Elementi di Probabilità   4/4/2010

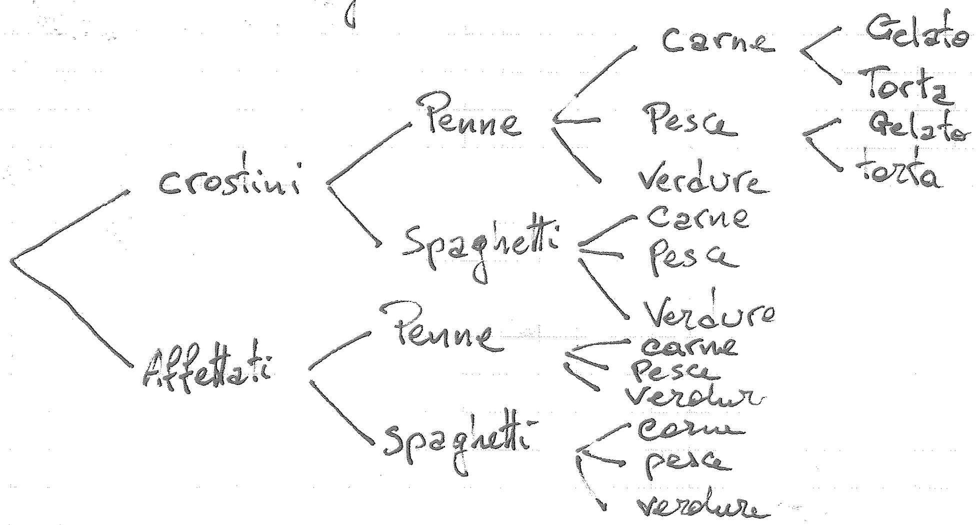
- Calcolo combinatorio

Impariamo a contare: Molti problemi di calcolo delle probabilità "elementare" richiedono il calcolo del numero di modi in cui un dato evento può verificarsi. Ci sono, in linea di principio, vari modi di fare questo calcolo (cioè di contarli).

Esempio 1. Scelta del menù. Quanti "pranzi" è possibile ordinare dal seguente menù?

Antipasto	Primo	Secondo	Dessert
- Crostini	- Penne	- Carne	- Gelato
- Affettati	- Spaghetti	- Pesce	- Torta
[niente]	[niente]	- verdura	[niente]
		[niente]	

Considero il seguente albero di scelte: ([niente] non è ammesso)



Ci sono 24 "foglie" cioè 24 possibili menu.

Se è possibile saltare una pietanza (cioè è possibile scegliere [niente]) ci sono $3 \times 3 \times 4 \times 3 = 108$ possibili menù compreso quello banale "sempre niente"

Esercizio: Quanti pranzi ci sono se non è consentito saltare più di 1 portata?

In generale, si può osservare che un processo che può essere completato in r stadi ognuno dei quali può essere fatto in $n_1, \dots n_r$ modi diversi, può essere eseguito in $n_1 \cdot n_2 \cdot \dots \cdot n_r$ modi. Questo può essere dimostrato per induzione, ma non lo faremo. Per il caso $r=2$ si vede subito che è così.

Esempio 2. I nomi degli americani hanno 3 iniziali (c'è il middle name). Consideriamo un paesino di 20.000 anime. In esso vivono almeno due persone (in realtà di più) con le stesse iniziali. Perché: in inglese ci sono 26 lettere quindi esistono $26^3 = 26 \cdot 26 \cdot 26 = 17526 < 20000$ terne possibili di iniziali.

Esempio 3. Supponiamo ci siano r persone in una stanza. Quante date (gg/mm) di compleanni possibili ci sono (escludiamo i nati bisestili). Elenchiamo le persone da 1 a r. Ci sono 365 possibilità per ciascuno quindi ci sono 365^r possibili date. Quante di queste sono distinte?

$$\underset{i=1}{\underset{\uparrow}{365}} \cdot \underset{i=2}{\underset{\uparrow}{364}} \cdot \underset{i=3}{\underset{\uparrow}{363}} \cdot \dots \cdot \underset{i=r}{\underset{\uparrow}{(365-r+1)}}$$

Questo numero si può anche scrivere così $\dfrac{365!}{(365-r)!}$

[Bisogna ricordarsi che $n! = 1 \cdot 2 \cdot \dots \cdot n$]

Se ci ricordiamo i coefficienti binomiali, possiamo anche scrivere questo numero così: $r!\binom{365}{r}$. Ricordo che $\binom{n}{k} := \dfrac{n!}{k!(n-k)!}$.

Consideriamo la seguente nozione intuitiva di probabilità:

$$\frac{\text{casi favorevoli}}{\text{casi possibili}}$$

e poniamoci il problema di stabilire, in una stanza di r persone quale sia la probabilità che ci siano almeno 2 persone con lo stesso compleanno. La probabilità che gli r compleanni siano tutti distinti è

$$P_{distinti}(r) = \frac{365 \cdot 364 \cdot \ldots \cdot (365-r+1)}{365^r}$$

Questo numero, approssimativamente vale

$r = 20 \quad \simeq 0{,}58$

$r = 30 \quad \simeq 0{,}29$

$r = 50 \quad \simeq 0{,}02$

La probabilità cercata vale $1 - P(r)$.

<u>Permutazioni</u> Una permutazione di un insieme di oggetti è un ordinamento di tali oggetti. Se ce ne sono molti è difficile elencare tutte le permutazioni possibili ce ne sono troppe anche per un computer! Fortunatamente però è facile trovare una formula per dire quante ce ne sono: se ci sono n oggetti ci sono n possibilità per il primo posto, $n-1$ per il secondo e così via. Quindi ci sono $n!$ permutazioni di n oggetti ($n!$ cresce molto rapidamente con n una formula asintotica è la seguente f. di Stirling

$$n! \simeq \sqrt{2\pi n}\; n^n e^{-n} = \sqrt{2\pi n}\; e^{n\log n - n}$$

che però vale solo in senso asintotico, infatti il valore assoluto dell'errore cresce con n)

Esempio 4. In quanti modi si possono disporre in fila 5 persone? Semplice, $5! = 120$.

