ma non l'asso di spade sono in numero di $\binom{38}{19}$; infatti all'asso di denari (che ci deve essere) debbono essere accompagnate 19 carte da scegliere fra le carte rimanenti nel mazzo, diverse dall'asso di spade; queste carte sono in numero di 38. I possibili insiemi di 20 carte che può ricevere Sally sono in numero di $\binom{40}{20}$; dunque $P(A \wedge B)$ vale $\frac{\binom{38}{19}}{\binom{40}{20}} = \frac{38!}{19! \cdot 19!} \cdot \frac{40!}{20! \cdot 20!} = \frac{10}{39}$.

■ 15 febbraio 2016 es.2 (Romeo e le sue amiche)

Romeo ha tre amiche: Anna, Maria, Giulietta. Ogni domenica Romeo esce con una di loro; se si trova bene, la domenica successiva esce nuovamente con la stessa amica, altrimenti ne sceglie una a caso delle altre due. La probabilità che Romeo sia contento quando esce con Anna è 0,8; quando esce con Maria è 0,6; sarà certamente contento se esce con Giulietta.
Scrivere la matrice di transizione della catena di Markov che rappresenta l'accompagnatrice di Romeo in una data domenica (stato 1=Anna, 2=Maria, 3=Giulietta) e studiare le sue proprietà. Verificare che un solo stato è ricorrente (e assorbente).
Calcolare infine il tempo medio di assorbimento in tale stato, nel caso che si parta da uno o dall'altro degli altri due.

Soluzione

La matrice di transizione è $P = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0 & 0 & 1 \end{pmatrix}$. Gli stati 1 e 2 sono transitori, perché entrambi comunicano con 3, il quale non comunica con nessuno degli altri due perché è assorbente.

I tempi medi τ_1, τ_2 di assorbimento nello stato 3, partendo rispettivamente dallo stato 1 o 2 soddisfano il sistema

$$\begin{cases} 1 = \tau_1 - p_{1,1}\,\tau_1 - p_{1,2}\,\tau_2 \\ 1 = \tau_2 - p_{2,1}\,\tau_1 - p_{2,2}\,\tau_2 \end{cases} \quad \text{cioè} \quad \begin{cases} 1 = \tau_1 - 0.8\,\tau_1 - 0.1\,\tau_2 \\ 1 = \tau_2 - 0.2\,\tau_1 - 0.6\,\tau_2 \end{cases}$$

risolvendo il quale si trova $\tau_1 = \frac{25}{3}$, $\tau_2 = \frac{20}{3}$.

15 febbraio 2016 es.3 (L’assicurazione per Eva)

Adamo ha 30 anni, come sua moglie Eva. Adamo stipula un’assicurazione a favore di Eva, con le seguenti caratteristiche: se Adamo morirà prima di compiere 40 anni, ed Eva gli sopravvivrà, ella riceverà C=100.000€ nel giorno del suo primo compleanno da vedova; nulla sarà dovuto in tutti gli altri casi, cioè se Adamo raggiungerà in vita il 40° compleanno, oppure Eva passerà a miglior vita prima di Adamo, o prima di compiere gli anni da vedova. Calcolare, esprimendoli attraverso i coefficienti delle tavole di mortalità e il coefficiente di sconto $v = \frac{1}{1+i}$ (i indica il tasso annuo di riferimento):

a) il premio unico teorico A che Adamo deve pagare oggi per l’assicurazione, cioè il valore attuale della speranza matematica della somma che nel caso descritto sarà corrisposta a Eva.

b) L’importo di una rata annua anticipata s che Adamo s’impegna a pagare in alternativa al premio unico di a), per 5 anni da oggi, a condizione che in ciascuna scadenza siano in vita sia lui sia Eva. (s deve essere tale che sia uguale ad A la speranza matematica dei valori attuali delle cinque rate).

