1 Distribuzione di Probabilità Ipergeometrica

Sia data un' urna contenente b biglie bianche ed r biglie rosse.
Si effettuano n estrazioni senza reimmissione. Calcolare la probabilità di estrarre esattamente k biglie rosse. Questo problema ha una soluzione non banale se

$$\begin{cases} n \leq b + r \\ k \leq r \\ k \leq n \\ n - k \leq b \end{cases} \quad \text{ovvero se} \quad \begin{cases} n \leq b + r \\ \max\{0, n - b\} \leq k \leq \min\{r, n\} \end{cases}$$

Per risolvere il problema posto costruiamo uno spazio di probabilità opportuno; a tale proposito associamo un numero naturale ad ogni pallina nell' urna in modo tale che contenga $r + b$ elementi distinti; per semplicità le palline rosse saranno identificate dai primi r numeri naturali e le palline bianche dai numeri da $r+1, r+2, ..., r+b$. Consideriamo dunque lo spazio campionario

$$\Omega = \{\omega = (\omega_1, ..., \omega_n) : \ \omega_i \neq \omega_j \text{ se } i \neq j, \ \omega_i \in \{1, 2, ..., r + b\}\}.$$

Per contare i punti di Ω possiamo procedere in due modi;

1 si possono considerare le sequenze ordinate di n elementi distinti estratti dall' insieme $\{1, 2, ..., r + b\}$ ed in tal caso $\#\Omega = \#D_n^{b+r}$;

2 si possono considerare i sottoinsiemi di cardinalità n di $\{1, 2, ..., r+b\}$ ed in tal caso $\#\Omega = \#C_n^{b+r}$.

Poiché il problema in esame *non dipende dall' ordine di estrazione*, possiamo scegliere come spazio campionario l' insieme dei sottoinsiemi di cardinalità n di $\{1, 2, ..., r + b\}$. In tal caso i punti di tale spazio possono essere messi in corrispondenza biunivoca con le combinazioni di $b + r$ elementi di classe n, C_n^{b+r}.
Quindi $\#\Omega = \#C_n^{b+r} = \binom{b+r}{n}$. Su questo spazio mettiamo la distribuzione di probabilità uniforme. Sia

$$A_k = \{\omega \in \Omega \text{ tale che } \omega \text{ contiene esattamente k biglie rosse. }\}$$

Questo evento si può anche scrivere come

$$A_k = \{\omega \in \Omega \text{ tale che } k \text{ elementi di } \omega \text{ appartengono all'insieme } \{1, 2, ..., r\}\}.$$

Dobbiamo calcolare $\#A_k$. Ogni punto di A_k contiene k palline rosse ed $n-k$ bianche. Come si conta l'insieme di punti che ha questa caratteristica? Osserviamo che ci sono $\#C^b_{n-k} = \binom{b}{n-k}$ punti di Ω che contengono un fissato sottoinsieme di cardinalità k di $\{1, 2, ..., r\}$. Infatti per ogni tale sottoinsieme esistono $\#C^b_{n-k} = \binom{b}{n-k}$ sottoinsiemi di cardinalità n di $\{1, 2, ..., r + b\}$ ottenuti associando al sottoinseme di palline rosse fissato un sottoinsieme di cardinalità $n-k$ di $\{r + 1, ..., r + b\}$.
Ricapitolando,

ad ogni fissato sottoinsieme di cardinalità k ***di*** $\{1, 2, ..., r\}$ ***si associano*** $\binom{b}{n-k}$ ***sottoinsiemi di cardinalità*** n ***di*** $\{1, 2, ..., r + b\}$.

Ma i sottoinsiemi di cardinalità k di $\{1, 2, ..., r\}$ sono $\#C^r_k = \binom{r}{k}$, e quindi

$$\#A_k = \binom{r}{k}\binom{b}{n-k}.$$

Concludiamo quindi che, per ogni $k \in \{\max\{0, n-b\}, ..., \min\{r, n\}\}$

$$P(A_k) = \frac{\#A_k}{\#\Omega} = \frac{\binom{r}{k}\binom{b}{n-k}}{\binom{b+r}{n}}. \tag{1}$$

La formula (1) definisce una distribuzione di probabilità sullo spazio $\Omega = \{\max\{0, n-b\}, ..., \min\{r, n\}\}$ detta DISTRIBUZIONE IPERGEOMETRICA.

Esempio 1.1. *Calcolare la probabilità di fare terno al gioco del lotto giocando i numeri* $\{3, 4, 22\}$.

