Le 64 Pillole

Statistica e altra Matematica

Alessandro Soranzo

Dipartimento di Matematica e Geoscienze

Università degli Studi di Trieste

vers. 1616

Nota 0. SI PREGA DI AVVERTIRE L'AUTORE DI OGNI ERRORE ANCHE SE MINUSCOLO, GRAZIE!

Nota 1. *Le parti di testo scritte in carattere slanted sono da considerare approfondimenti di livello superiore.*

Nota 2. Il testo è ipertestuale. I link $\leftarrow$ $_\downarrow$ $^\uparrow$ $^\nearrow$ rinviano all'Indice Analitico con Approfondimenti, rispettivamente per questioni di livello uguale, inferiore e superiore a quello che è obiettivo di questo testo, e per digressioni interessanti.

Nota 3. Attualmente la separazione della parole in fine riga, fatta automaticamente dall'editor di testo, non corrisponde sempre alla grammatica italiana (ma in questa frase sì).

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua]

Premessa

Si è voluto fare un testo ipertestuale su Statistica e altra Matematica. Soprattutto la Matematica utile per la Statistica.
Grande attenzione è stata posta agli errori più comuni: prevenire è meglio che curare. Esistono errori tipici, contro i quali è utile agire attivamente. Per esempio ritenere che l'arcotangente sia il reciproco della tangente, e invece è l'inversa. Questo errore consueto e altri simili sono dovuti ad ambiguità notazionali della matematica *de facto*, nello stato in cui si trova oggi scritta. Grande rilevanza è stata data a questo aspetto delle ambiguità notazionali. D'altra parte, la situazione oggi è tale, incredibilmente, che nemmeno è chiaro, quando in un testo si trova scritto

$$1{,}265$$

se si intenda quel numero, fra 1 e 2, "uno virgola duecentosessantacinque", oppure "milleduecentosessantacinque"! In questo testo col punto fra le cifre intendiamo il separatore decimale, come si usa nei testi scientifici avanzati e/o internazionali, per esempio $e \approx 2.718 < 3$. Non è un separatore delle migliaia, per favorire la lettura, come altri farebbe. Si noti che molti in Europa scrivono quel numero 2,718 ma la virgola internazionalmente di solito viene usata come separatore delle migliaia, per favorire la lettura.
D'altra parte a cosa servirebbe sapere i più sottili teoremi del calcolo infinitesimale, se poi si permetterà ad Excel, software usatissimo nelle scienze applicate, di calcolare

$$-3\hat{\ }2$$

in modo *erroneo* – o, volendo essere benevoli, attenendosi alla convenzione tutta sua, di quel software, opposta a quella di praticamente tutta la comunità scientifica? E' meglio avvertire di questo lo studente.
Nella simbologia ci si è attenuti in generale allo standard ISO 80000-2:2009[1], per esempio tan e non tang o tg per la funzione

[1] Si veda per esempio https://ja.wikipedia.org/wiki/ISO_80000-2

tangente. E uno spazietto, eventualmente, come separatore delle migliaia: 1 265.
Una certa attenzione è stata rivolta a Wikipedia. Lo studente dovrebbe abituarsi a consultarla. L'idea che non sia una fonte attendibile, per gli argomenti di questo testo elementare è – in generale, ovvio – senz'altro falsa. (Per questioni "sensibili", invece, ogni male si potrebbe dire di Wikipedia: si confrontino per esempio Wikipedia italiana e inglese, su come è iniziata la guerra Iran-Iraq...)
Ripetutamente si accenna a WolframAlpha: lo studente dovrebbe abituarsi a consultare questa potentissima intelligenza artificiale disponibile gratuitamente on-line.

Premesso che non tratteremo degli aspetti *materiali* della raccolta dati (questionari, interviste, eccetera), diciamo subito che la Statistica, così limitata, è una branca della Matematica, astratta come tutta la vera Matematica.
La Statistica verrà suddivisa in Statistica Descrittiva e Statistica Inferenziale.
In qualche modo possiamo dire che tutta la matematica che tratteremo, dai numeri alla Statistica Descrittiva, al Calcolo Infinitesimale e al Calcolo delle Probabilità, è propedeutica per la trattazione della Statistica Inferenziale.
La seconda Parte del corso riguarda le *Matematiche dell'Incertezza*.
Naturalmente, le Matematiche dell'Incertezza non fanno affermazioni incerte, bensì fanno affermazioni certe su fatti incerti, per esempio che conviene scommettere (alla pari) che un dado darà un numero primo piuttosto che un numero quadrato. Nelle *Matematiche della certezza*, d'altra parte, troviamo affermazioni come "la somma degli angoli interni di un triangolo è sempre 180°" e anche "la somma degli angoli interni di un triangolo è sempre minore di 180°", ciascuna *vera* in un diverso sistema assiomatico, senza tema che ne venga qualche antinomia nella realtà sensibile (dove esistono solo approssimazioni di *veri* triangoli, e

soprattutto delle misure dei loro angoli).
Nelle *Matematiche della certezza*, prima Parte del corso, è compresa la Statistica Descrittiva, che costituisce il Capitolo IV, e nelle *Matematiche dell'Incertezza* è compresa la Statistica Inferenziale, Capitolo XII, che in qualche modo rappresenta l'apice della trattazione.

Indice delle 4 Sezioni e dei 12 Capitoli

Segue l'Indice delle 64 lezioni.

Alla fine del testo si trova un capitolo di esercizi e poi l'Indice Analitico con Approfondimenti.

0 Indice delle 64 lezioni

PARTE A — MATEMATICHE DELLA CERTEZZA

Sezione A1 – Matematiche elementari

Sezione A2 – Calcolo infinitesimale

PARTE B — MATEMATICHE DELL'INCERTEZZA

Sezione B1 – Calcolo delle probabilità

PARTE A — MATEMATICHE DELLA CERTEZZA

Sezione A1 – Matematiche Elementari

I – Insiemistica e logica

1 I numeri, dai naturali ai reali

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri* – *work in progress* – continua]

1.1 Zero e successori, $\mathbb{N}$, operazioni, cenni a $\mathbb{Z}$

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri* – *work in progress* – continua]

0 e i suoi successori
Somma di Peano
Prodotto di Peano
Simbolo di sommatoria $\sum$: dati dei numeri a_3, a_4, a_5, la somma

$$a_3 + a_4 + a_5 \quad \text{la scriveremo anche} \quad \sum_{k=3}^{5} a_k$$

e i numeri 3 e 5 possono essere sostituiti da qualsiasi altri: una forma comoda per rappresentare somme di moltissimi numeri.
Da $\mathbb{N}$ in poi esiste nei numeri un *ordinamento* (che sostanzialmente supponiamo noto) per cui dati 2 numeri, o il primo è $<$ del secondo, o $>$ del secondo, o sono uguali (*tricotomia*). In modo ovvio si definiscono il *maggiore o uguale* e il *minore o uguale*.
Da $\mathbb{N}$ in poi si può sommare ad ambo i membri di un'uguaglianza, o di una disuguaglianza, una stessa *quantità* (fissa, cioè un numero, o variabile, come una funzione) conservando l'uguagliaza, o rispettivamente quella stessa disuguaglianza:

$$a < b \;\Leftrightarrow\; a + c < b + c \quad \textit{e analoghe con } \geq, <, \leq, = \;.$$

Volendo trattare la questione in termini elementari, diremo che $a + b = c$ lo scriveremo anche $c - a = b$ e pure $c - b = a$, ottenendosi così l'operazione della *sottrazione*, operazione binaria come

la somma, ma al contrario di quella nè associativa nè commutativa. Per esempio $10 - 3 = 7$ e $10 - 7 = 3$, perchè $3 + 7 = 10$. L'operazione non è sempre definita in $\mathbb{N}$, non esistendo per esempio $3 - 10$ perchè con nessun numero naturale x è $10 + x = 3$.
Risolveremo la questione ampliando $\mathbb{N}$ con $\mathbb{Z}$.
Anche nei successivi ampliamenti a $\mathbb{Q}$ ed $\mathbb{N}$, la sottrazione continuerà come in $\mathbb{Z}$ ad essere operazione binaria ovunque definita, nè associativa nè commutativa, priva di *elemento neutro*↑ (certo $x - 0 = x$ ma non coincide con $0 - x$).
Si noti comunque che da $\mathbb{N}$ in poi vale questa proprietà distributiva

$$(x - y) \cdot z = x \cdot z - y \cdot z \,.$$

1.2 $\mathbb{Z}$ e prodotto dei segni

Per risolvere l'equazione $10 + x = 3$ e altre analoghe, col primo addendo maggiore della somma, che si pongono in $\mathbb{N}$, costruiamo l'insieme $\mathbb{Z}$ dei *numeri interi (relativi).* I suoi elementi sono i numeri di $\mathbb{N}$, che anche potremo scrivere premettendovi un $+$, e anche nuovi elementi, che otteniamo da ciascun elemento di $\mathbb{N}$ premettendovi un segno $-$, e identificando $+0$ cioè 0 con -0:

$$\mathbb{Z} = \{\cdots - 3, -2, -1, 0, +1, +2, +3 \dots\} .$$

Il *valore assoluto* di -3 è 3, e anche di $+3$: $|-3| = |+3| = +3 = 3$. Una trattazione più precisa definisce un numero intero come *classe di equivalenza* di certe *coppie ordinate* di numeri naturali, per esempio $1 = \{(1, 0), (2, 1), (3, 2) \dots\}$, $-1 = \{(0, 1), (1, 2), (2, 3) \dots\}$. Ciò permette di definire l'*opposto* $-(a, b) := (b, a)$, la somma $(a, b) + (a', b') := (a + a', b + b')$, e più complessamente il prodotto. Noi però preferiremo più semplicemente dire che l'opposto di -3 è $+3$ e viceversa; e definire il prodotto come il prodotto in $\mathbb{N}$ dei valori assoluti dei 2 moltiplicandi, col segno $+$, in generale non trascritto, se di segno concorde e col segno $-$ se discorde *(prodotto dei segni*, che poi vale anche in $\mathbb{Q}$ ed $\mathbb{R}$):

$$2 \cdot 4 = 8 \qquad (-2) \cdot 4 = -8 \qquad 2 \cdot (-4) = -8 \qquad (-2) \cdot (-4) = 8 .$$

Per la somma, sommeremo i valori assoluti dei due numeri se di segno concorde e quello stesso segno daremo alla somma; sottrarremo i valori assoluti dei 2 numeri se di segno discorde, e alla somma daremo il segno del maggiore dei 2 in valore assoluto:

$$-2 - 3 = -5 \qquad 2 + (-3) = -1 .$$

Diciamo, semplificatamente, che $2+(-3)$ lo scriveremo in generale $2 - 3$. E col prodotto dei segni avremo la sottrazione:

$$-4 - (-3) = -4 + (+3) = -1 \qquad -4 - 3 = -4 + (-3) = -7 .$$

Da $\mathbb{Z}$ in poi vale questa molteplice relazione fra moltiplicazione e *ordinamento dei numeri*:

$$a > b \Rightarrow \begin{cases} ac > bc & se\ c > 0 \\ ac < bc & se\ c < 0 \end{cases} \qquad a \geq b \Rightarrow \begin{cases} ac \geq bc & se\ c > 0 \\ ac \leq bc & se\ c < 0 \end{cases}$$

cioè moltiplicando ambo i membri di $a < b$ oppure $a \leq b$ per un negativo si inverte l'ordinamento, e (basta scambiare i nomi delle variabili) ciò avviene anche per $>$ e $\geq$. Da $3 > 2$ segue per esempio, moltiplicando per $-10 < 0$, che $-30 < -20$. Invece moltiplicando per un positivo l'ordinamento si conserva: $30 > 20$.
Da $\mathbb{Z}$ in poi si può definire la funzione segno

$$\mathrm{sgn}(x) := \begin{cases} -1 & se\ \ x < 0 \\ 0 & se\ \ x = 0 \\ 1 & se\ \ x > 0 \end{cases}$$

che normalmente viene definita in $\mathbb{R}$ (con qualche ambiguità: per alcuni non è definita in 0); per esempio in $\mathbb{R}$ è $\mathrm{sgn}(3 - \pi) = -1$.
Da $\mathbb{Z}$ in poi si può definire la funzione *valore assoluto*

$$|x| := \begin{cases} x & se\ \ x \geq 0 \\ - & se\ \ x < 0 \end{cases}$$

che normalmente viene definita in $\mathbb{R}$, in vari modi equivalenti↑.

1.3 Divisione euclidea in *Z*

Come ci ricorda la canzoncina *Quarantaquattro gatti*, che si disponevano *in fila per 6*, ma *col resto di 2*, è $44 = 6 \cdot 7 + 2$. Più in generale, come leggiamo in Wikipedia, l'enciclopedia libera, alla voce *Divisione euclidea*

> Dati due numero interi a e b con $b \neq 0$ esiste un'unica coppia di interi q ed r detti *quoziente* e *resto* tali che
> $a = b \times q + r$
> $0 \leq r < |b|$

Diremo a il *dividendo* e b il *divisore*. Scriviamo per massima chiarezza:

$$dividendo = divisore \times quoziente + resto \quad 0 \leq resto < |divisore|.$$

Si noti che il resto è minore del divisore in valore assoluto, ma non necessariamente del quoziente in valore assoluto, per esempio $11 = 4 * 2 + 3$ (ed effettivamente il resto 3 è minore del divisore 4, ma non del quoziente 2), e stiamo dividendo 11 per 4, mentre se dividiamo 11 per 2 abbiamo $11 = 2 * 5 + 1$ (ed effettivamente il resto 1 è minore del divisore 2).
Si noti ancora che tutto ciò vale anche con numeri negativi, per esempio -21 diviso 9 dà quoziente -3 e resto 6, cioè $-21 = 9 \cdot (-3) + 6$, come troviamo subito online con wolframalpha con `Quotient[-21,9]` e `Remainder[-21,9]`.

1.4 Frazioni, Q, e cenni ai numeri decimali periodici

[*in fieri - work in progress* - continua] [*in fieri - work in progress* - continua] [*in fieri – work in progress* – continua]

1.5 L'insieme $\mathbb{R}$ dei numeri reali e i numeri decimali

[*in fieri - work in progress* - continua] [*in fieri - work in progress* - continua] [*in fieri – work in progress* – continua]

2 Logica delle proposizioni

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri – work in progress* – continua]

2.1 Proposizioni; vero e falso

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri – work in progress* – continua]

2.2 Alcuni connettivi logici e le tavole di verità

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri – work in progress* – continua]

2.3 Proposizioni composte

in fieri – continua

2.4 Tautologie, contraddizioni, proposizioni equivalenti

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri – work in progress* – continua]

2.5 Alcune proprietà delle operazioni logiche

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri – work in progress* – continua]

Altre questioni interessanti sono le implicazioni diretta, inversa, contraria, contronominale, e la dimostrazione per assurdo.

3 Prime nozioni sugli insiemi

3.1 Nozioni di insieme, elementi e appartenenza

In questa trattazione elementare, la nozione di *insieme* è supposta nota, ovvero è data come *concetto primitivo.* Possiamo dire che un insieme è, in qualche modo, una *raccolta* di elementi, o *classe* (di elementi). Ma anche la nozione di *elemento* è supposta nota, ovvero è data come *concetto primitivo*, e anche la nozione di *appartenenza* di un elemento ad un insieme.
Agli insiemi *di solito* si dà nome una lettera latina maiuscola, come A e X, ma per alcuni insiemi si usano grafie particolari, per esempio l'insieme dei numeri naturali viene denotato con $\mathbb{N}$.
L'*appartenenza* di un elemento ad un insieme si indica col simbolo $\in$, per esempio con $3 \in \mathbb{N}$, e la sua negazione con $\notin$: $-3 \notin \mathbb{N}$.
Come *variabili* atte a rappresentare un elemento indeterminato di un insieme, *di solito* si usano lettere latine minuscole: a, x...
è ovvio che un insieme si può considerare assegnato se è univocamente *chiaro* se un elemento vi appartiene o no, per esempio le persone *simpatiche* non costituiscono un insieme. La questione della menzionata *chiarezza* non è banale, ma qui non la approfondiremo; si noti comunque che, per esempio, i divisori di $1 + 2017^{2017}$ costituiscono un insieme, nonostante possa essere improbo stabilire se qualche determinato numero vi appartenga.
Un insieme può essere determinato elencandone gli elementi fra parentesi graffe, per esempio $\{a, b, c\}$, eventualmente con puntini di sospensione per indicare elementi non trascritti ma che si suppone il lettore possa capire quali sono, per esempio $\{0, 1, 2, 3...\}$, l'insieme $\mathbb{N}$ dei *numeri naturali.*
Tipici insiemi sono i *singoletti*, come $\{x\}$, $\{Milano\}$, $\{0\}$, eccetera, con un solo elemento.
Oltre che *per elencazione* (*principio di estensione*) gli insiemi si possono rappresentare (*principio di astrazione*) con la *proprietà caratteristica*, cioè *caratterizzante*, degli elementi, per esempio $\{x \in \mathbb{N} | x\, pari\}$.

3.2 Uguaglianza; insieme vuoto; inclusione; partizione

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri* – *work in progress* – continua]

3.3 Cardinalità, insiemi finiti e infiniti

In questa trattazione elementare la *cardinalità* di un insieme finito è il numero dei suoi elementi.
La cardinalità è indicata con $\#E$ (e da altri con card$E$ o $|E|$).
Per esempio $\#\{a, b, c\} = 3$. Si definisce che un insieme è infinito se è in corrispondenza biunivca con un suo sottoinsieme proprio; per esempio $\mathbb{N}$ è in corrispondenza biunivoca con l'insieme dei numeri pari:

$\mathbb{N}$		$\{numeri\, pari\}$
0	$\leftrightarrow$	0
1	$\leftrightarrow$	2
2	$\leftrightarrow$	4
3	$\leftrightarrow$	6
...	$\leftrightarrow$	...

e si dice che un insieme è finito se non è infinito. (Ad un livello più elementare si dice che un insieme è infinito se non ha un numero di elementi (cardinalità) finito).
Due insiemi con la stessa cardinalità si dicono *equipotenti*, oppure *in corrispondenza biunivoca*, oppure, parlando molto semplificatamente, si potrebbe dire che hanno la stessa quantità di elementi.
Fra le proprietà della cardinalità si segnalano in particolare la cardinalità dell'unione, la cardinalità del prodotto cartesiano e la cardinalità dell'insieme delle parti.
[***Questioni più avanzate.***
La cardinalità di un insieme E, finito o infinito (comunque indicata con gli stessi simboli sopra scritti) è la classe di equivalenza di tutti gli insiemi in corrispondenza biunivoca con E.
La cardinalità (infinita) degli insiemi $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$, che sono tutti in corrispondenza biunivoca fra loro, viene denotata $\aleph_0$ (cioè aleph zero*). (Non tutti gli insiemi infiniti hanno la stessa cardinalità e in particolare $\mathbb{R}$ ha cardinalità $\aleph_1 = 2^{\aleph_0}$, cardinalità dell'insieme delle parti di $\mathbb{N}$, maggiore di $\aleph_0$: ci sono più reali che naturali).*]

3.4 Insieme delle parti

L'insieme delle parti $\mathscr{P}(A)$ di un insieme A è l'insieme di tutti i sottoinsiemi di A. (Compresi ovviamente l'insieme vuoto e A stesso). In simboli

$$\mathscr{P}(A) := \{X : X \subseteq A\} \, .$$

L'insieme delle parti di A si chiama anche *insieme potenza* di A.
Trattazione elementare valida per i soli insiemi finiti.
Per esempio se $A := \{a, b, c\}$ allora

$$\mathscr{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\} \, .$$

Se un insieme ha n elementi allora (teorema) il suo insieme delle parti ha 2^n elementi. Cioè, usando il simbolo # della cardinalità, qua nel significato semplice di *numero di elementi*,

$$\#\mathscr{P}(A) = 2^{\#A} \, .$$

Nell'esempio $\mathscr{P}P(A)$ ha 8 elementi, cioè 2^3, perchè A ha 3 elementi.
[***Trattazione ad un livello superiore, per tutti gli insiemi.*** *(La definizione dell'insieme delle parti è ancora quella sopra scritta, e adesso l'insieme A può essere finito o infinito). La cardinalità dell'insieme delle parti di un insieme A è (teorema) 2 elevato alla cardinalità di A:*

$$\#\mathscr{P}(A) = 2^{\#A} \, .$$

(Questo teorema è vero anche per insiemi infiniti ma allora la cardinalità risultante è infinita e non ha più il significato elementare di numero di elementi*, e l'elevamento a potenza non è quello solito dei numeri naturali).*
(Nella teoria assiomatica degli insiemi di Zermelo-Fraenkel, l'esistenza dell'insieme delle parti di qualsiasi insieme è un assioma).]

3.5 Prodotto cartesiano

Si chiama *prodotto cartesiano* di 2 insiemi X e Y l'insieme denotato con $X \times Y$ delle *coppie ordinate* (x, y) con $x \in X$ e $y \in Y$. In simboli

$$X \times Y := \{(x, y) : x \in X, y \in Y\} .$$

L'insieme X può essere finito o infinito, e così pure Y. Se X e Y sono finiti allora (teorema)

$$\#(X \times Y) = (\#X) \cdot (\#Y)$$

cioè la cardinalità del prodotto cartesiano di 2 insiemi è il prodotto delle cardinalità dei 2 insiemi. (Questo teorema è vero anche per insiemi infiniti ma allora la cardinalità risultante è infinita e non ha più il significato elementare qua considerato di *numero di elementi*). Per esempio

$$\{a, b, c\} \times \{\alpha, \beta\} = \{(a, \alpha), (a, \beta), (b, \alpha), (b, \beta), (c, \alpha), (c, \beta)\}$$

ha $2 \cdot 3 = 6$ elementi, che sono coppie ordinate.
Similmente si definisce il prodotto cartesiano di 3 insiemi, composto dalle terne ordinate, e di n elementi, composto dalle n-uple; e si estende il teorema sulle cardinalità.
Esempio. Consideriamo un vecchio linguaggio di programmazione in cui le variabili possono avere un nome composto da una lettera (inglese) maiuscola o da una lettere seguita da una cifra (decimale). Quanti sono i possibili nomi di variabile?
L'insieme dei nomi di variabile di 2 caratteri è in corrispondenza biunivoca col prodotto cartesiano

$$\{A, B, ..., Z\} \times \{0, ..., 9\}$$

che ha $26 \cdot 10 = 260$ elementi. Considerando anche i 26 nomi di variabile di un carattere (lettera maiuscola) si hanno in tutto 286 possibili nomi di variabile.
[*Una definizione rigorosa di coppia ordinata* (x, y) *è* $\{\{x\}, \{x, y\}\}$, *ma in questa trattazione supponiamo di per sè chiaro il concetto.*]

4 Logica dei predicati

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri* – *work in progress* – continua]

4.1 Predicati, insieme di verità, quantificatori

4.2 Implicazione logica ed equivalenza logica

4.3 Regole di negazione

4.4 Teoremi, controesempi, dimostrazioni per assurdo

5 Logica e insiemistica di insiemi e funzioni

[*in fieri* - *work in progress* - continua] [*in fieri* - *work in progress* - continua] [*in fieri* – *work in progress* – continua]

5.1 Operazioni insiemistiche e logiche a confronto

5.2 Proprietà insiemistiche e logiche a confronto

5.3 Funzione, immagine, controimmagine, composta

Ad ogni funzione daremo un nome di 1 lettera, per esempio f oppure y oppure a, e la variabile che varia nel dominio la indicheremo con una lettera qualsiasi, per esempio scriveremo

$$f(x) := x^2$$

e poi sarà per esempio $f(2y) = 4y^2$, e $f(\alpha+1) = \alpha^2 + 2\alpha + 1$ con variabili y e poi α, mentre la f è la funzione di prima. Spesso useremo le variabili n, m, h, k se il dominio è un insieme di numeri interi, per esempio $\mathbb{N}$. In questo caso potremmo dare a tali funzioni nomi $y(n)$ o $a(n)$ ma più spesso useremo le notazioni y_n, a_n e in ogni caso le funzioni definite su $\mathbb{N}$ (o anche su una sua semiretta destra e cioè per $n \geq n_0$) le chiameremo *successioni*, per esempio la *successione di Fibonacci*

$$a_n := \frac{\phi^n - (-\phi)^{-n}}{\sqrt{5}}, \qquad n = 1, 2, ... \quad \phi := \frac{1+\sqrt{5}}{2}$$

(e ϕ è la nota *sezione aurea*), successione più spesso definita *per ricorrenza*: $a_1 := 1$, $a_2 := 1$, $a_n := a_{n-1} + a_{n-2}$, $n = 3, 4, ...$ Valori:

$$1, 1, 2, 3, 5, 8, 13, 21, 55...$$

5.4 Funzione iniettiva, suriettiva, biiettiva, inversa

II – Algebra e piano

6 Algebra dei numeri – prima parte

6.1 Operazioni unarie (interne) dell'algebra dei numeri

Nell'algebra (elementare) dei numeri (naturali, interi, razionali, reali), specializzazione dell'*algebra astratta*↑, le funzioni definite in un insieme numerico e a valori in quello stesso insieme si chiamano *operazioni unarie (interne)*. Da 1 *operando* producono 1 *risultato*. Ne considereremo alcune.

• *Passaggio all'opposto*, $x \mapsto -x$. Operazione unaria definita da $\mathbb{Z}$ in poi. Questo meno non indica affatto negatività ma passaggio all'opposto, per esempio l'opposto di -3 è il positivo 3. (Si noti allora che $-x$ può essere positivo).

• *Passaggio al reciproco*, $x \mapsto \frac{1}{x}$. Notazione deprecabile: x^{-1}. Operazione unaria definita per i numeri diversi da 0 da $\mathbb{Q}$ in poi.

• *Radice quadrata*, $x \mapsto \sqrt{x}$. Operazione unaria definita per **numeri non negativi** ma non tutti: in $\mathbb{R}$ tutti, in $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$ solo alcuni. Per esempio la radice quadrata di 9 esiste in tutti gli insiemi numerici considerati, ed è 3, **nel modo più assoluto la radice quadrata di 9 non è** ± 3. (Come invece affermano testi che seguono una definizione superata di radice quadrata). Invece la radice quadrata di 2 esiste solo in $\mathbb{R}$, ed è indicata con $\sqrt{2}$. Per $x \geq 0$ la $\sqrt{x}$ è definita come il numero $y \geq 0$ tale che $y^2 = x$.

• *Radice quarta, sesta, ottava...* $x \mapsto \sqrt[4]{x}$, $x \mapsto \sqrt[6]{x}$... Con considerazioni analoghe alla radice quadrata. Per esempio la radice quarta di 16 è 2, e quella di 9 esiste solo in $\mathbb{R}$ ed è $\sqrt{3}$.

• *Radice terza, quinta, settima...* $x \mapsto \sqrt[3]{x}$, $x \mapsto \sqrt[5]{x}$... Operazione unaria definita per tutti in numeri di $\mathbb{R}$ e alcuni di $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$. Per esempio la radice cubica di -8 è -2 ovunque, e la radice cubica di 12 è $2\sqrt{3}$ solo in $\mathbb{R}$. La $x \mapsto \sqrt[3]{x}$ è definita come il numero y tale che $y^3 = x$.

• *Valore assoluto*, $x \mapsto |x|$, che è l'opposto di x se x è negativo ed è x altrimenti. Per esempio $|-3| = 3$, $|5| = 5$ e $|0| = 0$.

Al di fuori dell'*algebra dei numeri*, il valore assoluto e le radici verranno sempre chiamate *funzioni*.

6.2 Operazioni binarie (interne) dell'algebra dei numeri

Operazioni binarie (interne) è sinonimo di *funzioni definite sul prodotto cartesiano del dominio per se stesso, a valori nell'insieme stesso*. Da 2 *operandi* producono 1 *risultato*.
Per esempio la somma nei numeri naturali $+ : I\!N \times I\!N \to I\!N$.
Ad un livello elementare possiamo considerare le seguenti 5.

- $+$, *somma* ovvero *addizione*.

- $-$, *differenza* ovvero *sottrazione*. Ad un livello superiore si potrebbe eliminare del tutto la definizione di sottrazione, e considerare $x - y$ come $x + (-y)$, riconducendosi ai soli *opposto*↑(che esiste da $\mathbb{Z}$ in poi) e *somma*, ma in una trattazione elementare conviene considerare la sottrazione come un'operazione a sè stante, coi suoi metodi, anche di calcolo pratico con la notazione posizionale decimale, a mano o con la calcolatrice.
Si calcoli per esempio $2017 - 1005$, che si fa "cifra per cifra", mentre il calcolo di $2017 - 1035$ risulterà meno immediato, richiedendo la tecnica del *riporto*, per la quale rinviamo alle trattazioni elementari.

- $\cdot$, *prodotto*, denotato in diversi contesti $x \cdot y$, $x \times y$, $x * y$, xy. L'ultima notazione purtroppo dà luogo ad un'ambiguità di scrittura: $y(x+1)$ denota sia una funzione y calcolata in $x+1$ che $y \cdot (x+1)$: perciò noi scriveremo sempre quest'ultimo $(x+1)y$.

- $/$, *divisione*, denotato in diversi contesti x/y oppure $x : y$ oppure $\frac{x}{y}$. Ad un livello superiore si potrebbe eliminare del tutto la definizione di divisione, e considerare x/y come $x \cdot \frac{1}{y}$, riconducendosi ai soli *reciproco*↓ (che esiste da $\mathbb{Q}$ in poi) e *prodotto*, ma in una trattazione elementare conviene considerare la divisione come un'operazione a sè stante.
Il secondo operando, *divisore*, dev'essere diverso da 0. Il primo, *dividendo*, può essere 0.

- $^\wedge$, *elevamento a potenza*, denotato $x^\wedge y$ o più spesso x^y.

6.3 Prime proprietà delle operazioni nei numeri

Il $-$ e il $/$ non sono nè associativi nè commutativi per esempio $4/2 = 2 \neq \frac{1}{2} = 1/2$.
Il $+$ e il $\cdot$ sono commutativi e associativi, in tutti gli insiemi numerici (precisazione che in generale ometteremo):

$$x + y = y + x \qquad x\,y = y\,x$$

$$(x + y) + z = x + (y + z) \qquad (x\,y)\,z = x\,(y\,z)$$

(Le parentesi indicano precedenze nel calcolo).
Proprietà distributive:

$$(x + y)\,z = x\,z + y\,z$$

$$(x - y)\,z = x\,z - y\,z$$

$$\forall z \neq 0 \qquad (x + y)/z = x/z + y/z \quad \textit{ossia} \quad \frac{x + y}{z} = \frac{x}{z} + \frac{y}{z}$$

$$\forall z \neq 0 \qquad (x - y)/z = x/z - y/z \quad \textit{ossia} \quad \frac{x - y}{z} = \frac{x}{z} - \frac{y}{z}$$

ma in generale $\frac{x}{y+z}$ è diverso da $\frac{x}{y} + \frac{x}{z}$, e similmente coi $-$.

$$-(-x) = x \qquad e\ \forall\, x \neq 0\ \ \frac{1}{\frac{1}{x}} = x$$

e più completamente

$$\frac{\frac{x}{y}}{\frac{z}{w}} = \frac{x}{y} \cdot \frac{w}{z} \qquad \forall\, y, z, w \neq 0\,.$$

Poi

$$x + (-x) = x - x = 0 \qquad e\,\forall\, x \neq 0\ \ x \cdot \frac{1}{x} = \frac{x}{x} = 1$$

$$x + 0 = 0 + x = x \quad (0\ \textit{elemento neutro rispetto al}\ +)$$

$$x \cdot 1 = 1 \cdot x = x \quad (1\ \textit{elemento neutro rispetto al}\ \cdot)$$

$$x \cdot 0 = 0 \cdot x = 0 \quad (0\ \textit{elemento assorbente rispetto al}\ \cdot)$$

$$\frac{0}{x} = 0\ \ \forall\, x \neq 0 \qquad \frac{x}{0}\quad \textbf{non esiste}$$

7 Algebra dei numeri – con proporzioni e %

7.1 Le potenze

Si definisce $a \in I\!R$ *elevato alla* $n \in I\!N$, $n > 0$, il numero

$$a^n = a \cdot ... \cdot a \quad (n\, volte)$$

e ponendo per $a \neq 0$

$$a^{-n} := \frac{1}{a^n}$$

la potenza è definita per tutti gli esponenti non nulli di $\mathbb{Z}$; e per l'esponente 0 si pone se $a \neq 0$ (rimanendo non definito 0^0)

$$a^0 := 1 \, .$$

Si definisce $a \in I\!R$ *elevato alla* $q = \frac{n}{m} \in Q$, $m, n > 0$, il numero

$$a^{\frac{n}{m}} :=^m\!\sqrt{a^n}$$

che esiste se $a \geq 0$ *vel* n è dispari.
Per la potenza con esponente $q \in I\!R$ si richiede $a \geq 0$, e se q è razionale vale quanto sopra, e se è irrazionale si dà una definizione complessa che si può quasi immaginare sostituendo q con una sua *"straordinariamente buona"* approssimazione razionale: $a^{\sqrt{2}}$ sarà *"circa"* $a^{1.41}$ a sua volta uguale a $\sqrt[100]{a^{141}}$. Ancor meglio $\sqrt[1000]{a^{1414}}$.
L'elevamento a potenza non è nè commutativo, per esempio $2^3 = 8 \neq 9 = 3^2$, nè associativo. Valgono invece le seguenti proprietà.

$$x^{y+z} = x^y \cdot x^z$$

$$(x \cdot y)^\wedge z = (x^\wedge z) \cdot (y^\wedge z) \qquad ossia\ (x\, y)^z = x^z y^z \ \ (distributiva)$$

$$x^{y \cdot z} = (x^y)^z \qquad 1^x = 1 \qquad x^1 = x$$

$$x^0 = 1 \quad \forall\, x \neq 0 \qquad 0^x = 0 \quad \forall\, x \neq 0 \qquad 0^0 \ \textbf{non esiste}$$

Si noti la notazione, ovvero convenzione sulla notazione:

$$usualmente\ scritto\ \ x^{y^z} := x^{(y^z)}\ \ raramente\ scritto$$

per esempio 2^{2^3} è 2^8 cioè 256, non è 4^3 cioè 64, che è $2^{(2^3)}$.

7.2 Proprietà dei radicali ovvero radici

Con una certa attenzione si possono far valere le considerazioni seguenti anche per *certi* numeri di $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, ma l'ambiente migliore per considerare le radici è $\{0\} \cup I\!R^+$, i reali non negativi. Se n è dispari, $n = 1, 3, 5, 7, 9...$, la radice n-esima di x è l'unico y tale che $y^n = x$. Se n è pari positivo, $n = 2, 4, 6, 8...$, la radice n-esima di $x \geq 0$ è l'unico $y \geq 0$ tale che $y^n = x$, e rimarchiamo che **per le radici di indice n pari sia x che $\sqrt[n]{x}$ sono non negativi e senz'ambiguità di segno: $\sqrt{-9}$ non esiste in $\mathbb{R}$, e $\sqrt{9} = 3$ e non ± 3.** L'indice 2 della radice quadrata di solito non viene scritto, cioè si scrive $\sqrt[2]{x}$ ma meglio $\sqrt{x}$.
Le radici possono essere intese come potenze con esponente frazionario e precisamente

$$\sqrt[1]{x} = x \qquad \sqrt{x} = x^{\frac{1}{2}} \qquad \sqrt[3]{x} = x^{\frac{1}{3}} \qquad \sqrt[4]{x} = x^{\frac{1}{4}}...$$

Alcune proprietà delle radici sono le seguenti.

$$x = \left(\sqrt{x}\right)^2 = \left(\sqrt[3]{x}\right)^3 = \left(\sqrt[4]{x}\right)^4...$$

$$\sqrt[n\cdot m]{x} = \sqrt[n]{\sqrt[m]{x}} \quad \textit{per esempio} \quad \sqrt[6]{x} = \sqrt{\sqrt[3]{x}}$$

$$\sqrt[n]{x^\alpha} = (\sqrt[n]{x})^\alpha, \quad x \geq 0,\ \alpha \in I\!R,\ n = 2, 3, 4...$$

$$\sqrt[n]{x} = \sqrt[nm]{x^m}, \quad n, m = 2, 3, 4...$$

Valgono poi con distinzione fra indici ed esponenti pari e dispari:

$$\sqrt{x\,y} = \sqrt{|x|}\,\sqrt{|y|} \quad \textit{e similmente con ogni indice pari}$$

$$\sqrt[3]{x\,y} = \sqrt[3]{x}\,\sqrt[3]{y} \quad \textit{e similmente con ogni indice dispari}$$

$$\sqrt{x/y} = \sqrt{|x|}/\sqrt{|y|} \quad \textit{e similmente con ogni indice pari}$$

$$\sqrt[3]{x/y} = \sqrt[3]{x}/\sqrt[3]{y} \quad \textit{e similmente con ogni indice dispari}$$

$$x = \sqrt[3]{x^3} \quad \textit{e similmente con ogni indice dispari}$$

$$(\textit{non } x) \quad |x| = \sqrt{x^2} \quad \textit{e similmente con ogni indice pari}$$

$$\textit{e anche} \quad x = \operatorname{sgn}(x)\sqrt{x^2} \quad \textit{e similmente con ogni indice pari}$$

Altre proprietà del valore assoluto verranno esposte in seguito.

7.3 Proporzioni; %; ridurre di un quarto e a un quarto

Proporzione. è la *relazione quaternaria* (cioè fra 4 numeri)

$$\frac{x}{y} = \frac{z}{w} \qquad \textit{classicamente scritta} \quad x : y = z : w \, .$$

Non sottolineeremo qua la corrispondenza del concetto con la realtà sensibile, dovuta essenzialmente alla natura del prodotto e della divisione, supponendola nota dagli studi elementari.
Daremo invece 2 esempi.
• In 2 diversi sistemi di misura, l'angolo piatto misura rispettivamente 180° e π radianti. Se per esempio volessimo sapere a quanti gradi corrisponde 1 radiante, avremmo, nel caso generale,

$$\pi_{rad} : 180^\circ = misura_{rad} : misura^\circ$$

e nel caso considerato $misura_{rad} = 1_{rad}$, trovando $1_{rad} \approx 57.3^\circ$.
• L'unico numero $x > 0$ che sia *medio proporzionale* fra 1 e $1 + x$

$$1 : x = x : (1 + x) \qquad \textit{ovvero} \quad \frac{1}{x} = \frac{x}{1+x}$$

è, come si trova risolvendo il sistema $x > 0 \wedge x^2 = 1 + x$,

$$\phi := \frac{1+\sqrt{5}}{2} \approx 1.618 \qquad (\text{mnemonico:} O_1\, numero_6!\, E'_1\, notevole_8.)$$

detto *sezione aurea*. Si noti che $\frac{1}{\phi} = \phi - 1 \approx 0.618$.
Percentuali. Ogni numero ≥ 0 può essere espresso in forma percentuale col simbolo % moltiplicandolo per 100, per esempio $0.125 = 12.5\%$. Per le probabilità, che sono numeri fra 0 e 1, questa è la norma. Ma si può farlo anche per numeri > 1, per esempio 1.4 è 140%. Si noti che portare una quantità al 140% del valore iniziale significa moltiplicarla per 1.4, mentre aumentarla del 140% significa moltiplicarla per 2.4.
Altra questione. Espressioni come *ridotto di un quarto* e *ridotto a un quarto*, corrispondono anch'esse alle 4 operazioni elementari considerate. Se a è la quantità considerata, le quantità ridotte sono rispettivamente $a - \frac{a}{4}$ e $\frac{a}{4}$ (ben diverse nonostante le espressioni verbali riportate, popolarmente confondibili).

7.4 Scrittura dei numeri e approssimazioni

0, raramente scritto $+0$ e talvolta fuorviantemente -0

-4

1265, raramente scritto $+1265$

1.265 che è $\frac{1265}{1000}$, cioè $\frac{253}{200}$

$1.26\overline{5}$ nel senso di 1.2655555555...

$1.2\overline{65}$ nel senso di 1.26565656565...

$\frac{3}{4}$ che a conti fatti è 0.75

$\frac{1}{3}$ che a conti fatti è $0.\overline{3} = 0.33333...$ variamente approssimabile

$\frac{\sqrt{2}}{2}$ variamente approssimabile

$\frac{1+\sqrt{5}}{2}$ variamente approssimabile

$\frac{1}{\log 2}$ variamente approssimabile

Se il risultato di un problema è un'espressione con (sole) frazioni e/o radici e/o altre funzioni elementari, i matematici preferiscono lasciarla stare com'è, ovviamente dopo aver fatto tutte le possibili semplificazioni, cioè – per esempio – mai $\frac{6}{8}$ ma $\frac{3}{4}$, non $\sqrt{12}$ ma $2\sqrt{3}$, non $|-\pi|$ ma π, possibilmente aggiungendo la sua scrittura decimale, finita o periodica o, altrimenti, bene approssimata. E se è una probabilità, in generale aggiungono anche la sua scrittura come percentuale. Per esempio scriveranno

$\frac{3}{4} = 0.75 = 75\%$

$\frac{1}{3} = 0.\overline{3} \approx 0.3333 = 33.33\%$, o spesso $\frac{1}{3} = 0.\overline{3} \approx 0.333 = 33.3\%$

$\frac{\sqrt{2}}{2} \approx 0.7071 = 70.71\%$, o spesso $\frac{\sqrt{2}}{2} \approx 0.707 = 70.7\%$

$\frac{1}{\log 2} \approx 1.442695$ o spesso $\frac{1}{\log 2} \approx 1.4427$, o perfino $\frac{1}{\log 2} \approx 1.443$

$2\pi = 6.2831853... \approx 6.283185 \approx 6.28319 \approx 6.2832 \approx 6.283 \approx 6.28.$

Invece nelle scienze applicate si preferisce evitare le frazioni, i decimali periodici, le radici e le altre funzioni elementari, e tutto si calcola esattamente, come nel caso di $\frac{3}{4}$, o si approssima come negli altri.

8 Il piano euclideo

8.1 Retta, cerchio e altre figure, e misura in radianti

Consideremo *concetti primitivi* il *punto* e anche:
la *retta*, che viene divisa in 2 *semirette* da 1 punto *origine*;
la *retta orientata*, dove i punti si *precedono* o *seguono* fra loro;
il *piano (euclideo)*; che una retta *origine* divide in 2 *semipiani*;
il *segmento* di *estremi* 2 punti e *lunghezza* la loro distanza;
per semplicitị trattazione, anche la *curva*.

Dati un punto P e numero $r > 0$, l'insieme dei punti che hanno distanza r da C si chiama *circolo* – da altri detto *cerchio* – di centro P e raggio r. Si noti che è in un certo senso "vuoto": il centro non appartiene al circolo. In questa trattazione chiameremo *cerchio* l'insieme dei punti che hanno distanza $\leq r$ da C (cioè il "cerchio pieno" per così dire). Un segmento nel cerchio lungo $2r$ è un *diametro* . Un diametro divide il cerchio in 2 *semicerchi*. Due semirette diverse con stessa origine P dividono il piano in 2 *angoli* di *vertice* P. Se essi sono uguali si dicono *piatti*. Se uno è "minore" dell'altro si dice *convesso* e l'altro *concavo*. Due rette del piano con intersezione vuota le diremo *parallele*. Se l'intersezione ha esattamente 1 punto si dicono *incidenti*. Se l'intersezione coincide con le rette stesse, si tratta di 2 rette *coincidenti* (talvolta dette *parallele coincidenti*): in effetti è 1 sola retta considerata 2 volte. Due rette incidenti possono formare 4 angoli uguali che si dicono *retti*, e le rette si dicono *ortogonali* o *perpendicolari*. Angoli "minori" di un angolo retto si chiamano *acuti* e angoli "maggiori" di un angolo retto e "minori" di un angolo piatto si dicono *ottusi*. L'intersezione di un angolo di vertice P con un circolo di centro P è un *arco di circolo*, di cui si suppone nota come concetto primitivo la nozione di *lunghezza*. Se il raggio del circolo è 1 quella lunghezza si chiama *misura in radianti* dell'angolo. La misura dell'angolo piatto (≈ 3.14) si indica con π. Si dice *figura* ogni insieme di punti: segmenti, rette, semirette, circoli, cerchi, semicerchi, sempiani, angoli...

8.2 Limitato, convesso, poligoni vari, perimetro, area

Una figura si dice *limitata* se esiste un cerchio che la contiene (per esempio i segmenti), altrimenti *illimitata* (per esempio le rette).
Una figura si dice *convessa* se contiene ogni segmento con estremi in quella figura (per esempio il cerchio), altrimenti non convessa.
L'unione di 2 segmenti AB e BC (da altri denotati $[AB]$ e $[BC]$) non *allineati*, con un estremo B in comune si chiama *linea spezzata* di 2 *lati* AB e BC e *vertice* B, e in modo analogo è definita quella di più *lati*. Non facciamo i dettagli, come pure per le seguenti intuitive denominazioni:

linea spezzata *chiusa*, o altrimenti *aperta*;
linea spezzata *intrecciata*, o altrimenti non intrecciata;
lati *consecutivi*;
lunghezza della linea spezzata.

Supponiamo noto il fatto (non banale da dimostrare) che una linea spezzata chiusa non intrecciata divide il piano in 2 parti. Quella limitata si chiama *poligono (semplice)* (o *non intrecciato*), coi lati e i vertici della linea spezzata, e *perimetro* la sua lunghezza. Il poligono si dice *equilatero* se ha i lati uguali ed *equiangolo* se ha gli angoli uguali, e *regolare* se è equilatero ed equiangolo.
I poligoni con 3 lati si chiamano *triangoli* e quelli con 4 *quadrilateri*. Un quadrilatero si chiama *rombo* se è equilatero, *parallelogramma* se ha i lati a 2 paralleli, *rettangolo* se ha i lati a 2 a 2 perpendicolari, *quadrato* se è equilatero equiangolo.
Il prodotto $b \cdot h$ delle lunghezze di 2 lati consecutivi di un rettangolo si chiama *area*. Ciò genera immediatamente il concetto di area di un parallelogramma, e per dimezzamento quello di area di un triangolo, $\frac{b \cdot h}{2}$, che a sua volta genera il concetto di area per ogni poligono, e con *passaggio al limite* (concetto non banale) l'area di ogni figura *sufficientemente regolare*, in particolare il cerchio, che si dimostra avere area πr^2. I punti, i segmenti, le linee spezzate, i circoli e gli archi di circolo, e perfino le rette hanno area 0. Non esiste (oppure è *infinita*) l'area di piano, semipiani e angoli.

