Appunti di Calcolo delle Probabilità

Riccardo Lo Iacono

Dipartimento di Matematica & Informatica
Università degli studi di Palermo
Sicilia
a.a. 2022-2023

Indice.

– 1 – Introduzione.

Si consideri il seguente caso: lanciando tre dati, è maggiore la probabilità che la somma dei valori ottenuti sia 9 o che sia 10?
Un metodo semplice per stabilire quanto richiesto, consiste nell'elencare tutti i possibili modi in cui è possibile ottenere 9 e 10.

SOMMA 9			
D_1	D_2	D_3	#
6	2	1	6
5	3	1	6
5	2	2	3
4	4	1	3
4	3	2	6
3	3	3	1
			25

SOMMA 10			
D_1	D_2	D_3	#
6	3	1	6
5	2	2	3
5	4	1	6
5	3	2	6
4	4	2	3
4	3	3	3
			27

Sia n il numero totale di casi possibili, nel caso in esame 216, la probabilità P di un evento[1] E risulta essere definita come segue.

$$P(E) = \frac{\sum \#}{n} \tag{1}$$

ove $\#$ indica il numero di casi favorevoli all'evento in questione.

Applicando l'Equazione (1) al caso in esame, segue

$$P(SOMMA = 9) = \frac{25}{216} \approx 0.12$$
$$P(SOMMA = 10) = \frac{27}{216} = 0.125$$

da cui risulta ovvio $P(SOMMA = 10) > P(SOMMA = 9)$.

[1] La definizione di evento sarà data in seguito.

– 2 – Criterio classico ed eventi.

Uno dei criteri alla base del calcolo probabilistico è il *criterio classico della probabilità*, il quale stabilisce che dato E un evento, la probabilità P che questi si verifichi è pari al rapporto di casi favorevoli su casi possibili, cioè

$$P(E) = \frac{\text{\# casi a favore ad E}}{\text{\# casi possibili}} = \frac{r_E}{m}$$

ove r_E sono i casi a favore di E , m tutti i casi possibili.

Definizione: una proposizione o un'affermazione di cui è possibile stabilire la veridicità è detto *evento*.

Di un evento è possibile stabilire l'indicatore, che segnala se l'evento è o non è verificato.

$$|E| = \begin{cases} 1, & \text{se E è vero} \\ 0, & \text{se E è falso} \end{cases},$$

inoltre se a priori è noto il valore di $|E|$, questi si dirà *certo* se $|E| = 1$ e lo si indicherà con Ω, *impossibile* se $|E| = 0$ e lo si indicherà con $\varnothing$.

– 2.1 – Proprietà degli eventi.

Gli eventi sono soggetti ad alcune proprietà, quali

- negazione;
- implicazione;
- uguaglianza;
- unione;
- intersezione.

– 2.1.1 – Negazione.

Dato un evento E, l'evento negato

$$E^C \begin{cases} \text{vero, se E è falso} \\ \text{falso, se E è vero} \end{cases}$$

– 2.1.2 – Implicazione.

Dati due eventi E_1 , E_2, si dirà che E_1 implica E_2 se

$$|E_1| = 1 \implies |E_2| = 1$$
$$|E_1| = 0 \implies |E_2| = 0, 1$$

– 2.1.3 – Uguaglianza.

Dati due eventi E_1 , E_2, si dice che $E_1 = E_2$ se ogni esito di E_1 è verificato in E_2, e viceversa.

– 2.1.4 – Unione.

Dati due eventi E_1 , E_2, si definisce l'evento $E_3 = E_1 \vee E_2$ che soddisfa gli esiti di E_1 o E_2, unione.

In particolare

- $E \vee E^C = \Omega$
- $E \vee \varnothing = E$
- $E \vee E = E$
- $E \vee \Omega = \Omega$

– 2.1.5 – Intersezione.

Dati due eventi E_1 , E_2, si definisce l'evento $E_3 = E_1 \wedge E_2$ che soddisfa gli esiti presenti sia in E_1 che in E_2, intersezione.

Nota: Se $E_1 \wedge E_2 = \varnothing$, i due eventi si dicono *incompatibili* (o disgiunti).

– 2.2 – Assiomi e proprietà derivate.

Il calcolo delle probabilità si fonda su alcune nozioni che, da ora in avanti, si daranno per assodate. Tali nozioni sono tre assiomi che di seguito saranno trattati.

Assioma 2.2.1.
Dato un evento E, la probabilità che questo si verifichi è sempre compresa tra 0 e 1. Cioè

$$0 \leq P(E) \leq 1$$

Assioma 2.2.2.
Considerato uno spazio campionato Ω, la probabilità che questi si verifichi è pari a 1. Cioè

$$P(\Omega) = 1$$

Assioma 2.2.3.
Dati un numero n di eventi, a due a due incompatibili, la probabilità che almeno uno tra gli n eventi si verifichi è pari alla somma delle rispettive probabilità. Cioè

$$P\left(\bigvee_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Si analizzano ora alcune proprietà derivate dagli assiomi precedentemente introdotti, spesso utili alla risoluzioni di problemi probabilistici.

