"Sapienza" Università di Roma
Ingegneria dell'Informazione,
Informatica e Statistica
Dipartimento di Informatica

Calcolo delle Probabilità

Appunti integrati con il libro "The Probability Tutoring Book", Carol Ash

Autore
Simone Bianco

16 settembre 2025

Indice

Informazioni e Contatti

Appunti e riassunti personali raccolti in ambito del corso di *Calcolo delle Probabilità* offerto dal corso di laurea in Informatica dell'Università degli Studi di Roma "La Sapienza".

Ulteriori informazioni ed appunti possono essere trovati al seguente link: **https://github.com/Exyss/university-notes**. Chiunque si senta libero di segnalare incorrettezze, migliorie o richieste tramite il sistema di Issues fornito da GitHub stesso o contattando in privato l'autore :

- Email: **bianco.simone@outlook.it**
- LinkedIn: **Simone Bianco**

Gli appunti sono in continuo aggiornamento, pertanto, previa segnalazione, si prega di controllare se le modifiche siano già state apportate nella versione più recente.

Prerequisiti consigliati per lo studio:

Apprendimento del materiale relativo al corso *Metodi Matematici per l'Informatica*

Licence:

1

Insieme ambiente, Esiti ed Eventi

Prima di poter parlare di probabilità, è necessario definire ciò di cui essa si occupa. Il **calcolo di una probabilità** corrisponde allo studio di un fenomeno osservabile esclusivamente dal punto di vista della possibilità o meno del suo verificarsi.

Consideriamo ad esempio il lancio di una moneta. Tale fenomeno può avere **solo due esiti**, ossia testa o croce. Possiamo rappresentare tale fenomeno sottoforma di insieme, dove i suoi elementi sono tutti gli esiti possibili:

$$S : \{T, C\}$$

Effettuando un esperimento su tale insieme, ossia un lancio, il risultato di tale esperimento rientrerà in un **numero finito di esiti**, rappresentabili tramite un insieme. Tale esperimento viene detto **aleatorio**, mentre l'insieme di tutti gli esiti possibili viene detto **insieme ambiente (o spazio campionario)**.

Consideriamo ora il lancio di un dado. Anche in questo caso, il numero di esiti risulta essere finito: può uscire solo una faccia avente da uno a sei pallini. **Enumeriamo** quindi tutti gli esiti possibili associando un numero ad ogni esito:

$$S : \{\bullet,\ \bullet\bullet,\ \bullet\bullet\bullet,\ \bullet\bullet\bullet\bullet,\ \bullet\bullet\bullet\bullet\bullet,\ \bullet\bullet\bullet\bullet\bullet\bullet\} \quad \longrightarrow \quad S : \{1, 2, 3, 4, 5, 6\}$$

Analogamente, possiamo enumerare gli esiti del lancio di una moneta:

$$S : \{T, C\} \quad \longrightarrow \quad S : \{0, 1\}$$

Consideriamo ora l'**insieme** A contenente le facce di un dado aventi un numero di pallini inferiore o uguale a tre. Possiamo rappresentare tale insieme in **tre modi**:

- Per **enumerazione**, ossia $A : \{1, 2, 3\}$
- Per **descrizione**, ossia $A : \{\texttt{facce di un dado il cui valore è al massimo 3}\}$
- Per **notazione matematica**, ossia $A : \{x \in S \mid x \leq 3\}$

Abbiamo quindi definito gli elementi di A come appartenenti ad S ($x \in S$), dove S è l'insieme ambiente contenente tutti gli esiti possibili del lancio di un dado. Dunque, ne segue che $A \subset S$ (dunque che $x \in B \implies x \in S$), ossia è un **sottoinsieme dell'insieme ambiente**, che definiamo come **evento**. L'insieme A, quindi, corrisponde all'evento in cui esce una faccia minore o uguale a tre.

Definizione 1.1: Evento

Un **evento** corrisponde ad un **sottoinsieme dell'insieme ambiente**, ossia dell'insieme contenente tutti i possibili esiti di un fenomeno.

1.1 Proprietà degli eventi

Consideriamo l'evento in cui esce una faccia pari. Definiamo tale evento come:

$$A : \{x \in S \mid x\%2 = 0\} : \{2, 4, 6\}$$

Riprendiamo anche l'evento già visto in cui esce una faccia minore o uguale a 3:

$$B : \{x \in S \mid x \leq 3\} : \{1, 2, 3\}$$

Definiti questi due eventi, possiamo prendere in considerazione l'**evento unione** tra i due, ossia l'evento in cui esce una faccia pari **oppure** minore o uguale a 3:

$$C : A \cup B : \{1, 2, 3, 4, 6\} \text{ dove } x \in A \cup B \iff x \in A \vee x \in B$$

Analogamente, possiamo prendere in considerazione l'**evento intersezione** tra i due, ossia l'evento in cui esce una faccia pari **e anche** minore o uguale a 3:

$$D : A \cap B : \{2\} \text{ dove } x \in A \cap B \iff x \in A \wedge x \in B$$

Notiamo come quest'ultimo evento corrisponda ad un **singleton** (o singoletto), ossia un insieme di un solo elemento. Tale evento viene detto **evento elementare**.

Immaginiamo ora di voler descrivere l'evento in cui esce una faccia dispari. Come sappiamo, un numero dispari non è altro che un numero non pari. Definiamo quindi tale evento come **evento complementare** dell'evento in cui escono facce pari:

$$A^c : \{x \in S \mid x \notin A\}$$

Attenzione: è necessario sottolineare come non basti definire l'evento delle facce dispari come l'evento contenente tutti gli esiti che non sono nell'evento delle facce pari (dunque $A^c \neq \{x \notin A\}$), poiché ciò includerebbe anche gli esiti esterni all'insieme ambiente. Dunque, quando si parla di **evento complementare**, tale evento deve sempre essere **rapportato all'insieme ambiente** (dunque $x \in A^c \implies x \in S$).

Ovviamente, da tale definizione di evento complementare ne segue che l'**evento complementare dell'evento complementare di A** sia l'evento A stesso:

$$(A^c)^c = A$$

Un ulteriore modo per poter definire un evento complementare è tramite l'**esclusione**: eliminando tutti gli esiti appartenenti all'evento A dall'insieme ambiente S, otteniamo l'evento complementare di A:

$$A^c : S - A$$

Volendo rappresentare l'evento contenente le facce minori o uguali a tre e non pari, possiamo definire tale evento in **due modi**:

- L'intersezione tra l'evento delle facce minori o uguali a tre e l'evento delle facce dispari (ossia il complementare delle facce pari)
$$E : B \cap A^c$$
- L'evento contenente gli esiti minori o uguali a tre esclusi gli esiti contenuti nell'evento delle facce pari
$$E : B - A$$

Dunque, ne traiamo che:

$$B - A = B \cap A^c$$

Trattandosi sostanzialmente di insiemi, gli eventi godono anche delle altre proprietà ad essi legati:

- **Proprietà disgiuntiva**
$$A \cap A^c = \emptyset$$
- **Proprietà associativa**
$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{e} \quad (A \cap B) \cap C = A \cap (B \cap C)$$
- **Proprietà distributiva**
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \quad \text{e} \quad (A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$
- **Legge di De Morgan**
$$(A \cup B)^c = A^c \cap B^c \quad \text{e} \quad (A \cap B)^c = A^c \cup B^c$$

Dimostrazione:

Ricordiamo che, nell'ambito dell'insiemistica, la notazione $A = B$ indica che l'insieme A coincide esattamente con l'insieme B. Tale affermazione può essere ricondotta alla condizione $A \subseteq B \wedge B \subseteq A$, poiché l'unico caso possibili in cui A è sottoinsieme proprio di B e B è sottoinsieme proprio di A è quando A e B coincidono.

Dunque, per dimostrare che $(A \cup B)^c = A^c \cap B^c$, è sufficiente dimostrare che:

- $(A \cup B)^c \subseteq A^c \cap B^c$
- $A^c \cap B^c \subseteq (A \cup B)^c$

Consideriamo la seguente unione:

$$A_1 \cup A_2 \cup ... \cup A_n = \bigcup_{i=1}^{n} A_i$$

Se un elemento x appartiene al complementare di tale unione, allora ne segue che esso non appartenga all'unione in se

$$x \in \left(\bigcup_{i=1}^{n} A_i \right)^c \implies x \notin \bigcup_{i=1}^{n} A_i$$

A sua volta, ciò è possibile solo se l'elemento x appartenga al complementare di qualsiasi insieme appartenente a tale unione:

$$x \notin \bigcup_{i=1}^{n} A_i \iff \forall A_i \text{ si ha che } x \in (A_i)^c$$

Quest'ultima condizione, infine, implica che:

$$\forall A_i \text{ si ha che } x \in (A_i)^c \implies x \in \bigcap_{i=1}^{n} (A_i)^c$$

Dunque, concludiamo che:

$$\left(\bigcup_{i=1}^{n} A_i \right)^c \subseteq \bigcap_{i=1}^{n} (A_i)^c$$

La stessa condizione, tuttavia, implica che non esiste un indice i tale che l'elemento x possa essere in A_i

$$\forall A_i \text{ si ha che } x \in (A_i)^c \implies \nexists i \mid x \in A_i$$

Dunque, considerando l'unione di tutte gli A_i insiemi, l'elemento x non può trovarsi in essa, dunque esso sarà necessariamente situato nel complementare di tale unione:

$$\nexists i \mid x \in A_i \implies x \in \left(\bigcup_{i=1}^{n} A_i\right)^c$$

Dunque, concludiamo che:

$$\bigcap_{i=1}^{n} (A_i)^c \subseteq \left(\bigcup_{i=1}^{n} A_i\right)^c$$

Poiché entrambe le condizioni sono verificate, otteniamo che:

$$\left[\left(\bigcup_{i=1}^{n} A_i\right)^c \subseteq \bigcap_{i=1}^{n} (A_i)^c\right] \land \left[\bigcap_{i=1}^{n} (A_i)^c \subseteq \left(\bigcup_{i=1}^{n} A_i\right)^c\right] \iff \left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} (A_i)^c$$

$\square$

- **Esclusione disgiuntiva (XOR)**

$$(A \cup B) - (A \cap B) = (A - B) \cup (B - A)$$

Dimostrazione:

$$(A \cup B) - (A \cap B)$$
$$(A \cup B) \cap (A \cap B)^c$$
$$[A \cap (A \cap B)^c] \cup [B \cap (A \cap B)^c]$$
$$[A \cap (A^c \cup B^c)] \cup [B \cap (A^c \cup B^c)]$$
$$[(A \cap A^c) \cup (A \cap B^c)] \cup [(B \cap A^c) \cup (B \cap B^c)]$$
$$[\emptyset \cup (A \cap B^c)] \cup [(B \cap A^c) \cup \emptyset]$$
$$(A \cap B^c) \cup (B \cap A^c)$$
$$(A - B) \cup (B - A)$$

$\square$

1.2 Cardinalità e Funzioni indicatrici

Analogamente agli insiemi, con il termine **cardinalità** indichiamo il **numero di esiti contenuti in un evento**. Per essere numerabile, ovviamente, un evento deve possedere una **quantità finita di eventi**.

Indichiamo la cardinalità di un evento con la notazione:

$$|A| = n$$

dove A è l'evento e n è la sua cardinalità.

Dato un evento A, invece, definiamo come **funzione indicatrice di A** (indicata come I_A) la funzione che preso un elemento x in input restituisce **1 se l'elemento appartiene all'evento**, oppure **0 altrimenti**.

Definizione 1.2: Funzione indicatrice

Dato un evento A, la sua **funzione indicatrice** corrisponde a

$$I_A : S \to \{0, 1\} : x \mapsto I_A(x)$$

$$I_A(x) = \begin{cases} 1 & \texttt{se } x \in A \\ 0 & \texttt{altrimenti} \end{cases}$$

Da tale definizione, quindi, ne segue logicamente che:

$$I_A(x) = 1 \iff I_{A^c}(x) = 0$$

$$I_A(x) = 0 \iff I_{A^c}(x) = 1$$

Consideriamo ora le due funzioni indicatrici I_A e I_B. La **funzione indicatrice dell'evento intersezione** $A \cap B$ può essere definita come:

$$I_{A\cap B}(x) = \begin{cases} 1 & \texttt{se } x \in A \cap B \\ 0 & \texttt{altrimenti} \end{cases} = \begin{cases} 1 & \texttt{se } x \in A \land x \in B \\ 0 & \texttt{altrimenti} \end{cases}$$

Poiché tale funzione deve valere 1 solo se $x \in A$ e $x \in B$, ne segue che ciò possa essere **possibile solo se per lo stesso elemento x si ha che** $I_A x = 1$ **e** $I_B(x) = 1$.

Possiamo quindi definire $I_{A\cap B}$ anche come il prodotto tra $I_A(x)$ e $I_B(x)$, poiché nel caso in cui **una (o entrambe) delle due funzioni restituisca 0** allora anche la funzione indicatrice dell'unione **restituirà 0**.

$$I_{A\cap B} = I_A \cdot I_B$$

Vediamo ora la **funzione indicatrice dell'evento unione** $A \cup B$, definita come:

$$I_{A\cup B}(x) = \begin{cases} 1 & \texttt{se } x \in A \cup B \\ 0 & \texttt{altrimenti} \end{cases} = \begin{cases} 1 & \texttt{se } x \in A \vee x \in B \\ 0 & \texttt{altrimenti} \end{cases}$$

Cerchiamo quindi un modo matematico per poter calcolare facilmente $I_{A\cup B}(x)$ tramite $I_A(x)$ e $I_B(x)$. Intuitivamente, si potrebbe pensare che la somma tra le due funzioni indicatrici di A e B corrisponda al valore dato da quella dell'unione. Tuttavia, notiamo che:

- Se $I_A(x) = 0$ e $I_B(x) = 0$, allora $I_A(x) + I_B(x) = 0$
- Se $I_A(x) = 1$ e $I_B(x) = 0$, allora $I_A(x) + I_B(x) = 1$
- Se $I_A(x) = 0$ e $I_B(x) = 1$, allora $I_A(x) + I_B(x) = 1$
- Se $I_A(x) = 1$ e $I_B(x) = 1$, allora $I_A(x) + I_B(x) = 2$

Notiamo quindi come l'ultimo caso dia un **risultato sbagliato** rispetto all'output che vorremmo (ossia 1). Il motivo di ciò può essere spiegato comodamente tramite il **principio di inclusione esclusione**:

- Consideriamo i due insiemi $A = \{1, 2, 3, 4\}, B = \{0, 2, 4, 5\}$
- L'unione tra i due insiemi risulterà essere $A \cup B = \{0, 1, 2, 3, 4, 5\}$
- Si ha quindi che $|A \cup B| \neq |A| + |B|$ (dunque che $6 \neq 4 + 4$). Ciò avviene poiché è stata **conteggiata due volte l'intersezione** $A \cap B$, poiché ogni elemento in tale intersezione (ossia $\{2, 4\}$) è stato contato sia nella **cardinalità di** A sia nella **cardinalità di B**.
- Per ri-bilanciare il conto, quindi, è necessario **sottrarre una volta tale intersezione**, in modo da conteggiarla in una sola delle due cardinalità

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Analogamente, quindi, il **calcolo esatto della funzione indicatrice unione** sarà:

$$I_{A\cup B} = I_A + A_B - I_{A\cap B}$$

Notiamo quindi una stretta relazione tra **cardinalità** e **funzione indicatrice**. Difatti, potremmo usare quest'ultima per descrivere la prima come la **somma di tutti i valori dati dalla funzione indicatrice dell'evento stesso** per ogni elemento dell'insieme:

$$A = \{x_1, x_2, ..., x_n\}$$

$$|A| = \sum_{i=1}^{n} I_A(x_i)$$

Nel caso in cui gli eventi da trattare siano **più di due**, possiamo utilizzare la **proprietà associativa** di cui essi godono per calcolare la loro cardinalità.

