# Probabilità condizionate

Dati due eventi $E$ ed $H$, con $H \neq \phi$, si definisce *evento condizionato* il seguente ente logico a tre valori

$$E|H = \begin{cases} \textit{vero} & \text{se } E \text{ ed } H \text{ sono entrambi veri} \\ \textit{falso} & \text{se } E \text{ è falso e } H \text{ è vero} \\ \textit{indeterminato} & \text{se } H \text{ è falso.} \end{cases}$$

L'evento $E$ si chiama *condizionando*, mentre l'evento $H$ si chiama *condizionante*.

Gli eventi sono una sottoclasse degli eventi condizionati (se $H = \Omega$, si ha $E|H = E|\Omega = E$).

$P(E|H) =$ misura del grado di fiducia nel verificarsi di $E$ *supposto vero* $H$.

*Criterio operativo:*
si considera una scommessa condizionata (valida se risulta vero $H$, annullata se $H$ risulta falso).

Se un individuo valuta $P(E|H) = p$, deve essere disposto a pagare (ricevere) $p$ per ricevere (pagare)

$$\begin{cases} 1, & \textit{se si verifica} \quad EH \\ 0, & \textit{se si verifica} \quad E^cH \\ p, & \textit{se si verifica} \quad H^c. \end{cases} \tag{1}$$

In generale, si paga (si riceve) $pS$, con $S \neq 0$, per ricevere (pagare)

$$\begin{cases} S, & \textit{se si verifica} \quad EH \\ 0, & \textit{se si verifica} \quad E^cH \\ pS, & \textit{se si verifica} \quad H^c. \end{cases}$$

Il guadagno aleatorio $\mathcal{G}$ coincide con l'espressione

$$S|H||E| - pS|H| = S|H|(|E| - p)$$

e i suoi possibili valori sono:

$$\begin{cases} S(1-p), & \textit{se si verifica} \quad E \cap H \\ -pS, & \textit{se si verifica} \quad E^c \cap H \\ 0, & \textit{se si verifica} \quad H^c \end{cases}$$

Il valore nullo del guadagno corrisponde al caso di scommessa annullata e, ai fini della verifica della coerenza, non deve giocare nessun ruolo.

Quindi, per verificare la coerenza di $p$, si considerano solo i valori della restrizione, $G|H$, di $G$ ad $H$ (ovvero, si assume che la scommessa sia valida).

Allora, la condizione di coerenza diventa

$$Min\, \mathcal{G}|H \ \cdot Max\, \mathcal{G}|H \ \leq 0\,.$$

I valori possibili di $\mathcal{G}|H$ sono $S(1-p)$ e $-pS$;
. . . . . . $p$ è coerente se e solo se $0 \leq p \leq 1$.

**Alcune proprietà.**

- $P(E|H) = 0$, per ogni $E$ ed $H$ tali che $EH = \emptyset$,

- $P(E|H) = 1$, per ogni $E$ ed $H$ tali che $H \subseteq E$,

  in particolare:

- $P(\emptyset|H) = 0,\ P(H|H) = P(\Omega|H) = 1.$

**Teorema delle probabilità composte.**
Dati due eventi $E$ ed $H$, con $H \neq \emptyset$, se le valutazioni di probabilità $P(EH), P(H), P(E|H)$ sono coerenti si ha

$$P(EH) = P(E|H)P(H). \tag{2}$$

*Dim.* Alla valutazione $P(E|H) = p$ corrisponde, nello schema della scommessa condizionata, il guadagno aleatorio

$$\mathcal{G} = |H||E| - p|H| = |H|(|E| - p) = X - p\,,$$

dove $X$ (che può definirsi come l'$indicatore$ di $E|H$) è la vincita aleatoria, corrispondente al pagamento dell'importo $p$, definita da

$$X = 1 \cdot |EH| + 0 \cdot |E^cH| + p \cdot |H^c|\,. \tag{3}$$

D'altra parte, l'importo da pagare per ricevere $X$ è la sua previsione $\mathbb{P}(X)$.

Pertanto, operativamente, risulta $\mathbb{P}(X) = p$, ovvero

$$1 \cdot P(EH) + 0 \cdot P(E^cH) + p \cdot P(H^c) = p\,,$$

da cui segue

$$P(EH) + p[1 - P(H)] = p\,,$$

e quindi

$$P(EH) = P(E|H)P(H)\,.$$

**Corollario** Se $P(H) > 0$, si ottiene

$$P(E|H) = \frac{P(EH)}{P(H)}\,. \tag{4}$$

(molti autori utilizzano la formula precedente come *definizione* della probabilità di $E$ condizionata ad $H$).

*Teorema delle probabilità composte nell'impostazione classica.*

. . . $m$ casi possibili, $r_H$ casi favorevoli ad $H$, $r_{EH}$ casi favorevoli ad $EH$.

Nota: tra i casi favorevoli ad $H$, quelli favorevoli ad $E$ sono $r_{EH}$.

. . . assumendo $r_H > 0$, si ha

$$P(H) = \frac{r_H}{m}\ , \quad P(EH) = \frac{r_{EH}}{m}\ , \quad P(E|H) = \frac{r_{EH}}{r_H}\ ,$$

pertanto

$$P(E|H) = \frac{r_{EH}}{r_H} = \frac{\frac{r_{EH}}{m}}{\frac{r_H}{m}} = \frac{P(EH)}{P(H)}\ .$$

**Corollario.**
Se $H = \Omega$ allora $EH = E$ e si ha

$$P(E) = P(E|\Omega)P(\Omega) = P(E|\Omega)\,. \qquad (5)$$

**Esempio** (*Estrazioni del lotto.*)

Sia $X = 1^0$ numero estratto,

$E = (X \leq 45)$ ed $H = (X > 30)$.

Calcolare $P(E|H)$.

**Generalizzazione del teorema delle probabilità composte.** Dati $n$ eventi arbitrari $E_1, E_2, \ldots, E_n$, iterando la formula relativa al caso di 2 eventi, si ha:

$$
\begin{aligned}
& P(E_1 E_2 \cdots E_n) = \\
& = P(E_n | E_1 E_2 \cdots E_{n-1}) P(E_1 E_2 \cdots E_{n-1}) = \\
& = \cdots\cdots\cdots\cdots = \\
& = P(E_1) P(E_2 | E_1) \cdots P(E_n | E_1 E_2 \cdots E_{n-1}).
\end{aligned}
\tag{6}
$$

Formule equivalenti si ottengono permutando in tutti i modi possibili l'ordine degli eventi $E_i$.

Ad esempio, per $n = 3$ si ha

$$
P(E_1 E_2 E_3) = P(E_1) P(E_2 | E_1) P(E_3 | E_1 E_2) =
$$

$$
= P(E_{i_1}) P(E_{i_2} | E_{i_1}) P(E_{i_3} | E_{i_1} E_{i_2}) \,,
$$

dove $\{i_1, i_2, i_3\}$ è una permutazione di $\{1, 2, 3\}$.

**Esempio** *(Gioco della roulette russa)*
In una pistola a 6 colpi viene inserito un solo proiettile e il tamburo viene fatto girare vorticosamente. Quindi 6 prigionieri sono costretti a sottomettersi alla prova della roulette russa.
Verificare che i 6 prigionieri hanno la stessa probabilità di morire, ovvero che gli eventi
$E_i =$ *il proiettile esplode all'i-mo colpo*, $i = 1, \ldots, 6$,
hanno tutti probabilità $\frac{1}{6}$.

Ovviamente, si ha $P(E_1) = \frac{1}{6}$. Inoltre

$$P(E_2|E_1^c) = \tfrac{1}{5}\,, \quad P(E_3|E_1^cE_2^c) = \tfrac{1}{4}\,,$$

$$\cdots\cdots\,, \qquad P(E_6|E_1^c \cdots E_5^c) = 1\,.$$

Allora, osservando che

$$E_2 = E_1^cE_2\,, \;\; E_3 = E_1^cE_2^cE_3\,, \;\; \cdots\,, \;\; E_6 = E_1^c \cdots E_5^cE_6\,,$$

applicando la (6) si ha

$$P(E_2) = P(E_1^c)P(E_2|E_1^c) = \tfrac{5}{6} \times \tfrac{1}{5} = \tfrac{1}{6}\ ;$$

$$P(E_3) = P(E_1^c)P(E_2^c|E_1^c)P(E_3|E_1^c E_2^c) = \tfrac{5}{6} \times \tfrac{4}{5} \times \tfrac{1}{4} = \tfrac{1}{6}\ ;$$

....................................................................

$$P(E_6) = P(E_1^c)P(E_2^c|E_1^c)P(E_3^c|E_1^c E_2^c)\cdots$$

$$\cdots P(E_5^c|E_1^c \cdots E_4^c)P(E_6|E_1^c \cdots E_5^c) =$$

$$= \tfrac{5}{6} \times \tfrac{4}{5} \times \tfrac{3}{4} \times \tfrac{2}{3} \times \tfrac{1}{2} \times 1 = \tfrac{1}{6}\ .$$

**Proprietà additiva per probabilità condizionate.**
Dati $A, B, H$, con $AB = \emptyset$, $P(H) > 0$, si ha

$$\begin{aligned} P(A \vee B|H) &= \tfrac{P(AH \vee BH)}{P(H)} = \tfrac{P(AH)}{P(H)} + \tfrac{P(BH)}{P(H)} = \\ &= P(A|H) + P(B|H)\ . \end{aligned} \tag{7}$$

La formula (7) può essere dimostrata anche nel caso $P(H) = 0$ (introducendo la condizione di coerenza nel caso generale di probabilità condizionate).

Applicando la (7), con $A = E$ e $B = E^c$, si ottiene

$$P(\Omega|H) = P(E \vee E^c|H) = P(E|H) + P(E^c|H) = 1\,,$$

e quindi

$$P(E^c|H) = 1 - P(E|H)\,. \tag{8}$$

*Nota:* in generale

$$P(E|H^c) \neq 1 - P(E|H)\,.$$

Data una partizione $\{H_1, \ldots, H_n\}$ e un evento qualsiasi $E$, ricordando che

$$E = EH_1 \vee \cdots \vee EH_n$$

ed applicando la proprietà additiva e il teorema delle probabiltà composte, si ottiene la seguente *formula di disintegrazione*

$$\begin{aligned} P(E) &= P(EH_1) + \cdots + P(EH_n) = \\ &= P(H_1)P(E|H_1) + \cdots + P(H_n)P(E|H_n). \end{aligned} \tag{9}$$

**Esempio** *(Problema del condannato)*
In un paese orientale un prigioniero è stato condannato a morte da uno sceicco.
Prima dell'esecuzione, lo sceicco offre una possibilità di salvezza al condannato, mettendogli a disposizione 2 urne $U_1$ e $U_2$, con 2 palline bianche e 2 nere.
Il condannato deve distribuire a suo piacere le palline nelle due urne, con la condizione che nessuna urna rimanga vuota.
Verrà poi scelta a caso un'urna da cui verrà estratta una pallina.
Se la pallina estratta sarà bianca il prigioniero sarà graziato.
Qual'è la migliore ripartizione delle palline nelle due urne?

Definiti gli eventi
$B =$ *la pallina estratta è bianca*,
$H_1 =$ *viene utilizzata l'urna* $U_1$,
$H_2 =$ *viene utilizzata l'urna* $U_2$,
la decisione migliore per il condannato è quella che rende massima la probabilità di $B$.

Supponiamo che il condannato inserisca $h$ palline bianche e $k$ palline nere in $U_1$, con $0 < h + k < 4$, e le rimanenti $4 - h - k$ palline in $U_2$.

Con tale strategia si ha

$$P(B) = P(H_1)P(B|H_1) + P(H_2)P(B|H_2) =$$

$$= \frac{1}{2}\left( \ \frac{h}{h+k} + \frac{2-h}{4-h-k} \ . \ \right)$$

Basta esaminare i seguenti casi (con il simbolo $\sim$ tra parentesi si indicano le decisioni equivalenti):

1. $h = 0 \ , \ \ k = 1 \ (\sim h = 2 \ , \ \ k = 1) \ ;$

2. $h = 0 \ , \ \ k = 2 \ \ (\sim h = 2 \ , \ \ k = 0) \ ;$

3. $h = 1 \ , \ \ k = 1 \ ;$

4. $h = 1 \ , \ \ k = 0 \ \ (\sim h = 1 \ , \ \ k = 2) \ .$

- Nel caso 1 si ha $P(B) = \frac{1}{2}(\frac{0}{1} + \frac{2}{3}) = \frac{1}{3}$ ;

- nel caso 2 si ha $P(B) = \frac{1}{2}(\frac{0}{2} + \frac{2}{2}) = \frac{1}{2}$ ;

- nel caso 3 si ha $P(B) = \frac{1}{2}(\frac{1}{2} + \frac{1}{2}) = \frac{1}{2}$ ;

- Nel caso 4 si ha $P(B) = \frac{1}{2}(\frac{1}{1} + \frac{1}{3}) = \frac{2}{3}$ .

La 4 è decisione migliore (una pallina bianca in un'urna e le rimanenti nell'altra).

Il problema si può generalizzare considerando $n$ palline bianche ed $n$ nere, con $n > 2$.

La decisione migliore è sempre quella di mettere una pallina bianca in un'urna e le rimanenti nell'altra.