Raggruppiamo le 90 palline nell' urna nei due gruppi disgiunti $\{3, 4, 22\}$, $\{1, 2, 5, ..., 21, 23, ..., 90\}$. L' evento $A = \{$ escono i numeri 3, 4, 22$\}$ coincide con l' evento $\{$ su 5 estrazioni si estraggono 3 elementi dal gruppo $\{3, 4, 22\}\}$

Quindi $P(A)$ si calcola utilizzando la distribuzione Ipergeometrica con $r = 3, b = 90-3, n = 5, k = 3$. In tal modo si ottiene

$$P(A) = \frac{\binom{3}{3}\binom{87}{2}}{\binom{90}{5}} = \frac{\frac{87!}{85!2!}}{\frac{90!}{85!5!}} = ...$$

Esercizio 1.1. *(Baldi 1.42, 2.42)*
Un gruppo di 10 turisti ha prenotato un volo con 120 passeggeri. Al momento dell' imbarco viene comunicato che sono state accettate troppe prenotazioni e che 10 passeggeri verranno estratti a sorte e spostati sul volo successivo.

1 *Qual è la probabilità che i 10 amici riascano a partire subito?*

2 *Qual è la probabilità che i 10 amici partano tutti con il volo successivo?*

Dividiamo i 120 passeggeri in due gruppi disgiunti, il primo composto dai 10 amici ed il secondo dai restanti 110 passeggeri.

1 L' evento $A = \{$ i 10 amici partono subito$\}$ coincide con l' evento $\{$ su 10 estrazioni casuali 10 elementi si estraggono dal primo gruppo$\}$ Quindi $P(A)$ si calcola utilizzando la distribuzione Ipergeometrica con $r = 10, b = 110, n = 10, k = 10$. In tal modo si ottiene

$$P(A) = \frac{\binom{10}{10}\binom{110}{0}}{\binom{120}{10}} = \frac{1}{\frac{120!}{110!10!}} = ...$$

2 L' evento $B = \{$ i 10 amici partono con il volo successivo$\}$ coincide con l' evento
$\{$ su 10 estrazioni casuali 10 elementi si estraggono dal secondo gruppo$\}$ Analogamente al primo punto $P(B)$ si calcola utilizzando la distribuzione Ipergeometrica con $r = 10, b = 110, n = 10, k = 0$:

$$P(B) = \frac{\binom{10}{0}\binom{110}{10}}{\binom{120}{10}} = \frac{\frac{110!}{100!10!}}{\frac{120!}{110!10!}} = ...$$

Per casa: Esercizio 1.32 Abundo.

2 Distribuzione di Probabilità Binomiale

Su uno spazio di probabilità (Ω, A, P) siano dati n eventi indipendenti ed equiprobabili $A_1, ..., A_n$. Indichiamo con $p = P(A_i),\ i = 1, ..., n$. Si vuole calcolare la probabilità dell' evento

$C_k = \{$ si verificano esattamente k eventi tra gli $A_1, ..., A_n\},\ k \in \{0, ..., n\}$.

Facciamo le seguenti osservazioni

1 Sia $i : \{1, ..., n\} \to \{1, ..., n\}$ una qualunque permutazione di $\{1, ..., n\}$. Consideriamo l' evento

$$B^i = A_{i(1)} \cap ... \cap A_{i(k)} \cap A^c_{i(k+1)} \cap ... \cap A^c_{i(n)}.$$

Allora $B^i \subseteq C_k$;

2 Se $\mathbb{1} : \{1, ..., n\} \to \{1, ..., n\}$ è la permutazione identica, allora

$$B^{\mathbb{1}} = A_1 \cap ... \cap A_k \cap A^c_{k+1} \cap ... \cap A^c_n$$

e per le assunzioni di indipendenza ed equiprobabilità si ha

$$P(B^{\mathbb{1}}) = p^k(1-p)^{n-k};$$

Inoltre questa è anche la probabilità di ogni evento B^i, qualunque sia la permutazione considerata.

3 Chiaramente C^k è unione di eventi come B^i, ma quanti sono al variare di tutte le permutazioni di $\{1, ..., n\}$ i B^i disgiunti? Osserviamo che ci sono delle permutazioni che lasciano invariati questi insiemi; ad esempio, se considero la permutazione che lascia fissi tutti gli indici tranni i primi 2, ovvero $i(1) = 2, i(2) = 1$ e $k \geq 3$ allora $B^{\mathbb{1}} = B^i$.

