Esercitazione 4 del corso di Statistica 2
Prof. Domenico Vistocco

Alfonso Iodice D'Enza

May 23, 2007

1 Esercizio

Si consideri un mazzo di carte francesi di 52 carte e si supponga di stare giocando a poker. Nel poker ogni giocatore ha 5 carte.

- Si trovi la probabilità di avere una *scala semplice*: si ottiene tale punto se le 5 carte di cui si dispone sono consecutive (la scala, appunto) e di seme diverso (esempio *3 di fiori, 4 di quadri, 5 di picche, 6 di fiori, 7 di cuori*), dal momento che la scala composta da elementi dello stesso seme è una *scala reale*, il punto più alto nel poker.

- Si trovi la probabilità di avere un *full*: si ottiene tale punto se 2 delle 5 carte sono uguali tra loro e contemporaneamente le rimanenti 3 carte sono uguali tra loro (esempio *3 di fiori, 3 di quadri, 10 di picche, 10 di fiori, 10 di cuori*).

1.1 Svolgimento

La prima cosa da fare è contare il numero di possibili combinazioni di 5 carte che si possono ottenere a partire dalle 52 del mazzo. Per fare ciò si ricorre al coefficiente binomiale $\binom{n}{k}$. Il numero di 'mani' (di poker) distinte che è pertanto possibile avere è $\binom{52}{5}$.

- Si trovi la probabilità di avere una *scala semplice*.

Si considerino le sequenze che vanno dall'asso al 5; considerando i quattro possibili semi (*cuori, quadri, fiori, picche*), il numero di sequenze è 4^5. Per provare tale affermazione si richiama il teorema fondamentale del conteggio; si consideri un problema composto da k sottoproblemi, che possono essere risolti rispettivamente in $n_1, n_2, \ldots, n_k$ modi: il numero di modi in cui il problema di partenza può essere risolto è dato da $n = n_1 \times n_2 \times \ldots \times n_k$. Tornando al poker, contare il numero di possibili sequenze dall'asso al 5 equivale a contare il numero di possibili scelte per l'asso, il numero di possibili scelte per il 2. il numero di possibili scelte per il 3. etc., si scompone il problema di partenza in 5 sottoproblemi;

è agevole pensare che il numero di possibili assi nella sequenza corrisponde al numero di semi , vale a dire 4, così come il numero di possibili 2, 3, 4 e 5; dunque $n_1 = n_2 = n_3 = n_4 = n_5 = 4$. Da cui

numero di sequenze dall'asso al $5 = n_1 \times n_2 \times n_3 \times n_4 \times n_5 = 4 \times 4 \times 4 \times 4 \times 4 = 4^5$

Nel poker, quando una scala è composta da carte dello stesso seme si chiama *scala reale* (ed è il punto più alto che si possa avere tra le mani!): tuttavia, si stanno contando solo le scale semplici e quindi bisogna sottrarre le 4 possibili scale reali dal totale delle 4^5 possibili scale dall'asso al 5. Si può concludere che il numero di possibili scale semplici dall'asso al 5 è $4^5 - 4$. Per ciascun seme, le carte nel mazzo sono 13, dall'asso al 10 più le tre figure J,Q e K. Le possibili scale semplici che si possono ottere a partire dalle 13 cartei è 10. In tavola 1.1 sono riportate le scale che, indipendentemente dal seme, è possibile ottenere.

asso	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8
5	6	7	8	9
6	7	8	9	10
7	8	9	10	J
8	9	10	J	Q
9	10	J	Q	K
10	J	Q	K	*asso*

Pertanto per ottenere il numero totale di scale, alla luce di quanto trovato, bisogna il totale di scale ottenibili è $(4^5 - 4) \times 10$. La probabilità di ottenere una scala semplice è pertanto

$$P(\text{scala semplice}) = \frac{\text{numero di possibili scale semplici}}{\text{numero di possibili mani}} = \frac{(4^5 - 4) \times 10}{\binom{52}{5}} = 0.0039$$