Soluzione

a) Alla fine del primo anno (31° compleanno di Adamo ed Eva) la Compagnia pagherà C se Adamo sarà morto ed Eva sarà viva; la probabilità di questo evento è $\frac{d_{30}}{\ell_{30}} * \frac{\ell'_{31}}{\ell'_{31}}$ (usiamo gli apici per indicare i dati relativo alla popolazione di sesso femminile). Il valore attuale di C pagato tra un anno è $C * v$; quindi la speranza matematica del valore attuale dell’eventuale pagamento dopo un anno è:

$$C * v * \frac{d_{30}}{\ell_{30}} * \frac{\ell'_{31}}{\ell'_{30}}$$

Analogamente, la speranza matematica del valore attuale dell’eventuale pagamento dopo due anni è

$$C * v * \frac{d_{30}}{\ell_{30}} * \frac{\ell'_{31}}{\ell'_{30}}$$

Analogamente, la speranza matematica del valore attuale dell’eventuale pagamento dopo due anni è

$$C * v^2 * \frac{d_{31}}{\ell_{30}} * \frac{\ell'_{32}}{\ell'_{30}}$$

E così via, fino al 40° compleanno di Eva. Il premio unico è la somma di questi addendi:

$$A = \sum_{k=1}^{10} C * v^k * \frac{d_{30+k-1}}{\ell_{30}} * \frac{\ell'_{30+k}}{\ell'_{30}}$$

b) Ora ci dobbiamo mettere nei panni della Compagnia assicuratrice, la quale da oggi e per cinque anni riceverà s alla scadenza di ciascun anno, se e solo se Adamo ed Eva saranno in vita.
s deve essere tale che sia uguale ad A la speranza matematica dei valori attuali delle cinque rate (l'espressione di A è quella calcolata al punto precedente).
La prima rata è certa ed immediata, quindi non va scontata né moltiplicata per la probabilità; la seconda rata, da pagarsi tra un anno, sarà effettivamente pagata con probabilità $\frac{l_{31}}{\ell_{30}} * \frac{\ell'_{31}}{\ell'_{30}}$, quindi la speranza matematica del valore attuale della rata da pagare tra un anno è $s * v * \frac{\ell_{31}}{\ell_{30}} * \frac{\ell'_{31}}{\ell'_{30}}$. Ragionando nello stesso modo si ottiene la relazione

$$A = \sum_{k=0}^{4} s * v^k * \frac{\ell_{30+k}}{\ell_{30}} * \frac{\ell'_{30+k}}{\ell'_{30}}$$

e quindi

$$s = A \Big/ \left(\sum_{k=0}^{4} v^k * \frac{\ell_{30+k}}{\ell_{30}} * \frac{\ell'_{30+k}}{\ell'_{30}} \right)$$

15 febbraio 2016 es.4 (Le uova di Nonna Papera)

Sei uova provenienti dal pollaio di Nonna Papera vengono pesate, rilevando i seguenti valori in grammi:

```
lista = {60, 65, 70, 80, 85, 90};
```

Si suppone che il peso di un generico uovo segua una distribuzione normale, con media μ e varianza σ^2 sconosciute. Determinare al livello del 95% un intervallo di confidenza per la media μ e uno per lo scarto quadratico medio σ, il primo centrato nel valore rilevato della media campionaria, il secondo del tipo $[0,b]$.