Proposizione 2.2.1.
Dati due eventi disgiunti E_1 , E_2, tali che $E_1 \implies E_2$, la probabilità di E_1 è minore o uguale a quella di E_2. Cioè

$$P(E_1) \leq P(E_2)$$

Dimostrazione: se $E_1 \implies E_2$, allora $E_2 = E_1 \vee E_1^C E_2$. Ma E_1 e $E_1^C E_2$ sono disgiunti, da cui per l'Assioma (2.2.3) segue

$$P(E_2) = P(E_1) + P(E_1^C E_2)$$

ma ciò implica, poiché $P(E_1^C E_2) \geq 0$, che

$$P(E_1) \leq P(E_2)$$

Proposizione 2.2.2.
Dati due eventi disgiunti E_1 , E_2, tali che $P(E_1 \vee E_2) = 1$, segue per gli Assiomi (2.2.2), (2.2.3) che

$$P(E_1 \vee E_2) = P(E_1) + P(E_2) = 1$$

ma da ciò

$$P(E_1) = 1 - P(E_2)$$

Proposizione 2.2.3.
Dati due eventi, la probabilità della loro unione è pari alla somma delle probabilità dei singoli eventi meno quella della loro intersezione. Cioè

$$P(E_1 \vee E_2) = P(E_1) + P(E_2) - P(E_1 E_2)$$

Dimostrazione: si nota che $E_1 \vee E_2 = E_1 \vee E_1^C E_2$, da cui per l'Assioma (2.2.3) segue

$$P(E_1 \vee E_2) = P(E_1) + P(E_1^C E_2)$$

ma a sua volta $E_2 = E_1 E_2 \vee E_1^C E_2$, da cui sempre per l'Assioma (2.2.3) segue

$$P(E_2) = (E_1 E_2) + P(E_1^C E_2)$$

da cui sostituendo si ottiene quanto si voleva dimostrare.

Nota: La proposizione (2.2.3) è generalizzabile ad n eventi applicando il principio dei cassetti.

– 3 – Variabili aleatorie.

Spesso quando si effettua lo studio di un fenomeno aleatorio, si è molto più interessati a una qualche funzione degli esiti che agli stessi.
Si prenda ad esempio il lancio di una moneta: ci si domanda quante volte esca testa.

Quantità come quella dell'esempio si definiscono *variabili aleatorie*. Conseguentemente, poiché il valore assunto da tali variabili è dipendente dall'esito del fenomeno, è possibile attribuire a queste una probabilità.

Esempio: Si supponga di lanciare tre monete. Sia γ il numero di volte in cui esce testa. Da ciò γ assume possibilmente i valori 0, 1, 2, 3, le cui rispettive probabilità sono

$$\begin{aligned}
&P\{\gamma = 0\} = P\{(C,C,C)\} = \tfrac{1}{8} \\
&P\{\gamma = 1\} = P\{(C,C,T),(C,T,C),(T,C,C)\} = \tfrac{3}{8} \\
&P\{\gamma = 2\} = P\{(C,T,T),(T,C,T),(T,T,C)\} = \tfrac{3}{8} \\
&P\{\gamma = 3\} = P\{(T,T,T)\} = \tfrac{1}{8}
\end{aligned}$$

Ma la probabilità di un singolo evento è pari a 1, da cui

$$P\left(\bigvee_{i=0}^{3} \{\gamma = i\}\right) = \sum_{i=0}^{3} P\{\gamma = i\} = 1$$

– 4 – Probabilità condizionata.

Si supponga di lanciare due dadi, si consideri che il primo dado dia 3, se si indica con E_2 tale evento e con E_1 l'evento *somma dei dadi uguale a otto*, qual è la probabilità di E_1?

Considerando che tutti gli esiti siano equiprobabili, poiché dato E_2, i possibili esiti

$$(3,1) \quad (3,2) \quad (3,3) \quad (3,4) \quad (3,5) \quad (3,6)$$

segue che ciascuno di essi ha probabilità $\frac{1}{6}$.

In generale $\forall E_1, E_2$ tali che E_1 è condizionato da E_2

$$P(E_1 \mid E_2) = \frac{P(E_1 E_2)}{P(E_2)} \qquad \text{tale che } P(E_2) > 0 \tag{2}$$

Generalizzando ulteriormente: dati $E_1 \cdots E_n$ eventi, si ha

$$P(E_1 \cdots E_n) = P(E_1)P(E_1 \mid E_2) \cdots P(E_n \mid E_1 \cdots E_{n-1})$$

la cui dimostrazione applicando l'Equazione risulta essere

$$P(E_1)\frac{P(E_1 E_2) \cdots P(E_1 \cdots E_n)}{P(E_1) \cdots P(E_1 \cdots E_{n-1})} = P(E_1 \cdots E_n)$$

– 4.1 – Formula di Bayes.

Siano E_1, E_2 eventi, con $0 < P(E_2) < 1$. Allora

$$E_1 = E_1 E_2 \cup E_1 E_2^C$$

poiché un esito di E_1 è, oppure no, in E_2.
Ma per l'Assioma (2.2.3) segue che, poiché $E_1 E_2, E_1 E_2^C$ sono incompatibili

$$\begin{aligned}
P(E_1) &= P(E_1 E_2) + P(E_1 E_2^C) \\
&= P(E_1 \mid E_2)P(E_2) + P(E_1 \mid E_2^C)P(E_2^C) \\
&= P(E_1 \mid E_2)P(E_2) + P(E_1 \mid E_2^C)[1 - P(E_2)]
\end{aligned}$$

cioè $P(E_1)$ è la media ponderata della probabilità di E_1 dato E_2 e della probabilità di E_1 dato E_2^C.