Esempio 5. Uno zio vuole fotografare le sue 4 nipotine e 3 nipotini in fila mettendo alternativamente un maschio e una femmina. Quante possibilità ci sono? Lo schema della foto è

$$F\,M\,F\,M\,F\,M\,F$$

Quindi ci sono $4!$ disposizioni possibili per le femmine e $3!$ per i maschi. In totale $4!\cdot 3! = 24\cdot 6 = 144$.

Supponiamo di avere un insieme di n oggetti, numerati per semplicità da 1 a n. Vogliamo sapere quanti sottoinsiemi <u>ordinati</u> <u>distinti</u> di k elementi si possono fare. Attenzione! Considero insiemi <u>ordinati</u>, per capirsi $(1,2,3)$ è diverso da $(2,3,1)$. Ci sono n modi possibili per scegliere il primo elemento, $n-1$ per il secondo e così via fino al k-simo per cui ci sono $n-k+1$ modi. Cioè ci sono

$$n(n-1)\cdot\ldots\cdot(n-k+1) = \frac{n!}{(n-k)!} = k!\binom{n}{k}$$

possibilità. Questa formula dovrebbe ricordarci qualcosa!

Esempio 6a. Ci sono 5 bagnini che possono essere utilizzati nel turno del sabato pomeriggio e ci sono 3 torrette di guardia. In quanti modi può essere organizzato il turno? (Il modo in cui sono distribuiti conta). La risposta è

$$5\cdot 4\cdot 3 = 60.$$

Supponiamo che Ada sia innamorata di Anselmo (bagnino n°1) tenta di fingere un annegamento. Qual è la probabilità che sia proprio Anselmo a soccorrerla? (Qui conta in quale torretta si trova Anselmo). Ovviamente $1^{2}/60^{*5}$. Bassino!

* Supponiamo, per fissare le idee che Ada simuli il suo annegamento presso la torretta n° 1. Per rispondere alla domanda dobbiamo dire quanti turni prevedono che Anselmo sia in servizio alla n° 1. Presto detto: La 1 è fissata ad Anselmo rimangono $12 = 4 \cdot 3$ possibilità. Quindi la probabilità cercata è $\frac{12}{60}$.

Supponiamo di considerare irrilevante l'ordine dei sottoinsiemi di k elementi considerati prima. Per capirsi $\{1,2,3\} = \{3,2,1\}$. In questo caso si parla di <u>combinazioni</u> di k elementi (non di permutazioni). Quante ce ne sono? Basta osservare che ci sono $\frac{n!}{(n-k)!}$ insiemi ordinati di k elementi e che ci sono $k!$ permutazioni di un insieme di k elementi. Quindi, visto che togliendo la distinzione d'ordine due diverse permutazioni danno lo stesso insieme, ci sono

$$\frac{n!}{k!(n-k)!} =: \binom{n}{k}$$

combinazioni di k elementi in un insieme di n elementi.

Esempio 6b. Supponiamo che sia irrilevante l'ordine in cui i bagnini sono disposti nelle torrette (si veda es. 6a). Per capire quanti turni sono possibili dobbiamo guardare le combinazioni ce ne sono

$$\frac{5\cdot 4\cdot 3}{3!} = \frac{5!}{3!(5-3)!} = 10$$

Ora, quale probabilità ha Ada che Anselmo sia in servizio quando tenta il suo annegamento simulato? $6/10$ *. Meglio no?

Esempio 6c. Se i nostri bagnini sono A, B, C, D, E supponiamo che ognuna delle nostre 10 combinazioni sia equiprobabile. Quale probabilità ha E di <u>non</u> lavorare? Le combinazioni che contengono E sono quelle di un gruppo di 2 oggetti scelti tra 4 (perché E è fissato) cioè sono $\frac{4!}{2!(4-2)!} = \frac{24}{2\cdot 2} = 6$. Quindi quelle che non contengono E sono $10-6=4$. Cioè E ha probabilità $4/10$ di non lavorare.

* Si devono calcolare le combinazioni di 2 oggetti scelti tra 4 (A è assegnato) cioè $\frac{4!}{2!(4-2)!} = 6$.

La probabilità cercata è $6/10$.

Esercizio - Come cambia il risultato precedente se Anselmo (che si è finalmente accorto di Ada) si offre volontario per lavorare?

Esercizio - Tra i 30 spettatori di un evento sportivo 5 vengono scelti a caso per avere una maglietta ricordo. Le magliette sono tutte uguali quindi l'ordine non conta. Quanti gruppi di 5 persone si possono formare?

Risp. $\binom{30}{5} = \frac{30!}{5!25!} = 142506$

Esempio 7. Consideriamo i numeri $1, \dots n$. In quanti modi si possono dividere in 4 gruppi non vuoti di numeri consecutivi? Mettiamoli in fila e scegliamo 3 numeri a_1, a_2, a_3 e consideriamo i gruppi

$$\underbrace{1, 2 \dots a_1}_{\#1}\,]\,\underbrace{a_1+1, \dots \quad a_2}_{\#2}\,]\,\underbrace{a_2+1, \dots, a_3}_{\#3}\,]\,\underbrace{, \dots \quad n}_{\#4}$$

quindi $a_1, a_2, a_3 \in \{1, \dots n-1\}$. Ci sono $\binom{n-1}{3}$ combinazioni.

Esempio 8 - 12 studenti vengono divisi in 3 gruppi di 5, 4, 3 studenti. In quanti modi può essere effettuata tale divisione? Ci sono $\binom{12}{5}$ gruppi di 5. Rimangono da selezionare 4 studenti tra i rimanenti 7. Questo si può fare in $\binom{7}{4}$ modi. Quindi

$$\binom{12}{5}\binom{7}{4} = \frac{12!}{5!7!} \cdot \frac{7!}{4!3!} = \frac{12!}{5!4!3!} = 27720$$

Questo fatto è generale. Il numero di modi di dividere n oggetti in r gruppi di $k_1, \dots, k_r$ oggetti (con $k_1 + \dots + k_r = r$) è dato da $\frac{n!}{k_1! \cdot k_2! \dots k_r!}$

Esempio 9. Una scatola per viti ne contiene 8 grosse, 5 medie e 3 piccole, una di dadi ne contiene 6 grosse 4 medie 2 piccole. Vengono scelti a caso [equiprobabili] una vite e 1 dado. Quale è la probabilità di formare un bullone? Si avrà che

$$\text{Probabilità (dado e vite vanno d'accordo)} = \frac{\#\,\text{coppie compatibili}}{\#\,\text{coppie possibili}}$$

Per calcolare questi numeri osserviamo che ci sono $8+5+3=16$ viti e $6+4+2=12$ dadi. Inoltre ci sono $16\cdot12=192$ coppie dado-vite e $6\cdot8=48$ dado-vite grossi, $4\cdot5=20$ medi e $2\cdot3=6$ piccoli per un totale di $48+20+6=74$ compatibili. La probabilità cercata è $74/192$.

