1 Distribuzione di probabilità Ipergeometrica

Da un'urna contenente b palline bianche ed r palline rosse si effettuano n estrazioni senza rimpiazzo. Calcolare la probabilità di estrarre esattamente k palline rosse.

Affinché questo esercizio abbia una soluzione non banale, restringiamo la nostra attenzione al caso in cui siano soddisfatti i seguenti vincoli naturali:

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

Descriviamo uno spazio di probabilità opportuno.

$$\Omega = \{\omega = (\omega_1, ..., \omega_n) : \ \omega_i \neq \omega_j \text{ per } \ i \neq j, \ \omega_i \in \{r_1, .., r_r, b_{r+1}, ..., b_{r+b}\}\}$$

A questo punto, poiché l'evento di cui dobbiamo calcolare la probabilità *non dipende dall'ordine di estrazione delle palline*, possiamo "contare" gli elementi di Ω in due modi:

1 I punti di Ω possono essere messi in corrispondenza biunivoca con le sequenze ordinate di lunghezza n di $b+r$ elementi distinti tra loro; in questo caso tale insieme risulta essere in corrispondenza biunivoca con l'insieme delle disposizioni di $b+r$ elementi di classe n; ovvero

$$\#\Omega = \#D_n^{b+r} = \frac{(b+r)!}{(b+r-n)!}$$

2 I punti di Ω possono essere messi in corrispondenza biunivoca con i sottoinsiemi di cardinalità n di $b+r$ elementi; in questo caso tale insieme risulta essere in corrispondenza biunivoca con l'insieme delle combinazioni di $b+r$ elementi di classe n; ovvero

$$\#\Omega = \#C_n^{b+r} = \frac{(b+r)!}{n!(b+r-n)!} = \binom{b+r}{k}$$

In entrambi i casi, aver distinto tutte le palline nell'urna (avendole numerate e non raggruppate per colore) mi permette di utilizzare la distribuzione di probabilità uniforme.

Utilizziamo il primo spazio di probabilità descritto, ovvero quello in corrispondenza biunivoca con D_n^{b+r}.

Consideriamo l'evento

$$A_k^{(1,...,k)} = \{ \ k \text{ palline rosse nelle prime } k \text{ estrazioni, } n\text{-}k \text{ palline bianche nelle successive estrazioni}\}$$

Per calcolare la probabilità di tale evento dobbiamo contare quante sono le disposizioni che hanno k palline rosse nelle prime k componenti e n-k palline bianche nelle successive n-k componenti.
A tale proposito il modo più semplice di procedere è chiedersi, a partire dalla prima componente, in quanti modi diversi la possiamo scegliere. Evidentemente la prima componente potrà essere scelta in r differenti modi (il numero totale delle palline rosse), la seconda in *r-1* differenti modi (il numero totale delle palline rosse rimaste dopo la scelta della prima rossa) e così via, fino ad arrivare alla componente k-sima, che potrà essere scelta in *r-k+1* differenti modi. A partire dalla componente $k+1$-sima si ragiona allo stesso modo considerando che adesso si devono scegliere le palline bianche, e quindi la $k+1$-sima componente potrà essere scelta in b differenti modi (il numero totale delle palline bianche), la seconda in *b-1* differenti modi e così via, fino ad arrivare alla componente n-sima, che potrà essere scelta in *b-(n-k)+1* modi diversi. Ovvero:

$$\#A_k^{(1,...,k)} = r(r-1)(r-2)\cdots(r-k+1)b(b-1)\cdots(b-(n-k)+1) = \frac{r!}{(r-k)!}\frac{b!}{(b-(n-k))!}$$

di conseguenza

$$P(A_k^{(1,\dots,k)}) = \frac{\frac{r!}{(r-k)!}\frac{b!}{(b-(n-k))!}}{\frac{(b+r)!}{(b+r-n)!}}$$