- Vogliamo calcolare la **cardinalità dell'insieme** $A \cap B \cap C$. Tramite la proprietà associativa, sappiamo che:

$$(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C$$
$$(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$$

  Dunque, ne traiamo che

$$|A \cap B \cap C| = |(A \cap B) \cap C| = |A \cap B| \cdot |C| = |A| \cdot |B| \cdot |C|$$

  Poiché ne faremo uso frequentemente, per comodità riscriviamo il **prodotto delle cardinalità** come
$$|A \cap B \cap C| = |A| \cdot |B| \cdot |C| = |ABC|$$

- Per la **cardinalità dell'insieme** $A \cup B \cup C$, le cose risultano un po' più complesse. Anche in questo caso, procediamo con la proprietà associativa:

$$\begin{aligned}
|A \cup B \cup C| &= |(A \cup B) \cup C| = |A \cup B| + |C| - |(A \cup B) \cap C| = \\
&= |A| + |B| - |A \cap B| + |C| - |(A \cup B) \cap C| = \\
&= |A| + |B| + |C| - |A \cap B| - |(A \cap C) \cup (B \cap C)| = \\
&= |A| + |B| + |C| - |A \cap B| - \{|A \cap C| + |B \cap C| - |(A \cap B) \cap (B \cap C)|\} = \\
&= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C| = \\
&= |A| + |B| + |C| - |AB| - |AC| - |BC| + |ABC|
\end{aligned}$$

  Notiamo come il risultato corrisponda ancora una volta al **ri-aggiustamento di un doppio conteggio**: sommando i tre insiemi contiamo 2 volte ognuna delle intersezioni a due, necessitando quindi che ognuno di essi venga sottratto una volta. In questo modo, però, abbiamo **aggiunto 3 volte** l'intersezione a tre (contando la somma tra i tre insiemi) e **sottratto 3 volte** la stessa intersezione (rimuovendo le tre intersezioni a due), necessitando quindi che essa venga **ri-conteggiata**

$$|A \cup B \cup C| = |A| + |B| + |C| - |AB| - |AC| - |BC| + |ABC|$$

- Analogamente, la **cardinalità dell'insieme** $A \cup B \cup C \cup D$, corrisponderà a:

$$\begin{aligned}
|A \cup B \cup C \cup D| &= |(A \cup B \cup C) \cup D| = \\
&= ... = \\
&= |A| + |B| + |C| + |D| - |AB| - |AC| - |AD| - |BC| - |BD| - |CD| + \\
&+ |ABC| + |ABD| + |ACD| + |BCD| - |ABCD|
\end{aligned}$$

Notiamo quindi la presenza di un certo **pattern** durante il calcolo della cardinalità di un'unione:

- **Aggiungiamo** gli n insiemi
- **Sottraiamo** tutte le intersezioni a due
- **Aggiungiamo** tutte le intersezioni a tre
- **Sottraiamo** tutte le intersezioni a quattro
- ...

Infatti, possiamo riscrivere la cardinalità delle due unioni viste precedentemente anche nel seguente **modo compatto**

$$|A \cup B \cup C| = \sum_{i=1}^{3} |A_i| - \sum_{1 \leq i < j \leq 3} |A_i A_j| + \sum_{1 \leq i < j < k \leq 3} |A_i A_j A_k|$$

$$|A \cup B \cup C \cup D| = \sum_{i=1}^{4} |A_i| - \sum_{1 \leq i < j \leq 4} |A_i A_j| + \sum_{1 \leq i < j < k \leq 4} |A_i A_j A_k| + \sum_{1 \leq i < j < k < h \leq 4} |A_i A_j A_k A_h|$$

dove, ad esempio, la notazione $1 \leq i < j \leq 3$ sottostante alla prima sommatoria indica **tutte tuple di valori possibili** in un range di numeri che va da 1 a 3, dove 3 è il **numero degli insiemi nell'unione**. Analogamente, la notazione $1 \leq i < j < k \leq 3$ indica tutte le triple di valori possibili e così via.

Notiamo anche come il segno di tali sommatorie sia alternato. Difatti, **aggiungiamo** tutte le m-uple in cui m **è un numero dispari**, mentre **sottraiamo** tutte le m-uple in cui m **è un numero pari**. Possiamo quindi generalizzare l'intero concetto a n **insiemi**, utilizzando la seguente notazione iper-compatta:

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{k=1}^{n} \left((-1)^{k+1} \cdot \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} |A_{i_1} \cdot A_{i_2} \cdot \ldots \cdot A_{i_k}| \right)$$

Alcune volte, però, è difficile calcolare l'esatta cardinalità di alcune unioni tra eventi. In questi casi si preferisce quindi un **calcolo approssimativo**, stabilendo un **limite inferiore** ed un **limite inferiore**, cercando di ridurre il più possibile la differenza tra di essi:

$$\underbrace{\sum_{i=0}^{n} |A_i| - \sum_{1 \leq i < j \leq n} |A_i A_j|}_{\text{Limite inferiore}} \leq \left| \bigcup_{i=1}^{n} A_i \right| \leq \underbrace{\sum_{i=0}^{n} |A_i|}_{\text{Limite superiore}}$$

2

Modello standard della probabilità

2.1 Probabilità di un evento

Considerato l'insieme ambiente "lancio di una moneta non truccata", definito come $S : \{T, C\}$, ci chiediamo quale sia la **probabilità** che si verifichi l'evento "Testa", dunque definito come $A : \{T\}$. Intuitivamente, concludiamo che la probabilità di tale evento sia 50%, dunque $\frac{1}{2}$, poiché tale evento può generare solo due risultati: T o C. Descriviamo quindi tale probabilità come:

$$P(T) = \frac{1}{2}$$

Analogamente, ci chiediamo quale sia la **probabilità** che si verifichi l'evento "faccia maggiore di 2", ossia $A : \{x \in S \mid x > 2\}$, nell'insieme ambiente "lancio di un dado", dunque $S : \{1, 2, 3, 4, 5, 6\}$. Anche in questo caso, il calcolo di tale probabilità risulta intuitivo, poiché vi sono 4 possibili esiti favorevoli rispetto ai 6 esiti totali:

$$P(\texttt{faccia maggiore di 2}) = \frac{4}{6} = \frac{2}{3}$$

Potremmo quindi pensare alla seguente **formula generica**:

$$P(x) = \frac{\texttt{Casi favorevoli}}{\texttt{Casi totali}}$$

Tuttavia, **tale formula risulta incorretta**, poiché non sempre tutti i casi risultano essere **equiprobabili**:

- Consideriamo il lancio di due monete, avente **tre esiti**: due volte testa, due volte croce, una volta testa ed una volta croce

$$S : \{\texttt{Due teste}, \texttt{Due croci}, \texttt{Una testa ed una croce}\}$$

- Se volessimo calcolare la probabilità dell'evento "Una volta testa ed una volta croce", otterremmo il seguente risultato:

$$P(\texttt{Una testa ed una croce}) = \frac{\texttt{Casi favorevoli}}{\texttt{Casi totali}} = \frac{1}{3}$$

- Tuttavia, tale calcolo risulta inesatto, poiché a livello intuitivo la probabilità dovrebbe essere più alta, poiché abbiamo due modi di verificare tale evento:
  - Primo Lancio: T - Secondo Lancio: C
  - Primo Lancio: C - Secondo Lancio: T
- Difatti, definendo l'insieme ambiente corrispondente al lancio di due monete, otteniamo **4 reali esiti possibili**:

$$S : \{TT, CC, TC, CT\}$$

- Analogamente, definendo rigorosamente l'evento "Una volta testa ed una volta croce", otteniamo **2 esiti possibili**:

$$A : \{TC, CT\}$$

- Dunque, il reale calcolo della probabilità richiesta corrisponderà a:

$$P(\texttt{Una testa ed una croce}) = \frac{|\texttt{Una testa ed una croce}|}{|\texttt{Lancio di due monete}|} = \frac{|A|}{|S|} = \frac{2}{4} = \frac{1}{2}$$

Definizione 2.1: Calcolo standard di una Probabilità

Dato un insieme ambiente S ed un evento A, dove $A \subseteq S$, la **probabilità** di tale evento corrisponde al **rapporto tra la cardinalità dell'evento e la cardinalità dell'insieme ambiente**:

$$P(A) = \frac{|A|}{|S|}$$

Esempi:

Considerando un mazzo di carte standard (52 carte), vogliamo calcolare la probabilità dell'evento "Pesco un asso o una carta con seme picche".

- Definiamo la situazione in modo rigoroso:

$$S : \{\texttt{Mazzo di carte standard}\}$$

$$A : \{\texttt{Carta asso}\} \text{ dove } A \subseteq S$$

$$B : \{\texttt{Carta picche}\} \text{ dove } B \subseteq S$$
$$A \cup B : \{\texttt{Carta asso o picche}\}$$

- Applichiamo la formula del **modello standard della probabilità**:

$$P(A \cup B) = \frac{|A \cup B|}{|S|} = \frac{|A| + |B| - |A \cap B|}{|S|}$$

- In un mazzo standard sono presenti 4 assi (uno per ogni seme) e 13 carte per ogni seme (dunque 13 picche). Ovviamente, questo implica che solo uno dei 4 assi sia un asso di picche.
- Il calcolo richiesto, quindi, si riconduce a:

$$P(A \cup B) = \frac{|A| + |B| - |A \cap B|}{|S|} = \frac{4 + 13 - 1}{52} = \frac{16}{52} = \frac{4}{13}$$

Consideriamo ora l'insieme ambiente "Lancio di due dadi", i cui elementi sono denotati come coppie XY, dove X è il primo lancio e Y il secondo:

$$S : \left\{ \begin{matrix} 11 & 12 & 13 & 14 & 15 & 16 \\ 21 & 22 & 23 & 24 & 25 & 26 \\ 31 & 32 & 33 & 34 & 35 & 36 \\ 41 & 42 & 43 & 44 & 45 & 46 \\ 51 & 52 & 53 & 54 & 55 & 56 \\ 61 & 62 & 63 & 64 & 65 & 66 \end{matrix} \right\}$$

1. Calcolare la probabilità dell'evento "Somma dei lanci uguale a 5"
   - Definiamo l'evento come:

   $$A : \{XY \in S \mid X + Y = 5\}$$

   - Analizzando tutti gli esiti possibili dell'insieme ambiente, gli esiti appartenenti all'evento corrispondono a:

$$S : \left\{ \begin{matrix} 11 & 12 & 13 & \color{red}{14} & 15 & 16 \\ 21 & 22 & \color{red}{23} & 24 & 25 & 26 \\ 31 & \color{red}{32} & 33 & 34 & 35 & 36 \\ \color{red}{41} & 42 & 43 & 44 & 45 & 46 \\ 51 & 52 & 53 & 54 & 55 & 56 \\ 61 & 62 & 63 & 64 & 65 & 66 \end{matrix} \right\} \implies A : \{14, 23, 32, 41\}$$

- La probabilità dell'evento, quindi corrisponde a:

$$P(A) = \frac{|A|}{|S|} = \frac{4}{36} = \frac{1}{9}$$

2. Calcolare la probabilità dell'evento "Secondo lancio maggiore del primo"

- Definiamo l'evento come:

$$A : \{XY \in S \mid Y > X\}$$

- Analizzando tutti gli esiti possibili dell'insieme ambiente, gli esiti appartenenti all'evento corrispondono a:

$$S : \left\{ \begin{array}{cccccc} 11 & \color{red}{12} & \color{red}{13} & \color{red}{14} & \color{red}{15} & \color{red}{16} \\ 21 & 22 & \color{red}{23} & \color{red}{24} & \color{red}{25} & \color{red}{26} \\ 31 & 32 & 33 & \color{red}{34} & \color{red}{35} & \color{red}{36} \\ 41 & 42 & 43 & 44 & \color{red}{45} & \color{red}{46} \\ 51 & 52 & 53 & 54 & 55 & \color{red}{56} \\ 61 & 62 & 63 & 64 & 65 & 66 \end{array} \right\} \Longrightarrow A : \left\{ \begin{array}{ccccc} 12 & 13 & 14 & 15 & 16 \\ 23 & 24 & 25 & 26 & 34 \\ 35 & 36 & 45 & 46 & 56 \end{array} \right\}$$

- La probabilità dell'evento, quindi corrisponde a:

$$P(A) = \frac{|A|}{|S|} = \frac{15}{36} = \frac{5}{12}$$

k

Vogliamo calcolare la probabilità dell'evento **"Almeno un lancio diverso"** effettuato su **"Cinque lanci di dadi"**. A differenza degli esempi della sezione precedente, risulterebbe estremamente lungo elencare ed analizzare tutti gli esiti contenuti nell'**insieme ambiente** poiché esso contiene 6^5 esiti possibili.

Possiamo definire quindi tale insieme come il **prodotto cartesiano ripetuto 5 volte sull'insieme "Un lancio di dado"**

$$I : \{1, 2, 3, 4, 5, 6\}$$
$$S : (I \times I \times I \times I \times I) \to (a, b, c, d, e) : \{\texttt{Cinque lanci di dadi}\}$$

Definiamo quindi l'evento A come "Almeno un lancio diverso". Ovviamente, calcolare la **cardinalità di tale evento** in modo diretto risulta complesso (ricordiamo che $|S| = 6^5$). Scegliamo quindi un **approccio più semplice**: per definizione stessa di **insieme complementare**, si ha che

$$|A^c| = |S| - |A| \iff |A| = |S| - |A^c|$$

Quindi, possiamo **riscrivere** il calcolo della **probabilità dell'evento A** come:

$$P(A) = \frac{|A|}{|S|} = \frac{|S| - |A^c|}{|S|} = 1 - \frac{|A^c|}{|S|}$$

A questo punto, ci basta calcolare la **cardinalità dell'evento complementare di A** (ossia l'evento "Lanci tutti uguali tra loro"), il quale possiede solo 6 esiti possibili:

$$A^c = \{11111, 22222, 33333, 44444, 55555, 66666\}$$

Dunque, la probabilità dell'evento A corrisponde a:

$$P(A) = 1 - \frac{|A^c|}{|S|} = 1 - \frac{6}{6^5} \approx 99.92\%$$

Teorema 2.1: Passaggio al complemento

Dato un insieme ambiente S ed un evento A, dove $A \subseteq S$, la probabilità dell'evento corrisponde a $1 - P(A^c)$, dove A^c è l'evento complementare ad A su S:

$$P(A) = \frac{|A|}{|S|} = 1 - \frac{|A^c|}{|S|} = 1 - P(A^c)$$

Esempio:

Considerando l'insieme ambiente "Tutte le coppie (carta rossa, carta blu)", vogliamo sapere la probabilità dell'evento "almeno un asso".

Definiamo quindi l'insieme ambiente come:

$$A : \{\texttt{Carte rosse}\} \longrightarrow |A| = 52$$
$$B : \{\texttt{Carte blu}\} \longrightarrow |B| = 52$$
$$S : \{\texttt{Coppie rosso-blu}\} : A \times B \longrightarrow |S| = 52 \cdot 52$$

Successivamente, definiamo l'evento "almeno un asso" come:

$$E : \{\texttt{coppie con asso rosso}\} \longrightarrow |E| = 4 \cdot 52$$
$$F : \{\texttt{coppie con asso blu}\} \longrightarrow |F| = 4 \cdot 52$$
$$C : \{\texttt{coppie con almeno un asso}\} : E \cup F \longrightarrow |C| = |E| + |F| - |E \cap F|$$
$$E \cap F : \{\texttt{coppie con entrambi gli assi}\} \longrightarrow |E \cap F| = 4 \cdot 4$$

La probabilità dell'evento, quindi, corrisponde a:

$$P(C) = \frac{|C|}{|S|} = \frac{|E| + |F| - |E \cap F|}{|S|} = \frac{4 \cdot 52 + 4 \cdot 52 - 4 \cdot 4}{52 \cdot 52} = \frac{400}{2704}$$

Di contraltare, il calcolo tramite il **passaggio al complemento** risulta **immediato** rispetto al precedente:

$$C^c = \{\texttt{coppie senza assi}\} \longrightarrow |C^c| = 48 \cdot 48$$

$$P(C) = 1 - P(C^c) = 1 - \frac{2304}{2704} = \frac{400}{2704}$$

2.2 Analisi combinatoria

2.2.1 Figure della combinatoria

Nella sezione precedente (e in quelle successive) abbiamo utilizzato le **figure della combinatoria** più comuni (disposizioni, anagrammi e combinazioni) per calcolare le cardinalità.