8.3 Congruenti e simili; Teoremi di Pitagora ed Euclide

Se per 2 figure F ed F' esiste una funzione biiettiva $f : F \to F'$ che conserva le distanze di tutte le coppie di punti corrispondenti

$$\overline{f(P)\, f(Q)} = \overline{P\, Q} \qquad \forall\, P, Q \in F$$

le 2 figure si dicono *congruenti* e si possono immaginare come *sovrapponibili*; ma si faccia attenzione che l'azione della sovrapposizione può necessitare di passare momentaneamente nello *spazio* in cui è contenuto il piano, come per i grafemi d e b. Tutte le rette sono congruenti, e anche le semirette e i semipiani.
Se F ed F' sono il piano euclideo stesso, f si chiama *isometria*↑.
Se per 2 figure F ed F' esistono un numero $\alpha > 0$ e una funzione biiettiva $g : F \to F'$ che moltiplica per α le distanze di tutte le coppie di punti corrispondenti ovvero precisamente

$$\overline{f(P)\, f(Q)} = \alpha \cdot \overline{P\, Q} \qquad \forall\, P, Q \in F$$

le 2 figure si dicono *simili*. Tutti i segmenti sono simili, e anche i circoli, i cerchi, i quadrati e i triangoli equilateri.
Si dimostra che se un triangolo ha gli angoli uguali a quelli di un altro, essi sono simili. Allora dette a, b, c le misure dei lati del primo, e a', b', c' quelle del secondo, vale la (doppia) proporzione

$$a : a' = b : b' = c : c' \qquad \textit{ossia} \;\; \frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}\,.$$

Un triangolo con un angolo retto si chiama *triangolo rettangolo*; *ipotenusa* il suo lato maggiore e *cateti* gli altri. Dette nell'ordine i, c_1 e c_2 le lunghezze di quei lati, h quella della distanza dell'ipotenusa dal vertice *opposto* (cioè h è la lunghezza dell'*altezza relativa all'ipotenusa*), e p_1 e p_2 le lunghezze delle *proiezioni dei cateti sull'ipotenusa* (nell'ordine), valgono questi 3 teoremi classici:

$$i^2 = c_1^2 + c_2^2 \qquad \textit{Teorema di Pitagora}$$

$$c_1^2 = h \cdot p_1, \quad c_2^2 = h \cdot p_2 \qquad \textit{Primo Teorema di Euclide}$$

$$p_1 : h = h : p_2 \qquad \textit{Secondo Teorema di Euclide.}$$

9 Piano cartesiano: punti e rette

Assi cartesiani e origine
coordinate del punto
distanza di 2 punti
orientazione del piano e degli angoli con vertice nell'origine
rette orizzontali, verticali, oblique
equazioni esplicita della retta
equazione implicita della retta
retta per 2 punti
distanza punto retta
rette parallele
rette perpendicolari
circolo
ellisse
iperbole
La parabola verrà trattata in seguito.

10 Piano cartesiano: coniche e altre figure

Premessa definizionale:

funzione: $f(x)$, p.es. $f(x) := x^2 - 2$, spesso scritta $y = x^2 - 2$;

equazione: $f(x) = g(x)$, p.es. $x^2 - 2 = 0$ ha *soluzione* $\pm\sqrt{2}$;

polinomio: $a_n x^n + ... + a_1 x + a_0$, p.es. $x^2 - 2$, ha *radici* $\pm\sqrt{2}$;

disequazione in 1 variabile: $f(x) > g(x)$ (o $<$ o $\geq$ o $\leq$), p.es. $x^2 > 2$, ha *soluzione* $x < -\sqrt{2} \vee x > \sqrt{2}$.

Rappresentazione cartesiana esplicita di funzioni e disequazioni

- retta
- altre funzioni

Rappresentazione cartesiana implicita di funzioni e disequazioni ed eventuale esplicitazione

- retta
- circolo, ellisse, iperboli

L'insieme dei punti per i quali è costante la distanza da un punto – detto *fuoco* – e da una retta – detta *direttrice* – si chiama *parabola*. La retta perpendicolare alla direttrice contenente il fuoco si chiama *asse* ed è asse di simmetria, e interseca la parabola nel *vertice*.

- folium di Cartesio

11 Valore assoluto; dis/equazioni; sistemi

11.1 Valore assoluto, e relative dis/equazioni

Il valore assoluto è una *funzione elementare* (già introdotta in Algebra come operazione unaria) definita da $\mathbb{Z}$ in poi$^{\uparrow}$.
Il valore assoluto verifica molte proprietà oltre a $|x| \geq 0$, fra cui

$$|x| = 0 \Leftrightarrow x = 0 \qquad e\, allora \quad |f(x)| = 0 \Leftrightarrow f(x) = 0$$

$$|x| = a \geq 0 \Leftrightarrow x = \pm a \qquad nel\ senso: \quad x = -a \vee x = a$$

$$|-x| = |x|$$

$$||x|| = |x|$$

$$|x| = \sqrt{x^2}$$

$$|x| = x \cdot \mathrm{sgn}(x)$$

$$x = |x| \cdot \mathrm{sgn}(x)$$

$$|x \cdot y| = |x| \cdot |y|$$

$$|x/y| = |x|/|y|, \qquad y \neq 0$$

$$|x^n| = |x|^n \qquad \forall\, x \in I\!R,\ \ \forall\, n \in \mathbb{Z}$$

$$|x^\alpha| = |x|^\alpha \qquad \forall\, x \in I\!R^+,\ \ \forall\, \alpha \in I\!R$$

$$|x + y| \leq |x| + |y| \quad disuguaglianza\, triangolare$$

$$\forall\, a \geq 0 \quad |f(x)| > a \Leftrightarrow f(x) < -a \vee f(x) > a \quad e\, analoga\, con\ \geq$$

$$\forall\, a > 0 \quad |x| < a \ \Leftrightarrow\ -a < x < a \qquad e\, anche\, con\ \leq\ .$$

La prima permette di risolvere alcune equazioni, e le ultime 2, duplici, alcune disequazioni. Dall'ultima, p. es. col $\leq$, si ha $\Big(\forall\, a > 0\Big) \quad |f(x)| \leq a \ \Leftrightarrow\ -a \leq f(x) \leq a$ che verrà in generale scritta con la notazione della parentesi graffa grande che vale *et*

$$\Big(\forall\, a > 0\Big) \quad |f(x)| \leq a \ \Leftrightarrow \begin{cases} f(x) \geq -a \\ f(x) \leq a \end{cases}$$

e l'ultimo si chiama *sistema*, in questo caso di disequazioni.
Le 5 dis/equazioni $|f(x)| \lesseqqgtr g(x)$ verranno considerate in seguito.

11.2 Sistemi di dis/equazioni e sistemi generalizzati

Un **sistema di n equazioni** in 1 incognita x, con la parentesi graffa (grande) che vale *et*, è il predicato (con opportune funzioni)

$$\begin{cases} f_1(x) = g_1(x) \\ \dots \\ f_n(x) = g_n(x) \end{cases}$$

e sostituendo gli = con segni di disuguaglianza si ha un **sistema di n disequazioni** in 1 incognita. Naturalmente possono considerarsi sistemi con 2 o più incognite, e anche sia con equazioni che disequazioni. (I predicati del tipo $f(x) \neq g(x)$, che potremmo chiamare *"inequazioni"*, non li nomineremo per nome considerandoli implicitamente compresi nelle disequazioni data l'equivalenza con $(f(x) - g(x))^2 > 0$). Per esempio questo **sistema di equazioni e disequazioni** in 2 incognite

$$\begin{cases} x\,y = 0 & \leftarrow \textit{rappresenta}\ 2\ \textit{rette} \\ x^2 + y^2 - 2x - 2y + 1 \leq 0 & \leftarrow \textit{rappresenta}\ 1\ \textit{cerchio} \end{cases}$$

$$\text{equivale a} \quad \begin{cases} x = 0 \vee y = 0 \\ (x-1)^2 + (y-1)^2 \leq 1 \end{cases} \qquad (*)$$

$$\Leftrightarrow \quad \begin{cases} x = 0 \\ (0-1)^2 + (y-1)^2 \leq 1 \end{cases} \vee \begin{cases} y = 0 \\ (x-1)^2 + (0-1)^2 \leq 1 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} x = 0 \\ (y-1)^2 \leq 0 \end{cases} \vee \begin{cases} y = 0 \\ (x-1)^2 \leq 0 \end{cases} \qquad (**)$$

ha soluzione, come anche si vedeva subito graficamente,

$$\Big(x = 0 \wedge y = 1\Big) \vee \Big(x = 1 \wedge y = 0\Big) .$$

Predicati come $(*)$ e $(**)$, ammettenti anche il *vel*, li chiameremo *sistemi generalizzati di equazioni e, in questo caso, disequazioni.* Ed ecco uno per la risoluzione di dis/equazioni col valore assoluto:

$$|f(x)| = g(x) \ \Leftrightarrow \ \begin{cases} f(x) \geq 0 \\ f(x) = g(x) \end{cases} \vee \begin{cases} f(x) < 0 \\ -f(x) = g(x) \end{cases}$$

e i 3 = possono sostituirsi con tutti e 4 i segni di disuguaglianza.

11.3 Funzioni e dis/equazioni di primo grado

Per ogni m la funzione $f(x) := m\,x$ si chiama *funzione lineare*, e in questo contesto la scriveremo $y = m\,x$. è una funzione crescente se $m > 0$, e *decrescente* se $m < 0$, *costantemente nulla* se $m = 0$. Il grafico è una retta passante per l'origine.
Fissato $m \neq 0$, l'equazione $m\,x = 0$ ha soluzione $x = 0$ (basta dividere per $m \neq 0$), mentre la disequazione in 1 variabile

$$m\,x > 0$$

si risolve dividendo per m ciò che, se e solo se $m < 0$, inverte l'ordinamento. Allo stesso modo si risolve se si aveva $\geq$, $<$ o $\leq$.
Per ogni $m, q \in \mathbb{R}$ la funzione $y = m\,x + q$ si chiama funzione *affine*, deprecabilmente detta lineare. è una funzione crescente se $m > 0$, e *decrescente* se $m < 0$, *costante* se $m = 0$. Il grafico è una retta, che *intercetta* (interseca) l'asse y in $(0, q)$.
Fissati $m \neq 0$ e q, l'equazione

$$m\,x + q = 0$$

ha soluzione $x = -\frac{q}{m}$. (Si sommi $-q$ e si divida per $m \neq 0$).
Le 4 disequazioni con $>$, $\geq$, $<$, $\leq$ si risolvono sommando $-q$ e poi dividendo per m invertendo l'ordinamento se $m < 0$.
Fissati m e q, la disequazione in 2 variabili

$$y \geq m\,x + q$$

rappresenta il *semipiano chiuso* "sopra" la retta $y = m\,x + q$, compresi i punti della retta. Col $>$, il *semipiano aperto*, esclusi i punti della retta.
Con $\leq$, e con $<$, si va "sotto" la retta, compresa (semipiano chiuso) o rispettivamente esclusa (semipiano aperto).
Ciò vale anche per la $y > m\,x$ e le 3 analoghe, che hanno $q = 0$.

III – Funzioni elementari, con equazioni e disequazioni

12 Funzioni e dis/equazioni di secondo grado

12.1 Funzioni e dis/equazioni di secondo grado

Le parabole con asse verticale hanno equazione

$$y = a\,x^2 + b\,x + c$$

con qualche $a, b, c \in I\!R$, e il loro asse ha equazione $x = \frac{b}{2a}$.
Definito il *discriminante* $\Delta := b^2 - 4ac$, la parabola interseca l'asse x in $x_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a}$ se $\Delta \geq 0$ e altrimenti mai. Se x_1 e x_2 sono le *radici* (eventualmente coincidenti) del polinomio $a\,x^2 + b\,x + c$, ovvero le *soluzioni* dell'equazione $a\,x^2 + b\,x + c = 0$, è

$$a\,x^2 + b\,x + c \equiv a(x - x_1)(x - x_2)\,.$$

Con uno *schema di prodotto dei segni* risolviamo la disequazione

$$a\,x^2 + b\,x + c > 0 \qquad \textit{oppure} \ \geq \quad \textit{oppure} \ < \quad \textit{oppure} \ \leq$$

Per esempio per $-x^2 + 2\,x + 8 \leq 0$ usando la *formula ridotta*

$$x_{1,2} = \frac{-\frac{b}{2} \pm \sqrt{\frac{\Delta}{4}}}{a} \qquad\qquad \frac{\Delta}{4} = \left(\frac{b}{2}\right)^2 - a\,c$$

utile se b è intero pari, si trova

$$\Delta = 1^2 - (-1)\cdot 8 = 1 + 8 = 9 > 0$$

$$x_{1,2} = \frac{-1 \pm 3}{-1} \qquad x_1 = -4 \quad x_2 = 2$$

$$-x^2 + 2\,x + 8 = -1\cdot(x+4)\cdot(x-2) \leq 0$$

$-1 > 0 \quad \textit{mai}$
$x + 4 > 0 \Leftrightarrow x > -4$
$x - 2 > 0 \Leftrightarrow x > 2 \qquad\qquad \textit{Soluzione}: \ x \leq -4 \vee x \geq 2\,.$

è meglio risolvere le 3 disequazioni con $>$ e poi individuare l'insieme soluzione a seconda se nella disequazione iniziale si aveva $>$, $\geq$, $<$ o $\leq$. (Altri Autori fanno diversamente). Questo è un procedimento generale che in seguito varrà anche per più di 3 termini.

13 Funzioni e dis/equazioni ir/razionali

13.1 Funzioni e dis/equazioni razionali intere facili

Limitandosi ad 1 variabile, un'**equazione razionale intera** è il predicato che uguaglia 2 polinomi in 1 variabile $P_1(x) = P_2(x)$ ovvero dopo le opportune riduzioni, che uguaglia un polinomio (in 1 variabile) a 0:

$$P(x) = 0$$

e la **disequazione razionale intera** si ottiene sostituendo il segno di uguaglianza con uno di disuguaglianza, per esempio

$$x^2 + 5x + 6 \leq 0\,.$$

Abbiamo visto come risolvere equazioni e disequazioni razionali intere di primo e secondo grado.
Per quelle di grado superiore, consideriamo per esempio

$$x^4(-x^2+2x+8)(2-x)(1-x+x^2) = 0 \qquad \textit{o, disequazione}, \leq 0\,.$$

Per risolvere *fattorizziamo* il polinomio in fattori di primo grado e/o di secondo grado *irriducibili* ovvero con discriminante negativo.
Il monomio x^4 si fattorizza in $x \cdot x \cdot x \cdot x$ ma in effetti i fattori che sono monomi x^n conviene lasciarli stare come sono. Il polinomio di secondo grado $-x^2 + 2x + 8$, con discriminante positivo, abbiamo visto che è $-1 \cdot (x+4)(x-2)$. Il fattore di primo grado $2-x$ è già a posto. E $1 - x + x^2$ è irriducibile avendo discriminante negativo, e allora non ha zeri. Si ha allora la fattorizzazione

$$x^4(-1)(x+4)(x-2)(2-x)(1-x+x^2)\,.$$

L'equazione con $= 0$ equivale, considerato che $x^4 = 0 \Leftrightarrow x = 0$,

$$x = 0 \;\vee\; x+4 = 0 \;\vee\; x-2 = 0 \;\vee\; 2-x = 0$$

che dà $x \in \{-4, 0, 2\}$. La disequazione si fa con lo schema di prodotto dei segni visto in precedenza trovandosi (per ≤ 0) la soluzione $x \leq -4 \vee x = 0 \vee x = 2$. Per > 0 si troverebbe $-4 < x < 0 \vee 0 < x < 2 \vee x > 2$.

13.2 Funzioni e dis/equazioni razionali intere difficili

L'equazione e le 4 disequazioni

$$x^9 - x^8 - 11x^7 + 28x^6 - 28x^5 + 16x^4 \quad = 0,\ > 0,\ \geq 0,\ < 0, \leq 0$$

si risolvono come prima una volta che si riconoscesse che questo è proprio il polinomio di prima, ma ora è molto lontano dalla fattorizzazione. Il problema di fattorizzare un polinomio può essere facile, come sarà in questo caso, o difficile, o impossibile.
Prima di tutto *raccogliamo* il fattore x^4, subito visto:

$$x^4\,(x^5 - x^4 - 11x^3 + 28x^2 - 28x + 16)$$

e ci resta da fattorizzare il polinomio $P_5(x)$ di $5^\wedge$ grado con *termine costante* non nullo. Cercheremo *solo* fattori $x - m$ con m un divisore intero del termine costante: ±1, ±2, ±4, ±8, ±16. Allora basta cercare un numero m fra quelli, il quale annulli il polinomio. è $P_5(2) = 0$ e allora dividiamo per $x - 2$ con la Regola di Ruffini:

i coefficienti $\rightarrow$	$+1$	-1	-11	$+28$	-28	$+16$
la radice $\rightarrow$ $+2$	$\downarrow$	$+2$	$+2$	-18	$+20$	-16
	$+1$	$+1$	-9	$+10$	-8	0 $\leftarrow$ sempre così

che dà i coefficienti del quoziente, e allora la fattorizzazione

$$(x-2)(x^4 + x^3 - 9x^2 + 10x - 8)$$

e riprendendo il monomio x^4

$$x^4(x-2)(x^4 + x^3 - 9x^2 + 10x - 8)\ .$$

Si continua cercando un fattore $x - m'$ con m' fra i divisori interi di -8, e si trova che 2 annulla il polinomio di quarto grado $P_4(x)$, e allora lo si divide per $x - 2$ con la Regola di Ruffini trovando

$$x^4(x-2)(x-2)(x^3 + 3x^2 - 3x + 4)$$

poi cercando fra i divisori interi di 4 si trova l'annullamento di $P_3(x)$ in -4 da cui con la Regola di Ruffini applicata a $P_3(x)$

$$x^4(x-2)(x-2)(x+4)(x^2 - x + 1)$$

che è come prima (un fattore è opposto ma congloba il (-1)).

13.3 Funzioni e dis/equazioni razionali fratte

Per risolvere l'**equazione razionale fratta**

$$\frac{N(x)}{Dx)} = 0$$

dobbiamo, coi metodi visti prima, fattorizzare il numeratore in fattori di primo o di secondo grado. Di essi sappiamo trovare gli eventuali zeri, che tutti insieme sono gli zeri del numeratore.
La soluzione dell'equazione è l'insieme di tutti gli zeri del numeratore che non sono zeri del denominatore (essi darebbero infatti l'espressione priva di significato $\frac{0}{0}$), e allora ogni zero del numeratore si dovrà verificare se annulla il denominatore, e in quel caso escluderlo.
Per risolvere le 4 **disequazioni razionali fratte**

$$\frac{N(x)}{D(x)} > 0, \geq 0, < 0, \leq 0$$

dobbiamo, coi metodi visti prima, fattorizzare il numeratore e il denominatore in fattori di primo grado e/o di secondo grado irriducibili. Si risolvono le corrispondenti disequazioni col > 0, e si fa uno schema di prodotto dei segni, con la seguente avvertenza. Mentre per gli intervalli fra i capisaldi non ci sono dubbi fra + e −, derivando dal banale prodotto dei segni, nei capisaldi possono verificarsi 2 casi: o 0, o non esiste.

13.4 Funzioni e dis/equazioni irrazionali

Per un'attenta considerazione della definizione della radice quadrata

$$\sqrt{f(x)} = g(x) \Leftrightarrow \begin{cases} g(x) \geq 0 \\ \\ f(x) = g^2(x) \end{cases}$$

e similmente con ogni indice pari. Invece con 3 e ogni dispari

$$\sqrt[3]{f(x)} = g(x) \Leftrightarrow f(x) = g^3(x) \,.$$

• Ancora, elevando ambo i membri al cubo, il che conserva l'ordinamento perchè t^3 è crescente, si ha questa, e tutte le analoghe con gli altri 3 segni di disuguagliaza,

$$\sqrt[3]{f(x)} > g(x) \Leftrightarrow f(x) > g^3(x)$$

estendibile sostituendo il 3 con ogni indice ed esponente dispari.
• Tutti gli indici pari (radici quadrate, quarte, seste...) si comportano in modo analogo a quello che ora scriveremo solo per il 2 (radici quadrate) ma il $<$, il $\leq$, il $>$ e il $\geq$ danno 4 casi diversi:

2 casi **"radice minore"** (che sono più semplici)

$$\sqrt{f(x)} < g(x) \;\Leftrightarrow\; \begin{cases} f(x) \geq 0 \\ g(x) > 0 \\ f(x) < g^2(x) \end{cases}$$

$$\sqrt{f(x)} \leq g(x) \;\Leftrightarrow\; \begin{cases} f(x) \geq 0 \\ g(x) \geq 0 \\ f(x) \leq g^2(x) \end{cases}$$

2 casi **"radice maggiore"** (che sono più complicati)

$$\sqrt{f(x)} > g(x) \;\Leftrightarrow\; \begin{cases} g(x) \geq 0 \\ f(x) > g^2(x) \end{cases} \vee \begin{cases} g(x) < 0 \\ f(x) \geq 0 \end{cases}$$

$$\sqrt{f(x)} \geq g(x) \;\Leftrightarrow\; \begin{cases} g(x) \geq 0 \\ f(x) \geq g^2(x) \end{cases} \vee \begin{cases} g(x) < 0 \\ f(x) \geq 0 \end{cases}$$

Si risolva per $_\mu$esempio: $\sqrt{x^2-9} \geq x-1$; $_\mu$e $x > \pi + \sqrt{x^2-10}$.

14 Le funzioni esponenziali e logaritmiche

14.1 Logaritmi naturali e decimali; esponenziale

Il logaritmo in base b di x è l'esponente da dare a b per avere x. Ecco due esempi:

$$\log_2 8 = 3 \Leftarrow 2^3 = 8 \qquad \log_{10} 100 = 2 \Leftarrow 10^2 = 100.$$

(Ma) dev'essere $x > 0$, e $0 < b \neq 1$. Considereremo quasi solo 2 basi: di grande importanza applicativa per secoli, e ancor oggi usato in Chimica, il *logaritmo decimale*, $\log_{10}$ ovvero lg, di cui si stampavano *tavole* numeriche; oggi in Matematica e Fisica si preferisce un'altra base: detta e, numero di Nepero, la "somma infinita" (serie) dei reciproci dei fattoriali dei numeri naturali

$$\mathrm{e} := \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots \approx 2.71828^{\nearrow}$$

e il logaritmo in base e si chiama *logaritmo naturale* e si denota

$$\ln x \qquad (\textit{oppure} \ \log_e x \qquad \textit{oppure} \ \log x)\ .$$

Tutti i logaritmi in 1 valgono 0. I logaritmi in base $b > 1$ sono funzioni crescenti ("verso $+\infty$"), e in base $0 < b < 1$ decrescenti ("verso $-\infty$"). In ogni caso il crescere in valore assoluto avviene con straordinaria lentezza, per esempio $\log_{10} 1000 = 3$. I limiti in 0 sono rispettivamente $-\infty$ e $+\infty$. In ogni caso è $\log_b : \mathbb{R}^+ \to \mathbb{R}$. (Per ogni possibile base b) il logaritmo è una funzione biiettiva, e l'inversa del logaritmo naturale si chiama *(funzione) esponenziale*, ed è definita su tutto $\mathbb{R}$, cioè $\exp : \mathbb{R} \to \mathbb{R}$:

$$\exp := \ln^{-1} \qquad \ln := \exp^{-1}\ .$$

Questi $^{-1}$ denotano l'inversa, assolutamente non il reciproco. L'esponenziale vale 1 in 0, ed è crescente "verso $+\infty$" con grande rapidità, per esempio, $\exp(20)$ è più o meno mezzo miliardo.
Un'attenta considerazione di quanto detto finora dà $\exp x = \mathrm{e}^x$, e

$$\ln \exp x = \ln \mathrm{e}^x = x,\ \forall x \qquad \exp \ln x = \mathrm{e}^{\ln x} = x,\ \forall x > 0$$

$$0 < b \neq 1 : \qquad \log_b b^x = x,\ \forall x \qquad b^{\log_b x} = x,\ \forall x > 0.$$

14.2 Proprietà dell'esponenziale e dei logaritmi

$$(\forall\, b > 0, b \neq 1)\ \log_b 1 = 0 \qquad \exp 0 = e^0 = 1 = b^0,\ \forall\, b > 0$$

$$(\forall\, x > 0)\, \log_b x^\alpha = \alpha\, \log_b x \ \Rightarrow\ \ln x^\alpha = \alpha \ln x,\ \log_{10} x^\alpha = \alpha \log_{10} x$$

$$\text{in particolare}\ \ \log_b \frac{1}{x} = -\log_b x\,, \qquad \log_b \sqrt[n]{x} = \frac{1}{n} \log_b x\,.$$

Formule per il *logaritmo del prodotto* ed *esponenziale della somma*:

$$\ln(x\, y) = \ln x + \ln y,\ \ \forall\, x, y > 0 \qquad\qquad \mathrm{e}^{x+y} = \mathrm{e}^x\, \mathrm{e}^y$$

$$\log_b(x\, y) = \log_b x + \log_b y,\ \ \forall\, x, y > 0 \qquad\qquad b^{x+y} = b^x\, b^y$$

e per il *logaritmo del quoziente* ed *esponenziale della differenza*:

$$\ln(x/y) = \ln x - \ln y,\ \ \forall\, x, y > 0 \qquad\qquad \mathrm{e}^{x-y} = \mathrm{e}^x / e^y$$

$$\log_b(x/y) = \log_b x - \log_b y,\ \ \forall\, x, y > 0 \qquad\qquad b^{x-y} = b^x / b^y\,.$$

Queste 2 formule approssimate consentono le 2 conversioni fra logaritmi naturali e decimali *[con circa $\frac{1}{79\,000}$ di errore relativo]*:

$$\log_{10} x \approx 0.4343 \ln x \qquad {\scriptstyle \textit{e reciprocamente}} \quad \ln x \approx \frac{\log_{10} x}{0.4343}.$$

Ecco due valori che dobbiamo imparare a memoria:

$$\log_{10} \approx \mathbf{0.3} \qquad [\textit{o meglio}\ \approx 0.301]$$

$$\ln 10 \approx \mathbf{2.3} \qquad [\textit{o meglio}\ \ \approx \frac{1}{0.4343} \approx 2.303]$$

I numeri in grassetto han errore relativo in valore assoluto $< 0.5\%$.
[Altre formule, più avanzate: formule di cambiamento di base:

$$\log_b x = \frac{\ln x}{\ln b} = \frac{\log_c x}{\log_c b} = (\log_c x) \cdot (\log_b c)$$

$$\textit{e in particolare} \quad \log_b x = \frac{1}{\log_x b} \qquad \log_{1/b} x = -\log_b x\ ]$$

15 Dis/equazioni esponenziali e logaritmiche

15.1 Dis/equazioni con exp e log

Il logaritmo è iniettivo e definito per ogni numero positivo e allora

$$0 < f(x) = g(x) \;\Leftrightarrow\; \log f(x) = \log g(x)$$

cioè si può applicare il logaritmo ad ambo i membri di un'equazione ottenendo un'equazione *equivalente*, cioè con le stesse (eventualmente esistenti) soluzioni. Di più: il logaritmo naturale e ogni logaritmo in base > 1 è crescente e allora si può applicarlo conservando l'ordinamento:

$$0 < f(x) \leq g(x) \;\Leftrightarrow\; \log f(x) \leq \log g(x) \quad \textit{e anche con} \quad <$$

mentre coi logaritmi in base minore di 1 si inverte l'ordinamento:

$$(\forall\, b,\ 0 < b < 1) \quad 0 < f(x) < g(x) \;\Leftrightarrow\; \log f(x) > \log g(x)$$

$$(\forall\, b,\ 0 < b < 1) \quad 0 < f(x) \leq g(x) \;\Leftrightarrow\; \log f(x) \geq \log g(x)\,.$$

15.2 Calcolo approssimato dei logaritmi

La considerazione dei logaritmi naturali o decimali è teoricamente equivalente per le formule di conversione date, del tipo $\log_a x = c \log_b x$, ma a livello pratico giova la $\log_{10}(10^m x) = m + \log_{10} x$.
Mentre i logaritmi decimali non hanno praticamente più alcuna rilevanza nella matematica pura, essi sono ancora ampiamente presenti nelle formule delle scienze applicate. Storicamente essi per secoli hanno consentito il calcolo di $x \cdot y$, x/y e $\sqrt[n]{x}$, operazioni computazionalmente complesse, con le molto più fattibili somme, sottrazioni e divisioni per n, grazie alle (a questo punto) ovvie

$$x \cdot y = 10^{\log_{10} x + \log_{10} y}, \quad x/y = 10^{\log_{10} x - \log_y}, \quad \sqrt[n]{x} = 10^{\frac{1}{n} \log_{10} x} .$$

Una volta, e ancor oggi talvolta in pratica, i (valori approssimati di) $\log_{10} x$ e 10^y si ricavavano da lunghe tavole numeriche, mentre oggi sono disponibili su tutte le calcolatrici scientifiche. Solo non si confonda il logaritmo in base e, di solito denotato ln o LN, col logaritmo decimale, spesso denotato con log o LOG.
Un'approssimazione calcolabile con le sole 4 operazioni e la radice quadrata è

$$\log_{10} x \approx (x^{1/256} - 1) \times 111$$

che su una calcolatrice che operi con almeno 7 cifre (decimali) significative dovrebbe avere non più di 0.004 di errore relativo (cioè 0.4%) e assoluto per $1.1 \leq x \leq 11$. Naturalmente $x^{1/256}$ si ottiene estraendo 8 volte consecutive la radice quadrata. Per i numeri $y < 1.1$ o $y > 11$, li si moltiplichi per un adeguato 10^m con $m \in \mathbb{Z}$ affinchè l'ottenuto $y = 10^m x$ ricada in $[1.1, 11]$ e poi

$$\log_{10} y = \log_{10}(10^m x) = m + \log_{10} x$$

per esempio $\log_{10} 1265 = \log_{10} 1.265 + 3$. Si approssimino così i logaritmi decimali dei numeri 2, 3,...,9, π, 1946; con la formula opportuna anche $\sqrt[3]{5}$. E poi si convertano in logaritmi naturali, con l'opportuna formula approssimata. Si verifichi – approssimatamente – che $\log_{10} 9 = 2 \log_{10} 3$ e similmente che $\ln 8 = 3 \ln 2$.

Un pregevole libro di testo

Nella didattica della matematica per le scienze applicate si fa apprezzare il libro di Marco Abate, *Matematica e Statistica – Le basi per le scienze della vita*, McGraw Hill Ed..
Vediamo qua alcuni esercizi che propone sui logaritmi.

Esercizio 6.20.
Usa le proprietà dei logaritmi [...] per esprimere nel modo più semplice in termini dei logaritmi naturali i seguenti numeri:
(a) $\log_2 3$;
(b) $\log_3 2$;
(c) $\log_2 64$;
(d) $\log_5 15$;
(e) $\log_2(1.5)$;
(f) $\log_3 \frac{1}{27}$.

Esercizio 6.21.
Per ciascuna delle funzioni seguenti, determina l'immagine, stabilisci se è iniettiva e, in tal caso, trova la funzione inversa:
(a) $f_1(x) = 5e^{4x}$;
(b) $f_2(x) = 3 + e^{x+2}$;
(c) $f_3(x) = \frac{1}{2}(e^x - e^{-x})$;
(d) $f_4(x) = (e^x - e^{-x})/(e^x + e^{-x})$. (*Suggerimento:* raccogli e^{-x} a numeratore e denominatore.)

Esercizio 6.22.
Risolvi le seguenti equazioni:
(a) $2^{3x+1} = 4$;
(b) $\log_3 x^2 - \log_3(2x) = 2$;
(c) $\log_3 5 + \log_3 7 = \log_3 x$;
(d) $\log_4 x + \log_4(x - 5) = 2$;
(e) $-3 \log_3 6 + \log_3 8 = \log_3 x$;
(f) $-\log_5(x - 3) + \log_5(x + 2) = 5$.

Esercizio 6.23.
Posto $f(x) = \log_3(3 \log_3 x - 1)$, determina il dominio di f e risolvi la disequazione $f(x) \geq 0$.

16 Funzioni trigonometriche – I

16.1 Definizioni

Nel piano cartesiano di assi X, Y e origine O consideriamo il circolo di centro O e raggio 1 (*circolo goniometrico).* Ad ogni *angolo goniometrico* (che è orientato e ha un lato sul semiasse delle ascisse positive) di misura in radianti $x \in [-2\pi, 2\pi]$, e poi in effetti *facendo anche più di 1 giro*, $x \in \mathbb{R}$, associamo il punto $P(x)$ corrispondente sul circolo goniometrico. Si definiscono le funzioni *seno* e *coseno*, $\sin : \mathbb{R} \to \mathbb{R}$, $\cos : \mathbb{R} \to \mathbb{R}$, e la *tangente*

$$\sin(x) := ordinata(P(x)) \qquad \cos(x) := ascissa(P(x))$$

$$\tan(x) := \frac{\sin(x)}{\cos(x)}, \qquad x \neq \frac{\pi}{2} + k\pi\,,\ k \in \mathbb{Z}$$

$$(\textit{e anche la cotangente } \cot(x) := \frac{\cos(x)}{\sin(x)},\ x \neq k\pi,\ k \in \mathbb{Z})$$

Per esempio si ha (banalmente le prime 2, difficilmente la terza)

$$\sin \pi = 0 \qquad \cos \pi = -1 \qquad \sin 1 \approx 0.84 \tag{1}$$

16.2 Periodicità, simmetrie, gradi e radianti

Il punto goniometrico è palesemente $2\pi-$periodico, $P(x+2\pi) = P(x)$ e allora $P(x+2k\pi) = P(x)$ per ogni k intero, dal che si vede subito la periodicità del seno e del coseno, mentre con considerazioni meno elementari si trova che la tangente è π-periodica:

$$\begin{aligned}\sin(x+2k\pi) &= \sin(x), \quad k \in \mathbb{Z},\\ \cos(x+2k\pi) &= \cos(x), \quad k \in \mathbb{Z},\\ \tan(x+k\pi) &= \tan(x), \quad k \in \mathbb{Z}\,.\end{aligned}$$

Dalla disegno del circolo goniometrico capiamo subito che

il seno è dispari: $\sin(-x) = -\sin x$
il coseno è pari: $\cos(-x) = -\cos x$
la tangente è dispari: $\tan(-x) = -\tan x$

La misura dell'angolo piatto di π radianti, in un altro sistema di misura – più antico – si dice di 180°, e vale la proporzione

$$\pi_{rad} : 180^\circ = x_{rad} : x^\circ \tag{2}$$

che consente di passare da un sistema all'altro. Ecco alcuni degli *angoli notevoli* che si considerano a un livello molto elementare:

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2}{3}\pi$	$\frac{3}{4}\pi$	$\frac{5}{6}\pi$	π	$\frac{3}{2}\pi$	2π
x°	0°	30°	45°	60°	90°	120°	135°	150°	180°	270°	360°

e un qualche interesse ha anche 1 rad $\approx 57^\circ$, e tutte queste corrispondenze si calcolano con la (2). Quasi mai si scrive *rad*.
Alcuni valori numerici interessanti:

$\frac{\pi}{4} \approx 0.785$ (che di solito useremo nel senso di $\frac{\pi}{4}\, rad \approx 0.785\, rad$)
$\frac{\pi}{3} \approx 1.05$ (che di solito useremo nel senso di $\frac{\pi}{3}\, rad \approx 1.05\, rad$)
$\frac{\pi}{2} \approx 1.57$ (che di solito useremo nel senso di $\frac{\pi}{2}\, rad \approx 1.57\, rad$).
$\frac{\sqrt{2}}{2} \approx 0.707$ (per esempio $\cos\frac{\pi}{4}$)
$\frac{\sqrt{3}}{2} \approx 0.866$ (per esempio $\cos\frac{\pi}{3}$)
$\frac{\sqrt{3}}{3} \approx 0.577$ (per esempio $\tan\frac{\pi}{3}$)

16.3 Alcuni valori notevoli

Dalle definizioni si ha subito

x	0	$\frac{\pi}{2}$	$-\pi$ π	$-\frac{\pi}{2}$ $\frac{3}{2}\pi$	2π
x°	0°	90°	180°	270°	360°
sin	0	1	0	-1	0
cos	1	0	-1	0	1
tan	0	$\nexists$	0	$\nexists$	0

Nello schema soprastante è da intendersi che

$$\sin\left(-\frac{\pi}{2}\right) = \sin\left(\frac{3}{2}\pi\right) = \sin(270^\circ)$$

assolutamente non si intenda che $-\frac{\pi}{2}$ **sia uguale a** 270°.
La considerazione del quadrato inscritto nel circolo goniometrico e coi lati paralleli agli assi, che per il Teorema di Pitagora ha lato $\sqrt{2}$, dà subito, con analogo *caveat* come sopra (e anche dopo),

x	$\frac{\pi}{4}$	$\frac{3}{4}\pi$	$-\frac{3}{4}\pi$ $\frac{5}{4}\pi$	$-\frac{\pi}{4}$ $\frac{7}{4}\pi$
x°	45°	135°	225°	315°
sin	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$
cos	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$
tan	1	-1	1	-1

La considerazione dei 2 esagoni regolari inscritti nel circolo goniometrico, con un vertice in $(0,1)$ e rispettivamente in $(1,0)$, dà subito, dividendoli in triangoli equilateri di lato 1 e altezza $\frac{\sqrt{3}}{2}$,

x	$\frac{\pi}{6}$	$\frac{5}{6}\pi$	$-\frac{5}{6}\pi$ $\frac{7}{6}\pi$	$-\frac{\pi}{6}$ $\frac{11}{6}\pi$
x°	30°	150°	210°	330°
sin	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$
cos	$\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$
tan	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$

x	$\frac{\pi}{3}$	$\frac{2}{3}\pi$	$-\frac{2}{3}\pi$ $\frac{4}{3}\pi$	$-\frac{\pi}{3}$ $\frac{5}{3}\pi$
x°	60°	120°	240°	300°
sin	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{3}}{2}$
cos	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$
tan	$\sqrt{3}$	$-\sqrt{3}$	$\sqrt{3}$	$-\sqrt{3}$

La considerazione di decagoni regolari darebbe ancora altri valori.

17 Funzioni trigonometriche – II

17.1 Alcune formule goniometriche notevoli

Nel corso dei secoli sono state dimostrate moltissime formule. Ecco alcune delle principali.
Identità goniometrica fondamentale:

$$\sin^2 x + \cos^2 x = 1 \tag{3}$$

Formule di addizione e sottrazione:

$$\begin{aligned}
\sin(x+y) &= \sin x \cos y + \cos x \sin y \\
\sin(x-y) &= \sin x \cos y - \cos x \sin y \\
\cos(x+y) &= \cos x \cos y - \sin x \sin y \\
\cos(x-y) &= \cos x \cos y + \sin x \sin y
\end{aligned}$$

Formule di duplicazione:

$$\sin(2\,x) = 2\,\sin x \cos x$$

$$\cos(2\,x) = \begin{cases} \cos^2 x - \sin^2 x \\ 2\,\cos^2 x - 1 \\ 1 - 2\,\sin^2 x \end{cases}$$

$$\tan(2\,x) = \tfrac{2\,\tan(x)}{1-\tan^2 x}, \qquad x \neq \tfrac{\pi}{4} + k\,\tfrac{\pi}{2}, \quad k \in \mathbb{Z}$$

Formule di prostaferesi:

$$\begin{aligned}
\sin x + \sin y &= 2\,\sin \tfrac{x+y}{2}\,\cos \tfrac{x-y}{2} \\
\sin x - \sin y &= 2\,\cos \tfrac{x+y}{2}\,\sin \tfrac{x-y}{2} \\
\cos x + \cos y &= 2\,\cos \tfrac{x+y}{2}\,\cos \tfrac{x-y}{2} \\
\cos x - \cos y &= -2\,\sin \tfrac{x+y}{2}\,\sin \tfrac{x-y}{2}
\end{aligned}$$

Formule di bisezione della tangente:

$$\tan \frac{x}{2} = \begin{cases} \tfrac{\sin x}{1+\cos x}, & x \neq (2\,k+1)\,\pi \\ \\ \tfrac{1-\cos x}{\sin x}, & x \neq k\,\pi \end{cases} \qquad k \in \mathbb{Z}$$

17.2 Funzioni goniometriche inverse

Arcoseno:

$$\arcsin := \left(\sin\big|_{[-\pi/2,\pi/2]}\right)^{-1}$$

Arcocoseno:

$$\arccos := \left(\cos\big|_{[0,\pi]}\right)^{-1}$$

Arcotangente:

$$\arctan := \left(\tan\big|_{]-\pi/2,\pi/2[}\right)^{-1}$$

Alcuni valori notevoli. Dalle definizioni si ha subito

x	0	$\frac{1}{2}$	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\sqrt{3}$
arcsin	0	$\frac{\pi}{6}$		$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\nexists$
arccos	$\frac{\pi}{2}$	$\frac{\pi}{3}$		$\frac{\pi}{4}$	$\frac{\pi}{6}$	0	$\nexists$
arctan	0		$\frac{\pi}{6}$			$\frac{\pi}{4}$	$\frac{\pi}{3}$

I valori mancanti in tabella non si considerano *valori notevoli*, ma esistono. (Per esempio $\arcsin\frac{\sqrt{3}}{3} \approx 0.61548$).

ESERCIZI

ESERCIZIO$_{\mu 2018}$ Risolvere la seguente equazione goniometrica:

$$\sin x = \frac{1}{2}$$

Svolgimento
Si risolve subito disegnando un circolo goniometrico e segnandovi gli angoli $\frac{\pi}{6}$ e $\frac{5}{6}\pi$, corrispondenti appunto a $\sin x = \frac{1}{2}$:

$$\boxed{x = \frac{\pi}{6} + 2k\pi \;\vee\; x = \frac{5}{6} + 2k\pi, \quad k \in \mathbb{Z}}$$

ESERCIZIO$_{\mu 2018}$ Approssimare $\cos(1)$ usando le (1) e (3).
Si verifichi poi su WolframAlpha con `Cos[1]`

ESERCIZIO$_{\mu 2018}$ Risolvere la disequazione

$$\sin x > 0.8$$

Svolgimento
(Aiutandoci senz'altro con un disegno) si trova subito

$$\boxed{\arcsin(0.8) + 2k\pi < x < \pi - \arcsin(0.8) + 2k\pi, \quad k \in \mathbb{Z}}$$

numericamente

$$0.9272... + 2k\pi < x < 2.2142... + 2k\pi, \quad k \in \mathbb{Z}$$

ESERCIZIO$_{\mu 2018}$ Risolvere la disequazione

$$\sin x \geq 0.8$$

(Aiutandoci senz'altro con un disegno) si trova subito

$$\boxed{-\pi - \arcsin(0.8) + 2k\pi \leq x \leq \arcsin(0.8) + 2k\pi, \quad k \in \mathbb{Z}}$$

numericamente

$$-4.0688... + 2k\pi < x < 0.9272... + 2k\pi, \quad k \in \mathbb{Z}.$$

ESERCIZIO$_{\mu 2018}$ Calcolare approssimatamente arcsin $\frac{21}{25}$ (a mano, ovvio, ma anche prendere confidenza con una calcolatrice scientifica sarà un bene).
Si cerchi poi il valore online su WolframAlpha con

`ArcSin[21/25]`

(Svolgimento in 65.12).

ESCERCIZIO$_{K2018}$ Risolvere la disequazione

$$\frac{\sqrt{3}}{2} - \cos x < 2\,\cos\left(x + \frac{2\pi}{3}\right)$$

Svolgimento

Con la formula di addizione del coseno

$$2\Big(\cos x\,\cos\left(\frac{2\pi}{3}\right) - \sin x\,\sin\left(\frac{2\pi}{3}\right)\Big) =$$

e coi valori notevoli ben conosciuti

$$= 2\left((\cos x)\left(-\frac{1}{2}\right) - (\sin x)\left(\frac{\sqrt{3}}{2}\right)\right) = -\cos x - \sqrt{3}\,\sin x$$

e con il secondo membro cosí riscritto la disequazione equivale a

$$\frac{\sqrt{3}}{2} - \cos x < -\cos x - \sqrt{3}\,\sin x$$

$$\frac{\sqrt{3}}{2} < -\sqrt{3}\,\sin x$$

$$-\sqrt{3}\,\sin x > \frac{\sqrt{3}}{2} \quad \Big/ : (-\sqrt{3}) < 0$$

$$\sin x < -\frac{1}{2}$$

che si risolve subito disegnando un circolo goniometrico e segnandovi l'arco da $-\frac{5}{6}\pi$ fino a $-\frac{\pi}{6}$, estremi esclusi, corrispondente appunto a $\sin x < -\frac{1}{2}$:

$$\boxed{-\tfrac{5}{6}\pi + 2k\pi < x < -\tfrac{\pi}{6} + 2k\pi, \quad k \in \mathbb{Z}}$$

equivalentemente scritta

$$\boxed{\tfrac{7}{6}\pi + 2k\pi < x < \tfrac{11}{6}\pi + 2k\pi, \quad k \in \mathbb{Z}}$$

IV – Statistica descrittiva

18 Medie e altri indici di posizione; outlier

Vogliamo riassumere molti valori (*dataset*) con 1 valore. Per comprenderli, divulgarli, confrontarli. Quest'opera di sintesi non è banale. Noi partiremo da dati numerici, ma a monte c'è il problema di quali numeri considerare, in una data situazione: se consideriamo un quadrato di lato 2 ovvero area 4, e 2 quadrati di lato 5 ovvero area 25, e volessimo rappresentare un "quadrato medio", facendo la media dei lati abbiamo 4 (con area 16), invece facendo la media delle aree abbiamo 18 (con lato $3\sqrt{2} \approx 4.24$).

18.1 Medie aritmetiche e geometrica; outlier

Media (aritmetica). Da ora consideriamo un dataset $\{x_1, ..., x_n\}$.

$$M(x_1, ..., x_n) := \frac{x_1 + x_2 + ... + x_n}{n}$$

Facile e bello, ottimo per i voti, ma il reddito medio di questi

$$3, 1, 2, 4, 6, 1000, 6$$

è 146, non poi così espressivo della situazione globale, a causa dell'*outlier* 1000, valore anomalo ovvero aberrante. (Vedi sotto).

Media geometrica. Per n numeri positivi (invece la media aritmetica non lo richiede), è la radice n-esima del loro prodotto:

$$GM(x_1, ..., x_n) := (x_1 \cdot ... \cdot x_n)^{\frac{1}{n}} .$$

Nell'esempio numerico soprastante ≈ 7.05.

Media (aritmetica) ponderata (o pesata). Dati dei *pesi* $a_1, ..., a_n$, di somma 1, la media pesata del dataset $\{x_1, ..., x_n\}$ è $a_1x_1 + ... + a_nx_n$.

Outlier. Leggiamo su Wikipedia, l'enciclopedia libera, alla voce "Outlier", e non è chi non veda le pericolose conseguenze di ciò:

> There is no rigid mathematical definition of what constitutes an outlier; determining whether or not an observation is an outlier is ultimately a subjective exercise.(...) Deletion of outlier data is a controversial practice

18.2 Mediana, moda, media interquartile↑

Mediana. Il numero centrale dei dati riordinati, adesso 4: $1, 2, 3, \underline{4}, 6, 6, 1000$. E se i dati sono in numero pari, si considera la media dei 2 centrali. La mediana è definita anche per valori non numerici che siano però almeno *ordinali*, come *sono completamente d'accordo*, *sono d'accordo*, ... *non sono assolutamente d'accordo.*
Moda. Il valore più frequente. Nel nostro esempio è 6 ma in generale non è definita perchè i numeri sono tutti diversi o perchè 2 o più valori sono ripetuti ugualmente. Ecco per esempio un dataset *bimodale*: $6, 6, 6, 6, 7, 7.5, 7.5, 8, 8, 8, 8, 8.5, 8.5, 9, 9.5, 10, 10$.
La moda è definita anche per dati *nominali*, neppure ordinali, per esempio in ogni nazione c'è un cognome più frequente.
Media interquartile Leggiamo su Wikipedia, l'enciclopedia libera, alla voce Central tendency:

> The method is best explained with an example. Consider the following dataset:
> 5, 8, 4, 38, 8, 6, 9, 7, 7, 3, 1, 6
> First sort the list from lowest-to-highest:
> 1, 3, 4, 5, 6, 6, 7, 7, 8, 8, 9, 38
> There are 12 observations (datapoints) in the dataset, thus we have 4 quartiles of 3 numbers. Discard the lowest and the highest 3 values:
> ~~1~~, ~~3~~, ~~4~~, 5, 6, 6, 7, 7, 8, ~~8~~, ~~9~~, ~~38~~
> We now have 6 of the 12 observations remaining; next, we calculate the arithmetic mean of these numbers:
> xIQM = (5 + 6 + 6 + 7 + 7 + 8) / 6 = 6.5
> This is the interquartile mean.
> For comparison, the arithmetic mean of the original dataset is
> (5 + 8 + 4 + 38 + 8 + 6 + 9 + 7 + 7 + 3 + 1 + 6) / 12 = 8.5
> due to the strong influence of the outlier, 38.

