

STATISTICA MATEMATICA

Ottavio Caligaris - Pietro Oliva

CAPITOLO 1

FUNZIONI DI DUE VARIABILI

I modelli matematici spesso devono tenere conto di molti parametri e per questa ragione non è sufficiente considerare funzioni di una sola variabile reale; spesso anzi il numero di parametri in gioco è molto alto e quindi bisogna ricorrere all'uso di funzioni di molte variabili reali.

Dal punto di vista concettuale non c'è grande differenza tra lo studio di una funzione di 2, 3 o 100 variabili reali, ma la differenza tra lo studio di una funzione di 1 variabile reale ed una funzione di 2 variabili reali è grande e va considerata attentamente.

Sviluppiamo pertanto lo studio di una funzione di 2 variabili reali per introdurre gli strumenti necessari al trattamento delle funzioni di più variabili reali a valori reali.

DEFINIZIONE 1.1. *Diciamo che è data una funzione di due variabili reali se sono assegnati un sottoinsieme $D \subset \mathbb{R}^2$ ed una corrispondenza f che ad ogni elemento $P = (x, y) \in D$ associa uno ed un solo elemento $z \in \mathbb{R}$.*

Diciamo che D è il dominio della funzione e denotiamo con

$$z = f(x, y) = f(P)$$

il corrispondente di $P = (x, y)$ secondo la legge assegnata f ; scriviamo anche

$$P = (x, y) \mapsto z = f(x, y) = f(P)$$

Chiamiamo rango di f l'insieme

$$R(f) = \{z \in \mathbb{R} : \exists (x, y) \in D, z = f(x, y)\}$$

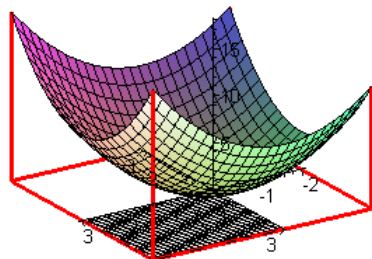
Chiamiamo grafico di f l'insieme

$$G(f) = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, z = f(x, y)\}$$

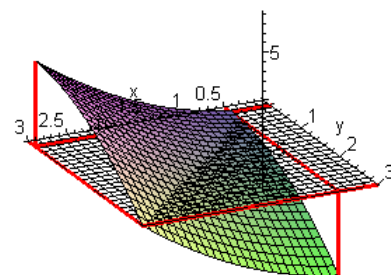
Osservazione. Il grafico di una funzione di 2 variabili è pertanto un sottoinsieme di $\mathbb{R}^3$ che descrive qualcosa che è immediato identificare come una superficie nello spazio. $\square$

Restrizione e composizione di funzioni sono definite come nel caso reale e parimenti simile è la definizione di iniettività, surgettività, bigettività.

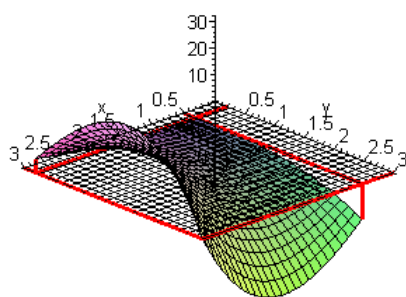
Per avere un'idea del comportamento della funzione sarebbe comodo poter disporre del suo grafico, che nel caso di funzioni di 2 variabili si rappresenta in uno spazio a 3 dimensioni $\mathbb{R}^3$; dobbiamo però tenere presente che:



1.1.1. Grafico 1



1.1.2. Grafico2



1.1.3. Grafico3

FIGURA 1.1. Grafici di funzioni di due variabili

- (1) Non è possibile rappresentare il grafico di funzioni che dipendano da 3 o più variabili
- (2) La rappresentazione in $\mathbb{R}^3$ di una funzione di due variabili passa attraverso tecniche di prospettiva.
- (3) La proprietà che risulta di maggiore interesse per tracciare il grafico qualitativo di una funzione di 1 variabile è la crescita o la decrescenza, che per le funzioni di 2 o più variabili non può più essere considerata dal momento che il dominio $\mathbb{R}^2$ (o $\mathbb{R}^n$) non ammette un ordine completo.

Non sarà pertanto semplice disegnare il grafico qualitativo di una funzione di 2 variabili e per farci un'idea del suo andamento dovremo ricorrere a rappresentazioni nel piano.

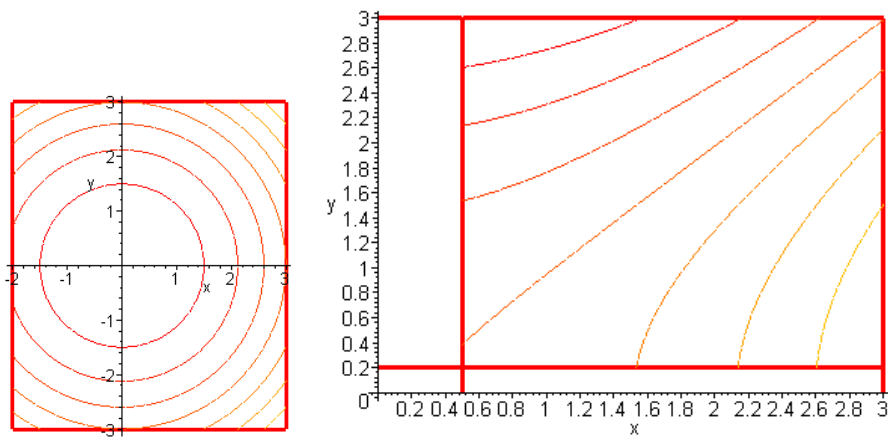
Un modo efficace di rappresentare una superficie è disegnare nel piano (x, y) le curve di livello della funzione.