- Si trovi la probabilità di avere un *full.*

Il full è un punto del poker in cui una carta si ripete due volte, un'altra carta si presenta tre volte. Si supponga un full composto da una coppia di assi e un tris di 8. Il numero di possibili combinazioni di una coppia di assi rispetto ai 4 semi è dato da $\binom{4}{2}$; analogamente i possibili tris di 8 sono $\binom{4}{3}$. Dunque, ripetendo il ragionamento per ciascuna delle 13 carte, si deduce che il numero di full possibili è dato da $13 * \binom{4}{2} * 12\binom{4}{3}$. Si noti che il 12 tiene conto del fatto che, se ad esempio l'asso compone la coppia, esso non potrà comporre il tris. Ricordando che il numero possibile di mani di poker è $\binom{52}{5}$, la probabilità di ottenere un full è

$$P(\text{full}) = \frac{\text{numero di possibili full}}{\text{numero di possibili mani}} = \frac{13 * \binom{4}{2} * 12\binom{4}{3}}{\binom{52}{5}} = 0.0014$$

2 Esercizio

Si consideri la prova consistente nel lancio di un dado equilibrato 200 volte. Qualè la probabilità che

- la faccia 4 si presenti fino a 28 volte
- la faccia 4 o la faccia 5 si presentano più di 60 volte e meno di 70

2.1 Svolgimento

- la faccia 4 si presenti fino a 28 volte

Per contare il numero di successi in n prove si utilizza la distribuzione binomiale. Volendo impostare il calcolo binomiale per l'evento *il numero di facce 3 ottenute in 200 lanci è minore o uguale al 28* si avrà

$$P(X \leq 28) = \binom{200}{0}\frac{1}{6}^{0}\frac{5}{6}^{200} + \binom{200}{1}\frac{1}{6}^{1}\frac{5}{6}^{199} + \ldots + \binom{200}{27}\frac{1}{6}^{27}\frac{5}{6}^{173} + \binom{200}{28}\frac{1}{6}^{28}\frac{5}{6}^{172}$$

Tuttavia, a meno di non avere a disposizione un computer, tale calcolo non è agevole. In questo caso si può utilizzare l'approssimazione della binomiale alla normle. In particolare, si calcolano media e varianza della binomiale, questi saranno i parametri della normale.

$$E(X) = \mu = np = 200 \times \frac{1}{6} = 33.3 \ , \ Var(X) = \sigma^2 = np(1-p) = 200 \times \frac{1}{6} \times \frac{5}{6} = 27.7$$

da cui $\sigma = \sqrt{27.7} = 5.26$. Bisogna considerare che si sta approssimando una distribuzione discreta con una continua, pertanto l'intervallo di valori da considerare è $[-0.5, 28]$. Bisogna dunque trasformare gli estremi dell'intervallo di riferimento in unità standard. In particolare

$$z_1 = \frac{-0.5 - 33.3}{5.26} = -6.42 \ , \ z_2 = \frac{28 - 33.3}{5.26} = -0.95$$

Dalle tavole si ha che

$$P(Z \leq -6.42) \approx 0 \ , \ P(Z \leq -0.95) = 1 - P(Z \leq 0.95) = 0.17$$

Quindi $P(X \leq 28) = 0.17$

- la faccia 4 o la faccia 5 si presentano più di 60 volte e meno di 70

Anche in questo si può approssimare la distribuzione binomiale di parametri $(n = 200), p = \frac{2}{6}$ ad una normale di parametri

$$E(X) = \mu = np = 200 \times \frac{2}{6} = 66.6 \ , \ Var(X) = \sigma^2 = np(1-p) = 200 \times \frac{2}{6} \times \frac{4}{6} = 44.4$$

da cui $\sigma = \sqrt{44.4} = 6.66$. Si passa ai valori standardizzati

$$z_1 = \frac{60 - 66.6}{6.66} = -1 \ , \ z_2 = \frac{70 - 66.6}{66.6} = 0.5$$

Dalle tavole si ha che

$P(Z \leq 1) = 0.84$, $P(Z \leq 0.5) = 0.69$ da cui $P(60 \leq X \leq 70) = 0.69+0.84-1 = 0.53$

3 Esercizio

Un aereo da turismo ha perso i contatti con la torre di controllo: le ricerche sono circoscritte a tre diverse zone, ognuna delle quali caratterizzata dalla stessa probabilità che l'aereo venga ritrovato. L'evento R_i l'*aereo viene ritrovato nella zona* i, per $i = 1, 2, 3$, ha probabilità $P(R_i) = \frac{1}{3}$. La probabilità che le ricerche nella zona i abbiano successo è data da $1 - \beta_i$.
Si consideri ancora l'evento E per il quale *le ricerche nella zona 1 non hanno prodotto risultati*, Come vengono aggiornate le probabilità di ritrovamento dell'aereo nelle tre zone?