Soluzione

La media campionaria per il nostro campione è

$$\overline{x} = \frac{1}{6}\sum_{k=1}^{6} \text{lista}[[k]]$$

```
75
```

La stima corretta della varianza è

$$s2 = \frac{1}{5}\sum_{k=1}^{6} (\text{lista}[[k]] - \overline{x})^2$$

```
140
```

$$s = N\left[\sqrt{s2}\right]$$

```
11.8322
```

L'intervallo di confidenza al livello del 95% (ossia 1-α con α=0.05) è
$\left[\overline{x} - \frac{s}{\sqrt{6}} t_{1-\frac{\alpha}{2}}, \overline{x} + \frac{s}{\sqrt{6}} t_{1-\frac{\alpha}{2}}\right]$
con $t_{1-\frac{\alpha}{2}}$ quantile di livello $1 - \frac{\alpha}{2} = 0.975$ per la distribuzione di Student con 5 gradi di libertà. Il risultato è il seguente:

```
t = Quantile[StudentTDistribution[5], 0.975]
```

```
2.57058
```

$$\text{Flatten}\left[\left\{\overline{x} - \frac{s}{\sqrt{6}} t,\ \overline{x} + \frac{s}{\sqrt{6}} t\right\}\right]$$

```
{62.5829, 87.4171}
```

Un intervallo di confidenza al livello 1-α per σ^2 è $\left[0, \frac{(n-1)\,s^2}{\chi^2_\alpha(n-1)}\right]$ con n taglia del campione e $\chi^2_\alpha(n-1)$ quantile di livello α per χ^2 con $n-1$ gradi di libertà. Qui ci serve $q = \chi^2_{0.05}(5)$ che vale

```
q = Quantile[ChiSquareDistribution[5], .05]
```

```
1.14548
```

e quindi il secondo estremo dell'intervallo di confidenza per σ^2 è

$$\frac{5\, s^2}{q}$$

```
611.1
```

che corrisponde a un maggiorante dello scarto quadratico medio uguale a

$$\sqrt{\frac{5\,s^2}{q}}$$

```
24.7204
```

15 febbraio 2016 es.5 (mangiasoldi binomiale)

In una macchina da gioco, ciascuna giocata costa 1€, e la macchina dichiara di elargire un premio aleatorio di 0€, 1€, 2€, 3€ secondo una distribuzione binomiale $B\left(3, \frac{1}{5}\right)$. Un accanito giocatore gioca per 400 volte, e prende nota della vincita lorda in ciascuna giocata. Alla fine i dati raccolti sono i seguenti:

$$\begin{pmatrix} \text{importo vincita} & 0 & 1 & 2 & 3 \\ \text{frequenza} & 188 & 174 & 33 & 5 \end{pmatrix}$$

a) Stabilire se al livello del 5% si può ritenere che quanto dichiarato sulla macchina sia vero, cioè che la macchina assegni i premi secondo una distribuzione $B\left(3, \frac{1}{5}\right)$.

b) calcolare qual è il numero minimo di giocate con le quali le stesse frequenze relative delle vincite condurrebbero al rifiuto dell'ipotesi.

Soluzione

a) La verifica dell'ipotesi avviene mediante il test χ^2 di Pearson. Le probabilità di ricevere 0, 1, 2, 3 secondo una $B\left(3, \frac{1}{5}\right)$ sono rispettivamente

```
p = Table[PDF[BinomialDistribution[3, 0.2], k], {k, 0, 3}]
```

```
{0.512, 0.384, 0.096, 0.008}
```

Le frequenze relative sperimentali sono rispettivamente

$$q = \frac{1}{400.}\{n0, n1, n2, n3\}$$

```
{0.47, 0.435, 0.0825, 0.0125}
```

La statistica-test è

$$t = 400 \sum_{k=1}^{4} \frac{(q[[k]] - p[[k]])^2}{p[[k]]}$$

```
5.85938
```

La regione di rigetto dell'ipotesi al livello α è $]\chi^2_{1-\alpha}(3), +\infty[$; con $\alpha = 0,05$ il quantile è

$\chi^2_{1-\alpha}(3)$ **= Quantile[ChiSquareDistribution[3], 0.95]**

```
7.81473
```

L'ipotesi non viene rifiutata.

b) Bisogna che n (numero delle giocate) sia tale che la statistica test

$$t1 = n \sum_{k=1}^{4} \frac{(q[[k]] - p[[k]])^2}{p[[k]]}$$

```
0.0146484 n
```

assuma valore ⩾ 7.81473

```
Solve[t1 == 7.814727903251178`, n]
```

{{n → 533.485}}

Il minimo numero di giocate affinché queste frequenze relative conducano al rifiuto dell'ipotesi è 534.

