Calcolo combinatorio

- Permutazioni di n oggetti: $P_n = n!$

- Disposizioni di k oggetti scelti tra n (numero di modi in cui si possono formare gruppi di k oggetti scelti tra n:

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

Possiamo ritrovare le permutazioni come caso particolare di disposizioni: $P_n = D_{n,n}$.

- Disposizioni con ripetizione: dati n oggetti distinti, quante liste di lunghezza k ~~p~~ distinte posso formare? (Es: colonna del totocalcio: oggetti $= (1, X, 2)$ $k = 13$ simboli

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

- Combinazioni semplici: dati n oggetti distinti sono i gruppi di k oggetti distinti che si possono formare con gli n oggetti. Due gruppi sono differenti se differiscono per la composizione (cioè se sono formati da oggetti diversi, indipendentemente dall'ordine in cui sono elencati).

Fissata una combinazione di k oggetti questa identifica $k!$ disposizioni semplici dei k oggetti che la compongono. Quindi:

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

$\binom{n}{k}$ si chiama coefficiente binomiale.

- Combinazioni con ripetizione:

Dati n oggetti, sono i possibili gruppi di k degli n oggetti distinti e con ripetizioni.

Osservazione: k può anche essere $> n$. In questo caso ci saranno necessariamente delle ripetizioni.

Due gruppi saranno considerati identici se e solo se sono formati dagli stessi elementi presi lo stesso numero di volte, indipendentemente dall'ordine. Si indicano con $C'_{n,k}$.

Indicando con $i_1 \dots i_k$ gli indici degli oggetti che fanno parte del gruppo, $C'_{n,k}$ è il numero di modi in cui si possono scegliere gli interi $i_1, \dots i_k \in \{1, 2, \dots, n\}$. Senza perdere in generalità possiamo supporre $1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n$. Questo è come dire:

$1 \leq i_1 < i_2+1 < i_3+2 < \dots < i_k+(k-1) \leq n+(k-1)$.

I valori $b_1 = i_1$, $b_2 = i_2+1$, $b_3 = i_3+2 \dots$, $b_k = i_k+(k-1)$ sono tutti distinti e variano nell'insieme $\{1, \dots, n+k-1\}$. Quindi:

$$C'_{n,k} = \binom{n+k-1}{k}$$

Esempio 1: 60 studenti devono sostenere un esame.
Per domani sono previsti 8 orali. In quanti modi può essere fatta la lista degli esaminandi?

Liste distinte se contengono nomi diversi o se contengono gli stessi nomi in un altro ordine. Sono disposizioni semplici:

$$D_{60,8} = \frac{60!}{52!} = \underbrace{60 \times 59 \times \cdots \times 53}_{8 \text{ fattori.}}$$

Esempio 2: Dopo l'esame i 60 studenti festeggiano con una partita a calcio contro i professori. In quanti modi diversi può essere scelta la formazione titolare degli studenti?

Formazioni diverse solo se differiscono per almeno un nome. Inoltre uno studente non può comparire più di una volta nella lista. Sono combinazioni semplici:

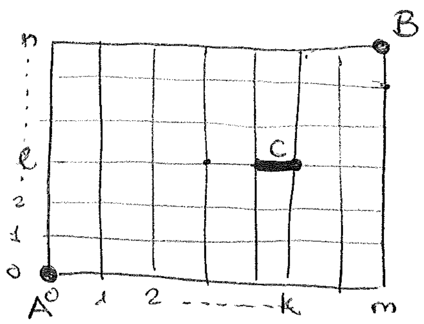
$$C_{60,11} = \binom{60}{11} = \frac{60!}{11!\,49!}$$

Esempio 3: I soliti 60 studenti partecipano ad alcune gare. Ci sono 5 premi uguali in palio. In quanti modi diversi possono essere assegnati? (I premi sono identici, conta solo quanti premi ha preso ciascuno studente).

Se numeriamo gli studenti da 1 a 60, l'elenco dei premiati è una lista di 5 numeri da 1 a 60 eventualmente con ripetizioni. Sono combinazioni con ripetizione:

$$C'_{60,5} = \binom{64}{5} = 64 \times 63 \times 62 \times 61 \times 60$$

Esercizio 1: Cartine di Manhattan (o, se preferite, di Viareggio). Isolati rettangolari. Voglio andare da A a B muovendomi solo verso ↑ o verso →.



Quanti sono i possibili percorsi distinti da A a B?

Sol: Tutti i percorsi hanno uguale lunghezza $n+m$. Se indico i movimenti con N = nord e E = est posso identificare un percorso con una stringa di N ed E in cui compaiono m "E" ed n "N". Due percorsi differiscono quindi solo per l'ordine in cui compaiono i simboli.

Quando scrivo un percorso devo riempire m caselle su $n+m$ con delle "E". Le altre seguono automaticamente. Si tratta quindi di Combinazioni senza ripetizioni di m oggetti (le caselle dove metto le "E") scelti tra $m+n$ (la lunghezza del percorso).

Sono quindi:
$$C_{m+n,m} = \binom{m+n}{m}.$$

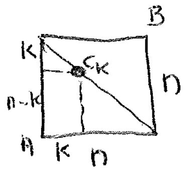
Nota: se avessi deciso di disporre le "N" e non le "E" avrei ottenuto, analogamente
$$C_{m+n,n} = \binom{m+n}{n}.$$

Questi due numeri rappresentano la stessa quantità e quindi devono essere uguali! Infatti:
$$C_{n+m,m} = \binom{n+m}{m} = \frac{(n+m)!}{m!\,n!} = \frac{(n+m)!}{n!\,m!} = \binom{m+n}{n} = C_{m+n,n}$$

Esercizio 2: Se, nella figura il punto C ha coordinate (k,ℓ), quanti sono i percorsi da A a B che passano per C?
E quelli che NON passano per il tratto in grassetto?

Proprietà dei coefficienti binomiali (n, k interi ≥ 0, $n \geq k$).

- $\binom{n}{k} = \binom{n}{n-k}$
- $\binom{n}{0} = \binom{n}{n} = 1$
- $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
- $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$
- Binomio di Newton: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.
- $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$ (vedi esercizio 2)



Dim: $\binom{2n}{n}$ = n° percorsi da A a B.

$\binom{n}{k}\binom{n}{n-k}$ = n° percorsi da A a B che passano per C_k.

Un percorso da A a B passa per uno ed un solo C_k per cui:

$$\sum_{k=0}^{n} \binom{n}{k}\binom{n}{n-k} = \binom{2n}{n}$$ Ma $\binom{n}{n-k} = \binom{n}{k}$ e segue la tesi.

Probabilità e statistica

Introduzione

Il calcolo delle probabilità è quel ramo della matematica che si occupa di calcolare la "probabilità" di certi eventi a partire da quelle di altri, in generale più semplici, per i quali si pensa di poterle valutare direttamente.

Esempio: Gioco del lotto

Generalmente pensiamo che tutti i 90 numeri hanno la stessa probabilità di essere estratti (cioè 1/90).
Possiamo utilizzare questo fatto per calcolare la probabilità che esca un certo ambo, terno, che un numero accumuli un certo ritardo e così via...

Ci sono vari punti di partenza possibili per definire una "Teoria della probabilità": a seconda della scelta iniziale avremo una impostazione "classica", "frequentista", "soggettiva",...

Impostazione classica

È la prima sviluppata in ordine di tempo (XVII secolo) (Pascal). La probabilità che accada un evento è definita come il rapporto tra il numero di casi favorevoli all'evento e il numero totale di casi possibili, purché questi siano tutti equiprobabili.

Si osservi che non può essere considerata una definizione vera e propria in quanto nasconde un loop. La probabilità è definita utilizzando la parola "equiprobabili".

La probabilità è comunque un numero compreso tra 0 ed 1 (nessun caso favorevole, tutti i casi favorevoli rispettivamente).
Avremo quindi probabilità 1 per l'evento certo, probabilità 0 per quello impossibile, almeno fin quando si parla di un numero finito di casi.

Una importante conseguenza della definizione è l'additività della probabilità (legge della probabilità totali):

Proposizione: Se due eventi sono INCOMPATIBILI (cioè non possono verificarsi contemporaneamente) la probabilità della loro unione è uguale alla somma delle probabilità dei singoli eventi.

Infatti, se i due eventi sono incompatibili nessun caso è favorevole ad entrambi contemporaneamente. I casi favorevoli ad $A \cup B$ sono esattamente la somma di quelli favorevoli solo ad A e di quelli favorevoli solo a B.

Tipicamente, i campi nei quali si può applicare direttamente la definizione classica di probabilità sono quelli dei giochi di dadi, di carte, estrazioni ecc ecc.
Si tratta, in generale, di calcolare il numero di casi possibili e di casi favorevoli all'evento in questione.
In questo gioca un ruolo fondamentale il calcolo combinatorio.

Esempio: Distribuendo le 52 carte in un mazzo mescolato, qual'è la probabilità che i 4 giocatori abbiano rispettivamente $13 \clubsuit$, $13 \diamondsuit$, $13 \heartsuit$, $13 \spadesuit$?

Soluzione: 1° modo: ogni giocatore ha 13 carte che possono essere pescate in 13! modi diversi (non ci interessa in che ordine ovvero in mano ma solo quali sono). [CASI FAVOREVOLI]
Il numero dei possibili modi in cui può essere mescolato un mazzo sono 52!. [CASI TOTALI].
Quindi la probabilità è $\dfrac{(13!)^4}{52!} \simeq 1{,}8 \cdot 10^{-29}$.

2° modo: le carte del 1° giocatore sono 13 scelte tra 52. Ci sono $\binom{52}{13}$ modi distinti di pescarle. Analogamente, per gli altri due giocatori ci sono $\binom{39}{13}, \binom{26}{13}, \binom{13}{13}$ modi possibili. Fra tutti questi casi possibili uno solo è favorevole, quindi la probabilità è $\left[\dfrac{52!}{13! \cdot 39!} \cdot \dfrac{39!}{13!\,26!} \cdot \dfrac{26!}{13! \cdot 13!}\right]^{-1} = \dfrac{(13!)^4}{52!}$.

Esempio: Distribuendo 13 carte a 4 giocatori da un mazzo di 52, qual'è la probabilità che NORD abbia 7 carte di picche?

Soluzione: I casi possibili sono ancora 52! Esaminiamo la mano di NORD:

Ha 7 picche su 13 : ci sono $\binom{13}{7}$ casi.

Ha 6 altre carte su 39 : ci sono $\binom{39}{6}$ casi.

1° modo: Dobbiamo tenere conto che non ci interessa l'ordine in cui sono distribuite le carte di NORD, così come non ci interessa l'ordine in cui sono distribuite le carte di SUD, EST, WEST.

Quindi

$$p = \frac{\binom{13}{7} \cdot 13! \cdot \binom{39}{6} \cdot 39!}{52!} = \simeq 0{,}0088$$

2° modo: Alternativamente, utilizzando le combinazioni, i modi possibili di scegliere le 13 carte di nord sono $\binom{52}{13}$ mentre i casi favorevoli sono $\binom{13}{7}\binom{39}{6}$. Quindi

$$p = \frac{\binom{13}{7}\binom{39}{6}}{\binom{52}{13}}.$$

Impostazione frequentista

L'impostazione frequentista si può sintetizzare nell'affermazione seguente:

"in una successione di prove fatte nelle stesse condizioni, la frequenza di un evento si avvicina alla probabilità dell'evento stesso e l'approssimazione tende a migliorare all'aumentare del numero di prove".

La probabilità di un evento è il limite della frequenza (relativa) dei successi quando il numero delle prove tende all'infinito.

Impostazione soggettiva

L'impostazione soggettiva intende la probabilità come il

"grado di fiducia che una persona ha nel verificarsi di un evento".

In questa impostazione la probabilità perde il significato assoluto di numero legato all'evento e dipende dall'opinione soggettiva di chi si trova di fronte all'evento.

Un modo per rendere utilizzabile questa definizione è di interpretarla riferendosi alle scommesse: è immaginare la probabilità come il prezzo che si è disposti a pagare (che si ritiene equo pagare) per ricevere 1 se l'evento si verifica, 0 altrimenti. Prezzo equo significa che non ci deve essere né un guadagno certo né una perdita certa.

Per chiarire questo concetto, cerchiamo di dedurre le 3 proprietà stabilite per la definizione classica:

i) $\mathbb{P}(A) \geq 0$ qualunque sia l'evento A.

ii) Se A è l'evento certo, $\mathbb{P}(A) = 1$

iii) Se A e B sono eventi incompatibili allora $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

ii) Se A è l'evento certo e $\mathbb{P}(A) > 1$ si ricaverà sicuramente un guadagno. Se fosse invece $\mathbb{P}(A) < 1$ si avrà sicuramente una perdita. Quindi $\mathbb{P}(A) = 1$.

i) In generale, pagando $\mathbb{P}(A)$ per la scommessa si avrà un guadagno $1 - \mathbb{P}(A)$ oppure una perdita $-\mathbb{P}(A)$. Se $\mathbb{P}(A)$ fosse negativa nel secondo caso avremmo un guadagno certo quindi $\mathbb{P}(A) \geq 0$.

iii) Consideriamo gli eventi $A_1, A_2, \dots, A_n$ incompatibili fra loro e tali che uno di essi si presenti necessariamente. Dimostriamo che la somma delle loro probabilità è 1:

consideriamo n scommesse su ciascuno degli eventi: paghiamo $\mathbb{P}(A_i)$ per ricevere 1 se A_i si verifica, 0 altrimenti.

Scommettendo su tutti gli eventi paghiamo $\sum_{i=1}^{n} \mathbb{P}(A_i)$ per ricevere sicuramente 1.

Per le condizioni di coerenza quello che paghiamo deve essere uguale a quello che riceviamo. Quindi $\sum_{i=1}^{n} \mathbb{P}(A_i) = 1$.