3.1 Svolgimento

Si vuole conoscere dunque quale sia la probabilità di avere l'evento R_i posto che si sia verificato E. In altre parole, si vuole aggiornare la probabilità di trovare l'aereo nell'i-ma zona, posto che le ricerche in zona 1 non abbiano prodotto risultati. Dalla definizione di probabilità condizionata si ha

$$\begin{aligned} P(R_1 \mid E) = \frac{P(R_1 \cap E)}{P(E)} &= \frac{P(E \mid R_1) \times P(R_1)}{\sum_{i=1}^{3} P(E \mid R_i) \times P(R_i)} = \\ &= \frac{\beta_1 \times \frac{1}{3}}{\beta_1 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3}} = \frac{\beta_1}{\beta_1 + 2} \end{aligned}$$

Ripetendo il ragionamento per gli eventi R_2 ed R_3 si ha

$$\begin{aligned} P(R_2 \mid E) = \frac{P(R_2 \cap E)}{P(E)} &= \frac{P(E \mid R_2) \times P(R_2)}{\sum_{i=1}^{3} P(E \mid R_i) \times P(R_i)} = \\ &= \frac{1 \times \frac{1}{3}}{\beta_1 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3}} = \frac{1}{\beta_1 + 2} \end{aligned}$$

$$\begin{aligned} P(R_2 \mid E) = \frac{P(R_3 \cap E)}{P(E)} &= \frac{P(E \mid R_3) \times P(R_3)}{\sum_{i=1}^{3} P(E \mid R_i) \times P(R_i)} = \\ &= \frac{1 \times \frac{1}{3}}{\beta_1 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3}} = \frac{1}{\beta_1 + 2} \end{aligned}$$

Notare che la probabilità di non trovare l'aereo nella zona 1 (E) posto che l'aereo si trovi nella zona 2 o nella zona 3 corrisponde all'evento certo, vale a dire $P(E \mid R_2) = P(E \mid R_3) = 1$.

4 Esercizio

Una classe di studenti di un corso di statistica dell'università di Oxford ha sostenuto una prova di statistica, valutata in decimi (voti da uno a dieci). I voti riportati dagli studenti si distribuiscono secondo una v.c. normale con media $\mu = 6.7$ e varianza $\sigma^2 = 1.2^2$. Sapendo che il voto finale é ottenuto arrotondando il punteggio riportato al test (es. $5.6 = 6$, $7.4 = 7$), determinare:

- la percentuale di studenti che hanno ottenuto il voto 6
- il voto migliore del peggior 10% degli studenti
- il voto peggiore del migliore 20% degli studenti

4.1 Svolgimento

- la percentuale di studenti che hanno ottenuto il voto 6

L'intervallo di punteggi riportati al test che producono il voto 6 è $[5.5, 6.5]$, pertanto la percentuale di studenti che hanno ottenuto un voto sei si ottiene integrando la f. di densità normale in tale intervallo. Occorre dunque standardizzare il problema:

$$z_1 = \frac{5.5 - 6.7}{1.2} = -1 \ , \ z_2 = \frac{6.5 - 6.7}{1.2} = -0.17$$

Sfruttando la simmetria della curva normale, la percentuale cercata è data da

$$P(z \leq 1) - P(z \leq 0.17) = 0.3413 - 0.0675 = 0.2738$$

- il voto migliore del peggior 10% degli studenti

Per trovare il voto migliore del peggiore 10% si consideri che il valore di z corrispondente al decimo percentile è uguale al valore di z corrispondente al novantesimo percentile moltiplicato per (-1). Dalle tavole risulta che il novantesimo percentile corrisponde a $z = 1.285$, quindi il decimo percentile è dato da $z_{10} = -1.285$. Il voto corrispondente al valore di si ottiene a partire dalla formula della standardizzazione.

$$x = \sigma z + \mu = 1.2 * (-1.285) + 6.7 = 5.16 \text{ arrotondando, } 5$$

- il voto peggiore del migliore 20% degli studenti

Il valore ricercato corrisponde in questo caso all'ottantesimo percentile. Sulle tavole va ricercato z cui corrisponde il tabulato pari a 0.8. In particolare $z = 0.845$. Quindi il valore di X cercato è

$$x = \sigma z + \mu = 1.2 * (0.845) + 6.7 = 7.71 \text{ arrotondando, } 8$$