DEFINIZIONE 1.2. Se $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ chiamiamo curve od insiemi di livello di f di altezza c gli insiemi

$$L_c = \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$$

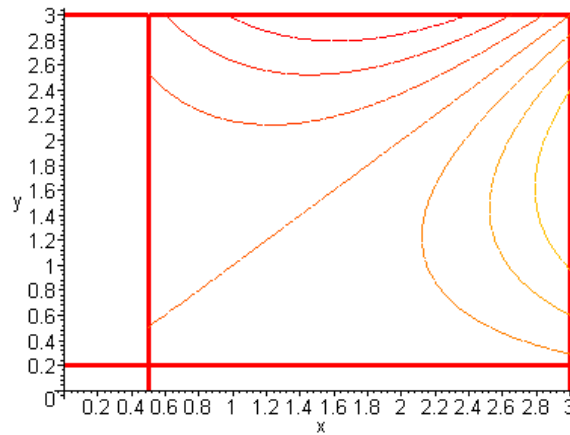
Le curve di livello di f consentono, in pratica, di rappresentare una mappa della superficie in esame. Esse definiscono i punti in cui la superficie assume quota costante uguale a c e, se le quote c sono scelte ad intervalli regolari, permettono di individuare le zone in cui la superficie è più ripida (le curve di livello sono più ravvicinate).

Le superfici prese in considerazione nella figura 1.1 hanno le curve di livello mostrate nella figura 1.2



1.2.1. Grafico 1

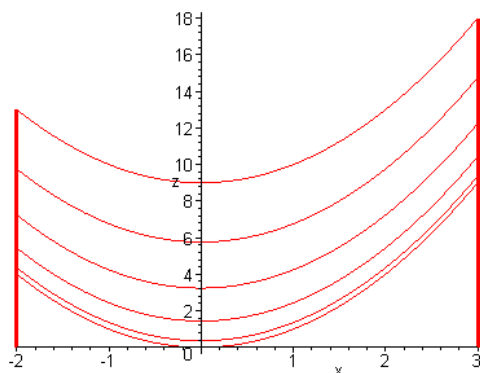
1.2.2. Grafico2



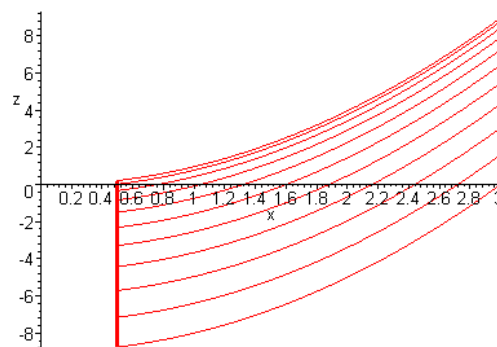
1.2.3. Grafico3

FIGURA 1.2. Curve di livello delle superfici in figura 1.1

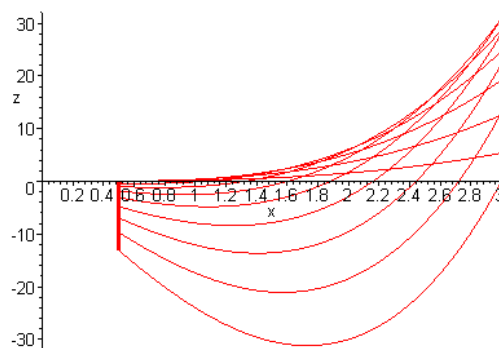
Per farci un'idea del grafico possiamo anche considerare l'andamento delle funzioni di x che si ottengono considerando fissati i valori di y ; chiamiamo questi grafici sezioni lungo l'asse x , si veda figura 1.3, e delle funzioni di y che si ottengono considerando fissati i valori di x ; chiamiamo questi grafici sezioni lungo l'asse y , si veda figura 1.4.



1.3.1. Grafico 1



1.3.2. Grafico 2



1.3.3. Grafico 3

FIGURA 1.3. Sezioni, per y fissato, dei grafici di figura 1.1

Come per le funzioni di una variabile è importante studiare la continuità e la derivabilità di una funzione di 2 o più variabili. Ovviamente per poter considerare la continuità è necessario conoscere la definizione di limite e ancora prima la definizione di intorno e la struttura dello spazio $\mathbb{R}^2$ in cui stiamo lavorando.

1. La struttura di $\mathbb{R}^2$.

Indichiamo con $\mathbb{R}^2$ lo spazio vettoriale costituito dalla coppie ordinate di numeri reali; in altre parole

$$P \in \mathbb{R}^2 \Leftrightarrow P = (x, y) \quad x, y \in \mathbb{R}$$

In $\mathbb{R}^2$ si definiscono le operazioni di somma e di prodotto per uno scalare mediante le