– 4.2 – Probabilità condizionata e assiomi.

Analogamente la probabilità finora trattata, quella condizionata soddisfa gli Assiomi (2.2.1), (2.2.2), (2.2.3). Segue quindi

- $0 \leq P(E_1 \,|\, F) \leq 1$

  Dimostrazione: la prima disuguaglianza è ovvia, la seconda discerne dal fatto che $EF \subset F \implies P(EF) \leq P(F)$.

- $P(\Omega \,|\, F) = 1$

  Dimostrazione: considerando l'Equazione (2) segue

$$P(\Omega \,|\, F) = \frac{P(SF)}{P(F)} = \frac{P(F)}{P(F)} = 1$$

- Dati E_i eventi, $i = \{1, 2, 3, \ldots\}$, a due a due disgiunti, allora

$$P\left(\bigvee_{i=1}^{\infty} E_i \,|\, F\right) = \sum_{i=1}^{\infty} P(E_i \,|\, F)$$

  Dimostrazione: considerando nuovamente l'Equazione (2) segue

$$\begin{aligned}
P\left(\bigvee_{i=1}^{\infty} E_i \,|\, F\right) &= \frac{P\left(\left(\bigvee_{i=1}^{\infty} E_i\right) F\right)}{P(F)} \\
&= \frac{P\left(\bigvee_{i=1}^{\infty} E_i F\right)}{P(F)} \\
&= \frac{\sum_{i=1}^{n} P(E_i F)}{P(F)} \\
&= \sum_{i=1}^{n} P(E_i \,|\, F)
\end{aligned}$$

– 4.3 – Esercizi esemplificativi.

Esercizio: siano date due urne U e V. Si sceglie casualmente tra le due, e da questa si effettuano estrazioni con ripetizione.

$$U = \{b, b, n\} \qquad V = \{n, n, b\}$$

Considerando pertanto

$$\begin{aligned} H &= \text{"l'urna scelta è U"} \\ E_i &= \text{"l'i-esima pallina è 'b'"} \end{aligned}$$

segue

$$P(H) = P(H^C) = \frac{1}{2}$$

inoltre

$$\begin{aligned} P(E_i) = P(E_i \cup \Omega) &= P(E_i H) + P(E_i H^C) \\ &= P(E_i \mid H)P(H) + P(E_i \mid H^C)P(H^C) \end{aligned}$$

Si calcolino

- $P(H \mid E_1)$: cioè, sapendo che la prima pallina estratta è bianca, qual è la probabilità che l'urna scelta sia U.
- $P(H \mid E_1 E_2^C E_3^C)$: cioè, sapendo che le palline estratte sono nell'ordine b, n, n, qual è la probabilità che l'urna scelta sia U.

Svolgimento: per la risoluzione degli esercizi si applicherà Bayes.

1.
$$P(H \mid E_1) = \frac{P(E_1 H)}{P(E_1)} = \frac{P(E_1 \mid H)P(H)}{P(E_1 \mid H) + P(E_1 \mid H^C)} = \frac{2}{3}$$

2.
$$P(H \mid E_1 E_2^C E_3^C) = \frac{P(E_1 E_2^C E_3^C H)}{P(E_1 E_2^C E_3^C)} = \frac{P(E_1 E_2^C E_3^C H)}{P(E_1 E_2^C E_3^C H) + P(E_1 \overline{E_2 E_3 H})}$$

ma

$$P(E_1 E_2^C E_3^C H) = P(\overline{E_3} \mid \overline{E_2} E_1 H) P(\overline{E_2} \mid E_1 H) P(E_1 \mid H) P(H) = \frac{1}{27}$$
$$P(E_1 E_2^C E_3^C H^C) = P(\overline{E_3} \mid \overline{E_2} E_1 H^C) P(\overline{E_2} \mid E_1 H^C) P(E_1 \mid H^C) P(H^C) = \frac{2}{27}$$

da cui

$$P(H \mid E_1 E_2^C E_3^C) = \frac{1}{3}$$

Esercizio: siano date tre urne X, Y, Z. Si sceglie una delle tre senza conoscerne il contenuto.

$$X = \{Premio\} \quad Y = \{Capra\} \quad Z = \{Capra\}$$

Supponendo si scelga inconsciamente l'urna X, ci si chiede se convenga cambiare la propria scelta se, una volta rivelato il contenuto dell'urna Y questa è una capra.

Svolgimento: si ha che

$$\forall i \in \{1,2,3\} \quad P(A_i) = \frac{1}{3}$$

inoltre

$$P(A_1 \mid A^C) = \frac{P(A_1 A^C)}{P(A^C)} = \frac{P(A_1)}{P(A_1) + P(A_3)} = \frac{1}{2}$$

cioè la scelta apparentemente sembra indifferente. Ma è davvero così?

Se si definisce $I =$ "la scatola due contiene una capra" si hanno due scenari

$$I \implies A_2 \text{ ? SI}$$
$$I \impliedby A_2 \text{ ? NO}$$

Pertanto è più corretto calcolare la probabilità condizionata ad I, segue

$$P(A_i \mid I) = \frac{P(I \mid A_1) P(A_1)}{\sum_{j=1}^{3} P(I \mid A_j) P(A_j)}$$

ma

$$P(I \mid A_1) = p$$
$$P(I \mid A_2) = 0$$
$$P(I \mid A_3) = 1$$

segue dunque

$$P(A_1 \mid I) = \frac{p\frac{1}{3}}{p\frac{1}{3} + \frac{1}{3}} = \frac{p}{p+1} \leq \frac{1}{p+1} = P(A_3 \mid I)$$

con

$$P(A_1 \mid I) = P(A_3 \mid I) \text{ se } p = 1$$
$$P(A_1 \mid I) < P(A_3 \mid I) \text{ se } p < 1$$

– 4.4 – Indipendenza stocastica.