Se ora A e B sono incompatibili, sia $C = \overline{A \cup B}$. A, B, C sono incompatibili e uno si presenta necessariamente. Segue $\mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) = 1$

Scommesse SNAI, casinò e TOTOCALCIO

Esempio: Quote SNAI del 29/1/06:

$$\boxed{\text{Tunisia - Guinea} \quad 1{,}75 \quad 3.25 \quad 4.50}$$ per 1X2 rispettivamente.

Nella nostra notazione significa che

$p_1 = \frac{1}{1.75} = 0{,}5728$; $\quad p_x = \frac{1}{3.25} = 0.3076$; $\quad p_2 = \frac{1}{4.50} = 0.2222$

Se il gioco fosse equo dovremmo avere $p_1 + p_x + p_2 = 1$.

Invece si ha: $p_1 + p_x + p_2 = 1{,}1026$

Altro esempio

$$\boxed{\text{Nigeria - Senegal} \quad 2{,}25 \quad 3.10 \quad 3.00}$$ Si ha:

$p_1 = \frac{1}{2.25} = 0{,}4444$ $\quad p_x = \frac{1}{3.10} = 0.3226$ $\quad p_2 = \frac{1}{3{,}00} = 0.3333$

Di conseguenza: $p_1 + p_x + p_2 = 1{,}1003$

$$\boxed{\text{Ghana - Zimbawe} \quad 1.22 \quad 5.00 \quad 12.00}$$

$p_1 = \frac{1}{1.22} = 0.8197$ $\quad p_x = \frac{1}{5.00} = 0.2000$ $\quad p_2 = \frac{1}{12.00} = 0.0833$

Quindi $p_1 + p_x + p_2 = 1{,}1030$

Questo significa che, mediamente, la SNAI ha garantito un guadagno del 10% sul monte scommesse sul lungo termine

Esempio: Roulette. 36 numeri più lo zero. Se esce 0 vince il banco.

~~$p_1 = p_2 = \dots = p_{36} = \frac{1}{37}$ (se gioco 1 euro ne ricevo 36)~~

Se gioco 1 euro ne ricevo 36 mentre ho $\frac{1}{37}$ di prob di vincere. Il casinò ha garantito il 2,70% del monte scommesse sul numero secco $\left(\frac{1}{37}\right)$.

Confronto ~~con~~ tra le varie impostazioni della probabilità.

<u>Valutazione classica:</u> i casi sono considerati equiprobabili.

Se abbiamo a che fare con una estrazione del lotto è ragionevole.
Se invece parliamo dell'esito della partita Fiorentina - Montevarchi è un'impostazione poco ragionevole.

Ci sono poi situazioni in cui il calcolo del numero di casi è difficile se non impossibile. Es: Dividiamo un segmento in 3 parti a caso. Qual'è la probabilità che formino un triangolo.

* Altro esempio: In un paese ci sono 5000 abitanti. Qual'è la probabilità che esattamente r siano fumatori?

La risposta classica è che $\mathbb{P}\left(\frac{r}{5000}\right) = \frac{r}{5001}$ per $r = 0, \dots, 5000$

<u>Valutazione frequentista:</u>

All'ultima domanda risponderò che $\mathbb{P}\left(\frac{r}{5000}\right) = \alpha$ se rilevo tale frequenza di risposte in un dialogo fatto nei paesi limitrofi.

Anche qui ci sono però dei punti a sfavore: se voglio valutare in maniera frequentista la probabilità con cui un giocatore di pallacanestro fa centro nei tiri liberi sono nei guai perché non è rispettata l'uniformità delle condizioni: il giocatore può essere più o meno stanco, più o meno concentrato ecc ecc.

Analogamente non posso utilizzare questo tipo di impostazione per eventi che si possono verificare ma che non si sono mai realizzati.

Es: Livorno campione d'Italia, latino scritto alla maturità scientifica ecc ecc

D'altra parte anche $A \cup B$ e C sono incompatibili e uno dei due si presenta certamente. Quindi $\mathbb{P}(A\cup B)+\mathbb{P}(C)=1$. Confrontando,
$\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)$.

<u>Osservazione</u>: nel corso della dimostrazione di iii) abbiamo dimostrato anche che $\mathbb{P}(\bar{A})=1-\mathbb{P}(A)$.

Definizione assiomatica della probabilità

È stata formulata negli anni '30. In essa si parte da un insieme di assiomi o postulati che definiscono gli oggetti dal punto di vista matematico. Si fissano le regole cui debbono obbedire senza riferirsi al loro contenuto sostanziale.
Da questi si traggono poi delle conseguenze. Tutta la costruzione della teoria si basa sugli assiomi di partenza.

ASSIOMA 1: Gli eventi sono sottoinsiemi di uno spazio Ω e formano una classe additiva (chiusa rispetto ad addizione e negazione).

ASSIOMA 2: Ad ogni evento A è assegnato un numero reale $\mathbb{P}(A)\geq 0$ detto probabilità di A.

ASSIOMA 3: $\mathbb{P}(\Omega)=1$

ASSIOMA 4: Se $A\cap B=\emptyset$ allora $\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)$

ASSIOMA 5: Se A_n $(n=1,2,\dots)$ è una successione di eventi e $\lim_{n\to\infty} A_n=\emptyset$ allora $\lim_{n\to\infty}\mathbb{P}(A_n)=0$.

A partire da questi 5 assiomi si possono dimostrare ad esempio le proprietà seguenti:

<u>Teorema 1</u>: $\mathbb{P}(A)+\mathbb{P}(\bar{A})=1$

Dimostrazione: $A\cup\bar{A}=\Omega$, $A\cap\bar{A}=\emptyset$ quindi $\mathbb{P}(A)+\mathbb{P}(\bar{A})=\mathbb{P}(A\cup\bar{A})=\mathbb{P}(\Omega)=1$

<u>Teorema 2</u>: $\mathbb{P}(\emptyset)=0$

Dimostrazione: $\mathbb{P}(\emptyset)=\mathbb{P}(\bar{\Omega})=1-\mathbb{P}(\Omega)=0$

Teorema 3: $\mathbb{P}(A) \leq 1$

Dimostrazione: $\mathbb{P}(A) = 1 - \mathbb{P}(\bar{A}) \leq 1$ dato che $\forall$ E evento, $\mathbb{P}(E) \geq 0$

Teorema 4: Se $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

Dimostrazione: Se $A \subset B$ allora $A = A \cap B$. Quindi $B = (A \cap B) \cup (A \cap \bar{B})$.
$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \bar{B}) = \mathbb{P}(A) + \mathbb{P}(A \cap \bar{B}) \geq \mathbb{P}(A)$ poiché $\mathbb{P}(A \cap \bar{B}) \geq 0$.
Segue $\mathbb{P}(B) \geq \mathbb{P}(A)$.

Osservazione: Sappiamo che $\mathbb{P}(\Omega) = 1$ ma non è detto che Ω sia l'unico evento ad avere probabilità 1. Analogamente $\mathbb{P}(\emptyset) = 0$ ma potrebbero esistere altri eventi non nulli di probabilità zero.
Questo significa che "evento certo" e "evento con probabilità 1" non sono le stesse cose.
Analogamente non sono sinonimi "evento impossibile" ed "evento con probabilità zero".

L'evento impossibile, se aggiunto ad un altro ~~evento~~ evento non lo modifica: $A \cup \emptyset = A$. Se sostituiamo $\emptyset$ con un evento B con $\mathbb{P}(B) = 0$ l'uguaglianza non sussiste più: $A \cup B \neq A$ ma l'uguaglianza continua a valere in probabilità: $\mathbb{P}(A \cup B) = \mathbb{P}(A)$.

Teorema 5: Se $\mathbb{P}(B) = 0$ allora: $\mathbb{P}(A \cup B) = \mathbb{P}(A)$, $\mathbb{P}(A \cap B) = 0$
Se $\mathbb{P}(C) = 1$ allora: $\mathbb{P}(A \cup C) = 1$, $\mathbb{P}(A \cap C) = \mathbb{P}(A)$

Dimostrazione: $A \cap B \subseteq B \Rightarrow \mathbb{P}(A \cap B) \leq \mathbb{P}(B) = 0$.
Analogamente: $C \subseteq A \cup C \Rightarrow 1 = \mathbb{P}(C) \leq \mathbb{P}(A \cup C). \Rightarrow \mathbb{P}(A \cup C) = 1$
Inoltre:
$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (\bar{A} \cap B)) = \mathbb{P}(A) + \mathbb{P}(\bar{A} \cap B) = \mathbb{P}(A) + 0 = \mathbb{P}(A)$$
$$\mathbb{P}(A) = \mathbb{P}((A \cap C) \cup (A \cap \bar{C})) = \mathbb{P}(A \cap C) + \mathbb{P}(A \cap \bar{C}) = \mathbb{P}(A \cap C) + 0$$

Esempio: Testa o Croce
Consideriamo una moneta regolare (cioè $\mathbb{P}(T) = \mathbb{P}(C) = 1/2$) e consideriamo una serie di lanci ripetuti.
In n lanci abbiamo 2^n risultati possibili (es: se $n=2$ ci sono 4 possibili risultati: TT, TC, CT, CC) ai quali, per motivi di simmetria attribuiamo le stesse probabilità 2^{-n}.

Se consideriamo una serie infinita di lanci i risultati possibili sono le successioni infinite di T e C.
Tali successioni avranno, sempre per motivi di simmetria tutte la stessa probabilità, cioè zero.
D'altra parte non possono essere tutti eventi impossibili dato che uno di essi si deve necessariamente verificare.

Esercizio: Quanti lanci si devono fare prima che compaia una testa?
Indichiamo con A_n, $n=1,2,\dots$ l'evento "T compare per la prima volta al lancio n-esimo".
Possiamo studiare l'evento A_n limitandoci a prove consistenti in n lanci.
A_n è costituito dalle sequenze che iniziano con $\underbrace{CC\dots C}_{n-1}T$. I casi favorevoli sono 1 sulle 2^n possibili sequenze di n lanci.
Di conseguenza $\mathbb{P}(A_n) = \frac{1}{2^n}$.

Se Ω è l'insieme composto da tutte le sequenze infinite, $\Omega \supseteq \bigcup_{n=1}^{\infty} A_n$.
Inoltre gli A_n sono tra loro incompatibili.
L'unica stringa che si trova in $\Omega \setminus \bigcup_{n=1}^{\infty} A_n$ è quella formata soltanto da C.
Chiamiamo A_0 l'evento consistente nel verificarsi della stringa formata solo da C. Allora:

$$\mathbb{P}(A_0) = \mathbb{P}(\Omega) - \sum_{n=1}^{\infty} \mathbb{P}(A_n) = 1 - \sum_{n=1}^{\infty} \frac{1}{2^n} = 1 - \frac{1}{2}\,\frac{1}{1-\frac{1}{2}} = 1 - 1 = 0$$

A_0 è un esempio di evento non impossibile ma di probabilità zero.

Esercizio: Stessa situazione precedente.
Sia B l'evento "T si presenta per la prima volta in un lancio di posto pari". Calcoliamo $\mathbb{P}(B)$:

$$\mathbb{P}(B) = \sum_{k=1}^{\infty} \mathbb{P}(A_{2k}) = \sum_{k=1}^{\infty} \frac{1}{2^{2k}} = \frac{1}{4}\,\frac{1}{1-\frac{1}{4}} = \frac{1}{4}\cdot\frac{4}{3} = 1/3$$

Esercizio: Sempre stessa situazione:
Sia D l'evento "T si verifica prima del k-esimo lancio".

$$\mathbb{P}(D) = 1 - \mathbb{P}\left(\bigcup_{n=k}^{\infty} A_n\right) = 1 - \sum_{n=k}^{\infty} \frac{1}{2^n} = 1 - \frac{1}{2^k}\,\frac{1}{1-\frac{1}{2}} = 1 - \frac{1}{2^{k-1}}$$

Probabilità condizionata

Consideriamo eventi del tipo:

- la durata di uno pneumatico sia superiore a 70000 km sapendo che ha già superato i 50000
- l'estrazione (senza reimmissione) di una pallina al decimo tentativo da un'urna contenente 100 palline di cui il 10% bianche sia una pallina bianca, sapendo che nelle 9 estrazioni precedenti sono state estratte già 5 palline bianche.
- in un controllo di qualità su un lotto di 150 pezzi con difettosità del 12.5 %, la sesta estrazione di un pezzo difettoso avendo già dato pezzi difettosi le 5 estrazioni precedenti.

Rispetto agli eventi visti finora intuitivamente forniamo probabilità diverse rispetto agli analoghi senza la condizione finale.
Vediamo un esempio:

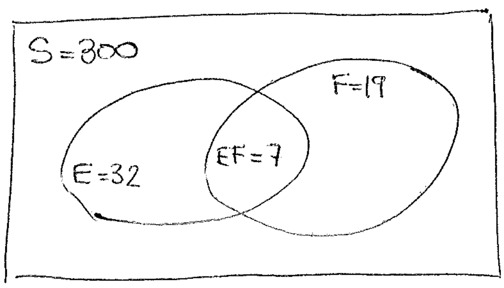
Esempio: In un lotto di 300 pneumatici, 25 hanno irregolarità nello spessore del battistrada, 12 hanno tracce non regolamentari e 7 hanno entrambi i difetti.
Supponiamo che ciascun pneumatico ~~ha~~ abbia le stesse probabilità di essere estratto. Sia E = "estrazione uno pneumatico con spessore irregolare", Sia F l'evento "estrazione di uno pneumatico con tracce non regolamentari".

$$\mathbb{P}(F) = \frac{19}{300}.$$

La probabilità di estrarre uno pneumatico avente le tracce non regolamentari, sapendo che ha delle irregolarità nello spessore è

$$\mathbb{P}(F|E) = \frac{7}{32}$$ ($F|E$ significa accade F sapendo che E è vero.)

