

DISTRIBUZIONE MULTINOMIALE

Consideriamo n prove indipendenti.

Per ogni prova abbiamo r risultati possibili

$\left\{ \begin{array}{ll} \text{RISULTATO } 1 & \text{con probabilità } p_1 \geq 0 \\ \vdots & \\ \text{RISULTATO } r & \text{con probabilità } p_r \geq 0 \end{array} \right.$

I valori p_1, \dots, p_r sono gli stessi per ogni prova; inoltre $p_1 + \dots + p_r = 1$.

In generale saremo interessati a calcolare probabilità di questo tipo:

$$P \left(\begin{array}{l} k_1 \text{ volte risultato 1} \\ \vdots \\ k_r \text{ volte risultato } r \end{array} \right) \text{ dove} \left\{ \begin{array}{l} k_1, \dots, k_r \geq 0 \text{ inten.} \\ k_1 + \dots + k_r = n \end{array} \right.$$

OSSERVAZIONE

Se si ha $r=2$ si ricupera il caso delle Bernoulli; ad esempio
RISULTATO 1 \longleftrightarrow SUCCESSO ($p_1 = p$)
RISULTATO 2 \longleftrightarrow FALLIMENTO ($p_2 = 1-p$)

(perché $p_1 + p_2 = 1$)

Consideriamo la seguente sequenza di risultati:

$$\underbrace{(e_1, \dots, R_1)}_{k_1 \text{ volte}}, \underbrace{(R_2, \dots, R_2)}_{k_2 \text{ volte}}, \dots, \underbrace{(R_r, \dots, R_r)}_{k_r \text{ volte}}).$$

Allora, per indipendenza delle prove, queste sequenze ha probabilità $p_1^{k_1} \cdots p_r^{k_r}$.

Ovviamente si ottiene lo stesso probabilità per qualsiasi altra sequenze con
" k_1 volte R_1, \dots, k_r volte R_r ". Quindi la grandezza che vogliamo calcolare è

$$P \left(\begin{array}{c} k_1 \text{ volte } R_1 \\ \vdots \\ k_r \text{ volte } R_r \end{array} \right) = \# \left\{ \text{sequenze con } "k_1 \text{ volte } R_1, \dots, k_r \text{ volte } R_r" \right\} \cdot p_1^{k_1} \cdots p_r^{k_r}$$

Si può verificare che

$$\frac{m!}{k_1! \cdots k_r!}$$

è detto COEFFICIENTE MULTINOMIALE
(per $r=2$ è un COEFFICIENTE BINOMIALE).

OSSERVAZIONE

Se in ogni prova i risultati R_1, \dots, R_r sono equiprobabili, cioè $P_1 = \dots = P_r = \frac{1}{r}$, si ha

$$P\left(\begin{array}{c} k_1 \text{ volte } R_1 \\ \vdots \\ k_r \text{ volte } R_r \end{array}\right) = \frac{m!}{k_1! \dots k_r!} \left(\frac{1}{r}\right)^{k_1 + \dots + k_r} =$$

$$= \frac{m!}{k_1! \dots k_r!} \left(\frac{1}{r}\right)^m$$

OSS Analogamente quanto visto per $X \sim \text{Bin}(n, p=1/2)$ dove si ha $P_X(k) = \binom{n}{k} \left(\frac{1}{2}\right)^n$ per $k=0, 1, \dots, n$

OSSERVAZIONE

Per $r=2$

$$P\left(\begin{array}{c} k_1 \text{ volte } R_1 \\ k_2 \text{ volte } R_2 \end{array}\right) = \frac{m!}{k_1! (m-k_1)!} P_1^{k_1} (1-P_1)^{m-k_1} = \binom{m}{k_1} P_1^{k_1} (1-P_1)^{m-k_1}$$

(con $0 \leq k_1 \leq m$).