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

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

Nei primi due casi si dirà che E è rispettivamente *correlato positivamente o negativamente* da H. Nel caso di $P(E|H) = P(E)$ invece, si dirà che E è *stocasticamente* indipendente da H.

Definizione: Dati E, H eventi, con $H \neq \emptyset$, si dirà E stocasticamente indipendente da H se

$$P(E\,|\,H) = P(E) \tag{3}$$

In generale, data una famiglia di eventi F, gli eventi di F si diranno indipendenti stocasticamente se, per ogni sottofamiglia $\{E_1, E_2, \cdots, E_n\}$ di F, $n \geq 2$ si ha

$$P(E_1 E_2 \cdots E_n) = P(E_1)P(E_2)\cdots P(E_n)$$

– 5 – Calcolo combinatorio e probabilità.

Prima di applicare alla probabilità il calcolo combinatorio, è bene fare un richiamo a concetti come permutazioni, disposizioni, combinazioni.

Si consideri il seguente esempio

Esempio: siano A, B, C località, siano p_1, p_2, p_3 percorsi da A a B e siano s_1, s_2 percorsi da B a C. Ci si chiede: quanti sono i percorsi distinti da A a C?

Si osserva che per ognuno dei percorsi da A a B, si hanno due possibili scelte da B a C. Si hanno dunque le seguenti possibilità

$$(p_1, s_1) \quad (p_2, s_1) \quad (p_3, s_1) \quad (p_1, s_2) \quad (p_2, s_2) \quad (p_3, s_2)$$

Quanto appena applicato è noto come *principio di moltiplicazione*, il quale stabilisce che: dati n modi per effettuare una scelta, per ciascuno di essi si hanno m modi per farne un'altra, esistono allora $n \cdot m$ diverse scelte.

– 5.1 – Disposizioni.

Definizione: dato $S = \{a_1, a_2, \dots, a_n\}$ insieme di n elementi, si definiscono disposizioni $D_{n,k}$, per un certo k intero positivo, i sottoinsiemi di k elementi distinti che si possono formare da S.

Si consideri il seguente esempio in due casi

1. si effettuano restituzioni;

2. non si effettuano restituzioni.

Esempio: da un urna contenente n palline, se ne estraggono $r \leq n$. Quante sono le possibili disposizioni?

1. Ogni pallina estratta è reinserita, segue

$$D_{n,r} = n^r \tag{4}$$

2. Una volta estratta la pallina, questa non viene reinserita, segue

$$D_{n,r} = n(n-1)(n-2)\cdots(n-r+1) \tag{5}$$

– 5.2 – Permutazioni.

Definizione: si definiscono permutazioni P_n di un insieme di n elementi, il numero di disposizioni per $r = n$. Cioè

$$P_n = D_{n,n} = n(n-1)(n-2)\cdots 1 = n! \tag{6}$$

Esempio: quanti sono i possibili anagrammi della parola "COSA", anche senza senso?
Applicando l'Equazione (6), segue

$$P_n = D_{n,n} = 4 \cdot 3 \cdot 2 \cdot 1 = 4! = 24$$

Osservazione: se la parola da anagrammare avesse delle lettere ripetute, l'Equazione (6) non sarebbe corretta. In casi simili si parla di permutazioni con ripetizioni. In generale se $k_1, k_2, \ldots, k_m$ sono elementi ripetuti, le permutazioni distinte sono

$$P_n^{k_1,k_2,\ldots,k_m} = \frac{P_n}{k_1!k_2!\ldots k_m!} = \frac{n!}{k_1!k_2!\ldots k_m!}$$

– 5.3 – Combinazioni.

Definizione: Si definiscono combinazioni $C_{n,r}$ di un insieme di n elementi, il numero di sottoinsiemi di r elementi che si possono ottenere dall'insieme, in modo che questi non siano ordinati.

Si distinguono due classico

1. combinazioni semplici: $\forall i, j \mid i \neq j \implies a_i \neq a_j$

$$C_{n,r} = \frac{D_{n,r}}{r!} = \frac{n!}{r!(n-r)!}$$

2. combinazioni con ripetizione: $\exists i, j \mid i \neq j \implies a_i = a_j$

$$C'_{n,r} = \frac{(n+r-1)!}{r!(n-r)!}$$

– 5.4 – Distribuzione binomiale.