In pratica è come se avessimo operato una restrizione del campione da tutto S ad E ed applicato la definizione di probabilità a priori



$\mathbb{P}(F|E)$ rappresenta la prob. che accade F sapendo che E si è verificato.

Possiamo scrivere:

$$\mathbb{P}(F|E) = \frac{\#FE}{\#E} = \frac{\frac{\#FE}{\#S}}{\frac{\#E}{\#S}} = \frac{\mathbb{P}(FE)}{\mathbb{P}(E)}$$

Questo ci porta a definire la probabilità condizionata:

Definizione: La probabilità condizionata dell'evento A dato l'evento B, con A e B eventi qualunque di uno spazio Ω è data da:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}, \quad \text{con} \quad \mathbb{P}(B) \neq 0$$

Similmente si può definire la probabilità dell'evento B dato A:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)}, \quad \text{con} \quad \mathbb{P}(A) \neq 0$$

Esempio: Lanciamo due dadi bilanciati. Lo spazio campione sarà dato dalle 36 coppie ordinate da $(1,1)$ a $(6,6)$: $S = \{(1,1), (1,2), \dots, (6,5), (6,6)\}$. Vogliamo calcolare la probabilità del verificarsi dell'evento A = {"la ~~probabilità~~ somma dei numeri usciti è maggiore di 8"} condizionata all'evento B = {uno dei numeri usciti è un 6}.

$\#S = 36$

$A = \{(3,6), (4,6), (4,5), (5,6), (5,5), (5,4), (6,6), (6,5), (6,4), (6,3)\}$

$B = \{(1,6), (2,6), (3,6), (4,6), (5,6), (6,6), (6,1), (6,2), (6,3), (6,4), (6,5)\}$

Si ha $\#A = 10$, $\#B = 11$. Inoltre

$AB = \{(3,6), (4,6), (5,6), (6,6), (6,5), (6,4), (6,3)\}$ e $\#AB = 7$.

Di conseguenza:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)} = \frac{7/36}{11/36} = 0{,}6363\ldots$$

Esempio: Calcolare la probabilità che sia estratto il 27 sulla ruota di Firenze due volte consecutive sapendo che è già stato estratto una volta.

Soluzione: Sia $F = \{$esce il 27 due volte consecutive$\}$ e sia

$E = \{$esce il 27 alla 1^a estrazione$\}$

$\mathbb{P}(F|E) = \dfrac{\mathbb{P}(FE)}{\mathbb{P}(E)}$. Ma $F = E \cap F$ quindi $\mathbb{P}(F) = \mathbb{P}(EF)$

Segue: $\mathbb{P}(F|E) = \dfrac{\mathbb{P}(F)}{\mathbb{P}(E)} = \dfrac{\frac{1}{90 \times 90}}{\frac{1}{90}} = \dfrac{1}{90}$

NOTA: Il fatto che sia già uscito una volta non condiziona la seconda estrazione!

Teorema delle probabilità totali. Teorema di Bayes

Vediamo ora come calcolare la probabilità di eventi collegati ad esperimenti che si attuano in più fasi

Esempio: Un produttore di automobili riceve da tre fornitori i dischi dei freni da installare sulle auto nelle seguenti percentuali: 65%, 25%, 10%. Sapendo che i 3 fornitori producono dischi con una difettosità dichiarata del 5%, 10%, 25% rispettivamente, qual'è la probabilità che il produttore di auto riceva un disco difettoso?

Soluzione: Siano B_1, B_2, B_3 i 3 fornitori e siano $\mathbb{P}(B_i)$ la probabilità ~~che~~ la ditta riceva un disco dal produttore B_i.

Essendo $B_i \cap B_j = \emptyset$ si ha $\mathbb{P}(B_1) + \mathbb{P}(B_2) + \mathbb{P}(B_3) = 1$.

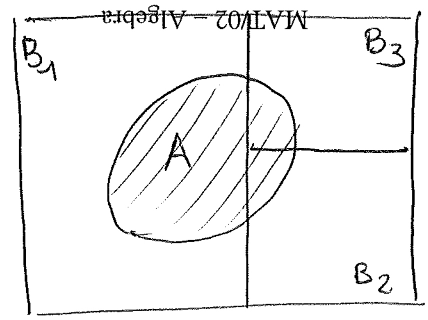
Si ha $\Omega = B_1 \cup B_2 \cup B_3$ e sia $A = \{$la ditta riceve un disco difettoso$\}$.

Allora:

$$A = A \cap S = A \cap (B_1 \cup B_2 \cup B_3) = \underbrace{(A \cap B_1)}_{C_1} \cup \underbrace{(A \cap B_2)}_{C_2} \cup \underbrace{(A \cap B_3)}_{C_3}$$

Si ha: $\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_2) + \mathbb{P}(A \cap B_3)$

dato che $C_i \cap C_j = \emptyset$ e $C_1 \cup C_2 \cup C_3 = A$



Ricordando la definizione di prob. condizionata

$$\mathbb{P}(A \cap B_i) = \mathbb{P}(A|B_i)\,\mathbb{P}(B_i)$$

Quindi
$$\mathbb{P}(A) = \sum_{i=1}^{3} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i) .$$

Nel nostro caso quindi $\mathbb{P}(A) = 0{,}65 \times 0{,}05 + 0.25 \times 0{,}10 + 0{,}10 \times 0{,}25 = 0.0825$ ∎

In pratica abbiamo dimostrato il seguente

Teorema delle probabilità totali

Se $B_1 \dots B_n$ sono una collezione di eventi mutuamente escludentesi tali che $\Omega = \bigcup_{j=1}^{3} B_j$ e tali che $\mathbb{P}(B_j) \neq \emptyset$ per ogni $j = 1, \dots, n$.
Qualsiasi sia $A \subseteq \Omega$ si ha:
$$\mathbb{P}(A) = \sum_{j=1}^{n} \mathbb{P}(A|B_j)\,\mathbb{P}(B_j)$$

Esempio (seconda parte)

Poniamoci un'altra domanda: "avendo scelto a caso un disco tra quelli ricevuti e avendolo riscontrato difettoso, qual'è la probabilità che provenga dal secondo fornitore?".
In pratica vogliamo conoscere $\mathbb{P}(B_2|A)$ (B_2 dato A).
Dal teorema delle probabilità totali si ha:
$$\mathbb{P}(B_2|A) = \frac{\mathbb{P}(B_2 \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_2)\,\mathbb{P}(B_2)}{\sum_{j=1}^{3} \mathbb{P}(A|B_j)\,\mathbb{P}(B_j)} = \frac{0{,}25 \times 0.10}{0.0825} = 0{,}303$$

In maniera analoga: $\mathbb{P}(B_1|A) = 0.394$, $\mathbb{P}(B_3|A) = 0.303$

La formula che abbiamo utilizzato si chiama Formula di Bayes.

FORMULA DI BAYES Siano $B_1, \dots B_n$ eventi ~~mut~~ incompatibili ed esaustivi e tali che $\mathbb{P}(B_j) \neq 0$ $\forall j$. Qualunque sia $A \subset \Omega$ si ha:
$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(A|B_j)\,\mathbb{P}(B_j)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i)}$$

Esercizio 1

Giocando a bridge, sulle linee N-S abbiamo 11 carte di picche.

a) Qual'è la probabilità che i resti (le 2 picche mancanti) siano una in est e una in ovest?

b) Qual'è la probabilità che i resti siano 2-0 (o 0-2)?

Soluzione.

Non ci interessa la suddivisione delle carte tra N e S. È sufficiente esaminare le mani di E e W; anzi, è sufficiente esaminare le mano di di uno solo dei due tra E e W dato che l'altra può essere dedotta per differenza.

Consideriamo dunque la composizione della mano di E: ci sono 3 possibilità:

A) Resti 1-1 : E ha 1 picche + 12 altre

B) Resti 2-0 : E ha 2 picche + 11 altre

C) Resti 0-2 : E ha 0 picche + 13 altre

$$\text{Casi totali: } A+B+C = \binom{2}{1}\binom{24}{12} + \binom{2}{2}\binom{24}{11} + \binom{2}{0}\binom{24}{13} =$$

$$= \frac{2!}{1!\,1!}\cdot\frac{24!}{12!\,12!} + \frac{2!}{0!\,2!}\cdot\frac{24!}{11!\,13!} + \frac{2!}{2!\,0!}\cdot\frac{24!}{13!\,11!} =$$

$$= \frac{24!}{13!\cdot 12!}\left(2\cdot 13 + 12 + 12\right) = 50\,\frac{24!}{13!\cdot 12!}$$

I casi favorevoli alla suddivisione 1-1 sono $A = 26\,\dfrac{24!}{13!\,12!}$, quindi

$\mathbb{P}(\text{resti } 1\text{-}1) = \dfrac{26}{50} = 0{,}52$. Di conseguenza $\mathbb{P}(2\text{-}0 \text{ oppure } 0\text{-}2) = 1 - \mathbb{P}(1\text{-}1) = 0{,}48$

Per simmetria $\mathbb{P}(2\text{-}0) = \mathbb{P}(0\text{-}2) = 0.24$.

Esercizio 2

Stessa situazione dell'esercizio 1. Supponiamo però di avere

2.1) Soltanto 10 carte di picche tra N e S. Calcolare $\mathbb{P}(3\text{-}0)$, $\mathbb{P}(2\text{-}1)$, $\mathbb{P}(1\text{-}2)$, $\mathbb{P}(0\text{-}3)$.

2.2) 9 carte di picche tra N e S. Calcolare $\mathbb{P}(4\text{-}0)$, $\mathbb{P}(3\text{-}1)$, $\mathbb{P}(2\text{-}2)$, $\mathbb{P}(1\text{-}3)$, $\mathbb{P}(4\text{-}0)$.

3.3) 8 carte di picche tra N e S. È più probabile che i resti siano (3-2 o 2-3) oppure (4-1 o 1-4)?

Esercizio 3

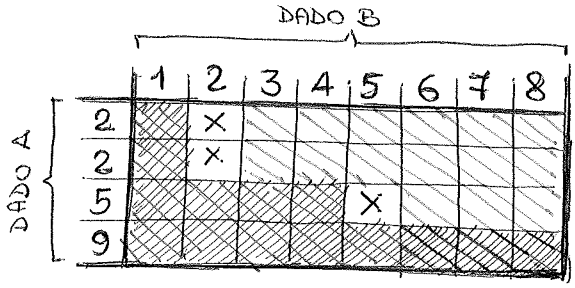
Due giocatori fanno il gioco seguente: lanciano ciascuno il proprio dado. Se esce lo stesso punteggio il lancio viene ripetuto, altrimenti vince chi ha ottenuto il punteggio più alto.

Il dado del giocatore A ha 4 facce numerate con 2,2,5,9.

Il dado del giocatore B ha 8 facce numerate con 1,2,3,4,5,6,7,8.

a) Se il vincitore riscuote un euro, quali sono gli importi pagati dai giocatori A e B per giocare in modo che il gioco sia equo?

b) Se il gioco fosse modificato in modo che in caso di pareggio vince il banco, qual'è la probabilità di successo del gestore del gioco?



Soluzione:

a) I possibili esiti del lancio dei due dadi sono $4 \times 8 = 32$ (non necessariamente distinti, dato che 2 delle 4 facce del dado A hanno segnato lo stesso numero)

I 32 casi sono equiprobabili. Esaminiamo i possibili esiti:

Esito pari = 3 casi $\Rightarrow \mathbb{P}(\text{pareggio}) = 3/32 = 0.09375$

Vince A = 14 casi $\Rightarrow \mathbb{P}(\text{vince } A^*) = 14/32 = 0.4375$

Vince B = 15 casi $\Rightarrow \mathbb{P}(\text{vince } B^*) = 15/32 = 0.46875$

Totale = 32 casi

(* = vince al 1° lancio!)

In caso di pareggio il lancio viene ripetuto quindi il 9.375% viene ripartito tra A e B in proporzione di 14:15.

In pratica è come se i 3 casi di pareggio non si presentassero!

$$\mathbb{P}(\text{vince } A) = \frac{14}{29}, \quad \mathbb{P}(\text{vince } B) = \frac{15}{29}.$$

Nota: Se calcoliamo il valore atteso dei due dadi abbiamo $E(A) = E(B) = 4.5$.

Il valore atteso non è una quantità idonea a giudicare l'equità di un gioco.

b) Il banco incassa la posta in caso di pareggio. Abbiamo visto che questo accade con probabilità 0,09375.

Esercizio 4:

Quante sono le permutazioni distinte della parola "TRENTATRE"? Scrivendo un anagramma a caso, qual'è la probabilità che le lettere siano in ordine alfabetico?

Soluzione: Analizziamo la composizione della parola: 1A, 1N, 2E, 2R, 3T. Abbiamo 9 lettere quindi le permutazioni possibili sono 9! Fra queste però dobbiamo identificare i "doppioni": ci sono 3! modi di permutare le T senza modificare la parola. Analogamente, ci sono 2! modi di permutare fra loro le due E o le due R.
In definitiva, ciascun anagramma ha $3! \cdot 2! \cdot 2!$ permutazioni che lo lasciano invariato (cioè i "doppioni").
Gli anagrammi distinti di "TRENTATRE" sono allora $\dfrac{9!}{3!\,2!\,2!} = \dfrac{362880}{24}$

$= 15120.$

Perchè un anagramma risulti in ordine alfabetico è necessario che figurino nell'ordine A, E, N, R, T.
Eliminati i doppioni l'unico caso favorevole è AEENRRTTT. La probabilità è allora $\dfrac{1}{15120} = 0.6614 \cdot 10^{-4}$.