Con artifici si definisce per un numero di dati non quadruplo.

19 Bar chart, istogrammi, e altri diagrammi

Bar chart (ovvero istogrammi a barre) e istogrammi – purtroppo con ambiguità nominalistiche nei vari testi – sono diagrammi per la visualizzazione di dati. Nei bar chart l'altezza di ogni colonna – o la lunghezza se disposta orizzontalmente – rappresenta un valore, negli istogrammi l'area rappresenta una frequenza assoluta o percentuale.

• Considerato il dataset

$$2, 3, 1, 4, 1, 5, 3, 4, 7, 1, 2, 4, 5, 9, 4, 8$$

se ne rappresenti il bar chart. Poi si rappresenti l'istogramma con intervalli $[0, 2.5[$, $[2.5, 5[$, $[5, 7.5[$, $[7.5, 10[$, e poi con intervalli $[0, 2.5[$, $[2.5, 5[$, $[5, 10[$.

• Considerato il dataset $\sin\big(k\frac{\pi}{6}\big)$, $k = 0, ..., 10$, se ne rappresenti ll bar chart. Poi si rappresenti l'istogramma (delle frequenze dei valori) con intervalli $[0, 0.3[$, $[0.3, 0.6[$, $[0.6, 0.9[$, $[0.9, 1.2[$, e poi con intervalli $[0, 0.6[$, $[0.6, 0.9[$, $[0.9, 1.2[$.

• Si rappresenti il diagramma a torta – di ovvio significato, almeno quando se ne fa uno solo: gli angoli o equivalentemente le aree sono proporzionali alla frazione percentuale da rappresentare – per questa situazione:

tifosi dei Gialli: 66
tifosi dei Neri: 132
tifosi dei Blu: 198.

20 Quartili e box-plot

Si abbia il dataset numerico $x_1, .., x_n$ e lo si riordini in modo crescente: $y_1, ..., y_n$. Il valore intermedio, o la media dei 2 intermedi se n è pari, sappiamo che è la mediana, che da adesso chiameremo anche *secondo quartile*, e lo indicheremo con $q_{0.5}$ oppure $q_{1/2}$ oppure $q_{50\%}$. Diviso così il dataset in 2 parti uguali, con eventuale ripetizione della mediana nelle 2 parti se n è dispari, la mediana della prima parte si chiama *primo quartile*, e lo indicheremo con $q_{1/4}$ oppure $q_{0.25}$ oppure $q_{25\%}$, e la mediana della seconda parte si chiama *terzo quartile*, e lo indicheremo con $q_{3/4}$ oppure $q_{0.75}$ oppure $q_{75\%}$. Si hanno poi i quartili di indice 0, che è il minimo del dataset, e quello di indice 1 ovvero 100%, che è il massimo del dataset. I 5 numeri detti, $q_{0\%}$,...,$q_{100\%}$, sono una statistica di sintesi vettoriale che si chiama *riassunto dei 5 numeri* (o *five number summary*). Leggiamo su Wikipedia, l'enciclopedia libera, alla voce "Diagramma a scatola e baffi":

> il diagramma a scatola e baffi (o diagramma degli estremi e dei quartili o box and whiskers plot o box-plot) è una rappresentazione grafica [...] Viene rappresentato (orientato orizzontalmente o verticalmente) tramite un rettangolo diviso in due parti, da cui escono due segmenti. Il rettangolo (la "scatola") è delimitato dal primo e dal terzo quartile, $q_{1/4}$ e $q_{3/4}$, e diviso al suo interno dalla mediana, $q_{1/2}$. I segmenti (i "baffi") sono delimitati dal minimo e dal massimo dei valori. In questo modo vengono rappresentati graficamente i quattro intervalli ugualmente popolati delimitati dai quartili.

Per esempio, per i primi 17 numeri naturali, i 5 numeri sono nell'ordine 0, 4, 8, 12, 16. Se ne disegni il box-plot, e quello dei primi 17 numeri primi, e quello dei residenti nelle regioni italiane. Talvolta gli outlier vengono rappresentati isolati e non computati. Si faccia così per il dataset 1, $\frac{1}{2}$, $\frac{1}{3}$,..., $\frac{1}{13}$, $\sqrt{7}$; e poi senza $\frac{1}{13}$.

21 Note sulla Statistica Descrittiva; skewness

21.1 Note sulla Statistica Descrittiva

La Statistica Descrittiva si occupa essenzialmente della rilevazione e sintesi di dati. L'insieme dei dati $x_1, ..., x_n$ (che spesso sono numeri ma non sempre) si chiama *dataset* e non è un insieme in senso matematico perchè le ripetizioni dei valori sono ammesse e non vanno elise.
Consideriamo dataset *nominali*, *ordinali* e *numerici*.
In questo testo elementare, tralasciata la questione della rilevazione dei dati, ci occupiamo della sintesi, ovvero della rappresentazione dei dati in una forma umanamente comprensibile e *trattabile*, cosa particolarmente utile se i dati sono più di una mezza dozzina. Consideriamo essenzialmente 2 capitoli della Statistica Descrittiva:
◇ le *rappresentazioni grafiche*, come gli istogrammi;
◇ le *statistiche di sintesi*, funzioni $f(x_1, ..., x_n)$ con un numero n a priori indeterminato di argomenti, le quali **vogliono riassumere il dataset con 1 o pochi valori**, per comprenderli, divulgarli, confrontarli. Ecco alcune statistiche di sintesi per un dataset:
• il minimo: $\min(x_1, ..., x_n) \leftarrow$ per un dataset almeno ordinale;
• il massimo: $\max(x_1, ..., x_n) \leftarrow$ per un dataset almeno ordinale;
• la somma: $x_1 + ... + x_n \leftarrow$ per un dataset numerico.
Il significato delle soprastanti statistiche di sintesi è ovvio. Si calcolino quei 3 valori per questo dataset: $\frac{1}{\pi}$, $\frac{2}{7}$, 0.3, $\frac{1}{\sqrt{10}}$, e, e^{-1}, $\frac{1}{3}$.
Più in dettaglio consideriamo queste altre statistiche di sintesi:
• gli indici di posizione, che vorrebbero riassumere in 1 solo valore il complesso dei dati;
• gli indici di dispersione ovvero variabilità, che vorrebbero quantificare con 1 numero la non omogeneità di dati numerici.
In una trattazione elementare della Statistica Descrittiva non si distingue fra *popolazione* e *campione*: abbiamo i dati che abbiamo, e quelli consideriamo popolazione, anche se in effetti sono un campione di una popolazione più ampia.

21.2 Eventuali asimmetrie nei dataset: skewness

Consideriamo per esempio i primi 17 numeri *triangolari*: 1, 3 = 1+2, 6 = 1+2+3,...,153 = 1+2+...+17. Il riassunto dei 5 numeri è 1, 15, 45, 91, 153. Il box-plot è molto approssimativamente

$.--|XXXXX|XXXXXXXXX|--------------...$

Oppure con un *istogramma dei quartili idealizzato*, necessariamente con 4 colonne di area 25%:
1 col. per 5 valori da 1 a 15 compresi: 1, 3, 6, 10, 15 (ampio 14)
1 col. per 5 valori da 15 a 45 compresi: 15, 21, 28, 36 45
1 col. per 5 valori da 45 a 91 compresi: 45, 55, 66, 78, 91
1 col. per 5 valori da 91 a 153 compresi: 91, 105, 120, 136, 153.
Possibili altezze: 100, ≈ 46.7, ≈ 30.4, ≈ 22.6. (Aree: ≈ 1400).
Se volessimo farne un bar chart, dovremmo dapprima decidere quanti intervalli usare, e se volessimo usare la regola di farne circa la radice quadrata della numerosità, avremmo 4 oppure 5 intervalli ugualmente ampi "intorno" a 1 e 153; scegliamo di fare intervalli ampi 45 da 0 a 180 escluso, ma non è l'unica possibilità:

da 0 a 45 escluso: 1, 3, 6, 10, 15, 21, 28, 36 → 8 valori
da 45 a 90 escluso:i 45, 55, 66, 78 → 4 valori
da 90 a 135 escluso: 91, 105, 120 → 3 valori
da 135 a 180 escluso: 136, 153. → 2 valori. Da cui il bar chart:

```
 [ 0, 45[ |XXXXXXXX
 [45, 90[ |XXXX
 [90,135[ |XXX
[135,180[ |XX
```

(Con 5 intervalli ampi 35, da 0, si ottengono numerosità 7, 4, 2, 3, 1, alquanto fuorviante: la distanza fra i dati aumenta sempre).
Sia dal box-plot che dall'istogramma che dal bar chart vediamo che i dati sono più addensati verso i valori bassi che quelli alti.
Questo tipo di distribuzione si dice *right skewed*, e si intuisce com'è una distribuzione *left skewed*. (Queste *non* sono definizioni rigorose ma permettono di decidere nella generalità dei casi non particolarmente "capricciosi"). Si provi con 0.3, 1, $\frac{1}{2}$, $\frac{1}{3}$,...,$\frac{1}{10}$, $\frac{1}{11}$.

22 Variabilità, covarianza e correlazione

22.1 Indici di dispersione ovvero variabilità

Gli indici dispersione ovvero variabilità cercano quantificare con 1 numero la non omogeneità di dati numerici. Sono statistiche di sintesi $f(x_1, ..., x_n)$ da $\mathbb{R}^n$ in $\mathbb{R}$, con n a priori indeterminato, proprio come le varie medie, viste in precedenza. è ovvio che dataset diversi possono avere diversi gradi di "disomogeneità", che vogliamo quantificare, anche se hanno la stessa media e/o mediana. Per esempio i 2 dataset, qua già ordinati,

$$2, 4, 6, 8, 10, 12, 14 \qquad 5, 6, 7, 8, 9, 10, 11$$

hanno uguale media, e perfino mediana, sempre 8, ma è chiaro che i valori del secondo dataset variano meno, ovvero sono più addensati intorno alla media. Si pensi a dei redditi per esempio.
Campo di variazione, o *range*.

$$range(x_1, ..., x_n) := \max(x_1, ..., x_n) - \min(x_1, ..., x_n) \,.$$

Che vale 12 e rispettivamente 6 nei 2 dataset considerati. Molto sensibile agli eventuali outlier, per quali si potrebbe ipotizzare la non considerazione – ovviamente avvertendo il lettore.
Differenza interquartile. Ben poco sensibile agli outlier:

$$iqr(x_1, ..., x_n) := q_{3/4} - q_{1/4} \,.$$

Varianza, deviazione standard, coefficiente di variazione. Detta $\bar{x}$ la media aritmetica del dataset $X : x_1, ..., x_n$

$$\sigma_X^2 = Var(x_1, ..., x_n) := \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2 \quad \text{(unità di misura "strana"!)}$$

$$\sigma_X = sd(x_1, ..., x_n) := \sqrt{Var(x_1, ..., x_n)}\,, \qquad cv(x_1, ..., x_n) := \tfrac{\sigma_X}{\bar{x}}\,.$$

Varianza e coefficiente di variazione sono sensibili agli outlier.

- Per i dataset visti in precedenza si calcolino i 5 indici sopradetti.

22.2 Covarianza; correlazione; retta di regressione

Dati 2 dataset numerici di uguale numerosità

$$X:\ x_1,...,x_n \qquad Y:\ y_1,...,y_n$$

si definisce la loro *covarianza* (ovvero *covarianza di 2 n-uple di numeri*, ovvero *covarianza osservata*)

$$Cov(x_1,...,x_n;y_1,...,y_n) := \frac{(x_1-\bar{x})(y_1-\bar{y})+...+(x_n-\bar{x})(y_n-\bar{y})}{n}$$

dove $\bar{x}$ e $\bar{y}$ sono le medie dei 2 dataset. Si trova denotata anche anche $\sigma_{X,Y}$, oppure con *cov*(... minuscolo. Si definisce anche l'*indice di correlazione di [Bravais-]Pearson*, o *indice di correlazione lineare* (e al posto della parola *indice* altri scrive *coefficiente*)

$$\rho_{X,Y} := \frac{Cov(x_1,...,x_n;y_1,...,y_n)}{\sqrt{Var(x_1,...,x_n)Var(y_1,...,y_n)}} \qquad = \frac{\sigma_{X,Y}}{\sigma_X\sigma_Y}$$

e questo dà un'idea di come varino le y_1 relativamente alle x_i: un indice prossimo a 1 suggerisce che al crescere delle x_i crescano linearmente le y_i, un indice prossimo a -1 suggerisce che al crescere delle x_i decrescano linearmente le y_i, e un indice prossimo a 0 indica la non presenza di tali relazioni approssimate. In ogni caso

$$- \leq \rho_{X,Y} \leq 1\ .$$

Coi dataset X : 2, 7, 4, 3, e Y : 3, -1, 0, 2 si calcoli l'indice di correlazione, e si rappresentino sul piano cartesiano i punti (x_i, y_i).

In un senso non banale, una retta $y = mx + q$ si avvicina meglio ai punti se ha coefficienti

$$m := \frac{\sigma_{X,Y}}{\sigma_X^2} \qquad q := \bar{y} - m\bar{x} \quad \text{equivalentemente: } m := \frac{\sigma_Y}{\sigma_X}\rho_{X,Y}$$

e in questo caso si chiama *retta di regressione* (dietro, vi è il *metodo dei minimi quadrati*, che non approfondiamo).

Per esempio per i punti $(2,1)$, $(1,3)$, $(3,-1)$, ovvero per i dataset $X:\ 2,1,3$ e $Y:\ 1,3,-1$, si trova $y = -2x+5$, che passa *esattamente* per i punti. è $\bar{x} = 2$, $\bar{y} = 1$, $\sigma_X^2 = \frac{2}{3}$, $\sigma_{X,Y} = -\frac{4}{3}$. E con $y_2 := 3.14159265$ si trova $y \approx -2.07x + 5.19$. Si facciano i disegni.

A1- ESERCIZI SULLE MATEMATICHE ELEMENTARI

ESERCIZIO 1 (Peano)

a) Calcolare $1 + 3$ con la definizione di Peano.

b) Calcolare $1 \cdot 3$ con la definizione di Peano.

ESERCIZIO 2 (Divisione euclidea)

a) Fare la divisione euclidea (quoziente e resto): 27 diviso 4.

b) Fare la divisione euclidea (quoziente e resto): -27 diviso 4.

c) Fare la divisione euclidea (quoziente e resto): -27 diviso -4.

ESERCIZIO 3 (Logica delle proposizioni)

a) Con le tavole di verità provare che è tautologia $(p \vee (\neg q)) \vee q$

b) Similmente, provare che è contraddizione $(p \wedge (\neg q)) \wedge q$

c) Similmente, provare che sono equivalenti $p \Rightarrow q$, $(\neg p) \vee q$

ESERCIZIO 4 (Insiemistica)

a) Determinare l'insieme delle parti dell'insieme delle parti dell'insieme delle parti dell'insieme $\{0\}$.

b) Determinare l'insieme delle parti dell'insieme delle parti dell'insieme delle parti dell'insieme vuoto.

c) Determinare il prodotto cartesiano dei 2 insiemi $\{x \in \mathbb{N} | x < 7 \wedge x\ e'\ primo\}$, $\{0, \emptyset, Roma\}$.

>Per generici insiemi X, Y e W in un insieme universo U, disegnare:

d) $C(X \cap Y)$, essendo C il passaggio al complementare

e) $X \Delta (Y \cup W)$

ESERCIZIO 5 (Logica dei predicati)

Si consideri questo predicato: $p(x, y)$:="la retta x passa per il punto y" (nel piano euclideo). Determinare il valore di verità delle seguenti proposizioni:

a) $(\forall x)(\forall y)p(x, y)$

b) $(\forall x)(\exists y)p(x, y)$

c) $(\exists x)(\forall y)p(x, y)$

ESERCIZIO 6 (Funzioni)

a) Trovare un'applicazione suriettiva $\mathbb{Z} \to \mathbb{N}$

b) Trovare 3 diverse applicazioni biiettive $\mathbb{Z} \to \mathbb{Z}$

ESERCIZIO 7 (Equazioni e disequazioni)

a) $(x^5 + x^4 + x^3)(x^2 - 3) < 0$

b) $x^3 - 5x^2 + 3x + 4 \leq 0$

c) $\frac{x^2-x-2}{x^2+x-6} = 0$

d) $\frac{x^3+3x^2-10x-24}{x^3+x^2-14x-24} \geq 0$

Sezione A2 – Calcolo Infinitesimale

Il Calcolo Infinitesimale comprende essenzialmente:
◇ La teoria dei limiti (delle successioni e delle funzioni)
◇ La teoria delle derivate, o Calcolo Differenziale
◇ La teoria delle serie (numeriche, e di funzioni)
◇ La teoria dell'integrale.

V – Limiti e derivate

23 Limiti di successioni

23.1 Limiti di successioni: introduzione

Consideriamo un numero, diciamo 2.017, e facciamone la radice quadrata, poi la radice quadrata del risultato, e così avanti, cioè consideriamo la *successione definita per ricorrenza*

$$\begin{cases} x_0 := 2.017 \\ x_{n+1} := \sqrt{x_n} \end{cases} \qquad n = 0, 1, ...$$

Con purtroppo inevitabili errori di approssimazione si trova

$x_0 = 2.017$	
$x_1 \approx 1.420\,21$	(Qua sono riportate solo approssimazioni
$x_2 \approx 1.191\,73$	dei risultati che via via la calcolatrice produce,
$x_3 \approx 1.091\,66$	facendo la radice quadrata del risultato
$x_4 \approx 1.044\,83$	precedente, ed essa fa i calcoli con più decimali).
$x_5 \approx 1.022\,17$	
$x_6 \approx 1.011\,02$	(Si noti la lieve spaziatura ogni 3 cifre decimali,
$x_7 \approx 1.005\,5$	che si può fare anche prima del punto decimale,
$x_8 \approx 1.002\,74$	conformemente allo standard)
...	
$x_{16} \approx 1.000\,01$	

Ci *sembra*, ed è vero e si potrebbe dimostrare (quando si definisce come più sotto), che x_n *tende a 1 per n che tende all'infinito*:

$$\lim_{n\to+\infty} x_n = 1\,. \qquad \text{(Altri scrive } {}_{n\to\infty} \text{ invece di } {}_{n\to+\infty}\text{).}$$

Più in generale si dà il caso, con altra successione x_n, e $L \in \mathbb{R}$,

$$\lim_{n\to+\infty} x_n = L \qquad \text{(o invece di n, qualunque variabile, p. es. } k\text{).}$$

Il significato è di un indefinito avvicinamento, con o senza raggiungimento del limite L. Più precisamente

$$(\forall\,\varepsilon > 0)\;(\exists\,\bar{n})\; n > \bar{n} \Rightarrow |x_n - L| < \varepsilon\,.$$

Cioè, fissato un numer(ett)o positivo, chiamiamolo ε, da un certo punto in poi x_n dista dal limite meno di ε. Si provi con stessa $x_{n+1} := \sqrt{x_n}$ ma con $x_0 := 0.2017$. O altro numero a piacere.

23.2 Limite di successioni: approfondimento

Invece di fare i limiti "a occhio" come prima, proviamo a dare delle regole teoriche. Consideriamo la successione $x_k := \frac{1}{k}$. è chiaro che quanto più k sarà grande, tanto più il suo reciproco x_k sarà vicino a 0, cioè, nel senso prima definito per il limite,

$$\lim_{k\to+\infty} \frac{1}{k} = 0\,.$$

Consideriamo ora la successione $x_m := m^2$. è chiaro che quanto più m sarà grande, tanto più il suo quadrato x_m sarà grande, e i valori supereranno qualunque soglia, crescendo indefinitamente. Scriveremo, usando per esempio la variabile m,

$$\lim_{m\to+\infty} m^2 = +\infty \qquad \lim_{m\to+\infty} -m^2 = -\infty$$

e il senso del limite $+\infty$ è (usando per esempio la variabile n) è

$$\lim_{n\to+\infty} x_n = +\infty \quad\Leftrightarrow\quad (\forall\, M)\;(\exists\,\bar{n})\; n > \bar{n} \Rightarrow x_n > M.$$

Allora adesso siamo in grado di *calcolare*

$$\lim_{h\to+\infty} \frac{1}{h^2} = 0\,.$$

Non solo n^2 ma anche n^3, eccetera, tendono a $+\infty$, e $\sqrt{n}$, $\sqrt[3]{n}$... E anche $7n^2 - n + 1$ perchè $7n^2 - n + 1 = n^2(7 - \frac{1}{n} - \frac{1}{n^2})$ che per "grandi" n è il prodotto di un numero "grande" per un numero prossimo a 7. E anche $n - \sqrt{n} = \sqrt{n}(\sqrt{n} - 1)$; e $\lg n$, $\ln n$, e^n. Tutto ciò costituisce il *calcolo dei limiti* delle successioni, che comunque può raggiungere più alti livelli di sottigliezza:

$$\lim_{n\to+\infty} \frac{3n^4 - 5n^3 + 6 - \mathrm{e}\sqrt{n}}{3n^3 - \frac{1}{n} - 2n^4} = \lim_{n\to+\infty} \frac{3 - \frac{5}{n} + \frac{6}{n^4} - \frac{\mathrm{e}}{n^3\sqrt{n}}}{\frac{3}{n} - \frac{1}{n^5} - 2} = -\frac{3}{2}\,.$$

Si dimostrano $\frac{e^n}{n} \to +\infty$ e $\frac{\ln n}{n} \to 0$. Ma non esistono, per esempio,

$$\lim_{n\to+\infty} (-1)^n \qquad \lim_{n\to+\infty} \sin n \qquad \lim_{n\to+\infty} \cos n \qquad \lim_{n\to+\infty} \tan n\,.$$

Però $\lim_{n\to+\infty} \frac{\sin n}{n} = 0$ perchè il numeratore è *limitato* (cioè si mantiene fra 2 numeri), e il denominatore $\to +\infty$. Così $\frac{(-1)^n}{\sqrt[3]{n}}$.

24 Limiti e continuità

Tranne che la questione di $(-1)^n$ e le successioni definite per ricorrenza, tutto quanto detto sui limiti di successioni si estende ai limiti di funzioni, per esempio $\lim_{x\to+\infty} \mathrm{e}^x = +\infty$, $\nexists \lim_{x\to+\infty} \sin x$,

$$\lim_{x\to+\infty} \frac{\sin x}{x} = 0 \qquad \lim_{x\to+\infty} \frac{3x^4 - 5x^3 + 6 - \mathrm{e}\sqrt{x}}{3x^3 - \frac{1}{x} - 2x^4} = ... = -\frac{3}{2}\,.$$

Solo che adesso x può andare anche a $-\infty$, p. es. $\lim_{x\to-\infty} \mathrm{e}^x = 0$,

$$\lim_{x\to-\infty} \frac{\sin x}{x} = 0 \qquad \lim_{x\to-\infty} \frac{3x^4 - 5x^3 + 6 - \mathrm{e}\sqrt{x}}{3x^3 - \frac{1}{x} - 2x^4} = ... = -\frac{3}{2}$$

o un numero $x_0 \in \mathbb{R}$, anche distinguendo x_0^+ e x_0^-, e ci sono 2 casi:
◇ $x_0 \notin \mathrm{dom} f$, p. es. $\lim_{x\to 0^-} \frac{x}{|x|} = -1$, $\lim_{x\to 0^-} \frac{1}{x} = -\infty$
◇ $x_0 \in \mathrm{dom} f$ e ci sono 2 sottocasi:
- $\lim_{x\to x_0} f(x) = f(x_0) \Rightarrow$ la f si dice continua in x_0 e si noti che le funzioni elementari sono continue nei domini, e allora per esse sempre limite=valore calcolato, p. es. $\lim_{x\to 0} \mathrm{e}^x = 1$;
- $\lim_{x\to x_0} f(x) \neq f(x_0) \Rightarrow$ la f si dice discontinua in x_0 e questo può essere con funzioni non elementari come $\mathrm{sgn}(x)$ e $[x]$.

Allora per le funzioni elementari sono significativi solo i limiti dove la funzione "smette di esistere": gli estremi, finiti o no, degli intervalli che compongono il dominio, non appartenenti ad esso:

$$\lim_{x\to 1} \frac{1-\sqrt{x}}{1-x} = \text{ [forma 0 su 0] } \lim_{x\to 1} \frac{1-\sqrt{x}}{(1-\sqrt{x})(1+\sqrt{x})} =$$

$$= \lim_{x\to 1} \frac{1}{1+\sqrt{x}} \text{ [funzione elementare, limite=valore] } = \frac{1}{2}\,.$$

Si dice *limitata* una f tale che esistono 2 numeri M e N tali che $M < f(x) < N$, per esempio $\sin x$. Limitata $\not\Rightarrow \exists$ limite.
Si dice *infinita* una funzione che tende a $+\infty$ o $-\infty$ e *infinitesima* una che tende a 0 (per $x \to u_0$ finito o no). Per esempio sono infinite per $x \to +\infty$ le a^x con $a > 1$, e infinitesime per $0 < a < 1$; viceversa per $x \to -\infty$. Infinite le $\log_b x$ per $x \to 0^+$ e $x \to +\infty$.

24.1 Teoremi sui limiti

Con le definizioni dei limiti (con ε, δ...) si dimostra che:

• Primo **limite fondamentale**: $\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = \mathrm{e}$.
• limite somma=somma limiti se finiti: $\lim_{x\to\pi}(\mathrm{sign}(x)+[x])=4$
• limite prodotto=prodotto limiti se finiti: $\lim_{x\to-\pi}\mathrm{sign}(x)[x]=4$
• limitata fratto infinita, $\to 0$, p.es. $\lim_{x\to+\infty}\frac{\sin x}{\log_{\frac{1}{2}} x}=0$
• limitata + infinita>0, $\to +\infty$, p.es. $\lim_{x\to+\infty}(\cos x+\sqrt{x})=+\infty$
• reciproca di tendente a $L\neq 0$ tende a $\frac{1}{L}$, p.es. $\lim_{x\to+\infty}\frac{1}{\log_2 8}=\frac{1}{3}$
• reciproca d'infinita $\to 0$, p.es. $\lim_{x\to+\infty}\frac{1}{\log_{10} x}=0$
• reciproca d'infinitesima >0, $\to +\infty$, p.es. $\lim_{x\to\pi^-}\mathrm{cosec}=+\infty$
• reciproca d'infinitesima <0, $\to -\infty$, p.es. $\lim_{x\to 1^-}\frac{1}{\ln x}=-\infty$
• per quanto gli infiniti non siano numeri e le espressioni seguenti siano scorrette matematicamente, valgono come mnemonici:
$+\infty+\infty=+\infty$ e analogamente $-\infty+(-\infty)=-\infty$
$+\infty\cdot(+\infty)=+\infty$ e analoghe 3 coi $-$ e il prodotto dei segni.
$\diamond$ Nelle prossime 3 indichiamo con L un limite finito:
$+\infty+L=+\infty$ p.es. $\lim_{x\to+\infty}(x^3+\arctan x)=+\infty$
$-\infty+L=-\infty$, p.es. $\lim_{x\to+\infty}(\log_{4-\pi} x+(4-\pi)^x)=-\infty$
$+\infty\cdot L=+\infty$ se $L>0$ e analoghe 3 per $L<0$, e $-\infty$.

Restano fuori queste **forme di indecisione**, da risolvere caso per caso: $+\infty-\infty$, $\frac{0}{0}$, $\frac{\infty}{\infty}$, $0\cdot\infty$, ∞^0, 1^∞. P.es. $x-\sqrt{x}=\sqrt{x}(\sqrt{x}-1)$.
Tutte le funzioni considerate in questa trattazione sono definite in singoletti o intervalli o unione di singoletti e/o intervalli.
I singoletti non hanno rilevanza per quanto riguarda i limiti.
Sono significativi solo gli intervalli del dominio delle funzioni.
P. es. $\frac{1}{x}$ è definita nell'unione dei 2 intervalli $]-\infty,0[$ e $]0,+\infty[$; considereremo i limiti a quegli estremi: $-\infty$, 0^-, 0^+, $+\infty$.

Esercizi. Per $x\sqrt[3]{x}$, $\log_2|2^x-1|$, $\log_{|x|}2$, $\sin 2^x$, $\cos\pi x$, $\mathrm{cosec}\pi x$,

$$\frac{x^3-8}{x^2-2x},\ \frac{4+3x-x^2}{x^2-2x-8}$$ questi 8 limiti: $\lim\limits_{x\to\pm\infty}$, $\lim\limits_{x\to-2^\pm}$, $\lim\limits_{x\to 0^\pm}$, $\lim\limits_{x\to 4^\pm}$.

Gli stessi 8 limiti per queste 3 funzioni: $\dfrac{2^x-2^{-x}}{3^x\pm 3^{-x}}$, $\log_{2^x}\left|1-\dfrac{2}{x}\right|$.

25 Derivata; teoremi algebrici sulle derivate

25.1 Derivata in un punto e funzione derivata

In questa trattazione elementare considereremo solo le derivate delle funzioni reali di variabile reale.
Essenzialmente, la derivata $f'(x)$ di una funzione f in un punto x, è un <u>numero</u> che se esiste rappresenta la pendenza del grafico in $(x, f(x))$; e se non esiste vuol dire che in quel punto il grafico non ammette retta tangente, per qualche sorta di sua non liscezza. La derivata positiva ($f'(x) > 0$) corrisponde alla crescenza e quella negativa alla decrescenza. Al variare di x nel dominio di f si ottiene una <u>funzione</u> $f'(x)$ che si chiama [funzione] derivata.
Tutto questo richiede molte precisazioni.
La derivata [come numero] è definita da

$$f'(x) := \lim_{z \to x} \frac{f(z) - f(x)}{z - x} \; \textit{ovvero} \; := \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

per ogni x per cui i limiti (equivalenti) esistono, anche se infiniti.
La *funzione* derivata ha la stessa definizione ma solo per ogni x per cui i limiti esistono *finiti*. L'argomento del secondo limite qua sopra si chiama *rapporto incrementale* di f in x.
(La maggioranza dei testi, trattando più separatamente le derivate come numero e come funzione, per il primo caso scrive $f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$ ovvero $:= \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$, derivata in x_0).
Si dimostra (teorema) che se in $(x, f(x))$ esiste la tangente al grafico essa forma con l'asse x un angolo orientato α tale che $\tan \alpha = f'(x)$. La derivata di $f'(x)$ è la *derivata seconda* $f''(x)$ o $f^{(2)}(x)$ eccetera.
Esempi. Si trova facilmente che la derivata di x^2 è $2x$. Dal che, la pendenza del grafico (parabola) in 0 è 0 (tangente orizzontale) e la funzione per $x < 0$ decresce e per $x > 0$ cresce. Si trova facilmente che la derivata di $|x|$ è $\frac{x}{|x|}$ corrispondentemente all'inesistenza della tangente al grafico in 0 (ovvero, in $(0, 0)$).

25.2 Derivate delle funzioni elementari reali

Tutte queste formule valgono nell'intersezione dei domini della funzione derivanda e della funzione derivata. Per esempio $D\ \ln x = \frac{1}{x}$ vale per $x > 0$ sebbene $\frac{1}{x}$ esista anche per $x < 0$, e $D\sqrt{x} = \frac{1}{2\sqrt{x}}$ vale per $x > 0$ sebbene la derivanda $\sqrt{x}$ esista anche per $x = 0$.

$$
\begin{array}{llcll}
\forall c & D\,c & = & 0 & \\
 & D\,x & = & 1 & \\
 & D\,x^{\alpha} & = & \alpha\, x^{\alpha-1} & \text{in particolare:} \\
 & & & & Dx^2 = 2x \\
 & & & & D\frac{1}{x} = -\frac{1}{x^2} \\
 & & & & D\sqrt{x} = \frac{1}{2\sqrt{x}} \\
 & & & & D\sqrt[3]{x} = \frac{1}{3\sqrt[3]{x^2}} \\
 & D\sin x & = & \cos x & \\
 & D\cos x & = & -\sin x & \\
 & D\tan x & = & \frac{1}{\cos^2 x} & = 1 + \tan^2 x [= \sec^2 x] \\
 & D\,\mathrm{cotan}\, x & = & -\frac{1}{\sin^2 x} & = -1 - \mathrm{cotan}^2 x [= -\mathrm{cosec}^2\, x] \\
 & D\ \arcsin x & = & \frac{1}{\sqrt{1-x^2}} & \\
 & D\ \arccos x & = & -\frac{1}{\sqrt{1-x^2}} & \\
 & D\arctan x & = & \frac{1}{1+x^2} & \\
 & D\,\mathrm{e}^x & = & \mathrm{e}^x & \\
\forall a > 0 & D\,a^x & = & a^x \ln a & \\
 & D\ \ln x & = & \frac{1}{x} & \text{vale anche } D\ln|x| = \frac{1}{x} \\
\forall\, 0 < b \neq 1 & D\ \log_b x & = & \frac{1}{x} \cdot \frac{1}{\log b} &
\end{array}
$$

Ad un livello superiore si considerano anche:

$$
\begin{array}{lcll}
D\ \sinh x & = & \cosh x & \\
D\ \cosh x & = & \sinh x & \\
D\tanh x & = & \frac{1}{\cosh^2 x} & [= 1 - \tanh^2 x = \mathrm{sech}^2\, x] \\
D\,\mathrm{arsinh}\, x & = & \frac{1}{\sqrt{x^2+1}} & \\
D\,\mathrm{arcosh}\, x & = & \frac{1}{\sqrt{x^2-1}} & \\
D\,\mathrm{artanh}\, x & = & \frac{1}{1-x^2} & [= D\,\mathrm{arcoth}\, x] \\
D\,|x| & = & \frac{x}{|x|} = \frac{|x|}{x} & [= \mathrm{sgn}\, x,\ \forall\, x \neq 0]
\end{array}
$$

25.3 Teoremi algebrici sulle derivate

In tutto questa Lezione, la somma è somma di funzioni, e così il prodotto e il quoziente: per esempio $(f+g)(x) := f(x)+g(x)$.
La derivata della somma [di 2 funzioni] è la somma delle derivate [di quelle 2 funzioni]:

$$(f+g)' = f' + g'$$

che potremmo anche scrivere $D(f+g) = Df + Dg$ ma continueremo con la notazione dell'apice per la derivata prima.
La derivata del prodotto non è il prodotto delle derivate bensì

$$(f \cdot g)' = f' \cdot g + f \cdot g' .$$

Derivata della [funzione] reciproca:

$$\left(\frac{1}{f}\right)' = -\,\frac{f'}{f^2}$$

e da questa e dalla precedente si ricava subito la derivata del quoziente:

$$\left(\frac{f}{g}\right)' = -\,\frac{f' \cdot g - f \cdot g'}{g^2} .$$

Derivata della funzione composta:

$$(f(g(x)))' = f'(g(x)) \cdot g'(x)$$

che consente la derivazione di f^g derivando l'equivalente $e^{g \cdot \ln f}$:

$$(f^g)' = f^g \cdot \left(g' \cdot \ln f + \frac{g \cdot f'}{f}\right)$$

ricordando la soprastante formula di derivazione del prodotto, e la derivata della funzione esponenziale e della funzione logaritmo.
Esempi. Con la prima, terza e quinta formula, ricordando la derivata dell'arcotangente, si troverà $D\left(\operatorname{arctg} x + \operatorname{arctg} \frac{1}{x}\right) = 0$.
Dalla quinta formula $Df^\alpha = \alpha \cdot f^{\alpha-1} \cdot f'$ per ogni $\alpha \in \mathbb{R}$.

● ● ●

Esempio. Trovare la derivata di $|x^3|$ in 0.
Non possiamo usare la
$D\,|x| = \frac{x}{|x|} = \frac{|x|}{x}$ [$= \operatorname{sgn} x,\ \forall\, x \neq 0$]
perchè non dà nulla in 0.
Il limite del rapporto incrementale di f in x è

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

e con la nostra f in $x = 0$

$$\lim_{h \to 0} \frac{|(0+h)^3| - |0^3|}{h} = \lim_{h \to 0} \frac{|h^3| - 0}{h} = \lim_{h \to 0} \frac{|h^2 \cdot h|}{h} = \lim_{h \to 0} \frac{|h^2| \cdot |h|}{h} =$$

ricordando che $|a^2| = a^2$ e $|a| = a \cdot \operatorname{sgn}(a)$

$$= \lim_{h \to 0} \frac{h^2 \cdot h \cdot \operatorname{sgn}(h)}{h} = \lim_{h \to 0} (h^2 \cdot \operatorname{sgn}(h)) = 0$$

(prodotto di infinitesima per limitata).

26 Prime applicazioni del Calcolo Differenziale

26.1 Tangente, de/crescenza, min/max, concavità, flessi

Definizioni. Si dice che *f è crescente in x_0* se in un intervall(in)o a sinistra di x_0 vale meno che in x_0 e in un intervall(in)o a destra di x_0 vale più che in x_0. (In simboli è alquanto complicato). Con ovvi mutamenti si definisce la funzione decrescente in un punto. Si dice che f è *crescente in un insieme E* se $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$. Se E è un intervallo ciò equivale (si dimostra) alla crescenza in ogni punto di E. Simile equivalenza con la decrescenza, $x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$, sugli intervalli.
Il punto x_0 si dice di *minimo relativo* per f se essa in un intervall(in)o a sinistra di x_0 e in un intervall(in)o a destra di x_0 vale *più* che in x_0. Se *meno*, si ha un *punto di massimo relativo.*
Il punto x_0 si dice di *flesso* per f se essa in un intervall(in)o a sinistra di x_0 *volge la concavità verso l'alto* e in un intervall(in)o a destra verso il basso, oppure viceversa; $(x_0, f(x_0)))$ si dice *flesso.*
Se $\forall\, x \in \mathrm{dom} f$ è $f(x) \geq f(x_0)$, $f(x_0)$ si dice *minimo assoluto* di f, e *massimo assoluto* se $\forall\, x \in \mathrm{dom} f$ è $f(x) \leq f(x_0)$.

Teorema 1. La **tangente** al grafico di $f(x)$ in $(x_0, f(x_0))$ è $y = f(x_0) + f'(x_0)(x - x_0)$ e allora si ha l'approssimazione

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \qquad \forall\, x \approx x_0\,.$$

Per esempio per $x \approx 0$ è $\sin x \approx x$, $\ln(1+x) \approx x$, $\sqrt{1+x} \approx 1 + \frac{x}{2}$.
Teorema 2. $\quad f'(x_0) > 0 \Rightarrow f$ **crescente** in x_0

$$f'(x_0) < 0 \Rightarrow f \text{ decrescente in } x_0\,.$$

(Se $f'(x_0) = 0$ allora f in x_0 può essere crescente o decrescente o nè crescente nè decrescente: si considerino x^3, $-x^3$, x^2).
Teorema 3. Se f è crescente prima di x_0 e decrescente dopo x_0 allora x_0 è un punto di massimo relativo.ì Similmente scambiando *crescente* e *decrescente* si ha un **punto di minimo relativo**.
Teorema 4. Se $f''(x) > 0$ in un intervallo, in esso $f(x)$ *volge la **concavità** verso l'alto*, e verso il basso se < 0.

26.2 Funzioni iperboliche; notazioni per le derivate

Le funzioni iperboliche hanno un interesse di per sè, tendendo a ricorrere nella Fisica, ma hanno anche uno speciale interesse nel Calcolo Differenziale a causa delle derivate di alcune di esse.

	$\sinh x := \frac{e^x - e^{-x}}{2}$	*seno iperbolico*
	$\cosh x := \frac{e^x + e^{-x}}{2}$	*coseno ip.*
	$\tanh x := \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	*tangente ip.*
sono inverse,	$\operatorname{arsinh} x := (\sinh)^{-1} x$	*arcoseno ip.*
assolutamente	$\operatorname{arcosh} x := \left(\cosh\Big\vert_{x\geq 0}\right)^{-1} x$	*arcocoseno ip.*
non reciproche	$\operatorname{artanh} x := (\tanh)^{-1} x$	*arcotangente ip.*

Primo *studio di funzione*, molto parziale: $f(x) := \sinh x$ ha dominio $\mathbb{R}$, limiti $\pm\infty$ a $\pm\infty$ rispettivamente, derivata $f'(x) = \cosh x = \frac{e^x + e^{-x}}{2} > 0$ da cui la crescenza ovunque. Derivata seconda $f''(x) = \sinh x = \frac{e^x - e^{-x}}{2}$ positiva per $x > 0$ (concavità verso l'alto) e negativa per $x < 0$ (concavità verso il basso), e punto di flesso in $x = 0$, flesso in $(0, 0)$ con tangente di flesso $y = x$ perchè $f(0) = 0$ e $f'(0) = \cosh 0 = 1$.
Un altro studio di funzione, per $\cosh x$ troverebbe $\operatorname{dom} f = \mathbb{R}$, limite $+\infty$ a $\pm\infty$, decrescenza per $x < 0$ e crescenza per $x > 0$, concavità verso l'alto $\forall\, x$, minimo assoluto 1 in 0, nessun flesso.

• • •

Notazioni. Funzione derivata di $f(x)$:

$$f'(x) \qquad Df(x) \qquad \dot{f}(x) \qquad \frac{d}{dx}f(x) \qquad \frac{df}{dx}(x)$$

Derivata di $f(x)$ in x_0:

$$f'(x_0) \qquad (Df)(x_0) \qquad (Df)_{x=x_0} \qquad \dot{f}(x_0) \qquad \left.\frac{df}{dx}\right|_{x=x_0}$$

Alcune delle numerose notazioni per le derivate successive: $f''(x)$

$$\ddot{f}(x) \quad f^{(2)}(x) \quad f'''(x) \quad f^{(3)}(x) \quad f^{(n)}(x) \quad \frac{d^2}{dx^2}f(x) \quad \frac{d^n}{dx^n}f(x).$$

27 Regola di de l'Hospital; asintoti

27.1 Regola di de l'Hospital

Detti u_0 e l due numeri o $+\infty$ o $-\infty$, se per $x \to u_0$

$$\left.\begin{array}{rl} & f(x) \to 0 \\ \wedge & g(x) \to 0 \\ \wedge & \frac{f'(x)}{g'(x)} \to l \end{array}\right| \quad \vee \quad \left|\begin{array}{rl} & f(x) \to \pm\infty \\ \wedge & g(x) \to \pm\infty \\ \wedge & \frac{f'(x)}{g'(x)} \to l \end{array}\right.$$

si dimostra (teorema detto Regola di de l'Hospital) che

$$\lim_{x \to u_0} \frac{f(x)}{g(x)} = l$$

se f e g sono "sufficientemente regolari" (come sono in genere le funzioni che ricorrono nelle Scienze Applicate, e negli esercizi elementari, anche di questa trattazione).
Si noti che l'eventuale inesistenza del limite del rapporto delle derivate non esclude l'esistenza del limite del rapporto iniziale.
Esempi.

$$\lim_{x \to 0} \frac{\sin x}{x} \left(= \frac{0}{0} \right) <=^H \lim_{x \to 0} \frac{\cos x}{1} = \frac{\cos 0}{1} = 1$$

che si chiama **secondo limite fondamentale** (e per esso non vale *come dimostrazione* il calcolo soprastante, esiste una dimostrazione specifica; e proprio da quel limite si dimostra che $D \sin x = \cos x$, che quassù viene utilizzato).

$$\lim_{x \to +\infty} \frac{e^x}{x} \left(= \frac{\infty}{\infty} \right) <=^H \lim_{x \to +\infty} \frac{e^x}{1} = +\infty\,.$$

$$\lim_{x \to +\infty} \frac{x}{\ln x} \left(= \frac{\infty}{\infty} \right) <=^H \lim_{x \to +\infty} \frac{1}{1/x} = \lim_{x \to +\infty} x = +\infty\,.$$

Da questi 2 limiti si ha questa sequenza di funzioni "sempre più infinite" in $+\infty$, cioè tali che il rapporto di una di esse con una precedente tende a $+\infty$: (lentissima) $\ln x$, x, e^x (velocissima).
Il $\lim_{x \to +\infty} \frac{x^2+5e^x+\pi}{7x^3+3x^2+1}$ si risolverà con 3 applicazioni successive del teorema. Invece $\lim_{x \to +\infty} \frac{2x+\sin x+\cos x}{5x-\sin x+3\cos x}$ semplificando per x.

27.2 Asintoti

Definizioni degli asintoti. Se x_0 è un numero [finito] e

$$\lim_{x\to x_0^-} f(x) = \pm\infty \quad \vee \quad \lim_{x\to x_0^+} f(x) = \pm\infty$$

allora la retta verticale $x = x_0$ si dice ***asintoto verticale*** per f. (Ma alcuni Autori ritengono inutile questa definizione).
Se esistono 2 numeri [finiti] m e q tali che

$$\lim_{x\to+\infty} \frac{f(x)}{x} = m, \qquad \lim_{x\to+\infty}(f(x) - mx) = q$$

allora la retta $y = mx + q$ si dice ***asintoto destro*** (oppure: *per* $x \to +\infty$) per f, in particolare *obliquo* se $m \neq 0$ e *orizzontale* se $m = 0$. Con $x \to -\infty$ si definisce l'eventuale ***asintoto sinistro***.
Esempi. La funzione $f(x) := \frac{1}{x}$ ha asintoto verticale $x = 0$ e asintoto orizzontale $y = 0$: si trovano coi limiti a 0^+, $-\infty$ e $+\infty$.
La funzione $\tan x$ ha gli infiniti asintoti verticali $x = k\frac{\pi}{2}$ per $k \in \mathbb{Z}$.
Si troverà che $y = \frac{\pi}{2}$ è asintoto orizzontale destro per $\arctan x$, e $y = -\frac{\pi}{2}$ è asintoto orizzontale sinistro.
Si troverà che $y = 1$ è asintoto orizzontale destro per $\tanh x$, e $y = -1$ è asintoto orizzontale sinistro.
Si trova che $y = 0$ è asintoto orizzontale sinistro per $\ln(1 + e^x)$ e $y = x$ è asintoto obliquo destro.
Per $f(x) := \sqrt{x}$ si troverebbe $m = 0$ ma q infinito e allora non esiste asintoto destro. (E sinistro non c'è perchè $\mathrm{dom} f = [0, +\infty[$).
Con la Regola di de l'Hospital per $\ln x$ si trova $m = 0$ ma q infinito e allora non esiste asintoto destro. (Sinistro escluso dal dominio).
La funzione $f(x) := \mathrm{e}^{\frac{1}{x}}$ ha asintoto verticale $x = 0$ che si trova col limite per $x \to 0^+$, e asintoto orizzontale [destro e sinistro] $y = 1$.
Esercizi. Si trovino i 2 asintoti di questa funzione considerata da Wikipedia, l'enciclopedia libera, alla voce *Asymptote*:

$$\frac{2x^2 + 3x + 1}{x}$$

e con essi e qualche valore si disegni un grafico approssimativo.
Similmente per la reciproca, che ha 3 asintoti.