Esempio 10. Un vaso contiene 6 palline bianche e 5 nere se ne estraggono due a caso. Quale è la probabilità di estrarne una bianca e una nera?. Ci sono $10\cdot11$ possibili estrazioni [la prima pallina è scelta tra 11 e la seconda tra le rimanenti 10], ci sono 6×5 casi in cui la prima è bianca e la seconda nera e 5×6 in cui si verifica il contrario. La probabilità cercata vale $(30+30)/110$.

Esercizio. Quale è la probabilità nell'esempio precedente di pescare 2 bianche? E di pescare 2 nere? Che succede se la prima pallina è riposta nel vaso prima di procedere alla seconda estrazione?

Esempio 11. Ci sono 10 studenti, 6 maschi e 4 femmine. All'esame ottengono tutti voti diversi. ① Quante sono le possibili classifiche? ② Se tutte le classifiche sono equiprobabili quale è la probabilità che le prime 6 posizioni siano occupate dalle studentesse? ③ Quale è la probabilità che le ultime posizioni siano occupate dalle studentesse?

① $10! = 3628800$ ② Ci sono $4!$ classifiche di femmine e $6!$ di maschi. Allora ci sono $4!6!$ classifiche in cui le femmine occupano i posti migliori. La probabilità cercata è $\frac{4!6!}{10!} = \frac{1}{210}$.
③ Esattamente come nel punto ②.

Esercizio. Nell'esempio precedente, quale è la probabilità che ci siano 2 femmine ai posti migliori? Risp. $\frac{\binom{4}{2}8!}{10!}$.

Esempio 12. Una commissione di 5 persone è scelta a caso tra un gruppo di 6 uomini e 9 donne. Quale è la probabilità che vengono scelti 3 uomini e 2 donne? "A caso" significa che le $\binom{15}{5}$ combinazioni sono equiprobabili. Le scelte che portano al risultato voluto sono $\binom{6}{3}\binom{9}{2}$. La probabilità cercata è

$$\binom{6}{3}\binom{9}{2}\Big/\binom{15}{5} = \frac{240}{1001}$$

Esempio 13. Si estrae da un insieme di n elementi un sottoinsieme di k elementi. Supponiamo di avere fissato in partenza un elemento dell'insieme. Quale è la probabilità che il sottoinsieme estratto contenga l'elemento fissato? Ci sono $\binom{n-1}{k-1}$ sottoinsiemi di cardinalità k che contengono l'elemento fissato e ci sono $\binom{n}{k}$ sottoinsiemi di cardinalità k. La probabilità cercata vale $\binom{n-1}{k-1}\Big/\binom{n}{k} = \frac{(n-1)!}{(n-k)!(k-1)!} \cdot \frac{k!(n-k)!}{n!} = \frac{k}{n}$.

Esempio 14. Una squadra è formata da 6 giocatori neri e 6 bianchi. Devono essere divisi a coppie per occupare camere doppie. Se la divisione è fatta a caso, quale è la probabilità che nessun nero divida la camera con un bianco? Ci sono $12/2^6$ modi di dividere i giocatori in coppie distinte (coppie possibili). Ci sono $\frac{6!}{2^3}$ possibili coppie di neri e altrettante di bianchi che possono essere distribuite in $\binom{6}{3}$ modi. Quindi la probabilità cercata è

$$\frac{\frac{6!}{2^3}\cdot\binom{6}{3}\cdot\frac{6!}{2^3}}{\frac{12!}{2^6}} = \frac{5}{231}.$$

- Definizioni e proprietà

Finora abbiamo considerato un approccio molto intuitivo alla probabilità. Lo studio di molti fenomeni non può essere intrapreso in modo così elementare per cui è necessario definire qualche nozione.

Sia Ω un insieme (l'insieme dei risultati possibili).

Una famiglia $\mathcal{A}$ di sottoinsiemi di Ω è detta un'algebra se

1) se $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$
2) se $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$
3) se $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

Per induzione segue che se $A_1 \dots A_n \in \mathcal{A}$ allora $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$ e $\bigcap_{i=1}^{n} A_i \in \mathcal{A}$.

La famiglia $\mathcal{A}$ è una σ-algebra se

1) $\emptyset, \Omega \in \mathcal{A}$
2) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
3) Se $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

(Una σ-algebra non vuota è un'algebra) Notiamo che dalle proprietà ② e ③ segue che $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$. Questo segue dalle leggi di De Morgan

$$\bigcap_{i \in I} A_i = \left(\bigcup_{i \in I} A_i^c \right)^c \; ; \; \bigcup_{i \in I} A_i = \left(\bigcap_{i \in I} A_i^c \right)^c$$

Nel caso di Ω finito c'è soltanto da considerare algebre finite. Niente paura noi ci limiteremo sempre al caso in cui Ω è un insieme finito. In una tale situazione si può considerare $\mathcal{A} = \mathcal{P}(\Omega)$ = insieme delle parti di Ω cioè la famiglia di tutti possibili sottoinsiemi di Ω. La maggiore generalità descritta sopra diventa necessaria quando Ω non

è un insieme finito. Infatti, in quei casi non sempre è possibile (anzi, quasi mai) definire una probabilità (vediamo come tra un attimo) su tutti i sottoinsiemi di Ω. Questo comunque va oltre gli scopi di questo corso.
Prima di definire cosa si intende per probabilità vediamo qualche esempio.