Di seguito, viene fornito un **breve ripasso** di tali figure. Per **approfondimenti**, si consiglia la sezione omonima degli appunti del corso di *"Metodi Matematici per l'Informatica"*.

Definizione 2.2: Disposizioni

Definiamo come **disposizione con ripetizione di ordine k di n oggetti** una sequenza ordinata $(x_1, x_2, ..., x_k)$ di k oggetti scelti tra gli n totali

$$D'_{n,k} = n \cdot n \cdot n \cdot ... \cdot n = n^k$$

Sia $1 \leq k \leq n$. Definiamo come **disposizione semplice di ordine** k **di** n **oggetti** una sequenza ordinata $(x_1, x_2, ..., x_k)$ di k oggetti distinti tra loro scelti tra gli n totali. Una disposizione semplice in cui $k = n$ viene detta **permutazione**.

$$D_{n,k} = n \cdot (n-1) \cdot ... \cdot (n-(k-1)) = \frac{n!}{(n-k)!}$$

Proposizione 2.1: Anagrammi

Gli **anagrammi** (o permutazioni con ripetizione) di un insieme di n lettere, in cui compaiono k gruppi di $n_1, n_2, ..., n_k$ lettere ripetute, corrispondono a

$$\#A = \frac{n!}{n_1! \cdot n_2! \cdot ... \cdot n_k!}$$

Definizione 2.3: Combinazione

Definiamo come **combinazione** un raggruppamento di k elementi, presi in qualsiasi ordine, formato a partire da n elementi distinti. Si distinguono in combinazioni **senza ripetizione**

$$C_{n,k} = \binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$$

e **con ripetizione**

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

Esempi:

- Dato un insieme di valori possibili $A : \{1, 2, 3, B, C, D\}$, le targhe da tre elementi possibili sono:

$$D'_{6,3} = 6 \cdot 6 \cdot 6 = 6^3$$

- Gli anagrammi della parola MISSISSIPPI sono:

$$\#A = \frac{11!}{4! \cdot 4! \cdot 2!}$$

- I modi di scegliere una mano di poker da 5 carte con un mazzo standard da 52 carte sono:

$$C_{52,5} = \binom{52}{5} = \frac{52!}{47! \cdot 5!} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5!}$$

- I sottoinsiemi possibili di un insieme di n elementi sono:

$$P = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} = \sum_{i=0}^{n} \binom{n}{i}$$

2.2.2 Estrazioni da un'urna di palline

Estrazioni con reinserimento

Immaginiamo di avere due tipologie di palline, alcune di **colore rosso** (la cui quantità viene indicata con R) ed altre di **colore bianco** (indicate con B). Considerando la somma N di tali palline, otteniamo che le rispettive **probabilità di estrazione** di ogni pallina, nel caso in cui esse vengano **reinserite nell'urna**, corrispondono a:

$$P(R) = \frac{R}{N} = 1 - P(B)$$
$$P(B) = \frac{B}{N} = 1 - P(R)$$

Poiché ad ogni estrazione le palline vengono reinserite nell'urna, la probabilità d'estrazione di ogni pallina rimane uguale. Dunque, la probabilità di **estrarre k volte una pallina dello stesso colore** corrisponde a:

$$P(\texttt{k volte pallina rossa}) = P\left(\bigcap_{i=1}^{k} R\right) = \frac{R}{N} \cdot \frac{R}{N} \cdot ... \cdot \frac{R}{N} = \left(\frac{R}{N}\right)^k$$

$$P(\texttt{k volte pallina bianca}) = P\left(\bigcap_{i=1}^{k} B\right) = \frac{B}{N} \cdot \frac{B}{N} \cdot ... \cdot \frac{B}{N} = \left(\frac{B}{N}\right)^k$$

Ad esempio, l'evento in cui viene estratta la sequenza $BBRRRBRR$ corrisponderà a:

$$P(BBRRRBRR) = P(B)^3 \cdot P(R)^3 \cdot P(B) \cdot P(R)^2 = P(R)^5 \cdot P(B)^3 = \left(\frac{R}{N}\right)^5 \cdot \left(\frac{B}{N}\right)^3 = \frac{R^5 \cdot B^3}{N^8}$$

Analogamente, la probabilità della sequenza $BRBRRBBB$ corrisponderà a:

$$P(BRBRRBBB) = P(B) \cdot P(R) \cdot P(B) \cdot P(R)^2 \cdot P(B)^3 = \left(\frac{R}{N}\right)^5 \cdot \left(\frac{B}{N}\right)^3 = \frac{R^5 \cdot B^3}{N^8}$$

Notiamo dunque come, effettuando n **estrazioni**, la probabilità di estrarre **una determinata sequenza ordinata** contenente k **elementi di un tipo** e $n - k$ **elementi di un altro**, corrisponde a:

$$P(\texttt{Seq. ord, con n estr. su 2 tipi}) = P(E_1)^k \cdot P(E_2)^{n-k}$$

Nel caso in cui, invece, **non ci interessi l'ordine** effettivo delle estrazioni, dunque qualsiasi sequenza di k elementi di un tipo e $n-k$ elementi di un altro, allora sarà necessario moltiplicare tale probabilità per il numero di **anagrammi di n elementi contenenti k e $n-k$ ripetizioni**, in modo da poter considerare tutte le sequenze possibili:

$$P(\texttt{k elem. 1, n-k elem. 2}) = \frac{n!}{(n-k)! \cdot k!} \cdot P(E_1)^k \cdot P(E_2)^{n-k}$$

Tuttavia, notiamo come il calcolo di tali anagrammi corrisponda ad una **scelta di k elementi su n**:

$$P(\texttt{k elem. 1, n-k elem. 2}) = \frac{n!}{(n-k)! \cdot k!} \cdot P(E_1)^k \cdot P(E_2)^{n-k} = \binom{n}{k} \cdot P(E_1)^k \cdot P(E_2)^{n-k}$$

Considerando la **somma di tutte le possibili k estrazioni non ordinate di elementi di un tipo**, dunque con $k = 0, 1, ..., n$, otteniamo la seguente sommatoria:

$$\texttt{P al variare di K} = \sum_{k=0}^{n} \binom{n}{k} \cdot P(E_1)^k \cdot P(E_2)]^{n-k}$$

Notiamo inoltre sommatoria corrisponde ad un **Binomio di Newton**, il cui teorema generale afferma che:

$$\sum_{k=0}^{n} \binom{n}{k} \cdot a^k \cdot b^{n-k} = (a+b)^n$$

Dunque, nel nostro caso otteniamo che:

$$\texttt{P al variare di K} = \sum_{k=0}^{n} \binom{n}{k} \cdot P(E_1)^k \cdot P(E_2)^{n-k} =$$
$$= [P(E_1) + P(E_2)]^n = [P(E_1) + (1 - P(E_1))]^n = 1^n = 1$$

Estrazioni senza reinserimento

Vediamo ora il caso in cui le palline estratte vengano rimosse dall'urna, diminuendo quindi il numero di palline totali.

Nella **prima estrazione**, la probabilità di estrarre una pallina rossa e la probabilità di estrarre una pallina bianca rimangono invariate:

$$P_1(R) = \frac{R}{N} = 1 - P_1 B)$$
$$P_1(B) = \frac{B}{N} = 1 - P_1(R)$$

Immaginiamo che nella prima estrazione sia stata estratta una pallina bianca. Di conseguenza, il numero delle palline bianche e delle palline totali viene ridotto di uno. Le nuove probabilità di estrazione sono:

$$P_2(R) = \frac{R}{N-1}$$
$$P_2(B) = \frac{B-1}{N-1}$$

Dunque, volendo calcolare la probabilità dell'**estrazione di due palline bianche di seguito**, otterremmo:

$$P(BB) = P_1(B) \cdot P_2(B) = \frac{B(B-1)}{N(N-1)}$$

E se l'**estrazione** delle due palline bianche avvenisse **in contemporanea**? Possiamo calcolare tale probabilità come il rapporto tra una scelta di 2 palline sulle B totali e una scelta di 2 palline sulle N totali:

$$P(\texttt{BB contemp}) = \frac{\binom{B}{2}}{\binom{N}{2}} = \frac{B!}{(B-2)! \cdot 2!} \cdot \frac{(N-2)! \cdot 2!}{N!} =$$
$$= \frac{B(B-1)(B-2)!}{(B-2)!} \cdot \frac{(N-2)!}{N(N-1)(N-2)!} = \frac{B(B-1)}{N(N-1)}$$

Notiamo quindi come la probabilità di un'**estrazione in contemporanea** sia **identica** a quella di un'**estrazione sequenziale**.

In linea più generale, quindi, possiamo dire che, se **non conta l'ordine** degli elementi estratti e se **non si hanno reinserimenti**, la probabilità di effettuare k estrazioni di un tipo e $n-k$ di un altro su un totale di n estrazioni corrisponde a:

$$P(\texttt{k estr. 1, n-k estr. 2}) = \frac{\binom{E_1}{k} \cdot \binom{E_2}{n-k}}{\binom{N}{n}}$$

Tale modello di calcolo viene detto **modello ipergeometrico**.

Metodo 2.1: Probabilità di n estrazioni da un'urna

Effettuando n estrazioni su un insieme di N elementi suddivisi in t tipologie ($\forall T_i$ dove $i \in [1, t]$), dove vengono estratti k_i elementi (dove $i \in [1, t]$) per ogni tipologia ($k_1 + k_2 + ... + k_t = n$) si ha:

- **Probabilità di estrarre una sequenza di ordine fissato (con reinserimenti)**:

$$P(\texttt{Seq. ord. fisso}) = \left(\frac{T_1}{N}\right)^{k_1} \cdot \left(\frac{T_2}{N}\right)^{k_2} \cdot ... \cdot \left(\frac{T_t}{N}\right)^{k_t} = \prod_{i=1}^{t} \left(\frac{T_i}{N}\right)^{k_i}$$

- **Probabilità di estrarre una sequenza di ordine casuale (con reinserimenti)**:

$$P(\texttt{Seq. ord. casuale}) = \frac{n!}{k_1! \cdot k_2! \cdot ... \cdot k_t!} \cdot P(\texttt{Seq. ord. fisso})$$

- **Probabilità di estrarre una sequenza di ordine casuale (senza reinserimenti)**:

$$P(\texttt{Seq. ord. cas. con rein.}) = \frac{\binom{T_1}{k_1} \cdot \binom{T_2}{k_2} \cdot ... \cdot \binom{T_t}{k_t}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \cdot \prod_{i=1}^{t} \binom{T_i}{k_i}$$

3

Definizione assiomatica della probabilità

Avendo sperimentato con alcune modalità di calcolo di alcune probabilità, possiamo ora dare una definizione assiomatica di essa.

Dato uno **spazio campionario** S, la probabilità che si verifichi un esito appartenente a tale spazio campionario corrisponde alla **somma delle probabilità di tutti gli eventi disgiunti** E **di** S (dove ricordiamo $E \subseteq S$).

$$P(S) = P(E_1) + P(E_2) + ... + P(E_n)$$

Dunque, possiamo definire la **probabilità di un evento** come una funzione del tipo

$$P : P(S) \to [0, 1]$$

Da tale definizione di probabilità possiamo ricavare una serie di **assiomi**:

- La **probabilità di un evento** E in S deve essere **compresa tra 0 ed 1**

$$0 \leq P(E) \leq 1, \forall E \subset S$$

- La **probabilità di** S **è sempre 1**

$$P(S) = 1$$

- Se $E_1, E_2, ..., E_n$ sono **eventi disgiunti due a due o mutualmente esclusivi** (ossia che $\forall i, j \in [0, n]$ dove $i \neq j$ si ha $A_i \cap A_j = \emptyset$), allora si ha che

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$$

Tramite tali assiomi, possiamo inoltre ricavare alcune **proprietà di calcolo delle probabilità**:

- La probabilità dell'**evento impossibile** è $P(\emptyset) = 0$:

$$\begin{aligned} 1 &= P(S) \\ 1 &= P(S \cup \emptyset) \\ 1 &= P(S) + P(\emptyset) \\ 1 &= 1 + P(\emptyset) \\ 0 &= P(\emptyset) \end{aligned}$$

- La probabilità dell'**evento complementare all'evento A** è $P(A^c) = 1 - P(A)$:

$$\begin{aligned} 1 &= P(S) \\ 1 &= P(A \cup A^c) \\ 1 &= P(A) + P(A^c) \\ 1 - P(A) &= P(A^c) \end{aligned}$$

- Se $A \subseteq B$, allora $P(A) \leq P(B)$:

$$\begin{aligned} B &= A \cup (A^c \cap B) \\ P(B) &= P(A) + P(A^c \cap B) \end{aligned}$$

- La probabilità dell'**evento unione** $A \cup B$ è $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$\begin{aligned} A \cup B &= A \cup (A^c \cap B) \\ P(A \cup B) &= P(A \cup (A^c \cap B)) \\ P(A \cup B) &= P(A) + P(A^c \cap B)) \\ P(A \cup B) &= P(A) + P(B) - P(A \cap B)) \end{aligned}$$

  Chiarimento: l'ultimo passaggio è dovuto a:

$$\begin{aligned} B &= (A \cap B) \cup (A^c \cap B) \\ P(B) &= P((A \cap B) \cup (A^c \cap B)) \\ P(B) &= P(A \cap B) + P(A^c \cap B) \\ P(B) - P(A \cap B) &= P(A^c \cap B) \end{aligned}$$

- In via generica, si ha che la probabilità dell'**unione di n eventi** corrisponde a (dimostrazione ricavabile tramite la precedente e le proprietà delle cardinalità mostrate nella sezione 1.2)

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}\left((-1)^{k+1} \cdot \sum_{1\leq i_1<i_2<\ldots<i_k\leq n} P(A_{i_1} \cdot A_{i_2} \cdot \ldots \cdot A_{i_k})\right)$$

3.1 Probabilità condizionata

Fino ad ora, abbiamo visto solo casi in cui la probabilità di un evento non è influenzata dalla presenza o meno di un altro evento appartenente allo stesso insieme ambiente.

Definiamo quindi la **probabilità di A dato B** come il **rapporto** tra la cardinalità dell'evento in cui l'evento A si verifichi in contemporanea all'evento B (dunque $|A \cap B|$) e la cardinalità totale dell'evento B:

$$P(A \mid B) = \frac{|A \cap B|}{|B|}$$

Possiamo, inoltre, riscrivere tale probabilità come:

$$P(A \mid B) = \frac{|A \cap B|}{|B|} \cdot \frac{|S|}{|S|} = \frac{|A \cap B|}{|S|} \cdot \frac{|S|}{|B|} = \frac{P(A \cap B)}{P(B)}$$

Definizione 3.1: Probabilità di A dato B

Dati due eventi $A, B \subseteq S$, la probabilità dell'evento A dato l'evento B corrisponde a:

$$P(A \mid B) = \frac{|A \cap B|}{|B|} = \frac{P(A \cap B)}{P(B)}, \text{ con } P(B) \neq 0$$

Esempi:

- Calcolare la probabilità che in una mano di poker da 5 carte utilizzando un mazzo standard da 52 carte escano 2 re sapendo che essa contiene un asso.

$$P(1A) = \frac{\binom{4}{1} \cdot \binom{48}{4}}{\binom{52}{5}}$$

$$P(2Re \cap 1A) = \frac{\binom{4}{1} \cdot \binom{4}{2} \cdot \binom{44}{2}}{\binom{52}{5}}$$

$$P(2Re \cap 1A \mid 1A) = \frac{P(2Re \cap 1A)}{P(1A)} = \frac{\binom{4}{2} \cdot \binom{4}{2} \cdot \binom{44}{2}}{\binom{52}{5}} \cdot \frac{\binom{52}{5}}{\binom{4}{1} \cdot \binom{48}{4}} = \frac{\binom{4}{2} \cdot \binom{44}{2}}{\binom{48}{4}} = \frac{473}{16215}$$

- Vengono lanciati due dadi equilibrati a sei facce. Calcolare la probabilità che la somma dei risultati sia uguale a 6 sapendo che uno dei due dadi abbia come risultato 2.

$$P(\texttt{almeno un 2} = 1 - P(\texttt{nessun 2}) = 1 - \frac{5^2}{6^2} = \frac{11}{36}$$

$$P(\texttt{somma 6} \ \cap \ \texttt{almeno un 2}) = \frac{2}{36}$$

$$P(\texttt{somma 6} \mid \texttt{almeno un 2}) = \frac{P(\texttt{somma 6} \ \cap \ \texttt{almeno un 2}}{P(\texttt{almeno un 2}} = \frac{2 \cdot 36}{36 \cdot 11} = \frac{2}{11}$$

3.2 Probabilità indipendenti

Consideriamo il lancio di due monete. Vogliamo sapere qual è la probabilità che al secondo lancio esca testa sapendo che al primo lancio è uscita testa.

$$P(\texttt{2° T | 1° T}) = \frac{P(\texttt{2° T} \cap \texttt{1° T})}{P(\texttt{1° T})} = \frac{1 \cdot 2}{4 \cdot 1} = \frac{1}{2}$$

Proviamo ora invece a calcolare direttamente la probabilità che al secondo lancio esca testa, senza tener conto del primo risultato:

$$P(\texttt{2° T}) = \frac{2}{4} = \frac{1}{2}$$

Notiamo quindi come la probabilità che al secondo lancio esca testa risulti essere **indipendente dal risultato del primo lancio**. A livello logico, ciò risulta intuitivo, poiché il lancio di una moneta equilibrata non ha alcun modo di poter influenzare un lancio successivo.