28 Teoria dello studio di funzione

28.1 La "ricetta" per lo studio di funzione; sup e inf

Fermo restando che una funzione molto "capricciosa" non può essere studiata con metodi elementari, è però vero che un'infinità di funzioni del tipo di quelle che tendono a capitare nelle Scienze Applicate può validamente studiarsi coi metodi visti finora.
Per quanto possibile si cercherà di seguire questa "ricetta":

1) Dominio
2) Simmetrie e periodicità
3) Zeri [equazione $f(x) = 0$]
4) Segni [disequazione $f(x) > 0$]
5) Limiti [negli "estremi" (in senso lato) del dominio]
6) Asintoti
7) Derivata prima
8) Limiti della derivata prima [non se $\lim_{x\to x_0} f(x) = \pm\infty$]
9) Crescenza/decrescenza, max/min, sup/inf
10) Derivata seconda, e magari suo studio [talvolta proibitivo]
11) Valori calcolati, ed eventualmente tangente in essi
12) Disegno del grafico

Le funzioni 2π-periodiche, si studino in $[0, 2\pi[$ o meglio $]-\pi, \pi]$. Analoghe riduzioni del dominio si attuino per altre periodicità.
Resta da vedere cosa sono $\sup f$, e $\inf f$. Se una funzione ha un massimo assoluto $\max f$, esso è l'*estremo superiore* $\sup f$. Tuttavia, consideriamo la funzione $\arctan x$. Essa per $x \to +\infty$ tende a $\frac{\pi}{2}$ ma senza mai raggiungerlo: vi si avvicina indefinitamente rimanendo sempre minore. In questo caso $\frac{\pi}{2}$ non è certo max – infatti per essere max ci vorrebbe un x_0 tale che arctan là vale prorio quel valore, il che non succede mai – ma si dice che è estremo superiore. Similmente, si potrebbe definire l'estremo inferiore $\inf f$; in questo caso $\inf f = -\frac{\pi}{2}$. Ovvio poi è il significato di $\sup f = +\infty$, e di $\inf f = -\infty$.
Esercizi. Si studino $\ln(1 + e^x)$, $\sin x + \cos x$, $\sqrt{x^2 - 1}$, $\frac{x-3}{x^2-4}$, $\frac{x^2-4}{x-3}$, $\sqrt{x} - x$, $x - \sqrt{2x + 3}$, $\ln\left(x^2 - x\right)$, $\frac{x+1}{x^2-2x+1}$, $(1 - x)\sqrt{x}$, $e^{\sqrt{x}}$.

29 Esempi di studio di funzione

- $$f(x) := \frac{x}{2 + x^2}$$
- $$f(x) := \ln(x^2 - 1)$$

VI – Serie numeriche e integrali

30 Serie numeriche

30.1 Introduzione alle serie numeriche

Essenzialmente le serie sono in qualche modo delle “somme infinite”. Estenderemo l'uso del simbolo di sommatoria $\sum$ dal caso finito a quello con infiniti termini: essenzialmente, vogliamo occuparci di $a_0 + a_1 + ... + a_n + ...$ con infiniti termini.
Definizione di serie (numerica). In questa trattazione elementare, una serie (numerica) è una scrittura come queste

$$a_0 + a_1 + ... + a_k + ... \quad \textit{il primo indice poteva essere diverso da } 0 \tag{4}$$

$$\text{o equivalentemente} \quad \sum_{k=0}^{\infty} a_k \quad \textit{anche qua, ovvio} \tag{5}$$

$$\text{per esempio} \quad \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n} + ... \text{ ovvero } \sum_{k=1}^{\infty} \frac{1}{k} \tag{6}$$

e questa dell'esempio si chiama *serie armonica* mentre quest'altra

$$3 + \frac{3}{2} + \frac{3}{4} + \frac{3}{8} + ... + \frac{3}{2^k} + ... \text{ ovvero } \sum_{k=0}^{\infty} 3\left(\frac{1}{2}\right)^k \tag{7}$$

è una (particolare) *serie geometrica* di *ragione* $\frac{1}{2}$.
Definizione di somma di una serie.

$$\sum_{k=0}^{\infty} a_k := \lim_{n \to +\infty} (a_0 + a_1 + ... + a_n) \quad \text{se } \exists\,. \tag{8}$$

Questa formula (8) ci dice che la *somma della serie* al primo membro è il **limite**, se esiste (finito o infinito) della

ridotta o *somma parziale* $a_0 + a_1 + ... + a_n$. ← (normale) numero.

Nota sull'ambiguità notazionale. Le 2 scritture equivalenti

$$\sum_{k=0}^{\infty} a_k \quad \text{e} \quad a_0 + a_1 + ... + a_n + ... \tag{9}$$

denotano sia la *serie* (scrittura), che l'eventuale sua *somma*.

30.2 Approfondimenti sulle serie (numeriche)

***Carattere* delle serie, "tipi" di serie, e serie particolari.**

• Serie *convergente*, con somma s: il limite delle somme parziali è il numero $s \in \mathbb{R}$. Per esempio la *serie ciclometrica*

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + ... + \frac{(-1)^k}{2k+1} + ... = \frac{\pi}{4}\,. \tag{10}$$

• *Serie non convergenti*: quelle *indeterminate* ($\nexists$ limite) come

$$\sum_{k=0}^{\infty}(-1)^k = 1 - 1 + 1 - 1 + ... + (-1)^k + ... \tag{11}$$

(la somma parziale vale alternatamente 1 e 0), quelle *divergenti a* $+\infty$ come $1 + 1 + ... + 1 + ...$ e quelle *divergenti a* $-\infty$.

$\star$ Serie *a termini positivi*: $a_k > 0$ per ogni k. (Similmente le serie a termini *non negativi* se ≥ 0, e negativi, e non positivi se ≤ 0).

$\star$ Serie *a termini di segno alternato*, con ovvio significato.

• *Serie geometrica di ragione* $r \in \mathbb{R}$, fissiamo le idee con $a > 0$:

$$a + ar + ar^2 + ... + ar^k + ... = \sum_{k=0}^{\infty} ar^k = \begin{cases} \frac{a}{1-r} & se\ \ |r| < 1 \\ +\infty & se\ \ r \geq 1 \\ \nexists & se\ \ r \leq -1 \end{cases} \tag{12}$$

Nel terzo caso la serie è indeterminata. Per $a<0$ vale il Teor. 1.

• *Serie esponenziale.* $1 + a + \frac{a^2}{2} + ... + \frac{a^k}{k!} + ... = \mathrm{e}^a$. (Si dimostra).

• La *serie armonica*, prima vista: diverge a $+\infty$.

Teorema 1. $ca_{n_0} + ca_{n_0+1} + ... + ca_k + ... = c(a_{n_0} + a_{n_0+1} + ... + a_k + ...)$ con l'ovvio significato se la serie in parentesi diverge a $+\infty$.

Teorema 2. Se $a_k \not\to 0$ allora la serie è non convergente.

Teorema 3. Se ogni $a_k \geq 0 \Rightarrow$ la serie converge o diverge a $+\infty$.

Teorema 4. (Di Leibniz). Se in una serie a termini di segno alternato $|a_k| \to 0$ e $|a_k|$ è decrescente anche solo in senso debole, cioè $|a_{k+1}| \leq |a_k|$, allora la serie converge.

In particolare $1 - \frac{1}{2} + \frac{1}{3} - ... + \frac{(-1)^{k+1}}{k} + ... = \ln 2$, *serie di Leibniz.*

Esercizi. Determinare il carattere delle serie, e ove possibile il "tipo" (a termini positivi, non negativi...), la denominazione (serie geometrica, eccetera), e se possibile la somma.

• $\arctan 0 + \arctan 1 + \arctan 2 + ... + \arctan k + ...$ ovvero $\sum_{k=0}^{\infty} \arctan k$ è serie a termini non negativi. Il termine generale $a_k \to \frac{\pi}{2}$ e allora $\not\to 0$. Allora la serie diverge a $+\infty$.

• $5+\frac{5}{2}+\frac{5}{3}+...+\frac{5}{k}+...$ ovvero $\sum_{k=1}^{\infty} \frac{5}{k}$ è serie a termini positivi. è 5 volte la serie armonica e allora diverge a $+\infty$.

• $3+\frac{3}{2}+\frac{3}{4}+\frac{3}{8}+...+\frac{3}{2^k}+...$ ovvero $\sum_{k=0}^{\infty} 3\left(\frac{1}{2}\right)^k$ è serie a termini positivi. è 3 volte la serie geometrica $\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k$ di ragione $\frac{1}{2}$ e allora converge a $\frac{3}{1-1/2} = 6$.

• $-7-\frac{7}{2}-\frac{7}{3}+...-\frac{7}{k}+...$ è serie a termini negativi. è -7 volte la serie armonica e allora diverge a $-\infty$.

• $3\arctan 0 - 3\arctan 1 + 3\arctan 2 + ... + (-1)^k 3\arctan k + ...$ è serie a termini di segno alternato. Il termine generale $\not\to 0$ perchè in valore assoluto $|a_k| \to \frac{3}{2}\pi$ e allora la serie non converge.

• $5-\frac{5}{2}+\frac{5}{3}-...+\frac{(-1)^{k+1}5}{k}+...$ è serie a termini di segno alternato, ed è 5 volte la serie di Leibniz, e allora converge a $5\ln 2$.

• • $-4+2-\frac{4}{3}-...+\frac{(-1)^k 4}{k}+...$ è serie a termini di segno alternato, ed è -4 volte la serie di Leibniz, e allora converge a $-4\ln 2$.

•• $1-\frac{1}{4}+\frac{1}{9}...+\frac{(-1)^k}{k^2}+...$ è serie a termini di segno alternato, in valore assoluto $|a_k| = \frac{1}{k^2}$ decrescenti e infinitesimi, e allora converge. (Con somma ≈ -0.8225 secondo Wolframalpha: `Sum[(-1)^k/k^2,{k,1,Infinity}]`).

•• $1+\frac{1}{4}+\frac{1}{9}...+\frac{1}{k^2}+...$ è serie a termini positivi, e allora converge o diverge a $+\infty$. Wolframalpha con `Sum[1/k^2,{k,1,Infinity}]` trova $\frac{\pi^2}{6}$, ma in questa trattazione elementare non lo sappiamo.

• • $\sum_{k=1}^{\infty} e^{\frac{1}{k}}$ è serie a termini positivi, che non tendono a 0, e allora diverge a $+\infty$. Si considerino • • • • $\sum_{k=1}^{\infty} \cos\left(3-\frac{1}{n}\right)$.
• • • • $\sum_{k=1}^{\infty} \frac{\cos k\pi}{k^3}$ • • • • $\sum_{k=0}^{\infty} \pi(\sin 1)^k$ • • • • $\sum_{k=1}^{\infty} \frac{\cos \pi k}{4k+2}$.

31 L'integrale indefinito

31.1 Introduzione all'integrale indefinito

L'integrale indefinito è in sostanza l'anti-derivata:

$$\int \frac{1}{1+x^2}\,dx = \arctan x + c$$
$$\Leftarrow \quad \frac{1}{1+x^2} = D\,(\arctan x + c).$$

(In qualche modo, il segno d'integrale può fare il "salto dell'uguale" trasformandosi in derivata, come un ln fa il "salto" trasformandosi in exp, per esempio $\ln x^2 = 3 \Leftrightarrow x^2 = \mathrm{e}^3$; e viceversa).
Il termine "$+c$" ci ricorda che non solo $\arctan x$ ha derivata $\frac{1}{1+x^2}$, ma anche $\arctan x + 99$ o più qualunque numero reale c. Allora l'integrale indefinito è un insieme di funzioni. Ciascuna di esse si chiama *primitiva* di f. Abbiamo subito questi integrali indefiniti:

	$\int \frac{1}{1+x^2}\,dx$	$=$	$\arctan x + c$	
$\forall\,\alpha$	$\int \alpha\,dx$	$=$	$\alpha x + c$	in particolare: $\int 0\,dx = c$
	$\int \mathrm{e}^x\,dx$	$=$	$\mathrm{e}^x + c$	
	$\int \frac{1}{x}\,dx$	$=$	$\ln\lvert x\rvert + c$	
$\forall\,\alpha \neq -1$	$\int x^\alpha\,dx$	$=$	$\frac{x^{\alpha+1}}{\alpha+1} + c$	in particolare $\int x\,dx = \frac{x^2}{2} + c$ $\int x^2\,dx = \frac{x^3}{3} + c$ $\int \frac{1}{x^2}\,dx = -\frac{1}{x} + c$ $\int \sqrt{x}\,dx = \frac{2}{3}x^{\frac{3}{2}} + c$
	$\int \cos x\,dx$	$=$	$\sin x + c$	*Nota: non diamo alcun*
	$\int \sin x\,dx$	$=$	$-\cos x + c$	*significato al termine dx.*

(Il $+c$ va inteso come la somma di una qualunque costante in ognuno degli intervalli massimali contenuti nel dominio dell'integranda ma senza errare troppo immaginiamo che sia 1 costante; per esempio l'integrale indefinito di $\frac{1}{x}$ è $\ln\lvert x\rvert + c$ e il dominio dell'integranda è costituito dai 2 intervalli $]-\infty, 0[$ e $]0, +\infty[$ e per ciascuno di essi si avrà una costante da sommare a $\ln x$, in modo indipendente).

31.2 Approfondimenti ed esercizi sull'integrale indefinito

Dalle proprietà (teoremi) delle derivate si hanno (non tutte in modo ovvio) queste 4 proprietà (teoremi) degli integrali indefiniti.

$$\int \alpha\, f(x)\, dx = \alpha \int f(x)\, dx \qquad \forall\, \alpha \in \mathbb{R} \tag{13}$$

$$\int (f(x) + g(x))\, dx = \int f(x)\, dx + \int g(x)\, dx \tag{14}$$

Integrazione per parti:

$$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx \tag{15}$$

Integrazione per sostutuzione $t := mx + q$:

$$\int f(m\,x + q)\, dx = \frac{1}{m}\Big(\int f(t)\, dt \Big)_{t = m\,x+q} \; \forall\, m, q \in \mathbb{R}, m \neq 0. \tag{16}$$

(Una formula più complicata che non diamo permette una sostituzione più generale $t := g(x)$, ed è là che si dà senso al dx).

Esempi.

$$\diamond \qquad \int (\sin x + 3\cos x + 7)\, dx =$$

$$=^{(14)} \int \sin x\, dx + \int 3\cos x\, dx + \int 7\, dx =$$

$$=^{(13)} (-\cos x + c_1) + 3\int \cos x\, dx + (7x + c_2) =$$

$$= (-\cos x + c_1) + 3\,(\sin x + c_3) + (7x + c_2) =$$

chiamiamo c la somma delle 3 costanti d'integrazione c_1, c_2 e c_3

$$= -\cos x + 3\sin x + 7x + c\,. \qquad \text{[Verificare derivando]}$$

$$\diamond \qquad \int x\, \mathrm{e}^x\, dx =$$

poniamo $f(x) := e^x$ e $g(x) := x$ per integrare per parti $\mathrm{e}^x x$

$$=^{(15)} x\,\mathrm{e}^x - \int 1 \cdot \mathrm{e}^x\, dx = x\,\mathrm{e}^x - \mathrm{e}^x + c. \qquad \text{[Verificare derivando]}$$

Esercizi.

$$\int \pi \ln(\sqrt{2}\, x)^3 \, dx \qquad \int \pi^2 \lg x^{\sqrt{2}} \, dx \qquad \int \frac{x^2+3}{2x^2+2} \, dx$$

$$\int \pi \ln(\sqrt{2}x - 6)^3 \, dx \qquad \int \mathrm{e}^2 \lg(1-3x)^{\frac{1}{\pi}} \, dx$$

32 L'integrale definito

Sino $F(x)$ e $f(x)$ due dunzioni definite fra a e b, cioè sull'intervallo $[a, b]$ se $a \leq b$ e sull'intervallo $[b, a]$ se $a > b$.
Se $F(x)$ è una primitiva di $f(x)$, cioè $F' = f$, e tale primitiva si può ottenere dall'integrale indefinito di f ponendo $c = 0$ o qualunque altro valore, in questa trattazione elementare **definiremo l'*integrale definito* di $f(x)$ da a a b in questo modo**

$$\int_a^b f(x)\,dx := F(b) - F(a)_{\mathit{scritto\ anche}}\,[F(x)]^b_{a \leftarrow \mathit{incremento\ di\ F\ da\ a\ a\ b}}$$

(è ovvio che qualunque primitiva si scelga, ovvero qualunque c, si ottiene lo stesso valore: c si elide nella sottrazione).
Esso uguaglia l'area (consueta) fra l'asse x e il grafico di f fra a e b se $a < b$ e su $[a, b]$ è $f(x) \geq 0$. *(Sottografico).* (A un livello superiore l'integrale viene definito sostanzialmente proprio con le aree e questo consente di averlo anche per funzioni prive di primitiva$^{\uparrow}$ come $\operatorname{sgn} x$ ma noi non considereremo tali integrali).
Si noti che l'integrale definito è un numero, la primitiva è una funzione, e l'integrale indefinito è un insieme di funzioni.
Esempio 1. Calcoliamo l'area sotto una "campata" della sinusoide, da 0 a π:

$$\int_0^\pi \sin x\,dx = [-\cos x]_0^\pi = -\cos\pi - (-\cos 0) = -(-1) - (-1) = 2.$$

Esempio 2. Calcoliamo l'area del sottografico dell'iperbole equilatera di equazione $y := \frac{1}{x}$ da 1 a 7:

$$\int_1^7 \frac{1}{x}\,dx = [\ln x]_1^7 = \ln 7 - \ln 1 = \ln 7$$

ed è ben evidente che il numero 7 si può sostituire con qualunque numero t, ottenendosi il significato geometrico del logaritmo. Se $0 < t < 1$ il logaritmo naturale è negativo, e infatti l'area del sottografico, da intendersi come *area con segno*, è negativa perchè la base viene percorsa da 1 a t in verso contrario all'orientazione dell'asse x. Si disegni l'iperbole e si stimi (dall'area) $\ln 10$.

33 Integrali: approfondimenti e applicazioni

Regola di Chasles. Per ogni a, b, c, in qualunque ordine,

$$\int_a^b = \int_a^c + \int_c^b \tag{17}$$

(dove si intendano completati i 3 simboli di integrale con $f(x)dx$, che per focalizzare il significato di questo teorema, non è stato scritto esplicitamente).

Esercizio. Con la Regola di Chasles si calcoli $\int_{-1}^{2} |x| \, dx$, si faccia un disegno, e si ricalcoli quell'integrale con le aree della geometria elementare. (Si noto che di $|x|$ non abbiamo dato una primitiva, ma quella funzione coincide con $-x$ fino a 0, e con x da 0 in poi).

PARTE B – MATEMATICHE DELL'INCERTEZZA

Sezione B1 – Calcolo delle probabilità

VII – Probabilita assiomatica ed elementare

34 Introduzione al Calcolo delle Probabilità

34.1 Inquadramento della questione

Il calcolo delle probabilità è una branca della matematica.
Come tale tratta oggetti astratti, ma a differenza di altre branche della matematica i modelli che crea sono in generale molto vicini o collegati alla realtà sensibile: troverete affermazioni su dadi, palline colorate, mortalità e lunghezza della vita, soldi...
Sebbene comprenda molti sottocapitoli, il suo prodotto principale è un numero che è *la probabilità di un evento*, come "la probabilità che la somma di 2 dadi sia 7 è $\frac{1}{6}$" che di per sè potrebbe voler dire poco ma può diventare preziosissimo per *scegliere* fra diverse alternative possibili, confrontandole numericamente; per esempio la probabilità che la somma dei punti di 2 dadi sia 8 è $\frac{5}{36}$ che è meno di $\frac{1}{6}$. Non solo per una scommessa sui dadi fra amici ma anche per delicate scelte mediche o economiche o personali.
Gli enti teorici di base sono gli eventi e la probabilità, definita in 4 modi:

- Concezione classica della probabilità
- Concezione frequentista della probabilità
- Concezione soggettiva della probabilità
- Concezione assiomatica della probabilità

Metodi del calcolo delle probabilità sono l'algebra e la geometria analitica, le funzioni elementari, il calcolo combinatorio, il calcolo differenziale, le serie numeriche, l'integrale definito.

34.2 Concezioni classica e frequentista della probabità

La concezione classica della probabilità è definita da

$$p = \frac{numero\ dei\ casi\ favorevoli}{numero\ dei\ casi\ possibili}$$

e vale se i casi possibili sono ragionevolmente da ritenere *equiprobabili* ovvero aventi la stessa probabilità.
Per esempio la probabilità che la somma dei punti di 2 dadi sia maggiore di 10 è $\frac{3}{36}$ cioè $\frac{1}{12}$ perchè dei 36 casi possibili equiprobabili $(1,1)$, $(1,2)$, $(2,1)$,..., $(6,6)$ sono 3 quelli favorevoli all'evento considerato (somma maggiore di 10) e cioè $(5,6)$, $(6,5)$, $(6,6)$.
Esercizio. Che probabilità c'è che la somma (dei punteggi) di 2 dadi sia un numero primo? E che non lo sia?
La concezione frequentista della probabilità è intuitiva e la vediamo con un esempio. Un farmaco è stato somministrato a 1000 persone con una certa diagnosi di malattia, e dopo 5 anni sono vive 700. Per l'uniformità delle condizioni e l'alto numero di *prove* si tende a ritenere che se quello stesso farmaco verrà somministrato a 1 persona con quella stessa diagnosi di malattia, c'è il 70% di probabilità che sia viva dopo 5 anni (ovviamente 70% cioè 0.7 è $\frac{700}{1000}$).
Se per quella stessa diagnosi di malattia un altro farmaco è stato somministrato a 2000 persone e dopo 5 anni sono vive 1200, si tenderà a ritenere che la sopravvivenza a 5 anni sia del 60%, meno del primo farmaco.
Se per quella stessa diagnosi di malattia un altro farmaco è stato somministrato a 4 persone e dopo 5 anni sono vive 3, ne verrebbe una probabilità di sopravvivenza a 5 anni del 75%, cioè 3 su 4, cioè perfino meglio del primo farmaco, che però è stato provato su 1000 persone, mentre questo solo su 4; e qui siamo giunti al limite della validità di questa concezione, se non vengono fatti ulteriori approfondimenti.
Le probabilità così ottenute si chiamano *probabilità a posteriori*.

35 Concezioni soggettiva e assiomatica

35.1 Concezioni soggettiva della probabilità

La definizione soggettiva della probabilità (di Leonard Jimmie Savage e Bruno de Finetti) la diamo implicitamente attraverso questo esempio. Supponiamo che io possa scommettere sulla vittoria della squadra dei Vispi Volpini della prima partita del campionato e abbia queste idee:
Metto 100 euro e se vince prendo 300 euro: accetto felicissimo
Metto 100 euro e se vince prendo 200 euro: accetto contento
Metto 100 euro e se vince prendo 126 euro: per un pelo accetto
Metto 100 euro e se vince prendo 125 euro: sono indifferente
Metto 100 euro e se vince prendo 124 euro: per un pelo ma rifiuto
Metto 100 euro e se vince prendo 110 euro: rifiuto
Metto 100 euro e se vince prendo 100 euro: ci mancherebbe!
Questo significa che io ritengo che la probabilità di vittoria della squadra è $\frac{4}{5} = 0.8 = 80\%$. In pratica nell'ipotesi intermedia in cui non ho nè interesse a scommettere nè a non scommettere ritengo che se l'evento si verificasse 5 volte, 4 volte la squadra vincerebbe e io avrei preso $4 \cdot 125$ euro cioè 500 euro, esattamente quanto avrei speso nelle 5 scommesse. è ovvio che a 126 euro considero vantaggiosa la scommessa (seppure di poco).
Il pareggiarsi della spesa con l'ipotetica vincita definisce la *probabiità soggettiva* che io attribuisco al verificarsi dell'evento considerato.
La formula è

$$p = \frac{\mathit{costo\ della\ scommessa}}{\mathit{vincita\ nel\ caso\ indifferente}}\,.$$

Nell'esempio $\frac{100}{125}$.
Con questa definizione, un esperto attuario può fissare il premio assicurativo per un evento per il quale non sia disponibile una casistica significativa, come l'immissione sul mercato di un farmaco da parte di una nuova azienda, o un viaggio su Marte; o esiste una casistica assolutamente non omogenea.

35.2 Concezione assiomatica della probabilità

La definizione assiomatica della probabilità è astratta e con definizioni e assiomi crea la probabilità, interpretabile in modo del tutto compatibile con le altre concezioni:
• Ω: *evento certo* (p.es. "il dado fa 1 o 2 o 3 o 4 o 5 o 6").
• $\mathbb{A}$: *σ-algebra degli eventi*, un sottoinsieme delle parti di Ω (cioè di $\mathbb{P}(\Omega)$) sufficientemente regolare, precisamente tale che:
 ⋄ $\emptyset,\ \Omega \in \mathbb{A}$
 ⋄ $(\forall A \in \mathbb{A})\ A^C \in \mathbb{A}$ (dove A^C è il complementare di A)
 ⋄ $(\forall A_1, ..., A_n, ... \in \mathbb{A})\ \cup_{n=1}^{\infty} A_n, \cap_{n=1}^{\infty} A_n \in \mathbb{A}$.
(In $\mathbb{A}$ stanno tutti gli eventi *ragionevolmente* considerabili).
• La *[funzione] probabilità* $P : \mathbb{A} \to [0, 1]$ tale che
 ⋄ $P(\Omega) = 1$ (l'evento certo ha probabilità 100%)
 ⋄ $P(A_1 \cup^* A_2 \cup^* ... \cup^* A_n \cup^* ...) = P(A_1) + ... + P(A_n) + ...$
dove $\cup^*$ indica l'unione disgiunta cioè con intersezione vuota.
• Lo *spazio di probabilità*: la terna ordinata $(\Omega, \mathbb{A}, P)$.
• Infine si definiscono la probabilità condizionata e gli eventi indipendenti, che vedremo.
• Si dimostra (sono teoremi) che valgono, $\forall A, B, A_1, A_2 \cdots \in \mathbb{A}$:
 ⋄ $P(B) = P(B \cap A) + P(B \cap A^C)$
 ⋄ $P(A^C) = 1 - P(A)$ da cui in particolare $P(\emptyset) = 0$
 ⋄ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 ⋄ $P(A \cup^* B) = P(A) + P(B)$
 ⋄ $P(A_1 \cup A_2 \cup \dots) = 1 - P(A_1^C \cap A_2^C \dots)$
 ⋄ $A \subseteq B \Rightarrow P(A) \leq P(B)$
 ⋄ Ls Formula di Bayes, che vedremo.

Per esempio uno spazio di probabilità si ottiene con un dado regolare, $\Omega = \{1, 2, 3, 4, 5, 6\}$, dove $\{1\}$,...,$\{6\}$ sono gli *eventi semplici*, possiamo fissare molto opportunamente $\mathbb{A} := \mathbb{P}(\Omega)$, e la funzione P vale costantemente $\frac{1}{6}$. Per esempio $A := \{2, 3, 5\}$ è l'evento "esce 1 o 3 o 5", ovvero "[esce] [un numero] dispari".
Si calcolino A^C, $P(A^C)$, $\{1, 2, 6\}^C =: B$, $P(A \cap B)$, $P(A \cup B)$, $P(B)$. Poi con lo stesso Ω si trovino altri 2 spazi di probabilità.

36 Probabilità combinatoria, prima parte

La probabilità combinatoria elementare si basa sulla concezione classica della probabilità (quella della probabilità come *casi favorevoli / casi possibili*) e sul calcolo combinatorio.

36.1 Introduzione al Calcolo Combinatorio

Il calcolo combinatorio è costituito da metodi per determinare il numero di elementi di un insieme finito E, cioè la sua cardinalità, indicata con $\#E$ (e da altri con $\mathrm{card}E$ o $|E|$).
Metodi del calcolo combinatorio che considereremo sono:

• L'elencazione con conteggio. Essa è adeguata quando è più semplice dell'applicazione di formule. Per esempio per determinare quanti sono i numeri primi minori di 25

$$2, 3, 5, 7, 11, 13, 17, 19, 23 \qquad \rightarrow \textit{ sono } 9$$

o (col computer) di 10 000.

• L'insiemistica, e in particolare il prodotto cartesiano, l'insieme delle parti e la cardinalità dell'unione;

• permutazioni, combinazioni, disposizioni, dismutazioni$^{\nearrow}$.

Osservato che il prodotto cartesiano di 2 insiemi finiti A e B ha (ovviamente) cardinalità

$$\#(A \times B) = \#A \cdot \#B$$

osserviamo che parallelamente, per così dire, se un'azione si può fare in n modi e una seconda azione si può fare in m modi, la sequenza delle 2 azioni si può fare in $n \cdot m$ modi. (Si consideri un esame A propedeutico a un esame B, il primo con esiti "superato" o "non superato", B coi voti da 18 a 30 e lode: in tutto 28 casi).

Tratteremo il calcolo combinatorio – che di per sè sarebbe una matematica elementare – nel contesto del calcolo delle probabilità. Per esempio, la probabilità che un numero primo minore di 25 scelto a caso sia dispari è

$$\frac{8}{9} \approx 0.889 = 88.9\%.$$

36.2 Cardinalità dell'unione; permutazioni

Cardinalità dell'unione. Se A è un insieme di a elementi e B uno di b elementi, e l'insieme $A \cap B$ ha $r \geq 0$ elementi, allora l'insieme $A \cup B$ ha $a + b - r$ elementi:

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

(formula della cardinalità dell'unione, valida così come è scritta, in questa trattazione elementare, solo per insiemi finiti).
In particolare se A e B sono disgiunti (e allora $r = 0$) la cardinalità dell'unione è la somma della cardinalità:

$$A \cap B = \emptyset \quad \Rightarrow \quad \#(A \cup B) = \#A + \#B\ .$$

In sostanza la seconda formula corrisponde al caso banale in cui semplicemente si sommano i numeri di elementi, mentre la prima formula ci dice che non dobbiamo conteggiare 2 volte gli elementi che sono sia in A che in B.
Esercizio. Quanti sono i numeri < 1000 multipli di 3 o 5?
Permutazioni. Ogni ordinamento totale di un insieme finito non vuoto A si dice *permutazione* degli elementi di A. Esso viene identificato con l'unica $(\#)$-upla di elementi di A che "rappresenta" quell'ordinamento totale. Per esempio dall'unica 5-upla – come (b, d, a, e, c) – che rappresenta un determinato ordinamento totale di A se esso ha 5 elementi.
Il numero delle permutazioni di un insieme di n elementi si indica talvolta con P_n e si pone anche $P_0 := 1$.
Il numero delle permutazioni di un insieme di n elementi è dato dal *fattoriale*↑ di n, che vale 1 se $n = 0$ e altrimenti $1 \cdot 2 \cdot \ldots \cdot n$:

$$P_n = n!\ .$$

Esempio. In quanti modi si possono ordinare le lettere C, R, O, N, I, S, T, A? Che probabilità c'è di ordinarle in quel modo estraendole a caso da un'urna?
è un problema di permutazioni, dell'insieme { C, R, O, N, I, S, T, A }, che ha 8 elementi. Le lettere si possono riordinare in $8! = 40\,320$ modi. La probabilità è $1/43\,320 \approx 0.0000231 = 0.00231\%$.

37 Probabilità combinatoria, seconda parte

37.1 Dismutazioni

Si dice *dismutazione* (o *permutazione completa* o *sconvolgimento*) di un insieme ogni sua permutazione in cui tutti gli elementi cambiano posizione. (Si immaginino degli studenti alle loro sedie, si alzano e "rimescolano", si siedono, ma nessuno sulla sedia che aveva prima). Per esempio per l'insieme ordinato $\{ 1, 2, 3, 4, 5 \}$ ovvero per la 5-upla ordinata $(1, 2, 3, 4, 5)$, due permutazioni sono $(4, 5, 1, 2, 3)$ e $(2, 5, 1, 4, 3)$, ma solo la prima è una dismutazione. Per il numero N di dismutazione di un insieme di n elementi, è (teorema) $N \approx \frac{n!}{e}$; l'errore assoluto è sempre < 0.5; e per $n > 5$ l'errore relativo è minore dello 0.05%, che in questo caso possiamo considerare una buona approssimazione, e $\to 0$ per $n \to \infty$:

$$\lim_{n \to +\infty} \frac{N}{n!/e} = 1.$$

Esempio. Si abbiano molte persone, e altrettanti foglietti coi loro nomi. Li distribuiamo a caso a quelle persone. Che probabilità c'è che nessuno riceva il foglietto col suo nome? E che qualcuno lo riceva? Ovviamente se non sappiamo il numero n di persone, non possiamo dare una risposta esatta. Tuttavia se, come d'usuale, ci accontentiamo di un valore ben approssimato della probabilità, in pratica possiamo, purchè n, che è stato garantito grande, sia ≥ 5 o meglio > 5, e sorprendentemente la risposta non dipende da quel numero esatto di persone. Consideriamo l'evento che nessuno riceva il foglietto con il suo nome. Così per il primo quesito i casi possibili sono le permutazioni e i casi favorevoli sono le dismutazioni, e allora

$$p \approx \frac{n!/e}{n!} = \frac{1}{e} \approx 0.368 = 36.8\% \qquad \scriptstyle (viene\ questo\ \forall n > 5,\ e\ per\ n=5\ viene\ \approx 0.367).$$

E allora la probabilità che qualcuno riceva il foglietto col suo nome (evento complementare) è $\approx 0.632 = 63.2\%$.

Esercizio. Elencare le dismutazioni di 2, 3 e 4 elementi.

37.2 Combinazioni semplici e disposizioni semplici

Consideriamo un insieme $n = 7$ elementi e scegliamone $k = 3$. In quanti modi si può fare? La risposta è data dalle *combinazioni (semplici)* che ora vedremo, ma se anche riteniamo ordinata la terna scelta, si può fare in molti più modi ($3! = 6$ volte tante) e sono le *disposizioni (semplici)*.

Combinazioni semplici. Dato un insieme E di $n > 0$ elementi (cioè non vuoto) e un numero naturale k tale che $0 \le k \le n$, i sottoinsiemi di E di k elementi si chiamano *combinazioni (semplici)* di n oggetti a k a k (per esempio a 3 a 3) e il numero di esse si chiama *coefficiente binomiale*, e si indica con $C_{n,k}$ o $\binom{n}{k}$ e si pone per convenzione $C_{n,0} := 1$. Questo numero si calcola col Triangolo di Tartaglia↓ oppure con una formula di esso

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$ *Numeri enormi li lasceremo indicati, p. es.* $\frac{365!}{23!\,342!}$ *per* $\binom{365}{23}$.

Disposizioni semplici. Dato un insieme E di $n > 0$ elementi (cioè non vuoto) e un numero naturale k tale che $0 \le k \le n$, i sottoinsiemi ordinati di E di k elementi si dicono *disposizioni (semplici)* di n oggetti a k a k (per esempio a 3 a 3), e il numero di esse si indica con $D_{n,k}$ e si pone per convenzione $D_{n,0} := 1$.
è (teorema)

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

Mnemonici:

co mbinazioni... **co**n $k!$ e **co**sì **co**mpare numero **co**rto

di spo sizioni... **di**menticati $k!$ e **dispo**ni ordinatamente

Esempi. 3 elementi ordinati si possono scegliere da 7 in $\frac{7!}{(7-3)!} = \frac{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7}{1\cdot2\cdot3\cdot4} = 5\cdot6\cdot7 = 210$ modi, e senza riguardo all'ordine in $\binom{7}{3} = \frac{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7}{1\cdot2\cdot3\cdot1\cdot2\cdot3\cdot4} = 35$ modi. (Riordinabili in $3!\cdot35 = 210$ modi). Si provi a elencare i 35 modi. (Si fissi per esempio $E := \{1, 2, 3, 4, 5, 6, 7\}$).

Esercizi. In quanti modi si possono scegliere 23 giorni diversi del 2018? Con o senza significato, quante "parole" di 4 lettere si possono comporre con le lettere C, R, O, N, I, S, T, A?

• • •

Esempio sulle combinazioni semplici. Che probabilità c'è di vincere giocando una cinquina su una ruota del lotto? Per essere sicuri di vincere giochiamo 1 euro su ciascuna di esse. Quanto guadagniamo?
Di casi favorevoli ce n'è 1, e i casi possibili sono le combinazioni semplici di 90 oggetti a 5 a 5, che sono in numero di $\binom{90}{5}$, e allora la probabilità è

$$p = \frac{1}{\binom{90}{5}} = \frac{1}{\frac{90!}{5!(90-5)!}} = \frac{5!\,85!}{90!} = \frac{1 \cdot \ldots \cdot 5 \cdot 1 \cdot \ldots \cdot 85}{1 \cdot \ldots \cdot 90} =$$

$$= \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{86 \cdot 87 \cdot 88 \cdot 89 \cdot 90} = \frac{1}{86 \cdot 87 \cdot 22 \cdot 89 \cdot 3} = \frac{1}{43\,949\,268}\,.$$

Abbiamo speso 43 949 268 euro e 1 cinquina da noi giocata vince; poichè si vince 6 milioni di volte la posta, vinciamo 6 milioni di euro, con un guadagno di −37 949 268 euro. (Ciclopica perdita)↗.
Esempio sulle disposizioni semplici. Quanti sono i numeri esadecimali di 5 "cifre" (da 00000 a FFFFF) con le "cifre" tutte diverse?
è un problema di disposizioni semplici. Consideriamo i sottoinsiemi ordinati di { 0, 1,..., 9, A,..., F } di 5 elementi, che – in base al teorema – sono in numero di

$$D_{16,5} = \frac{16!}{(16-5)!} = \frac{16!}{11!} =$$

$$= \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11 \cdot 12 \cdot 13 \cdot 14 \cdot 15 \cdot 16}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11} = 524\,160.$$

• • •

Esercizio risolto. In un'urna sono contenute queste palline numerate: 6 bianche, 4 nere, 3 rosse e 7 verdi. Se si estrae 1 pallina, che probabilità c'è che sia verde? In quanti modi diversi si possono estrarre 5 palline in modo che si abbiano fra esse palline di tutti i colori? (Si noti che le palline sono distinguibili perchè sono numerate).

Svolgimento.

I casi favorevoli sono 7 e i casi possibili sono 20=6+4+3+7 e allora

$$p = \frac{7}{20} = 0.35 = 35\%.$$

Per avere tutti i 4 colori con 5 palline, uno e uno solo dei colori, indicati con b, n, r e v è ripetuto 2 volte:

$b\,n\,r\,v\,b:\ \binom{6}{2} \cdot 4 \cdot 3 \cdot 7 = \frac{6 \cdot 5}{1 \cdot 2} \cdot 4 \cdot 3 \cdot 7 = 3 \cdot 5 \cdot 4 \cdot 3 \cdot 7$

$b\,n\,r\,v\,n:\ 6 \cdot \binom{4}{2} \cdot 3 \cdot 7 = 6 \cdot \frac{4 \cdot 3}{1 \cdot 2} \cdot 3 \cdot 7 = 6 \cdot 2 \cdot 3 \cdot 3 \cdot 7$

$b\,n\,r\,v\,r:\ 6 \cdot 4 \cdot \binom{3}{2} \cdot 7 = 6 \cdot 4 \cdot \frac{3 \cdot 2}{1 \cdot 2} \cdot 7 = 6 \cdot 4 \cdot 3 \cdot 7$

$b\,n\,r\,v\,v:\ 6 \cdot 4 \cdot 3 \cdot \binom{7}{2} = 6 \cdot 4 \cdot 3 \cdot \frac{7 \cdot 6}{1 \cdot 2} = 6 \cdot 4 \cdot 3 \cdot 7 \cdot 3.$

In tutto $2 \cdot 3 \cdot 7 \cdot (30 + 18 + 12 + 36) = 4032$.

Esercizi.

Che probabilità c'è che un numero < 10 sia primo? E < 25? E nei 2 casi, che sia quadrato? Triangolare? Pari?

38 Indipendenza e Formula di Bayes

38.1 Probabilità condizionata e indipendenza

Si definisce la *probabilità condizionata*

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

per la quale è apparsa conveniente l'interpretazione di probabilità che si verifichi B *sapendo* che si è verificato A. Nel caso che sapere che si è verificato A non muti per noi la probabilità che si verifichi B, cioè $P(B|A) := P(B)$, si ottiene, da quest'equazione e dalla precedente,

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

e quest'equazione definisce assiomaticamente il caso di 2 *eventi indipendenti*. E si estende a 3 o più eventi.↑

Fissiamo l'attenzione per esempio sullo *spazio di probabilità uniforme* $(\Omega, \mathbb{A}, P)$ del lancio di 1 dado, con 6 *eventi semplici* di probabilità $\frac{1}{6}$, e $64 = 2^6$ eventi nella σ-algebra $\mathbb{A}$ delle parti di Ω:

$$\Omega := \{1, 2, 3, 4, 5, 6\} \qquad P(k) = 1/6, \ k = 1, ..., 6$$

$$\mathbb{A} := \mathbb{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, ...\{6\}, \{1, 2\}, ..., \{1, 2, 3, 4, 5\}, \Omega\}.$$

Per esempio $\{1, 2\}$ è l'evento "esce 1" *vel* "esce 2"; $P(1, 2) = \frac{2}{6} = \frac{1}{3}$.

Naturalmente esistono spazi di probabilità finiti non uniformi, per esempio ottenuti con dadi truccati, e anche spazi di probabilità infiniti, per i quali nessuna uniformità è possibile.

Ricordiamo che $A \cap B$ è l'evento intersezione di 2 eventi A e B, per esempio (scrivendo in forma logica *et*)

$$\text{“}pari\text{”}\, et \,\text{“}primo\text{”} = \{pari\} \cap \{primo\} = \{2, 4, 6\} \cap \{2, 3, 5\} = \{2\}.$$

Quei 2 eventi non sono indipendenti perchè

$$P(\{2, 4, 6\} \cap \{2, 3, 5\}) = P(2) = \frac{1}{6} \neq \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(2, 4, 6) \cdot P(2, 3, 5).$$

Sono indipendenti per esempio "pari e "quadrato": si verifichi. Si trovi un'altra coppia di eventi indipendenti della stessa σ-algebra.

Si noti l'immediata $P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$.

38.2 Formula di Bayes

Si dimostra (teorema) la Formula di Bayes: dati un $B \in \mathbb{A}$ e una partizione di Ω, cioè degli insiemi $A_1, \ldots A_n$ disgiunti con unione Ω, vale

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)}\,.$$

Anche la sola uguaglianza dei denominatori è interessante:

$$P(B) = \sum_{k=1}^{n} P(A_k)P(B|A_k)$$

e in particolare con $n = 2$, scrivendo A invece di A_1 ed essendo necessariamente $A_2 = A^C$ (avendosi una partizione) si ha la

$P(B) = P(A)P(B|A) + P(A^C)P(B|A^C)$ (Legge delle Alternative)

Si faccia un disegno rappresentativo della situazione.

Esercizio$^{f}_{\mu}$ **risolto.** Un'urna U contiene 20 palline bianche e 40 nere, e un'urna V contiene 5 palline bianche e 6 nere. Si sceglie a caso un'urna e si estrae a caso una pallina che risulta bianca. Che probabilità c'è che sia stata scelta l'urna U?

Svolgimento$_{\mu}$

$U : 20b + 40n$

$V : 5b + 6n$

A_1: "scelta a caso l'urna U"

A_2: "scelta a caso l'urna V"

B: "estratta a caso una pallina bianca".

A_1 e A_2 costituiscono una partizione dell'evento certo.

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)} =$$

$$= \frac{\frac{1}{2} \cdot \frac{20}{60}}{\frac{1}{2} \cdot \frac{20}{60} + \frac{1}{2} \cdot \frac{5}{11}} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{5}{11}} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{5}{11}} =$$

moltiplicando numeratore e denominatore per $3 \cdot 11$

$$\frac{11}{11 + 5 \cdot 3} = \frac{11}{26} \approx 0.423 = 42.3\%.$$

(Osserviamo che allora con probabilità $\approx 58\%$ era stata scelta l'urna V e questo riflette il fatto che là c'erano proporzionalmente più palline bianche: visto che è venuta una pallina bianca, più probabilmente avevamo scelto quell'urna).

Esercizio$_{\mu}$ Costruire e risolvere un esercizio analogo.

Esercizio risolto$_L$ Tre urne A, B e C contengono palline bianche e nere secondo lo schema

$$A : 2b + 2n \qquad B : 3b + 1n \qquad C : 1b + 3n.$$

Si estrae una pallina da A e si vede il colore:
• se è bianca la si mette in C e poi si estrae una pallina da C e la si mette in A;
• se è nera la si mette in B e poi si estrae una pallina da B e la si mette in A.
Qual è la probabilità di ripristinare la situazione iniziale?
Quali altri casi si possono avere e con quali probabilità?

come all'inizio
$_{2/5}$ ↗○ A ○○●● B ○○○● C ○●●●
C ○○●●●
$_{3/5}$ ↘● A ○●●● B ○○○● C ○○●●
nuovo caso I
B ○○○●
B ○○○● $_{1/2}$ ↗○ A ○●●
A ○○●●
C ○●●● $_{1/2}$ ↘● A ○○●
nuovo caso II
$_{3/5}$ ↗○ A ○○○● B ○○●● C ○●●●
B ○○○●●
$_{2/5}$ ↘● A ○○●● B ○○○● C ○●●●
come all'inizio
C ○●●●

Ripristino della soluzione iniziale: può avvenire in 2 modi, eventi disgiunti; ciascuno dei 2 addendi è un prodotto perchè corrispende ad eventi composti (cioè $\cap$ ovvero *et*) indipendenti; probabilità: $\frac{1}{2} \cdot \frac{2}{5} + \frac{1}{2} \cdot \frac{2}{5} = \frac{2}{5} = 0.4 = 40\%$.
Nuovi casi e loro probabilità, valendo le osservazioni soprastanti:
I: $(A : 1b + 3n;\ B : 3b + 1n;\ C : 2b + 2n)$, $p' = \frac{3}{10} = 0.3 = 30\%$,
II $(A : 3b + 1n;\ B : 2b + 2n;\ C : 1b + 3n)$, $p'' = \frac{3}{10} = 0.3 = 30\%$.
Esercizio$_\mu$ Costruire e risolvere un esercizio analogo.

39 Sensibilità, specificità, predittività

Traiamo abbondantemente da Wikipedia, l'enciclopedia libera.
Con il termine sensibilità, in statistica, più precisamente nel campo della epidemiologia, si indica la capacità intrinseca di un test di screening di individuare in una popolazione di riferimento i soggetti malati. Essa è data dalla proporzione dei soggetti realmente malati e positivi al test (veri positivi) rispetto all'intera popolazione dei malati.

Un test sarà tanto più sensibile quanto più bassa risulterà la quota dei falsi negativi (cioè di soggetti malati erroneamente identificati dal test come sani). Un test molto sensibile, in definitiva, ci consente di limitare la possibilità che un soggetto malato risulti negativo al test.