Esempio 15. Sia Ω l'insieme dei possibili risultati del lancio di 3 monete (3 lanci succ. di una moneta). Cioè

$$\Omega = \{TTT, TTC, TCC, CCC, TCT, CTT, CTC, CCT\}$$

(l'ordine conta). Ovviamente $\# \Omega = 2^3 = 8$. Sia $\mathcal{A} = \mathcal{P}(\Omega)$. [Quale è la cardinalità di $\mathcal{P}(\Omega)$?]*3 Sia poi $A \in \mathcal{A}$ che corrisponde all'evento "viene almeno una testa", cioè

$$A = \{TTT, TTC, TCC, TCT, CTT, CTC, CCT\}$$

e B quello corrispondente a "vengono almeno due teste", cioè

$$B = \{TTT, TTC, TCT, CTT\}$$

Esercizio. Scrivere l'elemento di $\mathcal{A}$ corrispondente a "viene esattamente una testa".

Esercizio. Sia Ω lo spazio dei possibili risultati del lancio di 2 monete. Sia $\mathcal{A} = \mathcal{P}(\Omega)$, quale è $\# \mathcal{A}$?

Risp. $\binom{4}{4} + \binom{4}{3} + \binom{4}{2} + \binom{4}{1} + \binom{4}{0} = 16$

Esercizio. Sia Ω lo spazio dei possibili risultati del lancio di 2 dadi e $\mathcal{A} = \mathcal{P}(\Omega)$. Scrivere $A \in \mathcal{A}$ corrispondente a "la somma del valore dei dadi è divisibile per 3".

Diamo ora la definizione di probabilità di un evento. Qui, per "evento" si intende un elemento di $\mathcal{A}$.

* Se un insieme Ω ha cardinalità n il suo insieme delle parti ha cardinalità 2^n. Infatti

$$\# \mathbb{P}(\Omega) = \binom{n}{n} + \binom{n}{n-1} + \cdots + \binom{n}{1} + \binom{n}{0}$$

e, per il teorema del binomio di Newton

$$\left[(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \right]$$

si ha

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k}$$

Da cui segue l'affermazione.

Nel caso considerato $\# \Omega = 8$ quindi

$$\# \mathbb{P}(\Omega) = 2^8 = 256$$

Sia Ω un insieme e $\mathcal{A}$ una σ-algebra di sottoinsiemi di Ω. Una probabilità è un'applicazione $P:\mathcal{A}\to[0,1]$ tale che

1) $P(\Omega)=1$

2) Se $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ con $A_i\cap A_j=\phi$ per $i\neq j$, allora

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

(Quest'ultima proprietà si chiama σ-additività).

Chiamiamo <u>spazio di probabilità</u> una terna $(\Omega,\mathcal{A},P)$ con Ω un insieme, $\mathcal{A}$ una σ-algebra di parti di Ω e P una probabilità su $\mathcal{A}$. Si può dire che P "misura", e il termine non è scelto a caso, gli elementi di $\mathcal{A}$. Come ho già detto, una definizione così generale è necessaria per poter coprire le situazioni in cui Ω non è un insieme finito, magari neanche discreto. Noi non ci occuperemo di queste situazioni complicate. Ci limiteremo al caso $\#\Omega$ finito e $\mathcal{A}=\mathcal{P}(\Omega)$. Prima però conviene fare un'osservazione.

Notiamo che gli spazi di probabilità sono delle astrazioni, dei modelli matematici per dei fenomeni. La modellizzazione di un fenomeno naturale cioè la scelta di uno spazio di probabilità adeguato <u>non</u> è un procedimento matematico ma viene effettuato sulla base di considerazioni empiriche tenendo conto della natura del problema, magari delle esigenze di semplicità. Si porrà poi, a posteriori, il problema di verificare la correttezza, o meglio l'adeguatezza, delle scelte fatte. Questo sarà uno dei compiti della statistica matematica.

Esempio 16. Consideriamo il lancio di un dado a 6 facce

$$\Omega=\{1,2,3,4,5,6\}\ ,\quad \mathcal{A}=\mathcal{P}(\Omega)$$

se il dado è equo (non sbilanciato) poniamo

$$P(\{1\}) = P(\{2\}) = \dots = P(\{6\}) = \frac{1}{6}$$

(cioè supponiamo che tutti gli esiti siano equiprobabili e usiamo $P(\Omega)=1$). Sia A l'evento "otteniamo un numero pari", cioè $A = \{\{2\},\{4\},\{6\}\}$. Dalla σ-additività segue

$$P(A) = 3 \cdot \frac{1}{6} = \frac{1}{2}$$

Un esempio importante di spazi di probabilità sono gli spazi uniformi (quelli che considereremo noi).

Sia Ω un insieme finito, cioè t.c. $\#\Omega < +\infty$, e sia $\mathcal{A} = \mathbb{P}(\Omega)$. Poniamo $P(\{\omega\}) = p$ con p assegnato e $\{\omega\}$ un singoletto (cioè un sottoinsieme fatto da un singolo elemento di Ω). Si ha

$$1 = P(\Omega) = \sum_{\omega \in \Omega} P(\{\omega\}) = (\#\Omega) \cdot p$$

quindi $p = \frac{1}{\#\Omega}$. Se $A \in \mathcal{A}$ è un evento, $P(A) = \frac{\#A}{\#\Omega}$.

Esempio 17. Giochiamo i numeri 1, 2, 3 al lotto su una singola ruota. Calcoliamo la probabilità di fare un terno. Vengono estratte 5 palline senza reimbussolamento tra 90. Sia Ω l'insieme di tutte le cinquine. Supponiamo che tutte le cinquine siano equiprobabili. Sia A l'insieme delle cinquine che contengono 1, 2 e 3. Avremo

$$P(A) = \frac{\#A}{\#\Omega} = \frac{87 \cdot 86}{\binom{90}{5}} \simeq 0{,}00017$$

non giocare è meglio!