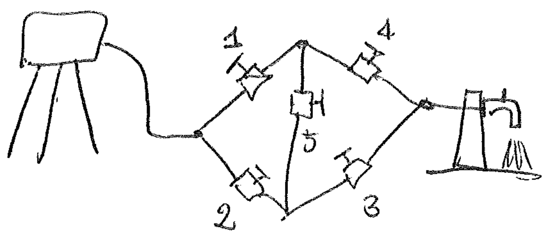
Contando anche i doppioni ovvero $3!\,2!\,2! = 24$ casi favorevoli su $9!$.
Il risultato, ovviamente, è lo stesso.

Esercizi Bayes

1) Supponendo equiprobabili maschi e femmine, qual'è la probabilità che una coppia con 3 figli abbia due figli maschi sapendo che ne ha uno?

2) Idem, che abbia un maschio sapendo che uno dei tre è femmina?

3) Se $\mathbb{P}(B) = \mathbb{P}(A|B) = \mathbb{P}(C|(A \cap B)) = \frac{1}{2}$, quanto vale $\mathbb{P}(A \cap B \cap C)$? $[1/8]$

4) 

$$\left[\begin{array}{l} a)\ p^5 + 5p^4(1-p) + 8p^3(1-p)^2 + 2p^2(1-p)^3 \\ b)\ (1-p)^4 + 4p(1-p)^3 + 5p^2(1-p)^2 + p^3(1-p) \end{array}\right]$$

Se i rubinetti 1...5 sono aperti con probabilità p e che i rubinetti sono tra loro indipendenti, calcolare:

a) $\mathbb{P}$(fontana abbia l'acqua)

b) $\mathbb{P}$(fontana secca sapendo che il rubinetto 1 è chiuso

5) Un'urna contiene 112 dadi di cui 56 equilibrati e 56 truccati in modo che $\mathbb{P}(1) = 1/2$ e $\mathbb{P}(X) = 1/10$, $x \neq 1$.

a) Estraggo un dado a caso e lo lancio: quanto vale $\mathbb{P}(3)$?

b) Estraggo un dado a caso, lo lancio 2 volte e ottengo 2 e 3. Qual è la probabilità che sia un dado truccato? $[9/34]$

6) lancio 2 dadi:

P(somma>8) dato che è uscito almeno un 6.

Regole per il calcolo delle probabilità

Regola moltiplicativa

Se $A_1 \dots A_n$ sono eventi appartenenti allo stesso spazio e $\mathbb{P}(A_1 \cap A_2 \cap \dots A_n) \neq 0$

Allora
$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \times \mathbb{P}(A_2 | A_1) \times \mathbb{P}(A_3 | (A_1 \cap A_2)) \times \dots \times \mathbb{P}(A_n | (A_1 \cap \dots \cap A_{n-1})).$$

Dimostrazione: è un esperimento che avviene in n fasi:

Si verifica A_1 con probabilità $\mathbb{P}(A_1)$, poi si verifica A_2 (ma sappiamo che A_1 è già accaduto!) con prob $\mathbb{P}(A_2 | A_1)$, poi si verifica A_3 (ma sappiamo che $A_1 \cap A_2$ è vero!) con prob $\mathbb{P}(A_3 | (A_1 \cap A_2))$ e così via.

Esempio: Apprendimento

Si vuole insegnare ad un animale a svoltare a sinistra in un labirinto con un bivio a Y. Per facilitare l'apprendimento si premia se va a sinistra, si punisce se va a destra.
Al primo tentativo la probabilità che vada a sinistra è $\frac{1}{2}$. Ai tentativi successivi però viene modificata dall'apprendimento. Si ha allora:

- prob $p_1 > \frac{1}{2}$ che vada a sinistra dopo aver ricevuto un premio al tentativo precedente.
- prob $p_2 > p_1$ che vada a sinistra dopo essere stato punito al tentativo precedente.

Si calcoli:

a) la prob. che giri a destra nei primi 3 tentativi

b) " che svolti a destra al 3^o tentativo

Sol: a) Indico con A_i l'evento "svolta a destra al tentativo i-esimo".

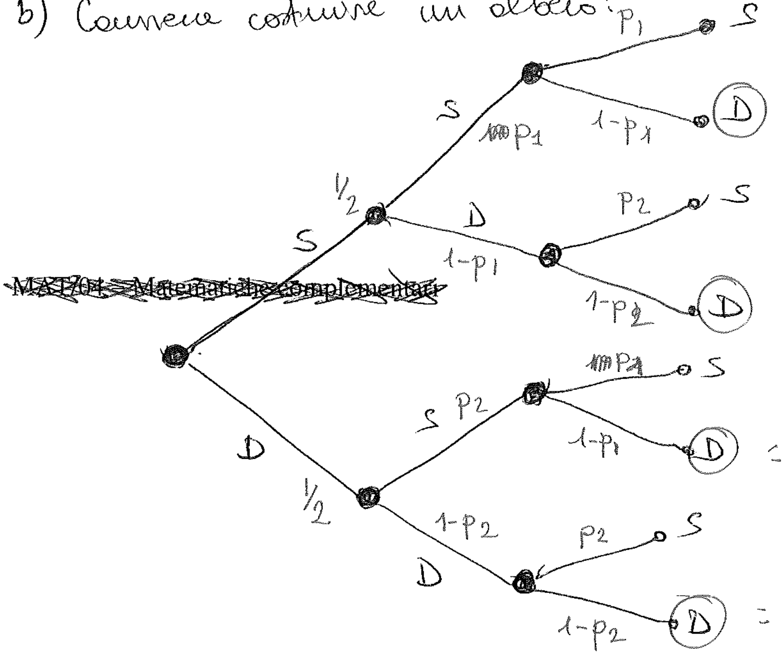
$$\mathbb{P}(A_1) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$\mathbb{P}(A_2 | A_1) = 1 - p_2$$

$$\mathbb{P}(A_3 | (A_1 \cap A_2)) = 1 - p_2$$

Quindi $\mathbb{P}(a)) = \frac{1}{2}(1-p_2)^2 = \mathbb{P}(A_1 \cap A_2 \cap A_3)$

b) Conviene costruire un albero:



$$= \frac{1}{2} p_1(1-p_1)$$

$$= \frac{1}{2}(1-p_1)(1-p_2)$$

$$= \frac{1}{2} p_2(1-p_1)$$

$$= \frac{1}{2}(1-p_2)^2$$

Quindi

$$\mathbb{P}(b)) = \frac{1}{2}\left(p_1(1-p_1) + (1-p_1)(1-p_2) + p_2(1-p_1) + (1-p_2)^2 \right) =$$

$$= \frac{1}{2}\left(\cancel{p_1} - p_1^2 + 1 - \cancel{p_1} - \cancel{p_2} + \cancel{p_1 p_2} + \cancel{p_2} - \cancel{p_1 p_2} + 1 + p_2^2 - 2p_2 \right) =$$

$$= \frac{1}{2}\left(p_2^2 - p_1^2 + 2 - 2p_2 \right) = 1 - p_2 + \frac{1}{2}p_2^2 - \frac{1}{2}p_1^2$$

Regola di addizione

Enunciamola nel caso di 2 o di 3 eventi. Siano $A_1 A_2 A_3$ eventi appartenenti allo stesso spazio. Allora:

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

$$\mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_2 \cap A_3) - \mathbb{P}(A_1 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3).$$

Osservazione: se $A_1 A_2 A_3$ sono incompatibili allora $\mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3)$ dato che $A_i \cap A_j = \emptyset$ se $i \neq j$.

Esempio: Viene fatto un controllo di qualità sulla produzione di un calzaturificio. I criteri sono:

A_1 = assenza di difetti nelle suole

A_2 = uniformità del colore

A_3 = cuciture regolari.

Le prob. dei 3 difetti sono 0,05; 0,03; 0,02 rispett.

Qual è la prob. che la produzione abbia almeno uno dei 3 difetti? $\mathbb{P}(A_1 \cup A_2 \cup A_3) = 0,097$

(Si faceva prima calcolandolo come $1 - \mathbb{P}(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3)$.)

Esempio: Quanti sono i numeri interi da 1 e 91 che non hanno divisori comuni con 91?

Esempio: E con 2013? $(2013 = 3 \times 11 \times 61)$ (numeri ≤ 2013)

Occhio a unioni e intersezioni

Indipendenza stocastica

<u>Definizione</u>: Due eventi A e B appartenenti allo stesso spazio degli eventi si dicono INDIPENDENTI (o STOCASTICAMENTE INDIPENDENTI, o STATISTICAMENTE INDIPENDENTI) se

~~MAT/06 - Probabilità e Statistica Matematica~~ vale l'uguaglianza

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$$

<u>Proposizione</u>: Le tre affermazioni seguenti sono equivalenti

1) $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$
2) $\mathbb{P}(A|B) = \mathbb{P}(A)$
3) $\mathbb{P}(B|A) = \mathbb{P}(B)$

Dimostrazione: $1) \rightarrow 2)$, $2) \rightarrow 3)$, $3) \rightarrow 1$ Facile.

<u>Esempio</u>: Consideriamo un esperimento consistente nel lancio di due dadi. Sia A = "la somma dei numeri è pari", B = "esce 6 sul primo dado". C = "la somma dei numeri usciti è 6".
Vediamo che A e B sono indipendenti mentre A e C, B e C non lo sono:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{3}{36}}{\frac{6}{36}} = \frac{1}{2} = \mathbb{P}(A) \Rightarrow \text{indipendenti}$$

$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(C)}{\mathbb{P}(C)} = 1 \neq \frac{1}{2} = \mathbb{P}(A) \Rightarrow \text{non indipendenti}$$

$$\mathbb{P}(C|B) = \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(B)} = 0 \neq \frac{5}{36} = \mathbb{P}(C) \Rightarrow \text{non indipendenti}$$

<u>Osservazione</u>: Non si confonda l'indipendenza con l'incompatibilità.
Si verifica abbastanza facilmente che se A e B si escludono mutuamente essi sono indipendenti se e solo se $\mathbb{P}(A)=0$ oppure $\mathbb{P}(B)=0$ (es: $\mathbb{P}(C|B)$).
Mentre se $\mathbb{P}(A) \neq 0$ e $\mathbb{P}(B) \neq 0$ l'indipendenza implica che essi non possono essere mutuamente esclusivi.

Esercizio: Abbiamo due urne che contengono palline colorate:

URNA 1 $= \{B_1, R_1, N_1\}$, URNA 2 $= \{B_2, R_2, N_2, G_2\}$.

Supponiamo che le estrazioni siano equiprobabili, cioè

$\mathbb{P}(B_1) = \mathbb{P}(R_1) = \mathbb{P}(N_1) = \frac{1}{3}$, $\mathbb{P}(B_2) = \mathbb{P}(R_2) = \mathbb{P}(N_2) = \mathbb{P}(G_2) = \frac{1}{4}$

a) Consideriamo gli eventi:

$E_1 = B_1 \cup R_1$, $H_1 = B_1 \cup N_1$. Si ha:

$\mathbb{P}(E_1)\mathbb{P}(H_1) = \frac{2}{3} \times \frac{2}{3} = \frac{4}{9}$ mentre $\mathbb{P}(E_1 \cap H_1) = \mathbb{P}(B_1) = \frac{1}{3}$.

Non sono indipendenti

b) Facciamo le stesse cose con la seconda urna:

$E_2 = B_2 \cup R_2$, $H_2 = B_2 \cup N_2$.

$\mathbb{P}(E_2)\mathbb{P}(H_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ mentre $\mathbb{P}(E_2 \cap H_2) = \mathbb{P}(B_2) = \frac{1}{4}$

Sono INDIPENDENTI!

L'indipendenza dipende dalla probabilità!

Definizione: Dati n eventi $A_1 \dots A_n$ appartenenti allo stesso spazio, si dice che $A_1 \dots A_n$ sono indipendenti se:

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(A_{i_k}) \quad \text{per } k = 1, 2, \dots n$$

NOTA: Se gli A_i sono a 2 a 2 indipendenti non significa che siano indipendenti complessivamente.

Esempio: lancio 2 dadi. Scelgo

$A_1 =$ "primo dado pari"

$A_2 =$ "secondo dado pari"

$A_3 =$ "somma dispari".

Si verifica A_1 e A_2 sono indipendenti. Idem per A_1 e A_3, e per A_2 e A_3.

D'altra parte però $\mathbb{P}(A_1 \cap A_2 \cap A_3) \neq \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)$ quindi non sono complessivamente indipendenti.

Distribuzioni di probabilità

VARIABILE CASUALE

Gli spazi campione possono essere insiemi di oggetti, di grandezze, di descrizioni, di proposizioni ecc ecc, non necessariamente trattabili con i comuni operatori matematici.
Nel momento in cui vogliamo analizzare dei fenomeni è opportuno rappresentare gli elementi degli spazi mediante numeri reali, stabilendo cioè una corrispondenza tra tutte le manifestazioni possibili del fenomeno casuale trattato e gli elementi di $\mathbb{R}$ (o $\mathbb{Z}$, o $\mathbb{N}$, o...)
In pratica dobbiamo costruire un modello matematico che rappresenti il sistema.

Definizione: Una VARIABILE CASUALE è definita come una funzione avente come dominio lo spazio dei campioni S e come codominio la retta reale con la condizione che tutti gli $s \in S$ tali che la loro immagine $X(s)$ sia minore di un certo $x \in \mathbb{R}$ sia un evento, cioè

$$A = \{ s : X(s) \leq x \ \forall x \in \mathbb{R} \}$$

sia un evento di S.