A tale decisione corrisponde per $P(B)$ il valore (massimo)

$$P(B) = \frac{1}{2}\left(1 + \frac{n-1}{2n-1}\right) ,$$

che all'aumentare di $n$ sale verso $\frac{3}{4}$.

**Indipendenza Stocastica.**

Dati $E$ ed $H$, per le probabilità $P(E)$ e $P(E|H)$ si può avere:

$$P(E|H) > P(E),\ P(E|H) < P(E),\ P(E|H) = P(E).$$

Nel primo caso $E$ è *correlato positivamente* con $H$, cioè assumendo $H$ vero la probabilità di $E$ aumenta.

Nel secondo caso $E$ è *correlato negativamente* con $H$.

Nel terzo caso si dice che $E$ è *stocasticamente indipendente* da $H$ : l'ipotesi che $H$ sia vero non fa nè aumentare nè diminuire la probabilità di $E$.

In questo caso si ha

$$P(EH) = P(H)P(E|H) = P(E)P(H)\ , \qquad (10)$$

cioè la probabilità dell'evento $EH$ coincide con il prodotto delle probabilità di $E$ ed $H$.

Se $P(E|H) = P(E) > 0$, segue $P(H|E) = P(H)$, cioè se $E$ è stocasticamente indipendente da $H$, si ha che $H$ è indipendente da $E$.

In tal caso segue anche

$$P(H^c|E) = 1 - P(H|E) = 1 - P(H) = P(H^c) .$$

Analoghe considerazioni si possono fare per le probabilità condizionate

$$P(H|E^c) , \;\; P(H^c|E^c) , \;\; P(E|H^c) , \;\; P(E^c|H^c) .$$

In generale, data una famiglia arbitraria $\mathcal{F}$ di eventi (tutti di probabilità positiva e minore di 1), gli eventi di $\mathcal{F}$ si dicono (**stocasticamente) indipendenti** se, per ogni sottofamiglia $\{E_1, \ldots, E_n\}$ di $\mathcal{F}$, $n \geq 2$, si ha

$$P(E_1 \cdots E_n) = P(E_1) \cdots P(E_n) . \tag{11}$$

**Osservazione.** Se la condizione (11) vale per un certo intero $n$, non è detto che valga per $n + 1$ (ad esempio, se vale per $n = 2$, può non valere $n+1 = 3$).

**Esempio.** Si effettua un'estrazione da un'urna contenente 4 palline, 1 rossa, 1 nera , 1 bianca, 1 gialla. Consideriamo gli eventi
$R$= "la pallina estratta è rossa",
$N$= "la pallina estratta è nera",
$B$= "la pallina estratta è bianca",
$G$= "la pallina estratta è gialla",

$E = R \vee B$, $H = N \vee B$, $A = G \vee B$.

Si può verificare che la (11) non vale per la terna $(E, H, A)$, mentre vale per ciascuna delle 3 coppie $(E, H)$, $(E, A)$, $(H, A)$.

**Teorema di Bayes**

Tale teorema evidenzia il ruolo delle probabilità condizionate come strumento teorico per *apprendere dall'esperienza*, attraverso l'aggiornamento delle valutazioni di probabilità di una o più *ipotesi*, quando lo stato di informazione cambia per effetto di nuovi dati o notizie.

Dati $E$ ed $H$, per il teorema delle probabilità composte si ha:

$$P(EH) = P(H)P(E|H) = P(E)P(H|E) \; ,$$

da cui, assumendo $P(E) > 0$, si ottiene

$$\begin{aligned} P(H|E) &= \tfrac{P(H)P(E|H)}{P(E)} = \\ &= \tfrac{P(H)P(E|H)}{P(H)P(E|H)+P(H^c)P(E|H^c)} \; . \end{aligned} \tag{12}$$

(*teorema o formula di Bayes* nella sua versione più semplice).

Supponiamo che $H$ rappresenti un'*ipotesi incerta*, di probabilità $P(H)$, ed $E$ un fatto osservabile, di probabilità $P(E)$ e di probabilità condizionata ad $H$ pari a $P(E|H)$.

La (12) ci mostra come il "grado di fiducia" nei riguardi di $H$ si modifica quando si suppone di aver osservato l'evento $E$.

**Esempio.** Una persona (Tizio) attende nella stazione ferroviaria di una località $A$ l'arrivo di un amico

(Caio) che dovrebbe essere partito dalla stazione di una località $B$.

Tizio sa che Caio, che è un tipo bizzarro, ha lanciato in aria una moneta, decidendo di partire (ipotesi $H$) oppure no (ipotesi $H^c$) a seconda che il risultato del lancio sia stato Testa oppure Croce.

Inoltre, nel caso che il risultato del lancio sia stato Testa (ipotesi $H$ vera), avendo a disposizione 6 treni Tizio ha lanciato in aria un dado, scegliendo il treno corrispondente al risultato ottenuto con il lancio del dado.

Considerati gli eventi
$A_1 =$ *Caio arriva con il primo treno*,
$\ldots$
$A_6 =$ *Caio arriva con il sesto treno*,
calcolare la probabilità che Caio sia partito supposto che non sia arrivato con nessuno dei primi cinque treni.
Posto

$$E = A_1^c A_2^c \cdots A_5^c = (A_1 \vee \cdots \vee A_5)^c = A_6 \vee H^c \ ,$$

si tratta di calcolare $P(H|E)$. Nel nostro caso si ha

$$P(H) = P(H^c) = \frac{1}{2}\,,\ \ P(A_i|H) = \frac{1}{6}\,,\ \ i = 1, \ldots, 6.$$

Osservando che

$$P(E|H) = P(A_6 \vee H^c|H) = P(A_6|H) + P(H^c|H) = \\ = \tfrac{1}{6} + 0 = \tfrac{1}{6}\,,$$

$$P(E|H^c) = P(A_6 \vee H^c|H^c) = P(A_6|H^c) + P(H^c|H^c) = \\ = 0 + 1 = 1\,,$$

segue

$$P(E) = P(E \mid H)P(H) + P(E \mid H^c)P(H^c) =$$

$$= P(A_6 \mid H)P(H) + P(H^c) = \tfrac{1}{12} + \tfrac{1}{2} = \tfrac{7}{12}\,.$$

In altro modo, $P(E)$ si può calcolare osservando che per ogni $i = 1, \ldots, 6$ si ha $A_i = A_i H$, da cui segue

$$P(A_i) = P(A_i H) = P(H)P(A_i|H) = \tfrac{1}{2} \times \tfrac{1}{6} = \tfrac{1}{12}\,.$$

Allora

$$P(E) = P(A_6 \vee H^c) = P(A_6) + P(H^c) = \frac{1}{12} + \frac{1}{2} = \frac{7}{12} \, .$$

Infine, dalla formula (12) si ottiene

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)} = \frac{\frac{1}{2} \times \frac{1}{6}}{\frac{7}{12}} = \frac{1}{7} \, .$$

Come si vede, supposto che sia rimasta possibile solo l'eventualità che Caio arrivi con l'ultimo treno, la probabilità che Caio sia partito è scesa da $\frac{1}{2}$ a $\frac{1}{7}$.

È da notare che con un ragionamento intuitivo errato si può essere indotti a valutare $P(H|E) = \frac{1}{2}$. Per convincersi che tale valutazione non è "ragionevole", basta considerare il caso in cui i treni anzichè 6 siano 100, oppure 1000, oppure ... e supporre che sia rimasto possibile solo l'arrivo con l'ultimo treno.

Con un ragionamento analogo a quello precedente si può verificare che

$$P(H|A_1^c) = \tfrac{5}{11}\,, \qquad P(H|A_1^cA_2^c) = \tfrac{4}{10}\,,$$

$$P(H|A_1^cA_2^cA_3^c) = \tfrac{3}{9}\,, \quad P(H|A_1^cA_2^cA_3^cA_4^c) = \tfrac{2}{8}\,.$$

Data una partizione $\{H_1, \ldots, H_n\}$ ed un evento $E$, per ogni fissato $i = 1, \ldots, n$, si ha

$$P(H_i|E) = \tfrac{P(H_i)P(E|H_i)}{P(E)} = \\ = \tfrac{P(H_i)P(E|H_i)}{P(H_1)P(E|H_1)+\cdots+P(H_n)P(E|H_n)}\,. \tag{13}$$

(*versione generale del Teorema di Bayes*)

Nelle applicazioni, $E$ rappresenta una "osservazione" in un dato esperimento, mentre $\{H_1, \ldots, H_n\}$ sono le possibili *ipotesi* che "spiegano" l'evento osservato.

$P(H_i),\ i = 1, \ldots, n,$ probabilità *iniziali*;
$P(H_i|E),\ i = 1, \ldots, n,$ probabilità *finali*;
$P(E|H_i),\ i = 1, \ldots, n,$ *verosimiglianze*.

La (13) ci mostra come $P(H_i)$ si modifica in $P(H_i|E)$ quando si suppone di aver osservato l'evento $E$.
In questo senso la formula di Bayes rappresenta lo strumento formale per *apprendere dall'esperienza*.

**Esempio.** Un diamante viene nascosto in una fra $n$ scatole. Successivamente, una persona (Tizio) sceglie a caso una scatola e se trova il diamante se ne impossessa. Siano definiti gli eventi:
$H =$ "Tizio sa quale scatola contiene il diamante",
$E =$"la scatola scelta da Tizio contiene il diamante".

Indichiamo con $p = P(H)$ la probabilità iniziale di $H$ e, inoltre, assumiamo che sia $P(E|H) = 1$ e $P(E|H^c) = 1/n$.
Applicando il teorema di Bayes, verificare che

$$P(H|E) = \cdots = \frac{np}{(n-1)p+1}\,.$$

Altri possibili aspetti da esaminare sono:

$$P(H|E) > p\,? \quad P(H|E) > P(H^c|E)\,?$$

**Nota** (*Test d'ipotesi*).
Se $P(H|E) > P(H^c|E)$ vuol dire che, assumendo vero $E$, l'ipotesi $H$ è più probabile dell'ipotesi $H^c$. Nel caso di $n$ ipotesi, applicando il teorema di Bayes, si può (soltanto) determinare quale delle ipotesi fatte è diventata più probabile rispetto alle altre, mentre non ha senso, in generale, parlare dell'ipotesi più probabile come se fosse (quella) *vera*.

**Esempio.** Un oggetto viene nascosto (da Tizio) in uno fra 7 contenitori, che successivamente vengono divisi in due gruppi di 3 e 4 contenitori rispettivamente.
Un amico (Caio), per cercare l'oggetto, può aprire solo 2 contenitori in uno dei due gruppi a sua scelta. Conviene scegliere a caso nel gruppo da 3 (strategia $A$) oppure nel gruppo da 4 (strategia $B$)?
Supposto di aver aperto 2 contenitori scelti a caso nel gruppo da 3 e di non aver trovato nulla, qual'è la probabilità che l'oggetto stia nel contenitore rimanente?
Si può dimostrare che le due strategie sono equivalenti, cosi come sarebbe indifferente aprire a caso

un contenitore del primo gruppo e uno del secondo gruppo.
Infatti, ognuno dei 7 contenitori ha probabilità $\frac{1}{7}$ di contenere l'oggetto nascosto. Allora, definiti gli eventi
$H =$ *l'oggetto è nascosto nel gruppo da 3*,
$H^c =$ *l'oggetto è nascosto nel gruppo da 4*,
$E_1 =$ *Caio apre a caso 2 contenitori nel gruppo da 3 e trova l'oggetto*,
$E_2 =$ *Caio apre a caso 2 contenitori nel gruppo da 4 e trova l'oggetto*, si ha

$$P(H) = \frac{3}{7}\,; \quad P(H^c) = \frac{4}{7}\,.$$

Inoltre

$$P(E_1|H) = \frac{\binom{1}{1}\binom{2}{1}}{\binom{3}{2}} = \tfrac{2}{3}$$

$$P(E_2|H^c) = \frac{\binom{1}{1}\binom{3}{1}}{\binom{4}{2}} = \tfrac{1}{2}\,.$$

Allora, con la strategia $A$ si ha

$$P(E_1) = P(E_1 H) = P(H) P(E_1|H) = \frac{3}{7} \times \frac{2}{3} = \frac{2}{7}\,,$$

mentre con la strategia $B$ si ha

$$P(E_2) = P(E_2 H^c) = P(H^c) P(E_2|H^c) = \frac{4}{7} \times \frac{1}{2} = \frac{2}{7}\,.$$

Pertanto le due strategie sono equivalenti. Inoltre, si può osservare che aprendo 2 contenitori scelti a caso fra i 7 iniziali la probabilità di trovare l'oggetto è ancora la stessa, essendo data da

$$\frac{\binom{1}{1}\binom{6}{1}}{\binom{7}{2}} = \frac{2}{7}\,.$$

Infine, con riferimento alla seconda domanda, occorre calcolare $P(H|E_1^c)$.
Osservando che $P(E_1^c|H^c) = 1$, dal teorema di Bayes

si ottiene

$$P(H|E_1^c) = \frac{P(H)P(E_1^c|H)}{P(E_1^c)} = \frac{P(H)P(E_1^c|H)}{P(H)P(E_1^c|H)+P(H^c)P(E_1^c|H^c)} =$$

$$= \frac{\frac{3}{7}\times\frac{1}{3}}{\frac{3}{7}\times\frac{1}{3}+\frac{4}{7}\times 1} = \frac{1}{5} \ .$$

Ovviamente risulta: $P(H^c|E_1^c) = \cdots = \frac{4}{5}$.
Inoltre

$$P(H^c|E_2^c) = \cdots = \frac{2}{5}\,, \quad P(H|E_2^c) = \cdots = \frac{3}{5} \ .$$

**Teorema.** Data una famiglia arbitraria $\mathcal{F}$ di eventi (tutti di probabilità positiva e minore di 1), se gli eventi di $\mathcal{F}$ sono (**stocasticamente) indipendenti** allora, per ogni sottofamiglia $\{E_1, \ldots, E_n\}$ di $\mathcal{F}$, con $n \geq 2$, e per ogni $1 \leq s \leq n$, si ha

$$P(E_1^c E_2^c \cdots E_s^c E_{s+1} \cdots E_n) =$$

$$= P(E_1^c)P(E_2^c) \cdots P(E_s^c)P(E_{s+1}) \cdots P(E_n) \ . \tag{14}$$

## Distribuzione Binomiale.

$E_1, E_2, \ldots, E_n$, eventi indipendenti ed equiprobabili, con $P(E_i) = p,\ P(E_i^c) = 1 - p = q$.
$X$ = numero aleatorio di successi sulle n prove, ovvero

$$X = |E_1| + |E_2| + \cdots + |E_n|.$$

Codominio di $X$: $\{0, 1, 2, \ldots, n\}$.