Affermiamo quindi che se la **probabilità di un evento A dato un evento B equivale alla probabilità dell'evento A stesso**, allora i due eventi sono **indipendenti tra loro**:

$$P(A \mid B) = P(A) \implies \texttt{A e B sono indipendenti tra loro}$$

Tuttavia, tale definizione di indipendenza presenta alcune **problematiche**, come ad esempio la stretta necessità che la **probabilità dell'evento B debba essere diversa da zero**, poiché:

$$P(A \mid B) = P(A) \implies \frac{P(A \cap B)}{P(B)} = P(A)$$

Possiamo quindi **riscrivere** tale definizione nel seguente modo, eliminando il vincolo richiesto:

$$P(A \cap B) = P(A) \cdot P(B)$$

Definizione 3.2: Eventi indipendenti

Considerati due eventi A e B, diciamo che tali eventi sono **indipendenti** l'uno dall'altro se e solo se si verifica che:

$$P(A \cap B) = P(A) \cdot P(B)$$

Osservazioni:

- Tale definizione implica che l'**evento impossibile $\emptyset$ sia indipendente da qualsiasi evento**, poiché si ha sempre che $\emptyset \cap B = \emptyset$ e che $P(\emptyset) = 0$:

$$\begin{gathered} P(\emptyset \cap B) = P(\emptyset) \cdot P(B) \\ 0 = 0 \cdot P(B) \\ 0 = 0 \end{gathered}$$

- Analogamente, anche l'**insieme ambiente S è indipendente da qualsiasi evento**, poiché si ha sempre che $S \cap B = B$ e che $P(S) = 1$:

$$\begin{gathered} P(S \cap B) = P(S) \cdot P(B) \\ P(B) = 1 \cdot P(B) \\ P(B) = P(B) \end{gathered}$$

- Se gli eventi **A e B sono indipendenti** tra loro, allora anche $\mathbf{A^c}$ **e B sono indipendenti** tra loro.

  Dimostrazione:

  - Per ipotesi, diamo per assunto che A e B siano indipendenti tra loro, dunque che valga $P(A \cap B) = P(A) \cdot P(B)$

- Dimostriamo quindi che anche A^c e B sono indipendenti tra loro:

$$P(S \cap B) = P(S) \cdot P(B)$$
$$P([A \cup A^c] \cap B) = 1 \cdot P(B)$$
$$P([A \cap B] \cup [A^c \cap B]) = 1 \cdot P(B)$$
$$P(A \cap B) + P(A^c \cap B) - P([A \cap B] \cup [A^c \cap B]) = 1 \cdot P(B)$$
$$P(A \cap B) + P(A^c \cap B) - 0 = 1 \cdot P(B)$$

- Siccome A e B sono disgiunti, sappiamo che $P(A \cap B) = P(A) \cdot P(B)$, dunque abbiamo che:

$$P(A) \cdot P(B) + P(A^c \cap B) - 0 = 1 \cdot P(B)$$
$$P(A^c \cap B) = 1 \cdot P(B) - P(A) \cdot P(B)$$
$$P(A^c \cap B) = (1 - P(A)) \cdot P(B)$$
$$P(A^c \cap B) = P(A^c) \cdot P(B)$$

- Dunque, concludiamo che anche A^c e B sono indipendenti tra loro

Esempio:

Dato il lancio di due gettoni aventi come facce (+1) e (-1), consideriamo i seguenti tre eventi:

$$A : \{\texttt{1° gettone è (+1)}\} \implies P(A) = \frac{1}{2}$$
$$B : \{\texttt{2° gettone è (+1)}\} \implies P(B) = \frac{1}{2}$$
$$C : \{\texttt{il prodotto dei due risultati è +1}\} \implies P(C) = \frac{2}{4} = \frac{1}{2}$$

Affinché i tre eventi siano indipendenti tra di loro, è necessario che:

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$
$$\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$$
$$\frac{1}{4} = \frac{1}{8}$$

Dunque, concludiamo che i tre eventi non siano indipendenti tra di loro. Difatti, strutturando il problema utilizzando la logica proposizionale, abbiamo che il verificarsi degli eventi A e B implichi anche il verificarsi dell'evento C ($A \wedge B \implies C$), rendendo quindi l'evento C dipendente anche dagli eventi A e B.

3.3 Probabilità totale e Formula di Beyes

Immaginiamo di avere un'urna, contenente **2 palline bianche e 3 palline nere**, e due mazzi di carte, **uno a cui sono state levate tutte le carte con seme cuori** (dunque contenente un seme rosso e due semi neri) ed **uno a cui sono state levate tutte le picche** (dunque contenenti due semi rossi e un seme nero). Le condizioni delle estrazioni sono le seguenti:

- Se la pallina estratta è bianca allora peschiamo una carta dal mazzo senza cuori
- Se la pallina estratta è nera allora peschiamo una carta dal mazzo senza picche

Vogliamo sapere qual è la **probabilità totale di pescare una carta rossa**. Definiamo quindi i tre eventi:

$$B : \{\texttt{pesco pallina bianca}\} \implies P(B) = \frac{2}{5}$$
$$N : \{\texttt{pesco pallina nera}\} \implies P(N) = \frac{3}{5}$$
$$R : \{\texttt{pesco carta rossa}\} : (R \cap B) \cup (R \cap N)$$

Tuttavia, ci risulta difficile calcolare in modo diretto la probabilità di R. Cerchiamo quindi **un altro approccio**: sapendo che $R = S \cap R$ e che $S = B \cup N$, poiché $N = B^c$ (dunque sono mutualmente esclusivi), abbiamo che:

$$P(R) = P(S \cap R)$$
$$P(R) = P([B \cup N] \cap R)$$
$$P(R) = P([(R \cap N) \cup (R \cap N)])$$
$$P(R) = P(B \cap R) + P(N \cap R)$$

Tramite la **formula della probabilità condizionata** ricaviamo che:

$$P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)} \implies P(X \mid Y) \cdot P(Y) = P(X \cap Y)$$

Dunque, riscriviamo la **probabilità totale di R** come:

$$P(R) = P(R \mid B) \cdot P(B) + P(R \mid B) \cdot P(N)$$

Poiché estraendo una pallina bianca peschiamo dal mazzo senza cuori, la probabilità di pescare una carta rossa data l'estrazione di una pallina bianca corrisponde a:

$$P(R \mid B) = \frac{13}{39} = \frac{1}{3}$$

Analogamente, la probabilità di pescare una carta rossa data l'estrazione di una pallina nera corrisponde a:

$$P(R \mid N) = \frac{26}{39} = \frac{2}{3}$$

Dunque, concludiamo che la **probabilità totale di R** è:

$$P(R) = P(R \mid B) \cdot P(B) + P(R \mid B) \cdot P(N) = \frac{1}{3} \cdot \frac{2}{5} + \frac{2}{3} \cdot \frac{3}{5} = \frac{8}{15}$$

Definizione 3.3: Probabilità totale

Dato un insieme ambiente S e un evento $A \subseteq S$, dove l'insieme S è partizionabile in n **partizioni**, ossia si ha:

$$S = \bigcup_{i=1}^{n} B_i$$

tali che ogni partizione è disgiunta l'una dall'altra:

$$\forall i, j \in [1, n] \mid i \neq j \implies B_i \cap B_j = \emptyset$$

Allora possiamo scrivere la **probabilità totale di A** come:

$$P(A) = \sum_{i=1}^{n} [P(A \mid B_i) \cdot P(B_i)]$$

Dimostrazione:

$$P(A) = P(A \cap S) = P\left(A \cap \bigcup_{i=1}^{n} B_i\right) = P\left(\bigcup_{i=1}^{n} (A \cap B_i)\right) =$$

$$\sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} [P(A \mid B_i) \cdot P(B_i)]$$

Esempio:

- Considerando gli stessi dati del problema precedente, ci chiediamo quale sia la **probabilità** che, avendo pescato una **carta rossa**, la pallina pescata sia **bianca**.

$$P(B \mid R) = \frac{P(B \cap R)}{P(R)}$$

Analogamente all'esempio visto precedentemente, poniamo $P(B \cap R) = P(R \mid B) \cdot P(B)$:

$$P(B \mid R) = \frac{P(R \mid B) \cdot P(B)}{P(R)} = \frac{\frac{1}{3} \cdot \frac{2}{5}}{\frac{8}{15}} = \frac{1 \cdot 2 \cdot 15}{3 \cdot 5 \cdot 8} = \frac{1}{4}$$

Definizione 3.4: Formula di Bayes

Dati due eventi A e B, la probabilità di A su B equivale a:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Esempio:

Abbiamo 10 confezioni di lampadine, ciascuna contenente 12 lampadine. Delle 10 totali, 9 contengono 10 lampadine funzionanti e 2 rotte, mentre la scatola rimanente contiene 2 lampadine funzionanti e 10 rotte. Ci chiediamo quale sia la probabilità di pescare due lampadine guaste da una qualsiasi scatola.

Definiamo gli eventi:

$$G : \{\texttt{peschiamo due guaste}\}$$

$$B : \{\texttt{peschiamo dalla confezione buona}\} \implies P(B) = \frac{9}{10}$$

Calcoliamo quindi le due probabilità richieste:

$$P(G \mid B) = \frac{|G \cap B|}{|B|} = \frac{\binom{2}{2}}{\binom{12}{2}} = \frac{1}{\binom{12}{2}} = \frac{10! \cdot 2!}{12!} = \frac{2}{12 \cdot 11} = \frac{1}{66}$$

$$P(G \mid B^c) = \frac{|G \cap B^c|}{|B^c|} = \frac{\binom{10}{2}}{\binom{12}{2}} = \frac{10!}{8! \cdot 2!} \cdot \frac{10! \cdot 2!}{12!} = \frac{10 \cdot 9}{12 \cdot 11} = \frac{15}{22}$$

$$P(G) = P(G \mid B) \cdot P(B) + P(G \mid B^c) \cdot P(B^c) = \frac{1}{66} \cdot \frac{9}{10} + \cdot \frac{15}{22} \cdot \frac{1}{9} = \frac{9}{110}$$

Ci chiediamo ora invece quale sia la probabilità che, avendo pescato due lampadine guaste, la scatola da cui abbiamo pescato sia la scatola contenente le 10 lampadine guaste, ossia $P(B^c \mid G)$. Usando la formula di Bayes, scriviamo tale probabilità come:

$$P(B^c \mid G) = \frac{P(G \mid B^c) \cdot P(B^c)}{P(G)} = \frac{15 \cdot 1 \cdot 100}{22 \cdot 10 \cdot 9} = \frac{5}{6}$$

4

Variabili aleatorie

Nel capitolo precedente abbiamo stabilito una definizione assiomatica di probabilità, indicando le sue principali componenti. Utilizzando la definizione data, cerchiamo di dare una definizione di **esperimento casuale misurabile**:

Definizione 4.1: Spazio di probabilità

La tripla $(S, \mathcal{P}(S), P)$, dove:

- S è un **insieme ambiente**
- $\mathcal{P}(S)$ è l'**insieme delle parti** di S
- P è la **funzione di probabilità** $P : \mathcal{P}(S) \to [0, 1]$

viene detta **spazio di probabilità** (o esperimento casuale)

Capita spesso che, durante lo studio di un fenomeno aleatorio (ossia casuale), si sia più interessati ad una qualche **funzione esprimente i probabili esiti** che agli esiti stessi.

Ad esempio, nel lancio di due dadi si è spesso interessati più alla somma dei valori risultanti che ai valori stessi. Specificatamente, potremmo essere più interessati a sapere se la somma dei due valori dia come risultato 7 piuttosto che sapere quale coppia di valori tra $(6, 1), (5, 2), (4, 3), (3, 4), (2, 5), (1, 6)$ si sia verificata.

Tali funzioni a valori reali definite sull'insieme ambiente vengono chiamate **variabili aleatorie**, definite come:

$$X : S \to \mathbb{R}$$

Poiché il **valore di una variabile aleatoria è determinato dall'esito dell'esperimento**, definito come $\omega \in S$, e dove all'esito ω è associato un valore numerico che lo rappresenta (ad esempio 1 se su un lancio di in dado esce la faccia avente un pallino), possiamo assegnare le **probabilità ai possibili valori ottenuti** dalla variabile aleatoria stessa:

$$X(\omega) \in \mathbb{R} \text{ dove } \omega \in S$$

Definiamo quindi l'evento $X = x$, ossia l'evento in cui la variabile aleatoria X corrisponda al valore x. La probabilità di tale evento corrisponde alla somma di tutti gli eventi $\omega \in S$ tali che $X(\omega) = x$:

$$P(X = x) = \sum_{\omega | X(\omega) = x} P(\{\omega\})$$

Esempi:

- Vengono lanciate 3 monete. Definiamo la variabile aleatoria X come il numero di teste risultanti dai lanci.

  In tal caso, si ha che:

$$P(X = 0) = \sum_{\omega | X(\omega) = 0} P(\{\omega\}) = P(CCC) = \frac{1}{8}$$

$$P(X = 1) = \sum_{\omega | X(\omega) = 1} P(\{\omega\}) = P(TCC) + P(CTC) + P(CCT) = \frac{3}{8}$$

$$P(X = 2) = \sum_{\omega | X(\omega) = 2} P(\{\omega\}) = P(TTC) + P(TCT) + P(CTT) = \frac{3}{8}$$

$$P(X = 3) = \sum_{\omega | X(\omega) = 3} P(\{\omega\}) = P(TTT) = \frac{1}{8}$$

  Notiamo, inoltre, come la **somma delle probabilità di tutti i valori assumibili dalla variabile X corrispondono alla probabilità totale**:

$$\sum_{i=0}^{3} P(X = i) = 1$$

- Vengono estratte 3 palline senza reinserimento da un'urna contenente 20 palline numerate da 1 a 20. Definiamo la variabile aleatoria X descrivente il massimo valore estratto. Vogliamo sapere la probabilità che il valore massimo estratto sia maggiore o uguale a 17.

  Prima di calcolare $P(X \geq 17)$, cerchiamo la probabilità generica che X possa valere k:

$$P(X = k) = \frac{\binom{1}{1} \cdot \binom{k-1}{2}}{\binom{20}{3}}$$

  valida per $3 \leq k \leq 20$.

