13 febbraio 2018 es.1 (La lotteria del Pedale Bolognese)

La prima domenica di dicembre i soci del Pedale Bolognese e i loro familiari s'incontrano per il pranzo sociale di fine stagione; in quell'occasione si svolge anche un sorteggio di premi tra i convitati: a ciascuno viene assegnato un biglietto numerato, e per sorteggio si assegnano premi ad alcuni di questi numeri. I convitati sono 99; i numeri vincenti sono n, con $33 \le n \le 96$.

a) Calcolare, in funzione di n, la probabilità che i due soci più giovani, Jacopo e Iuri, siano entrambi premiati.

b) I Consiglieri che organizzano la lotteria desiderano predisporre un numero n di premi tale che sia $\ge 90\,\%$ la probabilità che la famiglia del Presidente della Società, formata da tre persone, riceva almeno un premio. Verificare che il minimo valore di n che soddisfa questa condizione è $n = 53$.

Soluzione

a) Se i biglietti vincenti sono n, la probabilità che un numero scelto a caso tra 1 e 99, ad esempio quello di Iuri, sia vincente è $\frac{n}{99}$. Se il biglietto di Iuri risulta vincente, la probabilità che lo sia anche quello di Jacopo è $\frac{n-1}{98}$. La probabilità che entrambi i ragazzi siano premiati è $\frac{n}{99} \cdot \frac{n-1}{98}$.

b) Calcoliamo la probabilità dell'evento contrario, cioè che la famiglia del Presidente, formata da 3 persone, non riceva alcun premio. Ciò significa che tutti e tre i numeri posseduti da quella famiglia appartengono all'insieme dei $99 - n$ numeri non vincenti. La probabilità che ciò accada è

$$\texttt{pr[n\_]} := \frac{99-n}{99} * \frac{98-n}{98} * \frac{97-n}{97}$$

Si desidera che n sia tale che $\text{pr}[n] \le 0.10$; evidentemente $\text{pr}[n]$ diminuisce al crescere di n, e si ha

```
Print[{pr[52.], pr[53.]}]
```

```
{0.10338, 0.096781}
```

Quindi il minimo numero n di premi che soddisfa il requisito di (b) è $n = 53$.

13 febbraio 2018 es.2 (Tentativi senza memoria)

La password per utilizzare un dispositivo elettronico è una stringa ordinata di tre cifre decimali diverse tra loro; all'inizio appare questo display: ★★★; il dispositivo conferma la correttezza di ciascuna cifra una per volta, cioè, per esempio, se la password è 752 e digitiamo 7 come prima cifra, riceveremo subito la conferma della sua esattezza, e da quel momento vedremo: 7★★.
Un utente che non conosce la password la cerca per tentativi, sulla prima cifra, poi sulla seconda, poi sulla terza; costui è molto distratto, e se digita una cifra errata, subito dimentica qual era la cifra inserita, cosicché nel tentativo immediatamente successivo potrebbe ritornare a inserirla.

a) Descrivere con una catena di Markov a 4 stati (0,1,2,3), che indicano quante cifre sono state indovinate fino a quel momento, la sequenza dei tentativi per accedere al dispositivo; classificare gli stati, stabilire se la catena è irriducibile e se è regolare.

b) Calcolare il numero medio di tentativi che occorrono per indovinare la password (ossia, il numero medio di transizioni per passare dallo stato 0 allo stato 3)

Soluzione

a) La matrice di transizione si costruisce facilmente, tenendo presente le modalità con cui si svolgono i tentativi:

$$\begin{array}{c|cccc} & \text{(a)} \; 0 & 1 & 2 & 3 \\ \text{(da)} & & & & \\ 0 & \frac{9}{10} & \frac{1}{10} & 0 & 0 \\ 1 & 0 & \frac{8}{9} & \frac{1}{9} & 0 \\ 2 & 0 & 0 & \frac{7}{8} & \frac{1}{8} \\ 3 & 0 & 0 & 0 & 1 \end{array} \; ; \quad p = \begin{pmatrix} \frac{9}{10} & \frac{1}{10} & 0 & 0 \\ 0 & \frac{8}{9} & \frac{1}{9} & 0 \\ 0 & 0 & \frac{7}{8} & \frac{1}{8} \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