Distribuzione di probabilità di $X$:

si ha: $(X = 0) = E_1^c E_2^c \cdots E_n^c$, quindi

$$P(X = 0) = P(E_1^c E_2^c \cdots E_n^c) = \cdots = (1 - p)^n = q^n\,.$$

Simmetricamente: $(X = n) = E_1 E_2 \cdots E_n$, quindi

$$P(X = n) = P(E_1 E_2 \cdots E_n) = \cdots = p^n\,.$$

L'evento $(X = 1)$ coincide con la seguente unione di $n$ costituenti equiprobabili

$$E_1 E_2^c \cdots E_n^c \vee E_1^c E_2 E_3^c \cdots E_n^c \vee \cdots \vee E_1^c \cdots E_{n-1}^c E_n$$

e quindi

$$P(X = 1) = P(E_1 E_2^c \cdots E_n^c) + \cdots + P(E_1^c \cdots E_{n-1}^c E_n) =$$

$$= pq^{n-1} + \cdots + pq^{n-1} = npq^{n-1} = \tbinom{n}{1} p^1 q^{n-1}\,.$$

L'evento $(X = h)$ coincide con l'unione di $\binom{n}{h}$ costituenti, ognuno di probabilità $p^h q^{n-h}$, tra i quali ci sono, per esempio, i seguenti due

$$E_1 \cdots E_h E_{h+1}^c \cdots E_n^c\,, \quad E_1^c \cdots E_{n-h}^c E_{n-h+1} \cdots E_n\,.$$

Pertanto, per ogni $h \in \{0, 1, \ldots, n\}$, si ha

$$P(X = h) = \tbinom{n}{h} p^h q^{n-h}\,. \qquad (15)$$

$X$ ha una *distribuzione binomiale* di parametri $n$ e $p$, che si indica con il simbolo: $X \sim \mathbf{B}(n, p)$ .
Ovviamente

$$\sum_{h=0}^{n} P(X = h) = \sum_{h=0}^{n} \binom{n}{h} p^h q^{n-h} = (p + q)^n = 1 \,.$$

In particolare se $p = \frac{1}{2}$, come nel caso di $n$ lanci di una moneta ben costruita, si ha

$$P(X = h) = \frac{\binom{n}{h}}{2^n} \,, \quad h = 0, 1, 2, \ldots, n. \qquad (16)$$

In questo caso la distribuzione binomiale è simmetrica e si ha

$$P(X = 0) = P(X = n) = \frac{1}{2^n} \,,$$

$$P(X = 1) = P(X = n - 1) = \frac{n}{2^n} \,,$$

..................................................

Se $p \neq \frac{1}{2}$ la distribuzione è asimmetrica (come si può vedere dai grafici riportati di seguito).

Figura 1: Binomiale p=1/2, n=10

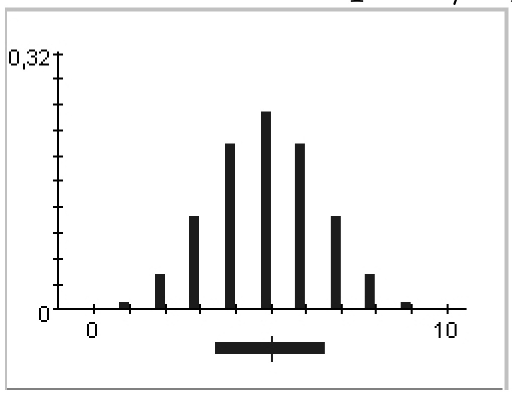
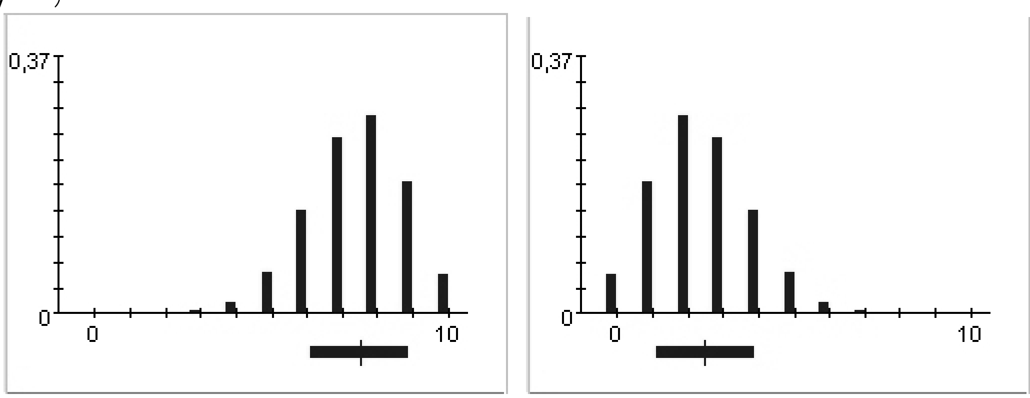



Figura 2: Binomiale p=2/3, n=10 - Binomiale p=1/3, n=10



*Previsione*

$$\mathbb{P}(X) = \sum_{i=1}^{n} \mathbb{P}(|E_i|) = \sum_{i=1}^{n} P(E_i) = np\,.$$

*Varianza*

Per l'indipendenza degli eventi $E_1, \ldots, E_n$, si ha

$Cov(|E_i|, |E_j|) = P(E_iE_j) - P(E_i)P(E_j) = 0\,, \ \ \forall\, i \neq j\,.$

Pertanto

$$Var(X) = \sum_{i=1}^{n} Var(|E_i|) = \sum_{i=1}^{n} pq = npq\,.$$

In sintesi, se $X \sim \mathbf{B}(n, p)$, allora

$$\mathbb{P}(X) = np\,, \quad Var(X) = npq\,, \quad \sigma_X = \sqrt{npq}\,.$$

Ad esempio:

$$\begin{array}{llll}
\mathbf{B}(10, \frac{1}{2}): & \mathbb{P}(X) = 5 & \sigma_X^2 = 2.50 & \sigma_X = 1.58 \\
\mathbf{B}(10, \frac{3}{4}): & \mathbb{P}(X) = 7.5 & \sigma_X^2 = 1.88 & \sigma_X = 1.37 \\
\mathbf{B}(10, \frac{1}{4}): & \mathbb{P}(X) = 2.5 & \sigma_X^2 = 1.88 & \sigma_X = 1.37
\end{array}$$

**Esempio***(distribuzione a caso di $n$ palline in $n$ urne)*
..............................
$X =$ "numero aleatorio di palline che capitano nella prima urna";

$X \sim ?$ Distribuzione di $X$ per $n$ molto grande?

**Estrazioni con restituzione da un'urna di composizione nota.**
Si effettuano $n$ estrazioni con restituzione da un'urna contenente $N$ palline, di cui $pN$ bianche e $qN$ nere ($pN + qN = N$).

$E_i =$ "l'i-esima pallina estratta è bianca",
$i = 1, \ldots, n$;

$X = \sum_{i=1}^{n} |E_i|$, numero aleatorio di palline bianche estratte nelle $n$ prove.

Gli eventi $E_i$ sono indipendenti ed equiprobabili, quindi $X \sim \mathbf{B}(n, p)$.

**Estrazioni senza restituzione da un'urna di composizione nota.**
Si effettuano $n$ estrazioni senza restituzione da un'urna contenente $N$ palline, di cui $pN$ bianche e $qN$ nere $(pN + qN = N)$.
Ovviamente, deve essere $n \leq N$.

$E_i =$ "l'i-esima pallina estratta è bianca",
$i = 1, \ldots, n$;

$X = \sum_{i=1}^{n} |E_i|$, numero aleatorio di palline bianche estratte nelle $n$ prove.

Valutiamo $P(E_i)$, per $i = 1, \ldots, n$.

$$P(E_1) = \tfrac{pN}{N} = p\,.$$

$$P(E_2) = P(E_2|E_1)P(E_1) + P(E_2|E_1^c)P(E_1^c) =$$

$$= \tfrac{pN-1}{N-1}p + \tfrac{pN}{N-1}q = p\,.$$

In alternativa, con il calcolo combinatorio si ha (*rapporto fra casi favorevoli e casi possibili*)

$$P(E_2) = \frac{(N-1)pN}{N(N-1)} = p\,.$$

In generale

$$P(E_i) = \frac{(N-1)(N-2)\cdots(N-i+1)pN}{N(N-1)\cdots(N-i+1)} = p\,.$$

In altro modo, se si fanno $N$ estrazioni, le $pN$ palline bianche usciranno in uno degli $\binom{N}{pN}$ sottoinsiemi di $pN$ prove fra le $N$. I casi favorevoli sono quelli in cui una pallina bianca esce nell'i-ma estrazione e le rimanenti $pN-1$ palline bianche escono in uno degli $\binom{N-1}{pN-1}$ sottoinsiemi di $pN-1$ prove fra le rimanenti $N-1$. Pertanto

$$P(E_i) = \frac{\overbrace{1}^{\textit{prova i-esima}} \cdot \binom{N-1}{pN-1}}{\binom{N}{pN}} = p\,.$$

Pertanto gli eventi sono equiprobabili anche nel caso di estrazioni senza restituzione.
Considerando adesso il costituente

$$E_1E_2\cdots E_hE^c_{h+1}\cdots E^c_n\,,$$

si ha

$$P(E_1E_2\cdots E_hE^c_{h+1}\cdots E^c_n) = \frac{D_{pN,h}D_{qN,n-h}}{D_{N,n}} =$$

$$= \frac{pN(pN-1)(pN-2)\cdots(pN-(h-1))qN(qN-1)\cdots(qN-(n-h)+1)}{N(N-1)(N-2)\cdots(N-(h-1))(N-h)\cdots(N-n+1)} =$$

$$= \frac{(pN)!(N-pN)!(N-n)!}{(pN-h)!(N-pN-n+h)!N!} = \frac{\frac{(N-n)!}{(pN-h)!(N-pN-n+h)!}}{\frac{N!}{(pN)!(N-pN)!}} = \frac{\binom{N-n}{pN-h}}{\binom{N}{pN}}\,.$$

Tale probabilità dipende solo da $h$ ed $n$, quindi per ogni scelta degli $h$ successi si ottiene

$$P(E_{i_1}E_{i_2}\cdots E_{i_h}E^c_{i_{h+1}}\cdots E^c_{i_n}) = \frac{\binom{N-n}{pN-h}}{\binom{N}{pN}}\,.$$

Indicando con $A_{h,n}$ il generico costituente

$$E_{i_1} E_{i_2} \cdots E_{i_h} E^c_{i_{h+1}} \cdots E^c_{i_n}$$

favorevole ad $(X = h)$ si ha

$$(X = h) = \bigvee_{\{i_1, i_2, \ldots, i_n\}} A_{h,n} \,,$$

dove l'unione è fatta rispetto a tutti gli $\binom{n}{h}$ costituenti favorevoli ad $(X = h)$. In definitiva si ha

$$P(X = h) = \frac{\binom{n}{h}\binom{N-n}{pN-h}}{\binom{N}{pN}} \,. \qquad (17)$$