$P_1 = 1 - P_2$ perché $P_1 + P_2 = 1$

$k_2 = m - k_1$ perché $k_1 + k_2 = m$

Si ricopre la BINOMIALE

ESERCIZIO

Un'urna contiene 3 palline bianche, 3 rosse e 2 nere. Si estraggono 4 palline a caso, una alla volta e con rimessaggio. Calcolare le probabilità dei seguenti eventi:

- 1) Viene estratta la sequenza di colori: (rosse, nere, nere, bianche).
- 2) Vengono estratte esattamente 2 palline rosse e 1 nera in qualsiasi ordine.
- 3) Vengono estratte esattamente 2 palline rosse in un qualsiasi ordine.

COMMENTO Negli esercizi generalmente s'intenderemo "esattamente" e "in qualsiasi ordine".

SVOLGIMENTO

Si hanno prove indipendenti perché le estensioni sono con reinamento.

In ogni prova abbiamo 3 risultati con probabilità $P_B = \frac{3}{8}$, $P_R = \frac{3}{8}$, $P_N = \frac{2}{8}$ (uso notazioni diverse da p_1, p_2, p_3)

1) Con notazioni: ovvie le probabilità richiesta è

$$P(R_1 \cap N_2 \cap N_3 \cap B_4) = P(R_1)P(N_2)P(N_3)P(B_4) = P_R P_N P_N P_B = P_B P_R P_N^2 = \boxed{\frac{3}{8} \cdot \frac{3}{8} \left(\frac{2}{8}\right)^2} = \frac{36}{4096}$$

OSS

L'espressione $P_B P_R P_N^2$ può essere vista come $P_1^{k_1} P_2^{k_2} P_3^{k_3}$ con $k_1=1$, $k_2=1$, $k_3=2$ $((1B, 1R, 2N)$ in un ordine fissato). Abbiamo una parte delle formule della multinomiale senza il coefficiente multinomiale (del resto è una sequenza fissata).

2) $P("2R \ e \ 1N" \ \text{in un qualsiasi ordine}) \Downarrow$ 4 palline estratte

$$= P("1B, 2R, 1N" \ \text{in qualsiasi ordine}) = \frac{4!}{1! 2! 1!} \left(\frac{3}{8}\right)^1 \left(\frac{3}{8}\right)^2 \left(\frac{2}{8}\right)^1 = \frac{648}{4096}$$

$\rightarrow = \frac{4 \cdot 3 \cdot 2}{2} = 12$

3) $P("2R" \ \text{in un qualsiasi ordine}) \Downarrow$ 4 palline estratte, contiene solo rosse e non rosse
(senza distinguere tra bianche e nere)

$$\Downarrow \binom{4}{2} \left(\frac{3}{8}\right)^2 \left(1 - \frac{3}{8}\right)^{4-2} = 6 \left(\frac{3}{8}\right)^2 \left(\frac{5}{8}\right)^2 = \frac{1350}{4096}$$

\exists #ROSSE estratte \sim BIN ($n=4, p = p_R = \frac{3}{8}$)

OSS. Se volessimo tenere conto dei 3 colori si ha

$$P("2R") = P("2R, 1B, 1N") + P("2R, 2B") + P("2R, 2N") = \frac{648 + 486 + 216}{4096} = \frac{1350}{4096}$$

\Downarrow

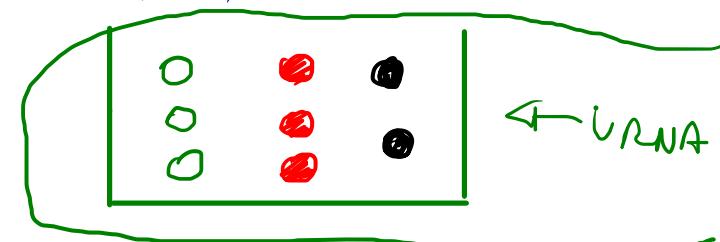
$$= \frac{4!}{2! 2! 0!} \left(\frac{3}{8}\right)^2 \left(\frac{3}{8}\right)^2 \left(\frac{1}{8}\right)^0 = \frac{4!}{0! 2! 2!} \left(\frac{3}{8}\right)^0 \left(\frac{3}{8}\right)^2 \left(\frac{2}{8}\right)^2$$