Supponiamo che un test di screening dia come risultato solamente due opzioni: positivo al test e negativo. Essere positivi al test equivale ad essere ammalato, ma indagini diagnostiche successive possono rivelare l'effettiva malattia o meno. Allora si otterranno 4 tipologie di osservati: Sani Negativi (veri negativi), Sani Positivi (falsi positivi), Malati Positivi (veri positivi) e Malati Negativi (falsi negativi), rappresentabili così in tabella:

/////// MALATI - SANI
POSITIVI Veri + Falsi +
NEGATIVI Falsi - Veri -
La sensibilità del test verrà così calcolata:

$$S = \frac{V_+}{totaleMALATI} = \frac{V_+}{V_+ + F_-}$$

ESEMPIO 25 2

4 55

86.2

Con il termine specificità, in medicina, si indica la capacità di un test di dare un risultato normale ("negativo") nei soggetti sani:

$$Sp = \frac{V_-}{totaleSANI} = \frac{V_-}{V_- + F_+}$$

ESEMPIO
25 2
4 55
96.5

Per predittività, in medicina, si intende la probabilità che un soggetto positivo ad un test di screening sia effettivamente malato. Il Valore Predittivo Positivo, che esprime numericamente la predittività, si calcola come quota di soggetti veri positivi sul totale dei positivi (veri e falsi positivi).

$$VPP = \frac{V_+}{totalePOSITIVI} = \frac{V_+}{V_+ + F_+}$$

ESEMPIO
25 2
4 55
92.6

Vediamo ora il caso in cui la prevalenza (frequenza) della malattia è decisamente minore, aumentando ad esempio di un fattore 100 le persone sane e lasciando inalterato il numero dei malati:
25 200
4 5500
11.1

Esercizio. Relativamente all'ultima tabella, si calcolino prevalenza, sensibilità e specificità.

Esempio dell'Excellent e del Modest [...]

VIII – Variabili aleatorie discrete

40 Introduzione alle variabili aleatorie

40.1 Variabili aleatorie discrete e continue

Una funzione (che ad eventi semplici associa numeri reali)

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

si chiama *variabile aleatoria* (purchè sia sufficientemente regolare – come sono tutte quelle che capitano ad un livello elementare – e precisamente deve essere $\{\omega | X(\omega) \leq t\} \in \mathbb{A}$ per ogni t reale). Per esempio una variabile aleatoria è il risultato del lancio di un dado, che potremo rappresentare così in 2 diversi casi:

$$X := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}^{\textit{dado equilibrato}}_{\textit{ovvero regolare}} \qquad Y := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{12} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{4} \end{pmatrix}^{\textit{qua un certo}}_{\textit{dado truccato}}$$

Un'altra variabile aleatoria, non così facilmente rappresentabile, è il peso in kg del primo bambino che nascerà vivo il 1 gennaio 2018, ora locale del luogo. Potrebbe essere 3.412..., 4.576..., eccetera. Ragionevolmente parlando, ognuno dei singoli valori ha probabilità 0: perchè mai il primo bambino dovrebbe pesare *esattamente* (con infiniti decimali) 3.18452785356... kg? Scriveremo

$\{X < 3.5\}$ intendendo l'evento $\{\omega | X(\omega) < 3.5\}$

$\{X \leq 3.5\}$ intendendo l'evento $\{\omega | X(\omega) \leq 3.5\}$

eccetera con $>$ e $\geq$, e ancora, con le possibili variazioni di $<$ e $\leq$ $\{2.5 < X \leq 3.5\}$ intendendo l'evento $\{\omega | 2.5 < X(\omega) \leq 3.5\}$ e per questi eventi, e altri del tipo $X \in I \subset \mathbb{R}$, I sufficientemente regolare, possiamo attenderci probabilità diverse da 0, per esempio $P(X < 20) = 1$ il bambino peserà sicuramente meno di 20 chili, $P(X < 0.1) = 0$ il bambino nato vivo non peserà meno di 100 g, e – ipotizziamo qua – con probabilità 50% peserà meno di 2.1 kg: $P(X \leq 2.1) = 0.5$. E magari ancora $P(X \leq 2.5) = 0.7 = 70\%$.

La funzione $F_X(t) := P(X \leq t)$ si chiama *funzione di ripartizione* della variabile aleatoria X, *cumulative distribution function.*

La prima v.a., quella del dado, si dice *discreta* (cioè a valori "ben separati", anche eventualmente infiniti) e la seconda *continua.*

Teorema (ovvio). $P\{a < X \leq b) = F_X(b) - F_X(a)$, $\forall\, a, b$.

41 Variabili aleatorie uniformi e geometriche

41.1 Densità di variabili aleatorie discrete

Data una v.a. discreta, cioè con valori "ben separati", come per esempio la X e la Y della lezione precedente, o queste

$$Z := \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix} \quad W := \begin{pmatrix} 1 & 2 & 3 & \dots & k & \dots \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \dots & \frac{1}{2^k} & \dots \end{pmatrix} \begin{matrix} \leftarrow infiniti\ valori\ possibili\ ma\ staccati \\ \\ \leftarrow qua\ si\ forma\ una\ serie\ geometrica \end{matrix}$$

individuiamo dei *valori* x_k, in numero finito o anche infiniti ma comunque "ben separati" (non necessariamente interi) e corrispondentemente le loro probabilità $p_k := P(v.a. = x_k)$:

$$X := \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_k & [\dots] \\ p_1 & p_2 & p_3 & \dots & p_k & [\dots] \end{pmatrix} \begin{matrix} \leftarrow fra\ le\ quadre\ possono\ esserci\ numeri,\ se\ i \\ \\ valori\ sono\ infiniti,\ oppure\ no\ nell'altro\ caso \end{matrix}$$

e ovviamente deve essere, nel senso di una somma o di una serie,

$$somma\ su\ tutti\ gli\ indici \rightarrow \sum_k p_k = 1 \quad (esistono\ serie\ con\_somma\ 1,\ non\ solo\ geom.)$$

La funzione di ripartizione $F_X(t) := P(X \leq t)$ di una v.a. discreta X è una *funzione a scala (step function)* costante fra x_{k-1} e x_k, con salti pari a p_k in x_1, x_2,...,x_k... e "pallini pieni" a sinistra (continuità a sinistra).

Da adesso consideriamo solo valori interi; allora $p_k = P(v.a. = k)$. La funzione

$$p_k := \begin{cases} P(X = k) & se\ k \in \mathbb{Z} \\ 0 & altrimenti \end{cases}$$

si chiama ***densità*** della variabile aleatoria X. Tenderemo a usare la lettera p coi pedici per le probabilità , mentre per i valori della v.a. possiamo usare x_1, x_2... per una v.a. X e y_1, y_2... per una v.a. Y, eccetera. O anche x_0, x_1... iniziando da 0, o da altro numero, a seconda dei casi. Per esempio per le X e W di prima

$$p_k = P(X = k) := 1/6, \quad k = 1, ..., 6$$

$$p_k = P(W = k) := 1/2^k, \quad k = 1, 2, 3...$$

Le densità si rappresentano graficamente coi bar chart.

41.2 Variabili aleatorie uniforme discreta e geometrica

Una **variabile aleatoria uniforme discreta** di parametri interi a e b ha n valori interi a, $a+1$, $a+2$,...,$b = a+n-1$, con le corrispondenti probabilità tutte uguali $p_k = \frac{1}{n}$ per $k = 1, ..., n$, e la sua distribuzione ovvero legge verrà indicata con $\mathbb{U}\{a, b\}$.
Per esempio il [risultato del lancio di un] dado (regolare) è una v.a. $X \sim \mathbb{U}\{1, 6\}$ e il simbolo $\sim$ lo leggeremo *"con legge"*:

$$X := \begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ \frac{1}{6}\,\frac{1}{6}\,\frac{1}{6}\,\frac{1}{6}\,\frac{1}{6}\,\frac{1}{6} \end{pmatrix} \qquad \textit{equivalentemente}: \; X \sim \mathbb{U}\{1, 6\}.$$

E associando a 0 e 1 le facce di una moneta

$$M := \begin{pmatrix} 0\,1 \\ \frac{1}{2}\,\frac{1}{2} \end{pmatrix} \sim \mathbb{U}\{0, 1\} \text{ rappresenta una moneta } \textit{equilibrata}.$$

Ecco un'altra, $\sim \mathbb{U}\{a, b\}$ con $a=-2$ e $b=4$ e allora $n=b-a+1=7$:

$$Y := \begin{pmatrix} -2 & -1\,0\,1\,2\,3\,4 \\ \frac{1}{7} & \frac{1}{7}\,\frac{1}{7}\,\frac{1}{7}\,\frac{1}{7}\,\frac{1}{7}\,\frac{1}{7} \end{pmatrix} \qquad \textit{equivalentemente}: \; Y \sim \mathbb{U}\{-2, 4\}.$$

Una **variabile aleatoria geometrica iniziante da 1** di parametro p (e dev'essere $0 \leq p \leq 1$) ha valori 1,2,3... con densità

$$p_k = p(1-p)^{k-1}, \qquad k = 1, 2, 3... \quad \textit{somma della serie}: 1$$

$$\text{per esempio con } p := \frac{1}{2} \text{ si ha quella di prima} \quad \begin{pmatrix} 1\,2\,3\,...\; k\; ... \\ \frac{1}{2}\,\frac{1}{4}\,\frac{1}{8}\,...\,\frac{1}{2^k}\,... \end{pmatrix}$$

$$\text{e con } p := \frac{1}{3} \text{ si ha} \quad \begin{pmatrix} 1 & 2 & 3\,...\; k & ... \\ \frac{1}{3} & \frac{2}{9} & \frac{4}{27}\,...\,\frac{1}{3}\left(\frac{2}{3}\right)^{k-1} & ... \end{pmatrix}$$

Una **variabile aleatoria geometrica iniziante da 0** di parametro p (e dev'essere $0 \leq p \leq 1$) ha valori 0,1,2... con densità

$$p_k = p(1-p)^{k}, \qquad k = 0, 1, 2... \quad \textit{somma della serie}: 1$$

$$\text{per esempio con } p := \frac{1}{2} \text{ si ha} \quad \begin{pmatrix} 0\,1\,2\,...\; k\; ... \\ \frac{1}{2}\,\frac{1}{4}\,\frac{1}{8}\,...\,\frac{1}{2^k}\,... \end{pmatrix}$$

$$\text{e con } p := \frac{1}{3} \text{ si ha} \quad \begin{pmatrix} 0 & 1 & 2\,...\; k & ... \\ \frac{1}{3} & \frac{2}{9} & \frac{4}{27}\,...\,\frac{1}{3}\left(\frac{2}{3}\right)^{k} & ... \end{pmatrix}.$$

Esempi ed esercizi

Esempio$^{f}_{\mu}$ Determinare la v.a. $\mathbb{U}\{0,7\}$.
Si ha subito in base alla definizione di $\mathbb{U}\{a,b\}$

$$U := \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \sim \mathbb{U}\{0,7\}.$$

Esercizio. Rappresentare graficamente densità e funzione di ripartizione della v.a. U del soprastante esempio.
Esempio$^{f}_{\mu}$ Determinare la v.a. $\mathbb{U}\{-3,1\}$.
Si ha subito in base alla definizione di $\mathbb{U}\{a,b\}$

$$V := \begin{pmatrix} -3 & -2 & -1 & 0 & 1 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix} \sim \mathbb{U}\{-3,1\}.$$

Esercizio. Rappresentare graficamente densità e funzione di ripartizione della v.a. V del soprastante esempio.
Esercizio. Determinare la v.a. $Z \sim \mathbb{U}\{-4,2\}$ e poi calcolare $P(Z \geq 1)$, $P(Z > 1)$, $P(Z \leq 1)$, $P(Z \geq -\pi)$, $P(Z \geq \sqrt{2})$, $P(-1 \leq Z < 1)$, $P(Z^2 = 1)$, $P(Z^2 \leq 1)$. [L'ultime vale $\frac{3}{7}$].
Esercizio. Determinare le v.a. geometriche dei 2 tipi, di parametri $\frac{2}{3}$ e $\frac{4}{5}$, e verificare le somme delle loro serie.
Esercizi$^{f}_{\mu}$ Rappresentare graficamente densità e funzioni di ripartizione delle v.a Z e W della lezione 41. Calcolare $P(W < 2)$, $P(W > 2)$, $P(W \geq 4)$, e con le serie $P(W > 100)$ e $P(W < 100)$.
Esempio$^{f}_{\mu}$ Per una v.a. a 3 valori con $p_1 = \frac{1}{\pi}$ e $p_2 = \frac{1}{\pi^2}$ determinare p_3.
Si tratta di trovare l'unico numero incognito p_3 in

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{1}{\pi} & \frac{1}{\pi^2} & p_3 \end{pmatrix}$$

e da $\frac{1}{\pi} + \frac{1}{\pi^2} + p_3 = 1 \qquad \leftarrow$ somma 1 delle probabilità
segue subito $p_3 = 1 - \frac{1}{\pi} + \frac{1}{\pi^2} \approx 0.5804 = 58.04\%$.
Esercizi. Per una v.a. a 3 valori con $p_3 = \frac{1}{\mathrm{e}}$ e $p_2 = \frac{1}{\mathrm{e}^2}$ determinare p_1. Per una v.a. A a 4 valori 0, 1, 2, 3 con $p_1 = \frac{1}{7}$ $p_2 = \frac{1}{11}$ e $p_0 = \frac{1}{5}$ determinare p_3, e poi $P(A \geq 1)$, $P(A > 1)$, $P(A \leq 1)$, $P(A \geq -\pi)$, $P(A \geq \sqrt{2})$, $P(-1 \leq A < 1)$, $P(A^2 \leq 1)$, $P(A \neq 1)$.

42 Variabili aleatorie binomiali

Si dice che una variabile aleatoria X è binomiale di parametri n e p (e dev'essere n intero ≥ 1 e $0 \leq p \leq 1$) e si scrive $X \sim B(n,p)$, se ha *densità binomiale*, cioè

$$p_k = P(X=k) := \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, ..., n. \quad (18)$$

Rappresenta il numero k di teste che si ottengono in n lanci di una moneta che ha $P(testa) = p$. Più in generale, dà la probabilità di un certo numero k di successi in uno schema successo-insuccesso con n prove (indipendenti) avendo il successo probabilità p ad ogni prova.

Esempio$^{f}_{\mu}$ Con $n = 2$ prove si possono ottenere 0 o 1 o 2 teste (che adesso consideriamo successi) e se la moneta è equilibrata è

$P(k=0) = \binom{2}{0}\left(\frac{1}{2}\right)^0\left(1-\frac{1}{2}\right)^{2-0} = \frac{1}{4} \quad \leftarrow P(0\,teste)$

$P(k=1) = \binom{2}{1}\left(\frac{1}{2}\right)^1\left(1-\frac{1}{2}\right)^{2-1} = \frac{1}{2} \quad \leftarrow P(1\,teste)$

$P(k=2) = \binom{2}{2}\left(\frac{1}{2}\right)^2\left(1-\frac{1}{2}\right)^{2-2} = \frac{1}{4} \quad \leftarrow P(2\,teste).$

Si disegnino densità e funzione di ripartizione.

Osservazione. La probabilità di "$X = h$ *vel* $X = k$", con $h \neq k$, essendo eventi disgiunti, è la somma delle 2 probabilità:

$$P(X{=}h \;\; vel \;\; X{=}k) = P(X{=}h) + P(X{=}k)$$

e similmente con 3 o più valori diversi, e questo vale per le v.a. discrete con qualunque distribuzione.

Esempio$^{f}_{\mu}$ Qual è la probabilità di ottenere [esattamente] 3 volte il numero 1 su 7 lanci di un dado [regolare]?

Il successo ad ognuna delle $n = 7$ prove ha probabilità $\frac{1}{6}$ e allora la probabilità cercata è:

$$P(k=3) = \binom{7}{3}\left(\frac{1}{6}\right)^3\left(1-\frac{1}{6}\right)^{7-3} = \frac{7!}{3!(7-3)!}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^4 =$$

$$= \frac{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7}{1\cdot2\cdot3\;\cdot1\cdot2\cdot3\cdot4}\cdot\frac{5^4}{6^7} = 5\cdot7\cdot\frac{625}{279936} = \frac{21875}{279936} \approx 0.0781 = 7.81\%.$$

Esempi ed esercizi

Esempio$^{f}_{\mu}$ Che probabilità c'è di ottenere almeno 2 volte il numero 7 con 5 lanci di un dado regolare a 8 facce?
Potremmo sommare le 4 probabilità

$$P(k=2)+P(k=3)+P(k=4)+P(k=5)$$

ma è ovvio che conviene considerare l'evento complementare e cioè calcolare

$$\begin{gathered}
1-(P(k=0)+P(k=1))= \\
=1-\binom{5}{0}\left(\frac{1}{8}\right)^{0}\left(1-\frac{1}{8}\right)^{5-0}-\binom{5}{1}\left(\frac{1}{8}\right)^{1}\left(1-\frac{1}{8}\right)^{5-1}= \\
=1-1\cdot 1\cdot\left(\frac{7}{8}\right)^{5}-5\cdot\left(\frac{1}{8}\right)^{1}\left(\frac{7}{8}\right)^{4}= \\
=1-\left(\frac{7}{8}\right)^{4}\cdot\left(\frac{7}{8}-\frac{5}{8}\right)=1-\left(\frac{7}{8}\right)^{4}\cdot\frac{2}{8}=1-\frac{2\,401}{16\,384}= \\
=\frac{13\,983}{16\,384}\approx 0.853=85.3\%.
\end{gathered}$$

Esercizio$_{\mu}$ Che probabilità c'è di ottenere 3 volte il numero 4 con 5 lanci di un dado regolare a 6 facce?
Esercizio$_{\mu}$ Che probabilità c'è di ottenere almeno 2 volte il numero 3 con 5 lanci di un dado regolare a 6 facce?
Esercizio$_{\mu}$ Che probabilità c'è di ottenere almeno 3 volte il numero 3 con 5 lanci di un dado regolare a 6 facce?
Esercizio$_{\mu}$ Che probabilità c'è di ottenere almeno 2 volte numeri primi con 4 lanci di un dado regolare a 6 facce?
Esercizio$_{\mu}$ Che probabilità c'è di ottenere un numero primo di volte testa con 5 lanci di una moneta regolare?
Esercizio$_{\mu}$ Che probabilità c'è di ottenere un numero primo di volte croce con 6 lanci di una moneta regolare?
Esercizio$_{\mu}$ Che probabilità c'è di ottenere un numero primo di volte testa con 7 lanci di una moneta regolare?
Esercizio$_{\mu}$ Che probabilità c'è di ottenere un numero dispari di volte testa con 5 lanci di una moneta regolare?

43 Leggi congiunte e indipendenza

43.1 Introduzione alle leggi congiunte

Considereremo solo coppie di variabili aleatorie discrete ma con qualche attenzione tutto può essere esteso a 3 o più variabili aleatorie. Date 2 variabili aleatorie discrete X e Y

$$X := \begin{pmatrix} x_1\, x_2\, x_3 \ldots x_k\, [\ldots] \\ p_1\, p_2\, p_3 \ldots p_k\, [\ldots] \end{pmatrix} \qquad Y := \begin{pmatrix} y_1\, y_2\, y_3 \ldots y_k\, [\ldots] \\ p_1\, p_2'\, p_3' \ldots p_k'\, [\ldots] \end{pmatrix}$$

si può dimostrare che la coppia ordinata (X, Y) è una v.a., detta *2-dimensionale*. La sua densità è la funzione **densità congiunta**

$$p(x, y) := P(X = x \wedge Y = y)$$

per esempio per i 2 dadi, o per la coppia di lanci di 1 dado,

$$(X, Y) = \begin{pmatrix} (1,1) & (1,2) & \ldots & (6,6) \\ \frac{1}{36} & \frac{1}{36} & \ldots & \frac{1}{36} \end{pmatrix}$$

$$p(x, y) = \begin{cases} \frac{1}{36} & \textit{per } x, y = 1, 2, 3, 4, 5, 6 \\ 0 & \textit{altrimenti} \end{cases} \qquad \textit{che scriveremo anche } p_{i,j}.$$

Detto V il prodotto cartesiano degli insiemi dei valori delle v.a.

$$V := \{x_1, x_2, x_3 \ldots x_k\, [\ldots]\} \times \{y_1, y_2, y_3 \ldots x_k\, [\ldots]\}$$

è ovviamente

$$\sum_{(x,y) \in V} p(x, y) = 1 \tag{19}$$

e se $A \subseteq V$, avendosi eventi disgiunti si ha ovviamente

$$P(A) = \sum_{(x,y) \in A} p(x, y) \tag{20}$$

per esempio per i 2 dadi regolari

$$P(X + Y = 8) = p(2,6) + p(3,5) + p(4,4) + p(5,3) + p(6,2) = \frac{5}{36}.$$

43.2 Densità marginali e indipendenza

Date le 2 v.a. discrete X e Y e la **v.a. bidimensionale** (X, Y) prima considerate, le densità p' di X e p'' di Y si chiamano **densità marginali** di (X, Y). Adesso indichiamo con p la densità congiunta. è (teorema)

$$p'_i = \sum_j p_{i,j} \qquad p''_j = \sum_i p_{i,j}$$

Le 2 v.a. X e Y si dicono **v.a. indipendenti** se

$$p_{i,j} = p'_i \cdot p''_j \qquad \forall\, i, j \quad \text{\small\textit{ovvero se congiunta=prodotto delle marginali.}}$$

Per esempio i 2 dadi o il doppio lancio di prima, con le marginali costantemente $\frac{1}{6}$ e la congiunta costantemente $\frac{1}{36}$. Invece questa densità congiunta

$$p_{i,j} := \begin{cases} \frac{1}{30} & \textit{per } i, j = 1, 2, 3, 4, 5, 6,\ i \neq j \\ 0 & \textit{altrimenti} \end{cases}$$

dà luogo a 2 v.a. (non determinate nei valori per adesso)

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \qquad \begin{pmatrix} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

che non sono indipendenti perchè $p_{1,2} = \frac{1}{30} \neq \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = p'_1 \cdot p''_2$.
Ecco un'altra densità $p_{i,j}$ di v.a. bidimensionale (X, Y)

$$\begin{array}{cccccccl}
 & y_1 & y_2 & y_3 & y_4 & & & \\
x_1 & \frac{4}{44} & \frac{3}{44} & \frac{2}{44} & \frac{2}{44} & \to & \frac{11}{44} & = p'_1 \\
x_2 & \frac{3}{44} & \frac{4}{44} & \frac{2}{44} & \frac{2}{44} & \to & \frac{11}{44} & = p'_2 \\
x_3 & \frac{2}{44} & \frac{2}{44} & \frac{5}{44} & \frac{5}{44} & \to & \frac{14}{44} & = p'_3 \\
x_4 & \frac{2}{44} & \frac{2}{44} & \frac{2}{44} & \frac{2}{44} & \to & \frac{8}{44} & = p'_4 \\
{\scriptstyle p_{4,1}} \nearrow & \downarrow & \downarrow & \downarrow & \downarrow & & & \\
 & \frac{11}{44} & \frac{11}{44} & \frac{11}{44} & \frac{11}{44} & & & \\
 & p''_1 & p''_2 & p''_3 & p''_4 & & &
\end{array}$$

con X e Y non indipendenti:

$$p_{1,2} = \frac{3}{44} \neq \frac{11}{44} \cdot \frac{11}{44} = p'_1 \cdot p''_2.$$

(Lasciamo tutto in 44-esimi sebbene qualche semplificazione sarebbe possibile).

$$X = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ \frac{11}{44} & \frac{11}{44} & \frac{14}{44} & \frac{8}{44} \end{pmatrix} \qquad Y = \begin{pmatrix} y_1 & y_2 & y_3 & y_4 \\ \frac{11}{44} & \frac{11}{44} & \frac{11}{44} & \frac{11}{44} \end{pmatrix}.$$

44 Speranza matematica e varianza

44.1 Speranza matematica e varianza di v.a. discrete

La speranza matematica essenzialmente è un numero che congloba la probabilità degli eventi col loro costo ovvero guadagno. La varianza estende il concetto già visto.
Supponiamo che ora lanceranno un dado, e io posso scegliere o di ricevere 2 euro se viene pari o 5 euro se viene 3: cosa mi conviene scegliere? Se spero nel pari, mediamente vincerei metà volte, e allora posso considerare che mediamente vincerò metà dei 2 euro in palio, cioè 1 euro; se spero nel numero 6, mediamente vincerei 5 volte su 6, e allora posso considerare che mediamente vincerò 5/6 dell'euro euro in palio. Allora conviene la scommessa sul pari.
Il concetto è immensamente più generale dei giochi d'azzardo, perchè non solo con le assicurazioni, ma anche coi fatti dell'economia e della salute pubblica, in ultima analisi, la situazione è analoga ad un gioco d'azzardo, con spese certe e ricavi incerti (purchè ovviamente si riesca a tradurre tutto in euro, il che per la salute e la vita umana è molto problematico, e ovviamente anche determinare le probabilità dei casi è problematico).
Un'attenta considerazione del caso numerico soprastante ci induce a definire come **speranza matematica** (o *media* o *valore atteso*) di una variabile aleatoria discreta $X := \binom{x_1\ x_2\ x_3\ \dots\ x_k\ [\dots]}{p_1\ p_2\ p_3\ \dots\ p_k\ [\dots]}$

$$\mu = E(X) := \sum_k x_k \cdot p_k \qquad \textit{somma su tutti i valori } x_k \textit{ della v.a.}$$

e questa è una somma finita o una serie, e in quest'ultimo caso a livello teorico si richiede anche una condizione di regolarità, $\sum_k |x_k| \cdot p_k < +\infty$ che in questa trattazione elementare diamo ma mai calcoleremo, e comunque è sempre verificata per le variabili aleatorie che considereremo. Se esiste la speranza matematica, si definisce la **varianza** di una v.a. discreta X in 2 modi equivalenti

$$\sigma^2 = Var(X) := \sum_k (x_k - E(X))^2 \cdot p_k = E(X^2) - E(X)^2.$$

44.2 Approfondimenti ed esempi

Teoremi.

Legge	$E(X)$ μ	$Var(x)$ σ^2
$X \sim \mathbb{U}\{a,b\}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
$X \sim B(n,p)$	$n{\cdot}p$	$np(1-p)$
X geometrica iniziante da 0	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
X geometrica iniziante da 1	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Esercizio$^f_\mu$ Qual è il valore atteso del punteggio di un dado? E la varianza?
Usando la definizione di speranza matematica si ha

$$X := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \quad E(X) = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6} =$$

$$= \frac{1}{6}\cdot(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2} = 3.5.$$

Usando il teorema sulla speranza matematica di $\mathbb{U}\{a,b\}$

$$X \sim \mathbb{U}\{1,6\} \Rightarrow E(X) = \frac{1+6}{2} = \frac{7}{2} = 3.5\ (come\ sopra).$$

Con le 2 definizioni di varianza si ha ugualmente

$$Var(X) = \left(1-\frac{7}{2}\right)^2\cdot\frac{1}{6} + \left(2-\frac{7}{2}\right)^2 + ... + \left(6-\frac{7}{2}\right)^2 = ... = \frac{35}{12}$$

$$X^2 = \begin{pmatrix} 1 & 4 & 9 & 16 & 25 & 36 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \quad Var(X) = 1^2\cdot\frac{1}{6} + ... + 6^2\cdot\frac{1}{2} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

e con il teorema sulla varianza di $\mathbb{U}\{a,b\}$ si ha ugualmente

$$X \sim \mathbb{U}\{1,6\} \Rightarrow Var(X) = \frac{(6-1+1)^2-1}{12} = \frac{36-1}{12} = \frac{35}{12}.$$

44.3 Esercizi sulla speranza matematica di v.a. discrete

Esercizio$^{f}_{\mu}$ Qual è la vincita netta attesa giocando 1 euro su una cinquina del lotto? (Con la parola “netta” intendiamo che conteggiamo la perdita sicura di 1 euro, il costo della giocata).
Detta V la v.a. “vincita netta”, ricordando che si vince 6 milioni di volte la giocata, con probabilità $1/\binom{90}{5} = 1/43\,949\,268$,

$$V = \begin{pmatrix} 5\,999\,999 & -1 \\ \frac{1}{43\,949\,268} & \frac{43\,949\,267}{43\,949\,268} \end{pmatrix}$$

$$E(V) = 5\,999\,99 \cdot \frac{1}{43\,949\,268} - \frac{43\,949\,267}{43\,949\,268} = -\frac{37\,949\,268}{43\,949\,268} =$$

adesso magari semplifichiamo per 2, poi ancora per 2, poi per 3

$$= -\frac{3\,162\,439}{3\,662\,439} \approx -0.8635\,euro$$

cioè una perdita media di circa 86 centesimi di euro. (In generale i giochi d’azzardo sono molto meno svantaggiosi di questo).
Esercizio$^{f}_{\mu}$ Qual è la vincita netta attesa giocando 1 euro su un numero della roulette europea?
La vincita avviene con probabilità $1/37$ perchè i numeri equiprobabili sono 37, da 0 a 36, e dà 36 volte la giocata:

$$V = \begin{pmatrix} 35 & -1 \\ \frac{1}{37} & \frac{36}{37} \end{pmatrix}$$

$$E(V) = 35 \cdot \frac{1}{37} - \frac{36}{37} = -\frac{1}{37} \approx 0.027\,euro$$

cioè mediamente perdiamo, molto approssimativamente, 3 centesimo ad ogni giocata di 1 euro.

45 V.a. continua, esponenziale, quantili

45.1 Introduzione alle variabili aleatorie continue

Come abbiamo visto, una v.a. è una funzione $X : \Omega \to \mathbb{R}$ che ad eventi associa numeri e si dice continua se i valori riempiono almeno un intervallo.
In questa trattazione elementare, le v.a. continue che consideremo sono **molto regolari**: **la f.r.** $F_X(z) := P(X \leq z)$ **è continua su tutto $\mathbb{R}$, e derivabile salvo al più in 1 o 2 punti**. La derivata della f.r., completata ad arbitrio in quegli eventuali punti particolari ovvero *singolari*,

$$f(z) := F_X'(z) = DP(X \leq z) \qquad {}_{\textit{spesso denotata } f_X(z)}$$

si chiama **densità** della variabile aleatoria.
Una funzione f è densità di una v.a. se e solo se

$$f(t) \geq 0 \quad \wedge \int_{-\infty}^{+\infty} f(t)\,dt = 1.$$

Se f e F_X sono rispettvamente densità e f.r. di X, allora (teorema)

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt \qquad \forall\, x$$

$$P(X = x) = 0 \qquad \forall\, x$$

$$P(a \leq x \leq b) = \int_a^b f(t)\,dt \qquad \forall\, a, b, \quad \textit{ogni} \leq \ \textit{sostituibile con} \ <$$

e in quest'ultima a e b possono essere infiniti (col $<$, ovvio).
Esempio 1: v.a. $\sim \mathbb{U}[a,b]$ oppure $\mathbb{U}(a,b)$, uniforme su $[a,b]$.

$$f(t) := \begin{cases} \frac{1}{b-a} & \textit{se } a \leq t \leq b \\ 0 & \textit{altrimenti} \end{cases} \qquad F_X(x) = \begin{cases} 0 & \textit{se } x < a \\ \frac{x-a}{b-a} & \textit{se } a \leq x \leq b \\ 1 & \textit{se } x > b \end{cases}.$$

Esempio 2: v.a. esponenziale di parametro λ.

$$f(t) := \begin{cases} \lambda \mathrm{e}^{-\lambda t} & \textit{se } t \geq 0 \\ 0 & \textit{altrimenti} \end{cases} \qquad F_X(x) = \begin{cases} 1 - \mathrm{e}^{-\lambda x} & \textit{se } x \geq 0 \\ 0 & \textit{altrimenti} \end{cases}.$$

45.2 Quantili e mediana di variabili aleatorie continue

Se l'insieme dove f è positiva è un intervallo, anche se illimitato, allora (si dimostra) per ogni $\alpha \in]0,1[$ esiste un'unica soluzione dell'equazione

$$F_X(t) = \alpha$$

che si chiama *quantile di ordine* α, indicato talvolta con q_α:

$$P(X \leq q_\alpha) = \alpha.$$

Si chiama *mediana* il quartile $q_{0.5}$ di ordine 0.5 ovvero 1/2, o 50%.
Esempio. Per una legge esponenziale di parametro 3 calcolare $P\Big(X \geq \frac{2}{3}\ln 2\Big)$.
La densità è

$$f(x) := \begin{cases} 3\mathrm{e}^{-3x} & \textit{se } x \geq 0 \\ 0 & \textit{altrimenti} \end{cases}$$

e allora

$$P\Big(X \geq \frac{2}{3}\ln 2\Big) = \int_{\frac{2}{3}\ln 2}^{+\infty} f(t)\,dt = \int_{\frac{2}{3}\ln 2}^{+\infty} 3\mathrm{e}^{-3t}\,dt = \star$$

calcoliamo l'integrale indefinito con la (16) con $z := -3t + 0$

$$\int 3\mathrm{e}^{-3t}\,dt = \frac{1}{-3}\Big(\int 3\mathrm{e}^{z}\,dz\Big)_{z=-3t} = -\frac{1}{3}\big(3\mathrm{e}^{z}\big)_{z=-3t} = -\mathrm{e}^{-3t}$$

$$\text{riprendiamo } \star = \Big[-\mathrm{e}^{-3t}\Big]_{\frac{2}{3}\ln 2}^{+\infty} = -\mathrm{e}^{-\infty} - \Big(-\mathrm{e}^{-3\frac{2}{3}\ln 2}\Big) =$$

$$_{\textit{intendendo } \mathrm{e}^{-\infty} \textit{ nel senso del limite e allora } 0} = 0 + \mathrm{e}^{-2\ln 2} = \mathrm{e}^{\ln 2^{-2}} = 2^{-2} =$$

$$= \frac{1}{4} = 0.25 = 25\%$$

(e allora $\frac{2}{3}\ln 2$ è il quantile di ordine $\frac{1}{4}$ ossia 0.25 ossia 25%).
Calcolando la probabilità p dell'evento complementare $X < \frac{2}{3}\ln 2$, e poi facendo $1 - p$, si potevano evitare gli infiniti.
Esercizi. Trovare le mediane dell'uniforme e dell'esponenziale. [Si troverà $q_{0.5} = \frac{a+b}{2}$ e $q_{0.5} = \frac{1}{\lambda}\ln 2$ rispettivamente].
Esercizi. Trovare i quartili di ordine $\frac{1}{4}$ e $\frac{3}{4}$ dell'uniforme e dell'esponenziale. Per un'esponenziale di parametro 2 calcolare $P(X \geq 3)$. (Si può calcolare come $1 - P(X < 3)$ evitando l'infinito).

Nota. *Distribuzione* è sinonimo di *legge*, e indica sia la *densità* che la *funzione di ripartizione* (e in via teorica anche ogni altra funzione ad esse equivalente ma in pratica useremo quelle due).

• • •

Proprietà di ogni f.r. $F(x)$ di v.a. continua:
1) ha valori fra 0 e 1, cioè $0 \leq F(x) \leq 1$;
2) è non decrescente: $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$;
3) tende a 0 per $x \to -\infty$ e tende a 1 per $x \to +\infty$;
4) è continua (in trattazioni di livello superiore continua a destra);
5) è crescente ovunque la densità $f = F'$ è positiva.

Proprietà di ogni densità $f(x)$ di v.a. continua:
1) ha valori ≥ 0, sappiamo;
2) ha integrale 1 su tutto $\mathbb{R}$, sappiamo; e poi si dimostra che
3) è continua salvo al più 1 o 2 punti (in questa trattazione).

• • •

3 bis) Tutte le densità che considereremo in questa trattazione tendono a 0 per $x \to -\infty$ e $x \to +\infty$ (ma densità particolarmente "capricciose" potrebbero non avere uno o entrambi quei limiti, con infinite oscillazioni con "campate" strette strette);

Esercizi.

•• I principali quantili considerati in statistica sono:

i *quartili* $q_{0.25}$, $q_{0.5}$, $q_{0.75}$, e il secondo è la mediana;

i *decili* $q_{0.1}$, $q_{0.2}$,...,$q_{0.9}$, corrispondenti a 10%, 20%,...,90%;

i *centili* o *percentili* $q_{0.01}$, $q_{0.02}$,...,$q_{0.99}$, corrispondenti a 1%, 2%,...,99%;

due *ventili*, precisamente $q_{0.05}$ e $q_{0.95}$.

Si calcolino relativamente alla densità esponenziale.
Per esempio per i decili si troverà

$$q_{k/10} = \frac{1}{\lambda} \ln \frac{10}{10-k} \qquad k = 1, ..., 9$$

(Si noti che per $k = 5$ si ottiene la mediana).

•• Per una v.a. X di densità esponenziale di parametro 5 calcolare $P(X < 2)$, $P(X^3 > 2)$, $P(X^2 \geq 4)$.

•• Si consideri una v.a. Z di densità

$$f(z) := \begin{cases} 6z(1-z) & se\ 0 \leq x \leq 1 \\ 0 & altrimenti \end{cases} .$$

Senza fare calcoli ma col disegno del grafico stabilire quanto vale la mediana e mettere in ordine crescente

$$P(Z < 0.4), \quad P(Z > 0.9), \quad P(Z \leq 0.2), \quad P(Z \geq 0.7).$$

46 Speranza matematica, varianza, covarianza

46.1 Speranza matematica, varianza, covarianza

Definizioni. Consideriamo una v.a. X continua di densità f.

$$\textbf{Speranza matematica: } \mu = E(X) := \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

se l'integrale con $|x|$ esiste finito, che non verificheremo mai.
La speranza matematica si chiama anche *valore atteso* o *media*.

$$\textbf{Varianza: } \sigma^2 = Var(X) := \int_{-\infty}^{+\infty} (x-\mu)^2\, f(x)\, dx.$$

Sia per v.a. discrete che continue:

$$\textbf{Deviazione standard: } \sigma := \sqrt{Var(X)} = \sqrt{\sigma^2}.$$

$$\textbf{Covarianza: } Cov(X,Y) := E(X\cdot Y) - E(X)\cdot E(Y)$$

e se la covarianza è 0 le v.a. si dicono *incorrelate*.
Diremo **indipendenti** 2 variabili aleatorie X e Y continue se informazioni sui valori assunti da una non modificano le nostre conoscenze probabilistiche sui valori dell'altra, ovvero e per ogni a, x, b, y con $a < x$ e $b < y$

$$P(a < X < x \wedge c < Y < y) = P(a < X < x) \cdot P(c < Y < y)$$

cioè $\{a < X < x\}$ e $\{c < Y < y\}$ sono eventi indipendenti.

46.2 Approfondimenti sulle v.a. discrete e continue

Tutte le cose che diremo qua valgono sia per variabili aleatorie discrete che continue.
Disuguaglianza di [Bienaymè-]Chebyshev (Čebyšëv):

$$P(|X - E(X)| > c) \leq \frac{Var(X)}{c^2} \qquad \forall\, c \tag{21}$$

che ci dice che minore è la variabilità della grandezza aleatoria X, minore è la probabilità che X assuma valori distanti dalla media. Equivalentemente con l'evento complementare

$$P(|X - E(X)| \leq c) \geq 1 - \frac{Var(X)}{c^2} \qquad \forall\, c. \tag{22}$$

Relazioni di 1 variabile aleatoria con 1 costante:

$$E(c + X) = c + E(X) \tag{23}$$

$$E(c \cdot X) = c \cdot E(X) \tag{24}$$

$$Var(c + X) = Var(X) \tag{25}$$

$$Var(c \cdot X) = c^2 \cdot Var(X) \tag{26}$$

Relazioni di 2 variabili aleatorie:

$$E(X + Y) = E(X) + E(Y). \tag{27}$$

$$X, Y \text{ indip. } \Rightarrow \text{ incorrelate } \Rightarrow E(X \cdot Y) = E(X) \cdot E(Y) \tag{28}$$

$$Var(X + Y) = Var(X) + Var(Y) + Cov(X, Y) \tag{29}$$

$$\text{indip. } \Rightarrow Var(X + ... + X_n) = Var(X_1) + ... + Var(X_n) \tag{30}$$

Esercizio. Calcolare la disuguaglianza di Chebyshev per $c = \sigma$, $c = 2\sigma$, $c = 3\sigma$, $c = 4\sigma$. (Per 4σ è $P(|X - E(X)| > 4\sigma) \leq \frac{1}{16}$).
Esercizi. Siano X e Y indipendenti di leggi esponenziali di parametri $\log 3$ $\log 4$ rispettivamente. Calcolare la loro covarianza, e la speranza matematica di X, $2Y$, $3 + Y$ e $\pi X - Y$.

47 Distribuzioni Gamma e del chi quadrato

47.1 Funzione Gamma, Legge Gamma, alcuni teoremi

La **funzione** $\Gamma(x)$, *funzione Gamma*, è una *funzione speciale* dell'Analisi Matematica, cioè, in pratica e semplificando, una funzione non elementare ma di notevole interesse, con una certa definizione che non daremo. I suoi valori approssimati vengono calcolati da vari software e molti si trovano elencati in tavole numeriche. Ma per i numeri $\frac{1}{2}, 1, 2, 3, 4...$ i suoi valori sono semplici:

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \qquad \Gamma(x) = (x-1)! \qquad \forall\, x \in \mathbb{N}^+$$

(si faccia un grafico) e quelli soli considereremo, per esempio

$$\Gamma(1) = 0! = 1, \quad \Gamma(2) = 1! = 1, \quad \Gamma(3) = 2! = 2, \quad \Gamma(4) = 3! = 6.$$

La **Legge** $\Gamma(\alpha, \lambda)$, Gamma di parametri $\alpha > 0$ e $\lambda > 0$, ha densità

$$f(t) := \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \mathrm{e}^{-\lambda t} & se\ \ t > 0 \\ \\ 0 & se\ \ t \le 0 \end{cases}. \tag{31}$$

Si scrivano esplicitamente e grafichino le densità per $\alpha := \frac{1}{2}, 1, 2, 3$ e $\lambda := 2$, e si calcolino e grafichino le funzioni di ripartizione.
Teoremi.

Legge	$E(X)$ μ	$Var(X)$ σ^2
$X \sim \mathbb{U}[a,b]$ ovvero $\mathbb{U}(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
X esponenziale di parametro λ	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$X \sim \Gamma(\alpha, \lambda)$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$

47.2 Densità e leggi del chi quadrato

La legge $\chi^2(n)$ [del] chi quadrato (chi-quadrato, chi quadro, chi-quadro, inglese *chi-square* o *chi-squared*) ovvero χ^2 di parametro n o come si meglio dice *a n gradi di libertà* ha densità

$$f(x) = f(x;n) := \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\, x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}} & \textit{se } \ x > 0 \\ 0 & \textit{se } \ x \leq 0 \end{cases} . \tag{32}$$

Si scrivano esplicitamente e grafichino le densità per $n = 1, 2, 4, 6$, e si calcolino e grafichino le ultime 3 funzioni di ripartizione.
Di questa famiglia (insieme) di leggi, in Statistica sono molto importanti i quantili. I valori approssimati vengono calcolati da vari software e molti si trovano elencati in tavole numeriche.
Se α è l'ordine di un quantile, $1-\alpha$ viene chiamato $p-$*value*, e il più tipico è 0.05, corrispondente ad $\alpha = 0.95$.
Purtroppo c'è confusione nelle notazioni, con il serio rischio di confondere nella pratica α con $1-\alpha$.
Il quantile q_α di parametro n di solito viene denotato $\chi^2_{1-\alpha}(n)$. In una tavola numerica, si cerchi il valore ≈ 3.84, quantile di ordine $\alpha = 0.95$ ovvero $p = 0.05$, relativo al parametro $n = 1$:

$$X \sim \chi^2(1) \qquad P(X \leq 3.84) \approx 0.95.$$

Ecco alcuni valori:

n	α *di* $\chi^2_{1-\alpha}$ $\to$	0.95	0.90	0.10	0.05	0.01
$\downarrow$	α *di* q_α $\to$	0.05	0.10	0.90	0.95	0.99
1		0.004	0.02	2.71	3.84	6.63
2		0.10	0.21	4.61	5.99	9.21
3		0.35	0.58	6.25	7.81	11.34
4		0.71	1.06	7.78	9.49	13.28
5		1.14	1.61	9.24	11.07	15.09
6		1.63	2.20	10.64	12.59	16.81
7		2.17	2.83	12.02	14.07	18.48
8		2.73	3.49	13.36	15.51	20.09
9		3.32	4.17	14.68	16.92	21.67
10		3.94	4.87	15.99	18.31	23.21

48 Distribuzione t di Student e altre leggi

48.1 Densità e quantili di Student, e legge di Cauchy

Come le leggi del chi quadrato, la t **di Student** è una famiglia di leggi, con un parametro, di solito indicato con n o ν, ancora detto *gradi di libertà*: $n = 1, 2, 3....$ Scriviamo le prime 2:

$$f(t; 1) = \frac{1}{\pi} \cdot \frac{1}{t^2 + 1} \qquad f(t; 2) = \frac{1}{2\sqrt{2}} \cdot \frac{1}{\left(\frac{t^2}{2} + 1\right)^{3/2}}$$

e nelle successive $\frac{t^2}{2}$ diventa $\frac{t^2}{3}$, $\frac{t^2}{4}$... e l'esponente $\frac{3}{2}$ diventa $\frac{4}{2}$, $\frac{5}{2}$...
I grafici un po' si assomigliano: simmetrici rispetto all'asse y, al crescere di n le campane diventano più alte e strette. La costante moltiplicativa coinvolge n e la funzione Γ, ma non t, ed è tale che l'area del sottografico da $-\infty$ a $+\infty$ è 1 per ogni n.
Di questa famiglia (insieme) di leggi, in Statistica sono molto importanti **i quantili**. I valori approssimati vengono calcolati da vari software e molti si trovano elencati in tavole numeriche.
Nella riga di testa si trovano alcuni valori, talvolta con la specificazione *"one tail"*, e per associarli correttamente ad α, evitando di confonderlo con $1-\alpha$, o $2(1-\alpha)$, o $2\alpha-1$, si cerchi nella prima riga (cioè per $n = 1$) il valore ≈ 6.31 corrispondente ad $\alpha = 0.95$:

$X \sim t$ di Student a 1 grado di libertà $\qquad P(X \leq 6.31) \approx 0.95.$

Ecco alcuni valori:

	two tails	$2\alpha - 1$	$\to$	0.9	0.95	0.98	0.99
	two tails	$2(1-\alpha)$	$\to$	0.1	0.05	0.02	0.01
	one tail	$1-\alpha$	$\to$	0.05	0.025	0.01	0.005
k	**one tail**	α	$\to$	0.95	0.975	0.99	0.995
1				$\underline{6.31}34$	12.706	31.820	63.657
...				...	...	...	...
100				1.6602	1.984	2.364	2.625

La $f(t; 1)$ si chiama **densità di Cauchy** e sorprendentemente non ha speranza matematica. Si faccia il grafico di essa, e della corrispondente f.r., che si ottiene integrando da $-\infty$ a x.

48.2 Approfondimenti su χ^2, t di Student, e altre leggi

Teoremi.