Siano dati $E_1, E_2, \ldots, E_n$ eventi stocasticamente indipendenti, con probabilità p.
Sia $q = 1 - p$.
Se si indica con X il numero di successi alle n prove, ne segue logicamente

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

Per calcolare la distribuzione di probabilità di X, occorre calcolare $P(X = i)$ per ogni $i \in \{1, 2, \ldots, n\}$. A tale scopo risulta utile considerare $X = i$ come la unione dei costituenti della famiglia $\{E_1, E_2, \ldots, E_n\}$ ad esso favorevoli, segue

$$(X = 0) = E_1^C E_2^C \cdots E_n^C$$

da cui

$$\begin{aligned} P(X = 0) = P(E_1^C E_2^C \cdots E_n^C) &= P(E_1^C) P(E_2^C) \cdots P(E_n^C) \\ &= (1 - p)(1 - p) \cdots (1 - p) \\ &= q^n \end{aligned}$$

Analogamente segue

$$(X = n) = E_1 E_2 \cdots E_n$$

per cui

$$\begin{aligned} P(X = n) = P(E_1 E_2 \cdots E_n) &= P(E_1) P(E_2) \cdots P(E_n) \\ &= p \cdot p \cdots p \\ &= p^n \end{aligned}$$

Si consideri adesso $X = 1$, segue

$$(X = 1) = E_1 E_2^C \cdots E_n^C \wedge E_1^C E_2 \cdots E_n^C \wedge \cdots E_1^C E_2^C \cdots E_n$$

si ha quindi

$$\begin{aligned} P(X = 1) &= P(E_1 E_2^C \cdots E_n^C \wedge E_1^C E_2 \cdots E_n^C \wedge \cdots \wedge E_1^C E_2^C \cdots E_n) \\ &= P(E_1 E_2^C \cdots E_n^C) + \cdots + P(E_1^C E_2^C \cdots E_n) \\ &= pq^{n-1} + pq^{n-1} + \cdots + pq^{n-1} \\ &= npq^{n-1} \\ &= \binom{n}{1} pq^{n-1} \end{aligned}$$

In generale fissato $i \in \{0, 1, \ldots, n\}$, vale

$$P(X = i) = \binom{n}{i} p^i q^{n-i} \tag{7}$$

Esempio: Data un'urna contenente sette palline, di cui tre bianche e le restanti nere. Si effettuano cinque estrazioni con restituzione: si calcolino

1. A = "escono almeno due bianche";
2. B = "escono esattamente due bianche";
3. C = "le prime due palline sono bianche".
4. Inoltre, supponendo di aver estratto almeno una bianca, si calcoli la probabilità condizionata α che sia uscita esattamente una bianca.

a. Si nota $A = (X \geq 2)$, ove X è il numero aleatorio di volte in cui la pallina estratta è bianca, segue

$$P(A) = P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$$

Se si ragiona per complementi si ha $P(A) = 1 - P(X = 0) - P(X = 1)$, segue

$$\begin{aligned} P(A) = 1 - P(X \leq 2) = 1 - \binom{5}{0}\left(\frac{3}{7}\right)^0\left(\frac{4}{7}\right)^5 - \binom{5}{1}\left(\frac{3}{7}\right)\left(\frac{4}{7}\right)^5 \\ = 1 - \left(\frac{4}{7}\right)^5 - \frac{5!}{1!4!}\left(\frac{3}{7}\right)\left(\frac{4}{7}\right)^4 \approx 0,71 \end{aligned}$$

b. Si nota che $B = (X = 2)$, ove X è il numero aleatorio di volte in cui la pallina estratta è bianca, segue

$$\begin{aligned} P(B) = P(X = 2) &= \binom{5}{2}\left(\frac{3}{7}\right)^2\left(\frac{4}{7}\right)^3 \\ &= \frac{5!}{2!3!}\left(\frac{3}{7}\right)^2\left(\frac{4}{7}\right)^3 \approx 0,343 \end{aligned}$$

c. Si osserva che $C = E_1 E_2 \overline{E_3 E_4 E_5}$, segue, poiché stocasticamente indipendenti, che

$$\begin{aligned} P(C) = P(E_1 E_2 \overline{E_3 E_4 E_5}) &= P(E_1)P(E_2)P(\overline{E_3})P(\overline{E_4})P(\overline{E_5}) \\ &= \left(\frac{3}{7}\right)^2\left(\frac{4}{7}\right)^3 \approx 0,034 \end{aligned}$$

d. Si osserva che $\alpha = p(X = 1 \mid X \geq 1)$, pertanto

$$\alpha = P(X = 1 \mid X \geq 1) = \frac{P((X = 1) \cap (X \geq 1))}{P(X \geq 1)} = \frac{P(X = 1)}{1 - P(X = 0)} \approx 0,243$$

– 5.5 – Distribuzione geometrica.

Definizione: un certo numero aleatorio $X \in \mathbb{N}$ si dice avere *distribuzione geometrica* di parametro $p \in (0,1)$, in simboli $X \sim G(p)$, se

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

Sia $E_1, E_2, \ldots, E_n$ una successione di eventi, tali che $P(E_i) = p \in (0,1), q = 1-p$.

$$X = \text{"Numero aleatorio di prove fino al primo successo"}$$

ossia

$$X = \inf\{n : n|E_n| = 1\}$$

Si ha dunque

$$\begin{aligned}
(X = 1) = E_1 &\implies P(X = 1) = P(E_1) = p \\
(X = 2) = E_1^C E_2 &\implies P(X = 2) = P(E_1^C E_2) = qp \\
(X = n) = E_1^C E_2^C \cdots E_{n-1}^C E_n &\implies P(X = n) = P(E_1^C E_2^C \cdots E_{n-1}^C E_n) = q^{n-1}p
\end{aligned}$$