Se si fanno estrazioni in blocco si ottiene la formula (equivalente alla (17), ma più facile da ricordare)

$$P(X = h) = \frac{\binom{pN}{h}\binom{qN}{n-h}}{\binom{N}{n}} \,. \qquad (18)$$

$X$ ha una distribuzione **Ipergeometrica**, che si indica con il simbolo

$$X \sim \mathbf{H}(N, n, p)\,.$$

Per quanto riguarda il codominio di $X$, si ha

$$\max\{0, n - qN\} \leq \; X \leq \min\{pN, n\}\,.$$

**Esempio.** Consideriamo 15 estrazioni senza restituzione da un'urna contenente 12 bianche e 8 nere. Si ha

$$N = 20, \quad pN = 12, \quad qN = 8, \quad n = 15, \quad p = \tfrac{3}{5}.$$

$$\max\{0, n - qN\} = \max\{0, 7\} = 7\,,$$

$$\min\{pN, n\} = \min\{12, 15\} = 12\,,$$

$$X \in \{7, 8, 9, 10, 11, 12\}\,, \quad X \sim \mathbf{H}(20, 15, 3/5)\,.$$

Fissato $h \ \in \{7, 8, \ldots, 12\}$, ad esempio $h = 8$, si ha

$$P(X = 8) = \frac{\binom{pN}{h}\binom{qN}{n-h}}{\binom{N}{n}} = \frac{\binom{12}{8}\binom{8}{7}}{\binom{20}{15}}\,.$$

*Previsione*

Essendo $P(E_i) = p$, come per la binomiale si ha

$$\mathbb{P}(X) = \sum_{i=1}^{n} \mathbb{P}(|E_i|) = \sum_{i=1}^{n} P(E_i) = np\,.$$

*Varianza*

La varianza, salvo casi banali, risulta più piccola rispetto a quella della binomiale. Infatti

$$Cov(|E_i|, |E_j|) = P(E_iE_j) - P(E_i)P(E_j) =$$

$$= \frac{pN-1}{N-1}p - p^2 = -\frac{pq}{N-1}\,.$$

Quindi

$$Var(X) = \sum_{i=1}^{n} Var(|E_i|) + 2\sum_{i<j} Cov(|E_i|, |E_j|) =$$

$$= npq + 2\binom{n}{2}\left(-\frac{pq}{N-1}\right) = npq\left(1 - \frac{n-1}{N-1}\right) \leq npq\,.$$

In sintesi se $X \sim \mathbf{H}(N, n, p)$ si ha

$$P(X = h) = \frac{\binom{n}{h}\binom{N-n}{pN-h}}{\binom{N}{pN}} = \frac{\binom{pN}{h}\binom{qN}{n-h}}{\binom{N}{n}}\,,$$

$$\mathbb{P}(X) = np\,, \quad Var(X) = npq\left(1 - \frac{n-1}{N-1}\right).$$

## Comportamento Asintotico.

$$\lim_{N\to\infty} P(A_{h,n}) =$$

$$= \lim_{N\to\infty} \frac{pN(pN-1)(pN-2)\cdots(pN-(h-1))qN(qN-1)\cdots(qN-(n-h)+1)}{N(N-1)(N-2)\cdots(N-(h-1))(N-h)\cdots(N-n+1)} =$$

$$= \lim_{N\to\infty} \frac{p^h N^h (1-\frac{1}{pN})\cdots(1-\frac{(h-1)}{pN}) q^{n-h} N^{n-h}(1-\frac{1}{qN})\cdots(1-\frac{(n-h-1)}{qN})}{N^n(1-\frac{1}{N})(1-\frac{2}{N})\cdots(1-\frac{h-1}{N})(1-\frac{h}{N})\cdots(1-\frac{n-1}{N})} =$$

$$= p^h q^{n-h} .$$

La distribuzione ipergeometrica, fissati i valori $n, p$, converge asintoticamente alla distribuzione binomiale di parametri $n, p$.

Infatti se la popolazione è molto numerosa ($N$ molto grande rispetto ad $n$), dal punto di vista numerico è praticamente la stessa cosa fare estrazioni *con* o *senza* restituzione.

**Estrazioni con restituzione da un'urna di composizione incognita.**

Consideriamo $n$ estrazioni *con restituzione* da un'urna contenente $N$ palline, di cui $r$ sono bianche, con $r$ incognito.

$E_i =$ "l'i-esima pallina estratta è bianca",
$i = 1, \dots, n$;

$X = \sum_{i=1}^{n} |E_i|$, numero aleatorio di palline bianche estratte nelle $n$ prove.

Valutiamo $P(E_i)$, per $i = 1, \dots, n$.

Introdotta la partizione $\{H_0, H_1, \dots, H_N\}$, con
$H_r =$ "*nell'urna ci sono* $r$ *palline bianche*",
per ogni $i = 1, \dots, n$, si ha

$$P(E_i|H_r) = \frac{r}{N}\,,\ r = 0, 1, \dots, N\,,$$

e quindi, per ogni $i$, si ottiene

$$P(E_i) = \sum_r P(E_i|H_r)P(H_r) = \sum_r \frac{r}{N} \cdot P(H_r)\,.$$

Come si vede, $E_1, \dots, E_n$ sono ancora equiprobabili. In particolare, se $P(H_r) = cost = \frac{1}{N+1}$, risulta

$$P(E_i) = \sum_{r=0}^{N} \frac{r}{N} \cdot \frac{1}{N+1} = \frac{1+2+\cdots+N}{N(N+1)} = \frac{1}{2}\,.$$

L'indipendenza stocastica, in generale, non sussiste più.
Ad esempio, per $N = 2$, assumendo equiprobabili le ipotesi $H_0, H_1, H_2$ e osservando che

$$P(E_1E_2 \,|\, H_r) = P(E_1|H_r)P(E_2|H_r) = \left(\tfrac{r}{N}\right)^2\,,$$

$$(\textit{indip. stocastica di } E_1, E_2 \textit{ condizionata ad } H_r)$$

si ha

$$P(E_2|E_1) = \frac{P(E_1E_2)}{P(E_1)} = \frac{\sum_{r=0}^{2} \left(\frac{r}{N}\right)^2 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{5}{6} > \frac{1}{2}\,,$$

e quindi $E_1$ ed $E_2$ sono correlati positivamente.

Infine, per ogni fissato $r$ e per ogni $h = 0, 1, \ldots, n$, si ha

$$P(X = h \,|\, H_r) = \binom{n}{h} \left(\frac{r}{N}\right)^h \left(\frac{N-r}{N}\right)^{n-h} ,$$

ovvero $X|H_r \sim B(n, \frac{r}{N})$. Allora

$$\begin{aligned} &P(X = h) = \textstyle\sum_{r=0}^{N} P(X = h \,|\, H_r) P(H_r) = \\ &= \textstyle\sum_{r=0}^{N} \binom{n}{h} \left(\frac{r}{N}\right)^h \left(\frac{N-r}{N}\right)^{n-h} P(H_r) \,, \end{aligned} \tag{19}$$

ovvero la distribuzione di probabilità di $X$ è una *mistura* di distribuzioni binomiali.

**Estrazioni senza restituzione da un'urna di composizione incognita.**

Se le estrazioni sono *senza restituzione*, per ogni $r$ e per ogni $i$, si ha ancora

$$P(E_i|H_r) = \frac{r}{N}\,, \quad P(E_i) = \sum_r \frac{r}{N} \cdot P(H_r)\,,$$

e quindi gli eventi sono sempre equiprobabili.
Anche in questo caso l'indipendenza stocastica, in generale, non sussiste più.
Ad esempio, per $N = 2$, assumendo equiprobabili le ipotesi $H_0, H_1, H_2$ e osservando che

$$P(E_1E_2 \,|\, H_r) = P(E_1|H_r)P(E_2|E_1H_r) = \frac{r(r-1)}{N(N-1)}\,,$$

(in particolare, nel nostro esempio:
$P(E_1E_2|H_0) = P(E_1E_2|H_1) = 0, P(E_1E_2|H_2) = 1$)
si ha

$$P(E_2|E_1) = \frac{\sum_{r=0}^{2} P(E_1E_2 \,|\, H_r) \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} > \frac{1}{2}\,.$$

*Nota:* $E_1$ ed $E_2$ sono ancora correlati positivamente.

Infine, per ogni $r$ e per ogni $h$, si ha

$$P(X = h \,|\, H_r) = \frac{\binom{n}{h}\binom{N-n}{r-h}}{\binom{N}{r}} \;=\; \frac{\binom{r}{h}\binom{N-r}{n-h}}{\binom{N}{n}}\,,$$

ovvero $X|H_r \sim H(N, n, \frac{r}{N})$. Allora

$$P(X = h) = \textstyle\sum_{r=0}^{N} P(X = h \,|\, H_r)P(H_r) =$$

$$= \textstyle\sum_{r=0}^{N} \frac{\binom{n}{h}\binom{N-n}{r-h}}{\binom{N}{r}} P(H_r) = \sum_{r=0}^{N} \frac{\binom{r}{h}\binom{N-r}{n-h}}{\binom{N}{n}} P(H_r)\,, \tag{20}$$

ovvero la distribuzione di probabilità di $X$ è una *mistura* di distribuzioni ipergeometriche.

**Esercizio:** Si considerino 2 estrazioni senza restituzione da un'urna contenente 2 palline, di cui $r$ bianche, con $r$ incognito. Posto $P(H_r) = \alpha_r$, con $\alpha_0 + \alpha_1 + \alpha_2 = 1\,,\ \alpha_r < 1\,, \forall\, r$, si può verificare che

$$P(E_1) = \frac{1}{2}\alpha_1 + \alpha_2\,, \quad P(E_2|E_1) = \frac{\alpha_2}{\frac{1}{2}\alpha_1 + \alpha_2}\,.$$

Si considerino le funzioni

$$f(\alpha_2) = 2(\sqrt{\alpha_2} - \alpha_2) = 2\sqrt{\alpha_2}(1 - \sqrt{\alpha_2})\,,$$

$$g(\alpha_2) = f(\alpha_2) + \alpha_2\,, \quad 0 \leq \alpha_2 \leq 1\,.$$

Si ha

$$0 \leq f(\alpha_2) \leq \frac{1}{2}\,, \quad 0 \leq g(\alpha_2) \leq 1\,.$$

Come si può verificare, per $\alpha_1 = f(\alpha_2)$, risulta

$$P(E_1) = P(E_2) = \frac{1}{2}\alpha_1 + \alpha_2 = \sqrt{\alpha_2}\,,$$

e quindi

$$P(E_2|E_1) = \sqrt{\alpha_2} = P(E_2) = P(E_1)\,.$$

Inoltre, per $0 \leq \alpha_1 < f(\alpha_2)$, si ha

$$P(E_2) < \sqrt{\alpha_2} < P(E_2|E_1)\,;$$

mentre, per $f(\alpha_2) < \alpha_1 \leq 1$, si ha

$$P(E_2) > \sqrt{\alpha_2} > P(E_2|E_1)\,.$$

Ad esempio, se

$$\alpha_2 = \tfrac{1}{k^2}\,, \quad \alpha_1 = 2\sqrt{\alpha_2}(1-\sqrt{\alpha_2}) = 2\,\tfrac{1}{k}\,\left(1-\tfrac{1}{k}\right)\,,$$

$$\alpha_0 = \left(1-\tfrac{1}{k}\right)^2\,,$$

si ha: $P(E_2|E_1) = P(E_2) = \frac{1}{k}$.

*Esercizio.* Dimostrare che, nel caso con restituzione, si ha

$$P(E_1) = P(E_2) = \frac{1}{2}\alpha_1 + \alpha_2\,, \;\; P(E_2|E_1) = \frac{\frac{1}{4}\alpha_1 + \alpha_2}{\frac{1}{2}\alpha_1 + \alpha_2}\,,$$

e quindi:

(i) se $\alpha_1 = 1$ (per cui $\alpha_0 = \alpha_2 = 0$), segue

$$P(E_2|E_1) = \frac{1}{2} = P(E_2) = P(E_1)\,;$$

(ii) se $\alpha_2 = 1$ (per cui $\alpha_0 = \alpha_1 = 0$), segue

$$P(E_2|E_1) = 1 = P(E_2) = P(E_1)\,;$$

(iii) se $0 < \alpha_1 < 1\,,\ 0 < \alpha_2 < 1$, posto $\beta = \alpha_1 + \alpha_2$, si può verificare che

$$\beta^2 + (2\alpha_2 - 1)\beta + \alpha_2^2 - 3\alpha_2 < 0\,,\ \ \forall \beta \in (0,1)\,,$$

che equivale a

$$P(E_2|E_1) = \frac{\frac{1}{4}\alpha_1 + \alpha_2}{\frac{1}{2}\alpha_1 + \alpha_2}\ >\ \tfrac{1}{2}\alpha_1 + \alpha_2 = P(E_2)\,,$$

$$\forall\,(\alpha_1, \alpha_2) \in (0,1) \times (0,1)\,.$$

*Problema dei carcerati.* Considerati $n$ prigionieri $a_1, \ldots, a_n$ condannati a morte, si supponga che uno di loro scelto a caso sia stato graziato. Definiti gli eventi

$$A_i = \text{"il prigioniero graziato è } a_i \,,\; i = 1, \ldots, n \,,$$

si ha $P(A_i) = \frac{1}{n}$, $\forall\, i$.
Supponiamo che il carceriere, su richiesta di $a_1$, vada a guardare i documenti applicando la seguente procedura:
(i) se il prigioniero graziato è $a_i$, con $i > 1$, il carceriere riferisce ad $a_1$ che $a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$ non sono stati graziati;
(ii) se il prigioniero graziato è $a_1$, il carceriere esclude a caso un prigioniero $a_i$ e riferisce ad $a_1$ che $a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$ non sono stati graziati.
Si consideri la partizione $\{H_2, \ldots, H_n\}$, dove $H_i =$ "il carceriere riferisce ad $a_1$ che

$$a_2 \,,\; \ldots \,,\; a_{i-1} \,,\; a_{i+1} \,,\; \ldots \,,\; a_n$$

non sono stati graziati", $i = 2, \ldots, n$.
Calcolare $P(A_1|H_i)\,,\; i = 2, \ldots, n$.