Esercizio. Qual è la probabilità di fare un ambo nell'esercizio precedente? Torneremo su questo esercizio. * vedere pag. 15

Vediamo di ricavare alcune formule utili per il calcolo della probabilità.

i) Osserviamo che se $A \in \mathcal{A}$ allora A^c è un evento incompatibile con A. Quindi $1 = P(\Omega) = P(A) + P(A^c)$ da cui segue

$$\boxed{P(A) = 1 - P(A^c)}$$

ii) Siano $A, B \in \mathcal{A}$ eventi. Allora $B \cap A$ e $B \cap A^c$ sono incompatibili e $B = (B \cap A) \cup (B \cap A^c)$ da cui segue

$$\boxed{P(B) = P(B \cap A) + P(B \cap A^c)}$$

iii) Siano $A, B \in \mathcal{A}$ eventi tali che $A \subseteq B$ allora, per la (ii),

$$P(B) = P(\underbrace{B \cap A}_{=A}) + P(B \cap A^c) = P(A) + \underbrace{P(B \cap A^c)}_{\geq 0} \geq P(A)$$

da cui segue

$$\boxed{\text{Se } A \subseteq B \text{ allora } P(B) \geq P(A)}$$

iv) Dalla (i) e dalle leggi di De Morgan, se $A_1, A_2 \in \mathcal{A}$ sono eventi segue che $P(A_1 \cup A_2) = 1 - P((A_1 \cup A_2)^c) = 1 - P(A_1^c \cap A_2^c)$. Più in generale, si ha che se $\{A_i\}_{i=1}^N$ è una famiglia di eventi

$$\boxed{P\left(\bigcup_{i=1}^N A_i\right) = 1 - P\left(\bigcap_{i=1}^N A_i^c\right)}$$

Esempio 18a. Probabilità di ottenere almeno un 6 lanciando 2 volte un dado (cubico a 6 facce). $\Omega = \{(1,1), (1,2), \dots, (1,6), (2,1), \dots (2,6), \dots\}$

Quindi $\#\Omega = 36$. $A_1 = \{$coppie il cui 1° elemento è 6$\}$

$A_2 = \{$coppie il cui 2° elemento è 6$\}$

L'evento cercato è $A = A_1 \cup A_2$. Notiamo che A_1 e A_2 <u>non</u> sono disgiunti. Tuttavia è facile calcolare $A_1^c \cap A_2^c = \{(\alpha, \beta) : \alpha \in \{1, \dots 5\}, \beta \in \{1, \dots, 5\}\}$ cioè l'insieme delle coppie di numeri interi da 1 a 5. Allora $\#(A_1^c \cap A_2^c) = 25$ dalla (iv) segue $P(A) = 1 - P(A_1^c \cap A_2^c) = 1 - \frac{25}{36} = \frac{11}{36}$

v) Siano A e B eventi. Si ha $A \cup (B \cap A^c) = A \cup B$. Osserviamo che A e $A^c \cap B$ sono incompatibili. Allora,

$$P(A \cup B) = P(A) + P(B \cap A^c)$$

e, per la (ii),

$$P(B \cap A^c) = P(B) - P(B \cap A)$$

ne segue che

$$\boxed{P(A \cup B) = P(A) + P(B) - P(B \cap A)}$$

Esempio 18b. Con le stesse notazioni dell'esercizio 18a, si ha $A_1 \cap A_2 = \{(6,6)\}$, quindi $P(A_1 \cap A_2) = \frac{1}{36}$. Allora, potevamo ricavare $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$.

In realtà vale la seguente formula generale che enunciamo senza dimostrazione: Siano $A_1 \dots A_n$ eventi

$$\boxed{P(A_1 \cup \dots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P(A_1 \cap \dots \cap A_n).}$$ *5

Esempio 19. Siano $Q = \{a_1 .. a_n\}$ simboli. Troviamo la probabilità che una permutazione casuale di Q ne lasci alcuno fisso. Sia A_i l'evento "la permutazione lascia fisso a_i". Ci sono $(n-1)!$ permutazioni con questa proprietà. Ognuna ha probabilità $\frac{(n-1)!}{n!} = \frac{1}{n}$, allora $\sum_{i=1}^{n} P(A_i) = n \cdot \frac{1}{n} = 1$. Ora $A_i \cap A_j$, se $i \neq j$, è l'evento che fissa la coppia (a_i, a_j). Di questi eventi ce ne sono $(n-2)!$ e $P(A_i \cap A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$. Ci sono $\binom{n}{2}$ eventi di questo tipo (cioè con $i \neq j$). Dunque $\sum_{1 \le i < j \le n} P(A_i \cap A_j) = \binom{n}{2} \frac{1}{n(n-1)} = \frac{1}{2!}$.
Similmente $P(A_i \cap A_j \cap A_k) = \frac{(n-3)!}{n!}$ se i, j, k sono tutti diversi e $\sum_{i<j<k} P(A_i \cap A_j \cap A_k) = \frac{1}{3!}$.
In definitiva $P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \frac{1}{2!} + \frac{1}{3!} - \dots \frac{(-1)^{n+1}}{n!}$ è la probabilità cercata.
Notiamo che per n crescente, questa probabilità si avvicina rapidamente a e^{-1}.

* Consideriamo il problema di fare un ambo giocando i numeri 1, 2, 3 al lotto su una singola ruota. Sia A l'evento $A = A_1 \cup A_2 \cup A_3$ con $A_1 =$ esce $\{1,2\}$, $A_2 =$ esce $\{1,3\}$, $A_3 =$ esce $\{2,3\}$
abbiamo che $A_1 \cap A_2 \cap A_3 =$ esce $\{1,2,3\}$ e $A_i \cap A_j =$ esce $\{1,2,3\}$

Si ha

$$P(A) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_1 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$$
$$= 3 \cdot \frac{5!}{90 \cdot 89} - 3 \cdot \frac{5!}{90 \cdot 89 \cdot 88} + \frac{5!}{90 \cdot 89 \cdot 88} \simeq 0.0446$$

Probabilità condizionata.

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Siano A ed E eventi. Ci si chiede quale sia la probabilità di A a condizione che si verifichi E. Si introduce l'evento $A|E$ (A subordinato ad E) che si verifica se si verificano E e A. non si verifica se E si verifica e A no. Non è definito se E non si verifica. (Nel senso che E è da considerare il nuovo spazio degli eventi)

Esempio 20. Prendiamo l'evento $A = \{\text{risultato} = 2\}$ nel lancio di un dado, condizionato dall'evento $E = \{\text{risultato pari}\}$.

Nel caso che stiamo considerando, cioè $\#\Omega < +\infty$ abbiamo naturalmente che $P(A|E) = \frac{\#(A\cap E)}{\#E} = \frac{\#(A\cap E)}{\#\Omega} \cdot \frac{\#\Omega}{\#E}$

da cui segue

$$\boxed{P(A|E) = \frac{P(A\cap E)}{P(E)}}$$

In generale, cioè per spazi di probabilità qualunque, questa relazione diventa una definizione.