Lo stato 3 è assorbente, quindi ricorrente; gli stati 0, 1, 2 sono transitori, perché ciascuno di essi comunica con 3, ma quest'ultimo non comunica con nessuno dei precedenti. La catena non è pertanto irriducibile, quindi nemmeno regolare.

b) Indicati con τ_i ($i = 0, 1, 2$) i valori medi della variabile "numero di transizioni necessarie per arrivare allo stato

3”, condizionate a $X_0 = 0$, valgono le relazioni

$$\begin{cases} \tau_0 = 1 + \frac{9}{10}\,\tau_0 + \frac{1}{10}\,\tau_1 \\ \tau_1 = 1 + \frac{8}{9}\,\tau_1 + \frac{1}{9}\,\tau_2 \\ \tau_2 = 1 + \frac{7}{8}\,\tau_2 \end{cases}$$

La richiesta è di calcolare τ_0 , ma per ottenerlo bisogna anche ricavare τ_1 e τ_2

```
Solve[{
  t0 == 1 + 9/10 t0 + 1/10 t1,
  t1 == 1 + 8/9 t1 + 1/9 t2,
  t2 == 1 + 7/8 t2
 }, {t0, t1, t2}]
{{t0 → 27, t1 → 17, t2 → 8}}
```

13 febbraio 2018 es.3 (Abramo, Sara e Isacco)

Abramo (70) e Sara (60) sono gli anziani genitori del neonato Isacco.

a) Calcolare, esprimendola mediante i dati delle tavole di sopravvivenza, la probabilità p_k che tra k anni accada *per la prima volta* che Isacco compie gli anni privo di entrambi i genitori.

b) Calcolare, ancora in funzione dei dati di sopravvivenza e anche del tasso di mercato i, il premio puro che Abramo e Sara debbono pagare oggi per garantire a Isacco la somma di 200 000 €, che gli saranno erogati in unica soluzione nel suo primo compleanno, fino al 18° compreso, in cui egli dovesse trovarsi in vita privo di entrambi i genitori; nel risultato si lasci per semplicità indicata p_k come tale, senza sostituire il simbolo p_k con l’espressione calcolata in (a).

Soluzione

a) Sia A_k l’evento descritto, di cui p_k è la probabilità. A_k si verifica se tra k anni Isacco è in vita ma non lo sono i suoi genitori, mentre un anno prima almeno uno di essi era vivente. Definiamo i seguenti eventi:

B_k : Isacco è in vita nel suo k – esimo compleanno

C_k : Abramo muore in età $70 + k - 1$ e Sara muore precedentemente al suo compleanno $60 + k - 1$

D_k : Sara muore in età $60 + k - 1$ e Abramo muore precedentemente al suo compleanno $70 + k - 1$

E_k : Abramo muore in età $70 + k - 1$ e Sara muore in età $60 + k - 1$

si ha $A_k = B_k \wedge (C_k \vee D_k \vee E_k)$

Gli eventi di cui si fa l’intersezione sono indipendenti, mentre i tre di cui si fa l’unione sono a due a due disgiunti. Poi si ha

$$P(B_k) = \frac{l_k}{l_0}\left(= \frac{l_k}{100\,000}\right);$$

$$P(C_k) = \frac{d_{69+k}}{l_{70}} \cdot \left(1 - \frac{l'_{59+k}}{l'_{60}}\right); \quad P(D_k) = \frac{d'_{59+k}}{l'_{60}} \cdot \left(1 - \frac{l_{69+k}}{l_{70}}\right); \quad P(E_k) = \frac{d_{69+k}}{l_{70}} \cdot \frac{d'_{59+k}}{l'_{60}}$$