*Nota:* con un ragionamento "intuitivo" si potrebbe essere indotti a dire che, se il carceriere riferisce ad $a_1$ che $a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$ non sono stati graziati, essendo rimasti solo in due ($a_1$ ed $a_i$) a poter essere graziati, la probabilità di $A_1$ (condizionata all'informazione data dal carceriere) è pari ad $\frac{1}{2}$.

È corretta tale valutazione ?

...... *la risposta è no!*

Infatti, $\frac{1}{2}$ è la probabilità dell'evento condizionato $A_1|(A_1 \vee A_i)$, il quale non coincide con $A_1|H_i$, in quanto $H_i \neq A_1 \vee A_i$.
In particolare $H_i \Rightarrow A_1 \vee A_i$, mentre $A_1 \vee A_i \not\Rightarrow H_i$.

Si può verificare che, per ogni $i = 2, \ldots, n$, risulta

$$P(H_i) = \frac{1}{n-1}, \quad P(A_1|H_i) = P(A_1) = \frac{1}{n}.$$

Infatti: $P(H_i|A_1) = \frac{1}{n-1}$, $P(H_i|A_i) = 1$,

$P(H_i|A_j) = 0\,,\ j > 1\,,\ j \neq i\,;$

allora

$$P(H_i) = \sum_j P(H_i|A_j)P(A_j) =$$

$$= P(H_i|A_1)P(A_1) + P(H_i|A_i)P(A_i) =$$

$$= \tfrac{1}{n-1} \cdot \tfrac{1}{n} + 1 \cdot \tfrac{1}{n} = \tfrac{1}{n-1} = P(H_i|A_1)\,;$$

quindi

$$P(A_1|H_i) = \tfrac{P(A_1 H_i)}{P(H_i)} = \tfrac{P(A_1)P(H_i|A_1)}{P(H_i)} = P(A_1) = \tfrac{1}{n}\,.$$

*Nota:* esaminare, in particolare, il caso $n = 3$, ponendo in generale $P(H_2|A_1) = p\,,\ P(H_3|A_1) = 1-p$, e verificando che

$$P(A_1|H_2) = \frac{p}{1+p} \in \left[0, \frac{1}{2}\right]\,,$$

con $P(A_1|H_2) = \frac{1}{3} = P(A_1)$, per $p = \frac{1}{2}$.

## Numeri aleatori discreti

Un numero aleatorio $X$ si dice discreto se l'insieme dei suoi possibili valori è finito o infinito numerabile, cioè

$$X \in \{x_1, x_2, \ldots, x_n, \ldots\}\,.$$

Sia $\{p_i\,,\ i \in \mathbb{N}\}$, con $p_i = P(X = x_i)$, la distribuzione di probabilità di $X$ e supponiamo che

$$\textstyle\sum_{i=1}^{\infty} p_i = 1\,, \quad \sum_{i=1}^{\infty} |x_i| p_i < \infty\,.$$

In questo caso la previsione si definisce nel seguente modo

$$\mathbb{P}(X) = \sum_{i=1}^{\infty} p_i x_i\,. \tag{21}$$

La funzione

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p_i\,, \quad x \in \mathbb{R}\,, \tag{22}$$

si chiama funzione di ripartizione di $X$.
Osserviamo che $F(x) \in [0,1]\,, \forall\, x \in \mathbb{R}$.

**Esempio.** La f.d.r di $X = |E|$, con $p = P(E)$, è

$$F(x) = \begin{cases} 0\,, & x < 0\,, \\ q\,, & 0 \le x < 1\,, \\ 1\,, & 1 \le x\,. \end{cases}$$

Basta osservare che per $x < 0$ si ha $(X \le x) = \emptyset$, mentre per $x \ge 1$ si ha $(X \le x) = \Omega$. Inoltre, per $0 \le x < 1$ si ha $(X \le x) = (X = 0)$. In particolare

$$P(X \le 0) = \overbrace{P(X = 0)}^{=P(E^c)} + \overbrace{P(X < 0)}^{=0} = P(E^c) = q$$
$$P(X \le 1) = P(X = 0) + P(X = 1) = q + p = 1$$

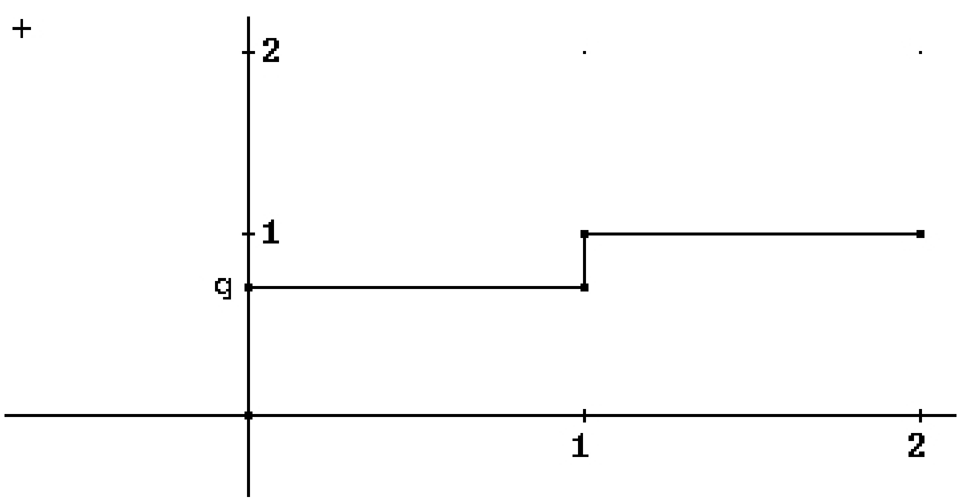



Figura 3: funzione di ripartizione di $|E|$

**Distribuzione Geometrica.**

$E_1, E_2, \ldots, E_n, \ldots,$ successione di eventi indipendenti ed equiprobabili, con $P(E_i) = p$;

$X$, n.a. di prove fino al primo successo;

$X$ è un n.a. discreto, infatti $X \in \mathbb{N}$.
(a rigore, andrebbe incluso $X = \infty$, che significa sempre insuccesso; come vedremo tale evento ha probabilità nulla).

Si ha

$$P(X = n) = q^{n-1}p, \quad n \in \mathbb{N}.$$

Infatti

$$P(X = 1) = P(E_1) = p\,,$$

$$P(X = 2) = P(E_1^c E_2) = qp\,,$$

..................................

$$P(X = n) = P(E_1^c E_2^c \cdots E_{n-1}^c E_n) = pq^{n-1}\,,$$

..................................

$X$ ha una distribuzione geometrica di parametro $p$, che si indica con

$$X \sim \mathbf{G}(p)\,.$$

La successione $\{p_n\}$, dove $p_n = P(X = n)$, è geometrica di ragione $(1-p)$. Infatti

$$p_1 = p \geq p_2 = p(1-p) \geq p_3 = p(1-p)^2 \geq \cdots\,.$$

$(X = 1)$ (successo alla prima prova) è il costituente più probabile.

Inoltre vale la $\sigma-$additività:

$$\textstyle\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} p(1-p)^{n-1} =$$

$$\textstyle= p\sum_{n=1}^{\infty}(1-p)^{n-1} = p\frac{1}{1-(1-p)} = 1\,.$$

Osserviamo che per ogni fissato $n$ si ha

$$(X > n) = E_1^c E_2^c \cdots E_n^c\,,$$

quindi

$$P(X > n) = q^n \, .$$

Inoltre $(X = \infty) \subseteq (X > n)$, $\forall\, n$, da cui segue $P(X = \infty) \leq P(X > n)$, $\forall\, n$.

Allora, assumendo $p > 0$, segue $q^n \to 0$ e quindi $P(X = \infty) = 0$.

**Proprietà di assenza di memoria.**
Un n. a. $X$, discreto e non negativo, ha una distribuzione geometrica se e solo se

$$P(X > n + m \,|\, X > m) = P(X > n), \;\; \forall\, n, m \in \mathbb{N} \, . \tag{23}$$

(*Significato:* "supposto di non aver avuto successo fino alla prova $m-$esima, la probabilità di non avere successo nelle successive $n$ prove è la stessa di non avere successo nelle prime $n$ prove".

dim. $(\Rightarrow)$ $Hp)$ $X \sim \mathbf{G(p)}$; $\quad Th)$ vale la (23).

$$P(X > n+m | X > m) = \tfrac{P(X>n+m, X>m)}{P(X>m)} =$$

$$= \tfrac{P(X>n+m)}{P(X>m)} = \tfrac{q^{n+m}}{q^m} = q^n = P(X > n)\,.$$

$(\Leftarrow)$ $Hp)$ vale la (23); $\quad Th)$ $X \sim \mathbf{G(p)}$.

Dalla (23) segue

$$P(X > n+m) = P(X > n) P(X > m)\,.$$

Introduciamo la funzione di sopravvivenza definita come

$$S(x) = 1 - F(x) = P(X > x)\,. \tag{24}$$

Per $m = 1$ si ha

$P(X > n+1) = P(X > 1) P(X > n), \;\; \forall n \in \mathbb{N}\,;$
allora
$S(n+1) = S(1)S(n) = [S(1)]^2 S(n-1) =$
$= \cdots = [S(1)]^{n+1}\,.$
Posto $p = P(X = 1)\,, q = 1 - p = P(X > 1) = S(1)$,
segue $S(n+1) = q^{n+1}$.

Osservando che

$$(X > n) \; = \; (X = n+1) \vee (X > n+1) \, ,$$

otteniamo

$$P(X = n+1) = P(X > n) - P(X > n+1) =$$

$$= S(n) - S(n+1) = q^n(1-q) = pq^n \, .$$

**Previsione e Varianza.**

$$\mathbb{P}(X) = \textstyle\sum_{n=1}^{\infty} npq^{n-1} = p \sum_{n=1}^{\infty} \frac{d(q^n)}{dq} =$$

$$= p \tfrac{d}{dq} \textstyle\sum_{n=1}^{\infty} q^n = \cdots = \frac{1}{p} \, .$$

Con un ragionamento analogo si può provare che

$$\mathbb{P}(X^2) = \textstyle\sum_{n=1}^{\infty} n^2 pq^{n-1} = \frac{1+q}{p^2} \, ,$$

$$Var(X) = \mathbb{P}(X^2) - [\mathbb{P}(X)]^2 = \tfrac{1+q}{p^2} - \tfrac{1}{p^2} = \tfrac{q}{p^2} \, .$$

## Distribuzione di Pascal

$E_1, E_2, \ldots, E_n, \ldots$, successione di eventi indipendenti ed equiprobabili, con $P(E_i) = p$;

$X$, n.a. di prove fino al $k^0$ successo;

$X \in \{k, k+1, \ldots, n, \ldots\}$.

Posto $P(X = n) = p_n$, si ha

$$p_k = P(E_1 E_2 \cdots E_k) = p^k \,,$$

$$p_{k+1} = P(E_1^c E_2 \cdots E_{k+1}) + \cdots + P(E_1 E_2 \cdots E_k^c E_{k+1}) =$$

$$= k p^k q \,,$$

..........................................................................

$$p_n = P(E_1^c \cdots E_{n-k}^c E_{n-k+1} \cdots E_n) + \cdots$$

$$\cdots + P(E_1 \cdots E_{k-1} E_k^c \cdots E_{n-1}^c E_n) = \binom{n-1}{k-1} p^k q^{n-k} \,,$$

..................................................................................

La probabilità $p_n$ si ottiene osservando che:
- i costituenti favorevoli all'evento $(X = n)$ sono tutti quelli del tipo

$$E_{i_1} \cdots E_{i_{k-1}} E_{i_k}^c \cdots E_{i_{n-1}}^c E_n \,,$$

dove $(i_1, \ldots, i_{n-1})$ è una generica permutazione di $(1, \ldots, n-1)$;
- di tali costituenti ce ne sono $\binom{n-1}{k-1}$;
- la probabilità di ciascuno di essi è $p^k q^{n-k}$.

$X$ ha una distribuzione di Pascal, di parametri $k, p$, che si indica con: $X \sim \mathbf{P}(k, p)$.

$X$ si può rappresentare come somma di numeri aleatori $X_1, \ldots, X_k$, dove:
$X_1$ è il n.a. di prove fino al $1^0$ successo;

$X_2$ è il numero di prove fino al $2^0$ successo, contate a partire dalla prima dopo il $1^0$ successo;

.............................................................

$X_k$ è il numero di prove fino al $k^0$ successo, contate a partire dalla prima dopo il $(k-1)^0$ successo.