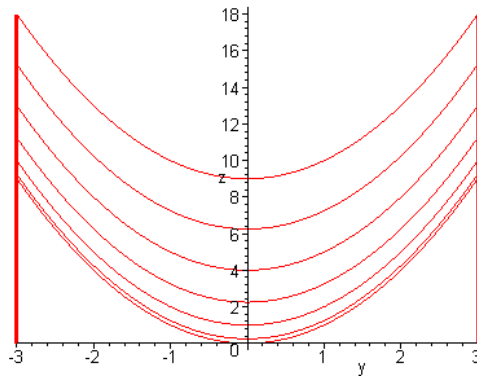
$$P_1 + P_2 = (x_1 + x_2, y_1 + y_2)$$

e, se $\alpha \in \mathbb{R}$,

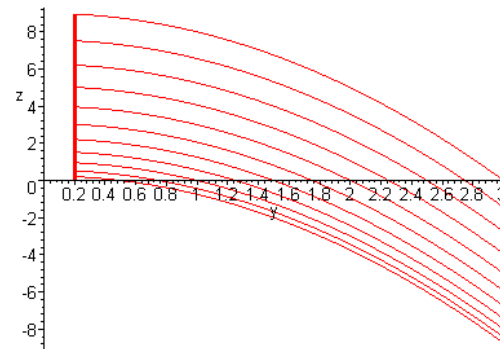
$$\alpha P = (\alpha x, \alpha y)$$

L'insieme dei vettori

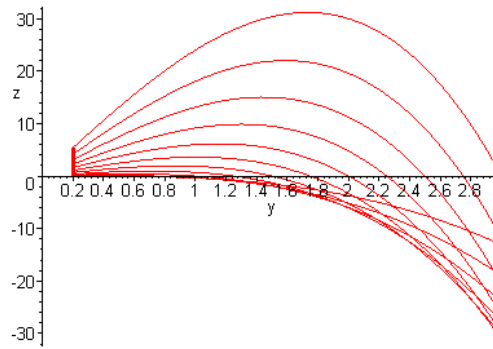
$$e_1 = (1, 0) \quad , \quad e_2 = (0, 1)$$



1.4.1. Grafico 1



1.4.2. Grafico 2



1.4.3. Grafico 3

FIGURA 1.4. Sezioni, per x fissato, dei grafici di figura 1.1

costituisce una base di $\mathbb{R}^2$; si avrà pertanto che, se $P \in \mathbb{R}^2$,

$$P = xe_1 + ye_2 = x(1, 0) + y(0, 1) = (x, y)$$

DEFINIZIONE 1.3. Si definisce norma in $\mathbb{R}^2$ una funzione che si indica con

$$\|\cdot\| : \mathbb{R}^2 \rightarrow \mathbb{R}$$

che verifica le seguenti proprietà:

- $\|P\| \geq 0 \quad \forall P \in \mathbb{R}^2$
- $\|P\| = 0 \Leftrightarrow P = 0$
- $\|\alpha P\| = |\alpha| \|P\| \quad \forall \alpha \in \mathbb{R}, \forall P \in \mathbb{R}^2$
- $\|P + Q\| \leq \|P\| + \|Q\| \quad \forall P, Q \in \mathbb{R}^2$

Si definisce prodotto scalare in $\mathbb{R}^2$ una funzione

$$\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$$

tale che

- $\langle P, P \rangle \geq 0 \quad \forall P \in \mathbb{R}^2$
- $\langle P, Q \rangle = \langle Q, P \rangle \quad \forall P, Q \in \mathbb{R}^2$
- $\langle P, P \rangle = 0 \Leftrightarrow P = 0$

$$\bullet \langle \alpha P + \beta Q, R \rangle = \alpha \langle P, R \rangle + \beta \langle Q, R \rangle \quad \forall P, Q, R \in \mathbb{R}^2, \forall \alpha, \beta \in \mathbb{R}.$$

Un esempio notevole di norma in $\mathbb{R}^2$ è

$$\|P\| = \sqrt{x^2 + y^2}$$

La norma di P indica la distanza di P dall'origine $O = (0, 0)$; se $P = (x, y)$, $P_0 = (x_0, y_0) \in \mathbb{R}^2$

$$\|P - P_0\|$$

indica la distanza tra i punti P e P_0 .

Un esempio notevole di prodotto scalare in $\mathbb{R}^2$ è definito da

$$\langle P_1, P_2 \rangle = x_1 x_2 + y_1 y_2$$

Se $\rho > 0$ chiamiamo intorno del punto $P_0 = (x_0, y_0)$, l'insieme

$$S(P_0, \rho) = \{P \in \mathbb{R}^2 : \|P - P_0\| < \rho\}$$

$S(P_0, \rho)$ è la sfera di centro P_0 e raggio ρ .

Definiamo inoltre intorno di ∞ il complementare di ogni sfera centrata nell'origine.

$$S(\infty, \rho) = \{P \in \mathbb{R}^2 : \|P\| > \rho\}$$

Diciamo che due vettori $P, Q \in \mathbb{R}^2$ sono ortogonali se $\langle P, Q \rangle = 0$.

Diciamo che sono paralleli se esiste $\lambda \in \mathbb{R}$ tale che $P = \lambda Q$.

Altri esempi di norme in $\mathbb{R}^2$ sono i seguenti

$$\|P\|_k = (|x|^k + |y|^k)^{1/k} \quad k \geq 1$$

$$\|P\|_\infty = \max\{|x|, |y|\}$$

Norme euclidea e prodotto scalare sono legati dalla seguente

Disuguaglianza di Schwarz

Per $P, Q \in \mathbb{R}^2$ si ha

$$|\langle P, Q \rangle| \leq \|P\| \|Q\|$$

La disuguaglianza di Schwarz può essere dedotta osservando che, per ogni $t \in \mathbb{R}$

$$0 \leq \|xP + tQ\|^2 = \langle P + tQ, P + tQ \rangle = t^2 \|Q\|^2 + 2t \langle P, Q \rangle + \|P\|^2$$

Ciò implica infatti che

$$\langle P, Q \rangle^2 - \|P\|^2 \|Q\|^2 \leq 0$$

Dalla disuguaglianza di Schwarz possiamo anche ricavare la disuguaglianza triangolare; infatti

$$\|P + Q\|^2 = \|P\|^2 + \|Q\|^2 + 2\langle P, Q \rangle \leq \|P\|^2 + \|Q\|^2 + 2\|P\| \|Q\|.$$

Osserviamo infine che

$$|\langle P, Q \rangle| = \|P\| \|Q\|$$

se e solo se esiste $t \in \mathbb{R}$ tale che $P + tQ = 0$, ovvero P e Q sono paralleli.