Esempio: Esperimento che consiste nel lancio di una moneta. Spazio campione $\{T, C\}$.
Se X è la variabile che indica il numero di esiti testa, allora:

$$X(s) = \begin{cases} 1 & \text{se } s = T \\ 0 & \text{se } s = C \end{cases}$$

Esempio: Esperimento che consiste nel lancio di due dadi.
Spazio campione $S = \{ (i,j) : i = 1, 2, \dots 6, \ j = 1, 2, \dots 6 \}$
Definiamo le variabili aleatorie X, Y, Z come segue:
X = la funzione casuale che associa all'esito del lancio la somma dei punteggi dei due dadi.

$$X((i,j)) = i + j \quad \forall (i,j) \in S.$$

Il codominio di X sono i numeri $\{2, 3, 4, \dots, 11, 12\}$ ovvero che assume i valori $2, 3, \dots, 11, 12$.

Y = la funzione casuale che associa ad ogni risultato il massimo tra i due ~~valori~~ punteggi dei dadi:

$$Y((i,j)) = \max(i,j) \quad \forall (i,j) \in S$$

Il codominio Y è costituito dall'insieme dei numeri $\{1, 2, 3, 4, 5, 6\}$.

Z = la funzione che associa ad ogni risultato dell'esperimento il valore assoluto della differenza tra i due punteggi apparsi:

$$Z((i,j)) = |i-j| \quad \forall (i,j) \in S$$

Il codominio di Z è costituito dai numeri $\{1,2,3,4,5\}$.

Nel caso in cui il codominio è composto da un numero finito di valori si dice che la variabile casuale è DISCRETA, così come quando il codominio è composto da una infinità numerabile di dati.
In caso contrario, si dice CONTINUA se può assumere tutti i valori di un intervallo di $\mathbb{R}$.

<u>Osservazione</u>: Indicheremo le variabili casuali con le lettere maiuscole, i loro valori con le minuscole.

<u>Funzione di distribuzione cumulativa</u>

<u>Definizione</u>: Data una variabile casuale X, si dice FUNZIONE DI DISTRIBUZIONE CUMULATIVA $F_X(x)$ quella funzione che ha per dominio l'asse reale e per codominio l'intervallo chiuso $[0,1]$ definita in questo modo:

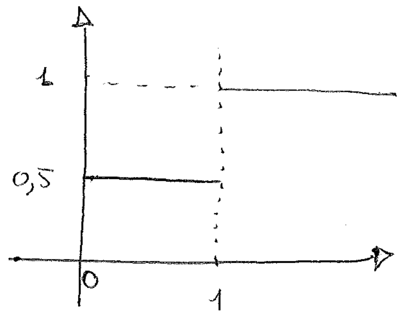
$$F_X(x) = \mathbb{P}[X \leq x] = \mathbb{P}\{s : X(s) \leq x\}$$

La distribuzione cumulativa spiega come sono distribuiti i valori della variabile casuale. Il termine "cumulativa" viene spesso omesso. La distribuzione cumulativa è individuata in modo univoco dalla variabile casuale.

<u>Esempio</u>: Lancio di una moneta, vedi sopra:

$$X = \begin{cases} 0 & \text{se } s = \text{Testa} \\ 1 & \text{se } s = \text{Croce} \end{cases}$$

$$F_X(x) = \begin{cases} 0{,}5 & \text{per } 0 \leq x < 1 & = \mathbb{P}(0) \\ 1 & \text{per } x \geq 1 & = \mathbb{P}(0) + \mathbb{P}(1) \\ 0 & \text{per } x < 0 & = \mathbb{P}(\emptyset) \end{cases}$$



Esempio: Sia $X((i,j)) = i+j$ dell'esempio precedente.

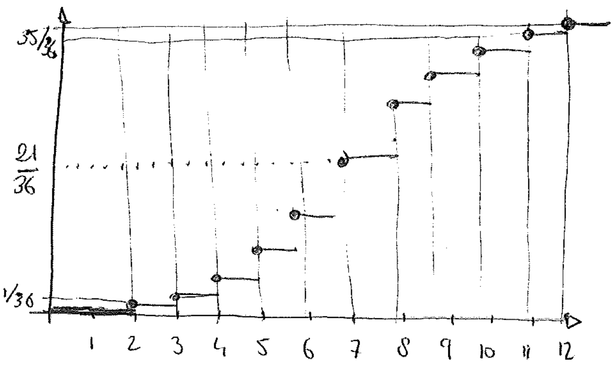
$\mathbb{P}(i+j=2) = \frac{1}{36}$ $\mathbb{P}(i+j=7) = \frac{6}{36}$ $\mathbb{P}(i+j=12) = \frac{1}{36}$

$\mathbb{P}(i+j=3) = \frac{2}{36}$ $\mathbb{P}(i+j=8) = \frac{5}{36}$

$\mathbb{P}(i+j=4) = \frac{3}{36}$ $\mathbb{P}(i+j=9) = \frac{4}{36}$

$\mathbb{P}(i+j=5) = \frac{4}{36}$ $\mathbb{P}(i+j=10) = \frac{3}{36}$

$\mathbb{P}(i+j=6) = \frac{5}{36}$ $\mathbb{P}(i+j=11) = \frac{2}{36}$



I valori di $F_X(x)$ sono quindi:

$\frac{1}{36}, \frac{3}{36}, \frac{6}{36}, \frac{10}{36}, \frac{15}{36}, \frac{21}{36}, \frac{26}{36}, \frac{30}{36}, \frac{33}{36}, \frac{35}{36}, 1$

Esercizio: Scrivere la distribuzione cumulativa per le variabili Y e Z.

Proprietà delle funzioni di densità

i) $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to \infty} F_X(x) = 1$

ii) $F_X(x)$ è una funzione monotona non decrescente, cioè, $\forall x_1, x_2 : x_1 < x_2$ si ha $F_X(x_1) \leq F_X(x_2)$.

iii) $F_X(x)$ è continua a destra (cioè $\lim_{h \to 0^+} F_X(x+h) = F_X(x)$).

La i) è ovvia. Dimostriamo la ii): consideriamo l'evento $\{s \in S : X(s) \leq x_2\}$ (più brevemente $\{X \leq x_2\}$). Possiamo scrivere:

$$\{X \leq x_2\} = \{X \leq x_1\} \cup \{x_1 < X \leq x_2\} \quad \text{(mutuamente esclusivi).}$$

Ma allora $F_X(x_2) = \mathbb{P}(\{X \leq x_2\}) = \mathbb{P}(\{X \leq x_1\}) + \mathbb{P}(\{x_1 < X \leq x_2\}) =$

$= F_X(x_1) + \mathbb{P}(\{x_1 < X \leq x_2\})$

Poiché $\mathbb{P}(\{x_1 < X \leq x_2\}) \geq 0$ segue la ii)

La iii) è una definizione diretta della definizione di distribuzione cumulativa e dell'assioma 5.

Funzioni di densità discrete

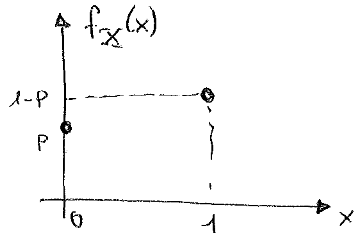
Definizione: Sia X una variabile aleatoria discreta il cui codominio sia $\{x_1, \dots, x_n, \dots\} \subset \mathbb{R}$

la funzione
$$f_X(x) = \begin{cases} \mathbb{P}(X = x_j) & \text{se } x = x_j,\ j = 1, 2, \dots, n, \dots \\ 0 & \text{altrimenti} \end{cases}$$

viene detta FUNZIONE DI DENSITÀ DISCRETA DI X (o funzione di probabilità).

Esempio: lancio moneta. In generale (moneta anche non bilanciata) avremo $\mathbb{P}(T) = p$, $\mathbb{P}(C) = 1-p$. Supponiamo che X conti il numero di teste (ovvero $T \rightsquigarrow 1$, $C \rightsquigarrow 0$). Allora

$$f_X(x) = \begin{cases} 1-p & \text{se } x = 0 \\ p & \text{se } x = 1 \\ 0 & \text{altrimenti} \end{cases}$$



Esempio: Esperimento casuale dato dal lancio di un dado.
Sia X la variabile che indica il numero che appare sulla faccia rivolta in alto.
La corrispondente funzione di densità discreta è $f_X(x) = \begin{cases} 1/6 & \text{se } x = 1, 2, \dots, 6 \\ 0 & \text{altrimenti} \end{cases}$

I punti in cui $f_X \neq 0$ sono detti "punti massa"; la $f_X(x)$ prende anche il nome di funzione di probabilità di massa.

Esercizio: Scrivere la funzione di densità discreta per le variabili X, Y, Z definite la volta scorsa:
$$X(i,j) = i+j, \quad Y(i,j) = \max\{i,j\}, \quad Z(i,j) = |i-j|.$$

Proprietà della funzione di densità discreta

(i) $f_X(x) \geq 0 \quad \forall x \in \mathbb{R} \quad (f_X(x_i) \neq 0\ \forall i,\ f_X(x) = 0\ \forall x \neq x_1 \dots x_n \dots)$

(ii) $\sum_j f_X(x_j) = 1$

Vale inoltre un'importante legame tra la $f_X(x)$ e la $F_X(x)$:

<u>Teorema</u>: Sia X una variabile aleatoria e siano $f_X(x)$ e $F_X(x)$ la densità discreta e la distribuzione cumulativa di X. Allora:

$$F_X(x) = \sum_{j:\, x_j \le x} f_X(x_j)$$

$$f_X(x) = \begin{cases} F_X(x_j) - \lim_{h \to 0^-} F_X(x_j - h) & \text{se } x = x_j \quad j = 1, 2, \dots, n, \dots \\ 0 & \text{altrimenti} \end{cases}$$

dove $x_1 \dots x_n \dots$ sono i punti massa.

Come conseguenza del teorema vale la seguente uguaglianza, utilissima quando si devono fare dei calcoli:

$$\mathbb{P}(a < x \le b) = \sum_{j:\, a < x_j \le b} f_X(x_j) = F_X(b) - F_X(a) \qquad (a < b)$$

Funzione di densità di probabilità

<u>Definizione</u>: Sia X una variabile aleatoria continua. La funzione $f_X(x)$ tale che

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \mathbb{P}\{X \le x\}$$

viene detta funzione di densità di probabilità di X

Proprietà delle densità di probabilità

(Analogamente al caso discreto)

(i) $f_X(x) \ge 0 \quad \forall x \in \mathbb{R}$

(ii) $\displaystyle\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

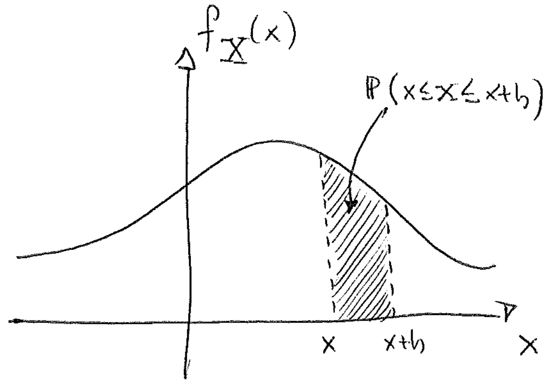
Continua anche a valere la formula $\mathbb{P}(a < x \le b) = F_X(b) - F_X(a) = \displaystyle\int_a^b f_X(t)\,dt$

OSSERVAZIONE

Dal teorema fondamentale del calcolo integrale segue $f_X(x) = \dfrac{d}{dx} F_X(x)$

NEI PUNTI IN CUI F_X è DERIVABILE

L'area sotto il grafico di $f_X(x)$ si può interpretare come la probabilità che X sia compresa tra x e $x+h$.



Osservazione

$f_X(x)$ NON È UNIVOCAMENTE DETERMINATA. Esempio:

$$F_X = \begin{cases} 0 & \text{se } x \leq a \\ \frac{1}{b-a}(x-a) & \text{se } a \leq x \leq b \\ 1 & \text{se } x > b \end{cases}$$

Si verifica facilmente che $f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{altrove} \end{cases}$ soddisfa $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$

Ma anche

$$g_X(x) = \begin{cases} \frac{1}{b-a} & \text{se } a < x < \frac{b-a}{2} \\ 3 & \text{se } x = \frac{b-a}{2} \\ \frac{1}{b-a} & \text{se } \frac{b-a}{2} \leq x < b \\ 0 & \text{altrove} \end{cases}$$

soddisfa la stessa relazione.
Questo perché modificando una funzione in un numero finito di punti o in una infinità numerabile di punti non viene modificato il valore dell'integrale.
In ogni caso, dato che la probabilità è determinata dall'integrale di f si può parlare di unicità della probabilità.

Valori attesi

Nella pratica, quando si tratta di costruire un modello matematico per un certo fenomeno si considerano delle densità di probabilità dipendenti da uno o più parametri, per poi sceglierne il valore o i valori in modo che il modello meglio si adatti al fenomeno da studiare.

Molto spesso i parametri sono quantità che hanno un significato statistico quali ad esempio medie, varianze, scarto quadratico ecc ecc.

Vediamo che significato si dà a queste parole nel caso delle variabili aleatorie:

Media

Definizione: Si definisce media o valore atteso della variabile casuale X la funzione:

$$E(X) = \sum_{x_j} x_j f_X(x_j) \qquad \text{se } X \text{ è discreta}$$

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx \qquad \text{se } X \text{ è continua con funzione di densità } f_X(x).$$

$E(X)$ viene anche chiamata "Speranza" matematica, da cui il simbolo E (expectation). Nelle formule si indica spesso con μ_X.

Si osservi che $E(X)$ non individua il valore più probabile di X anzi non è neppure detto che $E(X)$ sia un valore effettivamente assunto.

Può accadere che la serie o l'integrale che definiscono $E(X)$ non siano convergenti. In questo caso si dice che $E(X)$ non esiste.