$X_1, \dots, X_k$ hanno tutti una distribuzione geometrica di parametro $p$ e quindi

$$\mathbb{P}(X) = \mathbb{P}(X_1 + \cdots + X_k) = \mathbb{P}(X_1) + \cdots + \mathbb{P}(X_k) = \frac{k}{p} .$$

Inoltre, sfruttando l'indipendenza degli eventi $E_1.E_2, \dots, E_n, \dots$ si potrebbe verificare che

$$Var(X) = Var(X_1) + \cdots + Var(X_k) = \frac{kq}{p^2} .$$

## Distribuzione di Poisson

Un n.a. $X$ ha una distribuzione di Poisson di parametro $\lambda$, con $\lambda \in \mathbb{R}^+$, indicata con $X \sim \mathcal{P}(\lambda)$, se

$$1)\ X \in \mathbb{N}_0 = \{0, 1, 2, 3, \dots, n, \dots\};$$

$$2)\ P(X = n) = p_n = \tfrac{\lambda^n}{n!} e^{-\lambda},\ \ \forall\, n \in \mathbb{N}_0\,.$$

Si ha

$$\textstyle\sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1\,.$$

Nella distribuzione di Poisson la previsione e la varianza coincidono. Infatti

$$\mathbb{P}(X) = \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n \frac{\lambda^n}{n!} e^{-\lambda} = \cdots = \lambda\,,$$

$$\mathbb{P}(X^2) = \textstyle\sum_{i=0}^{\infty} n^2 p_n = \cdots = \lambda^2 + \lambda\,,$$

$$Var(X) = \lambda^2 + \lambda - \lambda^2 = \lambda\,.$$

La distribuzione di Poisson si usa in una grande varietà di applicazioni, ad esempio (sotto certe condizioni) si utilizza come distribuzione di probabilità di un n.a. $X$ che (in un certo *intervallo di tempo*) rappresenta

1. il num. di telefonate che giungono ad un centralino;

2. il num. di clienti che si presentano ad uno sportello;

3. il num. di automobili che transitano ad un casello autostradale;

4. il num. di particelle emesse da un corpo radioattivo;

5. ...................................

Una delle prime applicazioni si ebbe nell'esercito prussiano, per valutare la probabilità, molto piccola, che un soldato dell'esercito morisse in seguito ad un calcio di cavallo.

Osserviamo che, sotto certe ipotesi, la distribuzione Binomiale si può approssimare con la distribuzione di Poisson.

**Esempio**
(*Scelta a caso di* $n$ *punti in un intervallo* $[0, a]$)
Sia $I$ un sottointervallo di $[0, a]$ di ampiezza $t$, ad esempio $[0, t]$.
Indicando con $X_1, X_2, \ldots, X_n$ le ascisse aleatorie dei punti generati nell'esperimento, definiamo gli eventi

$$E_i = (X_i \in I), \quad i = 1, 2, \ldots, n.$$

Se i punti sono scelti a caso, si può valutare (come si vedrà nel caso continuo con la distribuzione uniforme)

$$P(E_i) = \frac{t}{a}, \quad i = 1, 2, \ldots, n,$$

(*la probabilità è il rapporto fra le ampiezze degli intervalli*).

Gli eventi $E_i$ si possono ritenere indipendenti ed equiprobabili.

Posto $Y_n = \sum_{i=1}^{n} |E_i|$, si ha

$$Y_n \sim \mathbf{B}(n, \frac{t}{a}), \;\; \forall n \in \mathbb{N}.$$

Al variare di $n$ si ha una successione di n.a. $\{Y_n\}$ con distribuzione binomiale.

Facciamo le seguenti assunzioni:

1. $a \to \infty$;

2. manteniamo fissa l'ampiezza $t$ di $I$;
(in questo modo, la probabilità di successo nella singola prova, $t/a$, diventa sempre più piccola e i successi diventano *eventi rari*);

3. Facciamo tendere $n$ a $+\infty$ in modo tale che il rapporto $\frac{n}{a}$ si mantenga costante ($\frac{n}{a} = cost = \alpha$).

Per la previsione di $Y_n$ si ha

$$\mathbb{P}(Y_n) = n \cdot \frac{t}{a} = \alpha t = cost = \lambda, \;\; \forall n\,.$$

Si ottiene

$$\lim_{n\to\infty} P(Y_n = h) = \lim_{n\to\infty} \binom{n}{h} (\tfrac{t}{a})^h (1 - \tfrac{t}{a})^{n-h} =$$

$$= \lim_{n\to\infty} \tfrac{n(n-1)\cdots(n-h+1)}{h!} (\tfrac{\lambda}{n})^h (1 - \tfrac{\lambda}{n})^{n-h} = \tfrac{\lambda^h}{h!} e^{-\lambda} \,.$$

Pertanto: $\mathbf{B}(n, \frac{\lambda}{n}) \to_{n\to\infty} \mathcal{P}(\lambda)\,,$

cioè, la distribuzione binomiale (sotto certe condizioni) si può approssimare con la distribuzione di Poisson.

**Altro esempio***(distribuzione a caso di $n$ palline in $n$ urne)*

..............................

$X =$ "numero aleatorio di palline che capitano nella prima urna";

$X \sim$ ? Distribuzione asintotica di $X$ ?

**Esempio.**

Un'azienda possiede un parco autobus costituito da 2000 unità. Da considerazioni statistiche si sa che in media il numero di autobus che vengono trovati guasti ogni mattina è pari a 2.

Indichiamo con $X$ il numero aleatorio di autobus guasti in una mattina specifica.

Se, per ogni $i = 1, 2, \ldots, 2000$, definiamo l'evento $E_i =$ *"l'i-mo autobus è difettoso"*, possiamo ritenere gli eventi $E_i$ equiprobabili e indipendenti, con $P(E_i) = p$.

Si ha quindi $X \sim \mathbf{B}(2000, p)$.

Calcolare $p$ e $P(X > 1)$.

## Eventi scambiabili

La nozione di scambiabilità gioca un ruolo fondamentale per mostrare sotto quali condizioni sia giustificato valutare la probabilità mediante la frequenza osservata.

**Definizione 1** Gli eventi $E_1, \ldots, E_n$ sono scambiabili se, per ogni fissato $h$, con $0 \leq h \leq n$, i costituenti del tipo

$$E_{i_1} \cdots E_{i_h} E^c_{i_{h+1}} \cdots E^c_{i_n} \qquad (25)$$

hanno la stessa probabilità (dipendente solo da $n$ e da $h$), per ogni permutazione $(i_1, \ldots, i_n)$ di $(1, \ldots, n)$.

Definiamo $P(E_{i_1} \cdots E_{i_h} E^c_{i_{h+1}} \cdots E^c_{i_n}) = p_{n,h}$.
Consideriamo i quattro casi di estrazioni con o senza restituzione da un'urna di composizione nota o incognita, contenente $N$ palline, e indichiamo con $E_i$ l'evento *"l'i-ma pallina estratta è bianca"*. Per semplicità , indichiamo con lo stesso simbolo gli eventi e i corrispondenti indicatori. Si può verificare che, in tutti e quattro i casi, gli eventi $E_1, \ldots, E_N$ sono scambiabili. Infatti:
- nel primo caso (*estraz. con rest. da un'urna di comp. nota*) si ha:

$$p_{n,h} = p^h q^{n-h}\,, \;\; \forall\,(i_1, i_2, \ldots, i_n)\,, \;\; 1 \leq n \leq N\,;$$

- nel secondo caso (*estraz. senza rest. da un'urna di* *comp. nota*) si ha:

$$p_{n,h} = \frac{\binom{N-n}{pN-h}}{\binom{N}{pN}}\,, \quad \forall\,(i_1, i_2, \ldots, i_n)\,, \quad 1 \le n \le N\,;$$

- nel terzo caso (*estraz. con rest. da un'urna di comp. incognita*) si ha:

$$p_{n,h} = \sum_{r=0}^{N} \left(\frac{r}{N}\right)^{h} \left(\frac{N-r}{N}\right)^{n-h} P(H_r)\,,$$

$$\forall\,(i_1, i_2, \ldots, i_n)\,, \quad 1 \le n \le N\,;$$

- nel quarto caso (*estraz. senza rest. da un'urna di comp. incognita*) si ha:

$$p_{n,h} = \sum_{r=0}^{N} \frac{\binom{N-n}{r-h}}{\binom{N}{r}} P(H_r)\,,$$

$$\forall\,(i_1, i_2, \ldots, i_n)\,, \quad 1 \le n \le N\,.$$

Come si vede, in tutti e quattro i casi $p_{n,h}$ non dipende dalla permutazione $(i_1, i_2, \ldots, i_n)$ degli indici $(1, 2, \ldots, n)$.

Posto $S_n = E_1 + \cdots + E_n$ , $P(S_n = h) = \omega_{n,h}$, si ha

**Teorema 1** $E_1, \ldots, E_n$ sono scambiabili se e solo se, per ogni $h$, vale

$$p_{n,h} = \frac{\omega_{n,h}}{\binom{n}{h}} . \tag{26}$$

*Dim.* $(S_n = h)$ è l'unione di $\binom{n}{h}$ costituenti del tipo (25). Se gli eventi sono scambiabili, tali costituenti hanno la stessa probabilità , $p_{n,h}$, e quindi

$$\omega_{n,h} = \binom{n}{h} p_{n,h} .$$

Viceversa, se vale (26), i costituenti del tipo (25) hanno la stessa probabilità , $p_{n,h}$, e quindi gli eventi sono scambiabili.

Una definizione equivalente di scambiabilità è la seguente:

**Definizione 2** Gli eventi $E_1, \ldots, E_n$ sono scambiabili se, per ogni fissato $h$, con $1 \leq h \leq n$, gli eventi del tipo

$$E_{i_1} \cdots E_{i_h} \tag{27}$$

hanno la stessa probabilità (dipendente solo da $h$), per ogni sottoinsieme $\{i_1, \ldots, i_h\} \subseteq \{1, \ldots, n\}$.

Infatti, si può dimostrare il seguente

**Teorema 2** Le Definizioni 1 e 2 sono equivalenti.

*Dim. (cenno).* Assumendo $E_1, \ldots, E_n$ scambiabili secondo la Definizione 1 e osservando che

$$(E_{i_{h+1}} \vee E^c_{i_{h+1}}) \cdots (E_{i_n} \vee E^c_{i_n}) = \Omega \,,$$

si ha

$$E_{i_1} \cdots E_{i_h} = E_{i_1} \cdots E_{i_h}(E_{i_{h+1}} \vee E^c_{i_{h+1}}) \cdots (E_{i_n} \vee E^c_{i_n}) =$$

$$= \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \,.$$

Sviluppando l'espressione precedente, per ogni $k = 0, 1, \ldots, n-h$, si ottengono $\binom{n-h}{k}$ costituenti del tipo:

$$E_{i_1} \cdots E_{i_h} E^c_{j_1} \cdots E^c_{j_k} E_{j_{k+1}} \cdots E_{j_{n-h}} ,$$

favorevoli a $E_{i_1} \cdots E_{i_h}$. Allora, definendo

$$P(E_{i_1} \cdots E_{i_h}) = q_h ,$$

risulta

$$P(E_{i_1} \cdots E_{i_h}) = q_h = p_{n,n} + (n-h)p_{n,n-1} + \cdots$$

$$\cdots + \tbinom{n-h}{k} p_{n,n-k} + \cdots + p_{n,h} ,$$

per ogni $\{i_1, \ldots, i_h\} \subseteq \{1, \ldots, n\}$; pertanto $E_1, \ldots, E_n$ sono scambiabili anche secondo la Definizione 2.

Viceversa, assumendo $E_1, \ldots, E_n$ scambiabili secondo la Definizione 2, si ha (utilizzando lo stesso simbolo per gli eventi e i loro indicatori)

$$E^c_{i_{h+1}} \cdots E^c_{i_n} = (1 - E_{i_{h+1}}) \cdots (1 - E_{i_n}) =$$

$$= 1 - \sum_{k=h+1}^n E_{i_k} + \sum_{r>k=h+1}^n E_{i_k} E_{i_r} - \cdots$$

$$\cdots + (-1)^{n-h} E_{i_{h+1}} \cdots E_{i_n} ,$$

e quindi

$$E_{i_1} \cdots E_{i_h} E^c_{i_{h+1}} \cdots E^c_{i_n} = \cdots =$$

$$= E_{i_1} \cdots E_{i_h} - \sum_{k=h+1}^n E_{i_1} \cdots E_{i_h} E_{i_k} +$$

$$+ \sum_{r>k=h+1}^n E_{i_1} \cdots E_{i_h} E_{i_k} E_{i_r} - \cdots$$

$$\cdots + (-1)^{n-h} E_{i_1} \cdots E_{i_h} E_{i_{h+1}} \cdots E_{i_n} ;$$

pertanto, si ha

$$P(E_{i_1} \cdots E_{i_h} E^c_{i_{h+1}} \cdots E^c_{i_n}) = p_{n,h} =$$

$$= q_h - (n-h) q_{h+1} + \binom{n-h}{2} q_{h+2} - \cdots + (-1)^{n-h} q_n ,$$

per ogni $\{i_1, \ldots, i_h\} \subseteq \{1, \ldots, n\}$; pertanto $E_1, \ldots, E_n$ sono scambiabili anche secondo la Definizione 1.