A questo punto, calcoliamo $P(X \geq 17)$ come la somma di tutti i casi validi:

$$P(X \geq 17) = \sum_{k=17}^{20} P(X = k) = \sum_{k=17}^{20} \frac{\binom{1}{1} \cdot \binom{k-1}{2}}{\binom{20}{3}} = \binom{20}{3}^{-1} \sum_{k=17}^{20} \binom{k-1}{2} \approx 0.508$$

In alternativa, avremmo potuto calcolare tale probabilità tramite la sua complementare:

$$P(X \geq 17) = 1 - P(X < 17) = 1 - \sum_{k=3}^{16} \frac{\binom{1}{1} \cdot \binom{k-1}{2}}{\binom{20}{3}}$$

- Abbiamo N palline, di cui K rosse. Estraendo n palline, vogliamo sapere la probabilità che ne escano k rosse.

  Definiamo quindi la variabile aleatoria X come il numero di palline rosse estratte e calcoliamo:

$$p(k) = P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

4.1 Densità discreta e Funzione di distribuzione

Definizione 4.2: Variabile aleatoria discreta

Una **variabile aleatoria** che possa assumere un'infinità al più numerabile di valori è detta **discreta**.

Per una variabile aleatoria discreta X, definiamo la sua **densità discreta** come:

$$p(x) = P(X = x)$$

dove quindi si ha che $p : \mathbb{R} \to [0, 1]$.

Definizione 4.3: Funzione di distribuzione

Data una variabile aleatoria X, definiamo come **funzione di distribuzione** la funzione $F : \mathbb{R} \to [0, 1]$ rappresentante, per ogni valore x, la probabilità che la variabile aleatoria X sia minore o uguale a x:

$$F(x) = P(X \leq x)$$

In termini di probabilità discreta, quindi abbiamo che:

$$F(x) = P(X \leq x) = \sum_{k \leq x} p(k)$$

Osservazione 4.1

Se la funzione F non è decrescente, allora si verifica che:

- $x_m = min\, X(\omega)$ dove $\omega \in S$
- $x_M = max\, X(\omega)$ dove $\omega \in S$
- $x < x_m \implies F(x) = 0$
- $x > x_M \implies F(x) = 1$

Osservazione 4.2

Inoltre, poiché $\{X = x\}$ e $\{X \leq x-1\}$ sono due eventi disgiunti tra loro per definizione stessa, si ha che:

$$P(X \leq x) = P(X = k) + P(X \leq x - 1)$$
$$F(x) = p(x) + F(x - 1)$$
$$p(x) = F(x) - F(x - 1)$$

4.2 Tipi principali di variabile aleatoria

Definizione 4.4: Variabile aleatoria costante

Una variabile aleatoria X viene detta **V.A. a distribuzione costante** se

$$P(X = k) = 1 \qquad\qquad P(X \neq k) = 0$$

Definizione 4.5: Variabile aleatoria di Bernoulli

Una variabile aleatoria $X : \{0, 1\}$ dove si verifica che

$$P(X = 1) = p \qquad\qquad P(X = 0) = 1 - p$$

viene detta **V.A. a distribuzione di Bernoulli**, indicata con la notazione $X \sim B(p)$.

Osservazione 4.3

Se una variabile aleatoria Y assume come valori a e b con probabilità rispettiva p e $1-p$, possiamo descrivere tale variabile come $Y = (a-b)X+b$, dove $X \sim B(p)$, avendo quindi che:

$$P(Y = a) = p \qquad\qquad P(Y = b) = 1 - p$$

Definizione 4.6: Variabile aleatoria binomiale

Una variabile aleatoria corrispondente a una serie di estrazioni da un'urna (senza ordine e con reinserimento), viene detta **V.A. a distribuzione binomiale**, indicata con la notazione $X \sim B(n, p)$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

dove n è il numero di estrazioni, k il numero di successi e p la probabilità che si verifichi un successo

Definizione 4.7: Variabile aleatoria ipergeometrica

Una variabile aleatoria corrispondente a una serie di estrazioni da un'urna (senza ordine e senza reinserimento), viene detta **V.A. a distribuzione ipergeometrica**, indicata con la notazione $X \sim H(n, m, r)$

$$P(X = k) = \frac{\binom{m}{k}\binom{n-m}{r-k}}{\binom{n}{r}}$$

dove n è il numero di elementi totali nell'urna, m il numero di elementi nell'urna del tipo interessato, r il numero di elementi estratti dall'urna e k è il numero di elementi estratti del tipo interessato

Esempi:

1. Lanciando 10 monete, vogliamo sapere la probabilità che escano 2 teste, ponendo $X =$ `numero di teste uscite`:

$$P(X = 2) = \binom{10}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^8$$

   In tal caso, si ha che $X \sim B(10, \frac{1}{2})$

2. Dobbiamo vendere 10 auto e sappiamo che la probabilità che un'auto sia guasta è $\frac{3}{10}$. Vogliamo sapere la probabilità che 2 auto siano rotte e la probabilità che almeno 6 auto funzionino.

   Poniamo quindi $Y =$ `numero di auto guaste`:

$$P(Y = 2) = \binom{10}{2} \left(\frac{3}{10}\right)^2 \left(\frac{7}{10}\right)^8$$

   Per calcolare la seconda probabilità richiesta, possiamo calcolare la probabilità che le auto rotte siano minori o uguali a 4:

$$P(Y \leq 4) = \sum_{k=0}^{4} P(Y = k) = \sum_{k=0}^{4} \binom{10}{k} \left(\frac{3}{10}\right)^k \left(\frac{7}{10}\right)^{10-k}$$

In tal caso, si ha che $X \sim B(10, \frac{3}{10})$

3. Vengono effettuate 5 estrazioni senza ripetizioni da un'urna contenente 7 palline nere e 8 palline bianche. Posto X il numero di palline nere estratte, vogliamo sapere la probabilità che vengano estratte 4 palline nere:

$$P(X = 4) = \frac{\binom{7}{4}\binom{8}{1}}{\binom{15}{5}} = \frac{40}{429}$$

In tal caso, si ha che $X \sim H(15, 7, 5)$

4.3 Valore atteso

Consideriamo un classico dado a sei facce. Il **valore medio dei possibili risultati** è 3.5, calcolato come:

$$\frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$$

Immaginiamo ora di avere un dado modificato, dove la faccia 4 è stata sostituita con un'altra faccia 3 e la faccia 5 è stata sostituita con un'altra faccia 2. In tal caso, il valore medio dei possibili risultati risulta essere:

$$\frac{1+2+3+3+2+6}{6} = \frac{17}{6} \approx 2.85$$

Notiamo come possiamo riscrivere tale somma anche sotto-forma di media pesata tramite la probabilità:

$$\frac{1+2+3+3+2+6}{6} = \frac{1+2\cdot 2+2\cdot 3+6}{6} = 1\cdot\frac{1}{6} + 2\cdot\frac{2}{6} + 3\cdot\frac{2}{6} + 4\cdot\frac{1}{6} + 6\cdot\frac{1}{6}$$

definiamo tale media pesata come **valore atteso** di una variabile aleatoria discreta.

Definizione 4.8: Valore atteso di una variabile aleatoria discreta

Dati $x_1, ..., x_n$ valori assumibili da una variabile aleatoria discreta X, definiamo come **valore atteso** (o speranza matematica) di X la somma dei prodotti tra ogni valore assumibile e la probabilità discreta di quel valore:

$$E[X] = x_1 \cdot P(X = x_1) + .. + x_n \cdot P(X = x_n) = \sum_{k=0}^{n} x_k \cdot P(X = k) = \sum_{k=0}^{n} x_k \cdot p(k)$$

Notiamo anche come la somma di tutti i prodotti tra tutti i possibili esiti $\omega \in S$ e la loro probabilità $P(\{\omega\})$ corrisponda anche al valore atteso della variabile aleatoria X:

$$\sum_{\omega\in S} X(\omega)\cdot P(\{\omega\}) = \sum_{x\in X}\left(\sum_{\omega \mid X(\omega)=x} X(\omega)\cdot P(\{\omega\})\right) = \sum_{x\in X}\left(\sum_{\omega \mid X(\omega)=x} x\cdot P(\{\omega\})\right) =$$

$$= \sum_{x \in X} x \left(\sum_{\omega | X(\omega) = x} P(\{\omega\}) \right) = \sum_{x \in X} x \cdot P(X = x) = E[X]$$

Esempi:

- Si lancia un dado con n facce e si definisce X la variabile aleatoria rappresentante il risultato del lancio. Poiché ogni faccia ha $\frac{1}{n}$ probabilità di uscire, il valore atteso sarà:
$$E[X] = \sum_{k=1}^{n} k \cdot p(k) = \sum_{k=1}^{n} k \cdot \frac{1}{n} = \frac{1}{n} \cdot \sum_{k=1}^{n} k = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$$

- Vi sono 120 persone divise su tre bus, il primo contenente 36 persone, il secondo 40 e il terzo 44. Scelta una persona a caso tra le 120, vogliamo sapere il valore atteso di X, ossia il numero di persone sull'autobus della persona scelta.
$$E[X] = 36 \cdot \frac{36}{120} + 40 \cdot \frac{40}{120} + 44 \cdot \frac{44}{120} = \frac{4832}{120} \approx 40.27$$

Proposizione 4.1: Valore atteso di una V.A. di Bernoulli

Se $X \sim B(p)$, allora:
$$E[X] = 1 \cdot p + 0 \cdot (1-p) = p$$

Proposizione 4.2: Valore atteso di una V.A. binomiale

Se $X \sim B(n, p)$, allora:
$$E[X] = np$$

Dimostrazione:

- Osserviamo che
$$E[X] = \sum_{k=0}^{n} k \cdot p(k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} = 0 + \sum_{k=1}^{n} k \cdot p(k) = \sum_{k=1}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

- Ricordando che $k\binom{n}{k} = n\binom{n-1}{k-1}$, si ha:
$$E[X] = \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} = np \cdot \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

- Poniamo quindi $m = n-1$ e $j = k-1$:
$$E[X] = np \cdot \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1-j)} =$$

$$= np \cdot \sum_{j=0}^{m} \binom{m}{j} p^j (1-p)^{(m-j)} = np(p+1-p)^m = np$$

Proposizione 4.3

Se $X \sim H(n, m, r)$, allora:

$$E[X] = \frac{mr}{n}$$

Dimostrazione:

- Osserviamo che

$$E[X] = \sum_{k=0}^{n} k \cdot p(k) = \sum_{k=0}^{n} k \cdot \frac{\binom{m}{k}\binom{n-m}{r-k}}{\binom{n}{r}} = 0 + \sum_{k=1}^{n} k \cdot \frac{\binom{m}{k}\binom{n-m}{r-k}}{\binom{n}{r}}$$

- Ricordando che $k\binom{m}{k} = m\binom{m-1}{k-1}$ e che $r\binom{n}{r} = n\binom{n-1}{r-1}$, si ha:

$$E[X] = \sum_{k=1}^{n} k \cdot \frac{\binom{m}{k}\binom{n-m}{r-k}}{\binom{n}{r}} = \sum_{k=1}^{n} k \cdot \frac{\binom{m}{k}\binom{n-m}{r-k} r}{r\binom{n}{r}} =$$

$$= \sum_{k=1}^{n} m \cdot \frac{\binom{m-1}{k-1}\binom{n-m}{r-k} r}{n\binom{n-1}{r-1}} = \frac{mr}{n} \sum_{k=1}^{n} \frac{\binom{m-1}{k-1}\binom{n-m}{r-k}}{\binom{n-1}{r-1}}$$

- Ponendo $j = k-1$, $x = m-1$, $y = r-1$ e $z = n-1$, si ha che

$$E[X] = \frac{mr}{n} \sum_{k=1}^{n} \frac{\binom{m-1}{k-1}\binom{n-m}{r-k}}{\binom{n-1}{r-1}} = \frac{mr}{n} \sum_{j=0}^{z} \frac{\binom{x}{j}\binom{z-x}{y-j}}{\binom{z}{y}}$$

- Notiamo come i valori interni alla sommatoria corrispondano a $P(Y = j)$, dove $Y \sim H(z, x, y)$, dunque si ha che:

$$E[X] = \frac{mr}{n} \sum_{j=0}^{z} \frac{\binom{x}{j}\binom{z-x}{y-j}}{\binom{z}{y}} = \frac{mr}{n} \sum_{j=0}^{z} P(Y = j) = \frac{mr}{n}$$

Proposizione 4.4: Valore atteso condizionato da un evento

Sia X una variabile aleatoria e sia A un evento. In tal caso, si verifica che:

$$E[X \mid A] = \sum_{x \in X} x \cdot P(X \mid A)$$

$$E[X] = E[X \mid A] \cdot P(A) + E[X \mid A^c] \cdot P(A^c)$$

Esempio:

- Lanciando un dado equilibrato a 6 facce, definiamo come X il punteggio del dado e come A l'evento in cui esce una faccia dispari. In tal caso, si ha che:

$$P(X = x \mid A) = \frac{P(X = x \cap A)}{P(A)} = \frac{P(X = x \cap A)}{\frac{1}{2}} = \begin{cases} \frac{1}{6} \cdot \frac{2}{1} = \frac{1}{3} & \text{se } x \text{ è dispari} \\ 0 & \text{se } x \text{ è pari} \end{cases}$$

$$E[X \mid A] = \sum_{k=1}^{6} k \cdot P_{X|A}(X = k \mid A) = 1 \cdot \frac{1}{3} + 2 \cdot 0 + 3 \cdot \frac{1}{3} + 4 \cdot 0 + 5 \cdot \frac{1}{3} + 6 \cdot 0 = \frac{9}{3} = 3$$

4.3.1 Definizione di Gioco Equo

Un **gioco** viene considerato come **equo** se il **prezzo per giocare** al gioco stesso coincide col **valore atteso del premio**.

- Viene lanciato un dado e, se il risultato del lancio è 2, allora il giocatore vince 30€. Qual è il prezzo giusto da chiedere al giocatore per effettuare un lancio?

  Il problema può essere modellato tramite una variabile aleatoria X rappresentante la vincita del giocatore. Di conseguenza, tale variabile può assumere solo 2 valori, ossia 30€ o 0€. Il valore atteso del premio sarà quindi:

$$E[X] = 30 \cdot \frac{1}{6} + 0 \cdot \frac{5}{6} = 5$$

  dunque il prezzo giusto da chiedere affinché si tratti di un gioco equo è 5€.

- Chiediamo ad uno sconosciuto di partecipare al seguente gioco: tirando un dado a 6 facce, se la faccia risultante è pari allora il giocatore vince una quantità di euro equivalente alla faccia risultante, altrimenti, se esce una faccia dispari, il giocatore sarà costretto a pagare una quantità di euro equivalente alla faccia risultante. Vogliamo inoltre chiedere al giocatore interessato di pagare una piccola somma per poter effettuare un lancio, chiedendoci quale sia la somma giusta da chiedere.

  Calcoliamo quindi il valore atteso della vincita affinché il gioco sia equo:

$$E[X] = (-1) \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + (-3) \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + (-5) \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

  La somma da chiedere al giocatore per effettuare un lancio, quindi, sarà 0.5€.