Legge	$E(X)$ μ	$Var(X)$ σ^2
$X \sim \chi^2(n)$ ovvero del χ^2 a n gradi di libertà	n	$2n$
$X \sim t$ di Student a n gradi di libertà	0 per $n > 1$	$\frac{n}{n-2}$ per $n > 2$

Osservazioni. Quelle che abbiamo visto, e le densità normale e log-normale che vedremo, sono fra le principali densità continue. Ma ne esistono infinite altre, alcune con un nome specifico, altre senza. Si considerino per esempio gli esercizi seguenti.

Esercizio 1$_\mu$ Calcolare la speranza matematica della densità

$$g(t) := \alpha\, \mathrm{e}^{-2|t|}$$

naturalmente dopo aver determinato la costante α. (La costante viene determinata dall'integrale unitario fra $-\infty$ e $+\infty$, che per la parità della densità è 2 volte l'integrale fra 0 e $+\infty$).

Esercizio 2$_\mu$ Calcolare la speranza matematica della densità

$$f(z) := \begin{cases} 0 & se\ \ z < \frac{1}{2} \\ c & se\ \ \frac{1}{2} \leq z < 1 \\ \frac{c}{z^3} & se\ \ z \geq 1 \end{cases} \qquad \textit{discontinua!}$$

naturalmente dopo aver determinato c. (Ovviamente bisognerà usare la Formula (14), Regola di Chasles, fatta valere su $]a, b[=]-\infty, +\infty[$, suddividendo 2 volte l'integrale, nei punti $\frac{1}{2}$ e 1).

Esercizio 3$_\mu$ Calcolare la speranza matematica della densità

$$\textit{discontinua!} \qquad u(x) := a(2 + \operatorname{sgn}(x))\, per\ -1 \leq x \leq 2,\ e\, 0\, altrimenti.$$

49 Legge e speranza matematica di $g(X)$

49.1 Legge di $g(X)$ e standardizzazione

Data una variabile aleatoria continua X e una funzione sufficientemente regolare g, anche $g(X)$ è una variabile aleatoria continua. Si pensi per esempio a $mX + q$, X^3, $-X^3$ e X^2.
Teoremi. La funzione di ripartizione di $g(X)$ è

$$F_{g(X)}(x) = F_X(g^{-1}(x)) \qquad \text{se } g \text{ crescente suriettiva} \tag{33}$$

$$F_{g(X)}(x) = 1 - F_X(g^{-1}(x)) \qquad \text{se } g \text{ decrescente suriettiva} \tag{34}$$

(le quali vanno bene per esempio per $mX + q$, X^3 e $-X^3$)

$$F_{X^2}(x) = F_X(\sqrt{x}) - F_X(-\sqrt{x}) \qquad \text{se } g(z) := z^2$$

e poi per tutte la densità si ottiene derivando, in particolare

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}}(f(\sqrt{x}) + f(-\sqrt{x})).$$

Esercizio. Si trovino densità e f.r. di $mX + q$ per $m > 0$ e < 0, X^3 e $-X^3$, e poi con una g non nulla a scelta.
Definizione. Per una variabile aleatoria X discreta o continua dotata di speranza matematica μ e varianza σ^2, scegliendo

$$m := \frac{1}{\sigma} \qquad q := -\frac{\mu}{\sigma}$$

la nuova variabile aleatoria $mX + q$

$$Y := \frac{X - \mu}{\sigma}$$

si chiama **standardizzazione** di X.
Teorema.
Se W è standardizzazione di una v.a. discreta o continua

$$E(W) = 0 \qquad Var(W) = 1.$$

Esercizio. Si scrivano le standardizzazioni dei 2 dadi di (40.1), con quella stessa notazione. E di una moneta regolare.

49.2 Speranza matematica di $g(X)$ e momenti

Teorema. Come detto, data una variabile aleatoria continua X e una funzione sufficientemente regolare g, anche $g(X)$ è una variabile aleatoria continua. In condizioni di ulteriore regolarità X e $g(X)$ hanno speranza matematica.
Se la v.a. continua X ha densità f_X allora

$$E(g(X)) = \int_{-\infty}^{+\infty} g(t) f_X(t)\, dt.$$

in particolare (con $g(z) := z^n$)

$$E(X^n)\, dx = \int_{-\infty}^{+\infty} x^n\, f(x)\, dx \quad \text{si chiama } \textit{momento } n\text{-esimo}.$$

Il momento primo è proprio la speranza matematica di X.
Esempi. Vediamo 2 esempi tratti da un libro che si distingue nella didattica del Calcolo delle Probabilità e della Statistica: *Calcolo delle probabilità*, di Paolo Baldi, ed. McGraw-Hill. Alle pp. 125-126:

> Supponiamo che X sia uniforme su $[0, 1]$. Quanto valgono $E[\sin(2\pi X)]$ e $E[e^X]$?(...)
>
> $$E[\sin(2\pi X)] = \int_0^1 \sin(2\pi x)\, dx = 0$$
>
> $$E[e^X] = \int_0^1 e^x\, dx = e - 1.$$

(L'uso delle soprastanti parentesi quadre non è conforme alle standard seguito in questa dispensa, ma si tenga presente che le notazioni fra i vari Autori matematici differiscono alquanto).
Esercizi. Si consideri una v.a. X con $f_X(x) := x$ fra 0 e $\sqrt{2}$ e 0 altrimenti. Trovare $E(\sin(2\pi X))$, $E(e^X)$ ed $E(\ln X)$, e anche un'altra matematica con una funzione a scelta. Poi si cerchino le stesse 4 speranze matematiche con una nuova densità a scelta.

49.3 Esercizi sulla legge di $g(X)$

Esercizio. Si trovi densità e f.r. di X^3.

$$P(X^3 \leq x) = P(X \leq \sqrt[3]{x}) \qquad \text{cioè}$$

$$\boxed{F_{X^3}(x) = F_X(\sqrt[3]{x})}$$

(che si poteva ottenere direttamente dalla (33) dopo aver riconosciuto che $g(x) := x^3$ è crescente suriettiva con inversa $\sqrt[3]{x}$) e derivando, essendo $D\sqrt[3]{x} = D\,x^{\frac{1}{3}} = \frac{1}{3}x^{-\frac{2}{3}} = \frac{1}{3\sqrt[3]{x^2}}$ per $x \neq 0$,

$$\boxed{f_{X^3}(x) = \frac{1}{3\sqrt[3]{x^2}}\, f_X(\sqrt[3]{x}) \quad x \neq 0}$$

e in 0 la densità può essere posta 0 o qualunque altro valore ≥ 0.
Esercizio. Si trovi densità e f.r. di $-X^3$.

$$P(-X^3 \leq x) = P(X^3 \geq x) = 1 - P(X^3 \leq x) = 1 - P(X \leq -\sqrt[3]{x}) \text{ cioè}$$

$$\boxed{F_{-X^3}(x) = 1 - F_X(-\sqrt[3]{x})}$$

(che si poteva ottenere direttamente dalla (34) dopo aver riconosciuto che $g(x) := -x^3$ è decrescente suriettiva con inversa $-\sqrt[3]{x}$) e derivando, essendo $D\sqrt[3]{x} = D\,x^{\frac{1}{3}} = \frac{1}{3}x^{-\frac{2}{3}} = \frac{1}{3\sqrt[3]{x^2}}$,

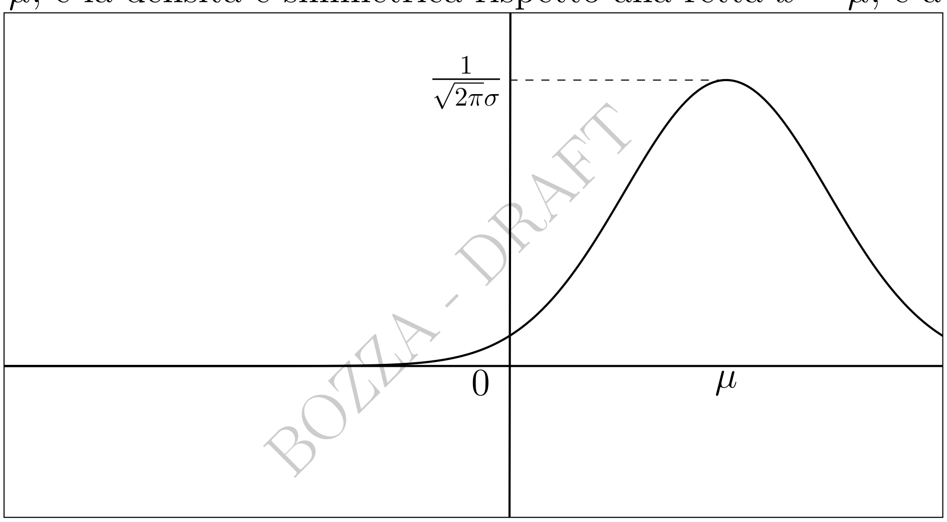
$$\boxed{f_{-X^3}(x) = -\frac{1}{3\sqrt[3]{x^2}}\, f_X(-\sqrt[3]{x})}$$

e in 0 la densità può essere posta 0 o qualunque altro valore ≥ 0.

50 Densità e variabile aleatoria normale

50.1 Introduzione alla densità e v.a. normale

La densità **normale** ovvero **gaussiana** ha un grafico detto "a campana", con limiti 0 a $+\infty$ e $-\infty$, prima crescente e poi decrescente, prima con la concavità verso l'alto, poi verso il basso e infine verso l'alto. La *moda* (cioè l'eventuale unico punto di massimo di una densità di v.a. continua) esiste e coincide con la media μ, e la densità è simmetrica rispetto alla retta $x = \mu$, e a



causa di questa simmetria anche la mediana è μ: la probabilità di un valore prima di μ è uguale alla probabilità di un valore dopo μ. Il massimo assoluto vale $\frac{1}{\sqrt{2\pi}\sigma}$ e allora minore è la varianza maggiore è il massimo e la strettezza della campana. Questa legge è denotata con $N(\mu, \sigma^2)$, ha 2 parametri (come la legge Gamma) ed essi sono proprio la media e la varianza. La densità è

$$f(t) := \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}. \tag{35}$$

Teorema. Se $X \sim N(\mu_1, \sigma_1^2)$ e $Y \sim N(\mu_1, \sigma_1^2)$ sono indipendenti

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

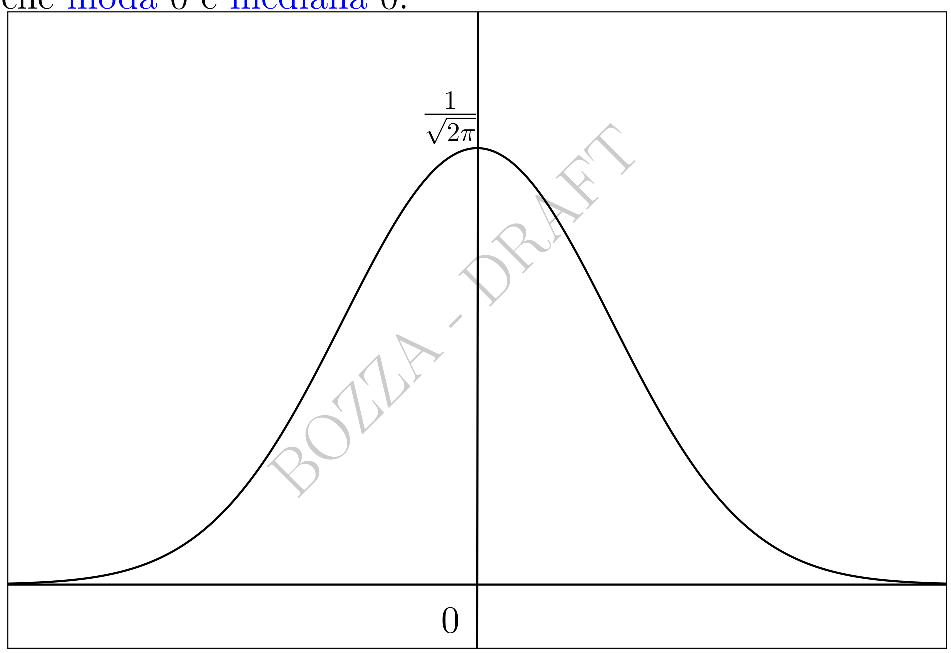
$$c + X \sim N(c + \mu_1, \sigma_1^2) \qquad c\,X \sim N(c\mu_1, c^2\sigma_1^2). \tag{36}$$

50.2 Variabile aleatoria normale standard

(A causa delle (36)) la standardizzazione di una qualunque variabile aleatoria normale è una variabile aleatoria normale $N(0, 1)$.
(Avendo media 0 e varianza 1, in base alla (35)) ha densità

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}} \quad =: \phi(x)$$

(*densità normale standard*, denotata con $\phi(x)$).
Ha anche moda 0 e mediana 0.



La sua funzione di ripartizione si indica con $\Phi(x)$ e si chiama *funzione di ripartizione normale standard*, in Inglese *(standard) normal cumulative distribution function*:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt \qquad \text{e derivando } \Phi' = \phi\,.$$

L'integrale che definisce questa *funzione speciale (dell'Analisi Matematica)* non può essere risolto in termini di funzioni elementari.
Valori numerici (approssimati) di $\Phi(x)$ si ottengono in vari modi.
La funzione inversa di $\Phi(x)$ dà i quantili normali ϕ_α.
Esercizio. Fare lo studio di funzione di $\phi(x)$. Dove sono i flessi?

50.3 Variabile aleatoria log-normale

Se X è una variabile aleatoria $N(\mu, \sigma^2)$ allora $Y := \mathrm{e}^X$ si dice *log-normale* di parametri μ e σ^2, che però non sono rispettivamente media e varianza della nuova variabile aleatoria.
Inversamente, se Y è log-normale di parametri di parametri μ e σ^2, allora $X := \ln Y$ è $N(\mu, \sigma^2)$.
Consideriamo ora solo il caso $\mu = 0$ e $\sigma^2 = 1$:

$$\forall x > 0 \qquad F_{\mathrm{e}^X}(x) = P(\mathrm{e}^X \leq x) = P(X \leq \ln x) = F_X(\ln x)$$

e per i non positivi, considerando i 2 casi disgiunti $x = 0$ e $x < 0$

$$\forall x \leq 0 \qquad F_{\mathrm{e}^X}(x) = P(\mathrm{e}^X \leq x) = P(\mathrm{e}^X = x) + P(\mathrm{e}^X < x) = 0+0$$

e in definitiva

$$F_Y(x) = F_{\mathrm{e}^X}(x) = \begin{cases} \Phi(\ln x) & se \;\; x > 0 \\ 0 & se \;\; x \leq 0 \end{cases} .$$

Derivando troviamo la densità log-normale di parametri $\mu = 0$ e $\sigma^2 = 1$, ricordando che $\Phi'(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, e naturalmente (derivata della funzione composta) deriviamo anche ln:

$$\forall x > 0 \qquad f_Y(x) = f_{\mathrm{e}^X}(x) = \phi(\ln x) \cdot \frac{1}{x}$$

$$f_Y(x) = f_{\mathrm{e}^X}(x) = \begin{cases} \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}} & se \;\; x > 0 \\ 0 & se \;\; x \leq 0 \end{cases} . \tag{37}$$

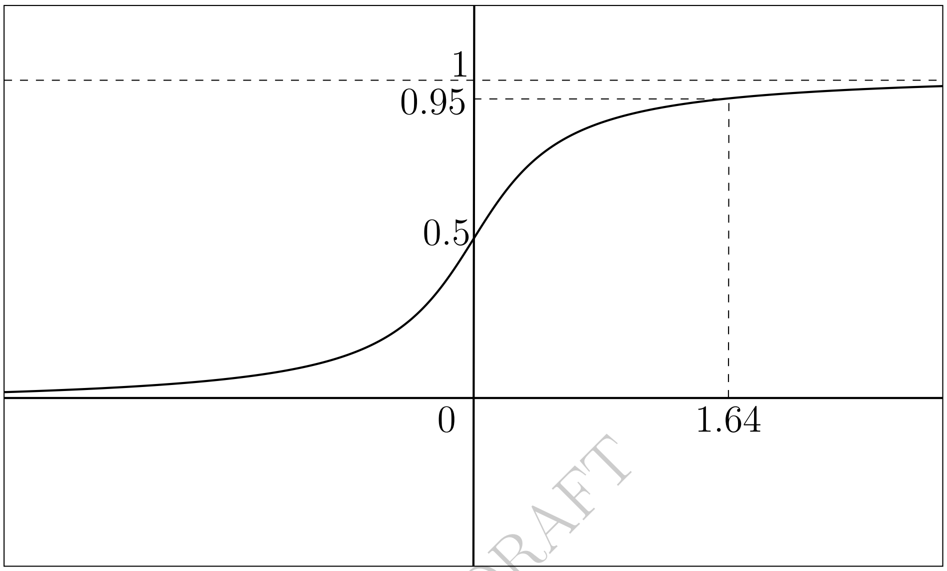
Con μ e σ^2 generici: $\dfrac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x - \mu)^2}{2}}$ se $x > 0$ e 0 altrimenti.

I 3 valori coincidenti per la normale standard, media moda e mediana, $E(X) = Mod(X) = \phi_{0.5}$, hanno 3 destini diversi: per la log-normale $Y := \mathrm{e}^X$ la mediana è e^μ, con $\mu = E(X)$:

$$P(Y \leq \mathrm{e}^\mu) = P(\mathrm{e}^X \leq \mathrm{e}^\mu) = P(X \leq \mu) = \frac{1}{2} \text{ per simmetria}$$

ma media e moda hanno diverse espressioni.
Esercizio. Si faccia lo studio di funzione di (37). Quanto vale la funzione di ripartizione in -1, 0, $\mathrm{e}^{1.64}$, $\mathrm{e}^{2.58}$, 100?

51 Approssimazione di $\Phi(x)$ e ϕ_α



Valori numerici (approssimati) della funzione di ripartizione normale standard $\Phi(x)$ e dei quantili normali ϕ_α si ottengono:
(1) con quelle (rare) calcolatrici scientifiche che la implementano;
(2) online in www.wolframalpha.com digitando
`CDF[NormalDistribution[0,1],`$_{\textit{valore di } x}$`]` per avere $\Phi(x)$,
`InverseCDF[NormalDistribution[0,1],`$_{\textit{valore di } \alpha}$`]` per avere ϕ_α;
(3) con molti software di manipolazione matematica, fra cui Maxima e Mathematica$^{(R)}$;
(4) a memoria per alcuni pochi valori speciali:

x	$\Phi(x)$
0	0.5
≈ 1.64	≈ 0.95
≈ 1.96	≈ 0.975
≈ 2.58	≈ 0.995
ϕ_α	α

(e si faccia attenzione che questi non sono i valori di $P(|X| < x)$);

(5) con apposite formule di approssimazione, per esempio[2]

$$\forall x \geq 0 \qquad \Phi(x) \approx 2^\wedge(-22^\wedge(1-41^\wedge(x/10)))$$

che ha errore assoluto $|\varepsilon(x)| < 0.00013$ per ogni $x \geq 0$, e la sua inversa (che si ottiene subito ricavando x, che è ϕ_α)

$$\forall \alpha \in \ [0, 0.5[\qquad \phi_\alpha \approx \frac{10}{\log 41} \log\Big(1 - \frac{\log((-\log\alpha)/\log 2)}{\log 22}\Big)$$

(che ha errori assoluto e relativo rispettivamente $|\varepsilon(\alpha)| < 5\,10^{-3}$ $\forall\alpha \in [0.5, 9925]$, $|\varepsilon_r(\alpha)| < 1\%$ $\forall\alpha \in [0.5, 0.99908]$);
(6) con le apposite tavole numeriche, che si trovano su internet cercando *normal table* o (e sono essenzialmente le stesse) *normal quantile table*.
Nota. Sia le tavole numeriche che, di solito, le formule di approssimazione danno (approssimano)
  $\Phi(x)$ solo per $x \geq 0$
  ϕ_α solo per $\alpha \geq 0.5$
e per i valori di $x < 0$ e $\alpha < 0.5$ si usano le formula di simmetria

$$\Phi(-x) = 1 - \Phi(x)$$

$$\phi_{1-\alpha} = -\phi_\alpha$$

(che seguono dalla parità della densità normale standard).

51.1 Le tavole numeriche di $\Phi(x)$ e ϕ_α

Del calcolo approssimato di $\Phi(x)$ si è detto, ma qui vogliamo approfondire la questione dell'uso delle tavole. Le stesse tavole si usano per trovare $\Phi(x)$ dato x e per trovare ϕ_α dato α. (...)???

[2]A. Soranzo, E. Epure (2014)
http://m-hikari.com/ams/ams-2014/ams-85-88-2014/epureAMS85-88-2014.pdf

52 Legge dei Grandi Numeri

52.1 Inquadramento euristico della situazione

Supponiamo di lanciare una moneta equilibrata un numero grandissimo di volte, e continuiamo a farlo, conteggiando il numero di teste e il numero di croci. Alcuni ingenui credono che i 2 numeri tendano a diventare sempre più simili, ma questo è falso. E' impensabile che lanciando un milione di volte la moneta siano venute esattamente 500mila teste e 500mila croci, o 500 001 o anche 500 002 o simili. Anzi si potrebbe dimostrare che la differenza fra teste e croci tenderà in generale ad essere sempre più grande, non più piccola! Invece quello che tende a succedere è che le proporzioni di teste e di croci tenderanno ad uguagliarsi, tendendo entrambe ad $\frac{1}{2}$. Quello che possiamo effettivamente aspettarci dopo un milione di lanci è una situazione di questo tipo:

teste: $500\,000 \pm$ qualche centinaio: $\#teste = 500\,000 + r := n_0$
croci: $500\,000 \mp$ qualche centinaio: $\#croci = 500\,000 - r := n_1$
r: qualche centinaio in positivo o in negativo, p.es. 424 o -723
frazione di teste: $\frac{500\,000+r}{1\,000\,000} = \frac{1}{2} + \frac{r}{1\,000\,000} \approx 0.5$
frazione di croci: $\frac{500\,000-r}{1\,000\,000} = \frac{1}{2} - \frac{r}{1\,000\,000} \approx 0.5.$

Le *proporzioni empiriche* tendono ad uguagliarsi, non le quantità! Questo diventerà ancora più evidente al crescere del numero di lanci, cioè l'approssimazione a 0.5 varrà con sempre più decimali, salvo casi sfortunatissimi, comunque sempre possibili.
Similmente avviene per qualunque p_1 fra 0 e 1 che sia la probabilità della testa della moneta (che se $p_1 \neq 0.5$ è non regolare): detto n il numero di lanci, e associato l'1 alla testa e 0 alla croce,

proporzione empirica di teste $\bar{p}_{n,1} = \frac{\#teste}{n} \to p_1$
proporzione empirica di croci $\bar{p}_{n,0} = \frac{\#croci}{n} \to p_0 := 1 - p_1.$

Similmente per un dado avremo 6 limiti p_1,...,p_6, cioè le proporzioni empiriche dei risultati tenderanno alle probabilità *vere* dei vari risultati, per esempio, per un dado regolare, sempre $\frac{1}{6}$:
$\forall\, k \in \{1, ..., m \leftarrow \text{è 6 per un dado}\} \quad \bar{p}_{k,n} \to p_k = P(X = k).$
Esercizio. Ipotizzare e graficare le $\bar{p}_{k,n}$ per $n = 100$, $k = 1, ..., 6$.

52.2 Limite in probabilità e Legge dei Grandi Numeri

In quanto detto, resta non definito cosa si intende per il "tendere" ai numeri p_k, e si è ben detto che ci possono essere casi sfortunatissimi. Si tratta di un tendere probabilistico, non deterministico com'è quello dei limiti delle funzioni reali di variabile reale. Esso è precisato e inquadrato dal concetto di *convergenza in probabilità* di una successione di variabili aleatorie X_n, che definiamo senza insistervi particolarmente. Si immagini la X_n di cui parliamo come la proporzione empirica $\bar{p}_{1,n}$ di teste dopo n lanci, che, sì, è una variabile aleatoria, "prima" di fare i lanci. Il limite della convergenza in probabilità di una successione di variabili aleatorie è esso stesso in generale una variabile aleatoria; solo che nell'esempio prima considerato è la variabile aleatoria discreta

$$X := \begin{pmatrix} 0.5 \\ 1 \end{pmatrix} = p$$

che vale 0.5 con probabilità 1. (Variabile aleatoria *costante*). Ma in generale il limite X di una convergenza in probabilità è proprio una variabile aleatoria con una funzione di ripartizione non banale, ed è una variabile aleatoria discreta o continua.

Definizione. Diremo che X_n converge in probabilità a X

$$X_n \to^P X \text{ se } \lim_{n\to\infty} P(|X_n - X| \geq \eta) = 0 \quad \forall\, \eta > 0$$

(o indifferentemente con $> \eta$).

Legge dei Grandi Numeri. Sia $(X_n)_n$ una successione di variabili aleatorie indipendenti di ugual legge con speranza matematica μ e varianza σ^2. Allora per la *media empirica*

$$\bar{X}_n := \frac{1}{n}(X_1 + ... + X_n)$$

$$\text{è} \quad \forall\, \eta > 0 \qquad \lim_{n\to\infty} P(|\bar{X}_n - \mu| \geq \eta) = 0$$

$$\text{ovvero equivalentemente } \bar{X}_n \to^P \mu.$$

Con $X_h := 1$ per testa e 0 altrimenti per $h = 1, ..., n$, si riottiene il primo caso considerato, con $\bar{X}_n$ la proporzione empirica $\bar{p}_{1,n}$.

•••

Possiamo fare i calcoli *esattamente*. La probabilità di k teste è

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

e ipotizzando una moneta regolare ($p = 1/2$, $p^k(1-p)^{n-k} = 2^{-n}$)

$$p_k = \binom{n}{k} 2^{-n}$$

e in particolare la probabilità i fare tante teste quante croci, che è 0 se n è dispari, per n pari è

$$\binom{n}{n/2} 2^{-n}$$

Con n piccolo è possibile fare facilmente il calcolo esatto, per esempio per $n := 6$ la probabilità è $\frac{5}{16}$ e per $n := 20$

$$\binom{20}{10} 2^{-20} \approx 0.176197 \quad \text{circa 1 su 6.}$$

La (38) con $n := 20$ ci dà $\quad * * * non\, fatto\, 2017 *$

$$10 - 2\sqrt{5} \leq X \leq 10 + 2\sqrt{5}$$

cioè

$$5.527... \leq numero\, di\, teste \leq 14.472...$$

ovvero

$$numero\, di\, teste = 6 \vee 7 \vee 8 \vee 9 \vee 10 \vee 11 \vee 12 \vee 13 \vee 14$$

e questo evento ha probabilità

$$p_6 + ... + p_{14} = 2^{-20}\left(\binom{20}{6} + ... + \binom{20}{14}\right) =$$

$$= \frac{125647}{131072} \approx 0.959 = 95.9\%.$$

Prima si era detto *almeno* 75%, ora si trova *esattamente* 0.95.... (Ma questo calcolo per $n := 1\,000\,000$ è improbo).

53 Scarti dalla media

53.1 Con la Disuguaglianza di Chebyshev

Consideriamo n lanci di moneta con probabilità p di fare testa. Useremo ora la Disuguaglianza di Cebyshev:

$$P(|X - E(X)| \leq c) \geq 1 - \frac{Var(X)}{c^2}.$$

Se X è il contatore di teste (successi) in n lanci, allora $X \sim B(n,k)$. Essendo per la $B(n,k)$ la varianza $np(1-p)$ e la speranza matematica np, con la moneta regolare $n/4$ e $n/2$ rispettivamente,

$$P(|X - n/2| \leq c) \geq 1 - \frac{n/4}{c^2}$$

e fissando $c := 2\sigma = 2\sqrt{VarB(n,k)} = 2\sqrt{n/4} = \sqrt{n}$

$$P(|X - n/2| \leq \sqrt{n}) \geq 1 - \frac{\sigma^2}{(2\sigma)^2}$$

cioè

$$P\Big((\Big|X - \frac{n}{2}\Big| \leq \sqrt{n}\Big) \geq \frac{3}{4} = 0.75 = 75\%$$

e ricordando che $|f(x)| \leq g(x)$ equivale a $-g(x) \leq f(x) \leq g(x)$, con probabilità almeno del 75% il numero di teste verifica

$$\frac{n}{2} - \sqrt{n} \leq X \leq \frac{n}{2} + \sqrt{n}. \tag{38}$$

Con un milione di lanci, al 75% il numero di teste sta fra 499 000 e 501 000.

Si noti che effettivamente $\frac{501000}{1000000} \approx 0.5$, e similmente con 499 000. Detto in altri termini

$$\frac{\frac{n}{2} \pm \sqrt{n}}{n} = \frac{1}{2} \pm \frac{1}{\sqrt{n}} \approx \frac{1}{2} \quad per \quad n \gg .$$

Esercizio. Si faccia l'analogo calcolo con $c := 3\sigma$, e poi di nuovo la particolarizzazione con $n := 1\,000\,000$. E poi con 4σ, 5σ e 6σ.

Analogamente si trova, con $c := m\,\sigma$, che per una **variabile aleatoria discreta o continua qualunque** purchè dotata di speranza matematica μ e varianza σ^2, per la Disuguaglianza di Chebyshev è, osservato che $1 - \frac{\sigma^2}{(m\sigma)^2} = 1 - \frac{1}{m^2}$,

$$P(|X - \mu| \leq m\,\sigma) \geq 1 - \frac{1}{m^2} \qquad m > 0$$

e con $m := 2, 3, 4, 5$ (il caso $m := 1$ non è significativo)

$$P(|X - \mu| \leq 2\sigma) \geq \frac{3}{4} = 0.75 = 75\%$$

$$P(|X - \mu| \leq 3\sigma) \geq \frac{8}{9} = 0.8888... \approx 88.9\%$$

$$P(|X - \mu| \leq 4\sigma) \geq \frac{15}{16} = 0.9375 = 93.75\%$$

$$P(|X - \mu| \leq 5\sigma) \geq \frac{24}{25} = 0.96 = 96\%.$$

53.2 Scarti dalla media per v.a. normale

Se si sa che la variabile aleatoria X è normale si ottengono disuguaglianze, che ora vediamo, molto più stringenti di quelle sopra, e che ovviamente non le contraddicono.
Una variabile aleatoria $X \sim N(\mu, \sigma^2)$ è in questa relazione con la sua standardizzazione $Y = \frac{X-\mu}{\sigma}$:

$$X = \sigma Y + \mu$$

e allora si ha, per ogni $\delta > 0$,

$$P(|X - \mu| \leq \delta) = P(|\sigma Y| \leq \delta) = P(\sigma |Y| \leq \delta) = P\Big(|Y| \leq \frac{\delta}{\sigma}\Big)$$

e con facili calcoli$^{\rightarrow}$ si conclude

$$P(|X - \mu| \leq m\,\sigma) = 2\,\Phi(m) - 1$$

e coi classici valori (che si trovano per esempio sulle tavole)

$$\Phi(1) \approx 0.8413$$
$$\Phi(2) \approx 0.9773$$
$$\Phi(3) \approx 0.9987$$

$$\boxed{\begin{array}{c} X \sim N(\mu, \sigma^2) \\ \\ P(|X-\mu| \leq \sigma) \approx 0.68.26 \approx 68.3\% \\ \\ P(|X-\mu| \leq 2\,\sigma) \approx 0.9546 \approx 95.5\% \\ \\ P(|X-\mu| \leq 3\,\sigma) \approx 0.9974 \approx 99.7\% \\ \\ P(|X-\mu| \leq 1.96\,\sigma) \approx 0.95 = 95\% \end{array}}$$

dove la quarta è una lieve modificazione della seconda per avere con più precisione 95%.

54 Approssimazione Normale

54.1 TLC e Approssimazione Normale

Oltre alla convergenza in probabilità esistono altri tipi di convergenza per successioni di variabili aleatorie e in questa trattazione se ne considererà una. Con ovvio significato dei simboli vale la:
Definizione. Diremo che X_n **converge in legge** a X

$$X_n \to^L X \text{ se } \lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

per ogni punto x in cui $F_X(x)$ è continua. (Ha valore teorico).
Osservazione. La forma a campana e in particolare la variabile aleatoria normale standard è una sorta di "attrattore" per le variabili aleatorie, perchè anche se ne sono alquanto diverse, la loro somma standardizzata, a certe condizioni che ora vedremo, tende proprio alla $N(0,1)$, in legge.
Teorema Limite Centrale. Sia $(X_n)_n$ una successione di variabili aleatorie indipendenti (discrete o continue) di ugual legge con speranza matematica μ e varianza σ^2. Allora la standardizzazione della somma delle prime n tende in legge ad una normale standard:

$$\exists\, X \sim N(0,1) \qquad S_n^* := \frac{X_1 + ... + X_n - n\,\mu}{\sigma\,\sqrt{n}} \to^L X_{\ \textit{risultato teorico}}$$

Per le variabili aleatorie sopra considerate vale (risultato più applicativo) l'**Approssimazione Normale**

$$\boxed{P\Big(X_1 + ... + X_n \le x\Big) \approx \Phi\Big(\tfrac{x - n\mu}{\sigma\sqrt{n}}\Big)} \tag{39}$$

e tradizionalmente questa approssimazione si ritiene sufficientemente buona per $n \ge 30$ (secondo altri Autori $n \ge 50$)

* se le $X_k \sim B(m,p)$ con $n\,p \ge 5$ e $n\,(1-p) \ge 5$

* per tutte le altre distribuzioni degli esercizi scolastici e anche di questa trattazione, e spessissimo anche della pratica.

●●●

Teorema. La convergenza in probabilità implica quella in legge.
Esempio. Calcoleremo la probabilità di ottenere più di 28 teste in 50 lanci di moneta equilibrata.
I risultati X_k hanno legge $B(1, \frac{1}{2})$, con speranza matematica $\mu = \frac{1}{2}$ e varianza $\sigma^2 = 1 \cdot \frac{1}{2}\big(1 - \frac{1}{2}\big) = \frac{1}{4}$, allora $\sigma = \frac{1}{2}$, e allora

$$P\Big(X_1 + ... + X_{50} > 28\Big) = 1 - P\Big(X_1 + ... + X_{50} \leq 28\Big) \approx$$

$$\approx 1 - \Phi\Big(\frac{28 - 50 \cdot \frac{1}{2}}{\frac{1}{2}\sqrt{50}}\Big) \approx 1 - \Phi(0.849) \approx 1 - 0.802 \approx 0.2 = 20\%.$$

B1 – ESERCIZI SUL CALCOLO DELLE PROBABILITà

Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 4 volte il testa con 5 lanci di una moneta regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 5 volte il testa con 6 lanci di una moneta regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 5 volte il croce con 7 lanci di una moneta regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 4 volte il numero 2 con 5 lanci di un dado regolare a 4 facce?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 1 volta il numero 3 con 7 lanci di un dado regolare a 4 facce?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 4 volte il numero 2 con 5 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 3 volte il numero 1 con 8 lanci di un dado regolaree?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 7 volte il numero 7 con 8 lanci di un dado regolare a 8 facce?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 3 volte un numero primo con 8 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 6 volte un numero primo con 7 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere almeno 6 volte un numero primo con 8 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere un numero primo di volte un numero primo con 6 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere un numero primo di volte un numero primo con 10 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere un numero quadrato di volte un numero primo con 8 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere un numero primo di volte un numero quadrato con 6 lanci di un dado regolare?
Esercizio$_\mu$ Che probabilità c'è di ottenere un numero triangolare di volte testa con 9 lanci di una moneta regolare?

55 Stimatori e stimatori non distorti

55.1 Parametri e stimatori; stimatori non distorti

Supponiamo che fra poco avremo le altezze di $n := 100$ persone, prese a caso dai 60milioni di italiani, e allora l'altezza di un italiano a caso è una variabile aleatoria X, e gli n numeri che stiamo per avere sono essi stessi variabili aleatorie X_1,...,X_n, e dopo che li avremo saranno numeri $x_1, ..., x_n$, detti *determinazioni* della variabile aleatoria. Chiameremo anche *campione aleatorio* le variabili aleatorie X_1,...,X_n e *campione* i numeri $x_1, ..., x_n$, ma non ci formalizzeremo su questa distinzione. La Statistica Descrittiva ci ha dato definizioni di media e varianza di quei numeri $x_1, ..., x_n$. Nel Calcolo delle Probabilità abbiamo definito la media ovvero speranza matematica e la varianza della variabile aleatoria X, ma qual è la relazione fra le 2 medie e le 2 varianze? Ottenere da n numeri una *stima* di un *parametro* incognito di una v.a. ovvero della sua legge, è un *problema di stima.* Ogni funzione

$$h(X_1, ..., X_n)$$

di un campione aleatorio si dice *stimatore*, in generale stimatore di un parametro incognito u della densità della variabile aleatoria, variabile aleatoria che sappiamo esistere ma non conosceremo mai in forma esatta (legge ovvero distribuzione, cioè densità o funzione di ripartizione), ma di cui avremo n determinazioni. Scriveremo

$$\hat{v} := h(X_1, ..., X_n)$$

e diremo che $\hat{v}$ è una stima di v. Per esempio per una v.a. $\mathbb{U}[0, a]$

$$\hat{a} := 2\,\frac{X_1 + ... + X_n}{n}$$

è lo *stimatore dei momenti* del parametro a. Esistono molti criteri di bontà di uno stimatore, e noi ne consideremo uno: diremo che lo stimatore $\hat{u}$ del parametro incognito u è *non distorto* se la speranza matematica di $\hat{u}$ è u:

$$\hat{u} \text{ stimatore non distorto di } u \text{ se } E(\hat{u}) = u.$$

55.2 Stimatori non distorti di media e varianza

La differenza basale fra la *statistica descrittiva* e la *statistica inferenziale* è che la prima opera (in generale) su numeri e la seconda su variabili aleatorie, con il che la prima ricade in quella che abbiamo chiamato *matematica della certezza* e la seconda nella *matematica dell'incertezza*.
Se abbiamo n numeri

$$x_1, x_2, ..., x_n$$

nella statistica descrittiva queste sono la media e la varianza degli n numeri:

$$\bar{x} := \frac{1}{n}\sum_{k=1}^{n} x_k \qquad Var(x_1, x_2, ..., x_n) := \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^2 \, .$$

Se invece supponiamo che $x_1, x_2, ..., x_n$ siano *determinazioni* di una variabile aleatoria X con una certa media μ e una certa varianza σ^2 *incognite*, dalle precedenti formule possiamo immediatamente definire questi (ragionevoli) stimatori – che sono variabili aleatorie – di μ e σ^2:

$$\hat{\mu} := \bar{X} := \bar{X}_n := \frac{1}{n}\sum_{k=1}^{n} X_k \qquad W := \frac{1}{n}\sum_{k=1}^{n}(X_k - \bar{X}_n)^2 \, .$$

Si dimostra che il primo è stimatore *non distorto* della media μ

$$E(\bar{X}_n) = \mu$$

ma il secondo non è stimatore non distorto di σ^2, cioè la sua speranza matematica è diversa dal parametro che si vuol stimare.
Si dimostra invece che (e si noti la differenza nel denominatore)

$$S^2 := \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X})^2 \;\left[:= \hat{\sigma}^2 \;{\scriptstyle \leftarrow \mathit{scrittura\ rara,\ scriveremo}\ S^2}\right]$$

è lo *stimatore non distorto della varianza*:

$$E(S^2) = \sigma^2 \, .$$

56 Stimatori dei momenti

56.1 Introduzione: stimatore col momento primo

Per una v.a. discreta o continua X, supponiamo di sapere che:
1) ha distribuzione di un certo tipo noto dipendente da un parametro incognito, sia esso a, per esempio $\mathbb{U}[-b, b]$ o $N(\mu, 1)$ o $N(0, \sigma^2)$ o $\Gamma(\alpha, 1)$ (dove a è rispettivamente b, μ, σ^2 e α);
2) avremo n determinazioni indipendenti x_1,...,x_n di X, che per ora, "prima" di averle, sono n variabili aleatorie X_1,...X_n.
Non sapremo mai quanto vale a ma vogliamo stimarlo con uno stimatore $\hat{a}$ coi dati x_1,...,x_n, ovvero (prima di averli) X_1,...,X_n.
Lo stimatore $\hat{a}$ *dei momenti, col momento primo*, attribuisce al parametro a quel valore che farebbe avere alla densità $f_a(x)$ come media, ovvero speranza matematica ovvero momento primo, il valore che è proprio la media di X_1,...X_n.
Esempio 1. Per una densità $N(\mu, \sigma^2)$ con σ^2 noto (ora $a := \mu$)

$$\hat{\mu} := \frac{1}{n}(X_1 + ... + X_n).$$

Esempio 2. Per una densità esponenziale di parametro λ (ora $a := \lambda$) è

$$E(X) = \frac{1}{\lambda} \stackrel{EQ}{=} \bar{X}_n$$

$$\Rightarrow \ \hat{\lambda} := \frac{1}{\bar{X}_n} = \frac{n}{X_1 + ... + X_n} \, .$$

Esempio 3. Per una densità $\Gamma(\alpha, \lambda)$ con α noto (ora $a := \lambda$) è

$$E(X) = \frac{\alpha}{\lambda} \stackrel{EQ}{=} \bar{X}_n$$

$$\Rightarrow \ \hat{\lambda} := \frac{\alpha}{\bar{X}_n} = \frac{n\,\alpha}{X_1 + ... + X_n} \, .$$

Esempio 4. Per una densità uniforme $\mathbb{U}[-u, u]$ è

$$E(X) = \frac{u - u}{2} = 0$$

che non dipende dal parametro incognito $a = u$: il metodo dei momenti col solo momento primo non si può applicare (vedi dopo).

56.2 Metodo di calcolo dello stimatore dei momenti

1) Si calcola la speranza matematica di $f_a(x)$ come funzione di a;
2) se dipende da a la si uguaglia a $\bar{X}_n$, media di X_1,...,X_n;
3) si risolve (se possibile) l'equazione in a;
Quella soluzione trovata, in termini di X_1,...,X_n è lo stimatore dei momenti, inteso come funzione (ed è una variabile aleatoria), per esempio

$$\hat{a} := \frac{n}{X_1 + ... + X_n}$$

e in termini di x_1,...,x_n è lo stimatore dei momenti inteso come valore dello stimatore, nel caso specifico di un campione x_1,...,x_n considerato, per esempio

$$\hat{a} := \frac{n}{x_1 + ... + x_n}$$

e quando a x_1,...,x_n si sostituiscano i loro valori, è un numero, per esempio

$$\hat{a} := \frac{3}{8 + 7 + 5} = \frac{3}{20} = 0.15.$$

4) Se $E(X)$ non dipende dal parametro incognito a, si usi allora il momento secondo uguagliandolo alla media quadratica.

Esempio 4 bis. Per una densità uniforme $\mathbb{U}[-u, u]$ è

$$E(X) = \frac{u - u}{2} = 0$$

che non dipende dal parametro incognito $a = u$. Con il momento secondo $\int_{-\infty}^{+\infty} x^2 \, f_a(x) \, dx$, ora che la densità è $\frac{1}{2\,u}$ fra $-u$ e u,

$$E(X^2) = \int_{-u}^{u} \frac{x^2}{2\,u} \, dx = \left[\frac{x^3}{6\,u} \right]_{-u}^{u} = \frac{u^3}{6\,u} - \frac{-u^3}{6\,u} = \frac{u^3}{3\,u} =$$

$$= \frac{u^2}{3} =^{EQ} \frac{1}{n}(X_1^2 + ... + X_n^2) \quad \leftarrow \textit{media quadratica}$$

$$\rightarrow \quad \hat{u} := \sqrt{\frac{3\,(X_1^2 + ... + X_n^2)}{n}}\,.$$

A un livello superiore si considerano parametri 2-dimensionali, come $a := (\mu, \sigma^2)$, ottenendosi un sistema di equazioni.

57 Stimatori di massima verosimiglianza

Per una v.a. discreta o continua X, supponiamo di sapere che:
1) ha distribuzione di un certo tipo noto dipendente da un parametro incognito, sia esso a, per esempio $\mathbb{U}[-b, b]$ o $N(\mu, 1)$ o $N(0, \sigma^2)$ o $\Gamma(\alpha, 1)$ (dove a è rispettivamente b, μ, σ^2 e α);
2) avremo n determinazioni indipendenti x_1,...,x_n di X, che per ora, "prima" di averle, sono n variabili aleatorie X_1,...,X_n.
Non sapremo mai quanto vale a ma vogliamo stimarlo con uno stimatore $\hat{a}$ coi dati x_1,...,x_n, ovvero (prima di averli) X_1,...,X_n.
Lo stimatore $\hat{a}$ *di massima verosimiglianza* "cerca di essere" quello che renderebbe massimamente probabile l'uscita di x_1,...,x_n, se a valesse $\hat{a}$. Se per esempio avremo una sola determinazione di $N(\mu, \sigma^2)$ con σ^2 noto, ovviamente stimeremo $\hat{a} := \hat{\mu} := X_1$. (Si disegni la campana gaussiana col massimo in μ).
Naturalmente, trattandosi di una densità continua, se $\mu = x_1$, la probabilità di quella particolare uscita x_1 è 0 comunque, come per qualunque altro valore e con qualunque valore del parametro μ, ma è chiaro che $\mu := x_1$ "cerca di essere" quello che renderebbe massimamente probabile l'uscita x_1 perchè almeno con quel valore di μ la densità ha un massimo in x_1.
Meno evidente è $\hat{\mu}$ se $n \geq 2$, ma si dimostra che è la media campionaria $\bar{X} := \frac{1}{n}(X_1+...+X_n)$ perchè la densità congiunta, prodotto delle densità $f_\mu(x)$ per l'indipendenza, calcolata in x_1,...,x_n,

$$f_\mu(x_1) \cdot ... \cdot f_\mu(x_n) = \frac{1}{\sqrt{2\pi}\,\sigma}\, \mathrm{e}^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot ... \cdot \frac{1}{\sqrt{2\pi}\,\sigma}\, \mathrm{e}^{-\frac{(x_n-\mu)^2}{2\sigma^2}}$$

ha un punto di massimo in $\bar{X}$ perchè lo ha il suo ln

$$n \ln \frac{1}{\sqrt{2\pi}\,\sigma} - \Big(\frac{(x_1-\mu)^2}{2\sigma^2} + ... + \frac{(x_n-\mu)^2}{2\sigma^2}\Big)$$

come si vede derivando rispetto al parametro, in questo caso μ,

$$0 - \Big(\frac{x_1-\mu}{\sigma^2} + ... + \frac{x_n-\mu}{\sigma^2}\Big)$$

e uguagliando a 0: $n\,\mu = x_1+...+x_n \Rightarrow \hat{\mu} = \frac{1}{n}(X_1+...+X_n) = \bar{X}_n$.