In generale

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

Esempio: sia $X > m$, ci si chiede $P(X > n+m \,|\, X > m)$.
Si osserva che

$$(X > n+m) \impliedby (X > m)$$

segue quindi

$$(X > n+m) \cap (X > m) = (X > n+m)$$

pertanto

$$\begin{aligned}
P(X > n+m \,|\, X > m) &= \frac{P(X > n+m)}{P(X > m)} \\
&= \frac{q^{n+m}}{q^n} = q^n = P(X > n)
\end{aligned} \tag{8}$$

Quanto appena affermato in Equazione (8), è noto come *proprietà di assenza di memoria*, per la quale in generale

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

Esempio: Una moneta viene lanciata X volte. Sia

$X =$ "Numero aleatorio di lanci utili ad ottenere testa per la prima volta"

e gli eventi

$$A = (X\ pari)$$
$$B = (X\ dispari)$$

Si stabilisca se

$$P(A) = P(B)$$
$$P(A) < P(B)$$
$$P(A) > P(B)$$

Supponendo inoltre di non aver ottenuto testa nei primi 1000 lanci, quanto vale la probabilità α che non esca testa nei successivi 10? Supposto invece di aver ottenuto testa per $X\ pari$, quanto vale la probabilità condizionata β di ottenere testa al secondo lancio.

1. Si parta col considerare $P(A)$, segue

$$\begin{aligned} P(A) &= P(X \in 2n, n \in \mathbb{N}) \\ &= P\left(\bigvee_{n=1}^{+\infty}(X = 2n)\right) \\ &= \sum_{n=1}^{+\infty} P(X = 2n) = \sum_{n=1}^{+\infty} q^{2n-1}p = \sum_{n=1}^{+\infty}\left(\frac{1}{4}\right)^n = \frac{1}{4}\frac{1}{1-\frac{1}{4}} = \frac{1}{3} \end{aligned}$$

Si consideri ora $P(B)$, segue

$$\begin{aligned} P(B) &= P(X \in 2n-1, n \in \mathbb{N}) \\ &= P\left(\bigvee_{n=1}^{+\infty}(X = 2n-1)\right) \\ &= \sum_{n=1}^{+\infty} P(X = 2n-1) = \sum_{n=1}^{+\infty} q^{2n-2}p = \frac{1}{2}\frac{1}{1-\frac{1}{4}} = \frac{2}{3} \end{aligned}$$

Segue ovviamente

$$P(A) < P(B)$$

2. Si osserva che $\alpha = P(X > 1010\,|\,X > 1000)$, ma poiché della forma $P(X > m + n\,|\,X > m)$, segue che

$$\begin{aligned} \alpha = P(X > 1010\,|\,X > 1000) &= \frac{P(X > 1010)}{P(X > 1000)} \\ &= \frac{\left(\frac{1}{2}\right)^{1010}}{\left(\frac{1}{2}\right)^{1000}} = \left(\frac{1}{2}\right)^{10} \end{aligned}$$

3. Si osserva che $\beta = P(X = 2 \,|\, X \ pari)$, segue

$$\begin{aligned}\beta = P(X = 2 \,|\, X \ pari) &= \frac{P((X = 2) \cap (X \ pari))}{P(X \ pari)} \\ &= \frac{P(X = 2)}{P(X \ pari)}\end{aligned}$$

ma se si considera E_i l'evento "esce testa all'i-esimo lancio", si ha $(X = 2) = E_1^C E_2$, si ha

$$P(X = 2) = P(E_1^C E_2) = P(E_1^C)P(E_2)$$

inoltre poiché $\forall i \in \{1, 2, \dots, n\}, P(E_i) = \frac{1}{2}$, segue

$$\beta = \frac{P(X = 2)}{P(X \ pari)} = \frac{P(E_1^C)P(E_2)}{P(X \ pari)} = \frac{\frac{1}{2}\frac{1}{2}}{\frac{1}{3}} = \frac{3}{4}$$

– 5.6 – Distribuzione ipergeometrica.

Si effettuano n estrazioni senza ripetizioni da un urna contenente N palline, di cui pN bianche e qN nere, $qN = N - pN$. Siano

$$E_i = \text{"esce bianca all'i-esima estrazione"}$$

$$X = \sum_{i=1}^{n} |E_i| = \text{"numero aleatorio di palline bianche estratte nelle } n \text{ estrazioni"}$$

Si consideri $P(E_i), i \in \{1, \ldots, n\}$, segue

$$\begin{aligned} P(E_1) &= \frac{pN}{N} = p \\ P(E_2) = P(E_1 \cap \Omega) &= P(E_2 \cap (E_1 \vee E_1^C)) \\ &= P(E_2 \mid E_1) + P(E_2 \mid E_1^C) \\ &= \frac{pN - 1}{N - 1} p + \frac{pN}{N - 1} q = p \end{aligned}$$

Analogamente se si applica il calcolo combinatorio, segue

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

In generale

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

Ma E_1, E_2 sono stocasticamente indipendenti?
Si osserva che

$$P(E_2 \mid E_1) = \frac{pN - 1}{N} \neq P(E_2) = \frac{pN}{N} = p$$

Sia $0 \le h \le n$, ci si chiede quanto valga $P(X = h)$.
Si ha che $\forall h \in \{0, \ldots, n\}, P(X = h) = P(E_1 E_2 \cdots E_h E_{h+1}^c \cdots E_n^C)$, vale cioé

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

il cui risultato unicamente dipendente da n, h.