e quindi

$$p_k = P(A_k) = P(B_k) \cdot (P(C_k) + P(D_k) + P(E_k)) = \frac{l_k}{l_0} \cdot \left(\frac{d_{69+k}}{l_{70}} \cdot \left(1 - \frac{l'_{59+k}}{l'_{60}}\right) + \frac{d'_{59+k}}{l'_{60}} \cdot \left(1 - \frac{l_{69+k}}{l_{70}}\right) + \frac{d_{69+k}}{l_{70}} \cdot \frac{d'_{59+k}}{l'_{60}}\right).$$

b) L’assicurazione è tenuta al pagamento se si verifica uno degli eventi A_k , con $1 \le k \le 18$. Se si verifica A_k l’assicurazione pagherà tra k anni da oggi 200 000 €, corrispondenti a un valore attuale $200\,000 \cdot \frac{1}{(1+i)^k}$. La probabilità che accada A_k è p_k ; quindi l’importo del premio puro è

$$200\,000 \cdot \sum_{k=1}^{18} \frac{1}{(1+i)^k} \cdot p_k \,.$$

13 febbraio 2018 es.4 (L’ascensore a Matematica)

L'ascensore del Dipartimento di Matematica è tarato per una portata massima di 700 Kg; se il peso viene superato suona un allarme e qualcuno deve scendere, altrimenti l'ascensore non parte.
È stato misurato il peso di otto studentesse di Matematica scelte a caso, con i seguenti risultati, che qui pensiamo come valori sperimentali di una variabile normale con media μ e varianza σ^2 sconosciute:

```
lista = {60, 62, 61, 56, 59, 58, 57, 59};
```

a) 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 misurato della media campionaria, il secondo della forma $[0,b]$.

b) Undici studentesse di Matematica sono entrate nell'ascensore. Utilizzando i risultati di (a), dare una stima dal basso alla probabilità p che l'ascensore possa partire (cioè stabilire una disuguaglianza della forma $p \geq \ldots$).

Soluzione

La media campionaria per i dati elencati è

$$\overline{\mathbf{x}} = \frac{1}{8} \sum_{k=1}^{8} \mathbf{lista[[k]]}$$

59

La stima corretta della varianza è

$$\mathbf{s2} = \frac{1}{7} \sum_{k=1}^{8} \left(\mathbf{lista[[k]]} - \overline{\mathbf{x}}\right)^2$$

4

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

2.

L'intervallo di confidenza al livello del 95% (ossia 1-α con α=0.05) è

$$\left[\overline{x} - \frac{s}{\sqrt{8}} t_{1-\frac{\alpha}{2}}, \overline{x} + \frac{s}{\sqrt{8}} t_{1-\frac{\alpha}{2}}\right]$$

con $t_{1-\frac{\alpha}{2}}$ quantile di livello $1 - \frac{\alpha}{2} = 0.975$ per la distribuzione di Student con 7 gradi di libertà. Il risultato è il seguente:

$$\mathbf{t_{0.975}} = \mathbf{Quantile[StudentTDistribution[7], 0.975]}$$

2.36462

$$\left\{\overline{x} - \frac{s}{\sqrt{8}} t_{0.975}, \overline{x} + \frac{s}{\sqrt{8}} t_{0.975}\right\}$$

{57.328, 60.672}

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 $q = \chi^2_{0.05}(7)$ che vale

```
q = Quantile[ChiSquareDistribution[7], .05]
```