Da quanto detto si può dedurre che

$$\|P\| = \sup\{\langle P, Q \rangle : \|Q\| \leq 1\} = \max\{|\langle P, Q \rangle| : \|Q\| \leq 1\}$$

2. Limiti e continuità per le funzioni di 2 variabili.

DEFINIZIONE 1.4. Sia $f : A \rightarrow \mathbb{R}$, $A \subset \mathbb{R}^2$ e sia P_0 un punto tale che ogni intorno di P_0 abbia intersezione non vuota con A (chiamiamo P_0 punto di accumulazione per A); diciamo che

$$\lim_{P \rightarrow P_0} f(P) = \ell$$

se $\forall \varepsilon > 0$ esiste $\delta(\varepsilon) > 0$ tale che per $P \in S(P_0, \delta(\varepsilon)) \cap A$, $P \neq P_0$ si ha

$$f(x) \in I(\ell, \varepsilon)$$

È possibile verificare che

- (1) ogni funzione che ammette limite finito è localmente limitata;
- (2) il limite di una funzione, se esiste, è unico;
- (3) vale il teorema della permanenza del segno;
- (4) il limite di una somma è uguale alla somma dei limiti, se questi esistono finiti;
- (5) il limite del prodotto di due funzioni è uguale al prodotto dei limiti, se questi esistono finiti;
- (6) il limite del reciproco di una funzione è uguale al reciproco del limite della funzione stessa, se non è nullo
- (7) valgono i risultati sul confronto dei limiti, in analogia a quanto già visto per le funzioni di una variabile
- (8) il limite di una funzione può essere caratterizzato per successioni
- (9) il limite di una funzione composta si calcola seguendo quanto fatto per le funzioni di una variabile

DEFINIZIONE 1.5. Diciamo che f è una funzione continua in P_0 se $\forall \varepsilon > 0$ esiste $\delta(\varepsilon) > 0$ tale che se $x \in A$, $\|P - P_0\| < \delta(\varepsilon)$ si ha

$$\|f(P) - f(P_0)\| < \varepsilon$$

Nel caso in cui $P_0 \in A$, sia un punto di accumulazione per A la condizione sopra espressa è equivalente alla

$$\lim_{P \rightarrow P_0} f(P) = f(P_0)$$

Ovviamente f si dice continua in A se è continua in ogni punto di A

Come nel caso delle funzioni reali di una variabile reale si prova che:

- (1) la somma di funzioni continue è continua;
- (2) il prodotto di una funzione a valori vettoriali per una funzione a valori scalari, entrambe continue, è continuo;
- (3) il reciproco di una funzione continua è continuo dove ha senso definirlo;

- (4) il prodotto scalare di due funzioni a valori vettoriali continue, è continuo;
- (5) vale la caratterizzazione della continuità per successioni
- (6) la composta di funzioni continue è una funzione continua.

La conoscenza della continuità delle funzioni elementari e le regole precedentemente enunciate permettono di stabilire in modo semplice la continuità in un gran numero di casi: ad esempio, poichè $(x, y) \mapsto x^2$ e $(x, y) \mapsto y^2$ sono continue possiamo anche affermare che

$$(x, y) \mapsto x^2 + y^2$$

è continua, se poi ricordiamo che l'esponenziale è continua avremo anche che

$$(x, y) \mapsto e^{x^2+y^2}$$

è continua.

Come per le funzioni continue di una variabile si possono provare importanti teoremi, tra i quali ricordiamo i seguenti risultati.

TEOREMA 1.1. - di Weierstraß - *Se f è una funzione continua su un insieme A che sia chiuso (contiene i limiti di ogni successione convergente di suoi punti) e limitato (è contenuto in una sfera) allora f ammette massimo e minimo assoluto su A*

TEOREMA 1.2. - degli zeri - *Se f è una funzione continua su un insieme A connesso (cioè, in parole semplici, fatto di un solo pezzo) e se esistono due punti $P_+, P_- \in A$ tali che*

$$f(P_+) > 0, \quad f(P_-) < 0$$

allora esiste un punto $P_0 \in A$ tale che

$$f(P_0) = 0$$

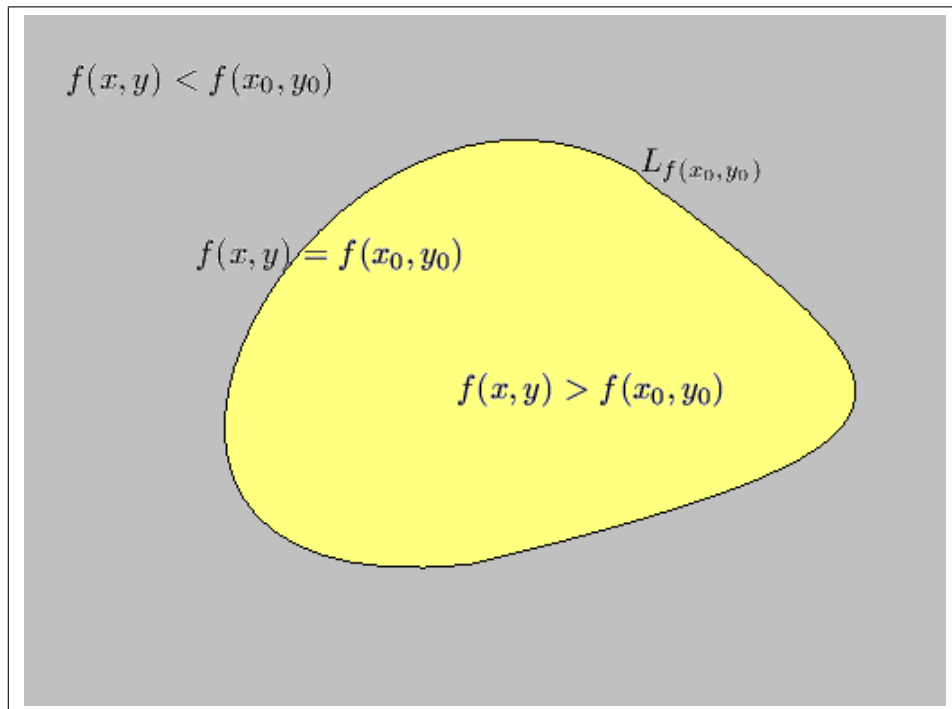
Un semplice ragionamento assicura, utilizzando il teorema degli zeri, che se una curva di livello di f

