Esercitazione 1 del corso di Statistica 2

Prof. Domenico Vistocco

Dott.ssa Paola Costantini

Esercizio n. 1

Estraendo due carte da un mazzo di carte napoletane con la reimmissione della carta nel mazzo dopo ciascuna prova (estrazione con ripetizione, gli eventi *sono indipendenti* in quanto, di prova in prova, il mazzo resta immutato), calcolare la probabilità che si presentino, nell'ordine,

1. Asso (*A*) e figura (*F*): $P(A \cap F) = P(A) \cdot P(F)$

2. Una carta di coppe (C) e una carta di denari (D): $P(C \cap D) = P(C) \cdot P(D)$

3. Un asso, una figura, un cinque ("5"): $P(A \cap F \cap "5") = P(A) \cdot P(F) \cdot P("5")$

Estraendo invece due carte dal mazzo, senza rimettere la prima carta estratta (estrazione senza ripetizione, *il secondo evento non è indipendente dal primo*), le stesse probabilità valgono:

4. Asso (*A*) e figura (*F*): $P(A \cap F) = P(A) * P(F|A)$

5. Una carta di coppe (C) e una carta di denari (D): $P(C \cap D) = P(C) \cdot P(D|C)$

6. Un asso, una figura, un cinque ("5"): $P(A \cap F \cap "5") = P(A) \cdot P(F|A) \cdot P("5"|A \cap F)$

Svolgimento

1. $P(A \cap F) = P(A) \cdot P(F) = \frac{4}{40} \cdot \frac{12}{40} = \frac{48}{1600}$

2. $P(C \cap D) = P(C) \cdot P(D) = \frac{10}{40} \cdot \frac{10}{40} = \frac{100}{1600}$

3. $P(A \cap F \cap "5") = P(A) \cdot P(F) \cdot P("5") = \frac{4}{40} \cdot \frac{12}{40} \cdot \frac{4}{40} = \frac{192}{64000}$

4. $P(A \cap F) = P(A) * P(F|A) = \frac{4}{40} \cdot \frac{12}{39} = \frac{48}{1560}$

5. $P(C \cap D) = P(C) \cdot P(D|C) = \frac{10}{40} \cdot \frac{10}{39} = \frac{100}{1560}$

6. $P(A \cap F \cap "5") = P(A) \cdot P(F|A) \cdot P("5"|A \cap F) = \frac{4}{40} \cdot \frac{12}{39} \cdot \frac{4}{38} = \frac{192}{59280}$

Esercizio n. 2

Estraendo a sorte una persona dalla popolazione degli attivi residenti in Italia al 25/10/1981 secondo la posizione nella professione e il titolo di studio (tab. 1) calcolare la probabilità

- **a.** che sia «imprenditore o professionista» e laureato;
- **b.** che sia «imprenditore o professionista» dato che è laureato;
- **c.** che sia almeno diplomato;
- **d.** che sia almeno diplomato, dato che è «lavoratore in proprio»;
- **e.** che abbia al più la licenza elementare;
- **f.** che abbia al più la licenza elementare, dato che è «altro lavoratore dipendente»;
- **g.** che sia coadiuvante;

Tab.1

Grado di istruzione / Posizione Nella professione	Analfabeti o Alfabeti senza titolo	Licenza elementare	Licenza media	Diploma	Laurea	Totale
Imprenditori o professionisti	7	147	155	190	182	681
Lavoratori in proprio	360	1.978	783	216	28	3.365
Coadiuvanti	48	300	205	65	5	623
Dirigenti e impiegati	5	471	1.594	2.530	939	5.539
Altri lavoratori dipendenti	972	5.105	3.413	528	22	10.040
Totale	1.392	8.001	6.150	3.529	1.176	20.248