2.16735

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

$$\frac{\mathbf{7\, s^2}}{\mathbf{q}}$$

12.919

che corrisponde a un maggiorante dello scarto quadratico medio uguale a

$$\mathbf{r} = \sqrt{\frac{\mathbf{7\, s^2}}{\mathbf{q}}}$$

3.5943

b) Il peso complessivo Y delle 11 studentesse è una variabile aleatoria somma di 11 normali indipendenti, ciascuna con media μ e scarto quadratico medio σ stimati in (a). Allora $Y \sim N(11\,\mu,\ 11\,\sigma^2)$ e perciò la variabile $Z = \frac{Y-11\,\mu}{\sigma\sqrt{11}}$ ha distribuzione $N(0,\ 1)$; l'evento "L'ascensore può partire" significa «$Y \le 700$», cioè $Z \le \frac{700-11\,\mu}{\sigma\sqrt{11}}$. Se $\Phi(\bullet)$ è la funzione ripartizione di $N(0,\ 1)$ allora, per definizione è

$$P(Y \le 700) = P\left(Z \le \frac{700-11\,\mu}{\sigma\sqrt{11}}\right) = \Phi\left(\frac{700-11\,\mu}{\sigma\sqrt{11}}\right)$$

Abbiamo trovato in (a) che con probabilità 0.95 la media μ è minore o uguale di

```
m = x̄ + s/√8 t0.975
```

```
60.672
```

e, indipendentemente dalla stima di μ, e ancora con probabilità 0.95 lo scarto quadratico medio σ è minore o uguale di

```
r = √(7 s²/q)
```

```
3.5943
```

e quindi con probabilità

```
0.95²
```

```
0.9025
```

valgono entrambe le maggiorazioni, e in tal caso si ha $\frac{700-11\,\mu}{\sigma\sqrt{11}} \ge \frac{700-11\,m}{\sigma\sqrt{11}} \ge \frac{700-11\,m}{r\sqrt{11}}$; quest'ultima espressione vale

```
(700 - 11 m)/(r √11)
```

```
2.73531
```

Siccome Φ è strettamente crescente, ***se è vero che*** $\frac{700-11\,\mu}{\sigma\sqrt{11}} \ge \frac{700-11\,m}{r\sqrt{11}}$ allora è

$$P(Y \le 700) = \Phi\left(\frac{700-11\,\mu}{\sigma\sqrt{11}}\right) \ge \Phi\left(\frac{700-11\,m}{r\sqrt{11}}\right)$$

Quest'ultimo valore è

```
p1 = CDF[NormalDistribution[0, 1], (700 - 11 m)/(r √11)]
```

```
0.996884
```

$$p = P(Y \le 700) \ge P\left(\left(\frac{700-11\,\mu}{\sigma\sqrt{11}} \ge \frac{700-11\,m}{r\sqrt{11}}\right) \wedge Y \le 700\right) =$$

$$= P\left(\frac{700-11\,\mu}{\sigma\sqrt{11}} \ge \frac{700-11\,m}{r\sqrt{11}}\right) \cdot P\left(Y \le 700 \;\middle|\; \left(\frac{700-11\,\mu}{\sigma\sqrt{11}} \ge \frac{700-11\,m}{r\sqrt{11}}\right)\right)$$

cioè si ha che p è maggiore o uguale di

```
p1 * 0.95²
```

```
0.899688
```

13 febbraio 2018 es.5 (Riduzione del consumo di benzina)

Un'industria specializzata produce un olio lubrificante di nuova concezione per automobili, e sostiene che l'impiego di quel lubrificante diminuisce il consumo di benzina. Per verificare l'attendibilità di questa affermazione, sette collaudatori percorrono 1000 Km ciascuno, con sette automobili dello stesso modello, che impiegano olio convenzionale; poi ripetono il percorso dopo il cambio dell'olio. I consumi di benzina di ciascun pilota, prima con l'olio "vecchio", poi con quello "nuovo" sono qui sotto riportati, espressi in litri:

$$\text{tabella} = \left(\frac{38\ \ 41\ \ 40\ \ 41\ \ 38\ \ 42\ \ 41\quad (\text{con olio convenzionale})}{39\ \ 40\ \ 41\ \ 40\ \ 37\ \ 40\ \ 37\quad (\text{con olio di nuova concezione})}\right);$$

a) Stabilire se si può ritenere al livello 5% che il nuovo olio mantiene la promessa; cioè, detti μ_1, μ_2 rispettivamente i consumi medi per 1000 Km con lubrificante convenzionele e con lubrificante di nuovo tipo, verificare se i dati sperimentali consentono di rifiutare al livello 5% l'ipotesi $\mu_1 \leq \mu_2$ (test unilaterale).