Diremo che due eventi A e B sono indipendenti se non si "influenzano" l'uno l'altro. Più precisamente se $P(A|B) = P(A)$ (o, equivalentemente $P(B|A) = P(B)$). In modo equivalente, si ha che A e B sono indipendenti se $P(A\cap B) = P(A)P(B)$

Questa si chiama anche formula delle probabilità composte.

In generale, la legge delle probabilità composte diventa per una famiglia di eventi $A_1, \dots A_n$,

$$\boxed{\begin{array}{l}\text{Se } P(A_1 \cap \dots \cap A_{n-1}) > 0 \text{ allora} \\ \quad P(A_1\cap\dots\cap A_n) = P(A_1)\cdot P(A_2|A_1)\cdot \dots\ P(A_n|A_1\cap\dots A_{n-1})\end{array}}$$

Nel caso di eventi indipendenti, la legge diventa

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot \dots \cdot P(A_n)$$

Osservazione. Gli eventi Ω e $\emptyset$ sono indipendenti da qualunque altro evento. Infatti, se $A \in \mathcal{A}$ è un evento

$$0 = P(A) \cdot P(\emptyset) \qquad \text{perchè } P(\emptyset) = 1 - P(\Omega) = 0$$

ma $\quad 0 = P(\emptyset) = P(A \cap \emptyset) \quad$ perchè $A \cap \emptyset = \emptyset$.

Dunque $\quad P(A \cap \emptyset) = P(A) P(\emptyset)$.

Similmente, $\quad P(A) \cdot P(\Omega) = P(A) \quad$ perchè $P(\Omega) = 1$

inoltre $\quad P(\Omega \cap A) = P(A) \quad$ perchè $\Omega \cap A = A$

Dunque $\quad P(\Omega \cap A) = P(\Omega) P(A)$.

Osservazione. Se A, B sono eventi indipendenti allora A^c e B sono indipendenti. Infatti, $P(B) = P(A^c \cap B) + P(A \cap B) = P(A^c \cap B) + P(A)P(B)$

ne segue che $\quad P(A^c \cap B) = P(B)(1 - P(A)) = P(A)P(B)$.

Similmente si dimostra che anche A^c, B^c sono indipendenti.

Esercizio. Verificare che se A, B sono indipendenti allora

$$P(A \cup B) = 1 - P(A^c) P(B^c)$$

Si ha che, per eventi A e B,

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

se $P(A)P(B) > 0$ si ottiene

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

o, equivalentemente

$$P(A|B) = P(A) \cdot \frac{P(B|A)}{P(B)} \qquad \text{Formula di Bayes}$$

Questa formula si può interpretare come la probabilità che si sia verificata l'ipotesi A sapendo che si è verificato l'evento B.

Esempio 21. Due quesiti in una prova sono selezionati casualmente in un archivio contenente 50 domande di attualità, 30 di storia e 20 di letteratura. Quale è la probabilità che vengono selezionate 2 domande di attualità? Siano B_1 e B_2 gli eventi in cui il 1° e 2° rispettivamente è di attualità.

Si ha $P(B_1 \cap B_2) = P(B_1) \cdot P(B_2 | B_1) = \frac{50}{100} \cdot \frac{49}{99}$

Esempio 22. Per vincere ad un gioco si devono superare 3 prove indipendenti. La probabilità di superarle è rispettivamente 0,5, 0,3 e 0,1. Quale è la probabilità di vincere? Se A_i denota l'evento di vincere alla i-sima prova, si ha

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3) = 0{,}015.$$

Esempio 23. Si consideri il gioco in cui 1 numero è estratto tra $\{1, 2, \dots, 1000\}$. Sia E l'evento: è estratto un numero da 1 a 100. Se A è l'evento: viene estratto 1. Sapendo che si verifica A quale è la probabilità che si verifichi E? Ovviamente 1 (è un evento certo). Notiamo che questo è coerente con la formula di Bayes.

Esempio 24. Siano M_1 e M_2 due macchine che producono lo stesso tipo di oggetto. Questi oggetti sono indistinguibili per ciò che riguarda la loro origine. In un'ora queste macchine producono rispettivamente 100 e 50 pezzi, ma 8 sono difettosi nel caso di M_1 e solo 2 nel caso di M_2. Supponiamo di trovare un prodotto difettoso. Quale è la probabilità che sia stato prodotto da M_1? Usando la formula di Bayes, posto A_1 l'evento essere prodotto da M_1 e B l'evento essere difettoso, si ha

$$P(A_1|B) = P(A_1) \cdot \frac{P(B|A_1)}{P(B)}$$

$$= \frac{100}{150} \cdot \frac{\frac{8}{100}}{\frac{10}{150}} = \frac{8}{10}.$$

La formula di Bayes può essere generalizzata come segue: Sia $H_1 \dots H_n$ una partizione di Ω nel senso che $H_1 \dots H_n$ sono mutualmente incompatibili e $H_1 \cup H_2 \cup \dots \cup H_n = \Omega$. Sia E un evento. Allora, per ogni $i=1, \dots n$,

$$P(H_i|E) = P(H_i) \frac{P(E|H_i)}{P(H_1)P(E|H_1) + \dots + P(H_n)P(E|H_n)}$$

L'interpretazione è la seguente: Se $H_1 \dots H_n$ sono ipotesi. Si conduce un esperimento il cui risultato è l'evento E. Sappiamo che H_i produce E con probabilità $P(E|H_i)$. La formula sopra ci dice con quale probabilità è vera H_i.

Esempio 25. Il Dr. House riceve un paziente con un sintomo E che lui sa essere prodotto con probabilità p_1 dalla condizione H_1, p_2 dalla H_2 e p_3 dalla H_3. Sa che non ci sono altre possibilità. Sa inoltre che la probabilità di contrarre H_i, $i=1, \dots 3$, è f_i. Quale è la probabilità che il paziente sia affetto da H_1? Per la formula di Bayes,

$$P(H_1|E) = f_1 \cdot \frac{p_1}{p_1 f_1 + p_2 f_2 + p_3 f_3}$$ è la probabilità cercata. Inutile dire che questo esempio non è troppo realistico!