$$L_c = \{(x, y) \in U : f(x, y) = c\} = \{(x, y) \in \mathbb{R}^2 : x \in I, y = \varphi(x)\}$$

divide il piano in due parti connesse allora $f(x, y) > 0$ in una delle due parti e $f(x, y) < 0$ nell'altra.

Se infatti in una parte connessa ci fossero due punti P_+, P_- tali che

$$f(P_+) > 0, \quad f(P_-) < 0$$

FIGURA 1.5. Curve di livello e segno di f

esisterebbe in quella parte P_0 tale che

$$f(P_0) = 0$$

ma in quella parte si può solo avere $f(P) > 0$ oppure $f(P) < 0$.

3. Derivabilità e differenziabilità per funzioni di 2 variabili.

Come per le funzioni di 1 variabile è necessario considerare il problema della approssimazione mediante funzioni lineari, cioè il problema della differenziazione.

È molto naturale porre la seguente definizione

DEFINIZIONE 1.6. Diciamo che f è derivabile parzialmente se le funzioni

$$\phi(x) = f(x, y) \quad \psi(y) = f(x, y)$$

sono derivabili.

Chiamiamo $\phi'(x) = f_x(x, y)$ derivata parziale rispetto ad x e $\psi'(y) = f_y(x, y)$ derivata parziale rispetto ad y ; definiamo inoltre gradiente di f e scriviamo $\nabla f(x, y)$ il vettore (punto di $\mathbb{R}^2$) definito da

$$\nabla f(x, y) = (f_x(x, y), f_y(x, y))$$

Di fatto in tal modo si opera derivando rispetto ad x (o ad y) con y (o x) fissati.

Va osservato che, pur essendo molto naturale, l'uso delle derivate parziali non consente, da solo, di ricavare informazioni utili sulla funzione in esame.

Si pensi ad esempio che la funzione

$$f(x, y) = \begin{cases} 1 & \text{se } xy = 0 \\ 0 & \text{se } xy \neq 0 \end{cases}$$

il cui grafico, si veda la figura ?? è costituito dal piano $z = 0$ privato degli assi x ed y e dalle due rette parallele agli assi x ed y poste a quota $z = 1$, non è continua in $(0, 0)$ pur avendo derivate parziali nulle in $(0, 0)$.

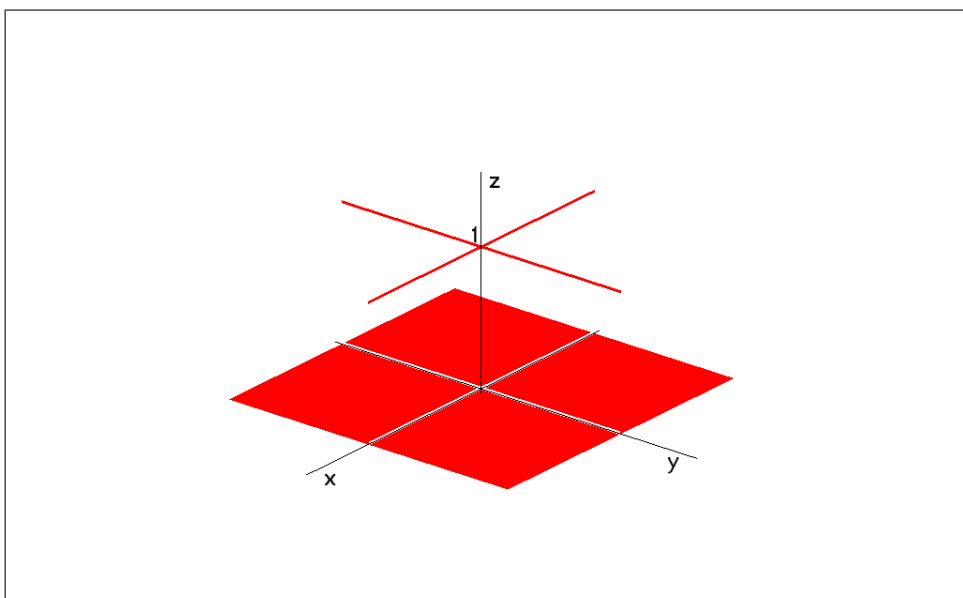


FIGURA 1.6. Il grafico di una funzione derivabile parzialmente, non continua.

Occorre quindi definire cosa si intende per differenziabile e per questo serve parlare di applicazioni lineari.

DEFINIZIONE 1.7. Si chiama *applicazione lineare in $\mathbb{R}^2$* una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che

$$f(\alpha P + \beta Q) = \alpha f(P) + \beta f(Q) \quad \forall P, Q \in \mathbb{R}^2, \quad \forall \alpha, \beta \in \mathbb{R}$$

L'insieme delle applicazioni lineari su $\mathbb{R}^2$ si chiama anche spazio duale di $\mathbb{R}^2$.