Osservazione: La definizione di $E(X)$ richiama strettamente quella di baricentro di un sistema materiale, a masse concentrate o distribuite.

Es: nel caso discreto, il baricentro di un sistema di masse complessiva 1 è dato da:

$$x_B = \frac{\sum_{x_j} x_j f_X(x_j)}{\sum_{x_j} f_X(x_j)} = \sum_{x_j} x_j f_X(x_j) = E[X]$$

dato che $\sum_{x_j} x_j = 1$.

I valori di f_X giocano il ruolo delle masse mentre x_j sono le posizioni in cui si trovano le masse.

Se non conosciamo la densità di probabilità, la media può essere calcolata anche a partire dalla $F_X(x)$ mediante la formula:

$$E(X) = \int_0^{\infty} (1 - F_X(x))\,dx - \int_{-\infty}^{0} F_X(x)\,dx$$

Esempio: lancio di un dado. X indica il numero uscito.

Si ha: $$E(X) = \sum_{j=1}^{6} j\, f_X(j) = \sum_{j=1}^{6} j\, \mathbb{P}(X = j) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} =$$

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

Esempio: Calcolare la speranza matematica per le variabili casuali X, Y, Z introdotte precedentemente, a partire dal lancio di due dadi.

$$X = i + j, \quad E(X) = \sum_{i=2}^{12} i\, f_X(i) = \sum_{i=2}^{12} i\, \mathbb{P}[X = j] = 2 \times \frac{1}{36} + 3 \times \frac{2}{3} + \dots + 12\, \frac{1}{36} = 7$$

$$Y = \max\{i,j\}, \quad E(Y) = \sum_{i=1}^{6} i\, f_Y(i) = 1 \times \frac{1}{36} + 2 \times \frac{3}{36} + \dots + 6 \times \frac{11}{36} = \frac{161}{36} \simeq 4{,}47$$

$$Z = |i - j|, \quad E(Z) = \sum_{i=0}^{5} i\, f_Z(i) = \sum_{i=0}^{5} i\, \mathbb{P}[X = j] = 0 \cdot \frac{6}{36} + 1 \cdot \frac{10}{36} + \dots + 5\, \frac{2}{36} = \frac{70}{36} = 1{,}94$$

Esempio: Sia X la variabile casuale la cui densità di prob. è:

$$f_X(x) = \begin{cases} 1/x^2, & \text{per } 1 < x < \infty \\ 0, & \text{altrove.} \end{cases}$$ Allora:

$$E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_{1}^{\infty} \frac{x}{x^2}\, dx = \ln x \Big|_1^{\infty} = +\infty \qquad E(x) \text{ non esiste.}$$

Moda

Si determina facilmente a partire dalla densità $f_X(x)$

Nel caso di densità continue corrisponde alla x : $f_X(x)$ è massimo.

Varianza

<u>Definizione</u>: Data una variabile casuale X con media μ_X, si definisce varianza di X la funzione:

$$\mathrm{var}(X) = \sum_{x_j} (x_j - \mu_X)^2 f_X(x_j) \qquad \text{se } X \text{ è discreta con punti massa } x_1, \dots x_n, \dots$$

$$\mathrm{var}(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx \qquad \text{se } X \text{ è continua con densità } f_X(x).$$

<u>Definizione</u>: Si definisce deviazione standard o scarto quadratico medio la quantità $\sigma_X = \sqrt{\mathrm{var}(X)}$

<u>Osservazione</u>: Così come c'è una similitudine tra media e baricentro, la varianza si può assimilare al momento d'Inerzia.

<u>Esercizio</u>: Calcolare la varianza per le solite X, Y, Z relative al lancio dei dadi.

$$\mathrm{var}(X) = \frac{210}{36} = 5.83, \quad \mathrm{var}(Y) = \frac{91980}{1296} = 1.97, \quad \mathrm{var}(Z) = \frac{95760}{1296} = 2.05$$

<u>Esercizio</u>: Calcolare speranza matematica e varianza della variabile casuale X con densità cumulativa $F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{se } x < 0 \end{cases}$.

(Si ha $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{se } x < 0 \end{cases}$, $E(X) = \dfrac{1}{\lambda}$, $\mathrm{Var}(X) = \dfrac{1}{\lambda^2}$

<u>Osservazione</u>: Anche la varianza può essere calcolata a partire dalla distribuzione cumulativa:

$$\mathrm{var}(X) = \int_0^{\infty} 2x \left[1 - F_X(x) + F_X(-x)\right] dx - \mu_X^2$$

~~Esercizio: Calcolare la varianza della variabile casuale~~

Valore atteso di funzioni di una variabile casuale

Spesso è utile fare uso di trasformazioni casuali ottenute a partire da una funzione g della variabile casuale X, con $g:\mathbb{R}\to\mathbb{R}$. e spesso si pone il problema di calcolare il valor medio di $g(X)$.

<u>Definizione</u>: Si definisce valore atteso di $g(X)$ la quantità:

$$E(g(X)) = \sum_{x_j} g(x_j) f_X(x_j)$$ con X discreta, x_j punti massa

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$ con X continua di densità $f_X(x)$.

Si osservi che media e varianza non sono altro che due scelte particolari di $g(x)$: se $g(x)=x$ si ottiene la media, se $g(x) = (x-\mu_X)^2$ la varianza.

<u>Proprietà del valore atteso</u>:

(i) $E(c) = c$

(ii) $E(cg(X)) = cE(g(X))$

(iii) $E(c_1g_1(X)+c_2g_2(X)) = c_1E(g_1(X)) + c_2E(g_2(X))$

(iv) $E(g_1(X)) \leq E(g_2(X))$ se $g_1(x) \leq g_2(x)$ $\forall x\in\mathbb{R}$

(dimostrazioni: provate a farle)

<u>Proprietà della varianza</u>:

Valgono le proprietà già viste per gli insiemi di dati:

(v) $\mathrm{var}(X) = E(X^2) - E(X)^2$ (se $E(X)$ esiste!)

(vi) $\mathrm{var}(cX) = c^2\mathrm{var}(X)$

DISTRIBUZIONI DI PROBABILITÀ

A seconda del fenomeno che si intende analizzare, utilizzeremo un modello matematico piuttosto che un'altro.
Nel seguito presenteremo un catalogo di distribuzioni di probabilità insieme alle loro tipiche situazioni di utilizzo.

DISTRIBUZIONE BINOMIALE (o di BERNOULLI) Jakob, 1654-1705

Consideriamo una serie di n eventi indipendenti ed equiprobabili (es: lancio di una moneta truccata) con $\mathbb{P}(E_i)=p$ e $\mathbb{P}(E_j^c)=1-p=q$.

Sia $E_{i_1,i_2,\dots,i_k} = E_{i_1}\cap E_{i_2}\cap\dots\cap E_{i_k}\cap E^c_{i_{k+1}}\cap\dots\cap E^c_n$. Allora $\mathbb{P}(E_{i_1,\dots,i_n}) = p^k q^{n-k}$

Sia ora $E=\{$si verificano k degli n eventi$\}$. $\mathbb{P}(E)=\binom{n}{k}p^kq^{n-k}$.

Indichiamo infine con X la variabile aleatoria che conta quanti eventi tra gli n si verificano. Al variare di $k=0,1,\dots,n$ si ha:

$$\mathbb{P}(E)=\mathbb{P}(X=k)=f_X(k)=\binom{n}{k}p^kq^{n-k}.$$

Si vede facilmente che $f_X(k)\geq 0$ e che $\sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k}=1$ (è $(p+q)^n=1^n$)

La funzione $F_X(k)$ si dice si dice DISTRIBUZIONE BINOMIALE (o DI BERNOULLI)

Esempio: $\mathbb{P}$(almeno k successi su n prove), $= \sum_{j=0}^{k}\binom{n}{j}p^jq^{n-j}$
$\mathbb{P}$(al più k successi su n prove) $= \sum_{j=k}^{n}\binom{n}{j}=p^jq^{n-j}$

Esempio: Se $\alpha\in(0,1)$ determinare n tale che
$\mathbb{P}$(almeno 1 successo su n prove) $>\alpha$
Se ad esempio $p=0.25$, $\alpha=0.90$: $\mathbb{P}$(almeno 1 succ) $=1-\mathbb{P}$(nessun successo).
$\mathbb{P}$(nessun successo) $=\binom{n}{0}p^0q^n=(1-p)^n$. Cerchiamo n tale che
$(1-p)^n<1-\alpha$.
Segue $e^{n\log(1-p)}<e^{\log(1-\alpha)} \Rightarrow n\log(1-p)<1-\alpha \Rightarrow n=\dfrac{1-\alpha}{\log(1-p)}$
Nel nostro caso $n=\dfrac{0.10}{\log(0,75)}$. Se $n>8$ la condizione è soddisfatta.

Esempio: In un'urna ci sono N palline di cui r sono rosse e $N-r$ sono verdi.

Si fanno n estrazioni con restituzione.

Sia $E_i = \{\text{la i-esima estrazione è rossa}\}$. $\mathbb{P}(E_i) = \frac{r}{N} = p$.

Gli eventi sono equiprobabili e indipendenti.

Il numero di palline estratte ha distribuzione binomiale

Esempio: Stessa situazione SENZA restituzione.

Si ha ancora $\mathbb{P}(E_i) = \frac{r}{N}$, $i = 0 \dots n$, infatti:

$$\mathbb{P}(E_2) = \mathbb{P}(E_1)\mathbb{P}(E_2|E_1) + \mathbb{P}(E_1^c)\mathbb{P}(E_2|E_1^c) = p\frac{r-1}{N-1} + (1-p)\frac{r}{N-1} =$$

$$= \frac{1}{N-1}(\cancel{pr} - p + r - \cancel{pr}) = \frac{1}{N-1}\left(r - \frac{r}{N}\right) = \frac{(\cancel{N-1})\,r}{N(\cancel{N-1})} = \frac{r}{N} = p$$

Gli eventi però NON sono indipendenti dato che

$$\mathbb{P}(E_2|E_1) = \frac{r-1}{N-1} = \frac{Np-1}{N-1} \neq p = \mathbb{P}(E_2)$$

DISTRIBUZIONE IPERGEOMETRICA

È la distribuzione adatta per le estrazioni senza restituzione.

Se $Y = \{\text{n° di successi su } n \text{ prove}\}$ in una sequenza di estrazioni senza restituzione.

Allora:
$$\mathbb{P}(Y=k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \quad \text{dove } k \in [\max\{0, n-N+r\}, \min\{r,n\}]$$

(r = n° palle rosse in un'urna di N palle).

Nota: Tenendo fissi N e $p = \frac{r}{N}$ si può dimostrare che

$\lim_{r \to \infty} \mathbb{P}(Y=k) = \mathbb{P}(X=k)$ cioè "quando ci sono tante palle l'urna non si accorge della restituzione".

Si ha $E(X) = E(Y) = np$.

Per la varianza bisogna ricordare che $Var(U+V) = Var(U) + Var(V) + 2Cov(U,V)$.
(dimostrarlo per esercizio).

$Var(Y) = Var\left(\sum_{k=1}^{n} Z_k\right)$ (Z_k = var. aleatoria esito di 1 singola estrazione).

$$Var(Y) = \sum_{k=1}^{n} Var(Z_k) + 2\sum_{\substack{k=1\dots n\\ j=1\dots n\\ k\neq j}} Cov(Z_k, Z_j)$$

Si ha $Var(Z_k) = pq$ (verificatelo!)

$Cov(Z_k, Z_j) = p\frac{r-1}{N-1} - p^2$ (idem!)

Segue
$$Var(Y) = npq + 2\underbrace{\left[\frac{n(n-1)}{2}\right]}_{\substack{\text{n° coppie}\\ \text{distinte} = \binom{n}{2}}}\left(p\frac{r-1}{N-1} - p^2\right) = \dots npq\left(1 - \frac{n-1}{N-1}\right)$$

Osservazione: $Var(Y)$ decresce al crescere di n. Se $n = N$ abbiamo la certezza che i successi siano r e quindi $Var = 0$.
In generale $Var(X) \leq Var(Y)$.

Esercizio: Da un insieme di 50 oggetti buoni e 5 difettosi se ne estraggono 10. Calcolare $\mathbb{P}(A)$ dove $A = \{2 \text{ dei } 10 \text{ sono difettosi}\}$.
Si ha: $n = 10$, $N = 55$, $r = 5$, $k = 2$ quindi: $\mathbb{P}(A) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} = \frac{\binom{5}{2}\binom{50}{8}}{\binom{55}{10}} =$
$= \simeq 0.1836$

Potremo usare la distribuzione binomiale per approssimare l'ipergeometrica (considerando 55 "grande"). Allora $p = 5/55 \simeq 0.0909$
e $\mathbb{P}(A) \simeq \binom{10}{2} p^2 q^8 = 45 \cdot \left(\frac{1}{11}\right)^2 \left(\frac{10}{11}\right)^8 \simeq 0.1735$

La distribuzione ipergeometrica è utilizzata nei rilevamenti statistici tipo cattura/liberazione di animali

Esercizio: Prendo le 12 figure da un mazzo da 52 carte. Sia X la variabile aleatoria che conta il numero di re estratti estraendo 5 carte senza reimm.
$$\left(\mathbb{P}(X = r) = \frac{\binom{4}{r}\binom{8}{5-r}}{\binom{12}{5}} \right) \quad r = 0, 1, 2, 3, 4.$$
Disegnare densità di probabilità e distribuzione cumulativa.

DISTRIBUZIONE DI POISSON (1781-1840)

È utilizzata quando si parla di eventi "rari" (es: errori di stampa in un testo, ultracentenari in una popolazione ecc...)