Considerando il generico costituente

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

favorevole ad $(X = h)$, segue

$$(X = h) = \bigvee_{\{i_1,\dots,i_n\} \subseteq \{1,\dots,n\}} E_{i_1} E_{i_2} \cdots E_{i_h} E_{i_{h+1}} \cdots E_{i_n}$$

ossia l'unione degli $\binom{n}{h}$ costituenti a favore di $(X = h)$. In definitiva

$$P(X = h) = \binom{n}{h} \frac{\binom{N-n}{pN-h}}{\binom{N}{pN}} \tag{9}$$

Definizione: si definisce la distribuzione relativa ad X *ipergeometrica*. In simboli

$$X \sim H(N, n, pN)$$

Esempio: un urna contiene dodici palline, di cui cinque bianche e le restanti nere. Si effettuano tre estrazioni, se

$$E_i = \text{"l'i-esima pallina estratta è bianca"}$$

$$X = \sum_{i=1}^{3} |E_i|$$

quanto vale $P(X = 2)$?

Si osserva che

$$P(X = 2) = P(E_1 E_2 \overline{E_3} \vee E_1 \overline{E_2} E_3 \vee E_1^C E_2 E_3)$$

da cui applicando l'Equazione (9), segue

$$P(X = 2) = \frac{\binom{3}{2}\binom{9}{3}}{\binom{12}{5}} \approx 0,318$$

verificando anche con l'Equazione (7), si ha

$$P(X = 2) = \binom{3}{2} \left(\frac{5}{12}\right)^3 \left(\frac{7}{12}\right) = 0,303 \approx 0,318$$

– 5.7 – Distribuzione di Poisson.

Sia $X = 0, 1, 2, \dots$ numero aleatorio, questa si dice avere distribuzione di Poisson, di parametro λ se

$$P(X = i) = e^{-\lambda} \frac{\lambda}{i!} \qquad i = 0, 1, 2, \dots \ , \lambda > 0$$

La distribuzione di Poisson ha diverse applicazioni, tra queste l'approssimazione di un numero aleatorio con distribuzione binomiale a parametri n, p.

Dimostrazione: Sia X numero aleatorio a parametri n, p, sia $\lambda = np$, segue

$$\begin{aligned} P(X = i) &= \frac{n!}{(n-1)!i!} p^i (1-p)^{n-1} \\ &= \frac{n!}{(n-1)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-1} \\ &= \frac{n(n-1)(n-2)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^i} \end{aligned}$$

Per n molto piccoli e p sufficientemente piccoli tali che λ sia costante

$$\begin{aligned} \left(1 - \frac{\lambda}{n}\right)^n &\approx e^{-\lambda} \\ \frac{n(n-1)(n-2)\cdots(n-i+1)}{n^i} &\approx 1 \\ \left(1 - \frac{\lambda}{n}\right)^i &\approx 1 \end{aligned}$$

segue

$$P(X = i) \approx e^{-\lambda} \frac{\lambda^i}{i!}$$

Inoltre $\forall i \in \mathbb{N}$, si ha che

$$P(X = i + 1) = \frac{\lambda}{i + 1} P(X = i)$$

– 5.8 – Distribuzione di Pascal e binomiale inversa.

Siano $E_1, E_2, \ldots, E_n$ successione di eventi stocasticamente indipendenti. Posto T_k il numero di prove fino al k-esimo successo, segue

$$\begin{aligned} P(T_k = i) &= \binom{i-1}{k-1} p^{k-1} q^{i-k} p \\ &= \binom{i-1}{k-1} p^k q^{i-k} \end{aligned}$$

poiché si ha che nelle $i-1$ prove si hanno $k-1$ successi.

Esempio: Viene lanciata una moneta fino ad ottenere testa per la seconda volta: si stabilisca

1. $P(T_2 = 3)$
2. $P(T_2 = 4)$

1. Si consideri l'evento associato a $(T_2 = 3)$, si ha

$$(T_2 = 3) = TCT \vee CTT$$

questo perché al terzo lancio necessariamente si ottiene testa, dunque

$$P(T_2 = 3) = P(TCT) + P(CTT)$$

2. Analogamente a prima, si consideri l'evento associato a $(T_2 = 4)$, segue

$$(T_2 = 4) = CCTT \vee CTCT \vee TCCT$$

da cui pertanto

$$P(T_2 = 4) P(CCTT) + P(CTCT) + P(TCCT)$$

– 5.8.1 – Distribuzione binomiale inversa.

Siano $E_1, E_2, \ldots, E_n$ successione di eventi stocasticamente indipendenti. Posto Y_n il numero di insuccessi all'i-esima prova, si ha

$$P(Y = n) = P(X = n+1) = pq^n \tag{10}$$

Nota: Y non gode di assenza di memoria. Non vale cioè

$$P(Y > n+m \mid Y > n) = P(Y > m)$$

Si dimostra infatti che $P(Y > n+m \mid Y > n) = q^m$ e $P(Y > m) = q^{m+1}$.

– 5.9 – Valore atteso e varianza di numeri aleatori discreti.

Due concetti fondamentali del calcolo probabilistico sono quelli di *valore atteso* e *varianza*.

Definizione: Dato X un numero aleatorio che per ogni valore assunto ha una certa probabilità p_i, si definisce valore atteso $\mathbb{E}(X)$ di X, la media pesata dei valori assunti da X, ciascuno con peso p_i. Cioè

$$\mathbb{E}(X) = \sum_{i=0}^{+\infty} x_i p_i$$

Definizione: Dato X un numero aleatorio di media μ, si definisce varianza

$$Var(X) = \mathbb{E}[(X - \mu)^2]$$

Più spesso la varianza è definita come

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

– 6 – Variabili aleatorie e distribuzioni continue.