Svolgimento

a. $P(I \cap L) = \frac{182}{20.248} = 0{,}009$

b. $P(I \mid L) = \dfrac{P(I \cap L)}{P(L)} = \dfrac{182}{20.248} \Big/ \dfrac{1176}{20.248} = \dfrac{182}{1176} = 0{,}155$

c. $P(D \cup L) = P(D) + P(L) = \dfrac{3529}{20248} + \dfrac{1176}{20248} = 0.174 + 0.058 = 0.232$

d. $P[(D \cup L) \ / \ LP] = \dfrac{P[(D \cup L) \cap LP]}{P(LP)} = \dfrac{216 + 28}{20.248} \Big/ \dfrac{3365}{20.248} = \dfrac{244}{3365} = 0{,}0725$

e. $P(A \cup E) = P(A) + P(E) = \dfrac{1392}{20248} + \dfrac{8001}{20248} = 0.0687 + 0.395 = 0.4638$

f. $P[(A \cup E) \ / \ AD] = \dfrac{P[(A \cup E) \cap AD]}{P(AD)} = \dfrac{972 + 5105}{20.248} \Big/ \dfrac{10040}{20248} = \dfrac{6077}{10040} = 0{,}06053$

g. $P(C) = \dfrac{623}{20248} = 0{,}0308$

Esercizio n. 3

In un'urna vi sono tre palline blu e cinque palline rosse. Si vuole determinare la probabilità che, estraendo due palline in blocco (o senza rimessa), esse

a. siano entrambe rosse;
b. sia una rossa e una blu

Svolgimento

Quando l'estrazione avviene in blocco vi è *dipendenza fra le prove* e il calcolo della probabilità di eventi definiti in funzione del risultato di due o più estrazioni richiede il ricorso alla probabilità condizionata.
Caso in cui si *prescinde dall'ordine.* Si indichi con R_1 l'evento "risulta una pallina rossa alla prima estrazione" e con R_2 l'evento "risulta una pallina rossa alla seconda estrazione"

a. $P((R_1 \cap R_2) = P(R_1) * P(R_2 | R_1) = \dfrac{5}{8} \times \dfrac{4}{7} = 0{,}357$
b. $P((B_1 \cap R_2) \cup (R_1 \cap B_2)) = P(B_1) * P(R_2 | B_1) + P(R_1) * P(B_2 | R_1) = \dfrac{3}{8} * \dfrac{5}{7} + \dfrac{5}{8} * \dfrac{3}{7} = 0.5357$

Caso in cui *interessa l'ordine.* Si vuole determinare la probabilità che, estraendo due palline in blocco (o senza rimessa), esse

a. siano entrambe rosse;
b. sia una rossa e una blu

c. estrazione di una pallina blu alla prima estrazione e di una pallina rossa alla seconda estrazione.

Svolgimento

a. $\mathrm{P}((R_1 \cap R_2) = P(R_1) * P(R_2|R_1) = \frac{5}{8} \times \frac{4}{7} = 0{,}357$

b. $P((B_1 \cap R_2) \cup (R_1 \cap B_2)) = P(B_1) * P(R_2|B_1) + P(R_1) * P(B_2|R_1) = \frac{3}{8} * \frac{5}{7} + \frac{5}{8} * \frac{3}{7} = 0.5357$

c. $P(B_1 \cap R_2) = P(B_1) * P(R_2|B_1) = \frac{3}{8} * \frac{5}{7} = 0{,}2678$

Questo stesso risultato può essere ottenuto sia utilizzando la probabilità condizionata, come nel caso precedente, sia mediante il rapporto fra casi favorevoli e casi possibili.
Se si estraggono due palline da un'urna che ne contiene otto e *si prescinde dall'ordine*, si utilizza il coefficiente binomiale; il numero di eventi elementari che possono risultare dall'esperimento è dato dal numero di combinazioni di due elementi scelti fra otto.

$$\textit{casi possibili} = \binom{8}{2} = \frac{8!}{2!(8-2)!} = \frac{8!}{2!6!} = \frac{8 \times 7}{2} = 28$$

I casi favorevoli sono dati dal numero dei combinazioni con le quali si possono scegliere due palline rosse fra le cinque disponibili

$$\textit{casi favorevoli} = \binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = \frac{5 \times 4}{2} = 10$$