4.3.2 Linearità del valore atteso

Il valore atteso di una variabile aleatoria risulta essere una **funzione lineare**, ossia avente le seguenti proprietà:

- Valore atteso con **costante moltiplicativa**

$$E[aX] = \sum_{\omega \in S} (aX)(\omega) \cdot P(\{\omega\}) = a \sum_{\omega \in S} X(\omega) \cdot P(\{\omega\}) = a \cdot E[X]$$

- Valore atteso con **costante additiva**

$$E[X + b] = \sum_{\omega \in S} (X + b)(\omega) \cdot P(\{\omega\}) = \sum_{\omega \in S} [X(\omega) \cdot P(\{\omega\}) + b \cdot P(\{\omega\})] =$$

$$= \sum_{\omega \in S} X(\omega) \cdot P(\{\omega\}) + \sum_{\omega \in S} b \cdot P(\{\omega\}) = E[X] + b$$

- Valore atteso con **costante additiva e moltiplicativa**

$$E[aX + b] = E[aX] + b = a \cdot E[X] + b$$

- Valore atteso di una **funzione di variabile aleatoria**:

$$E[f(X)] = \sum_{y \in f(X)} y \cdot P(f(X) = y) = \sum_{y \in f(X)} y \cdot \left(\sum_{x \mid f(X) = y} P(X = x) \right) =$$

$$= \sum_{y} \left(\sum_{x \mid f(X) = y} y \cdot P(X = x) \right) = \sum_{y} \left(\sum_{x \mid f(X) = y} f(x) \cdot P(X = x) \right) =$$

$$= \sum_{x \in X} f(x) \cdot P(X = x) = \sum_{x \in X} f(x) \cdot p(x)$$

Proposizione 4.5: Linearità del valore atteso

Data una variabile aleatoria X, il valore atteso di tale variabile assume le seguenti **proprietà lineari**:

- $E[aX] = a \cdot E[X]$
- $E[X + b] = E[X] + b$
- $E[aX + b] = a \cdot E[X] + b$
- $E[f(X)] = \sum_{x \in X} f(x) \cdot p(x)$

Esempi:

- Siano $-1, 0, 1$ i valori assumibili da una variabile aleatoria X, dove $P(X = -1) = P(X = 0) = P(X = +1) = \frac{1}{3}$. Il valore atteso di X^2 sarà:

$$E[X^2] = \sum_{k=-1}^{1} k^2 p(k) = (-1)^2 \cdot \frac{1}{3} + (0)^2 \cdot \frac{1}{3} + (+1)^2 \cdot \frac{1}{3} = \frac{2}{3}$$

  Notiamo inoltre come, considerando direttamente i valori assumibili da X^2, ossia 1 e 0, si ha che $P(X^2 = 0) = \frac{1}{2}$ e $P(X^2 = 1) = \frac{2}{3}$, implicando quindi che:

$$E[X^2] = \sum_{y=0}^{1} y \cdot P(X^2 = y) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$

- Lanciando tre monete, vogliamo sapere il valore atteso del quadrato del numero di teste uscite. Notiamo come:

	TTT	TTC	TCC	TCT	CTC	CTT	CCT	CCC
X	3	2	1	2	1	2	1	0
X^2	9	4	1	4	1	4	1	0

  dunque si ha che:

$$E[X^2] = 9 \cdot \frac{1}{8} + 4 \cdot \frac{3}{8} + 1 \cdot \frac{3}{8} + 0 \cdot \frac{1}{8} = \frac{24}{8} = 3$$

4.4 Variabili aleatorie congiunte e condizionate

Siano X e Y due variabili aleatorie relative ad un esperimento. La probabilità che entrambe le variabili aleatorie assumano un determinato valore contemporaneamente viene detta **distribuzione congiunta di X ed Y**:

$$P_{XY}(x, y) = P(X = x, Y = y) = P(\{X = x\} \cap \{Y = y\})$$

Tramite una distribuzione congiunta possiamo risalire alle **probabilità marginali** di X ed Y, ossia:

$$P_X(x) = P(X = x) = \sum_y P_{XY}(x, y)$$

$$P_Y(y) = P(Y = y) = \sum_x P_{XY}(x, y)$$

Esempi:

1. Consideriamo il lancio di tre monete, dove viene assegnato un punto per ogni testa risultante. Sia X la somma dei punteggi ottenuti e sia Y il prodotto del primo e del terzo risultato.

   - Si ha quindi che:

	000	001	010	011	100	101	110	111
X	0	1	1	2	1	2	2	3
Y	0	0	0	0	0	1	0	1

   - Stiliamo quindi la seguente tabella della distribuzione congiunta delle variabili aleatorie X ed Y, a seconda dei valori che esse possano assumere:

$Y \backslash X$	0	1	2	3
0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{2}{8}$	0
1	0	0	$\frac{1}{8}$	$\frac{1}{8}$

   ad esempio, quindi si ha che:

$$P_{XY}(3,1) = \frac{1}{8}$$

   - Tramite tale tabella, inoltre, possiamo facilmente calcolare le probabilità marginali, ad esempio:

$$P_Y(1) = \sum_{x=0}^{3} P_{XY}(x,1) = 0 + 0 + \frac{1}{8} + \frac{1}{8} = \frac{2}{8} = \frac{1}{4}$$

2. Siano $X : \{1,2,3,4\}$ e $Y : \{1,2,3,4\}$ due variabili aleatorie dove si ha che

$$P_{XY}(i,j) = \frac{c}{i+j}$$

   Quanto deve valore la **costante di normalizzazione** c affinché si tratti di una distribuzione congiunta valida?

   - Per definizione stessa di probabilità, sappiamo che

$$\sum_{i=1}^{4}\sum_{j=1}^{4} P_{XY}(i,j) = 1 \implies \sum_{i=1}^{4}\sum_{j=1}^{4} \frac{c}{i+j} = 1 \implies$$

$$\implies c \cdot \sum_{i=1}^{4}\sum_{j=1}^{4} \frac{1}{i+j} = 1 \implies c\left(\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \frac{4}{5} + \frac{3}{6} + \frac{2}{7} + \frac{1}{8}\right) = 1 \implies$$

$$\implies c \cdot \frac{3087}{840} = 1 \implies c = \frac{840}{3087}$$

Come per le normali variabili aleatorie, definiamo la funzione di distribuzione come

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

In questo caso, tale funzione di distribuzione assume delle particolari proprietà, ad esempio:

$$\lim_{x \to -\infty} F_{XY}(x, y) = 0 \qquad \lim_{y \to -\infty} F_{XY}(x, y) = 0$$
$$\lim_{x \to +\infty} F_{XY}(x, y) = F_Y(y) \qquad \lim_{y \to +\infty} F_{XY}(x, y) = F_X(x)$$

Analogamente, definiamo due variabili aleatorie congiunte come **indipendenti** se

$$P_{XY}(x, y) = P_X(x) \cdot P_Y(y)$$

Inoltre, nel caso in cui le variabili siano indipendenti, si verifica che:

$$F_{XY}(x, y) = P_{XY}(X \leq x, Y \leq y) = \sum_{i \leq x} \sum_{j \leq y} P_{XY}(i, j) = \sum_{i \leq x} \sum_{j \leq y} P_X(i) \cdot P_Y(j) =$$
$$= \sum_{i \leq x} P_X(i) \cdot \sum_{j \leq y} P_Y(j) = P_X(X \leq i) \cdot P_Y(Y \leq j) = F_X(x) \cdot F_Y(y)$$

La **densità condizionata** di una variabile aleatoria X in base all'evento $Y = y$ corrisponde a:

$$P_{X|Y=y}(x \mid y) = P(X = x \mid Y = y) = \frac{P_{XY}(x, y)}{P_Y(y)}$$

Di conseguenza, si ha che:

$$P_{XY}(x, y) = P_{X|Y=y}P(X \mid Y) \cdot P_Y(y)$$

Esempio:

- Dato il lancio di dado a 4 facce equilibrato, definiamo come X il punteggio del dado, dunque $P_X(x) = \frac{1}{4}, \forall x \in [1, 4]$.

  Successivamente, vengono lanciate X monete, dove Y è il numero di teste risultanti. Vogliamo sapere la probabilità di $Y = 2$.

  Dato che:

  $$P_Y(y) = \sum_{i=1}^{4} P_{XY}(i, y) = \sum_{i=1}^{4} P_{Y|X=i}(y \mid i) \cdot P_X(i)$$

  E poiché $Y \mid X = x \sim B(x, \frac{1}{2})$, si ha che:

  $$P_Y(2) = \sum_{i=1}^{4} P_{Y|X=i}(2 \mid i) \cdot P_X(i) = \sum_{i=1}^{4} P_{Y|X=i}(2 \mid i) \cdot \frac{1}{4} = \frac{1}{4} \sum_{i=1}^{4} P_{Y|X=i}(2 \mid i) =$$

$$= \frac{1}{4}\left(P_{Y|X=1}(2 \mid 1) + \sum_{i=2}^{4} P_{Y|X}(2 \mid i)\right) = \frac{1}{4}\left(0 + \sum_{i=2}^{4} P_{Y|X=i}(2 \mid i)\right) =$$

$$= \frac{1}{4}\sum_{i=2}^{4}\binom{i}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{i-2} = \frac{1}{4}\sum_{i=2}^{4}\binom{i}{2}\left(\frac{1}{2}\right)^{i} =$$

$$= \frac{1}{4}\left(1\cdot\frac{1}{4} + 3\cdot\frac{1}{8} + 6\cdot\frac{1}{16}\right) = \frac{1}{4}\cdot\frac{4+6+6}{16} = \frac{1}{4}$$

4.5 Valore atteso di somma e prodotto tra V.A.

Date due variabili aleatorie $X : \{a_1, \dots, a_n\}$ e $Y : \{b_1, \dots, b_m\}$, si ha che:

$$X + Y : \{a_1 + b_1, a_1 + b_2, \dots, a_n + b_{m-1}, a_n + b_m\}$$

Di conseguenza, si verifica che il valore atteso di $X + Y$ corrisponde a:

$$E[X+Y] = \sum_{\omega \in S}(X+Y)(\{\omega\})P(\{\omega\}) = \sum_{\omega \in S}[X(\{\omega\}) + Y(\{\omega\})]P(\{\omega\}) =$$

$$= \sum_{\omega \in S} X(\{\omega\})P(\{\omega\}) + \sum_{\omega \in S} Y(\{\omega\})P(\{\omega\}) = E[X] + E[Y]$$

Tuttavia, nel caso del valore atteso di $X \cdot Y$, si verifica che:

$$E[XY] = \sum_{\omega \in S}(XY)(\{\omega\})P(\{\omega\}) = E[X]E[Y] \iff P_{XY}(x,y) = P_X(x)P_Y(y)$$

ossia se e solo se le due variabili sono **indipendenti tra loro** *(dimostrazione omessa)*

Esempi:

1. Vengono lanciati due dadi equilibrati a 4 facce. Definiamo come X il risultato del primo dado e come Y il risultato del secondo dado.

   - Il valore atteso di $X + Y$ sarà:

   $$E[X+]) = E[X]+E[Y] = \sum_{x=1}^{4}\frac{x}{4} + \sum_{y=1}^{4}\frac{y}{4} = \frac{1}{4}\left(\sum_{x=1}^{4} x + \sum_{y=1}^{4} y\right) = \frac{2}{4}\sum_{k=1}^{4} k = \frac{4\cdot 5}{4} = 5$$

   - Siccome X e Y sono indipendenti l'una dall'altra, il valore atteso di XY sarà:

   $$E[XY] = E[X]E[Y] = \left(\sum_{x=1}^{4}\frac{x}{4}\right)\left(\sum_{y=1}^{4}\frac{y}{4}\right) = \frac{5}{2}\cdot\frac{5}{2} = \frac{25}{4}$$

2. Sia X una variabile aleatoria che assume il valore 1 con probabilità p e il valore 0 con probabilità $1-p$ e sia $Y = 1 - X$.

   - Considerando $X+Y$ e XY, la tavola delle probabilità corrisponde a:

X	Y	$P(X \cap Y)$	$X+Y$	XY
0	0	0	0	0
0	1	1-p	1	0
1	0	p	1	0
1	1	0	2	1

   - Di conseguenza, si verifica che:

$$E[X+Y] = E[X] + E[Y] = p + 1 - p = 1$$

$$E[XY] = 0 \cdot 0 + 0 \cdot (1-p) + 0 \cdot p + 1 \cdot 0 = 0 \cdot (0 + 1 - p + p) + 1 \cdot 0 = 0$$

   - Difatti, essendo X e Y dipendenti, si ha che $E[XY] \neq E[X]E[Y]$

Lemma 4.1

Se X e Y sono due variabili aleatorie indipendenti, allora

$$E[f(X)g(Y)] = E[f(X)] \cdot E[g(Y)]$$

4.6 Valore atteso condizionato

Data una variabile aleatoria X, il valore atteso di X dato il verificarsi di un evento A corrisponde a:

$$E[X \mid A] = \sum_{x \in X} x \cdot P_{X|A}(X = x \mid A) = \sum_{x \in X} x \cdot \frac{P(X = x \cap A)}{P(A)} = \frac{E[X \cap A]}{P(A)}$$

Teorema 4.1

Date due variabili aleatorie X ed Y, esiste un'unica variabile aleatoria $Z = f(X)$, dunque dove $Z(\{\omega\}) = f(X(\{\omega\}))$, tale che:

$$f(x) = E[Z \mid X = x] = E[Y \mid X = x]$$

In particolare, si ha che $Z = f(X) = E[Y \mid X]$

Dimostrazione:

- Dato $f(x) := E[Y \mid X = x]$, si ha che:

$$f(x) = E[Y \mid X = x] = \frac{E[Y \cap X = x]}{P_X(x)}$$

- Poiché $Z = F(X)$, verifichiamo che:

$$E[Z \mid X = x] = \frac{E[Z \cap X = x]}{P_X(x)} = \frac{E[f(X) \cap X = x]}{P_X(x)} = $$
$$= \frac{E[f(x) \cap X = x]}{P_X(x)} = \frac{f(x) \cdot P_X(x)}{P_X(x)} = f(x)$$

- Quindi, si ha che:

$$Z = F(X) = E[Y \mid X] \iff E[Y \mid X = x] = f(x) \iff$$
$$\iff E[Y \mid X = x] = \sum_{y \in Y} y \cdot P_{Y|X}(y \mid x) = \sum_{y \in Y} \frac{y \cdot P_{YX}(y, x)}{P_X(x)}$$

$\square$

Corollario 4.1: D

ate due variabili aleatorie indipendenti X ed Y, si verifica che

$$E[Y \mid X = x] = E[Y]$$

Dimostrazione:

- Se X e Y sono indipendenti allora

$$E[Y \mid X = x] = \sum_{y \in Y} \frac{y \cdot P_{YX}(y, x)}{P_X(x)} = \sum_{y \in Y} y \frac{P_Y(y) P_X(x)}{P_X(x)} = \sum_{y \in Y} y P_Y(y) = E[Y]$$

$\square$

Corollario 4.2

Date due variabili aleatorie X ed $Y = g(X)$, si verifica che

$$E[Y \mid X = x] = g(X)$$

Dimostrazione:

$$E[Y \mid X = x] = \sum_{y \in Y} y \cdot P(Y = y \mid X = x) = \sum_{x' \in X} g(x') \cdot P(X = x' \mid X = x) = g(x)$$

□

Corollario 4.3

Date due variabili aleatorie X e Y, si verifica che:

$$E[E[Y \mid X]] = E[Y]$$

Dimostrazioni:

- Posto $f(X) = E[Y \mid X]$ dove $f(x) = E[Y \mid X = x]$, si ha che:

$$E[E[Y \mid X]] = E[f(X)] = \sum_{x \in X} f(x) \cdot P_X(x) = \sum_{x \in X} E[Y \mid X = x] \cdot P_X(x) =$$

$$= \sum_{x \in X} \sum_{y \in Y} y \cdot P(Y = y \mid X = x) \cdot P_X(x) = \sum_{x \in X} \sum_{y \in Y} y \cdot P_{YX}(y, x) =$$

$$= \sum_{y \in Y} y \sum_{x \in X} P_{YX}(y, x) = \sum_{y \in Y} y \cdot P_Y(y) = E[Y]$$

□

Esempi:

- Partecipiamo al seguente gioco: si hanno n punti, tra cui uno speciale. Definiamo come N la variabile aleatoria indicante che all'N-esima estrazione è stato trovato il punto speciale, dunque si ha che $P_N(k) = P(N = k) = \frac{1}{n}$.

  Definiamo la variabile aleatoria $T_k = 1 + X_k$, dove $X_k \sim B(\frac{1}{2})$ (dunque $P_{X_k}(0) = P_{X_k}(1) = \frac{1}{2}$), come il tempo impiegato a verificare che il k-esimo punto estratto sia quello speciale.