57.1 Massima verosimiglianza: il metodo generale

Generalizzeremo come segue il procedimento visto nell'esempio.
In questa trattazione elementare, prenderemo come stimatore di massima verosimiglianza $\hat{a}$ (di un parametro incognito a di una densità $f_a(x)$ per il resto nota) lo zero, se esiste unico (ad un livello superiore si considera il caso di non unicità) della derivata rispetto ad a del logaritmo naturale

$$\ln f_a(x_1) + ... + \ln f_a(x_n)$$

della funzione di verosimiglianza

$$f_a(x_1) \cdot ... \cdot f_a(x_n).$$

Procederemo allora con questi passaggi:
1) calcolo di $f_a(x_1) \cdot ... \cdot f_a(x_n)$ lasciando indicati x_1,...,x_n (cioè considerandoli variabili reali e non costanti numeriche);
2) calcolo del logaritmo naturale della predetta funzione;
3) derivazione della predetta funzione rispetto al parametro a;
4) uguagliamento a 0 della predetta funzione;
5) risoluzione dell'equazione.
Quello zero trovato, in termini di X_1,...,X_n è lo stimatore di massima verosimiglianza inteso come funzione (ed è una variabile aleatoria), per esempio

$$\hat{a} := \frac{1}{n}(X_1 + ... + X_n)$$

e in termini di x_1,...,x_n è lo stimatore di massima verosimiglianza inteso come valore dello stimatore, nel caso specifico di un campione x_1,...,x_n considerato, per esempio

$$\hat{a} := \frac{1}{n}(x_1 + ... + x_n)$$

e quando a x_1,...,x_n si sostituiscano i loro valori, è un numero, per esempio

$$\hat{a} := \frac{1}{3}(8 + 7 + 5) = \frac{20}{3} \approx 6.67.$$

Ad un livello superiore, si considerano anche parametri 2-dimensionali, per esempio $a := (\mu, \sigma^2)$. (Là sono incogniti sia μ che σ^2).

57.2 Esempio: parametro λ dell'esponenziale

Ora il parametro incognito è $a := \lambda$ e la densità

$$f_\lambda(x) = \begin{cases} \lambda\, \mathrm{e}^{-\lambda\, x} & \textit{se } x \geq 0 \\ 0 & \textit{altrimenti} \end{cases} .$$

La parte della definizione per $x < 0$ non è rilevante e similmente avviene per tutti i casi di interesse pratico in cui la densità è nulla per $x < 0$. (E per tutti i casi di questa trattazione elementare).

1) calcolo di $f_a(x_1) \cdot ... \cdot f_a(x_n)$ lasciando indicati x_1,...,x_n:

$$\forall\, x_1, ..., x_n \geq 0 \qquad f_\lambda(x_1) \cdot ... \cdot f_\lambda(x_n) = \lambda\, \mathrm{e}^{-\lambda\, x_1} \cdot ... \cdot \lambda\, \mathrm{e}^{-\lambda\, x_n} =$$

$$= \lambda^n\, \mathrm{e}^{-\lambda\,(x_1 + ... + x_n)}$$

2) Calcolo del logaritmo naturale della predetta funzione:

$$n \ln \lambda - \lambda\,(x_1 + ... + x_n)$$

3) derivazione della predetta funzione rispetto al parametro λ:

$$\frac{n}{\lambda} - (x_1 + ... + x_n)$$

4) uguagliamento a 0 della predetta funzione:

$$\frac{n}{\lambda} - (x_1 + ... + x_n) =^{EQ} 0$$

5) risoluzione dell'equazione:

$$\lambda = \frac{n}{x_1 + ... + x_n} .$$

Infine, lo stimatore del parametro λ della legge esponenziale risulta

$$\hat{\lambda} = \frac{n}{X_1 + ... + X_n} = \frac{1}{\bar{X}_n}$$

dove si è riconosciuto il reciproco della media campionaria $\bar{X}_n$.

•••

57.3 Esempio: stimatore di σ^2 di $N(\mu, \sigma^2)$ con μ noto

Ora il parametro incognito è $a := \sigma^2$ e la densità

$$f_{\sigma^2}(x) = \frac{1}{\sqrt{2\pi\,\sigma^2}}\,\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Si è scritto σ^2 sotto radice invece di σ fuori dalla radice, e così c'è solo σ^2, in tutto 2 volte, e mai σ da solo.

1) Calcolo di $f_a(x_1) \cdot ... \cdot f_a(x_n)$:

$$f_{\sigma^2}(x_1) \cdot ... \cdot f_{\sigma^2}(x_n) = \frac{1}{\sqrt{2\pi\,\sigma^2}}\,\mathrm{e}^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot ... \cdot \frac{1}{\sqrt{2\pi\,\sigma^2}}\,\mathrm{e}^{-\frac{(x_n-\mu)^2}{2\sigma^2}}$$

2) calcolo del logaritmo naturale della predetta funzione

$$\frac{n}{2}\ln\frac{1}{2\pi} - \frac{n}{2}\ln\sigma^2 - \Big(\frac{(x_1-\mu)^2}{2\sigma^2} + ... + \frac{(x_n-\mu)^2}{2\sigma^2}\Big)$$

3) derivazione della predetta funzione rispetto al parametro σ^2:

$$0 - \frac{n}{2}\cdot\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}((x_1-\mu)^2 + ...(x_n-\mu)^2)$$

4) uguagliamento a 0 della predetta funzione:

$$\frac{1}{2\sigma^2}\Big(\frac{1}{\sigma^2}((x_1-\mu)^2 + ...(x_n-\mu)^2) - n\Big) =^{EQ} 0$$

5) risoluzione dell'equazione:

$$\sigma^2 = \frac{1}{n}((x_1-\mu)^2 + ...(x_n-\mu)^2).$$

Infine, lo stimatore di σ^2 di $N(\mu, \sigma^2)$ con μ noto risulta

$$\hat{\sigma}^2 = \frac{1}{n}((X_1-\mu)^2 + ...(X_n-\mu)^2).$$

Ad un livello superiore, si dimostra che lo stimatore del parametro 2-dimensionale $a := (\mu, \sigma^2)$ di $N(\mu, \sigma^2)$ è

$$\hat{a} = \Big(\frac{1}{n}(X_1 + ... + X_n),\ \frac{1}{n}((X_1-\bar{X})^2 + ...(X_n-\bar{X})^2\Big).$$

Si noti che la seconda coordinata è stimatore distorto di σ^2.

58 Intervalli di fiducia e caso di μ di $N(\mu, \sigma^2)$

58.1 Introduzione agli intervalli di fiducia

Se X è una variabile aleatoria normale $N(\mu, \sigma^2)$ con μ incognito, e ne avremo un campione aleatorio X_1,...,X_n (variabili aleatorie indipendenti tutte con la stessa legge $N(\mu, \sigma^2)$) oppure ne abbiamo delle determinazioni (numeri) x_1,...,x_1, possiamo stimare μ con la media campionaria $\hat{\mu} := \bar{X}_n$ e allora numericamente $\hat{\mu} = \bar{x}_n$. Potrebbe essere per esempio $\bar{x}_n \approx 2.018$ sia se $n = 1$ sia se $n = 100$: la *stima puntuale* ottenuta non distingue i due casi mentre è chiaro che nel secondo caso il valore 2.018 è molto più "sicuro" se non nella sua esattezza (che comunque ha probabilità 0) almeno nella sua vicinanza al valore vero.

La *stima intervallare* è costituita invece da 2 stimatori $\hat{u}$ e $\hat{v}$ tali che il parametro incognito, sia ora esso a (per esempio μ di $N(\mu, \sigma^2)$) sia nell'*intervallo aleatorio* $[\hat{u}, \hat{v}]$ con probabilità $\geq 95\%$

$$P(a \in [\hat{u}, \hat{v}]) \geq 95\%$$

fino a che $\hat{u}$ e $\hat{v}$ sono variabili aleatorie dipendenti da X_1,...,X_n, cioè prima di venire determinate coi dati numerici x_1,...,x_n: dopo, μ o sta o non sta nell'intervallo determinato, e sarebbe arduo anche solo definire in che senso si potrebbe attribuire una probabilità a quel fatto; si potrebbe farlo solo con la concezione soggettiva della probabilità ma con essa si può proporre *qualunque* valore.

(è un po' come chiedersi che probabilità ha 2017 di essere primo: o è primo o non lo è, non c'è una grande questione probabilistica).

Fissati i valori x_1,...,x_n e conseguentemente i valori di $\hat{u}$ e $\hat{v}$ non si parla più di probabilità ma di *(livello di) confidenza.*

Tutto questo si estende mutando la soglia 95% in qualunque altro *livello* $1 - \alpha$, e normalmente si usano anche 90% e 99%, e si estende a qualunque parametro incognito di una densità per il resto nota. Il complemento α di $1 - \alpha$ si chiama *(livello di) significatività*, e in generale è $5\% = 0.05$ o $10\% = 0.1$ o $1\% = 0.01$.

Purtroppo α e $1 - \alpha$ vengono scambiati fra loro nei vari testi.

58.2 Intervalli di fiducia per μ per campioni gaussiani

Precisiamo dapprima che *intervallo di fiducia* e *intervallo di confidenza* sono sinonimi.
In questo paragrafo considereremo:
• la stima intervallare di μ essendo noto σ^2;
• la stima intervallare di μ essendo ignota anche σ^2.
Tutte le formule di questo paragrafo valgono per campioni gaussiani e (approssimatamente) anche per campioni non gaussiani se la numerosità è $n \geq 30$ e la densità non è "troppo" asimmetrica.
Per μ con σ^2 <u>nota</u>, intervallo di fiducia bilatero centrato al livello (di confidenza) $1-\alpha$, dove α è "piccolo", come $0.05 = 5\%$ (da non confondersi con $1-\alpha$, "grande"):

$$\bar{X}_n \pm \frac{\sigma}{\sqrt{n}}\,\phi_{1-\frac{\alpha}{2}} \quad \textit{ovvero} \quad \left[\bar{X}_n - \frac{\sigma}{\sqrt{n}}\,\phi_{1-\frac{\alpha}{2}},\, \bar{X}_n + \frac{\sigma}{\sqrt{n}}\,\phi_{1-\frac{\alpha}{2}}\right]$$

e in particolare ricordiamo il valore $\phi_{0.975} \approx 1.96$ da cui l'intervallo classico con livello di confidenza del 95%, cioè $\alpha = 0.05$,

$$\left[\bar{X}_n - 1.96\,\frac{\sigma}{\sqrt{n}},\, \bar{X}_n + 1.96\,\frac{\sigma}{\sqrt{n}}\right] \quad (\alpha = 0.05)$$

$$\text{ovvero} \quad \bar{X}_n \pm 1.96\,\frac{\sigma}{\sqrt{n}} \quad (\alpha = 0.05)$$

e spesso nelle scienze applicate 1.96 viene approssimato con 2.
Come sopra, per μ con σ^2 <u>non nota</u>:

$$\left[\bar{X}_n - \frac{S_n}{\sqrt{n}}\,t_{1-\frac{\alpha}{2}}(n-1),\, \bar{X}_n + \frac{S_n}{\sqrt{n}}\,t_{1-\frac{\alpha}{2}}(n-1)\right]$$

$$\text{ovvero} \quad \bar{X}_n \pm \frac{S_n}{\sqrt{n}}\,t_{1-\frac{\alpha}{2}}(n-1)$$

con i quantili di Student e la radice quadrata S_n dello stimatore S_n^2 della varianza. Si noti che esistono infiniti intervalli bilateri allo stesso livello, non centrati in $\bar{X}_n$, e 2 unilateri.
Per μ, un intervallo <u>unilatero</u> al livello di confidenza $1-\alpha$:

$$\left]-\infty,\, \bar{X}_n + \frac{S_n}{\sqrt{n}}\,t_{1-\alpha}(n-1)\right]. \quad \textit{(Qua ovviamente } \sigma^2 \textit{ non nota).}$$

59 Intervalli di fiducia per la varianza

Per σ^2 con μ <u>non nota</u>, (un) intervallo di fiducia <u>bilatero</u> al livello (di confidenza) $1-\alpha$, dove α è "piccolo", come $0.05 = 5\%$ (da non confondersi con $1-\alpha$, "grande"):

$$\left[\frac{n-1}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\,S_n^2\,,\ \frac{n-1}{\chi^2_{\frac{\alpha}{2}}(n-1)}\,S_n^2\right]. \qquad (\textit{Non centrato in } S_n^2).$$

Come sopra, un intervallo <u>unilatero</u>:

$$\left[0\,,\ \frac{n-1}{\chi^2_{\alpha}(n-1)}\,S_n^2\right].$$

Nota 1. Le 2 formule soprastanti valgono per campioni gaussiani e (approssimatamente) anche per campioni non gaussiani se la numerosità è $n \geq 30$ e la densità non è "troppo" asimmetrica.
Nota 2. I quantili del chi quadrato, oltre a venire calcolati (approssimatamente) da molti software, si trovano (approssimati) per alcuni tipici valori di α e piccoli valori di n su apposite tavole numeriche, e poi vale (teorema) l'approssimazione

$$\forall\, n \geq 30 \qquad \chi^2_{\alpha}(n) \approx \frac{1}{2}\Big(\phi_\alpha + \sqrt{2\,n-1}\Big)^2.$$

Nota 3. Si possono considerare anche intervalli di fiducia per altri parametri e altri tipi di variabili aleatorie, per esempio↑ per p di $B(1,p)$.

Esercizio. (RISOLTO A LEZIONE) Sia
5 8.5 12 15 7 9 7.5 6.5 10.5
un campione $N(\mu, \sigma^2 = 4)$. Determinare l'intervallo di fiducia usuale al livello $1 - \alpha = 0.99$ (da altri detto al livello 0.01). Quanto dovrebbe essere numeroso un analogo campione affinchè l'ampiezza dell'intervallo di fiducia al livello 0.05 sia ≤ 0.4?
Nota. Usuale = quello bilatero centrato.

XII—— Test statistici

60 Test ed errori di prima e seconda specie

Dalle determinazioni x_1,...,x_n di un campione aleatorio X_1,...,X_n tratto da una v.a. X di densità nota salvo un suo parametro a, vogliamo rispondere con "sì" o "no" , con "ragionevole certezza statistica" ovviamente, a una domanda sul parametro incognito. La domanda potrebbe essere per esempio $\mu > 0$, oppure $p = \frac{1}{2}$, che nella realtà sensibile può significare per esempio la regolarità di una moneta, modellizzata con una v.a. $B(1, p)$.
Il test statistico si preordina – prima di avere i dati i mano ovvero prima di fare un esperimento nella realtà sensibile – formulando un'ipotesi statistica, indicata con H, *ipotesi nulla*, e una ipotesi *alternativa*, indicata con A, per esempio

$$H : p = \frac{1}{2} \qquad A : p \neq \frac{1}{2} \, .$$

Anticipiamo che l'ipotesi nulla H va identificata in generale col caso che si spera che non sia. ("Vogliamo A!").
Un esempio minimo potrebbe essere così: lanceremo 5 volte la moneta e rifiuteremo l'ipotesi di regolarità se viene testa 0 o 5 volte, perchè se la moneta è regolare quei risultati hanno complessivamente probabilità 1/16, un po' pochino. (In realtà la statistica usuale viene fatta "a 1/20", cioè al 5 ovvero 95%). L'insieme $\{0, 5\}$ è la *regione critica* ovvero di rigetto dell'ipotesi (nulla). La regione critica di solito viene espressa come un sottoinsieme D di $\mathbb{R}$ in cui una certa funzione del campione aleatorio può cadere (e allora rifiutiamo H) o non cadere. In questi termini, potremmo porre la regione critica $\{0, 5\}$ e verificare se vi cade $X_1 + ... + X_5$ o più usualmente porre $D := \{0, 1\}$ e verificare se vi cade $\bar{X}_5$. In casi più significativi di questo microscopico esempio la regione critica di solito ha una forma del tipo $x > x_0$ (test unilatero) oppure $x < x_1 \vee x > x_2$ (test bilatero).
Vediamo un altro esempio. Sia U la variabile aleatoria che è la glicemia in Italia e V la glicemia dopo la somministrazione di un certo farmaco. Ci potrebbe interessare se mediamente la glicemia

diminuisce con quel farmaco cioè se la media (parametro incognito) della variabile aleatoria $X := U - V$ è > 0.
Si formula l'ipotesi nulla: il farmaco non riduce la glicemia (in realtà spero che riduca la glicemia: ipotesi alternativa). Misuriamo la glicemia in 20 soggetti (campione, o più precisamente determinazione $u_1, \ldots, u_{20}$ di un campione aleatorio $U_1, \ldots, U_{20}$). Diamo ai 20 soggetti il farmaco. Misuriamo di nuovo la glicemia dei 20 soggetti ottenendo così 20 determinazioni $v_1, \ldots, v_{20}$ della variabile aleatoria V. Facciamo 20 sottrazioni ottenendo 20 determinazioni $x_1, \ldots, x_{20}$ della variabile aleatoria X. Della variabile aleatoria X vogliamo sapere se la media (speranza matematica) è > 0. L'idea ingenua è fare la media aritmetica dei 20 numeri e concludere che se è > 0 il farmaco ha diminuito la glicemia. Se fosse così in questo e analoghi casi, la statistica inferenziale non servirebbe, ma non è così: quella verifica non dice di per sè sostanzialmente nulla perchè non distingue l'effetto del farmaco dalle inevitabili fluttuazioni casuali di X, che, non per niente, è da considerarsi una variabile *aleatoria*. (Non possiamo certo aspettarci un effetto deterministico del farmaco, che *sempre* riduca la glicemia). è invece necessario applicare un opportuno test statistico, cioè di fatto applicheremo una non banale formula che ci potrà dire, nel caso che la media aritmetica degli x_i sia < 0, che quell'effetto con ragionevole plausibilità non è casuale. Se invece la media aritmetica viene positiva l'esperimento è andato male e la statistica non ci aiuta ulteriormente. Alla fine rifiutiamo o non rifiutiamo l'ipotesi nulla. Speriamo di rifiutarla.

Media degli X_i

Statistica ingenua:

il farmaco funziona — il farmaco non funziona

0

Statistica inferenziale:

rifiutiamo l'ipotesi che la media sia $\geq$ a prima — non rifiutiamo l'ipotesi che la media sia $\geq$ a prima

soglia 0

Come detto, prima di eseguire un test statistico dobbiamo formulare un'ipotesi nulla H e la complementare ipotesi alternativa A. Per esempio

$$H: \quad \mu > 0$$
$$A: \quad \mu \leq 0$$

dove μ potrebbe essere la media di una variabile aleatoria X di cui disporremo di un campione aleatorio $X_1, \ldots, X_n$. I ruoli delle 2 ipotesi non sono interscambiabili: in linea generale come ipotesi alternativa va fissata quella che speriamo vera (e come ipotesi nulla quella che ci avrebbe fatto perdere tempo).
Tratto ovvero prodotto ovvero rilevato il campione x_1,...,x_n, ne calcoliamo l'opportuna funzione che la statistica ci insegnerà a seconda del tipo di test, sia essa ora $g(x_1, .., x_n)$, per esempio la media $\bar{x}_n$, ci sono 4 casi relativamente alla regione critica D, a $g(x_1, .., x_n)$, e alla verità dell'ipotesi H:

1) $g(x_1, .., x_n) \in D$ ed è vera H:
male respingo ipotesi vera: errore di prima specie.

2) $g(x_1, .., x_n) \in D$ ed è vera A:
bene respingo ipotesi falsa: è il caso sperato.

3) $g(x_1, .., x_n) \notin D$ ed è vera H:
non respingo ipotesi vera: ho perso tempo.

4) $g(x_1, .., x_n) \notin D$ ed è vera A:
male non respingo ipotesi falsa: errore di seconda specie.
L'errore di prima specie è considerato molto più grave. Ad esempio per un farmaco si può mettere l'ipotesi che non curi, sperando di falsificarla con l'esperimento. I 4 casi in dettaglio sono questi.
1) e 3) è vera l'ipotesi H: il farmaco è inutile o dannoso (esempio: sopravvivenza a 5 anni uguale o diminuita).
2) e 4) è vera l'alternativa A: il farmaco è utile (esempio: sopravvivenza a 5 anni aumentata).

1) e 2) La sperimentazione dà esito buono.

1) Per caso i soggetti trattati sono vissuti a lungo e il farmaco dannoso fa bella figura e magari si diffonde: GRAVISSIMO.

2) I soggetti trattati sono vissuti a lungo grazie al farmaco che giustamente fa bella figura

3) e 4) La sperimentazione dà esito cattivo.

3) non respingo l'ipotesi che il farmaco sia inutile o dannoso; il farmaco inutile o nocivo è stato correttamente riconosciuto tale sperabilmente non verrà commercializzato. Ho perso tempo, speravo fosse utile.

4) Per caso i soggetti trattati sono vissuti poco ma il farmaco in generale funziona: il farmaco utile viene purtroppo abbandonato. Peccato.
Questo è un modello molto semplificato perchè tiene conto di un solo parametro, comunque di altissimo valore. Nella pratica si usano talvolta parametri molto più rapidi da verificare, come la diminuzione della massa tumorale in un certo tempo; se poi in un tempo doppio il paziente muore, questo non rientra nella statistica. è ovvio che la sopravvivenza a 5 anni richiede una sperimentazione lunghissima, e costosa.
Ma per quanto di altissimo valore, anche quel solo parametro è poco, nella complessità della realtà. Non si è tenuto conto della qualità della vita in seguito alla somministrazione del far-

maco, nè del costo economico: gli stessi soldi il Sistema Sanitario potrebbe spenderli con maggiore beneficio complessivo per la cura di un'altra malattia.

61 Test di Student per la media $\leq$ e $=$

Teorema. Sia X_1,...,X_n un campione gaussiano cioè $X_i \sim N(\mu, \sigma^2)$ sono indipendenti (o anche non gaussiano se n è sufficientemente grande), sia μ_0 è un numero reale, e $T(X_1, ..., X_n)$ sia definito da

$$T := \sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n}$$

(coi soliti stimatori $\bar{X}_n$ e S_n^2 di media evarianza). Allora si hanno questi 2 test con le rispettive regioni critiche al livello $1 - \alpha$:

$$H : \mu \leq \mu_0 \quad A : \mu > \mu_0 \qquad T > t_{1-\alpha}(n-1)$$

$$H : \mu = \mu_0 \quad A : \mu \neq \mu_0 \qquad |T| > t_{1-\frac{\alpha}{2}}(n-1).$$

(Altri dicono al livello α; in ogni caso α è "piccolo").
(Si calcola T coi valori x_1,..,x_n e se supera il quantile si respinge l'ipotesi, altrimenti "non la si respinge").
Esercizio del libro di Paolo Baldi già citato:

> L'altezza media delle reclute alla visita di leva del 1970 era di 169 cm; 121 reclute vengono scelte a caso nel 1980. I valori di media varianza del campione sono:
>
> $$\bar{X} = 171 \quad \leftarrow \textit{intende } \bar{X}_n \textit{ con } n{:=}121$$
>
> $$S^2 = 85 \quad \leftarrow \textit{intende } S_n^2 \textit{ con } n{:=}121$$
>
> Si può affermare che l'altezza media delle reclute è aumentata?

(Siccome si vuole dimostrare che l'altezza è aumentata si metta come ipotesi che non è aumentata, e siccome nulla è detto sul livello del test si usi la classica soglia del 5%).

62 Campioni indipendenti: un Test di Student

Ecco un test di Student per le medie di 2 campioni indipendenti.
Teorema. Siano $X_1,...,X_n$ un campione aleatorio (gaussiano oppure con n sufficientemente grande) di media μ_X e varianza σ_X^2 sconosciute e $Y_1,..,Y_m$ un campione aleatorio (gaussiano oppure con m sufficientemente grande) di media μ_Y e varianza $\sigma_Y^2 = \sigma_X^2 =: \sigma^2$ (uguale a quella del primo campione e sconosciuta), con le X_i e le Y_i indipendenti fra loro. Siano

$$S_{tot}^2 := \frac{1}{n+m-2}\Big(\sum_{i=1}^{n}(X_i-\bar{X}_n)^2+\sum_{i=1}^{m}(Y_i-\bar{Y}_m)^2\Big)$$

$$\Big[= \frac{1}{n+m-2}\Big((n-1)S_X^2+(m-1)S_Y^2\Big)\Big]$$

$$T := \frac{\bar{X}_n-\bar{Y}_m}{S_{tot}\sqrt{\frac{1}{n}+\frac{1}{m}}}\,.$$

Allora

(1) (test bilatero)

$$|T| > t_{1-\frac{\alpha}{2}}(n+m-2)$$

definisce una regione critica al livello [di significatività] α per il test

$$H : \mu_X = \mu_Y \qquad A : \mu_X \neq \mu_Y.$$

(2) (test unilatero)

$$T < t_{1-\alpha}(n+m-2)$$

definisce una regione critica al livello [di significatività] α per il test

$$H : \mu_X > \mu_Y \qquad A : \mu_X \leq \mu_Y.$$

??? RIVEDERE

63 Il Test del χ^2, quello basico

63.1 La teoria del Test del χ^2, quello basico

Il primo test del chi quadrato, dei due che considereremo, lo chiameremo semplicemente *Test del chi quadrato*.
Si riferisce a una variabile aleatoria discreta, sia essa X, con un numero finito k di valori che possiamo identificare con k numeri interi, per esempio 0,1 oppure 1,...,6; per fissare le idee diciamo 1,...,k. Nel discorso che segue si possono immaginare, per fissare le idee, $n = 2000$ lanci di un dado. Se abbiamo n determinazioni x_1,...,x_n di X, ce ne sono n_1 che valgono 1, n_2 che valgono 2,..., n_k che valgono k. Sono le *"frequenze assolute (osservate ovvero empiriche)"*. (Dividendole per n si ottengono le *"frequenze relative (osservate ovvero empiriche)"* $\bar{p}_i := \frac{n_i}{n}$, che tendenzialmente dovrebbero assomigliare alle p_i, le probabilità *"teoriche"*, e con esse si ottiene la formula alternativa che daremo in parentesi).
Il test vuole respingere l'ipotesi che la densità di X sia data dai valori $(p_1, ..., p_k)$**.** Per esempio escludere la regolarità di un dado, cioè $H: \ (p_1, ..., p_6) = (\frac{1}{6}, ..., \frac{1}{6})$, $A: \ (p_1, ..., p_6) \neq (\frac{1}{6}, ..., \frac{1}{6})$. (Naturalmente gli indici potrebbero essere 0,...,$k-1$).
Fissato *prima* di vedere i dati (come sempre si deve fare in un test statistico serio) il livello di significatività α, tipicamente 0.05, il rifiuto si ha se

$$T_n := \sum_{i=1}^{k} \frac{(n_i - n\,p_i)^2}{n\,p_i} > \chi^2_{1-\alpha}(k-1) \quad \left(\text{o con: } n \sum_{i=1}^{k} \frac{(\bar{p}_i - p_i)^2}{p_i}\right)$$

purchè $n\,p_i \geq 5$ per tutti gli $i = 1, ..., k$.
Se con l'ultima classe si ha $n\,p_k < 5$, si associano le classi $k-1$ e k, e se col nuovo $\tilde{p}_{k-1} = p_{k-1} + p_k$ si ha $n\,\tilde{p}_{k-1} \geq 5$ si fa il test per la nuova v.a. a $k-1$ valori. Per esempio si considera un dado "a 5 valori": 1, 2, 3, 4, $5 \vee 6$, con frequenze assolute n_1, n_2, n_3, n_4, $n_5 + n_6$, e nell'ipotesi p_1, p_2, p_3, p_4, $p_5 + p_6$. Similmente se la prima, o più raramente un'altra probabilità, è troppo piccola, $n\,p_m < 5$, si può provare ad associare 2 classi consecutive.

63.2 Esempio di applicazione del Test del χ^2

Si lancia una moneta 3 volte, e questo lo si fa per 100 volte, ottenendo

0 teste 31 volte
1 testa 36 volte
2 teste 32 volte
3 teste 1 volta.

Eventualmente associando opportunamente due classi verificare l'ipotesi che la moneta dia testa con probabilità $\frac{1}{3}$. [...]

Esercizio. Nelle stesse ipotesi dell'esempio visto si verifichi l'ipotesi della regolarità della moneta.

• • •

Esempio di Wikipedia tratto dal già citato libro di P. Baldi. Con 2000 lanci di un dado con frequenze assolute osservate 388, 322, 314, 316, 344, 316, si rifiuta, al livello del 5%, ovvero del 95%, la regolarità perchè

$$T \approx 12.6 > 11.07 \approx \chi^2_{0.95}(5).$$

(Naturalmente T è T_{2000}).

64 Test del χ^2 di indipendenza

In questa trattazione elementare, verrà considerato solo il caso più semplice, con 4 valori/dati. Fissiamo subito le idee su un esempio. Supponiamo di avere 280 monete uguali di sconosciuta probabilità di dare testa, e che a 50 di esse venga fatta meccanicamente una piccola incisione al centro sulla faccia della testa. Le monete trattate hanno tutte un'uguale sconosciuta nuova probabilità di dare testa. (Di maggiore interesse potrebbe essere il trattamento con un farmaco, e l'eventuale morte a 5 anni, ma le monete e i dadi restano modelli generali ottimali, anche perchè da essi *a priori ci aspettiamo* di default un comportamento casuale, mentre per un farmaco saremmo portati a supporre causalità – contenendo la tal molecola *dovrebbe* fare un certo effetto – eventualmente inesistenti: nella *statistica medica* dobbiamo appunto verificare se funziona, *contando i morti*, per così dire). Lanciamo una volta ciascuno delle 280 monete e contiamo le teste nelle 2 classi. Supponiamo di avere questa situazione (risultato=viene 1)

	non risultato	risultato
non trattati	188	42
trattati	43	7

Naturalmente $43+7$ sono le 50 monete trattate, e $188+42$ sono le rimanenti 230, non trattate. Il trattamento ha diminuito la probabilità di dare testa? Verrebbe da dire di sì perchè fra i trattati la frequenza relativa di 1 è $7/50 = 0.14$ mentre fra i non trattati è $42/230 = 0.183$ ma dobbiamo ragionevolmente escludere un effetto casuale: tale miglioramento potrebbe apparire per caso anche se a nostra insaputa il tecnico preposto non avesse affatto compiuto l'incisione dichiarata.

Abbiamo una variabile aleatoria X con 2 valori, "trattate" e "non trattate", e una variabile aleatoria Y con 2 valori, "testa" e "croce". Il test statistico è definito da

$$H : \text{indipendenza} \qquad A : \text{non indipendenza}$$

(vorremmo escludere l'indipendenza, cioè che l'incisione non abbia prodotto risultato).
Prima di tutto bisogna passare dalle frequenze empiriche assolute a quelle relative dividendo i $2 \cdot 2$ numeri (in una forma più generale il test si fa con $m \cdot n$ dati) per la loro somma N, cioè qua i 280 soggetti:

frequenze empiriche relative		
	$\bar{r}_{1,1} = 0.671429$	$\bar{r}_{1,2} = 0.15$
	$\bar{r}_{2,1} = 0.153571$	$\bar{r}_{2,2} = 0.025$

Poi si calcolano le marginali relative sommando per righe e per colonne i predetti numeri:

marginali			
			0.821429
			0.178571
	0.825	0.175	(= 1)

Poi si calcolano le frequenze relative attese/teoriche, nell'ipotesi dell'indipendenza, ottenute come prodotti delle predette marginali:

frequenze relative attese		
	$a_{1,1} = 0.677679$	$a_{1,2} = 0.14375$
	$a_{2,1} = 0.147321$	$a_{2,2} = 0.0312499$

Il test per funzionare ragionevolmente bene richiede che tutte le 4 frequenze assolute attese siano ≥ 5, ed esse si ottengono moltiplicando per N le frequenze relative attese $a_{h,k}$, e all'atto pratico basta fare la moltiplicazione solo per la minima frequenza attesa relativa, qua $a_{2,2}$:

frequenze assolute attese		
	≥ 5	≥ 5
	≥ 5	$8.74997 \geq 5$

Fisseremo allora un livello di confidenza $1-\alpha$, per esempio 95%. La regione critica (di respingimento dell'ipotesi nulla) al livello [di significatività] α del test è definita da

$$T := n \sum_{h=1}^{m=2} \sum_{k=1}^{n=2} \frac{(a_{h,k} - \bar{r}_{h,k})^2}{\bar{r}_{h,k}} > \chi^2_{1-\alpha}(g)$$

col numero g di gradi di libertà

$$g := (m-1)\cdot(n-1) = (\#righe - 1) \times (\#colonne - 1) \qquad \text{qua } 1$$

e se [lo stimatore/statistica] T supera il quantile si respinge l'ipotesi, altrimenti non la si respinge (che alcuni dicono "si accetta").
Coi dati considerati:

$$T = \Big(\frac{(a_{1,1} - \bar{r}_{1,1})^2}{\bar{r}_{1,1}} + \frac{(a_{1,2} - \bar{r}_{1,2})^2}{\bar{r}_{1,2}} + \frac{(a_{2,1} - \bar{r}_{2,1})^2}{\bar{r}_{2,1}} + \frac{(a_{2,2} - \bar{r}_{2,2})^2}{\bar{r}_{2,2}}\Big) \approx$$

$$\approx 280\,(\frac{(0.677679\ - 0.671429)^2}{0.671429} + \frac{(0.14375\ - 0.15)^2}{0.15} +$$

$$+ \frac{(0.147321\ - 0.153571)^2}{0.153571} + \frac{(0.0312499\ - 0.025)^2}{0.025}) \approx$$

$$\approx 0.597914 \not> 3.84 \approx \chi^2_{0.95}(1)$$

e allora l'ipotesi dell'indipendenza (ovvero della casualità della diminuzione percentuale di teste nelle monete trattate) non è respinta al livello di confidenza del 95% (ovvero di significatività del 5%). Non è respinta neanche al miserevole livello di significatività del 30%. Il valore soglia fra l'accettazione e il respingimento nei test statistici si chiama p-value e tanto più è piccolo tanto più "statisticamente sicura" è l'affermazione conclusiva.

Nota 1. Lo stimatore T è tanto più piccolo, e quindi porta a non respingere l'ipotesi di indipendenza, quanto più le frequenze empiriche relative assomigliano al prodotto delle marginali, ciò che caratterizza appunto l'indipendenza.
E viceversa, con effetto opposto, tanto più grande è la numerosità N tanto più grande è T.
Nota 2. Tutti questi calcoli statistici con somme di quadrati di differenze vanno fatti con molti decimali – diciamo qua 5 – altrimenti si possono ottenere risultati non significativi, a causa delle approssimazioni successive.
Nota 3. Si può evitare l'approssimazione che è intrinseca nel test del chi quadrato d'indipendenza col *Test di Fischer esatto*, che ha una formula molto più complicata che viene comunque calcolata in un'istante dai computer: qua si è voluto *far capire* la natura della questione, per premere tasti c'è sempre tempo.

— ◦ —

65 Indice Analitico con Approfondimenti

65.1 Alcuni valori di $P(\lvert X\rvert < x)$ per X normale standard

x	$P(\lvert X\rvert < x)$
1	≈ 0.68
1.28	≈ 0.8
1.64	≈ 0.9
1.96	≈ 0.95
2	≈ 0.9544
2.58	≈ 0.99
3	≈ 0.9973

(Si noti che questi non sono i valori di $P(X \leq x)$) per X normale standard, cioè $\Phi(x)$).

65.2 Algebra astratta

Nell'algebra astratta si considerano operazioni generiche definite in insiemi generici, non identificando le operazioni come somma o prodotto per esempio, nè gli insiemi come $\mathbb{N}$ o $\mathbb{R}$ per esempio. Facendo quelle scelte di specifici insiemi e specifici insiemi, si ottiene l'*algebra dei numeri*; si veda qua`<--` e qua`<--`. Nell'algebra astratta per esempio si possono considerare, fra le moltissime altre cose, l'operazione binaria (interna), l'elemento neutro, l'elemento assorbente, l'elemento simmetrico, la proprietà associativa, la proprietà commutativa, la proprietà commutativa. Concetti di livello superiore sono il monoide, il gruppo, l'anello, il corpo e il campo, lo spazio vettoriale.

65.3 Ambiguità notazionale del segno $=$

• Il segno $=$ viene usato per indicare un'uguaglianza che sussiste fra 2 funzioni eventualmente per 1 o più valori di x, y... i quali vengono ricercati. Si ha cioè un'equazione, per esempio:

$$\sin^2 x + \cos^2 x = 0.1 \qquad x^2 + 5x + 6 = 0 \qquad x\,y = 0\,.$$

• Il segno $=$ viene usato per indicare un'uguaglianza fra valori di 2 funzioni che vale $\forall\, x, y ...$ del loro dominio, per esempio

$$\sin^2 x + \cos^2 x = 1 \qquad (x+y)^2 = x^2 + 2\,x\,y + y^2$$

cioè per un'identità (funzionale). In questo caso talvolta si usa il simbolo $\equiv$

$$\sin^2 x + \cos^2 x \equiv 1 \qquad (x+y)^2 \equiv x^2 + 2\,x\,y + y^2$$

• Il segno $=$ viene usato per indicare una definizione: il termine a destra del segno di uguaglianza ha già senso nella trattazione finora svolta, e il termine a sinistra lo acquisisce in base all'espressione stessa, per esempio:

$$f(x) = x^2 + 1 \qquad \cosh x = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2} \qquad x_{n+1} = 2\,x_n\,.$$

Talvolta in questo senso si scrive $:=$ invece di $=$

$$f(x) := x^2 + 1 \qquad \cosh x := \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2} \qquad x_{n+1} := 2\,x_n\,.$$

65.4 Ambiguità notazionale dell'esponente -1

Una delle pù gravi ambiguità notazionali è quella dell'esponente -1:

$$a^{-1}$$

indica inequivocabilmente il reciproco $\frac{1}{a}$ di a se a è un numero, ma se è una funzione – e può ben esserlo – cominciano problemi:

$$f^{-1}(x)$$

in questa trattazione significherà sempre l'inversa di $f(x)$ mentre la reciproca la denoteremo $\frac{1}{f(x)}$ e casomai, volendoci del male, $(f(x))^{-1}$. Il problema maggiore si ha quando – e concretamente ciò viene fatto spesso – in lunghi calcoli per brevità di scrittura si omette di indicare la variabile indipendente, e si scrive f intendendo $f(x)$, lasciando nell'incertezza riguardo la scrittura f^{-1}.

Così all'atto pratico ci si ritrova sempre con qualcuno che pensa che l'arcotangente sia il reciproco della tangente, o il logaritmo il reciproco dell'esponenziale (invece è l'inversa).
Resta il fatto che in un testo diverso da questo, la scrittura

$$\ln^{-1}(x)$$

è ambigua, e può indicare sia $\exp(x)$ che $\frac{1}{\ln(x)}$. Qua solo $\exp(x)$.

65.5 Ambiguità notazionali della parentesi graffa grande

La parentesi graffa grande – di quella sola ci si occuperà qui – viene usata con almeno 3 diversi significati che vediamo con 3 esempi.
Parentesi graffa (grande) *"selettiva dei casi"*:

$$|x| := \begin{cases} x & se\ \ x \geq 0 \\ \\ -x & se\ \ x < 0 \end{cases}$$

Parentesi graffa (grande) *"onnicomprensiva dei casi"*:

$$\cos(2x) = \begin{cases} \cos^2 x - \sin^2 x \\ 2\cos^2 x - 1 \\ 1 - 2\sin^2 x \end{cases}$$

Parentesi graffa (grande) *dei sistemi*, con significato di *et*:

$$\begin{cases} x^2 > 9 \\ x + 1 < 0 \end{cases}$$

è chiaro che in tutti questi significati, c'è un *et* "retrostante", ma non in modo immediato.

65.6 Ambiguità notazionale della virgola nei numeri

In Italia un numero come

$$148,128$$

spesso sarà da intendere come 148 virgola 128, poco meno di 150. In generale nella letteratura scientifica internazionale invece sarà da intendere come centoquarantottomilacentoventotto, e la virgola è scritta solo per facilitare la lettura, raggruppando le cifre a 3 a 3. Certo il dubbio ci potrebbe rimanere. Questo numero esprime il volume di Giove a meno di un fattore $10^{10} km^3$, come leggiamo su un sito internet della NASA, e la massa della Terra risulta invece 108.321: ecco che *in questo caso* non ci sono dubbi: quella che spesso in Italia denotiamo con la virgola qua la scrivono col punto.

65.7 Ambiguità notazionali per i logaritmi

Da Wikipedia, l'enciclopedia libera, alla voce "Logaritmo naturale":

> • I matematici sono soliti utilizzare la scrittura "`log(x)`" per intendere $\texttt{log}_e\texttt{(x)}$; altrimenti si è soliti specificare la base nella scrittura (es. $\texttt{log}_{10}\texttt{(x)}$ è il logaritmo in base 10 di x).
> • Ingegneri, biologi e altre professioni generalmente scrivono "`ln(x)`" o (raramente) "$\texttt{log}_e\texttt{(x)}$" per intendere il logaritmo naturale di x, mentre per "`log(x)`" sottintendono $\texttt{log}_{10}\texttt{(x)}$.
> • Nei più comuni linguaggi di programmazione, tra cui C, C++, Fortran, e BASIC, "`log`" o "`LOG`" sottintendono il logaritmo naturale.
> • Nelle calcolatrici il logaritmo naturale è "`ln`", mentre "`log`" è il logaritmo in base `10`.

Da Wikipedia, l'enciclopedia libera, alla voce "Logaritmo":

> base **10** [...] li si indica con log_{10}, o con *Log* [...] simbolo ISO *lg* [...]
> base 2 [...] li si indica con log_2, oppure con *log* quando la base a cui ci si riferisce è chiara dal contesto (simbolo ISO *lb*).

Pregiate calcolatrici denotano LOG o log il logaritmo in base 10, e LN o ln quello naturale. Si aggiunga che nei linguaggi di programmazione si usano altre scritture ancora.
Lo standard ISO 80000-2:2009 per il logaritmo decimale è lg, e per il logaritmo naturale ln.

65.8 Calcolo Combinatorio: qualche nota

Esistono un'infinità di metodi e formule nel calcolo combinatorio che non sono considerati in questa trattazione elementare. Esistono addirittura riviste scientifiche interamente dedicate alla questione, come il Journal of Combinatorics.
Nel 2017 Wikipedia elenca in https://en.wikipedia.org/wiki/List_of_unsolved_problems_in_mathematics#Combinatorics ben 10 problemi irrisolti di calcolo combinatorio e ne esistono moltissimi altri.

65.9 Confronto fra media, mediana e moda

Su Wikipedia, l'enciclopedia libera, alla voce "Mode (statistics)" troviamo questo significativo esempio: 1, 2, 2, 3, 4, 7, 9; media=4, mediana=3, moda=2.

65.10 Elemento assorbente

Sia $\star$ un'operazione binaria (interna) definita in un insieme E. Si definisce **elemento assorbente** (rispetto all'operazione $\star$) un elemento, indichiamolo qua con α, tale che $x \star \alpha = \alpha \star x = \alpha$ per ogni x [di E]. Per esempio lo 0 per il $\cdot$, in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

65.11 Elemento neutro

Sia $\star$ un'operazione binaria (interna) definita in un insieme E. Si definisce **elemento neutro** (rispetto all'operazione $\star$) un elemento, indichiamolo qua con e, tale che $x \star e = e \star x = e$ per ogni x [di E]. Per esempio lo 0 per il $+$, e 1 per il $\cdot$, in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

65.12 Esercizio sull'arcoseno

ESERCIZIO$_{\mu 2018}$ Calcolare approssimatamente $\arcsin \frac{21}{25}$ (a mano, ovvio, ma anche prendere confidenza con una calcolatrice scientifica sarà un bene).
Si cerchi poi il valore online su WolframAlpha con

`ArcSin[21/25]`

Svolgimento
Calcoliamo $21 : 25 = 0.84$ (con la divisione a mano, o meglio osservando che $\frac{21}{25} = \frac{21}{25} \cdot \frac{4}{4} = 0.84$) e insomma cerchiamo un angolo, fra $-\frac{\pi}{2}$ e $\frac{\pi}{2}$, il cui seno sia 0.84. Ricordando (vedi (1)) che $\sin 1 \approx 0.84$ concludiamo $\arcsin \frac{21}{25} \approx 1$.
WolframAlpha con `ArcSin[21/25]` ci dà ≈ 0.997.

65.13 Eventi indipendenti

Nella concezione assiomatica della probabilità, nella σ-algebra $\not{A}$
• 2 eventi A e B si dicono indipendenti se e solo se

$$P(A \cap B) = P(A)P(B)\,;$$

• n eventi $A_1 \dots A_n$ si dicono a 2 a 2 indipendenti se e solo se

$$P(A_i \cap A_j) = P(A_i)P(A_j) \qquad \forall\, i, j = 1, \dots, n,\ i \neq j\,;$$

• n eventi $A_1 \dots A_n$ si dicono indipendenti se e solo se

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k})$$

per ogni scelta di indici distinti $i_1, \dots, i_k$.
Queste definizioni sono fatte in modo che l'indipendenza assiomatica (insieme alla probabilità assiomatica) corrisponda bene all'indipendenza degli eventi nel linguaggio comune, riferito alla realtà sensibile.
Per esempio è $\frac{1}{18}$ la probabilità che esca il 5 al lotto sulla ruota di Venezia alla prossima estrazione, ed è $\frac{1}{6}$ la probabilità che venga 5 al prossimo lancio di un dado regolare, e la probabilità che venga 5 al lotto e 5 sul dado (eventi indipendenti) è

$$P(\{5\ sulla\ ruota\ di\ Venezia\} \cap \{5\ sul\ dado\}) =$$

$$= P(5\ sulla\ ruota\ di\ Venezia) \cdot P(5\ sul\ dado) =$$

$$= \frac{1}{18} \cdot \frac{1}{6} = \frac{1}{108}\,.$$

Si noti che

$$P(B > 0) \wedge A, B\ indipendenti \Rightarrow P(A|B) = P(A)$$

conformemente alla nozione comune di indipendenza e probabilità condizionale.

65.14 Fattoriale

Per ogni $n \in I\!N$, si dice *fattoriale* di n e si indica con $n!$ il numero $1 \cdot 2 \cdot ... \cdot n$ se $n > 0$ e 1 se $n = 0$.
Per esempio $5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$; e $11! = 39\dot{}916\dot{}800$.
Esistono estensioni della definizione ai numeri reali e perfino ai numeri complessi mediante la funzione Gamma:

$$z! := \Gamma(z-1)\,, \quad z \neq -1, -2, -3, ...$$

Valgono le approssimazioni di Stirling

$$x! \approx \sqrt{2\pi x}\left(\frac{x}{e}\right)^x$$

e con maggiore precisione

$$x! \approx \sqrt{2\pi x}\left(\frac{x}{e}\right)^x\left(1 + \frac{1}{12x}\right)$$

per i numeri interi o reali, purchè sufficientemente grandi – diciamo maggiori di 8 per avere un'approssimazione all'1% per la prima e maggiori di 1 per avere un'approssimazione all'1 per mille per la seconda.
Sono approssimazioni asintotiche e allora il simbolo $\approx$ può essere opportunamente sostituito da $\sim$.
Si noti la grande rapidità con cui cresce la funzione fattoriale in $I\!N$, e in $I\!R$ per $x \geq 1$. Per esempio 100! ha 158 cifre.

65.15 Funzioni, tipi di

In questa trattazione elementare consideriamo funzioni di 5 tipi:
• $\mathrm{sgn}(x)$ che vale 0 in 0 e altrove $\frac{x}{|x|}$
• $\lfloor x \rfloor$ (scritta anche $[x]$ ma non è conforme allo standard ISO) parte intera di x, che per $x \geq 0$ "taglia i decimali": $[\pi] = 3$. (Ma $\lfloor -\pi \rfloor = -4$).
• $\Gamma(x)$, la funzione gamma.
• $\Phi(x)$, funzione utile in statistica.
• ϕ_α, funzione utile in statistica.