\hookrightarrow vedere risposte a domande precedente

Colevoli più
Complicati
Metodo meno
conveniente

ESERCIZIO PRECEDENTE CON ESTRAZIONI "SENZA REINSEGUIMENTO" (ANZICHE "CON REINSEGUIMENTO")

$$1) P(R_1 \cap N_2 \cap N_3 \cap B_4) = \underbrace{P(R_1)}_{= \frac{3}{8}} \underbrace{P(N_2 | R_1)}_{= \frac{2}{7}} \underbrace{P(N_3 | R_1 \cap N_2)}_{= \frac{1}{6}} \underbrace{P(B_4 | R_1 \cap N_2 \cap N_3)}_{= \frac{3}{5}} = \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{1}{6} \cdot \frac{3}{5} = \frac{3}{280}$$



$$2) P("2R \text{ e } 1N") = P("1B, 2R, 1N") =$$

$$= \frac{\binom{3}{1} \binom{3}{2} \binom{2}{1}}{\binom{8}{4}} = \frac{3 \cdot 3 \cdot 2}{70} = \frac{18}{70} = \frac{9}{35}$$

R
S.
B.
ROSSE E NON ROSSE
Senza distinguere
tra bianche e nere

$$3) P("2R") = \frac{\binom{3}{2} \binom{5}{2}}{\binom{8}{4}} = \frac{3 \cdot 10}{70} = \frac{30}{70} = \frac{3}{7}$$

$$P("2R") = P("2R, 1B, 1N") + P("2R, 2B") + P("2R, 2N") = \frac{18+9+3}{70} = \frac{30}{70} = \frac{3}{7}$$

$$= \frac{18}{70}$$

$$= \frac{\binom{3}{2} \binom{3}{2} \binom{2}{0}}{\binom{8}{4}} = \frac{9}{70}$$

$$= \frac{\binom{3}{0} \binom{3}{2} \binom{2}{2}}{\binom{8}{4}} = \frac{3}{70}$$

$$\binom{8}{4} = \frac{8!}{4! (8-4)!} =$$

$$= \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4!}{4! \cdot 4!} =$$

$$= \frac{8 \cdot 7 \cdot 6 \cdot 5}{4 \cdot 3 \cdot 2} = 70$$

ESERCIZIO

Si lancia 5 volte un dado equo. Calcolare la probabilità che escano (esattamente) "2 volte $\boxed{1}$ e 1 volta $\boxed{3}$ " (in un qualsiasi ordine).

RISPOSTA

Abbiamo 5 prove indipendenti (i 5 lanci) e 3 risultati in ogni prova:

$\boxed{1}$ con prob. $P_1 = 1/6$ (2 volte)

$\boxed{3}$ con prob. $P_2 = 1/6$ (1 volta)

$\boxed{2|4|5|6}$ con prob. $P_3 = 4/6$ (2 volte)
↓ dedotto

La probabilità richiesta è (applicazione diretta delle formule della multinomiale)

$$\frac{5!}{2! \cdot 1! \cdot 2!} \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^1 \left(\frac{4}{6}\right)^2 = \frac{5 \cdot 4 \cdot 3 \cdot 2}{2 \cdot 2} \cdot \frac{1}{6^3} \cdot \frac{4}{9} =$$

$$= \frac{5 \cdot 4}{36 \cdot 9} = \frac{5}{81}$$

osservazione

Vediamo un altro modo di ottenere lo stesso risultato (molto più complicato, non conveniente).