■ 15 febbraio 2016 es.6 (secondo con contorno)

Il proprietario di una trattoria offre tre alternative come "secondo": bistecca, pollo arrosto, cotechino e tre come "contorno: patate fritte, purè, spinaci. Controllando gli ordini di (n=)500 clienti egli osserva le seguenti frequenze per gli accoppiamenti di (secondo+contorno):

$$\begin{pmatrix} 70 & 30 & 50 & \text{bistecca} \\ 80 & 90 & 80 & \text{pollo} \\ 50 & 30 & 20 & \text{cotechino} \\ \text{patate} & \text{purè} & \text{spinaci} & 0 \end{pmatrix}$$

a) Stabilire se, al livello 1%, si può ritenere che la scelta del secondo e quella del contorno siano indipendenti.
b) La risposta cambierebbe se le stesse frequenze relative per le scelte dei diversi accoppiamenti fossero state osservate per un campione di soli 100 clienti, anziché 500?

Soluzione

La seguente matrice aggiunge ai dati del problema le distribuzioni marginali relative a "secondo" e "contorno":

□	**patate**	**purè**	**spinaci**	***marg.secondo***
bistecca	**70**	**30**	**50**	**150**
pollo	**80**	**90**	**80**	**250**
cotechino	**50**	**30**	**20**	**100**
marg.contorno	**200**	**150**	**150**	**500**

Questa corrisponde alle seguenti frequenze relative (accoppiate e marginali)

```
Print[MatrixForm[m1f]]
```

$$\begin{pmatrix} 0.14 & 0.06 & 0.1 & 0.3 \\ 0.16 & 0.18 & 0.16 & 0.5 \\ 0.1 & 0.06 & 0.04 & 0.2 \\ 0.4 & 0.3 & 0.3 & 1. \end{pmatrix}$$

Invece la matrice con le frequenze relative accoppiate teoriche in caso di indipendenza, cioè i prodotti delle distribuzioni marginali è

```
Print[MatrixForm[m2f]]
```

$$\begin{pmatrix} 0.12 & 0.09 & 0.09 & 0.3 \\ 0.2 & 0.15 & 0.15 & 0.5 \\ 0.08 & 0.06 & 0.06 & 0.2 \\ 0.4 & 0.3 & 0.3 & 1. \end{pmatrix}$$

La statistica-test per l'indipendenza delle due variabili "secondo" e "contorno" è

$$t = n * \sum_{h=1}^{3} \sum_{k=1}^{3} \frac{(m1f[[h]][[k]] - m2f[[h]][[k]])^2}{m2f[[h]][[k]]}$$

20.3889

Se le due variabili considerate sono indipendenti, *t* ha aprossimativamente distribuzione $\chi^2(4)$ (i gradi di libertà sono (3-1)× (3-1)). Il quantile che limita inferiormente la regione di rifiuto del'ipotesi di indipendenza è

```
q = Quantile[ChiSquareDistribution[4], .99]
```

```
13.2767
```

L'ipotesi viene rifiutata.

b) Se *n* passa da 500 a 100 lasciando invariate le frequenze relative, il solo cambiamento nel calcolo del valore sperimentale della statistica *t* è appunto il fattore *n*; si ottiene in questo caso

$$\texttt{t = 100} * \sum_{h=1}^{3} \sum_{k=1}^{3} \frac{(\texttt{m1f[[h]][[k]] - m2f[[h]][[k]]})^2}{\texttt{m2f[[h]][[k]]}}$$

```
4.07778
```

In questo caso non ci sono elementi sufficienti per il rifiuto dell'ipotesi.