Riassumendo:

$$q_h = \sum_{k=0}^{n-h} \binom{n-h}{k} p_{n,n-k} = \sum_{k=0}^{n-h} \frac{\binom{n-h}{k}}{\binom{n}{n-k}} \omega_{n,n-k}\,;$$

$$\omega_{n,h} = \binom{n}{h} p_{n,h} = \binom{n}{h} \sum_{k=0}^{n-h} (-1)^k \binom{n-h}{k} q_{h+k}\,.$$

**Corollario.** Dati $n$ eventi $E_1, \ldots, E_n$ scambiabili, per ogni $\{i_1, \ldots, i_h\} \subset \{1, \ldots, n\}$ gli eventi $E_{i_1}, \ldots, E_{i_h}$ sono scambiabili.

La dimostrazione segue dalla Definizione 2.

*Nota:* Se gli eventi $E_1, \ldots, E_h$ sono scambiabili, non è detto che $E_1, \ldots, E_h, \ldots, E_n$ siano scambiabili, come mostrato nell'esempio che segue.

*Esempio.* Da un'urna contenente una pallina bianca e una nera si effettuano due estrazioni senza restituzione. Gli eventi $E_1 =$ *"la prima pallina estratta è bianca"*, $E_2 =$ *"la seconda pallina estratta è bianca"*, sono scambiabili e si ha

$$P(E_1) = P(E_2) = P(E_1 E_2^c) = P(E_1^c E_2) = \tfrac{1}{2},$$

$$P(E_1 E_2) = P(E_1^c E_2^c) = 0;$$

quindi, per ogni altro evento $E_3$, dalle relazioni

$$P(E_1^c E_2^c) = P(E_1^c E_2^c E_3^c) + P(E_1^c E_2^c E_3),$$

$$P(E_1 E_2) = P(E_1 E_2 E_3^c) + P(E_1 E_2 E_3),$$

segue

$$P(E_1^c E_2^c E_3^c) = P(E_1^c E_2^c E_3) = P(E_1 E_2 E_3^c) =$$

$$= P(E_1 E_2 E_3) = 0.$$

Se $E_1, E_2, E_3$ fossero scambiabili, si avrebbe

$$P(E_1^c E_2^c E_3) = P(E_1 E_2 E_3) = P(E_1 E_2^c E_3) =$$

$$= P(E_1^c E_2 E_3) = 0,$$

ed essendo

$$E_3 = E_1 E_2 E_3 \vee E_1 E_2^c E_3 \vee E_1^c E_2 E_3 \vee E_1^c E_2^c E_3,$$

si otterrebbe $P(E_3) = 0 \neq \frac{1}{2} = P(E_1) = P(E_2)$.

Pertanto, in relazione ai due eventi scambiabili $E_1, E_2$, non esiste alcun evento $E_3$ tale che $E_1, E_2, E_3$ sono scambiabili.

In generale, quindi, degli eventi possono essere scambiabili ad $n$ ad $n$, senza esserlo ad $n+1$ ad $n+1$.

Mostriamo un altro esempio con $n = 2$.

*Esempio.* Dati tre eventi $E_1, E_2, E_3$, supponiamo che le probabilità dei costituenti

$$E_1^c E_2^c E_3^c\,,\quad E_1 E_2^c E_3^c\,,\quad E_1^c E_2 E_3^c\,,\quad E_1^c E_2^c E_3\,,$$

$$E_1 E_2 E_3^c\,,\quad E_1 E_2^c E_3\,,\quad E_1^c E_2 E_3\,,\quad E_1 E_2 E_3\,,$$

siano valutate rispettivamente

$\frac{1}{10}\,,\quad \frac{1}{10}\,,\quad \frac{1}{20}\,,\quad \frac{1}{20}\,,\quad \frac{3}{20}\,,\quad \frac{3}{20}\,,\quad \frac{1}{5}\,,\quad \frac{1}{5}.$

Allora

$P(E_1 E_2^c) = P(E_1^c E_2) = \frac{1}{4}\,,\quad P(E_1 E_3^c) =$

$= P(E_1^c E_3) = \frac{1}{4}\,,\quad P(E_2 E_3^c) = P(E_2^c E_3) = \frac{1}{5}\,;$

pertanto $E_i, E_j$ sono scambiabili, per ogni $\{i, j\} \subset \{1, 2, 3\}$, ma $E_1, E_2, E_3$ non sono scambiabili.

In questo esempio si ha:

$$P(E_1) = P(E_2) = P(E_3) = \tfrac{3}{5}\,,$$
$$P(E_1 E_2) = P(E_1 E_3) = \tfrac{7}{20} \neq P(E_2 E_3) = \tfrac{2}{5}\,.$$

Dati $N$ eventi $E_1, \ldots, E_N$ ed un sottoinsieme $I_n = \{i_1, \ldots, i_n\} \subseteq J_N = \{1, \ldots, N\}$, con $1 \leq n \leq N$, consideriamo i seguenti numeri aleatori (*frequenze assolute di successo*)

$$S_n = \sum_{r=1}^{n} E_{i_r}\,, \qquad S_{N\setminus n} = \sum_{j \in J_N \setminus I_n} E_j\,,$$

$$S_N = \sum_{i=1}^{N} E_i = S_n + S_{N\setminus n}\,.$$

Osserviamo che il numero di costituenti favorevoli all'evento $(S_N = k)$ è pari a $\binom{N}{k}$, mentre il numero di costituenti favorevoli all'evento $(S_n = h) \wedge (S_N = k)$, con $0 \leq h \leq k$, è pari a $\binom{n}{h}\binom{N-n}{k-h}$.
Inoltre (*formula di disintegrazione o teorema delle probabilità totali*)

$$P(S_n = h) = \sum_{k=h}^{h+N-n} P(S_n = h \,|\, S_N = k) P(S_N = k)\,;$$

ovvero

$$\omega_{n,h} = \sum_{k=h}^{h+N-n} P(S_n = h \,|\, S_N = k)\, \omega_{N,k}\,.$$

Si ha il seguente

**Teorema 3** Se $E_1, \ldots, E_N$ sono scambiabili, segue, per ogni $0 \leq h \leq n\,,\ h \leq k \leq N-(n-h)$,

$$P(S_n = h \,|\, S_N = k) = \frac{\binom{n}{h}\binom{N-n}{k-h}}{\binom{N}{k}} = \frac{\binom{k}{h}\binom{N-k}{n-h}}{\binom{N}{n}}\,.$$

*Dim.* Essendo gli eventi scambiabili, ogni costituente relativo ad $E_1, \ldots, E_N$ contenente $k$ successi ed $N-k$ insuccessi ha una probabilità pari a $p_{N,k}$. Pertanto

$$P(S_n = h \,|\, S_N = k) = \tfrac{P[(S_n=h)\wedge(S_N=k)]}{P(S_N=k)} =$$

$$\tfrac{\binom{n}{h}\binom{N-n}{k-h}p_{N,k}}{\binom{N}{k}p_{N,k}} = \tfrac{\binom{n}{h}\binom{N-n}{k-h}}{\binom{N}{k}} = \tfrac{\binom{k}{h}\binom{N-k}{n-h}}{\binom{N}{n}}\,.$$

In particolare, ponendo

$$C_{n,h} = E_{i_1} \cdots E_{i_h} E^c_{i_{h+1}} \cdots E^c_{i_n} \,,$$

si ha

$$P(C_{n,h} \,|\, S_N = k) = \frac{P[C_{n,h} \wedge (S_N = k)]}{P(S_N = k)} = \frac{\binom{N-n}{k-h} p_{N,k}}{\binom{N}{k} p_{N,k}} = \frac{\binom{N-n}{k-h}}{\binom{N}{k}} \,.$$

*Nota:* come si vede, il numero aleatorio condizionato $S_n \,|\, (S_N = k)$ ha una distribuzione di probabilità ipergeometrica $H(N, n, p)$, con $p = \frac{k}{N}$.

Dal Teorema 3 segue immediatamente

**Teorema 4** Se $E_1, \ldots, E_N$ sono scambiabili, allora per ogni $0 \leq h \leq n\,,\ 1 \leq n \leq N$ vale

$$\begin{aligned} \omega_{n,h} &= \sum_{k=h}^{h+N-n} \frac{\binom{k}{h}\binom{N-k}{n-h}}{\binom{N}{n}} \, \omega_{N,k} = \\ &= \binom{n}{h} \sum_{k=h}^{h+N-n} \frac{\binom{N-n}{k-h}}{\binom{N}{k}} \, \omega_{N,k} \,. \end{aligned} \tag{28}$$

In particolare,

$$P(C_{n,h}) = p_{n,h} = \frac{\omega_{n,h}}{\binom{n}{h}} = \sum_{k=h}^{h+N-n} \frac{\binom{N-n}{k-h}}{\binom{N}{k}} \omega_{N,k}\,. \quad (29)$$

Cosa si può dire degli eventi $E_{n+1}, \ldots, E_N$ avendo osservato gli eventi $E_1, \ldots, E_n$ ?
In particolare, come si modificano le valutazioni di probabilità per le possibili frequenze "future" $S_N = k$, avendo osservato una frequenza "passata" $S_n = h$?
Si ha

**Teorema 5** Se $E_1, \ldots, E_N$ sono scambiabili, allora

$$\forall\, 0 \leq h \leq n,\ h \leq k \leq N,\ n-h \leq N-k,\ 1 \leq n \leq N,$$

indicando con $\{i_1, \ldots, i_n\}$ una permutazione di $\{1, \ldots, n\}$, con $\{i_{n+1}, \ldots, i_N\}$ una permutazione di $\{n+1, \ldots, N\}$ e ponendo

$$D_{N-n,k-h} = E_{i_{n+1}} \cdots E_{i_{n+k-h}} E^c_{i_{n+k-h+1}} \cdots E^c_{i_N}\,,$$

vale

$$P(D_{N-n,k-h}\,|\,C_{n,h}) = P(D_{N-n,k-h}\,|\,S_n = h) =$$

$$= \frac{p_{N,k}}{\sum_{r=h}^{h+N-n}\binom{N-n}{r-h}p_{N,r}} = \frac{\omega_{N,k}}{\binom{N}{k}\sum_{r=h}^{h+N-n}\frac{\binom{N-n}{r-h}}{\binom{N}{r}}\,\omega_{N,r}}\,.$$

*Dim.* La prima uguaglianza segue osservando che i costituenti relativi a $E_1,\dots,E_N$ e favorevoli all'evento $D_{N-n,k-h}\wedge(S_n = h)$ sono $\binom{n}{h}$ e si ha

$$P[D_{N-n,k-h}\wedge(S_n=h)] = \binom{n}{h}P(D_{N-n,k-h}\wedge C_{n,h})\,,$$

da cui segue

$$P(D_{N-n,k-h}\,|\,C_{n,h}) = \frac{P(D_{N-n,k-h}\wedge C_{n,h})}{P(C_{n,h})} =$$

$$= \frac{\binom{n}{h}P(D_{N-n,k-h}\wedge C_{n,h})}{P(S_n=h)} = P(D_{N-n,k-h}\,|\,S_n = h)\,.$$

La seconda uguaglianza si ottiene osservando che

$$C_{n,h} = C_{n,h} \wedge (E_{i_{n+1}} \vee E^c_{i_{n+1}}) \cdots (E_{i_N} \vee E^c_{i_N}),$$

e quindi, per ogni $r = h, \dots, h + N - n$, ci sono $\binom{N-n}{r-h}$ costituenti del tipo $C_{n,h} \wedge D_{N-n,r-h}$ favorevoli a $C_{n,h}$.

Pertanto

$$P(C_{n,h}) = \sum_{r=h}^{h+N-n} \binom{N-n}{r-h} p_{N,r}$$

e quindi

$$P(D_{N-n,k-h} \,|\, C_{n,h}) = \frac{p_{N,k}}{\sum_{r=h}^{h+N-n} \binom{N-n}{r-h} p_{N,r}}.$$

Allora, ricordando la relazione

$$p_{n,h} = P(C_{n,h}) = \frac{P(S_n = h)}{\binom{n}{h}} = \frac{\omega_{n,h}}{\binom{n}{h}},$$

si ottiene

$$P(D_{N-n,k-h}\,|\,C_{n,h}) = \frac{\omega_{N,k}}{\binom{N}{k}\sum_{r=h}^{h+N-n}\frac{\binom{N-n}{r-h}}{\binom{N}{r}}\,\omega_{N,r}},$$

che permette (utilizzando solo le probabilità $\omega_{N,h}, \omega_{N,h+1}, \ldots, \omega_{N,h+N-n}$) di calcolare la probabilità del costituente $D_{N-n,k-h}$ relativo agli eventi $E_{n+1}, \ldots, E_N$, supposto di aver osservato il costituente $C_{n,h}$ relativo agli eventi $E_1, \ldots, E_n$.