Pensiamo a 6 risultati possibili tutti con probabilità $\frac{1}{6}$. Abbiamo le due seguenti situazioni:

$$\rightarrow \text{per } \underbrace{h=2,4,5,6}_{4 \text{ casi}} \quad P(\text{"2 volte } \boxed{1}, 1 \text{ volta } \boxed{3}, 2 \text{ volte } \boxed{h}") = \frac{5!}{2!1!2!0!0!0!} \left(\frac{1}{6}\right)^5 = \frac{5 \cdot 4 \cdot 3 \cdot 2}{2 \cdot 2} \left(\frac{1}{6}\right)^5 =$$

$$\rightarrow \text{per } \{k, h\} \subset \{2, 4, 5, 6\} \quad P(\text{"2 volte } \boxed{1}, 1 \text{ volta } \boxed{3}, 1 \text{ volta } \boxed{k}, 1 \text{ volta } \boxed{h}") = \frac{5!}{2!1!1!1!0!0!} \left(\frac{1}{6}\right)^5 = \frac{5 \cdot 4 \cdot 3 \cdot 2}{2} \left(\frac{1}{6}\right)^5 = \frac{10}{6^4} = \frac{5}{1296}$$

$$\left(\frac{5}{1296}\right) = 6 \text{ casi} \quad P(\text{"2 volte } \boxed{1}, 1 \text{ volta } \boxed{3}, 1 \text{ volta } \boxed{k}, 1 \text{ volta } \boxed{h}") = \frac{5!}{2!1!1!1!0!0!} \left(\frac{1}{6}\right)^5 = \frac{5 \cdot 4 \cdot 3 \cdot 2}{2} \left(\frac{1}{6}\right)^5 = \frac{10}{6^4} = \frac{10}{1296}$$

$$\text{la prob. richiesta è } \left(4 \cdot \frac{5}{1296}\right) + \left(6 \cdot \frac{10}{1296}\right) = \frac{20+60}{1296} = \frac{80}{1296} = \frac{5}{81} \quad (\text{che è lo stesso risultato visto prima})$$

PROSSIMI ARGOMENTI (ALTRI DI DISTRIBUZIONI DI SERIE TE NOTEVOLI)

→ UNIFORME DISSETTA

→ POISSON (nome di un matematico francese)

→ GEOMETRICA E DISTRIBUZIONI COLLEGATE
(queste nelle prossime lezioni)

ess. Spesso saranno definite a partire dalla espressione della distribuzione (o delle densità discrete) senza fare riferimento a (R, \mathcal{A}, P) .

DISTRIBUZIONE UNIFORME DISSETTA

Si tratta del caso in cui, per un insieme finito $E = \{x_1, \dots, x_n\} \subset \mathbb{R}$, si ha

$$P(X \in B) = \frac{\#(B \cap E)}{\# E} = \frac{\#(B \cap E)}{n} \quad \forall B \subset \mathbb{R}.$$

ESEMPIO: Si ha questa distribuzione con $E = \{1, 2, 3, 4, 5, 6\}$ se X è la v.a. che indica il numero che esce lanciando un dado equip.

DISTRIBUZIONE DI POISSON

Una v.a. X ha distribuzione di Poisson con parametro $\lambda > 0$ (in qualche caso sciviamo $X \sim \text{Poisson}(\lambda)$)
se si ha

$$P_X(k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \forall k \in \{0, 1, 2, 3, \dots\}$$

OSS.

La definizione è ben posta se

$$\sum_{k=0}^{\infty} P_X(k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \underbrace{\frac{\lambda^k}{k!}}_{= e^{\lambda}} = e^{-\lambda} \cdot e^{\lambda} = 1.$$

per
definizione

ESERCIZIO

Sia $X \sim \text{Poisson}(\lambda=4)$.

1) Calcolare $P(X > 2)$

2) Calcolare $P(X=k | X \leq 2)$ per ogni $k \geq 0$ intero.