Ma qual è la probabilità dell'evento $A_k = \{$ escono esattamente k palline rosse in n estrazioni$\}$? Sicuramente $A_k^{(1,\dots,k)} \subset A_k$. Inoltre

$$A_k = \bigcup_{i_1,\dots,i_k} A_k^{(i_1,\dots,i_k)}$$

essendo

$$A_k^{(i_1,\dots,i_k)} = \{\ k \text{ palline rosse nelle posizioni individuate dagli indici } i_1,\dots,i_k\}$$

Tali eventi sono *disgiunti*, *equiprobabili* ed in *numero pari a* $\binom{n}{k}$, ovvero in numero pari a tutti i modi di poter allocare k palline rosse su n posizioni (ogni k-pla di indici corrisponde ad un sottoinsieme di cardinalità k di un insieme di n indici....). Pertanto

$$P(A_k) = \binom{n}{k}\frac{\frac{r!}{(r-k)!}\frac{b!}{(b-(n-k))!}}{\frac{(b+r)!}{(b+r-n)!}} = \frac{\binom{r}{k}\binom{b}{n-k}}{\binom{b+r}{n}}$$

Al variare di k nell'insieme descritto all'inizio del paragrafo, si ottiene una distribuzione di probabilità nota come *distribuzione di probabilità ipergeometrica.*

2 Distribuzione di probabilità Binomiale

Su uno spazio di probabilità $(\Omega, \mathcal{A}, P)$ siano dati n eventi $A_1, \dots, A_n$ indipendenti ed equiprobabili. Indichiamo co $p = P(A_1) = \dots = P(A_n)$. Evidentemente $1 - p = P(A_1^c) = \dots = P(A_n^c)$. Calcoliamo la probabilità dell'evento $A = \{$ si verificano esattamente k eventi tra $A_1, \dots, A_n\}$, $k = 0, \dots, n$.
Osserviamo che

1 Per ogni $i_1, \dots, i_n$ permutazione di $1, \dots, n$ l'evento $B^i = \{A_{i_1} \cap \dots \cap A_{i_k} \cap A^c_{i_{k+1}} \cap \dots \cap A^c_{i_n}\} \subseteq A$;

2 In particolare $B^1 = \{A_1 \cap \dots \cap A_k \cap A^c_{k+1} \cap \dots \cap A^c_n\} \subseteq A$;

3 Di conseguenza, quale che sia la permutazione i in oggetto, gli eventi B^i hanno probabilità pari a $P(B^i) = P(A_{i_1} \cap \dots \cap A_{i_k} \cap A^c_{i_{k+1}} \cap \dots \cap A^c_{i_n}) = P(A_{i_1}) \cdots P(A_{i_k}) P(A^c_{i_{k+1}}) \cdots P(A^c_{i_n}) = p^k(1-p)^{n-k}$;

4 $A = \bigcup_i B^i$ dove l'unione è estesa a tutte le possibili permutazioni di $\{1, \dots, n\}$. Quante sono le permutazioni che generano eventi B^i disgiunti tra loro? Prendiamo ad esempio l'evento B^1. Ci sono permutazioni che lo lasciano invariato, e sono quelle che permutano i primi k elementi tra loro e i restanti $n - k$ elementi tra loro senza mischiare i due gruppi. Queste permutazioni sono $k!(n-k)!$. Ovvero ogni evento di tipo B^i è mutato in se stesso da $k!(n-k)!$ permutazioni. Quindi il numero di eventi disgiunti che compongono A è uguale al numero di tutte le permutazioni di $\{1, \dots, n\}$ diviso $k!(n-k)!$, ovvero è $\frac{n!}{k!(n-k)!} = \binom{n}{k}$ Pertanto:

$$P(A) = \#\{B^i \text{ distinti}\}P(B^i) = \#\{B^i \text{ distinti}\}p^k(1-p)^{n-k} = \binom{n}{k}p^k(1-p)^{n-k}, \quad k = 0, \dots, n.$$

Al variare di $k = 0, \dots, n$ si ottiene una distribuzione di probabilità nota come *distribuzione di probabilità binomiale di parametri n e p.*