b) Ripetere la verifica supponendo ora che i consumi con olio tradizionale siano gli stessi di (a), mentre quelli con olio nuovo siano tutti inferiori di un litro rispetto a quelli della seconda riga della tabella di (a), cioè siano 38, 39, 40, 39, 36, 39, 36. Non occorre rifare tutti i calcoli!

Soluzione

a) La variabile su cui dobbiamo lavorare è $Z = X - Y$, essendo X la quantità di benzina occorrente per 1000 Km con lubrificante convenzionale, Y la benzina occorrente con lubrificante di nuovo tipo. I valori campionari Z_i di Z sono le differenze $X_i - Y_i$, essendo X_i e Y_i i valori della tabella data sopra:

```
ListaZ = Table[tabella[[1]][[k]] - tab[[2]][[k]], {k, 1, 7}]
```

```
{-1, 1, -1, 1, 1, 2, 4}
```

La media campionaria e la varianza corretta corrispondenti a questi valori sperimentali sono

$\bar{z}$ = N[Mean[ListaZ]]

```
1.
```

```
s2 = N[Variance[ListaZ]]
```

```
3.
```

cosicché lo stimatore dello scarto quadratico medio è

$s = \sqrt{s2}$

```
1.73205
```

È ragionevole supporre X e Y distribuite normalmente; detta $\mu_Z = \mu_1 - \mu_2$, la variabile $\sqrt{7}\ \frac{\bar{z}-\mu_Z}{s}$ ha distribuzione di Student con 6 gradi di libertà.

L'ipotesi che si vorrebbe confutare è: $\mu_Z \leq 0$. La statistica da usare è il numero $T = \sqrt{7}\ \frac{\bar{z}}{s}$. Un intervallo di rifiuto al livello α per questa ipotesi, relativo alla statistica T, è $]\,t_{1-\alpha}(6), +\infty[$. Nel caso attuale abbiamo $\alpha = 0.05$, quindi

$t_{1-\alpha}(7)$ = Quantile[StudentTDistribution[6], 0.95]

```
1.94318
```

Il valore di T corrispondente ai dati sperimentali è

$T = \sqrt{7}\ \frac{\bar{z}}{s}$

```
1.52753
```

il quale non appartiene all'intervallo di rigetto; l'ipotesi non viene rifiutata, al livello 5%, vale a dire che a questo livello, non si può escludere che la diminuzione complessiva di consumo sia dovuta a fattori casuali (vento, condizioni di traffico...).

b) Chiamati Y_i' i nuovi valori, è $Y_i' = Y_i - 1$ quindi $Z_i' = X_i - Y_i' = Z_i + 1$ e quindi la media campionaria cresce di 1 mentre la varianza stimata rimane invariata. La statistica-test con i nuovi dati diventa quindi

$T_1 = \sqrt{7}\ \frac{\bar{z}+1}{s}$

```
3.05505
```

questa volta appartenente all'intervallo di rifiuto dell'ipotesi (che è lo stesso di (a)); quindi in questo caso l'ipotesi viene respinta.