Definizione: Sia $\lambda \in \mathbb{R}$, $\lambda \geq 0$ e sia X_λ una variabile aleatoria tale che
$$\mathbb{P}(X_\lambda = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad (x \in \mathbb{N}).$$
Si dice che X_λ segue la distribuzione di Poisson.
In altre parole $f_{X_\lambda}(x) = \frac{\lambda^x e^{-\lambda}}{x!}$.

Osservazione: Si ha che $f_{X_\lambda}(x) \geq 0 \ \forall x \in \mathbb{N}$ e $\sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1$
Quindi f_{X_λ} è effettivamente una densità di probabilità.

Osservazione: La distribuzione di Poisson può essere considerata come il limite per $N \to \infty$ della distribuzione binomiale, mantenendo $np = \lambda$ costante.

Esercizio: Calcolare $E(X_\lambda) = \lambda$.

Esempio: Il n° di incidenti su una certa strada tra le 16 e le 20 segue la distribuzione di Poisson con media 2.
Qual'è la probabilità che in un giorno non ci siano incidenti?

Soluzione: $\mathbb{P}(X_2 = 0) = \dfrac{2^0 e^{-2}}{0!} = e^{-2} \simeq 0{,}1353$

DISTRIBUZIONE GEOMETRICA

È utilizzata quando si tratta di descrivere il numero di prove eseguite fino al 1° successo in una successione di eventi indipendenti ed equiprobabili.
Se X è una variabile aleatoria che conta le prove fino al primo successo,

$$f_X = \mathbb{P}(X = x) = (1-p)^{x-1} p \quad (x-1 \text{ insuccessi} + 1 \text{ successo, nell'ordine})$$

dove $p = \mathbb{P}(\text{successo})$. $(x \in \mathbb{N})$.

Esercizio: Verificare che $\lim\limits_{x\to\infty} F_X(x) = 1$ (cioè $\sum\limits_{x=1}^{\infty} f_X(x) = 1$).

Calcolare $E(X) = \dfrac{1}{p}$.

Osservazione: Se X è una variabile aleatoria con distribuzione geometrica allora $\mathbb{P}((X>k+n) \mid (X>k)) = \mathbb{P}(X>n)$

In altre parole l'attesa del 1° successo non dipende (e non deve!) dal tempo già trascorso. (cfr. numeri ritardatari al lotto).

$$\mathbb{P}((X>k+n) \mid (X>k)) = \frac{\mathbb{P}((X>k+n) \cap (X>k))}{\mathbb{P}(X>k)} = \frac{\mathbb{P}(X>k+n)}{\mathbb{P}(X>k)} =$$

$$= \frac{\sum\limits_{x=k+n+1}^{\infty} p(1-p)^{x-1}}{\sum\limits_{x=k+1}^{\infty} p(1-p)^{x-1}} = \frac{p}{p} \cdot \frac{\sum\limits_{x=k+1}^{\infty} (1-p)^{x-1}}{\sum\limits_{x=k+1}^{\infty} (1-p)^{x-1}} (1-p)^n = \boxed{(1-p)^n}$$

D'altra parte $\mathbb{P}(X>n) = \sum\limits_{x=n+1}^{\infty} p(1-p)^{x-1} = p(1-p)^{n} \sum\limits_{x=1}^{\infty} (1-p)^{x-1} = p(1-p)^n \dfrac{1}{1-(1-p)} = \boxed{(1-p)^n}$

Esempio: Urna con 100 palline di cui 20 bianche.
Sia X = n° di estrazioni necessarie per avere la prima bianca.
Calcolare $E(X)$ e $\mathbb{P}(X \geq 3)$.

Soluzione: $E(X) = \frac{1}{p}$ e $p = \frac{20}{100} = \frac{1}{5}$. Segue $E(X) = 5$

$$\mathbb{P}(X \geq 3) = \sum_{x=3}^{\infty} p(1-p)^{x-1} = p(1-p)^2 \sum_{x=0}^{\infty} (1-p)^x = p(1-p)^2 \frac{1}{1-(1-p)} = (1-p)^2 = \frac{16}{25}$$

DISTRIBUZIONI CONTINUE

DISTRIBUZIONE UNIFORME

Una distribuzione uniforme su un intervallo $[a,b]$ ha densità data da:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{se } x \in [a,b] \\ 0 & \text{altrove} \end{cases}.$$

$$E(X) = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a} \left.\frac{x^2}{2}\right|_a^b = \frac{b^2-a^2}{b-a} \cdot \frac{1}{2} = \frac{a+b}{2}$$

$$Var(X) = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a}\, dx \quad \text{oppure} \quad = E(X^2) - E(X)^2. \quad = \frac{(b-a)^2}{12}$$

Osservazione: $Var(X)$ dipende solo da $b-a$ e non da a e b separatamente e conferma che $Var(X)$ è invariante per traslazione.

Esercizio: Determinare la densità di probabilità di X^2.
In ogni caso $E(X^2) = \int_a^b g(x) f_X(x)\, dx$ con $g(x) = x^2$.

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a}\, dx = \frac{b^3-a^3}{b-a} \frac{1}{3} = \frac{1}{3}(b^2+ab+b^2) \quad \text{da cui}$$

$$E(X^2) - E(X)^2 = \frac{1}{3}(b^2+ab+b^2) - \frac{a^2+2ab+b^2}{4} = \frac{4b^2+4ab+4b^2-3a^2-3b^2-6ab}{12} = \frac{a^2-2ab+b^2}{12}.$$

Esempio: Una resistenza R ha valore incognito distribuito in modo uniforme tra 850 e 1000 ohm.
Calcolare funzione di ripartizione e densità della Conduttanza $\left(\frac{1}{R}\right) = Z$.

Si ha: $F_R(x) = \mathbb{P}(R \leq x) = \frac{x-850}{150}$ se $x \in [850, 1000]$. Di conseguenza:

$$F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}\left(\frac{1}{R} \leq z\right) = \mathbb{P}\left(R \geq \frac{1}{z}\right) = 1 - \mathbb{P}\left(R \leq \frac{1}{z}\right) = 1 - \frac{1/z - 850}{150} = \frac{20}{3} - \frac{1}{150z}.$$

Infine: $f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{1}{150 z^2}$, $z \in \left[\frac{1}{1000}, \frac{1}{850}\right]$.

DISTRIBUZIONE ESPONENZIALE

Ha come densità di probabilità:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Abbiamo già visto che se X ha distribuzione esponenziale, $E(X) = \lambda^{-1}$, $Var(X) = \lambda^{-2}$

È la versione continua della distribuzione di Poisson.

Tipica per il tempo di attesa di un evento ($n°$ prove $\sim$ discreto, tempo $\sim$ continuo).

Esercizio: ~~Esempio~~: Sia X il tempo di attesa fino al 1° guasto di un certo dispositivo, X segue la distribuzione esponenziale.
Determinare la distribuzione cumulativa di X e verificare che anche in questo caso

$$\mathbb{P}((X \geq s+t) \mid (X \geq s)) = \mathbb{P}(X \geq t).$$

DISTRIBUZIONE NORMALE

Si parla di distribuzione normale (o gaussiana, o degli errori) se X ha densità data da

$$N_{0,1} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad \text{normale standard}$$

Più in generale,

$$N_{m,\sigma} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Si ha che $\int_{-\infty}^{\infty} N_{m,\sigma}(x)\,dx = 1$ cioè N è effettivamente una densità di prob.

Inoltre $E(X) = \int_{-\infty}^{\infty} x\, N_{m,\sigma}(x)\,dx = m$, $\quad Var(X) = \int_{-\infty}^{\infty} (x-m)^2 N_{m,\sigma}(x)\,dx = \sigma$

Si osservi infine che $N(x) = N(-x)$.

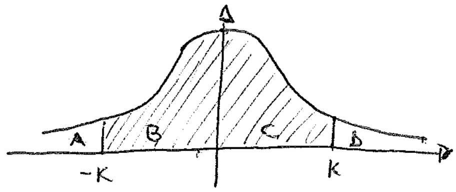
Poiché $N_{m,\sigma}(x) = \frac{1}{\sigma} N_{0,1}(x-m)$ si possono dedurre le proprietà di $N_{m,\sigma}$ a partire da quelle di $N_{0,1}$.

Solitamente la distribuzione cumulativa di $N_{0,1}$ si indica con Φ:

$$\Phi(x) = \int_{-\infty}^{x} N(s)\,ds$$

Osservazione: $\mathbb{P}\{|X_{m,\sigma}-m|\le k\sigma\} = \mathbb{P}\{|X_{0,1}|\le k\} = 2\Phi(k)-1$

Infatti: $\Phi(k) = A+B+C$



ma anche, per simmetria,

$\Phi(k) = B+C+D$

D'altra parte $1 = A+B+C+D$.

A noi interessa $B+C = \underbrace{A+\boxed{B+C}}_{\Phi(k)} + \underbrace{(B+C)+D}_{\Phi(k)} - \underbrace{(A+B+C+D)}_{1}$

Osservazione: $\phi(-k)+\phi(k)=1$

Infatti: $\phi(k)=A+B+C=B+C+D$, $\phi(-k)=A \Rightarrow \phi(k)+\phi(-k)=1$.

Osservazione: Per conoscere tutti i valori delle distribuzioni ~~gaussiane~~ cumulative delle gaussiane è sufficiente conoscere

1) soltanto i valori di $\phi(x)$ densità cumulativa di $N_{0,1}$.
2) conoscerle soltanto per $x<0$ o ~~per~~ solo per $x>0$ dato che $\Phi(-x)=1-\Phi(x)$

A questo scopo ci sono delle apposite tavole numeriche.

Ad esempio si ha:

$$\mathbb{P}\{|X_{m,\sigma}-m|\le\sigma\}=0,6826$$

$$\mathbb{P}\{|X_{m,\sigma}-m|\le 3\sigma\}=0,9974$$

Esercizio: Carlo abita in periferia. Per andare al lavoro può seguire 3 strade:

X_C 1) In macchina via centro: 28 min con scarto di 7'30''.
X_T 2) In macchina via tangenziale: 32 min con scarto di 4'.20''.
X_P 3) In macchina da paese vicino: 33 min con scarto di 1'10''.

Quale strada conviene prendere ~~per~~ se si vuole arrivare entro 30 min?

Siano X_C, X_P, X_T le 3 variabili aleatorie relative ai 3 itinerari.

$$\mathbb{P}(X_C<30) = \Phi\left(\frac{30-m}{\sigma}\right) = \Phi\left(\frac{30-28}{7,5}\right) = \Phi(0,27) = 0.6064.$$

$$\mathbb{P}(X_T<30) = \Phi\left(\frac{30-32}{4.333}\right) = \Phi(-0,46) = 0.3228$$

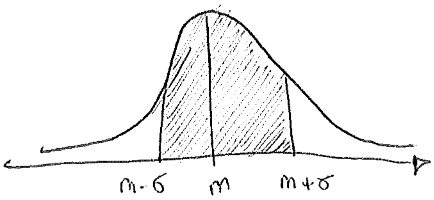
$$\mathbb{P}(X_P<30) = \Phi\left(\frac{30-33}{1,167}\right) = \Phi(-2,286) = 0,005$$

Quindi conviene passare per il centro.

Esercizio:

1) $N_{m,\sigma} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$ ha moda m ed ha due flessi obliqui in $x = m+\sigma$ e $m-\sigma$.

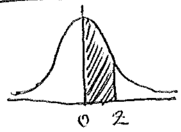
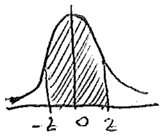
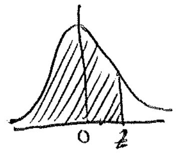
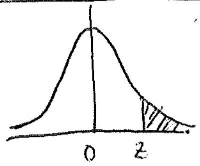
La ~~distr~~ densità di probabilità è simmetrica rispetto a $x=m$.



2) Se $\sigma_1 > \sigma_2$ allora $N_{m,\sigma_1}(m) < N_{m,\sigma_2}(m)$.

3) Se $\Phi_{m,\sigma}(x) = \int_{-\infty}^{x} N_{m,\sigma}(x)dx$, fissato m tutte le distribuzioni cumulative passano tutte per $(m, \frac{1}{2})$, cioè $\Phi_{m,\sigma}(m) = \frac{1}{2}$.

Tabelle: Sui libri si trovano sostanzialmente 4 tipi di tabelle:

z				
	$\Phi(z) - \frac{1}{2}$	$2\Phi(z) - 1$	$\Phi(z)$	$1 - \Phi(z)$
1,00	0.3413	0.6826	0,8413	0.1587
2,00	0.4772	0.9544	0,9772	0.0228
3,00	0.4986	0.9972	0,9986	0.0014
3.30	0,4995	0.9990	0.9995	0.0005

Il 68,27% delle misure è in $[m-\sigma, m+\sigma]$

Il 90,00% " in $[m-1.645\sigma, m+1.645\sigma]$

Il 95,00% " in $[m-1.960\sigma, m+1.960\sigma]$

Il 99.00% " " $[m-2,576\sigma, m+2.576\sigma]$

Esempi di uso delle tavole:

a) Sia X una variabile casuale con densità $N_{131,25,\ 156.25^2}$.
Calcolare il valore della normale standardizzata corrispondente a $x=97,6$

$$z = \frac{x-m}{\sigma} = \frac{97.6-131.25}{156.25} = -0,215$$

b) Sia X una variabile casuale con densità $X \sim N_{32,8,\ \sigma^2}$.
Se 1,39 è il valore standardizzato corrispondente a $x=42.4$ determinare la deviazione standard.

$$1.39 = \frac{42.4-32.8}{\sigma} \rightsquigarrow \sigma = 6,906$$

Esercizio: Il tempo necessario per risolvere un certo test è modellato secondo una distribuzione gaussiana con media 38,9 minuti e deviazione standard di 9.7 minuti.

Determinare la probabilità che uno studente termini il test in un intervallo compreso tra 30 e 45 minuti.