Elementi di statistica descrittiva

In questa breve lezione introduttiva ci occuperemo di quella parte della statistica (N.b. statistica = scienza dello stato) che si occupa di descrivere sinteticamente un insieme (potenzialmente anche molto grande) di dati.

L'idea è che se si hanno a disposizione molte osservazioni, può trattarsi di misurazioni, rilevamenti demografici, registrazioni di azioni di borsa, log di dati, si ha bisogno di strumenti che aiutino a riassumere in pochi numeri questa massa di dati. La mente umana non è fatta per gestire grandi elenchi e il processo decisionale beneficia della disponibilità di <u>indicatori sintetici</u> che descrivano adeguatamente la situazione senza <u>sommergerci</u> di informazioni.

Il problema è che ogni sintesi porta necessariamente a scartare alcune informazioni. Si pone dunque il problema di trovare degli indicatori che mantengano le informazioni "importanti" e scartino le altre.

Un po' di terminologia:

- Un <u>fenomeno statistico</u> riguarda una pluralità di individui o eventi, deve essere analizzato in circostanze omogenee e deve essere riscontrabile in più occasioni o individui.
- Una <u>popolazione</u> è un insieme di elementi che sono omogenei rispetto ad una (almeno) caratteristica. Per esempio sono popolazioni Gli italiani, gli infermieri, le temperature massime registrate a Firenze nel 2001.
- Le <u>unità statistiche</u> sono i soggetti o elementi (i singoli dati) che sono oggetto di un'indagine. Sono indivisibili nell'ambito dell'indagine condotta

- Un campione rappresentativo è una parte della popolazione che la rappresenta bene rispetto alle caratteristiche studiate

Elementi importanti della statistica descrittiva sono

i) Le tabelle
ii) Le rappresentazioni grafiche (i grafici)
iii) Gli indici sintetici

In breve, le tabelle servono a ordinare e presentare i dati; i grafici servono a visualizzarli e gli indici sintetici servono a riassumerli.

Esempio 26a. La seguente tabella rappresenta i tempi di percorrenza di una certa tratta cittadina dell'autobus osservati in 10 giorni

Giorno	1	2	3	4	5	6	7	8	9	10
tempo (min.)	26	23	29	21	19	23	18	20	25	22

Un indicatore di quanto "solitamente" si impiega a percorrere quel tratto è la media. Semplicemente la media aritmetica delle osservazioni. Nel nostro caso $\frac{1}{10}(26+23+29+\dots+22) = 22{,}6$ minuti.

La media aritmetica delle osservazioni è uno degli indicatori di cui parlavamo:

Se $x_1 \dots x_n$ sono valori osservati, la media è

$$\boxed{\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

Ci chiediamo però quanto la media aritmetica sia una buona indicazione.

Esempio 26b. Nella situazione dell'esempio 26a, supponiamo di estendere le nostre osservazioni anche all'11° giorno in cui si verifica uno sciopero e impieghiamo 70 minuti per percorrere il tratto in esame. La media calcolata tenendo conto di questo valore diventa circa 26,9 minuti. Questa non è una buona indicazione della durata tipica di un viaggio. Forse è meglio cercare un altro indicatore. Disponiamo le nostre rilevazioni in ordine crescente

18 19 20 21 22 23 23 25 26 29 70

il valore centrale rispetto a quest'ordine (il 6°) è 23 minuti che è una indicazione migliore della durata tipica del nostro viaggio. E' detta <u>mediana</u>. Notiamo che se non teniamo conto dell'11ª osservazione, non troviamo un valore centrale tra le osservazioni fatte (perchè sono in numero pari) ne inseriamo uno artificialmente mettendo la media delle due centrali (la 5ª e la 6ª) ottenendo 22,5 minuti. Si vede che la mediana è meno sensibile della media alla presenza di osservazioni "sballate".

Per definire la mediana si fa così:

- Caso numero n <u>dispari</u> di osservazioni.

  Si ordinano le osservazioni $\{x_i\}_{i=1}^{n}$ in modo tale che

  $$x_1 \leq x_2 \leq \dots \leq x_n$$

  La mediana μ è data da $\mu = x_N$ con $N = \frac{n+1}{2}$

- Caso numero n <u>pari</u> di osservazioni.

  Si ordinano le osservazioni $\{x_i\}_{i=1}^{n}$ in modo che

  $$x_1 \leq x_2 \leq \dots x_n$$

  La mediana μ è data da $\mu = \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n+1}{2}}\right)$.

Osserviamo anche che ci sono vari tipi di media. Una che riveste particolare importanza è la <u>media pesata</u>. Supponiamo di avere rilevazioni $x_1 \dots x_n$ in ognuna delle quali riponiamo fiducia diversa. Diciamo che i numeri $p_1 \dots p_n$ rappresentano il grado di fiducia che riponiamo in $x_1 \dots x_n$ rispettivamente. La media pesata di $x_1 \dots x_n$ è data da

$$\boxed{m = \frac{p_1 x_1 + \cdots + p_n x_n}{p_1 + p_2 + \cdots + p_n}}$$

Un caso tipico è il seguente.

Esempio 27. Supponiamo che due gruppi uno di 10 e uno di 15 persone abbiano misurato indipendentemente la lunghezza di una certa strada. La lunghezza stimata dal primo gruppo è di 11,2 km e quella stimata dal secondo gruppo è da 10,9 km. È naturale, in assenza di altri fattori, assegnare peso 10 alla lunghezza stimata dal primo gruppo e 15 a quella stimata dal secondo. La media pesata è allora $m = \frac{11,2 \cdot 10 + 10,9 \cdot 15}{25} = 11,02$.

C'è un altro indicatore sintetico utile: la <u>moda</u>. Questa è definita come i (ce ne può essere più di una) valore di massima frequenza tra quelli osservati.

Esempio 28. Nel caso dell'esempio 26a e 26b la moda è 23 minuti. Infatti il 23 si presenta due volte nella tabella. Osserviamo che la moda è applicabile anche a caratteri non numerici (cioè per valori non esprimibili con un numero).

Esempio 29. La seguente tabella rappresenta il colore degli occhi dei bambini nati in un certo ospedale in un mese dato.

Colore	numero di bambini
Azzurro	13
Azzurro chiaro/grigio	6
Blu	12
Marrone chiaro	23
Marrone scuro/Nero	19
Verde	18

La moda è ovviamente "marrone chiaro".