13 febbraio 2018 es.6 (Taratura della *slot-machine* del bar)

La *slot-machine* collocata in un bar dichiara, in un avviso scritto in caratteri di piccole dimensioni, le possibili vincite lorde e le rispettive probabilità, per ogni giocata del costo di 1€:

$$\begin{pmatrix} \text{vincite €} & 50 & 10 & 5 & 0 \\ \text{probabilità} & 0.01 & 0.02 & 0.03 & 0.94 \end{pmatrix}$$

Per controllare la correttezza di quanto indicato, il barista prende nota dei risultati di 500 giocate di avventori del suo bar, ottenendo

$$\begin{pmatrix} \text{vincite €} & 50 & 10 & 5 & 0 \\ \text{frequenze} & 3 & 11 & 16 & 470. \end{pmatrix}$$

a) Stabilire se, al livello del 5%, si può ritenere che la macchina non funzioni correttamente, cioè se si deve respingere l'ipotesi che le probabilità delle possibili vincite siano quelle dichiarate.

b) Assumendo come veritiere le probabilità dichiarate sulla macchina, determinare *M* tale che: con probabilità 99% sia almeno *M*€ il guadagno per il barista relativamente a 10 000 giocate.

Soluzione

a) Le probabilità dichlarate per le diverse vincite sono, come riporta il testo,

In[4]:= **{pa, pb, pc, pd}**

Out[4]= {0.01, 0.02, 0.03, 0.94}

Queste vanno confrontate con le frequenze relative delle vincite per i 500 giocatori osservati:

In[9]:= $\{\overline{pa}, \overline{pb}, \overline{pc}, \overline{pd}\} = \frac{1}{n}\{qa, qb, qc, qd\}$

Out[9]= {0.006, 0.022, 0.032, 0.94}

La statistica-test è

In[10]:= $$T = n * \left(\frac{(pa - \overline{pa})^2}{pa} + \frac{(pb - \overline{pb})^2}{pb} + \frac{(pc - \overline{pc})^2}{pc} + \frac{(pd - \overline{pd})^2}{pd}\right)$$

Out[10]= 0.966667

Se l'ipotesi è vera, *T* ha approssimativamente distribuzione $\chi^2(3)$; l'intervallo di rifiuto dell'ipotesi al livello 5% è $]\chi^2_{0.95}(3), +\infty] =]7.8147, +\infty]$. Il valore osservato di *T* non appartiene all'intervallo di rifiuto; l'ipotesi non può essere respinta.

b) Sia *X* la variabile "vincita lorda in una giocata". La media di *X* è

In[11]:= $\mu = 50 * pa + 10 * pb + 5 * pc$

Out[11]= 0.85

Calcoliamo la varianza; la media di X^2 è

In[12]:= $50^2 * pa + 10^2 * pb + 5^2 * pc$

Out[12]= 27.75

cosicché lo scarto quadratico medio di *X* è

In[13]:= $\sigma = \sqrt{\% - \mu^2}$

Out[13]= 5.1988

La somma delle vincite elargite su 10 000 giocate è $S = X_1 + \ldots + X_{10\,000}$, con le X_k indipendenti ed equidistribuite a *X*.

Per il Teorema limite centrale, la variabile $Z = \frac{S - 10\,000\,\mu}{\sigma\sqrt{10\,000}}$ ha in pratica esattamente distribuzione $N(0, 1)$, grazie al valore assai grande $n = 10\,000$. Si ha $0.99 = P(Z \leq \phi_{0.99})$, con

```
ϕ0.99 = Quantile[NormalDistribution[0, 1], 0.99]
```

```
2.32635
```

In[14]:= $\text{Solve}\left[\frac{S - 10\,000\,\mu}{\sigma\sqrt{10\,000}} == \phi_{0.99},\ S\right]$

Out[14]= $\{\{S \to 519.88\ (16.3499 + 1.\ \phi_{0.99})\}\}$

Ciò significa che con probabilità 0.99 l'importo complessivo delle vincite distribuite relativamente a 10 000 giocate non sarà superiore a 8583.07 € e quindi il guadagno del barista sarà almeno (€)

In[15]:=
```
10 000 - 519.879793798528`
```

Out[15]=
```
9480.12
```