Soluzione: Dobbiamo calcolare $\mathbb{P}[30<X<45]$. Se Z è la variabile standardizzata si tratta di trovare i corrispondenti z_1 e z_2 di 30 e 45.

$$z_1 = \frac{30-38.9}{9.7} = -0.917 \qquad z_2 = \frac{45-38.9}{9.7} = 0.629$$

Ergo:

$$\mathbb{P}[30<X<45] = \mathbb{P}[z_1<Z<z_2] = \mathbb{P}[-0.917<Z<0.629] = \Phi(0.629) - \Phi(-0.917) =$$
$$= 0,7357 - 0.1788 = 0.557.$$

Approssimazione della distribuzione binomiale mediante la gaussiana

Sia X una variabile casuale binomiale con parametri n e p (n prove, p prob di successo). Se il numero delle prove tende all'infinito allora si ha:

$$\frac{n!}{x!(n-x)!}p^x q^{n-x} \rightarrow \frac{1}{\sqrt{2\pi[npq]}} e^{-\frac{(x-np)^2}{2npq}} = N_{np,npq}(x)$$

per ogni x appartenente ad un intervallo centrato sul valor medio np e di ampiezza pari alla deviazione standard $\sqrt{npq}$.

(TEOREMA DI BERNOULLI)

Per la dimostrazione si utilizza $n! \sim \sqrt{2\pi n}\,\frac{n^n}{e^n}$ (Stirling) e tanti calcoli.

Esempio: Determinare la probabilità che in 1200 lanci di un dado si ottenga 1 esattamente 190 volte.

Sol con la binomiale: $n = 1200$ $p = \frac{1}{6}$. si ha:

$$f_X(190) = \binom{1200}{190}\left(\frac{1}{6}\right)^{190}\left(\frac{5}{6}\right)^{1200-190} = 0.0233$$

ma è difficile da calcolare per problemi numerici.

Sol. con l'approx col teorema di Bernoulli:

Approssimiamo $f_X(190)$ con una distrib. gaussiana con

$m = np = \frac{1200}{6} = 200$ e $\sigma = \sqrt{npq} = \sqrt{1200 \cdot \frac{1}{6} \cdot \frac{5}{6}} = \sqrt{166,66\ldots}$

$$f_X(190) \sim N_{200,\,166,6}(190) = 0,02289 = \frac{1}{\sqrt{400\pi}}\, e^{-\frac{(190-200)^2}{(2\sqrt{166,66})^2}}$$

Esercizio: Calcolare la prob che su 1200 lanci di un dado si ottenga 1 tra **190 e 210** volte (nel senso che $190 < X \leq 210$).

Questa volta dobbiamo calcolare $\mathbb{P}[190 < X < 210]$ che possiamo approssimare con $\mathbb{P}\left[\frac{190-200}{\sqrt{166,66}} < Z \leq \frac{210-200}{\sqrt{166,66}}\right]$ dove Z = variabile standardizzata di X.

Dalle tavole calcoliamo $\Phi\left(\frac{10}{\sqrt{166..}}\right) - \Phi\left(-\frac{10}{\sqrt{166..}}\right) = 2\Phi(0,775) - 1 =$

$= 0,522$

DISTRIBUZIONI CONGIUNTE

__Definizione:__ Siano X ed Y due variabili casuali definite nello stesso spazio di probabilità.
Si definisce FUNZIONE DI DISTRIBUZIONE CUMULATIVA CONGIUNTA di X ed Y (o della variabile casuale bidimensionale (X,Y)) la funzione $F_{X,Y}(x,y) : \mathbb{R}^2 \to [0,1]$ definita da

$$F_{X,Y}(x,y) = \mathbb{P}[X \le x, Y \le y] \qquad \forall (x,y) \in \mathbb{R}^2.$$

(congiunta perché $X \le x$, $Y \le y$ si devono verificare contemporaneamente).

__Esempio:__ Esperimento che consiste nel lancio di 2 dadi con facce numerate da 1 a 6.
Sia $X = \{$esito lancio 1° dado$\}$ e $Y = \{\min$ fra i due numeri$\}$.
Spazio dei campioni $S = \{(i,j) : i,j = 1,2,..6\}$. Supponiamo che i punti di S siano equiprobabili cioè $\mathbb{P}((i,j)) = 1/36$ $\forall (i,j) \in S$
Le variabili X ed Y congiuntamente hanno codominio:

$X \le 3$

$F(x,y)$	1	2	3	4	5	6	F_X
1	6	6	6	6	6	6	6
2	7	12	12	12	12	12	12
3	8	14	18	18	18	18	18
4	9	16	21	24	24	24	24
5	10	18	24	28	30	30	30
6	11	20	27	32	35	36	36
F_Y	11	20	27	32	35	36	

$Y \le 4$

X \ Y	1	2	3	4	5	6	f_X
1	6	0	0	0	0	0	6
2	1	5	0	0	0	0	6
3	1	1	4	0	0	0	6
4	1	1	1	3	0	0	6
5	1	1	1	1	2	0	6
6	1	1	1	1	1	1	6
f_Y	11	9	7	5	3	1	

(21 valori)

Es: $(X,Y) = (3,3)$ in 4 casi su 36, quando gli esiti sono: $[3,3]$, $[3,4]$, $[3,5]$, $[3,6]$ (Ricordo che $Y = \min(i,j)$).

Per calcolare $F_{X,Y}$ si procede contando in quanti modi ogni evento si può manifestare: $F_{X,Y}(3,4) = \mathbb{P}(X \le 3, Y \le 4) = \mathbb{P}((1,1)) + \mathbb{P}((2,1)) + \mathbb{P}((2,2)) + \mathbb{P}(3,1) + \mathbb{P}(3,2) + \mathbb{P}(3,3) = \frac{6}{36} + \frac{1}{36} + \frac{5}{36} + \frac{1}{36} + \frac{1}{36} + \frac{4}{36} = \frac{18}{36} = \frac{1}{2}$

Per esercizio: completare i valori mancanti di $F_{X,Y}(x,y)$

PROPRIETÀ DELLA DISTRIBUZIONE CUMULATIVA CONGIUNTA

(i) $\lim_{x \to -\infty} F_{X,Y}(x,y) = F_{X,Y}(-\infty, y) = 0 \qquad \forall y \in \mathbb{R}$

$\lim_{y \to -\infty} F_{X,Y}(x,y) = F_{X,Y}(x, -\infty) = 0 \qquad \forall x \in \mathbb{R}$

$\lim_{\substack{x \to +\infty \\ y \to +\infty}} F_{X,Y}(x,y) = F_{X,Y}(+\infty, +\infty) = 1$.

(ii) Se (x_1, y_1), (x_2, y_2) sono tali che $x_1 < x_2$, $y_1 < y_2$ si ha:

$$P[x_1 < X \leq x_2, y_1 < Y \leq y_2] = F_{XY}(x_2, y_2) - F_{XY}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{XY}(x_1, y_1) \geq 0$$

(iii) $\lim_{h \to 0^+} F_{XY}(x+h, y) = \lim_{h \to 0^+} F_{XY}(x, y+h) = F_{XY}(x, y)$ ("continuità" a destra)

NOTA: <u>NON</u> è una continuità in $\mathbb{R}^2$.

<u>Definizione</u>: Se $F_{XY}(x,y)$ è la funzione di distribuzione cumulativa congiunta di X ed Y, le funzioni:

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y) \quad ; \quad F_Y(x) = \lim_{x \to \infty} F_{X,Y}(x,y)$$

si dicono DISTRIBUZIONI CUMULATIVE MARGINALI

<u>Osservazione</u>: $F_X(x) = \mathbb{P}[X \leq x, Y \leq +\infty] = \mathbb{P}((X \leq x) \cap (Y \leq +\infty)) = \mathbb{P}(X \leq x)$

$F_Y(y) = \mathbb{P}(Y \leq y)$

<u>Esercizio</u>: Calcolare le distribuzioni cumulative marginali per l'esempio precedente

FUNZIONE DI DENSITÀ CONGIUNTA

<u>Definizione</u>: Siano X e Y variabili casuali discrete definite sullo stesso spazio di probabilità.
La funzione di densità discreta congiunta di X ed Y (o di (X,Y)) è definita come

$$f_{X,Y}(x,y) = \begin{cases} \mathbb{P}(X = x, Y = y) & \text{se } (x_i, y_j) = \text{punto massa,} \\ 0 & \text{altrove.} \end{cases}$$

PROPRIETÀ DELLA DENSITÀ CONGIUNTA

(i) $f_{X,Y}(x,y) \geq 0$

(ii) $\sum_i \sum_j f_{X,Y}(x_i, y_j) = 1$ dove (x_i, y_j) sono i punti massa.

<u>Definizione</u>: Se X, Y variabili casuali discrete con densità $f_{X,Y}(x,y)$, le funzioni:

$$f_X(x_i) = \sum_{y_j} f_{X,Y}(x_i,y_j) \qquad (x_i \text{ fissato})$$

$$f_Y(y_j) = \sum_{x_i} f_{XY}(x_i,y_j) \qquad (y_j \text{ fissato})$$

si dicono FUNZIONI DI DENSITÀ DISCRETE MARGINALI

Esercizio: Calcolare $f_X(x_i)$ e $f_Y(y_j)$ per l'esempio precedente

CASO CONTINUO

<u>Definizione</u>: la variabile (X,Y) si dice continua se e solo se $\exists\, f_{X,Y}(x,y)$ tale che

$$F_{XY}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u,v)\,du\,dv$$

e $f_{XY}(x,y)$ si dice densità di probabilità congiunta.

<u>Proprietà</u>: (i) $f_{XY}(x,y) \geq 0$

(ii) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(u,v)\,dx\,dy = 1$$

<u>Osservazione</u>: $$\mathbb{P}(a \leq X \leq b,\ c \leq Y \leq d) = \int_a^b \left\{ \int_c^d f_{XY}(x,y)\,dy \right\} dx = \mathbb{P}\big((X,Y) \in [a,b]\times[c,d]\big).$$

In generale parleremo di eventi del tipo $(x,y) \in D$ con $D \subset \mathbb{R}^2$ e avremo $\mathbb{P}((x,y) \in D) = \iint_D f_{xy}(x,y)\,dx\,dy$.

<u>Teorema</u>: Data $f_{XY}(x,y)$ densità di prob. congiunta si ricava F_{XY} con la relazione

$$F_{XY}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u,v)\,du\,dv.$$

Viceversa, data $F_{XY}(x,y)$ si può ricavare f_{XY} con la formula

$$f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)$$

in tutti i punti in cui F è derivabile due volte.

Definizione: Se X, Y sono due variabili continue con densità congiunta $f_{XY}(x,y)$, le funzioni

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

si dicono FUNZIONI DI DENSITÀ MARGINALI di X e Y rispett.

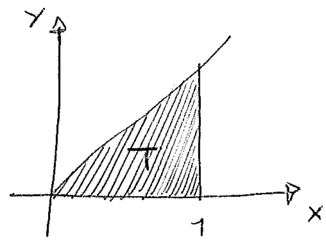
Osservazione:

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx}\left\{ \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{X,Y}(u,v)\,dv \right) du \right\} = \int_{-\infty}^{\infty} f_{XY}(x,v)\,dv$$

Esempio:

$$f_{XY}(x,y) = \begin{cases} k(x-y) & \text{per } 0 \le y \le x \le 1 \\ 0 & \text{altrove} \end{cases}$$

Determinare k in modo che sia una densità di probabilità e determinare le densità marginali.



(i) $f_{XY}(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2 \Rightarrow k \geq 0$

(ii) $\displaystyle\int_{\mathbb{R}^2} f_{XY}(x,y)\,dx\,dy = 1 : \int_0^1 \int_0^x k(x-y)\,dy\,dx = 1$

$$\int_0^1 \left\{ k\left(xy - \frac{y^2}{2}\right)\Bigg|_{y=0}^{y=x} \right\} dx = \int_0^1 k\left[x^2 - \frac{x^2}{2}\right] dx = k\int_0^1 \frac{x^2}{2}\,dx = k\,\frac{x^3}{6}\Bigg|_{x=0}^{x=1} = \frac{k}{6} = 1$$

da cui $k = 6$. Quindi $f_{XY}(x,y) = \frac{1}{6}(x-y)$ in T, 0 altrove

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \begin{cases} 0 & x < 0 \\ \displaystyle\int_0^x 6(x-y)\,dy = 6\left[xy - \frac{y^2}{2}\right]\Bigg|_{y=0}^{y=x} = 6\left(x^2 - \frac{x^2}{2}\right) = 3x^2 \\ 0 & x > 1 \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \begin{cases} 0 & y < 0 \\ \displaystyle\int_y^1 6(x-y)\,dx = 6\left[\frac{x^2}{2} - xy\right]\Bigg|_{x=y}^{x=1} = 6\left[\frac{1}{2} - y - \frac{y^2}{2} + y^2\right] = 3(1-y)^2 \\ 0 & y > 1 \end{cases}$$

~~Calcoliamo anche $F_{XY}(x,y)$: Nei punti di T si ha:~~

$$\cancel{F_{XY}(x,y) = \int_{-\infty}^{x}\left\{\int_{-\infty}^{y} 6(u-v)\,dv\right\}du = \int_0^x\left\{\int_0^y 6(u-v)\,dv\right\}du = \int_0^x\left\{6\left(uv - \frac{v^2}{2}\right)\Bigg|_{v=0}^{v=y}\right\}du =}$$

$$\cancel{= \int_0^x 6\left[uy - \frac{y^2}{2}\right]du = 6\left[y\frac{u^2}{2} - \frac{y^2}{2}u\right]\Bigg|_{u=0}^{u=x} = 6\left[\frac{yx^2}{2} - \frac{y^2}{2}\right] = 3y(}$$