*Teorema di Bayes per eventi scambiabili.*
L'ultima uguaglianza permette di calcolare la probabilità della frequenza "futura" $S_N = k$, supposto di aver osservato la frequenza "passata" $S_n = h$. Si ha

$$P(S_N = k\,|\,S_n = h) = \binom{N-n}{k-h} P(D_{N-n,k-h}\,|\,C_{n,h}) =$$

$$= \frac{\frac{\binom{N-n}{k-h}}{\binom{N}{k}}\,\omega_{N,k}}{\sum_{r=h}^{h+N-n}\frac{\binom{N-n}{r-h}}{\binom{N}{r}}\,\omega_{N,r}} = \frac{\binom{k}{h}\binom{N-k}{n-h}\,\omega_{N,k}}{\sum_{r=h}^{h+N-n}\binom{r}{h}\binom{N-r}{n-h}\,\omega_{N,r}} =$$

$$= c\,P(S_n = h \,|\, S_N = k) P(S_N = k) = c\,\frac{\binom{n}{h}\binom{N-n}{k-h}}{\binom{N}{k}}\,\omega_{N,k}\,,$$

dove per la costante $c$ si ha

$$c^{-1} = P(S_n = h) = \sum_{r=h}^{h+N-n} \frac{\binom{n}{h}\binom{N-n}{r-h}}{\binom{N}{r}}\,\omega_{N,r}\,.$$

Osserviamo che

$$P(S_N = k \,|\, S_n = h) = \binom{N-n}{k-h} P(D_{N-n,k-h} \,|\, C_{n,h}) =$$

$$= \frac{\binom{N-n}{k-h} P(D_{N-n,k-h} \wedge C_{n,h})}{P(C_{n,h})} = \frac{P((S_N = k) \wedge C_{n,h})}{P(C_{n,h})} =$$

$$= P(S_N = k \,|\, C_{n,h}) = a\,\frac{\binom{N-n}{k-h}}{\binom{N}{k}}\,\omega_{N,k}\,, \tag{30}$$

dove per la costante $a$ si ha

$$a^{-1} = P(C_{n,h}) = \sum_{r=h}^{h+N-n} \frac{\binom{N-n}{r-h}}{\binom{N}{r}}\,\omega_{N,r}\,.$$

Le uguaglianze

$$P(D_{N-n,k-h} \,|\, C_{n,h}) = P(D_{N-n,k-h} \,|\, S_n = h)\,,$$

$$P(S_N = k \,|\, C_{n,h}) = P(S_N = k \,|\, S_n = h)$$

ci dicono che nel fare inferenza sugli eventi $D_{N-n,k-h}$ ed $(S_N = k)$ non conta l'ordine con cui si sono verificati i successi e gli insuccessi, ma conta soltanto il totale dei successi ottenuti;
in questo senso $S_n = \sum_{r=1}^{n} E_{i_r}$ è una *statistica sufficiente*: ai fini inferenziali l'evento $(S_n = h)$ fornisce un'informazione uguale a quella contenuta nell'evento $C_{n,h}$.

La (30) ci permette di determinare (in modo approssimato) l'ipotesi $(S_N = k)$ che si "rinforza" di più , avendo osservato $(S_n = h)$.
Infatti:
(i) il massimo della funzione

$$f(x) = x^h(1-x)^{n-h}\,,\;\; 0 \leq x \leq 1\,,$$

si ha per $x = \frac{h}{n}$;
(ii) per $N >> n$ si ha: $\frac{\binom{N-n}{k-h}}{\binom{N}{k}} \simeq \left(\frac{k}{N}\right)^h \left(1 - \frac{k}{N}\right)^{n-h}$;
(iii) allora, dalla (30) segue che l'ipotesi $(S_N = k)$ che si "rinforza" di più (condizionatamente all'evento $S_n = h$) corrisponde al valore $k$ tale che $\frac{k}{N}$ è uguale (o all'incirca uguale) ad $\frac{h}{n}$;
(iv) in particolare, per $N = n + 1$, assumendo $\omega_{n+1,h} \simeq \omega_{n+1,h+1}$, risulta

$$P(E_{n+1} \,|\, E_{i_1} \cdots E_{i_h} E^c_{i_{h+1}} \cdots E^c_{i_n}) = P(E_{n+1} \,|\, S_n = h) =$$

$$= P(D_{1,1} \,|\, C_{n,h}) = \frac{p_{n+1,h+1}}{p_{n+1,h} + p_{n+1,h+1}} = \frac{\frac{\omega_{n+1,h+1}}{\binom{n+1}{h+1}}}{\frac{\omega_{n+1,h}}{\binom{n+1}{h}} + \frac{\omega_{n+1,h+1}}{\binom{n+1}{h+1}}} =$$

$$= \frac{(h+1)\,\omega_{n+1,h+1}}{(n-h+1)\,\omega_{n+1,h} + (h+1)\,\omega_{n+1,h+1}} \simeq \frac{h+1}{n+2} \simeq \frac{h}{n}\,;$$

ovvero, la probabilità di successo nell'$(n+1)^a$ prova, condizionata ad una osservata frequenza di successi $\frac{h}{n}$, sotto opportune condizioni (di scambiabilità , ...) è all'incirca uguale alla frequenza osservata $\frac{h}{n}$.

*Esempio.* Un lotto è costituito da $N$ pezzi e non si ha nessuna informazione sulla composizione del lotto. Supponiamo di valutare equiprobabili gli eventi: $H_r$ = *"il lotto contiene* $r$ *pezzi difettosi"*, $r = 0, 1, \ldots, N$,

Dal lotto vengono estratti a caso dei pezzi *senza restituzione*. Siano definiti gli eventi:

$E_i$ = *"l'*$i$*−mo pezzo estratto è difettoso"*, $i = 1, \ldots, n, \ldots, N$,

$S_k = E_1 + \cdots + E_k\,,\ k = 1, \ldots, N$.

(i) dimostrare che $E_1, \ldots, E_N$ sono scambiabili; inoltre, calcolare:

(ii) la probabilità di osservare $h$ pezzi difettosi nelle prime $n$ estrazioni, supposto che il lotto contenga $k$ pezzi difettosi;

(iii) la probabilità di osservare $h$ pezzi difettosi nelle prime $n$ estrazioni;

(iv) la probabilità di osservare un pezzo difettoso nell'$(n+1)$−ma estrazione, supposto di aver osservato la frequenza $\frac{h}{n}$ nelle prime $n$ prove.

*Soluzione.* (i) Si ha

$$P(H_r) = P(S_N = r) = \omega_{N,r} = \frac{1}{N+1}\,,\ \forall\, r\,.$$

Inoltre, per ogni $n \le r$, si ha

$$P(E_{i_1} \cdots E_{i_n} \,|\, H_r) = \frac{\binom{N-n}{r-n}}{\binom{N}{r}}$$

e quindi

$$P(E_{i_1} \cdots E_{i_n}) = \sum_{r=n}^{N} \frac{\binom{N-n}{r-n}}{\binom{N}{r}} \cdot \frac{1}{N+1}\,,\ 1 \le n \le N\,,$$

per ogni $\{i_1, \ldots, i_n\} \subseteq \{1, \ldots, N\}$. Pertanto $E_1, \ldots, E_N$ sono scambiabili.
In particolare

$$P(E_{i_1}) = \sum_{r=1}^{N} \frac{\binom{N-1}{r-1}}{\binom{N}{r}} \cdot \frac{1}{N+1} = \sum_{r=1}^{N} \frac{r}{N} \cdot \frac{1}{N+1} = \frac{1}{2}\,,$$

$$P(E_1 \cdots E_N) = P(H_N) = \frac{1}{N+1}\,.$$

(ii) si ha

$$P(S_n = h \,|\, H_k) = P(S_n = h \,|\, S_N = k) =$$

$$= \frac{\binom{n}{h}\binom{N-n}{k-h}}{\binom{N}{k}} = \frac{\binom{k}{h}\binom{N-k}{n-h}}{\binom{N}{n}}\,.$$

(iii) si ha

$$P(S_n = h) = \omega_{n,h} = \frac{1}{n+1}\,,\ n = 1, 2, \ldots, N\,.$$

Infatti, applicando la (28) con $n = N - 1$, si ha

$$\omega_{N-1,h} = \tfrac{N-h}{N}\,\omega_{N,h} + \tfrac{h+1}{N}\,\omega_{N,h+1} =$$

$$= \tfrac{N-h}{N} \cdot \tfrac{1}{N+1} + \tfrac{h+1}{N} \cdot \tfrac{1}{N+1} = \tfrac{1}{N}\,;$$

allora, assumendo

$$\omega_{n+1,h} = \frac{1}{n+2}\,, \quad h = 0, 1, \ldots, n+1\,,$$

e applicando la (28) con $N = n + 1$, si ha

$$\omega_{n,h} = \tfrac{n-h+1}{n+1}\,\omega_{n+1,h} + \tfrac{h+1}{n+1}\,\omega_{n+1,h+1} =$$

$$= \tfrac{n-h+1}{n+1} \cdot \tfrac{1}{n+2} + \tfrac{h+1}{n+1} \cdot \tfrac{1}{n+2} = \tfrac{1}{n+1}\,.$$

(iv) infine, si ha

$$P(E_{n+1} \,|\, S_n = h) = \frac{(h+1)\,\omega_{n+1,h+1}}{(n-h+1)\,\omega_{n+1,h} + (h+1)\,\omega_{n+1,h+1}} =$$

$$= \frac{(h+1)\cdot\frac{1}{n+2}}{(n-h+1)\cdot\frac{1}{n+2} + (h+1)\cdot\frac{1}{n+2}} = \frac{h+1}{n+2} \simeq \frac{h}{n}\,.$$

*Esercizio.* Da un'urna contenente una pallina bianca e una nera si effettuano estrazioni *con restituzione*, aggiungendo ogni volta una pallina dello stesso colore di quella estratta. Dimostrare che gli eventi $E_i =$ *"la pallina estratta nell'$i-$ma prova è bianca"* $(i = 1, \ldots, n, \ldots)$ sono scambiabili.

*Dim.* Osserviamo che la probabilità del particolare costituente

$$C_{N,k} = E_1 \cdots E_k E^c_{k+1} \cdots E^c_N$$

è pari a

$$\frac{1}{2} \cdot \frac{2}{3} \cdots \frac{k}{k+1} \cdot \frac{1}{k+2} \cdot \frac{2}{k+3} \cdots \frac{N-k}{N+1} = \frac{1}{(N+1)\binom{N}{k}} \, .$$

Si può verificare che la stessa cosa vale per il generico costituente contenente $k$ successi ed $N-k$ insuccessi. Infatti, considerati $k$ indici

$$1 \le j_1 < j_2 < \cdots < j_k \le N \, ,$$

siano $E_{j_1}, \ldots, E_{j_k}$ i $k$ eventi veri fra gli $N$ eventi $E_1, \ldots, E_N$. Allora il costituente corrispondente è

$$C_{N,k} = E^c_1 \cdots E^c_{j_1-1} E_{j_1} E^c_{j_1+1} \cdots E^c_{j_2-1} E_{j_2} E^c_{j_2+1} \cdots$$

e si ha

$$P(C_{N,k}) = \frac{1}{2} \cdot \frac{2}{3} \cdots \frac{j_1-1}{j_1} \cdot \frac{1}{j_1+1} \cdot \frac{j_1}{j_1+2} \cdots = \frac{1}{(N+1)\binom{N}{k}} \, ,$$

da cui segue anche:

$$P(S_N = k) = \frac{1}{N+1}\,,\ k = 0, 1, \ldots, N\,.$$

Osserviamo adesso che all'evento $E_{j_1} \cdots E_{j_k}$ sono favorevoli i costituenti $C_{N,r}$, con $r \geq k$ eventi veri, fra i quali $E_{j_1}, \ldots, E_{j_k}$, ed $N - r$ eventi falsi. Inoltre, per ogni fissato $r$ di tali costituenti ce ne sono $\binom{N-k}{r-k}$ ed ognuno ha probabilità $\frac{1}{(N+1)\binom{N}{r}}$.
Pertanto, osservando che $\sum_{r=k}^{N} \binom{r}{k} = \binom{N+1}{k+1}$, si ha

$$P(E_{j_1} \cdots E_{j_k}) = \sum_{r=k}^{N} \frac{\binom{N-k}{r-k}}{(N+1)\binom{N}{r}} = \frac{1}{N+1} \sum_{r=k}^{N} \frac{\binom{r}{k}}{\binom{N}{k}} =$$

$$= \frac{\sum_{r=k}^{N} \binom{r}{k}}{(N+1)\binom{N}{k}} = \frac{\binom{N+1}{k+1}}{(N+1)\binom{N}{k}} = \frac{1}{k+1}\,,$$

e, ricordando la Definizione 2, gli eventi $E_1, \ldots, E_n, \ldots$ sono scambiabili. In particolare, gli eventi $E_i$ hanno tutti probabilità $\frac{1}{2}$.

*Applicazione.* Una popolazione è inizialmente composta da due individui di opinioni politiche diverse $F_1, F_2$. In fasi successive nella popolazione si aggiunge ogni volta un individuo, il quale acquisisce l'opinione politica del primo individuo che incontra a caso nella popolazione. Definiamo gli eventi:
$E_i =$ *"l'i-mo individuo che si aggiunge nella popolazione acquisisce l'opinione* $F_1$", $i = 1, 2, \ldots$;
in base all'esempio di prima gli eventi $E_1, E_2, \ldots$ sono scambiabili; inoltre, per ogni fissato $n$, dopo $n$ ingressi le $n+1$ possibili composizioni della popolazione sono

$$(1, n+1)\,,\ (2, n)\,, \ldots\,,\ (n, 2)\,,\ (n+1, 1)\,,$$

dove $(h+1, n-h+1)$ significa che nella popolazione si sono aggiunti $h$ individui di tipo $F_1$ ed $n-h$ individui di tipo $F_2$. Tali composizioni sono tutte equiprobabili, di probabilità $\frac{1}{n+1}$.