Definizione: Sia X un numero aleatorio con densità $f(x)$, si definisce valore atteso di X come

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x \tag{11}$$

Esempio: Sia X un numero aleatorio con densità $f(x)$ descritta come a seguito

$$\begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{altrimenti} \end{cases}$$

Applicando l'Equazione (11), poiché per $x > 1 \land x < 0, f(x) = 0$, segue

$$\mathbb{E}(X) = \int_{0}^{1} 2x^2\,\mathrm{d}x = \frac{2}{3}$$

La varianza di un numero aleatorio continuo è definita analogamente ai numeri aleatori discreti: cioè

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

– 6.1 – Distribuzione uniforme.

Sia X un numero aleatorio, questi si dice avere distribuzione uniforme $X \sim U$ in un intervallo (a, b) se, considerata la sua densità $f(x)$ si ha che

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{altrimenti} \end{cases}$$

Inoltre la funzione di distribuzione $F(x)$ è definita come

$$F(x) = \begin{cases} 0, & x \leq a \\ \frac{1}{b-a}, & a < x < b \\ 1, & x \geq b \end{cases}$$

– 6.2 – Distribuzione esponenziale.

Sia X un numero aleatorio. Questi si dice avere distribuzione esponenziale, di parametro $\lambda > 0$, se la densità $f(x)$ di X è definita come segue.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{se } x \geq 0 \\ 0, & \text{altrimenti} \end{cases}$$

Considerando invece la distribuzione $F(x)$, si ha che

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_0^x \lambda e^{\lambda x} \, \mathrm{d}t = 1 - e^{\lambda x}, \quad x \geq 0 \end{aligned}$$

Nota: Da opportuni calcoli si può dimostrare che $\lambda = \frac{1}{\mathbb{E}(X)}$. Si dimostra inoltre che se $X \sim Exp(\lambda)$, questi gode di assenza di memoria, e vale il viceversa.

– 6.2.1 – Funzione di sopravvivenza.

Sia Xun numero aleatorio continuo; si definisce $S(X)$ la probabilità che $X > x, \forall x \in \mathbb{R}$. Cioè

$$S(X) = P(X > x), \qquad \forall x \in \mathbb{R}$$

Da ciò segue

$$S(X) = \int_x^\infty f(t) \, \mathrm{d}t$$

Inoltre se $X \sim Exp(\lambda)$, segue

$$S(X) = \begin{cases} 1, & \text{se } x \leq 0 \\ e^{-\lambda x}, x > 0 \end{cases}$$

cioè

$$S(X) = 1 - F(X)$$

– 6.3 – Funzione e distribuzioni Gamma.

Funzione Gamma.

La funzione Gamma è definita come

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} \, \mathrm{d}y$$

Se opportunamente integrata si ha, per $\alpha \in \mathbb{Z}$, che

$$\Gamma(\alpha) = (\alpha - 1)!$$

Distribuzione Gamma.

Sia X un numero aleatorio. Questa si dice avere distribuzione Gamma, di parametri (α, λ), se

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, x \geq 0 \\ 0, \text{ altrimenti} \end{cases}$$

Inoltre se $\alpha \in \mathbb{N}$, la distribuzione Gamma è utilizzata per quantizzare il tempo di attesa prima che si verifichino n eventi.

Sia T_n il tempo necessario al verificarsi dell'n-esimo evento. Segue

$$P(T_n \leq x) \iff P(N(x) \geq n)$$

ove $N(x)$ è il numero di eventi verificatesi in $[0, x]$.

Da cui

$$\begin{aligned} P(T_n \leq x) &= P(N(x) \geq x) \\ &= \sum_{i=n}^{\infty} P(N(x) = i) \\ &= \sum_{i=n}^{\infty} \frac{e^{-\lambda x}(\lambda x)^{\lambda i}}{i!} \end{aligned}$$

da cui derivando per x, si dimostra che

$$P(T_n \leq x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}$$

Esempio: Sia T_n il tempo di attesa utile all'arrivo dell'n-simo cliente, con distribuzione Gamma $\lambda = 2$. Si calcoli

1. $T_1 > 10$;
2. $T_5 \leq 30$;
3. $T_3 > 10$.

1. Si osserva che
$$\begin{aligned} P(T_1 < 10) &= \int_0^\infty \frac{2xe^{-2x}}{0!}\,\mathrm{d}x \\ &= \int_0^\infty 2xe^{-2x}\,\mathrm{d}x = e^{-20} \end{aligned}$$

2. Ricordando che
$$\int_0^\infty \frac{2xe^{-2x}}{(n-1)!}\,\mathrm{d}x = 1$$
segue
$$P(T_5 \leq 30) = 1 - P(T_5 > 30) = P(N_{30} \geq 5)$$
da cui
$$P(T_5 \leq 30) = \sum_{i=0}^{4} \frac{(2 \cdot 30)^i}{i!} e^{-60}$$

3. Applicando un ragionamento analogo al punto precedente
$$\begin{aligned} P(T_3 > 10) &= P(N_{10} < 3) \\ &= \sum_{i=0}^{2} \frac{(2 \cdot 30)^i}{i!} e^{-20} = 221e^{-20} \end{aligned}$$