  - Il valore atteso di T_j sarà:

  $$E[T_j] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2}$$

  - Il tempo necessario a trovare il punto speciale all'N-esimo tentativo sarà:

  $$T = \sum_{k=1}^{N} T_k = E[T \mid N]$$

- Il tempo atteso, invece, sarà:

$$E[T] = E[E[T \mid N]] = \sum_{h=1}^{n} P_N(k) \cdot E[T \mid N = h] =$$

$$= \sum_{h=1}^{n} \frac{1}{n} \cdot E[T \mid N = h] = \frac{1}{n} \sum_{h=1}^{n} E\left[\sum_{j=1}^{N} T_j \mid N = h\right] =$$

Siccome $N = i$, allora

$$E[T] = \frac{1}{n} \sum_{h=1}^{n} E\left(\sum_{j=1}^{N} T_j \mid N = h\right) = \frac{1}{n} \sum_{h=1}^{n} E\left[\sum_{j=1}^{h} T_j \mid N = h\right] =$$

$$= \frac{1}{n} \sum_{h=1}^{n} \sum_{j=1}^{h} E\left[T_j \mid N = h\right]$$

Siccome T_j e N sono indipendenti, allora

$$E[T] = \frac{1}{n} \sum_{h=1}^{n} \sum_{j=1}^{h} E[T_j \mid N = h] = \frac{1}{n} \sum_{h=1}^{n} \sum_{j=1}^{N} E[T_j] = \frac{1}{n} \sum_{h=1}^{n} \sum_{j=1}^{h} \frac{3}{2} =$$

$$= \frac{1}{n} \sum_{h=1}^{n} \frac{3}{2} h = \frac{3}{2n} \sum_{h=1}^{n} h = \frac{3}{2n} \cdot \frac{n(n+1)}{2} = \frac{3}{4}(n+1)$$

- Supponiamo ora di aver già calcolato un tempo atteso. Vogliamo sapere il valore atteso di N dato $T = 2$.
  - Notiamo come:

$$E[N \mid T = 2] = \sum_{k=1}^{n} k \cdot P(N = k \mid T = 2) = \sum_{k=1}^{n} k \cdot \frac{P(T = 2 \mid N = k) \cdot P(N = k)}{P(T = 2)}$$

  - Siccome

$$P(T = 2) = \sum_{j=1}^{N} P(T = 2 \mid N = j) \cdot P(N = j) = \frac{1}{n} \sum_{j=1}^{N} P(T = 2 \mid N = j)$$

e siccome

$$P(T = 2 \mid N = k) = \begin{cases} \frac{1}{2} & k = 1 \\ \frac{1}{4} & k = 2 \\ 0 & k > 2 \end{cases}$$

allora ne segue che:

$$P(N = k \mid T = 2) = \frac{P(T = 2 \mid N = k) \cdot P(N = k)}{P(T = 2)} =$$

$$= \frac{P(T = 2 \mid N = k) \cdot P(N = k)}{\sum_{j=1}^{N} P(T = 2 \mid N = j) \cdot P(N = j)} = \frac{\frac{1}{n} \cdot P(T = 2 \mid N = k)}{\frac{1}{n} \sum_{j=1}^{n} P(T = 2 \mid N = j)} =$$

$$= \frac{P(T = 2 \mid N = k)}{\frac{1}{2} + \frac{1}{4} + 0} = \frac{4}{3} \cdot P(T = 2 \mid N = k) = \begin{cases} \frac{2}{3} & k = 1 \\ \frac{1}{3} & k = 2 \\ 0 & k > 2 \end{cases}$$

- Di conseguenza, si ha che:

$$E[N \mid T = 2] = \sum_{k=1}^{n} k \cdot P(N = k \mid T = 2) = \sum_{k=1}^{n} k = 1 \cdot \frac{2}{3} + 2 \cdot \frac{1}{3} + 0 = \frac{4}{3}$$

4.7 Varianza e Covarianza

Definizione 4.9: Varianza

Definiamo come **varianza** di una variabile aleatoria X l'**indice di dispersione** dei valori assunti da X attorno al suo valore medio $E[X]$), definita come:

$$Var[X] = E[(X - E[X])^2)$$

In poche parole, più i valori assunti da X sono concentrati attorno al valore medio, minore sarà la varianza.

Osservazione 4.4

Data una variabile aleatoria X, si verifica che:

$$Var[X] = E[X^2] - E[X]^2$$

Dimostrazione:

$$Var[X] = E[(X - E[X])^2) = E[X^2 - 2X \cdot E[X] + E[X]^2] = E[X^2] - E[2X \cdot E[X]] + E[X]^2] =$$

$$= E[X^2] - 2 \cdot E[X] \cdot E[X] + E[X]^2] = E[X^2] + E[X]^2 - 2E[X]^2 = E[X^2] - E[X]^2$$

□

Definizione 4.10: Covarianza

Definiamo come **covarianza** di due variabili aleatorie X e Y l'**indice** di quanto le due **varino assieme**, definito come:

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

Se si verifica che $Cov(X,Y) = 0$, allora tali variabili vengono detto **decorrelate**.

Osservazione 4.5

Data una variabile aleatoria X, si verifica che:

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

Dimostrazione:

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY - X \cdot E[Y] - Y \cdot E[X] + E[X]E[Y]] =$$
$$= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] = E[XY] - E[X]E[Y]$$

Osservazione 4.6

Se due variabili aleatorie sono indipendenti, allora esse sono anche decorrelate (non è detto il contrario)

$$P_{XY}(x,y) = P_X(x)P_Y(y) \implies Cov(X,Y) = 0$$

Dimostrazione:

- Supponendo che X e Y siano due variabili aleatorie indipendenti, si ha che

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$$

$\square$

Esempio:

- Data la variabile aleatoria $X : \{-1, 0, 1\}$, dove $P_X(1) = P_X(-1) = \frac{1}{5}$ e $P_X(0) = \frac{3}{5}$, definiamo una variabile aleatoria $Z : \{-1, 1\}$ indipendente da X, dove $P_Z(1) = P_Z(-1) = \frac{1}{2}$, e una variabile aleatoria $Y = XZ$.
  - Si verifica che:

$$E[X] = (-1) \cdot \frac{1}{5} + 0 \cdot \frac{3}{5} + 1 \cdot \frac{1}{5} = 0$$
$$E[Z] = (-1) \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$$
$$E[Y] = E[XZ] = E[X]E[Z] = 0 \cdot 0 = 0$$

– Inoltre, notiamo che:

$$E[XY] = E[X^2 Z] = E[X^2]E[Z] = E[X^2] \cdot 0 = 0$$

– Di conseguenza, ne segue che:

$$Cov(X, Y) = E[XY] - E[X]E[Y] = 0 - 0 = 0$$

dunque, X e Y sono decorrelate tra loro, tuttavia, esse non sono indipendenti, poiché ad esempio:

$$P_{XY}(0, 1) = 0 \neq \frac{3}{2} = P_X(0)P_Y(1)$$

Osservazione 4.7

Data una variabile aleatoria Xsi verifica che:

$$Var[\alpha X] = \alpha^2 Var[X]$$

Dimostrazione:

$$\begin{aligned} Var[\alpha X] &= E[(\alpha X)^2] - E[\alpha X]^2 = E[\alpha^2 X^2] - (\alpha E[X])^2 = \\ &= \alpha^2 E[X] - \alpha^2 E[X]^2 = \alpha^2 (E[X^2] - E[X]^2) = \alpha^2 Var[X] \end{aligned}$$

□

Teorema 4.2

Se $X_1, \ldots, X_n$ variabili aleatorie indipendenti, allora

$$Var[X_1 + \ldots + X_n] = \sum_{k=1}^{n} Var[X_k]$$

Dimostrazione:

- Dimostriamo il caso base utilizzando due variabili aleatorie:

$$Var[X + Y] = E[(X + Y)^2] - E[X + Y]^2 = E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 =$$

$$\begin{aligned} &= E[X^2] + 2E[X]E[Y] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2 = \\ &= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 = Var[X] + Var[Y] \end{aligned}$$

- Per ipotesi induttiva, affermiamo che

$$Var[X_1 + \ldots + X_n] = \sum_{k=1}^{n} Var[X_k]$$

- Effettuiamo quindi il passo induttivo

$$Var[X_1+\ldots+X_n+X_{n+1}] = Var[(X_1+\ldots+X_n)+X_{n+1}] = Var[X_1+\ldots+X_n]+Var[X_{n+1}] =$$

$$= \sum_{k=1}^{n} Var[X_k] + Var[X_{n+1}] = \sum_{k=1}^{n+1} Var[X_k]$$

□

Proposizione 4.6

Se $X \sim B(p)$, allora:

$$Var[X] = p(1-p)$$

Dimostrazione:

$$Var[X] = E[X^2] - E[X]^2 = E[X^2] - p^2 = 1^2 \cdot p + 0^2 - p^2 = p - p^2 = p(1-p)$$

□

Proposizione 4.7

Se $X \sim B(n,p)$, allora:

$$Var[X] = np(1-p)$$

Dimostrazione:

- Analogamente alla dimostrazione già vista nella sezione 4.3 per $E[X] = np$ dove $X \sim B(n,p)$, otteniamo che:

$$E[X^2] = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} = np \cdot \sum_{k=1}^{n} k \binom{n-1}{k-1} p^{k-1}(1-p)^{(n-1)-(k-1)}$$

- Ponendo $m = n-1$ e $j = k-1$, si ha che:

$$E[X^2] = np \cdot \sum_{k=1}^{n} k \binom{n-1}{k-1} p^{k-1}(1-p)^{(n-1)-(k-1)} = np \cdot \sum_{j=0}^{m} (j+1) \binom{m}{j} p^j (1-p)^{m-j} =$$

$$= np \left[\sum_{j=0}^{m} j \binom{m}{j} p^j (1-p)^{m-j} + \sum_{j=0}^{m} \binom{m}{j} p^j (1-p)^{m-j} \right]$$

- Notiamo come i termini interni alle due sommatorie corrispondano a $P(Y = j)$, dove $Y \sim B(m,p)$, dunque si ha che:

$$E[X^2] = np\left[\sum_{j=0}^{m} j\binom{m}{j}p^j(1-p)^{m-j} + \sum_{j=0}^{m}\binom{m}{j}p^j(1-p)^{m-j}\right] = np[E[Y]+1] =$$

$$= np(mp+1) = np((n-1)p+1) = np(np-p+1)$$

- Infine, concludiamo che:

$$Var[X] = E[X^2] - E[X]^2 = np(np-p+1) - (np)^2 = np(1-p)$$

Proposizione 4.8

Se $X \sim H(n, m, r)$, allora:

$$Var[X] = \frac{rm(n-r)(n-m)}{n^2(n-1)} = E[X]\frac{(n-r)(n-m)}{n(n-1)}$$

Dimostrazione:

- Analogamente alla dimostrazione già vista nella sezione 4.3 per $E[X] = \frac{mr}{n}$ dove $X \sim H(n, m, r)$, otteniamo che:

$$E[X^2] = \frac{mr}{n}\sum_{k=1}^{n} k\frac{\binom{m-1}{k-1}\binom{n-m}{r-k}}{\binom{n-1}{r-1}}$$

- Ponendo $j = k-1$, $x = m-1$, $y = r-1$ e $z = n-1$, si ha che:

$$E[X^2] = \frac{mr}{n}\sum_{j=0}^{z}(j+1)\frac{\binom{x}{j}\binom{z-x}{y-j}}{\binom{z}{y}} = \frac{mr}{n}\left[\sum_{j=0}^{z} j\frac{\binom{x}{j}\binom{z-x}{y-j}}{\binom{z}{y}} + \sum_{j=0}^{z}\frac{\binom{x}{j}\binom{z-x}{y-j}}{\binom{z}{y}}\right]$$

- Notiamo come i termini interni alle due sommatorie corrispondano a $P(Y = j)$, dove $Y \sim H(z, x, y)$, dunque si ha che:

$$E[X^2] = \frac{mr}{n}\left[\sum_{j=0}^{z} j\frac{\binom{x}{j}\binom{z-x}{y-j}}{\binom{z}{y}} + \sum_{j=0}^{z}\frac{\binom{x}{j}\binom{z-x}{y-j}}{\binom{z}{y}}\right] = \frac{mr}{n}[E[Y]+1] = \frac{mr}{n}\left[\frac{xy}{z}+1\right] =$$

$$= E[X]\left[\frac{(m-1)(r-1)}{n-1}+1\right]$$

- Infine, concludiamo che:

$$Var[X] = E[X^2] - E[X]^2 = E[X]\left(\frac{(m-1)(r-1)}{n-1}+1\right) - E[X]^2 =$$

$$= E[X]\left(\frac{(m-1)(r-1)}{n-1} + 1 - E[X]\right) =$$

$$= \frac{mr}{n}\left(\frac{(m-1)(r-1)}{n-1} + 1 - \frac{mr}{n}\right) =$$

$$\frac{rm(n-r)(n-m)}{n^2(n-1)} = E[X]\frac{(n-r)(n-m)}{n(n-1)}$$

□

4.8 Spazi di probabilità infiniti

Fino ad ora, abbiamo considerato solo casi in cui lo spazio di probabilità S assuma un numero finito di valori. Tuttavia, è possibile lavorare anche **spazi di probabilità di dimensione infinita**.

Consideriamo il seguente esempio:

- In una gara di pattinaggio si stima che la probabilità che un pattinatore riesca a completare un percorso è pari a $\frac{2}{5}$. Definiamo come X il numero di tentativi effettuati da un pattinatore per riuscire a completare il percorso.
- Notiamo subito come i valori assunti da X corrispondano a $X : \{1, 2, 3, \ldots\}$, poiché, potenzialmente, potrebbero volerci infiniti tentativi al pattinatore per poter completare il percorso.
- Proviamo quindi a calcolare la probabilità di $X = 10$, equivalente a calcolare la probabilità che si verifichino 9 esiti negativi e 1 esito positivo finale:

$$P(X = 10) = \left(\frac{3}{5}\right)^9 \frac{2}{5} = \frac{3^9 \cdot 2}{5^10}$$

Definizione 4.11: Variabile aleatoria geometrica

Una variabile aleatoria con probabilità p di verificarsi all'k-esimo tentativo dopo $k-1$ esiti negativi viene detta **V.A. a distribuzione geometrica**, indicata come $X \sim G(p)$:

$$P_X(k) = (1-p)^{k-1}p$$

Nell'esempio precedente, quindi, si ha che $X \sim G(\frac{2}{5})$

Proposizione 4.9

Se $X \sim G(p)$, allora

$$E[X] = \frac{1}{p}$$

Dimostrazione:

- Ricordando la serie notevole

$$\sum_{k=1}^{\infty} k\alpha^k = \frac{\alpha}{(1-\alpha)^2}$$

verifichiamo che:

$$E[X] = \sum_{k=1}^{\infty} k \cdot P_X(k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = \frac{p}{p-1}\sum_{x=1}^{\infty} k(1-p)^k =$$

$$= \frac{p}{p-1} \cdot \frac{p-1}{(1-(p-1))^2} = \frac{p}{p^2} = \frac{1}{p}$$

□

Proposizione 4.10

Se $X \sim G(p)$, allora

$$Var[X] = \frac{1-p}{p^2}$$

Dimostrazione:

- Ricordando la serie notevole

$$\sum_{k=1}^{\infty} k^2\alpha^k = -\frac{\alpha(\alpha+1)}{(\alpha-1)^3}$$

verifichiamo che:

$$E[X^2] = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = \frac{p}{p-1}\sum_{x=1}^{\infty} k^2(1-p)^k =$$

$$= \frac{p}{p-1} \cdot \frac{(p-1)(2-p)}{p^3} = \frac{2-p}{p^2}$$

- Dunque, concludiamo che:

$$Var[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

□

4.9 Variabili aleatorie continue

Definizione 4.12: Variabile aleatoria continua

Definiamo X come una **V.A. continua** se esiste una funzione $\rho : \mathbb{R} \to [0, +\infty)$ detta **densità di probabilità** tale che

$$P(X \in [a, b]) = \int_a^b \rho(t)\ dt$$

Osservazione 4.8

Se X è una V.A. continua, allora

$$P(X \in (-\infty, +\infty)) = \int_{-\infty}^{\infty} \rho(t)\ dt = 1 \qquad E[X] = \int_{-\infty}^{\infty} t \cdot \rho(t)\ dt$$

Definizione 4.13: Variabile aleatoria continua uniforme

Una variabile aleatoria continua X che attribuisce la stessa probabilità a tutti i valori assumibili in un intervallo $[a, b]$ viene detta **V.A. a distribuzione uniforme continua**, indicata come $X \sim U(a, b)$:

$$\rho(t) = \begin{cases} \frac{1}{b-a} & \text{se } t \in [a, b] \\ 0 & \text{se } t \notin [a, b] \end{cases}$$