SVOLGIMENTO

1) Per eventi di questo tipo si deve passare alla probabilità dell'evento complementare;

$$\begin{aligned} P(X > 2) &= 1 - P(X \leq 2) = 1 - \sum_{k=0}^2 P_X(k) = 1 - \left(\frac{4^0}{0!} e^{-4} + \frac{4^1}{1!} e^{-4} + \frac{4^2}{2!} e^{-4} \right) = \\ &= 1 - (1 + 4 + 8) e^{-4} = 1 - 13 e^{-4}. \end{aligned}$$

$$2) P(X=k | X \leq 2) = \frac{P(\{X=k\} \cap \{X \leq 2\})}{P(X \leq 2)} = \begin{cases} \frac{P(X=k)}{P(X \leq 2)} & \text{per } k \in \{0, 1, 2\} \\ 0 & \text{per } k \geq 3 \end{cases}$$

(per $k \in \{0, 1, 2\}$ $\{X=k\} \subset \{X \leq 2\}$)
(per $k \geq 3$ $\{X=k\} \cap \{X \leq 2\} = \emptyset$)

ed inoltre (per $k=0, 1, 2$)

$$\frac{P(X=k)}{P(X \leq 2)} = \frac{P_X(k)}{\sum_{j=0}^2 P_X(j)} = \frac{\frac{4^k}{k!} e^{-4}}{\sum_{j=0}^2 \frac{4^j}{j!} e^{-4}} = \frac{\frac{4^k}{k!} e^{-4}}{\left(\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!}\right) e^{-4}} = \begin{cases} \text{per } k=0 & \frac{1}{13} \\ \text{per } k=1 & \frac{4}{13} \\ \text{per } k=2 & \frac{8}{13} \end{cases}$$

OSS. Gli eventi $\{\{X=k\}: k \geq 0\}$ costituiscono una partizione numerabile

Quindi si dava ovvero $\sum_{k=0}^{\infty} P(X=k | X \leq 2) = 1$.

In effetti si ha $\sum_{k=0}^{\infty} P(X=k | X \leq 2) = \sum_{k=0}^2 P(X=k | X \leq 2) = \frac{1+4+8}{13} = 1$.

trasemo gli
escludendo per $k \geq 3$
tutti uguali a zero

APPROSSIMAZIONE DELLA BINOMIALE CON LA POISSON

Sia $\lambda > 0$. Sia $n \geq \lambda$ intero; in questo modo $\frac{\lambda}{n} \in [0, 1]$ e non è restrittivo perché poi siamo interessati a considerare un limite per $n \rightarrow \infty$.

Prendiamo le densità di $X \sim \text{BIN}(n, p_n = \frac{\lambda}{n})$.

Allora, per $k \in \{0, 1, \dots, n\}$, si ha

$$\begin{aligned}
 p_X(k) &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n!}{k! (n-k)!} \cdot \frac{\lambda^k}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} = \\
 &= \frac{\lambda^k}{k!} \cdot \underbrace{\frac{n(n-1)\dots(n-k+1)}{n n \dots n}}_{\text{distribuzione binomiale}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{e^{-\lambda}} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{1} \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda}
 \end{aligned}$$

$$\frac{n!}{(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{(n-k)!}$$

OSS.
Questo numero
entra in $\frac{n!}{k!}$ a $n \geq k$
e quindi
questo limite
vale
questo numero

distribuzione di
una POISSON (λ)

Ricapitolando, per $X = X_m \sim \text{BIN}(n, p_m = \frac{\lambda}{m})$ e $Z \sim \text{Poisson}(\lambda)$, si ha

$$\lim_{m \rightarrow \infty} P_{X_m}(k) = P_Z(k) \quad \text{per } k \geq 0 \text{ intero}$$

(limite delle densità discrete)

Questo limite ha un interesse teorico che non approfondiremo.