Ogni applicazione lineare in $\mathbb{R}^2$ si può identificare con un punto P^* di $\mathbb{R}^2$ mediante la seguente uguaglianza

$$f(P) = \langle P, P^* \rangle$$

In altre parole le applicazioni lineari su $\mathbb{R}^2$ sono tutte e sole le funzioni che si possono scrivere nella forma

$$f(P) = \langle P, P^* \rangle \quad \text{con} \quad P^* \in \mathbb{R}^2$$

È anche utile ricordare che per funzioni lineari possiamo provare che

Se f è una applicazione lineare su $\mathbb{R}^2$ allora

$$|f(P)| = |\langle P, P^* \rangle| \leq \|P\| \|P^*\|$$

Diciamo che $f \in \mathcal{C}^1(A)$ se f ammette derivate parziali continue in A .

DEFINIZIONE 1.8. Diciamo infine che f è differenziabile in P_0 se esiste $(\alpha, \beta) \in \mathbb{R}^2$ tale che

$$\lim_{P \rightarrow P_0} \frac{f(P) - (f(P_0) + \alpha(x - x_0) + \beta(y - y_0))}{\|P - P_0\|} = 0$$

Pertanto una funzione è differenziabile se

$$f(P) = f(P_0) + \alpha(x - x_0) + \beta(y - y_0) + \|P - P_0\| \omega(P - P_0)$$

dove ω è una funzione infinitesima per $P \rightarrow P_0$

$$\omega(P - P_0) = \frac{f(P) - (f(P_0) + \alpha(x - x_0) + \beta(y - y_0))}{\|P - P_0\|}$$

Questa proprietà si esprime dicendo che $f(P)$ si può approssimare con una funzione lineare affine

$$t(P) = f(P_0) + \alpha(x - x_0) + \beta(y - y_0)$$

a meno di un infinitesimo

$$\|P - P_0\| \omega(P - P_0)$$

di ordine superiore al primo rispetto alla distanza $\|P - P_0\|$.

La funzione $t(p)$ si definisce piano tangente al grafico di f nel punto P_0

Se f è differenziabile in P_0 allora f è anche derivabile parzialmente e si può verificare che risulta

$$\alpha = f_x(P_0) \quad \beta = f_y(P_0)$$

pertanto

Il piano tangente al grafico di una funzione f in P_0 è dato da

$$t(P) = f(P_0) + f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0)$$

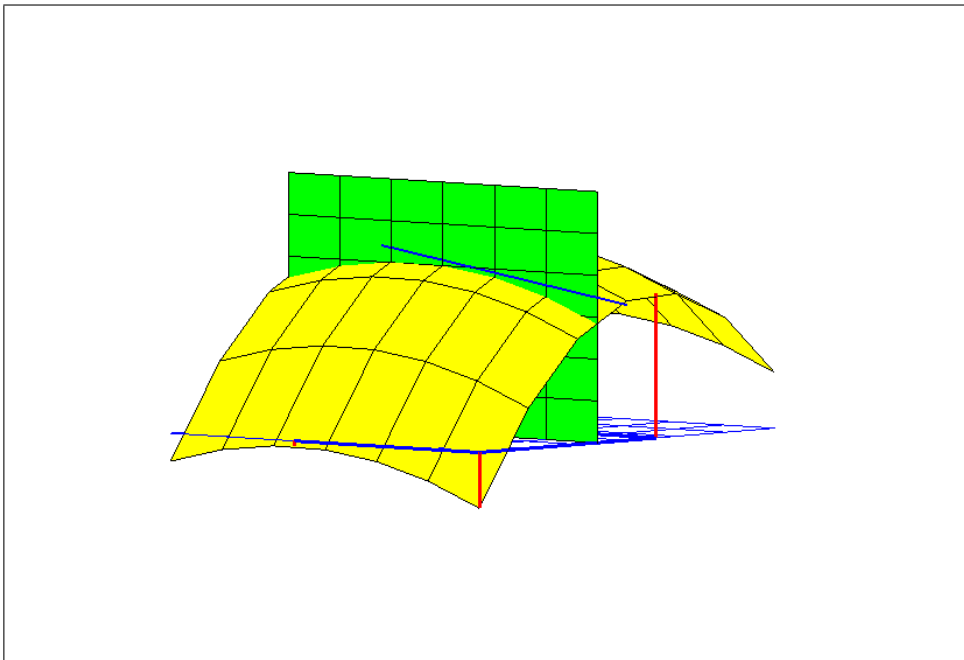


FIGURA 1.7. Definizione di derivata direzionale

DEFINIZIONE 1.9. Se $Q \in \mathbb{R}^2$, diciamo che f è derivabile in P_0 rispetto al vettore Q o che f ammette derivata in P_0 lungo la direzione Q se

$$\lim_{t \rightarrow 0^+} \frac{f(P_0 + tQ) - f(P_0)}{t}$$

esiste finito. In tal caso denotiamo il valore di tale limite con $f'(P_0, Q)$ e lo chiamiamo derivata direzionale di f in P_0 lungo la direzione Q .