(Queste 5 sopra non sono funzioni elementari, le prime 2 sono discontinue e le ultime 3 continue).
• Le funzioni elementari: $x^2, \sin x$... e loro somme, composizioni...

65.16 Integrale definito di funzione priva di primitiva

In questa trattazione elementare, l'integrale definito è stato definito tramite la funzione primitiva, e allora non esiste per funzioni prive di primitiva. Ad un livello superiore, l'integrale $\int_a^b f(x)\,dx$ viene definito con un procedimento che sostanzialmente equivale a considerare l'area *con segno* fra l'asse x e il grafico di $f(x)$, con un'eventuale inversione di segno se $a > b$. Vediamo un esempio con una funzione, che essendo non continua nell'intervallo di integrazione, là è priva di primitiva.
Calcolare

$$\int_{-1}^{1} \operatorname{sgn} x\,dx\,.$$

Non abbiamo dato una primitiva (ovvero equivalentemente un integrale indefinito) di $\operatorname{sgn} x$, incertezza rig e in effetti si potrebbe dimostrare che non esiste, perchè solo le funzioni continue possono avere primitiva. In questo caso l'area di interesse esiste senz'altro come si vede dal disegno del grafico. Allora possiamo calcolare l'integrale così:

$$\int_{-1}^{1} \operatorname{sgn} x\,dx =_{\textit{per la Regola di Chasles}} \int_{-1}^{0} \operatorname{sgn} x\,dx\ + \int_{0}^{1} \operatorname{sgn} x\,dx\ =$$

ricordando che la funzione segno vale -1 sui negativi e 1 sui positivi (e quanto fa nel singolo valore 0 non importa: l'integrale definito di qualunque funzione non cambia modificando il valore della funzione integranda in un numero finito punti)

$$= \int_{-1}^{0} (-1)\,x\,dx\ + \int_{0}^{1} 1\,dx\ = [-x]_{-1}^{0} + [x]_0^1 =$$

$$= -0 - (-1) + 1 - 0 = 2\,.$$

Si faccia un disegno riconoscendo i vari elementi in questione.

65.17 Intervallo di fiducia per la v.a. bernoulliana

Per p di $B(1,p)$, un intervallo di fiducia <u>bilatero centrato</u> **al livello (di confidenza)** $1-\alpha$, dove α è "piccolo", come $0.05 = 5\%$ (da non confondersi con $1-\alpha$, "grande"), è (approssimativamente) <u>per n sufficientemente grande</u>

$$\left[\bar{X}_n - \phi_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}\,,\ \bar{X}_n + \phi_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}\right]$$

$$\text{ovvero} \quad \bar{X}_n \pm \phi_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}\,.$$

65.18 Involucro convesso

Diamo un'idea di cos'è l'*involucro convesso* di n punti, oppure infiniti, del piano o dello spazio (euclidei). Consideriamo dapprima 3 punti del piano: si immagini di piantare 3 pioli nei 3 punti, e di passare una corda intorno ad essi: la zona "entro" la corda è un triangolo, involucro convesso dei 3 punti. L'involucro convesso naturalmente esiste anche per 1 o 2 o più di 3 punti, anche infiniti, per esempio una figura non convessa. Per esempio l'involucro convesso di una stella ad m punte è un poligono ad m lati. Nello spazio euclideo si dovrà immaginare non un una corda ma un tessuto.

65.19 Isometrie e congruenza in piano e spazio euclidei

Detto E il piano euclideo, o lo spazio euclideo, ogni funzione da E in E che *conserva le distanze* si chiama *isometria* o *trasformazione rigida*.
Più formalmente, una funzione

$$f : E \to E$$

essendo E il piano euclideo o lo spazio euclideo, si dice *isometria* se

$$\Big(\forall P, Q \in E\Big) \quad d(f(P), f(Q)) = d(P, Q)$$

essendo d la *distanza* (euclidea).
Due sottoinsiemi del piano (o spazio) euclideo dei quali uno sia il trasformato isometrico dell'altro si dicono *congruenti*.
Le isometrie conservano gli angoli e le aree, e nello spazio i volumi.
Da adesso consideriamo solo il piano euclideo.
Nel piano euclideo esistono 5 tipi di isometrie:

- d – l'identità o trasformazione identica
- dd – traslazioni
- dp – rotazioni
- db – le simmetrie assiali (o riflessioni)
- dq – antitraslazioni (anche dette glissosimmetrie, glissoriflessioni o simmetrie con scorrimento)

cioè date 2 figure congruenti, ciascuna si trasforma nell'altra con 1 delle 5 isometrie dette.
Tutte le isometrie del piano euclideo (compresa l'identità) o sono 1 riflessione o sono composizioni di 2 o 3 riflessioni.
Escludento le identità, si hanno i seguenti 4 casi:

```
.................Conserva l'orientazione?
.................SI'..........NO
Ha punti...SI'...rotazione.....riflessione
fissi?.....NO....traslazione...antitraslazione
```

65.20 Legge congiunta

Definizioni. Per 2 variabili aleatorie X e Y discrete o continue, si chiama **funzione di ripartizione congiunta** di X e Y ovvero del *vettore aleatorio* (X, Y) – si immagini una coppia di dadi – la

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

e nel caso che X e Y siano continue (molto regolari come sempre) la derivata rispetto a y della derivata rispetto a x di quella

$$f_{X,Y}(x, y) := D_y D_x F_{X,Y}(x, y)$$

si chiama densità congiunta di X e Y.
Teorema 1. Dalla densità congiunta di X e Y continue si trova la densità di X, detta *densità marginale*, con

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy \quad \text{e analogamente per } Y.$$

Teorema 2.

$$\text{indipendenti } \Rightarrow F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y) \text{ (discrete o continue)}$$

$$\text{indipendenti } \Rightarrow f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) \text{ (continue).}$$

Teorema 3. Densità della somma di 2 variabili aleatorie continue di cui sia nota la densità congiunta:

$$f_{X+Y}(t) = \int_{-\infty}^{+\infty} f_{X,Y}(x, t-x)\, dx = \text{ se indip. } \int_{-\infty}^{+\infty} f_X(x) f_Y(t-x)\, dx.$$

Per esempio, si calolerà la densità della somma di 2 variabili aleatorie indipendenti uniformi su $[0, 1]$, siano esse U e V:

$$f_U(t) := f_V(t) := \begin{cases} 1 & se\ 0 \leq t \leq 1 \\ 0 & altrimenti \end{cases} .$$

65.21 Mnemonici per le cifre di e

Associando ordinatamente ad ogni parola di questa frase

> *"La loquela è vincente* $\leftarrow 2, 7, 1, 8$
> *ma talvolta è migliore il silenzio,* $\leftarrow 2, 8, 1, 8, 2, 8$
> *come disse Aristocle."* $\leftarrow 4, 5, 9$

i numeri delle lunghezze delle parole, si ottengono le prime cifre significative di e: $2, 7, 1, 8, 2, 8, 1, 8, 2, 8, 4, 5, 9$.
Fra le molte molte frasi analoghe, una famosa con le stesse cifre è

> *"Ai modesti o vanitosi*
> *ai violenti o timorosi*
> *do, cantando gaio ritmo,*
> *logaritmo."*

65.22 Mnemonici per le cifre di π: un *piem*

Associando ordinatamente ad ogni parola di queste frasi

> *"Non è dato a tutti ricordare il numero aureo del sommo filosofo Archimede. Certuni sostengon che si può ricordare tale numero, ma questi poi non recitano che un centone insensato."*

i numeri delle lunghezze delle parole, si ottengono le prime cifre significative di π: 3, 1, 4, 1, 5, 9, 2, 6, 5, eccetera. Esistono molte frasi analoghe, anche in altre lingue.

65.23 Moda di una variabile aleatoria continua

è il valore che ha la massima densità. Semprechè ce ne sia solo uno; se ce ne sono 2 entrambi si chiamano mode e la variabile aleatoria si chiama bimodale. Non considereremo il caso in cui ce ne siano più di 2.
La moda di una variabile aleatoria X si indica con $\mathrm{Mod}(X)$.

65.24 Operazione binaria (interna)

In sostanza, è una legge che a 2 elementi di un insieme associa un elemento di quello stesso insieme, come sono per esempio la somma e la moltiplicazione nei consueti insiemi numerici. Più formalmente, se E è un insieme, si chiama operazione binaria (interna) una funzione $* : E \times E \to E$. (è una funzione definita sul prodotto cartesiano $E \times E$). Il suo valore quando viene calcolata in 2 elementi x e y si indica con $x * y$.

65.25 Permutazioni complete

Vedi dismutazioni.

65.26 Poetica osservazione malevola sulla Statistica

[...] seconno le statistiche d'adesso
risurta che te tocca un pollo all'anno:
e, se nun entra nelle spese tue,
t'entra ne la statistica lo stesso
perchè c'è un antro che ne magna due. [Trilussa]

Naturalmente il classico problema qua evidenziato viene superato dalla considerazione degli indici di dispersione ovvero variabilità.

65.27 Precedenze algebriche e parentesi

Le parentesi indichino *precedenze* diverse da quelle *implicite*, grazie alle quali per esempio abbiamo scritto $(x + y)z = xz + yz$ e non $(x + y)z = (xz) + (yz)$, perchè il $\cdot$ ha la *precedenza* sul $+$.
In una sequenza di somme e sottrazioni, come $2 + 3 - 7 - 4 + 5$, faremo i calcoli nell'ordine in cui si presentano, e per il resto ecco le precedenze da osservare in mancanza di parentesi che le modifichino, dalla più debole:

$+$ *e* $-$

$\cdot$ anche se non trascritto: $x + 3y$ è $x + (3 \cdot y)$ e non $(x + 3) \cdot y$

^ anche se non scritto esplicitamente: $2x^3$ è $2(x^3)$ e non $(2x)^3$. Molti testi danno anche altre regole di precedenza, che permettono di stabilire se $2/x \cdot y$ è $(2/x) \cdot y$ o $2/(x \cdot y)$, ma in questa trattazione elementare rifiuteremo categoricamente scritture come $2/x \cdot y$, e scriveremo con le parentesi $(2/x) \cdot y$ oppure $2/(x \cdot y)$ a seconda di cosa intendiamo. **Sempre meglio parentesi che incertezze.** Si possono usare quante si vogliano coppie di parentesi "annidate" una dentro l'altra

$$(...(...(...)...)...)$$

ma spesso per chiarezza si usano anche parentesi quadre e graffe:

$$\{...[...(...)...]...\}$$

Rimarchiamo lo sconcertante fatto che il software Excel, di ampio uso nelle scienze applicate, contro le convenzioni di tutta la comunità matematica, per -3^2 dà 9 invece che -9, come se fosse da intendere $(-3)^2$.

65.28 Proprietà associativa

Sia $\star$ un'operazione binaria (interna) definita in un insieme E. Si dice che l'operazione $\star$ verifica la **proprietà associativa** se per ogni $x, y, z \in E$ è:

$$(x \star y) \star z = x \star (y \star z) \ .$$

Per esempio il $+$ e il $\cdot$ sono associativi in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

65.29 Proprietà commutativa

Sia $\star$ un'operazione binaria (interna) definita in un insieme E. Si dice che l'operazione $\star$ verifica la **proprietà commutativa** se per ogni $x, y \in E$ è

$$x \star y = y \star x \ .$$

Per esempio il $+$ e il $\cdot$ sono commutativi in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ ed $\mathbb{R}$.

65.30 Proprietà distributiva

Sia $\star$ un'operazione binaria (interna) definita in un insieme E. Se $\diamond$ è un'altra operazione binaria in E, si dice che l'operazione $\star$ verifica la **proprietà distributiva** rispetto all'operazione $\diamond$ se per ogni $x, y \in E$ è

$$x \star (y \diamond z) = (x \star y) \diamond (x \star z) \quad et \quad (y \diamond z) \star x = (y \star x) \diamond (z \star x)\ .$$

Per esempio il $\cdot$ è distributivo rispetto al $+$, in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

65.31 Rette parallele e rette sghembe

Due rette del piano con intersezione vuota le diremo *parallele*. Due rette dello spazio con intersezione vuota e *complanari* (cioè appartenenti ad uno stesso piano) le diremo *parallele*. Due rette dello spazio con intersezione vuota e (ma) non *complanari* le diremo *sghembe*.

65.32 Scarti dalla media per una v.a. normale

Una variabile aleatoria $X \sim N(\mu, \sigma^2)$ è in questa relazione con la sua standardizzazione $Y = \frac{X-\mu}{\sigma}$:

$$X = \sigma Y + \mu$$

e allora si ha

$$P(|X - \mu| \leq \delta) = P(|\sigma Y| \leq \delta) = P(\sigma |Y| \leq \delta) = P(|Y| \leq \frac{\delta}{\sigma}) =$$

e considerando l'evento complementare

$$= 1 - P(|Y| > \delta) =$$

per le proprietà del valore assoluto

$$= 1 - P\Big(Y < -\frac{\delta}{\sigma} \vee Y > \frac{\delta}{\sigma}\Big) =$$

e trattandosi di eventi disgiunti

$$= 1 - \left(P\left(Y < -\frac{\delta}{\sigma}\right) + P\left(Y > \frac{\delta}{\sigma}\right)\right) =$$

e per la simmetria della v.a. normale standard Y

$$= 1 - \left(P\left(Y > \frac{\delta}{\sigma}\right) + P\left(Y > \frac{\delta}{\sigma}\right)\right) =$$

$$= 1 - 2P\left(Y > \frac{\delta}{\sigma}\right) =$$

e passando all'evento complementare

$$= 1 - 2\left(1 - P\left(Y \leq \frac{\delta}{\sigma}\right)\right) = 2P\left(Y \leq \frac{\delta}{\sigma}\right) - 1 = 2\,\Phi\left(\frac{\delta}{\sigma}\right) - 1$$

cioè in conclusione

$$P(|X - \mu| \leq \delta) = 2\,\Phi\left(\frac{\delta}{\sigma}\right) - 1$$

e in particolare con con $\delta := m\,\sigma$

$$P(|X - \mu| \leq m\,\sigma) = 2\,\Phi(m) - 1$$

e coi classici valori (che si trovano per esempio sulle tavole)

$$\begin{aligned} \Phi(1) &\approx 0.8413 \\ \Phi(2) &\approx 0.9773 \\ \Phi(3) &\approx 0.9987 \end{aligned}$$

$$P(|X - \mu| \leq \sigma) = 2\,\Phi(1) - 1 \approx 0.68.26 \approx 68.3\%$$

$$P(|X - \mu| \leq 2\,\sigma) = 2\,\Phi(2) - 1 \approx 0.9546 \approx 95.5\%$$

$$P(|X - \mu| \leq 3\,\sigma) = 2\,\Phi(3) - 1 \approx 0.9974 \approx 99.7\%$$

e per avere con più precisione 95% si sostituisca $2\,\sigma$ con $1.96\,\sigma$:

$$P(|X - \mu| \leq 1.96\,\sigma) = 2\,\Phi(2) - 1 \approx 0.95 = 95\%$$

tutte e 4 valide per X normale di media μ e varianza σ^2.

65.33 Sconvolgimento

Vedi dismutazioni.

65.34 Simmetrico di un elemento

Sia $\star$ un'operazione binaria (interna) definita in un insieme E. Se $x \star y = y \star x = e$, allora x si dice **simmetrico** di y (rispetto all'operazione $\star$). (E ovviamente y è simmetrico di x). Per esempio -3 (che esiste da $\mathbb{Z}$ in poi) è simmetrico di 3 rispetto al $+$ e $\frac{1}{3}$ (che esiste da $\mathbb{Q}$ in poi) è simmetrico di 3 rispetto al $\cdot$.
Il simmetrico rispetto al $+$ si chiama *opposto* e il simmetrico rispetto al $\cdot$ si chiama *reciproco*.

65.35 Tassa degli stolti

Nel 2015 gli italiani hanno speso 55 miliardi per beni durevoli, come automobili ed elettrodomestici, e 88 miliardi in giochi d'azzardo, cioè mediamente circa 1470 euro a testa, diventati 1600 nel 2016; in parte persi, ovvio, mica organizzano i giochi solo per farci divertire; si dice che il lotto è una *tassa per gli stolti*. Bisogna comunque dire che la cinquina del lotto è un caso particolare, gli altri premi del lotto non sono così non equi; e altri giochi sono ancor meno non equi, dando premi più equilibratamente; e in alcuni conta anche l'abilità, per esempio la schedina del calcio. Inoltre c'è da dire che coi giochi d'azzardo lo Stato raccoglie cifre enormi senza suscitare avversione: rinunciando ai giochi lo Stato potrebbe dover esigere quei soldi forzatamente. Una cosa triste è che a pagare sono in larga misura – difficile da quantificare nei dettagli – i meno abbienti. Altra cosa triste è che per molti diventa una malattia, la ludopatia, ma questa può esistere anche indipendentemente dai giochi organizzati dallo Stato, giocando fra privati. Altra questione ancora, è che le sale slot offrono alla malavita un facile modo di riciclare il denaro sporco: incassano denaro pulito e per le vincite danno quello proveniente da illeciti,

che si disperde nei mille rivi delle spese personali della gente. Tutto questo, ovviamente, fermo restando che di per sè non è immorale giocare d'azzardo, nè fra amici nè con lo Stato, e può essere una cosa sana, se esercitata con l'opportuna misura. Certo, 1600 euro l'anno a testa, danno da pensare. Si potrebbero fare anche molte altre cose con tanti soldi.
Una questione molto interessante, invece, è che i giochi d'azzardo forniscono un campo di studi per il calcolo delle probabilità, anzi ne sono perfino una delle origini storiche.

65.36 Triangolo di Tartaglia e potenza del binomio

(Evitando complesse definizioni formali, in questa trattazione elementare) il *Triangolo di Tartaglia* (o *di Pascal*, in ogni caso sostanzialmente noto ben prima di quegli Autori) è la *scrittura*

$$\begin{array}{c} 1 \\ 1 \quad 1 \\ 1 \quad 2 \quad 1 \\ 1 \quad 3 \quad 3 \quad 1 \\ 1 \quad 4 \quad 6 \quad 4 \quad 1 \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{array}$$

variamente estesa con la Formula di Pascal

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

(e cioè ogni numero è la somma dei 2 soprastanti, a parte gli 1 ai margini) dove n è il numero di riga a partire dalla 0-esima e k è la posizione nella riga a partire dalla 0-esima.
Il Triangolo di Tartaglia consente il calcolo dei *coefficienti binomiali* $\binom{n}{k}$ mediante sole somme. Ciò è utile sia nel problema delle combinazioni semplici che nella *potenza del binomio*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \, b^{n-k} \qquad \text{ovvero}$$

$$= \binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \binom{n}{3} a^{n-3} b^3 + \ldots + \binom{n}{n} b^n$$

(si noti che nella seconda formula gli a e b nel secondo membro sono scambiati fra loro rispetto alla prima formula, come se là fosse scritto $a^{n-k} \, b^k$, ma è equivalente). I vari coefficienti dei monomi si trovano semplicemente sulla n-esima riga del Triangolo di Tartaglia (intendendo come 1-esima quella con due 1), per esempio (e si provi poi con la quinta potenza)

$$(a+b)^4 = a^4 + 4a^3 b + 6a^2 b^2 + 4ab^3 + b^4.$$

65.37 Trigonometria – la goniometria dei triangoli

Teorema dei seni.[3] In un triangolo di lati a, b, c, con angoli rispettivamente opposti α, β, γ

$$\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} = \frac{c}{\sin\gamma}.$$

Teorema del Coseno.[4] In un triangolo di lati a, b, c, con angolo γ opposto al lato di misura c

$$c^2 = a^2 + b^2 - a\,b\,\cos\gamma.$$

Ed esistono poi innumerevoli altre formule.

65.38 Valore assoluto: 4 definizioni equivalenti

Da $\mathbb{Z}$ in poi si può considerare il *valore assoluto* di un numero

$$|x| := \sqrt{x^2}$$

$$|x| := x\,\mathrm{sgn}(x)$$

$$|x| := \max\{x, -x\}$$

Definizione consigliabile:

$$|x| := \begin{cases} x & \textit{se } x \geq 0 \\ -x & \textit{se } x < 0 \end{cases}.$$

La soprastante parentesi graffa (grande) è *“selettiva dei casi”*, ma quel simbolo ha anche altri usi affini.
In $\mathbb{R}$ il valore assoluto è una *funzione elementare*, mentre nell'algebra dei numeri può considerarsi *operazione unaria*.

65.39 Varianza di un solo numero

Si noti che nella statistica descrittiva una lista composta da un solo numero ha (ovviamente) varianza 0 in base alle formule viste

$$\bar{x} := x_1 \qquad Var(x_1) := 0$$

[3] Nasir al-Din al-Tūsī, persiano, XIII sec. d.C. “He is often considered the creator of trigonometry as a mathematical discipline in its own right”, Wikipedia, l'enciclopedia libera.

[4] Euclide, greco, III sec. a.C., ben noto, e Jamshīd al-Kāshī, persiano, XIII sec. d.C. “Much of al-Kāshī's work was not brought to Europe, and much, even the extant work, remains unpublished in any form”, Wikipedia, l'enciclopedia libera.

mentre nella statistica inferenziale un campione con un solo dato non consente (ovviamente) alcuna stima sulla varianza della variabile aleatoria "retrostante", infatti nel calcolo si avrebbe $\frac{0}{0}$:

$$\bar{X} := X_1 \qquad S^2 := \frac{(X_1 - \bar{X})^2}{0} \, .$$

66 Paragrafi da rivedere

66.1 Carattere

Sinonimo usato dagli economisti per variabile.

66.2 Insieme potenza

Vedi insieme delle parti.

66.3 Triangolo

Si può definire il triangolo in molti modi, in particolare come poligono (però da definirsi) con 3 lati (comunque da definire), o intersezione di 3 (molto opportuni) semispazi (comunque da definire). In questa trattazione elementare, **definiremo *triangolo (non degenere)* l'involucro convesso di 3 punti distinti del piano euclideo**, supponendo noto il concetto di involucro convesso. (Comunque, diamo un'idea di cos'è: si immagini di piantare 3 pioli nei 3 punti, e di passare una corda intorno ad essi: la zona "entro" la corda è il triangolo, involucro convesso dei 3 punti. L'involucro convesso naturalmente esiste anche per più di 3 punti, e nello spazio euclideo si dovrà immaginare non un una corda ma un tessuto).
I 3 punti inizialmente considerati sono i *vertici*. I 3 segmenti che li congiungono sono i *lati*. (Il *segmento* che *congiunge* 2 punti è naturalmente l'involucro convesso dei 2 punti).
Il *perimetro* è la somma delle lunghezze dei lati.
L'*area* è la misura
Criteri di congruenza per i triangoli (da Wikipedia, l'enciclopedia libera):

I: Due triangoli sono congruenti se hanno congruenti due lati l'angolo tra essi compreso
II: Due triangoli sono congruenti se hanno un lato e due angoli a esso adiacenti rispettivamente congruenti
III: Due triangoli sono congruenti se hanno tutti i lati ordinatamente congruenti
IV: Due triangoli sono congruenti se hanno ordinatamente congruenti due angoli e un lato.

66.4 Punti, segmenti e involucro convesso

66.5 Modalità

Sinonimo usato dagli economisti per valore.

66.6 Rango

- Significato 1
Dato un campione $X_1, \dots, X_n$, si dice rango del campione la sua numerosità, cioè il numero n.
- Significato 2
Dato un campione $X_1, \dots, X_n$, il rango di un numero X fra quei valori è il suo indice (k) nel campione ordinato in senso crescente $X_{(1)}, \dots, X_{(k)} = X, \dots, X_{(n)}$. (Oppure partendo dall'indice 0).
Per esempio nel campione

$$20\ 100\ 9\ 8\ 20\ 30\ 35$$

il valore 9 ha rango 2:

$$8\ 9\ 20\ 20\ 30\ 35\ 100.$$

Per il 20 che è ripetuto 2 volte in posizione (3) e (4) una definizione ragionevole del rango è dire che ha rango 3.5.
Da Wikipedia, l'enciclopedia libera:
"In statistics, "ranking" refers to the data transformation in which numerical or ordinal values are replaced by their rank when the data are sorted. For example, the numerical data 3.4, 5.1, 2.6, 7.3 are observed, the ranks of these data items would be 2, 3, 1 and 4 respectively. For example, the ordinal data hot, cold, warm would be replaced by 3, 1, 2. In these examples, the ranks are assigned to values in ascending order. (In some other cases, descending ranks are used.) Ranks are related to the indexed list of order statistics, which consists of the original dataset rearranged into ascending order.
Some kinds of statistical tests employ calculations based on ranks"

66.7 Frequenza

La frequenza o più esattamente frequenza campionaria di x è il numero di volte che il valore x è fra gli $x_1, \dots, x_n$ di un campione.

66.8 Stimatori

Dato il campione aleatorio $X_1, \ldots, X_n$, una variabile aleatoria

$$T = t(X_1, \ldots, X_n)$$

(è una funzione delle variabili aleatorie $X_1, \ldots, X_n$) si chiama stimatore del parametro θ della distribuzione delle X_i; in teoria una funzione qualunque, ma nella pratica verrano considerate solo funzioni con proprietà apprezzabili; per esempio gli stimatori non distorti.
Questa definizione può essere generalizzata sostituendo il parametro θ con una sua qualunque funzione $\psi(\theta)$.

66.9 Stimatori non distorti

Uno stimatore si chiama stimatore non distorto del parametro θ della distribuzione delle X_i se

$$E(T) = \theta$$

(Cioè se la sua speranza matematica coincide con il parametro che si vuole stimare. Che potrebbe essere per esempio la media o la varianza.)
Questa definizione può essere generalizzata sostituendo il parametro θ con una sua qualunque funzione $\psi(\theta)$.
Esempi:

- La media campionaria
- La varianza campionaria

66.10 Media

Possiamo considerare 4 medie:
- La media aritmetica
- La media campionaria
- La media di una variabile aleatoria discreta
- La media di una variabile aleatoria continua

66.11 Media aritmetica

Dati gli n numeri $x_1, \ldots, x_n$ si definisce la media aritmetica

$$m = \frac{x_1 + \cdots + x_n}{n}\,.$$

(Dati n numeri produce 1 numero).

66.12 Media campionaria

Dato il campione aleatorio $X_1, \ldots, X_n$, se si sa a priori che sono dotate di speranza matematica finita, allora

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n}$$

è stimatore non distorto della media.

66.13 Media di una variabile aleatoria

Vedi Speranza matematica di una variabile aleatoria

66.14 Speranza matematica di una variabile aleatoria

Classicamente si considerano questi 2 casi:
- La speranza matematica di una variabile aleatoria discreta
- La speranza matematica di una variabile aleatoria continua

66.15 Media di una variabile aleatoria discreta

Vedi Speranza matematica di una variabile aleatoria discreta

66.16 Speranza matematica di una variabile aleatoria discreta

Data una variabile aleatoria discreta X

$$X = \begin{pmatrix} x_1 \ldots x_n \,[\ldots] \\ p_1 \ldots p_n \,[\ldots] \end{pmatrix}$$

si definisce la media o speranza matematica

$$E(X) = \sum_i x_i \, p_i$$

e questa è una somma finita oppure una serie ma in quest'ultimo caso si richiede non solo la convergenza ma anche l'assoluta convergenza. (Cioè $\sum_{i=1}^{\infty} |x_i| \, p_i < +\infty$).

66.17 Media di una variabile aleatoria continua

Vedi Speranza matematica di una variabile aleatoria continua

66.18 Mediana di un campione

Nella sostanza la mediana di un campione $x_1, \ldots, x_n$ è il valore x_i tale che una metà dei valori del campione sono minori e una metà sono maggiori, con qualche precisazione. Che ora vediamo.
Se n è dispari, $n = 2k + 1$, si ordinano i valori del campione in ordine crescente

$$x_{(1)}, \ldots, x_{(k)}, x_{(k+1)}, x_{(k+2)}, \ldots, x_{(n)}$$

e la mediana è $x_{(k+1)}$, che a sinistra (cioè minori o uguali) ha k valori, e a destra (cioè maggiori o uguali) ha k valori.
Per esempio la mediana di

20 100 9 8 20 30 35

è 20:

8 9 20 20 30 35 100.

Se n è pari, $n = 2k$, come mediana si prende la media dei 2 valori intermedi $x_{(k)}$ e $x_{(k+1)}$, dopo aver ordinato il campione come prima:

$$x_{(1)}, \ldots, x_{(k-1)}, x_{(k)}, x_{(k+1)}, x_{(k+2)}, \ldots, x_{(n)}$$

la mediana è

$$\frac{x_{(k)} + x_{(k+1)}}{2} \, .$$

Per esempio per

20 100 9 8 20 30 35 50

è 25:

8 9 20 20 30 35 50 100 .

66.19 Moda

è, molto approssimativamente, il valore più probabile, con varie precisazioni a seconda dei casi.
Classicamente si considerano queste 3 mode:

- La moda di n numeri
- La moda di una variabile aleatoria discreta
- La moda di una variabile aleatoria continua

66.20 Moda di n numeri

Dati n numeri $x_1, \ldots, x_n$ la moda è il valore più frequente cioè il punto di massimo dell'istogramma densità. Semprechè ci sia un solo punto di massimo: se ce ne sono 2 ci sono 2 mode e la distribuzione che ha quell'istogramma densità si chiama bimodale.
Per esempio per

27 28 26 28 30 25 26 27 26 26 28 27 27 25 27

la moda è 27 che ha frequenza 5, come si vede dall'istogramma:
25 25
26 26 26 26
27 27 27 27 27
28 28 28
|
30

66.21 Moda di una variabile aleatoria discreta

è il valore più probabile della variabile aleatoria. Semprechè ce ne sia solo uno; se ce ne sono 2 entrambi si chiamano mode e la

variabile aleatoria si chiama bimodale.
Cioè, nel caso semplice, data la variabile aleatoria discreta

$$X = \begin{pmatrix} [x_0]\, x_1 \dots x_n\, [\dots] \\ [p_0]\, p_1 \dots p_n\, [\dots] \end{pmatrix}$$

x_n è moda se p_n è il massimo dei numeri $p_1, ..., p_n$ [...].
La moda di una variabile aleatoria X si indica con Mod(X).
Per esempio si troverà che la moda della variabile aleatoria binomiale $B(3, \frac{2}{3})$ è 2 che ha probabilità $p_2 = \frac{4}{9}$.

66.22 Varianza

Possiamo considerare 5 varianze:
- La varianza di n numeri
- La varianza di una variabile aleatoria discreta
- La varianza di una variabile aleatoria continua
- I 2 stimatori non distorti della varianza da alcuni impropriamente chiamati varianza

66.23 Varianza di n numeri

Dati n numeri $x_1, \dots, x_n$, si definisce varianza il numero

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - media)^2 \,.$$

(Ovviamente *media* è la media aritmetica).

66.24 Varianza di una variabile aleatoria

Classicamente si considerano principalmente questi 2 casi:
- La varianza di una variabile aleatoria discreta
- La varianza di una variabile aleatoria continua

66.25 Stimatori non distorti della varianza

Si consideri una variabile aleatoria X discreta o continua, dotata di varianza incognita.
Dato il campione aleatorio $X_1, \dots, X_n$

$$\bar{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

è stimatore non distorto della varianza se è nota la media μ di X (caso relativamente raro nella pratica).
E poi c'è la varianza campionaria, che è stimatore non distorto della varianza se non è nota la media.

66.26 Varianza campionaria

Si consideri una variabile aleatoria X discreta o continua, dotata di varianza incognita.
Dato il campione aleatorio $X_1, \dots, X_n$, la varianza campionaria

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

è stimatore non distorto della varianza se non è nota la media di X. (Ovviamente $\bar{X}$ è lo stimatore non distorto della media).
(Se si usa n invece di $n-1$ nel denominatore, come spesso si fa nella pratica, per n grande il risultato non cambierà poi di molto).

66.27 Varianza di una variabile aleatoria discreta

Data una variabile aleatoria discreta X

$$X = \begin{pmatrix} x_1 \dots x_n \, [\dots] \\ p_1 \dots p_n \, [\dots] \end{pmatrix}$$

dotata di speranza matematica si definisce la varianza di X

$$\mathrm{Var}(X) = \sigma^2 = \sum_i (x_i - E(X))^2 \, p_i$$

che può essere una somma finita o una serie.

66.28 Teorema sulle varianze (delle variabili aleatorie discrete e delle variabili aleatorie continue)

La varianza è la differenza tra la media dei quadrati dei valori, e il quadrato della media dei valori:

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

66.29 Skewness

La skewness misura l'asimmetria della distribuzione ed è positiva o negativa a seconda se è "sbandata" a destra o a sinistra rispetto alla media.

66.30 Kurtosis

Vedi curtosi.

66.31 Kurtosi

Vedi curtosi.

66.32 Curtosi

(Variante grafica: Kurtosi; in Inglese *k*urtosis).
Si può dire che, in qualche modo, la curtosi misura di quanto è lontana una distribuzione dalla distribuzione normale.
La formula della curtosi dà il numero 3 per ogni distribuzione normale.
Talvolta purtroppo equivocata con il coefficiente di curtosi.

66.33 Coefficiente di curtosi

(Sinonimo di indice di curtosi, con purtroppo equivoci tra gli Autori; internazionalmente in Inglese *excess kurtosis*).
è

$$\text{curtosi} - 3$$

(Vedi curtosi).

66.34 Excess kurtosis

(Vedi Coefficiente di curtosi).

66.35 Test di normalità

Come si sa se un campione proviene da una variabile aleatoria normale? Con il normal plot a occhio, oppure con il test di Shapiro-Wilk, che invece è una precisa formula, con la sua significatività statistica.

66.36 Normal plot

Si calcolano la media campionaria $\bar{X}$ e la varianza campionaria S^2 dei dati, si standardizzano con la formula

$$Z_i = \frac{X_i - \bar{X}}{S} ,$$

e si graficano le coppie (X_i, Z_i). Se i punti sono più o meno allineati si conclude che (plausibilmente) la variabile è normale.
Per esempio col campione 1, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 7, 8 si riscontrerà un notevole allineamento apparente.

66.37 Deviazione standard

Sinonimo di scarto quadratico medio

66.38 Scarto quadratico medio

E' la radice quadrata σ della varianza σ^2 di una variabile aleatoria. (Che può essere discreta o continua).
Ha la stessa unità di misura dei valori della variabile aleatoria. (Invece la varianza non ha questa apprezzabile proprietà).
Un suo stimatore ragionevole (non approfondiamo qua se non distorto) è la radice quadrata S della varianza campionaria S^2.

66.39 Standardizzazione

Se X è una variabile aleatoria, la variabile aleatoria

$$Z = \frac{X - E(X)}{\sqrt{\mathrm{Var(X)}}}$$

si dice standarizzazione di X.
Ha media 0 e varianza 1:

$$E(Z) = 0 \qquad \mathrm{Var}(Z) = 1\,.$$

Se $X_1, \ldots, X_n$ è un campione aleatorio, si standardizza con la formula

$$Z_i = \frac{X_i - \bar{X}}{S}\,.$$

(Ovviamente $\bar{X}$ è la media campionaria e S è la radice quadrata della varianza campionaria S^2).
è

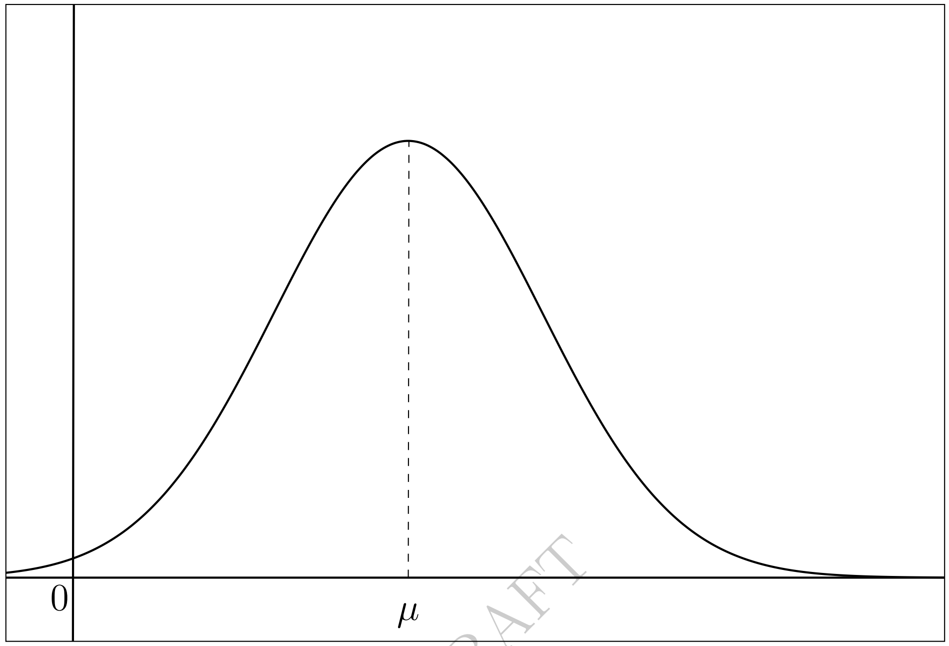
$$E(Z_i) = 0 \qquad \mathrm{Var}(Z_i) = 1\,.$$

66.40 Variabile aleatoria normale

è qualunque variabile aleatoria avente densità

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

con $\mu \in I\!R$ e $\sigma > 0$.
Ha speranza matematica μ e varianza σ^2.
(Ha anche moda μ e mediana μ).
è simmetrica (e allora ha skewness 0). (Ha curtosi 3).



è indicata con $N(\mu, \sigma^2)$.
La variabile aleatoria normale $N(0, 1)$ con media 0 e varianza 1 si chiama (variabile aleatoria) normale standard.

66.41 Teorema Limite Centrale

La variabile aleatoria

$$Y := X_1 + \cdots + X_n$$

somma di n variabili aleatorie indipendenti con la stessa distribuzione con varianza positiva, per n molto grande è approssimativamente normale.
Standardizzando, possiamo dire che se le variabili aleatorie $X_1, \ldots,$ X_n (discrete o continue) hanno la stessa distribuzione con varianza positiva (il che esclude le variabili aleatorie costanti), detta μ la loro media e σ^2 la varianza, la variabile aleatoria

$$S_n^* := \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

ha per n molto grande una distribuzione approssimativamente $N(0,1)$, cioè normale standard.
(Più precisamente si ha una *convergenza in legge*).

66.42 Normalizzazione di dati non normali

In caso che dei dati risultino non gaussiani (vedi test di normalità), si può provare a sostituirli con le loro radici quadrate, oppure logaritmi, eccetera. Se con la radice quadrata risultano normali si continua coi dati così modificati. Similmente coi logaritmi, eccetera.

66.43 Distribuzione t-Student

è una famiglia di distribuzioni continue, con un indice $n = 1, 2, \ldots$, la cui densità ha la forma di una campana tanto più schiacciata quando più basso è n. Per $n = +\infty$, e in pratica già per $n \geq 100$, si ottiene la densità normale standard. è la base del test di Student.

66.44 Chi quadrato

è una famiglia di distribuzioni continue, con un parametro $n = 1, 2, \ldots$ e un parametro $\alpha \in]0,1[$, indicate con $\chi^2(n)$. è la distribuzione somma di n quadrati della variabile aleatoria normale standard.
è la base del test del Chi quadrato.
La densità del Chi quadrato ha una forma vagamente a campana asimmetrica, che comincia in 0, mentre per ascisse negative la densità è nulla.
Valori approssimati di molti quantili del Chi quadrato si trovano su apposite tavole numeriche, che si trovano su internet cercando *Chi square table*.
Per $n > 30$ si usa la formula approssimata

$$\chi^2_\alpha(n) \approx \frac{1}{2}(\phi_\alpha + \sqrt{2n-1})^2$$

essendo ovviamente ϕ_α il quantile normale.

66.45 Distribuzione *F* di Fisher

è una famiglia di distribuzioni continue, con 2 parametri $k = 1, 2, \ldots$ e $m = [1], 2, 3, \ldots$ e un parametro p oppure $\alpha \in]0, 1[$.
è la base del test di Fisher.
Molti valori approssimati si trovano su apposite tavole numeriche.
Si trovano su internet cercando *Fisher table*.
è

$$F_\alpha(k, m) = \frac{1}{F_{1-\alpha}(k, m)} \, .$$

66.46 Statistica Inferenziale

Essenzialmente, la statistica inferenziale, fa affermazioni sulla popolazione a partire dai dati che costituiscono un campione (aleatorio).
Sostanzialmente fa affermazioni su qualche parametro incognito della distribuzione di una variabile aleatoria; per esempio un'affermazione sulla varianza di una variabile aleatoria che si sa essere normale (come: "è la varianza maggiore di 8?") a partire da una trentina di determinazioni (valori numerici) di quella variabile aleatoria.
La questione si chiarisce meglio andando all'apposito esempio.
Questioni fondamentali della Statistica Inferenziale sono:

- stima intervallare di parametri con gli intervalli di confidenza
- stima puntuale di parametri con gli stimatori
- i test statistici

66.47 Soglia del 5%

Il 5% è la soglia tradizionale di *significatività statistica*, nel senso che se stabiliamo coi calcoli che una certa cosa può essere successa per caso con probabilità del 2%, ed è effettivamente successa (per esempio glicemia diminuita) riteniamo che non è stato un caso (ma piuttosto il farmaco). Se invece quella variazione poteva succedere per caso con probabilità dell'8%, allora consideriamo che può ben essere stato un caso. Ottenere risultati con p-value ancora minore, è ancor più preferibile.

66.48 Estrazione bernoulliana ed esaustiva

Supponiamo di avere una popolazione di 60 milioni di persone con reddito medio μ e scarto quadratico medio σ. Estraggo a caso un campione di 20 persone, chiedo i loro redditi, e faccio la media m_1. Rimetto le 20 persone fra le 60 milioni. Estraggo a caso un campione di 20 persone (e qualcuna potrebbe essere quella di prima), chiedo i loro reddito e faccio la media m_2; e così via. Questa è l'estrazione bernoulliana. I numeri $m_1, m_2, \dots$ sono determinazioni di una variabile aleatoria M che ha la stessa media μ della popolazione dei 60 milioni, e varianza $\frac{\sigma^2}{20}$. Con l'estrazione esaustiva, in cui i campioni vengono via via eliminati dalla popolazione la varianza è un po' minore ma non useremo questo metodo. Se invece delle medie faccio le mediane dei campioni, ottengo delle determinazioni $n_1, n_2, \dots$ di una variabile aleatoria N che ha la stessa media μ ma una varianza più grande e allora in genere se estraiamo un solo campione di 20 persone conviene fare la media e non la mediana dei loro redditi per stimare il reddito medio dei 60 milioni di persone.

66.49 Statistica medica

è una disciplina scientifica, sottocapitolo della Statistica, atta ad analizzare quantitativamente fenomeni della medicina.
Fra i suoi ambiti principali:

• epidemiologia, per esempio lo studio del tasso di mortalità infantile.

• ambito clinico sperimentale: efficacia e sicurezza dei farmaci, delle terapie, efficienza delle diagnosi... Si veda un esempio apposito.

66.50 Atti preliminari all'esecuzione di una ricerca statistica

Consideriamo questi 3 atti preliminari all'esecuzione di una ricerca statistica

• Scelta della variabile, o delle variabili. La glicemia nell'esempio altrove considerato. (Si noti la drammaticità di questo passaggio: un farmaco che riducesse del 10% la glicemia ma nel contempo abbreviasse di 5 anni l'aspettativa di vita residua – cosa che non emergerà certo da uno studio clinico sperimentale su una ventina di soggetti – risulterà *efficace*).

• Determinazione della popolazione di riferimento e del campione. Se vogliamo ottenere un risultato valido per l'Italia – per esempio – dovremmo scegliere un campione *a caso* in tutta Italia. Se prendiamo una ventina di persone tutte in una stessa città, cosa immensamente più facile nella pratica (si consideri per esempio una ricerca fatta in un certo ospedale) i risultati della ricerca si applicheranno solo a quella città. Certo è ben possibile che senza grave pregiudizio si estenda a tutta Italia, ma non è detto e ci vuole molta cautela. Grande insistenza deve essere posta sulla casualità del campione. Classico è il caso storico di molti decenni fa delle interviste telefoniche per prevedere l'esito di una votazione (negli Stati Uniti). Il campione era sì casuale ma solo fra le persone presenti nell'elenco telefonico: mediamente più ricche della media, visto che avevano un telefono quando ancora pochi ce l'avevano. Questo bias distrusse la previsione.

• Con apposite formule si determina la dimensione del campione a seconda della precisione (sicurezza, affidafilità) con cui si vogliono ottenere i risultati. (Vedi in particolare la soglia del 5%) Con opportuni sorteggi casuali si estrae il campione; oppure in altri casi lo si aspetta, per esempio in una sezione specifica di ospedale (cardiologia, reumatologia...).

66.51 Z scores

Alcuni Autori invece dei quantili normali usano gli equivalenti Z-scores:

$$Z_\alpha = \phi_{1-\alpha} .$$

Per esempio $Z_{0.025} = \phi_{0.975} \approx 1.96$.

66.52 Intervallo di confidenza

Si abbia una variabile aleatoria la cui distribuzione (in pratica la densità) dipende da un parametro θ incognito. Per esempio una variabile aleatoria normale di media nulla e varianza $\theta = \sigma^2$. Il parametro θ può essere anche vettoriale, per esempio per una variabile aleatoria normale con incognite media e varianza: $\theta = (\mu, \sigma^2)$. Per esempio la variabile aleatoria X potrebbe essere il reddito dei italiani, supponendo per esso in via molto semplificata una distribuzione normale $N(\mu, \sigma^2)$. Se stabiliamo il reddito di 100 italiani a caso, la media campionaria $\bar{X}_{100}$ ci dà immediatamente una stima puntuale $\hat{\mu}$ del reddito medio μ degli italiani. Si noti però che questo singolo valore numerico può venire da 10 o 100 o 1000 o qualsiasi numero n di rilevazioni, mentre è ovvio che da più rilevazioni si ottiene un grado di certezza maggiore. Gli intervalli di confidenza sono il modo in cui questa maggiore certezza viene tenuta in considerazione. Per esempio per X normale $N(\mu, \sigma^2)$ l'intervallo di confidenza per μ supponendo nota σ^2 è

$$[\bar{X}_{100} - \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}, \bar{X}_{100} + \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}] \quad \textit{ovvero} \quad \bar{X}_{100} \pm \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}$$

ovvero in termine di Z-scores

$$\bar{X}_{100} \pm \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} .$$

Si osservi che al crescere della numerosità n del campione l'ampiezza dell'intervallo di confidenza si riduce intorno al stimatore puntuale. Anche una eventuale varianza ridotta nella popolazione

partecipa a rendere piccolo l'intervallo. Anche un valore grande di α riduce l'ampiezza dell'intervallo.
Più in generale per una variabile aleatoria con una densità dipendente da un parametro incognito θ eventualmente vettoriale un intervallo di fiducia ovvero confidenza per θ di livello α ovvero secondo altri Autori $1-\alpha$ è un intervallo aleatorio dipendente da 2 estremi aleatori $[t_1(X_1,\ldots,X_n), t_2(X_1,\ldots,X_n)]$ (essendo t_1 e t_2 due funzioni) tale che qualunque sia θ l'intervallo aleatorio detto contiene θ con probabilità $\geq 1-\alpha$.
Tipicamente si usa la soglia $\alpha = 5\%$.