4 Ogni B^i è identificato da un sottoinsieme di indici di cardinalità k di $\{1, ..., n\}$ che rappresenta gli indici corrispondenti ai k eventi che si verificano. Per chiarezza, per l' evento $B^{\mathbb{1}}$ tale sottoinsieme è $\{1, ..., k\}$; in generale per per l' evento B^i tale sottoinsieme è $\{i(1), ..., i(k)\}$.

5 A sottoinsiemi di cardinalità k di $\{1, ..., n\}$ diversi corrispondono B^i disgiunti.
In sostanza l' insieme di tutti i B^i distinti è in corrispondenza biunivoca con le combinazioni di n elementi di classe k e quindi tutti i B^i disgiunti che compongono C^k sono pari a $\#C^n_k = \binom{n}{k}$ e sono tutti tra loro equiprobababili. Di conseguenza

$$P(A) = \binom{n}{k} p^k(1-p)^{n-k}. \tag{2}$$

La formula (2) definisce una distribuzione di probabilità sullo spazio $\Omega = \{0, ..., n\}$ che dipende da due parametri, n **e** p detta Distribuzione Binomiale di parametri n, p.

La distribuzione binomiale è ricavata nell' ambito di uno modello centrale nel calcolo delle probabilità, noto con il nome di Schema di Bernoulli o delle prove ripetute (o ancora schema successo-insuccesso). In questo modello si prende in considerazione una sequenza (talvolta infinita) di eventi (o prove) indipendenti e tutti con la stessa probabilità p. Esempi di Schema di Bernoulli sono

- lanci successivi di un dado o di una moneta;
- Estrazioni con reimmissione da un' urna che contiene due soli tipi di oggetti;
- In generale sequenze di esperimenti indipendenti e binari, in cui ad ogni prova si è interessati al verificarsi o meno di un evento (successo) o del suo complementare (insuccesso).

Esercizio 2.1. *(Baldi, 2.2, 1.3)*
Da un' urna contenente 4 palline bianche e 3 nere si effettuano 2 estrazioni con rimpiazzo. Calcolare

1 *la probabilità che le due palline estratte siano dello stesso colore;*

2 *la probabilità che almeno una delle due palline estratte sia nera.*

Consideriamo gli eventi $A_i = \{\text{lı-sima pallina estratta è nera}\}, i = 1, 2$.

1 L'evento $B = \{\text{le due palline estratte sono dello stesso colore}\}$, si esprime come
$$B = (A_1 \cap A_2) \cup (A_1^c \cap A_2^c).$$
Pertanto, essendo gli eventi $A_1 \cap A_2$ e $A_1^c \cap A_2^c$ disgiunti,
$$P(B) = P(A_1 \cap A_2) + P(A_1^c \cap A_2^c).$$
Essendo le estrazioni con ripiazzo, gli eventi A_1 ed A_2 sono indipendenti (DIMOSTRARE), ed inoltre $P(A_1) = P(A_2) = \frac{3}{7}$. Quindi posso utilizzare la distribuzione di probabilità binomiale per calcolare entrambi gli addendi. Precisamente il primo addendo corrisponde alla distribuzione binomiale di parametri $n = 2, p = \frac{3}{7}$ calcolata per $k = 2$ (escono 2 palline nere) mentre il secondo addendo alla stessa distribuzione binomiale calcolata per $k = 0$ (escono 2 palline bianche):
$$\begin{aligned} P(B) =& \binom{2}{2} \left(\frac{3}{7}\right)^2 \left(1 - \frac{3}{7}\right)^{2-2} + \binom{2}{0} \left(\frac{3}{7}\right)^{2-2} \left(1 - \frac{3}{7}\right)^2 \\ =& \frac{9}{49} + \frac{16}{49} = \frac{25}{49}. \end{aligned}$$

2 L'evento $C = \{\text{almeno una delle due palline estratte sia nera}\}$, si esprime come

$C =$ esattamente una delle palline estratte è nera$\}\cup\{$entrambe le palline estratte

Gli eventi in cui si decompone C sono disgiunti ed anche in questo caso si utilizza la distribuzione di binomiale di parametri $n = 2, p = \frac{3}{7}$ per calcolare entrambi gli addendi, il primo essendo il valore di tale distribuzione per $k = 1$ (esce esattamente 1 pallina nera) mentre il secondo per $k = 2$ (escono 2 palline nere).
Alternativamente, poiché $C^c = A_1^c \cap A_2^c$, si ottiene più agevolmente

$$P(C) = 1 - P(A_1^c \cap A_2^c) = 1 - \left(\frac{4}{7}\right)^2 = \frac{33}{49}.$$