Si può vedere che f è derivabile rispetto alla prima variabile se e solo se $f'(P_0, e_1)$ ed $f'(P_0, -e_1)$ esistono finiti e

$$f'(P_0, e_1) = -f'(P_0, -e_1)$$

Analogamente f è derivabile rispetto alla seconda variabile se e solo se $f'(P_0, e_2)$ ed $f'(P_0, -e_2)$ esistono finiti e

$$f'(P_0, e_2) = -f'(P_0, -e_2)$$

Si dimostra che

TEOREMA 1.3. *Se f è differenziabile in P_0 ; allora f è derivabile in P_0 lungo ogni direzione Q e si ha*

$$f'(P_0, Q) = \langle \nabla f(P_0), Q \rangle$$

È utile estendere alle funzioni di più variabili la regola di derivazione delle funzioni composte; ci limitiamo qui a considerare solo due casi particolari.

Siano

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad , \quad g : \mathbb{R} \rightarrow \mathbb{R}^2$$

$$\mathbb{R} \ni t \mapsto g(t) = (x(t), y(t)) \mapsto f(g(t)) = f(x(t), y(t)) \in \mathbb{R}^2$$

Se f e g sono differenziabili (non solo derivabili!) allora

$$\frac{d}{dt} f(g(t)) = f_x(x(t), y(t)) \dot{x}(t) + f_y(x(t), y(t)) \dot{y}(t)$$

Se viceversa consideriamo

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad , \quad g : \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$\mathbb{R}^2 \ni (x, y) \mapsto g(x, y) \mapsto f(g(x, y)) \in \mathbb{R}$$

e se f e g sono anche qui differenziabili avremo che

$$\frac{\partial f}{\partial x} = f'(g(x, y)) g_x(x, y)$$

$$\frac{\partial f}{\partial y} = f'(g(x, y)) g_y(x, y)$$

Abbiamo già visto che

se f è differenziabile in $P_0 \in \mathbb{R}^2$ e se Q è una direzione in $\mathbb{R}^2$, allora

$$f'(P_0, Q) = \langle \nabla f(P_0), Q \rangle = \|\nabla f(P_0)\| \|Q\| \cos \alpha$$

dove α è l'angolo formato dai vettori $\nabla f(P_0)$ e Q nel piano da essi individuato.

Ne possiamo dedurre che

la derivata direzionale è

- **massima quando $\cos \alpha = 1$ e cioè quando $\alpha = 0$ e $Q = \nabla f(P_0)$,**
- **nulla quando $\cos \alpha = 0$ e cioè quando $\alpha = \frac{\pi}{2}$ e $Q \perp \nabla f(P_0)$,**
- **minima quando $\cos \alpha = -1$ e cioè quando $\alpha = \pi$ e $Q = -\nabla f(P_0)$.**

Consideriamo ora una curva di livello di f

$$L_c = \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$$

e supponiamo che sia rappresentabile, almeno localmente, mediante il grafico di una funzione $y = \varphi(x)$. In termini un po' più precisi supponiamo che

$$L_c = \{(x, y) \in U : f(x, y) = c\} = \{(x, y) \in \mathbb{R}^2 : x \in I, y = \varphi(x)\}$$

o più semplicemente

$$f(x, y) = c \iff f(x, \varphi(x)) = c \iff y = \varphi(x)$$

Da $f(x, \varphi(x)) = c$, derivando e tenendo presenti le regole di derivazione delle funzioni composte, otteniamo che:

$$f_x(x, \varphi(x)) + f_y(x, \varphi(x))\varphi'(x) = 0$$

da cui

$$\langle \nabla f(x, \varphi(x)), (1, \varphi'(x)) \rangle = 0$$

e possiamo ricavare che

$$\nabla f(x, \varphi(x)) \perp (1, \varphi'(x))$$

D'altro canto la retta tangente τ al grafico di φ nel punto $P_0 = (x_0, y_0)$ è data da