Di conseguenza la probabilità desiderata, calcolata come rapporto fra casi favorevoli su casi possibili, è

$$\mathrm{P}(\text{“}\textit{due palline rosse}\text{”}) = \frac{\binom{5}{2}}{\binom{8}{2}} = \frac{10}{28} = \frac{5}{14} = 0.357$$

Per calcolare la probabilità di avere una pallina rossa e una pallina blu è opportuno osservare che i casi favorevoli sono dati dal prodotto del numero di combinazioni con le quali si può scegliere una pallina blu fra tre. Di conseguenza la probabilità di estrarre due palline di diverso colore è

$$P(\text{"}una\ pallina\ rossa\ e\ una\ blu\text{"}) = \frac{\binom{5}{1}\binom{3}{1}}{\binom{8}{2}} = \frac{5 \times 3}{28} = \frac{15}{28} = 0.5357$$

Nel caso in cui invece si è *interessati all'ordine*, la regola di conteggio da considerare è $n_1 \times n_2$, dove n_1 indica il numero di possibili esiti della prima sottoprova e n_2 il numero di esiti positivi sulla seconda sottoprova.
In questo caso i *casi possibili* risulterebbero: 8 x 7 (una qualunque delle otto palline alla prima estrazione e una qualunque delle restanti sette alla seconda sottoprova), mentre i *casi favorevoli* sono 5 x 4 (una qualunque delle cinque palline rosse alla prima estrazione e una qualunque delle restanti quattro alla seconda estrazione).

$$P(\text{"}due\ palline\ rosse\text{"}) = \frac{casifavorevoli}{casipossibili} = \frac{5 \times 4}{8 \times 7} = \frac{20}{56} = 0.357$$

Come si può osservare, il risultato in probabilità è lo stesso della regola di conteggio che prescinde dall'ordine (coefficiente binomiale), ma il numero di casi possibili e favorevoli risultano essere esattamente il doppio.

$$P(\text{"}una\ pallina\ rossa\ e\ una\ blu\text{"}) = \frac{5 \times 3}{56} + \frac{3 \times 5}{56} = 0.5357$$

Esercizio n. 4

Il 70% dei messaggi in arrivo in una casella e-mail è costituito da messaggi illegittimi, usualmente indicati con il termine "spam". Il software antispam riconosce correttamente i messaggi illegittimi con probabilità 0,97. Si possono però verificare dei falsi positivi, per i quali un messaggio legittimo è 0,05. Si vuole calcolare la probabilità che un messaggio sia legittimo dato che è stato classificato come spam. Sulla base di questa probabilità si potrà decidere se cancellare un messaggio prima di ancora di aprirlo. Si indichi con A = "il messaggio è legittimo", sicchè $\overline{A}$ è l'evento "il messaggio è illegittimo". Gli eventi A e $\overline{A}$ sono le ipotesi formulate sulla natura del messaggio. Le loro probabilità a priori sono $P(A)$= 0.3 e $P(\overline{A})$= 0.7. Sia B l'evento "il messaggio è classificato come spam"; le probabilità probative sono $P(B|A) = 0.05$ e $(B|\overline{A}) = 0.97$.

Soluzione

Applicando la formula di Bayes : $P(A|B) = \frac{P(A)P(B|A)}{P(B)}$ si ottiene la probabilità a posteriori che il messaggio sia legittimo dato che è stato classificato come spam.

$$P(A|B) = \frac{0.3 \times 0.05}{0.3 \times 0.05 + 0.7 \times 0.97} = 0.022$$

La probabilità a posteriori che il messaggio sia illegittimo dato che è stato classificato come spam è

$$P(\overline{A}|B) = 1 - P(A|B) = 1 - 0.022 = 0.978$$

Il confronto fra le probabilità delle due ipotesi, A "il messaggio è legittimo" e $\overline{A}$ "il messaggio è illegittimo" , dato B, induce ragionevolmente il destinatario a cancellare il messaggio senza aprirlo.