Al contrario c'è un interesse pratico: serve per calcolare valori approssimati

di probabilità di eventi legati a v.a. con distribuzione Binomiale

"con n grande e p piccolo". Presenteremo di seguito due esempi illustrativi.

ESTRATTO / ESEMPIO

Sia $X \sim \text{BIN}(n=1000, p=\frac{3}{500})$. Calcolare $P(X \geq 5)$



OSS, si può pensare di fare riferimento ad un'urna con 500 palline numerate da 1 a 500.

Si compiono 1000 estrazioni casuali con reinserimento. Vogliamo calcolare le probabilità di estrarre almeno 5 volte uno dei numeri 1, 2, 3.

RISPOSTA

Il valore esatto delle probabilità da calcolare è

$$P(X \geq 5) = \sum_{k=5}^{1000} \binom{1000}{k} \left(\frac{3}{500}\right)^k \left(1 - \frac{3}{500}\right)^{1000-k} = 1 - \sum_{k=0}^4 \binom{1000}{k} \left(\frac{3}{500}\right)^k \left(1 - \frac{3}{500}\right)^{1000-k}.$$

Non è semplice calcolare un valore approssimato di queste probabilità.

Per trovare un valore numerico approssimato, facciamo riferimento alle approssimazioni

Portoname delle Binomiale. Si ha $P_X(k) \approx P_Z(k)$ (per $k \geq 0$ intero)

Allora

dove $Z \sim \text{Poisson}$ ($\lambda = np = \cancel{1000} \cdot \frac{3}{\cancel{500}} = 6$)

approssimazione con

$$\begin{aligned} P(X \geq 5) &= 1 - P(X \leq 4) = 1 - \sum_{k=0}^4 P(X = k) \approx 1 - \sum_{k=0}^4 P(Z = k) = \\ &= 1 - \sum_{k=0}^4 \frac{6^k}{k!} e^{-6} = 1 - \left(\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!} + \frac{6^4}{4!} \right) e^{-6} \\ &= 1 - (1 + 6 + 18 + 36 + 54) e^{-6} = 1 - 115 \cdot e^{-6}. \end{aligned}$$

Altro ESEMPIO

Stessa urne dell'esempio precedente.

1, 2, ..., 500

Si estraiano 200 palline, una alla volta e con rimessaggio.

Calcolare le probabilità di estrarre al più 2 volte uno dei numeri 1, 2, 3, 4, 5, 6, 7.

RISPOSTA

La probabilità richiesta è $P(Y \leq 2)$ dove $Y \sim \text{BIN}(n=200, p=\frac{7}{500})$

$$\text{Si ha } P(Y \leq 2) = \sum_{k=0}^2 \binom{200}{k} \left(\frac{7}{500}\right)^k \left(1 - \frac{7}{500}\right)^{200-k}.$$

Vogliamo trovare un valore approssimato come nell'esempio precedente.

Si ha $P_Y(k) \approx P_Z(k)$ ($\forall k \geq 0$ interi) dove $Z \sim \text{Poisson} (\lambda = np = 200 \cdot \frac{7}{500} = \frac{14}{5})$.

Allora

$$\begin{aligned}
 P(Y \leq 2) &= \sum_{k=0}^2 P(Y=k) \stackrel{\text{approximazione qui}}{\cong} \sum_{k=0}^2 P(Z=k) = \sum_{k=0}^2 \frac{\left(\frac{14}{5}\right)^k e^{-14/5}}{k!} = \\
 &= \left(\frac{\left(\frac{14}{5}\right)^0}{0!} + \frac{\left(\frac{14}{5}\right)^1}{1!} + \frac{\left(\frac{14}{5}\right)^2}{2!} \right) e^{-14/5} = \left(1 + \frac{14}{5} + \frac{\frac{196}{25}}{\frac{50}{25}} \right) e^{-14/5} = \\
 &= \frac{193}{25} e^{-14/5}.
 \end{aligned}$$