A proposito della frequenza che compare nella definizione di moda, osserviamo che questa può comparire in due forme sostanzialmente equivalenti. La frequenza (assoluta) di una modalità è il numero che questa viene rilevata. La frequenza relativa è la frequenza diviso il numero totale di osservazioni. Spesso quest'ultima è riportata in percentuale.

Esercizio. Scrivere la tabella delle frequenze relative dei dati osservati nell'esempio 28.

Notiamo che le frequenze relative e assoluta sono caratteri numerici anche se la modalità osservata poteva non esserlo.

Un altro modo per distillare informazioni utili da una massa di dati è stabilire quale sia la posizione relativa, per esempio del dato che ci interessa, rispetto a tutti i dati. Si pensi alla situazione di un manager che vuole stabilire quale è la posizione relativa della sua industria tra altre 500 che producono cose simili. Una classifica può contenere troppe informazioni

La nozione che viene spesso utilizzata a questo scopo è quella di quantile. Per una popolazione finita $x_1 \dots x_n$ ordinata in modo tale che $x_1 \leq x_2 \leq \dots \leq x_n$ il k-simo q-quantile ($k<q$) si calcola così: Calcoliamo il numero $I = n \cdot \frac{k}{q}$. Si considerano due casi

i) I non è intero. In questo caso sia J il primo intero maggiore di I. Il k-simo q-quantile è x_J.

ii) I è intero. In questo caso il k-simo q-quantile è dato da $\frac{1}{2}(X_I + X_{I+1})$.

L'idea è di dividere l'insieme dei dati in q parti essenzialmente uguali. Notiamo che per $q=2$ il 1° quantile è la mediana

Esempio 30. Considero una popolazione di 10 valori

$$3, 6, 7, 8, 8, 10, 13, 15, 16, 20$$

Il 1° 4-quantile (detto quartile) è in posizione 3 cioè 7.

il 2° quartile (è la mediana) è dato da 9. Il 3° quartile è in posizione 8 cioè è 15.

Tutto quanto visto fino ad ora è utile fino ad un certo punto: dobbiamo anche sapere come sono dispersi i dati.

Esempio 31. La seguente tabella raccoglie le performance di due azioni, A e B, in 8 giorni

A	2	4	2	4	2	4	2	4
B	0,5	5,5	0,5	5,5	0,5	5,5	0,5	5,5

La performance media è la stessa ma è ovvio che il comportamento è molto diverso.

Ci interessiamo ora ad alcuni indicatori di quanto un insieme di dati è disperso.

Nel caso dell'esempio 31 l'intervallo di variazione di A è $[2,4]$ mentre quello di B è $[0{,}5, 5{,}5]$ quindi l'ampiezza della variazione nel caso di A è 2 e di 5 nel caso di B.

Una misura della dispersione dei dati è la <u>varianza</u> definita da

$$\boxed{\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

dove $x_1, \dots x_N$ sono le osservazioni.

Il problema della varianza è che non rispetta le unità di misura. Per esempio, se gli x_i sono come nell'esempio 26a σ^2 viene misurata in minuti^2 che è una strana unità di misura.

Meglio è usare la <u>deviazione standard</u>

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

che mantiene le unità di misura.

E' interessante osservare che se avessimo considerato semplicemente la media degli scarti, cioè $\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})$ avremmo ottenuto sempre 0 indipendentemente da come sono distribuiti gli x_i. Infatti:

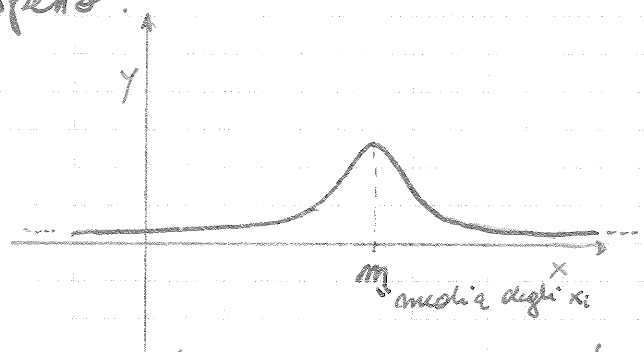
$$\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x}) = \frac{1}{N}\sum_{i=1}^{N}x_i - \frac{1}{N}\sum_{i=1}^{N}\bar{x} = \bar{x} - \frac{1}{N}(N\bar{x}) = 0.$$

La presenza del quadrato nella definizione di σ e σ^2 fa sì che questo fenomeno non si verifichi.

La deviazione standard ha un significato fisico molto preciso nel caso in cui gli x_i siano distribuiti secondo una curva normale.

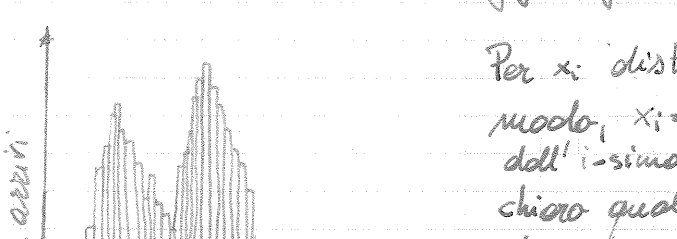
Vediamo di che si tratta:

La curva normale è una curva secondo la quale sono distribuite molte popolazioni in natura. Per esempio si distribuiscono così le misure dirette di una lunghezza assumendo che esse siano sottoposte ad un errore casuale.
Ha una equazione del tipo $y = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$
e ha quest'aspetto:



La variazione standard σ ci dice quanto è "aperta" la campana. Se la x_i è distribuita secondo questa curva ci possiamo aspettare di trovare circa il 65% degli x_i nell'intervallo $[m-\sigma, m+\sigma]$ mentre l'intervallo $[m-2\sigma, m+2\sigma]$ ne conterrà circa il 98%.
Non tutte però le popolazioni sono distribuite in questo modo. Per esempio se gli x_i sono distribuiti come nel seguente esempio il significato di media e deviazione standard è molto sminuito

Esempio 32. Il seguente grafico rappresenta i tempi impiegati dai partecipanti ad una maratona a raggiungere il traguardo.



Per x_i distribuiti in questo modo, x_i = tempo impiegato dall'i-simo concorrente, non è chiaro quale sia l'interpretazione di σ.