Infatti, si ha che:

$$\int_{-\infty}^{\infty} \rho(t)\ dt = \int_a^b \rho(t)\ dt = \int_a^b \frac{1}{b-a}\ dt = \frac{1}{b-a} \int_a^b 1\ dt = \frac{b-a}{b-a} = 1$$

Proposizione 4.11

Se $X \sim U(a, b)$, allora:

$$F_X(x) = P(X \in (-\infty, x]) = P(X \in [a, x]) = \frac{x-a}{b-a}$$

Dimostrazione:

$$F_X(x) = P(X \in (-\infty, x]) = \int_{-\infty}^{x} \rho(x)\ dt = \int_a^x \rho(x)\ dt = \frac{1}{b-a} \int_a^x 1\ dt = \frac{x-a}{b-a}$$

$\square$

Proposizione 4.12

Se $X \sim U(a,b)$, allora:
$$E[X] = \frac{a+b}{2}$$

Dimostrazione:

$$E[X] = \int_{-\infty}^{\infty} x\rho(x)\,dx = \int_a^b x\frac{1}{b-a}\,dx = \frac{1}{b-a}\int_a^b x\,dx =$$
$$= \frac{1}{b-a}\cdot\frac{b^2-a^2}{2} = \frac{(b+a)(b-a)}{2(b-a)} = \frac{a+b}{2}$$

□

Esempio:

- Una società di servizi elettrici fornisce livelli di tensione uniformemente distribuiti, compresi tra $123.0V$ e $125.0V$.
  - Definiamo dunque la variabile aleatoria $X \sim U(123.0, 125.0)$, rappresentante la tensione fornita, dunque
  $$\rho(t) = \begin{cases} \frac{1}{1.5} = \frac{2}{3} & \text{se } t \in [123.0, 125.0] \\ 0 & \text{se } t \notin [123.0, 125.0] \end{cases}$$
  - La probabilità che l'azienda invii una tensione inferiore o uguale a $123.5V$ corrisponde a:
  $$P(x \in (-\infty, 123.5]) = F_X(123.5) = \frac{123.5-123.0}{125.0-123.0} = \frac{0.5}{1.5} = \frac{1}{3}$$
  - La probabilità che l'azienda invii una tensione maggiore di $124.0V$ corrisponde a:
  $$P(x \in (124.V, +\infty)) = 1 - F_X(124.0) = 1 - \frac{124.0-123.0}{125.0-123.0} = 1 - \frac{1}{2} = \frac{1}{2}$$

4.10 Variabili aleatorie di Poisson

Sia $X \sim B(n,p)$, dove n è un numero molto grande e p un numero molto piccolo, dunque

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- Posto $p = \frac{\lambda}{n}$, si ha che:

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{(n-k)! \cdot k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{n! \cdot \lambda^k}{(n-k)! \cdot k! \cdot n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

- Siccome

$$\lim_{n\to+\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

$$\lim_{n\to+\infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = (1-0)^{-k} = 1$$

$$\lim_{n\to+\infty} \frac{n!}{(n-k)! \cdot n^k} = \lim_{n\to+\infty} \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n \cdot n \cdot n \ldots \cdot n} = 1$$

- Allora ne segue che

$$\lim_{n\to+\infty} P_X(k) = \lim_{n\to+\infty} \frac{n! \cdot \lambda^k}{(n-k)! \cdot k! \cdot n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

Definizione 4.14: Variabile aleatoria di Poisson

Una variabile aleatoria X che esprime le probabilità per il numero di eventi che si verificano successivamente ed indipendentemente in un dato intervallo di tempo, avente densità

$$P_X(k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad \text{per } k = 0, 1, \ldots$$

viene detta **V.A. a distribuzione di Poisson** (o legge degli eventi rari) indicata come $X \sim P(\lambda)$

Proposizione 4.13

Sia $X \sim P(\lambda)$, allora
$$E[X] = \lambda$$

Dimostrazione:

- Osserviamo che:
$$E[X] = \sum_{k=0}^{\infty} kP_X(k) = \sum_{k=0}^{\infty} ke^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\lambda\sum_{k=0}^{\infty} k\frac{\lambda^{k-1}}{k!} =$$
$$= e^{-\lambda}\lambda\sum_{k=1}^{\infty} k\frac{\lambda^{k-1}}{k!} = e^{-\lambda}\lambda\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda}\lambda\sum_{j=0}^{\infty} \frac{\lambda^j}{j!}$$
- Ricordando che:
$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$
Ne concludiamo che:
$$E[X] = e^{-\lambda}\lambda\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = e^{-\lambda}\lambda e^{\lambda} = \lambda$$

Proposizione 4.14

Sia $X \sim P(\lambda)$, allora:
$$Var[X] = \lambda$$

Dimostrazione:

- Sappiamo già che se $X \sim P(\lambda)$ allora $E[X] = \lambda$, dunque
$$Var[X] = E[X^2] - E[X]^2 = E[X^2] - \lambda^2$$
- Osserviamo ora che:
$$E[X^2] = \sum_{k=0}^{\infty} k^2e^{-\lambda}\frac{\lambda^k}{k!} = \lambda\sum_{k=0}^{\infty} k^2e^{-\lambda}\frac{\lambda^{k-1}}{k!} = \lambda\sum_{k=1}^{\infty} ke^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!} =$$
$$= \lambda\sum_{j=0}^{\infty}(j+1)e^{-\lambda}\frac{\lambda^j}{j!} = \lambda\left(\sum_{j=0}^{\infty} je^{-\lambda}\frac{\lambda^j}{j!} + \sum_{j=0}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!}\right) =$$
$$= \lambda\left(E[X] + 1\right) = \lambda^2 + \lambda$$

- Dunque, concludiamo che:

$$Var[X] = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

□

Proposizione 4.15

Se $X = X_1 + X_2$, dove $X_1 \sim P(\lambda_1)$ e $X_2 \sim P(\lambda_2)$ con X_1 e X_2 indipendenti, allora $X \sim P(\lambda_1 + \lambda_2)$

Dimostrazione:

- Osserviamo che

$$P(X = k) = \sum_{h=0}^{\infty} P(X = k \mid X_1 = h) \cdot P(X_1 = h) =$$

$$= \sum_{h=0}^{\infty} P(X_1 + X_2 = k \mid X_1 = h) \cdot P(X_1 = h) =$$

$$= \sum_{h=0}^{\infty} P(h + X_2 = k \mid X_1 = h) \cdot P(X_1 = h)$$

- Siccome X_1 e X_2 sono indipendenti, allora:

$$P(X = k) = \sum_{h=0}^{\infty} P(X_2 = k-h \mid X_1 = h) \cdot P(X_1 = h) = \sum_{h=0}^{\infty} P(X_2 = k-h) \cdot P(X_1 = h) =$$

$$= \sum_{h=0}^{\infty} e^{-\lambda_2} \frac{\lambda_2^{k-h}}{(k-h)!} e^{-\lambda_1} \frac{\lambda_1^h}{h!} = e^{-(\lambda_1+\lambda_2)} \sum_{h=0}^{\infty} \frac{\lambda_2^{k-h}}{(k-h)!} \frac{\lambda_1^h}{h!} =$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{h=0}^{\infty} \frac{\lambda_2^{k-h}}{(k-h)!} \frac{\lambda_1^h}{h!} \frac{k!}{k!} = \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{h=0}^{\infty} \binom{k}{h} \lambda_2^{k-h} \lambda_1^h =$$

$$\frac{e^{-(\lambda_1+\lambda_2)}}{k!} (\lambda_1 + \lambda_2)^k \implies X \sim P(\lambda_1 + \lambda_2)$$

□

4.11 Formulario variabili aleatorie

- Data una **variabile aleatoria di Bernoulli** $X \sim B(p)$, si ha che:

$$P_X(k) = \begin{cases} p & \text{se } k = 1 \\ 1 - p & \text{se } k = 0 \end{cases}$$

$$E[X] = p$$

$$Var[X] = p(1 - p)$$

- Data una **variabile aleatoria binomiale** $X \sim B(n, p)$, si ha che:

$$P_X(k) = \binom{n}{k} p^k (1 - p)^n - k$$

$$E[X] = np$$

$$Var[X] = np(1 - p)$$

- Data una **variabile aleatoria geometrica** $X \sim G(p)$, si ha che:

$$P_X(k) = (1 - p)^{k-1} p$$

$$E[X] = \frac{1}{p}$$

$$Var[X] = \frac{1 - p}{p^2}$$

- Data una **variabile aleatoria ipergeometrica** $X \sim H(n, m, r)$, si ha che:

$$P_X(k) = \frac{\binom{m}{k}\binom{n-m}{r-k}}{\binom{n}{r}}$$

$$E[X] = \frac{mr}{n}$$

$$Var[X] = \frac{rm(n - r)(n - m)}{n^2(n - 1)} = E[X]\frac{(n - r)(n - m)}{n(n - 1)}$$

- Data una **variabile aleatoria di Poisson** $X \sim P(\lambda)$, si ha che:

$$P_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

$$E[X] = \lambda$$

$$Var[X] = \lambda$$

- Data una **variabile aleatoria continua uniforme** $X \sim U(a,b)$, si ha che:

$$P_x(k) = \begin{cases} \frac{1}{b-a} & \text{se } k \in [a,b] \\ 0 & \text{se } k \notin [a,b] \end{cases}$$

$$E[X] = \frac{a+b}{2}$$

$$Var[X] = \frac{(b-a)^2}{12}$$

5

Teoremi limite

5.1 Disuguaglianza di Markov

Definizione 5.1: Disuguaglianza di Markov

Se X è una variabile aleatoria che assume solo valori non negativi, allora $\forall \ell > 0$ si ha che:

$$P(x \geq \ell) \leq \frac{E[x]}{\ell}$$

Dimostrazione:

- Sia
$$I = \begin{cases} 1 & \text{se } X \geq \ell \\ 0 & \text{se } X < \ell \end{cases}$$

- Siccome $X > 0$, allora sicuramente si ha che
$$I \leq \frac{X}{\ell}$$

- Calcolando il valore atteso di entrambi i valori, si ha che:
$$E[I] \leq E\left[\frac{X}{\ell}\right] \implies E[I] \leq \frac{E[X]}{\ell} \implies$$
$$\implies (0 \cdot P(X < \ell) + 1 \cdot P(X \geq \ell)) \leq \frac{E[X]}{\ell} \implies P(X \geq \ell) \leq \frac{E[X]}{\ell}$$

□

Corollario 5.1

Se $f : \mathbb{R} \to \mathbb{R}$, allora
$$P(f(X) \geq f(\ell)) \leq \frac{E[f(X)]}{f(\ell)}$$

Esempio:

- Sia $X \sim B(n,p)$, si verifica che:

$$P(X \geq \ell) \leq \frac{E[X]}{\ell} = \frac{np}{\ell}$$

5.2 Disuguaglianza di Čebyšëv

Definizione 5.2: Disuguaglianza di Čebyšëv

Se X è una variabile aleatoria che assume solo valori non negativi, allora $\forall \ell > 0$ si ha che:
$$P(|X - E[X]| \geq \ell) \leq \frac{Var[X]}{\ell^2}$$

Dimostrazione:

- Poniamo $Y = |X - E[X]|$ e $f(x) = x^2$. Sappiamo che:

$$P(f(Y) \geq f(\ell) \leq \frac{E[f(X)]}{f(\ell)} \implies P(|X - E[X]|^2 \geq \ell^2) \leq \frac{E[|X - E[X]|^2]}{f(\ell)} \implies$$
$$P(|X - E[X]|^2 \geq \ell^2) \leq \frac{E[(X - E[X])^2]}{f(\ell)} = \frac{Var[X]}{f(\ell)}$$

- Inoltre, sappiamo che:
$$P(Y \geq \ell) \leq P(f(Y) \geq f(\ell))$$

- Dunque, concludiamo che:

$$P(Y \geq \ell) \leq P(f(Y) \geq f(\ell)) \leq \frac{Var[X]}{f(\ell)} \implies P(Y \geq \ell) \leq \frac{Var[X]}{f(\ell)} \implies$$
$$\implies P(|X - E[X]| \geq \ell) \leq \frac{Var[X]}{\ell^2}$$

□

Esempio:

- Sia X una variabile aleatoria dove $E[X] = 50$ e $Var[X] = 25$. Vogliamo sapere il massimo valore della probabilità che $X \geq 75$ e il massimo valore della probabilità che $40 < X < 60$.
- Per la disuguaglianza di Markov, sappiamo che:
$$P(X \geq 75) \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3}$$
- Riformulando la seconda richiesta, si ha che:
$$P(40 < X < 60) = P(|X - 50| \leq 10) = 1 - P(|X - 50| \geq 10)$$
Per la disuguaglianza di Čebyšëv, sappiamo che:
$$P(P(|X - 50| \geq 10) \leq \frac{Var[X]}{10^2} = \frac{25}{100} = \frac{1}{4}$$
Dunque concludiamo che:
$$P(40 < X < 60) = 1 - \frac{1}{4} = \frac{3}{4}$$

5.3 Leggi dei grandi numeri

Definizione 5.3: Legge debole dei grandi numeri

Siano $X_1, \ldots, X_n$ una successione di variabili aleatorie indipendenti aventi tutte la stessa distribuzione, dove $E[X_i] = \mu, \forall i \in [1, n]$.

Posto:
$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_i = \frac{X_1 + \ldots + X_n}{n}$$
per ogni $\varepsilon > 0$, si ha che:
$$\lim_{n \to +\infty} P\left(\left|\overline{X_n} - \mu\right| \geq \varepsilon\right) = 0$$
E quindi che:
$$\lim_{n \to +\infty} P\left(\left|\overline{X_n} - \mu\right| \leq \varepsilon\right) = 1$$

Dimostrazione:

- Si verifica che:

$$Var[\overline{X_n}] = Var\left[\frac{X_1 + \ldots + X_n}{n}\right] = \frac{1}{n^2} \cdot Var[X_1 + \ldots + X_n]$$

- Siccome $X_1, \ldots, X_n$ sono tutte indipendenti ed hanno la stessa varianza $Var[X_1] = \ldots = Var[X_n]$, poiché seguono la stessa identica legge di distribuzione, allora

$$Var[\overline{X_n}] = \frac{1}{n^2} \cdot Var[X_1 + \ldots + X_n] = \frac{1}{n^2} \cdot \sum_{k=1}^{n} Var[X_k] = \frac{1}{n^2} \cdot n \cdot Var[X_1] = \frac{Var[X_1]}{n}$$

- Inoltre, notiamo come:

$$E[\overline{X_n}] = \frac{E[X_1 + \ldots + X_n]}{n} = \frac{E[X_1] + \ldots + E[X_n]}{n} = \frac{n\mu}{n} = \mu$$

- Per la disuguaglianza di Čebyšëv, si ha che:

$$P(\left|\overline{X_n} - E[\overline{X_n}]\right| \geq \varepsilon) \leq \frac{Var[\overline{X_n}]}{\varepsilon^2} = \frac{Var[X_1]}{n\varepsilon} \implies$$

$$\lim_{n\to+\infty} P(\left|\overline{X_n} - \mu\right] \geq \varepsilon\right|) \leq \lim_{n\to+\infty} \frac{Var[X_1]}{n\varepsilon} \implies$$

$$\implies \lim_{n\to+\infty} P(\left|\overline{X_n} - \mu \geq \varepsilon\right|) \leq 0$$

$\square$

Definizione 5.4: Legge forte dei grandi numeri

Siano $X_1, \ldots, X_n$ una successione di variabili aleatorie indipendenti aventi tutte la stessa distribuzione, dove $E[X_i] = \mu, \forall i \in [1, n]$.

Posto:

$$\overline{X_n} = \frac{1}{n} \sum_{k=1}^{n} X_i = \frac{X_1 + \ldots + X_n}{n}$$

Si ha che:

$$P\left(\lim_{n\to+\infty} \overline{X_n} = \mu\right) = 1$$