$$y - y_0 = \varphi'(x_0)(x - x_0)$$

e si può scrivere nella forma

$$\langle (x - x_0, y - y_0), (\varphi'(x_0), -1) \rangle = 0$$

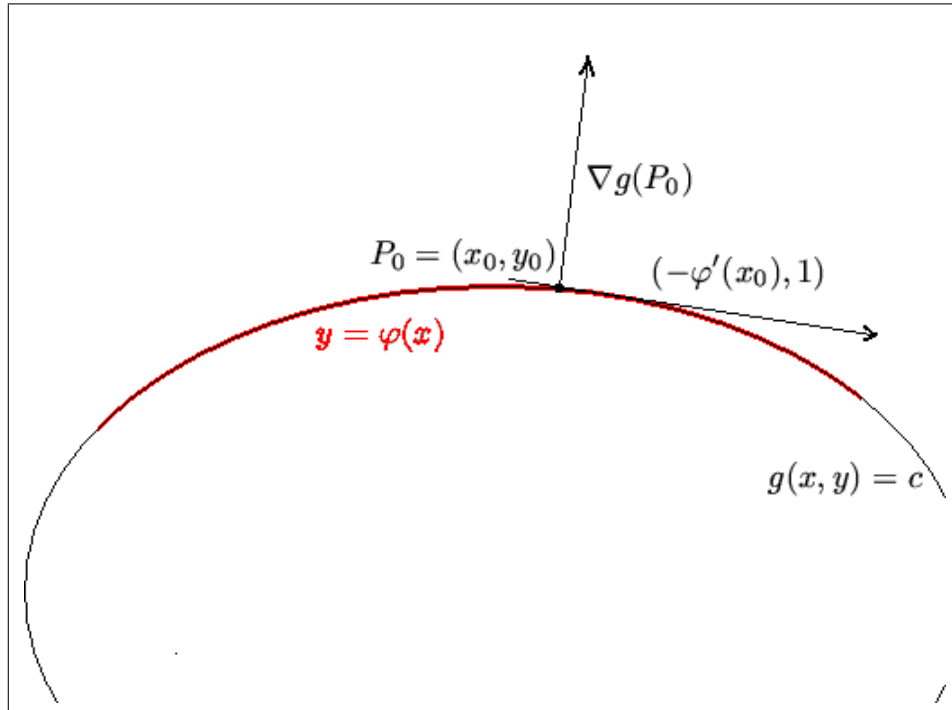


FIGURA 1.8. Curve di livello e gradiente

dalla quale risulta evidente che

$$(\varphi'(x_0), -1) \perp P \quad \forall P \in \tau$$

Se ora teniamo conto che, evidentemente,

$$\langle (\varphi'(x_0), -1), (1, \varphi'(x_0)) \rangle = 0$$

e quindi

$$(\varphi'(x_0), -1) \perp (1, \varphi'(x_0))$$

possiamo ricavare che

$$(1.1) \quad \nabla f(x, \varphi(x)) \perp P \quad \forall P \in \tau$$

Poichè τ è la retta tangente in P_0 al grafico della funzione φ che rappresenta vicino al punto P_0 (localmente in P_0) la curva di livello L_c , esprimeremo la 1.1 dicendo che

il gradiente di f , cioè il vettore $\nabla f(x, y)$, è ortogonale alle curve di livello di f ($L_c = \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$)

4. Derivate del secondo ordine: forma quadratica Hessiana.

Possiamo anche considerare le derivate seconde rispetto ad x due volte, ad y due volte, ad x e ad y , ad y e ad x ; chiamiamo queste derivate

$$f_{xx}(P_0) \quad f_{yy}(P_0) \quad f_{xy}(P_0) \quad f_{yx}(P_0)$$

Si può dimostrare che, nel caso in cui $f_{x,y}(P_0)$, o $f_{y,x}(P_0)$ sia continua allora (**teorema di Scharwz**)

$$\boxed{f_{x,y}(P_0) = f_{y,x}(P_0)}$$

Ciò si esprime dicendo che le derivate seconde miste sono uguali.

Chiamiamo matrice Hessiana la matrice i cui elementi sono le derivate seconde di f . Cioè

$$Hf(P_0) = \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix}$$

Nel caso in cui le derivate miste siano uguali, la matrice Hessiana è simmetrica.

Ad ogni matrice simmetrica, e quindi anche alla matrice Hessiana, possiamo associare un polinomio di secondo grado in 2 variabili (e.g. h, k) omogeneo che chiamiamo forma quadratica associata.

La forma quadratica Hessiana è, posto $R = \begin{pmatrix} h \\ k \end{pmatrix}$

$$Q(R) = Q(h, k) = \begin{pmatrix} h & k \end{pmatrix} \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} =$$

$$R^T Hf(P_0) R = f_{xx}(P_0)h^2 + 2f_{xy}(P_0)hk + f_{yy}(P_0)k^2$$

Diciamo che la forma quadratica Q è semidefinita positiva se

$$Q(h, k) = f_{xx}(P_0)h^2 + 2f_{xy}(P_0)hk + f_{yy}(P_0)k^2 \geq 0$$

per ogni $(h, k) \in \mathbb{R}^2$.

Diciamo che Q è definita positiva se

$$Q(h, k) = f_{xx}(P_0)h^2 + 2f_{xy}(P_0)hk + f_{yy}(P_0)k^2 > 0$$

per ogni $(h, k) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

Ovviamente per identificare una forma quadratica semidefinita o definita negativa è sufficiente cambiare il segno delle disuguaglianze.

Semplici considerazioni sul segno di un trinomio di secondo grado permettono di ottenere condizioni per studiare il carattere di una forma quadratica.

La forma quadratica Q è definita positiva se

$$\det \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix} > 0$$

e $f_{xx}(P_0) > 0$, oppure $f_{yy}(P_0) > 0$

Osservazione. Se

$$\det \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix} = f_{xx}(P_0)f_{yy}(P_0) - (f_{xy}(P_0))^2 > 0$$

allora

$$f_{xx}(P_0)f_{yy}(P_0) \geq (f_{xy}(P_0))^2 > 0$$

e quindi $f_{xx}(P_0)$ ed $f_{yy}(P_0)$ hanno lo stesso segno □

La forma quadratica Q è semidefinita positiva se

$$\det \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix} \geq 0$$

e $f_{xx}(P_0) \geq 0$, o equivalentemente $f_{yy}(P_0) \geq 0$

Si può inoltre dimostrare che

Se λ_1, λ_2 sono gli autovalori della matrice

$$\det \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix} \geq 0$$

allora, per la simmetria della matrice, essi sono reali ed inoltre

- **se λ_1, λ_2 sono entrambi positivi (negativi) la forma quadratica Q è definita positiva (negativa)**
- **se λ_1, λ_2 sono entrambi positivi (negativi) o nulli la forma quadratica Q è semidefinita positiva (negativa)**
- **se λ_1, λ_2 hanno segni discordi la forma quadratica Q è non definita**

Osservazione. Se

$$\det \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix} < 0$$

la forma quadratica può assumere sia valori positivi che negativi e quindi non è definita. □

5. Massimi e minimi per le funzioni di 2 variabili.

DEFINIZIONE 1.10. Diciamo che P_0 è un punto di minimo (massimo) relativo per f se esiste una sfera $S(P_0, \rho)$, $\rho > 0$, tale che

$$f(P) \geq f(P_0) \quad (f(P) \leq f(P_0))$$

per ogni $P \in S(P_0, \rho)$

Utilizzando tecniche che sfruttano i risultati noti per le funzioni di una variabile possiamo provare le seguenti condizioni necessarie per l'esistenza di un punto di minimo o massimo relativo.

TEOREMA 1.4. Se P_0 è un punto di minimo (massimo) relativo per f interno al suo dominio ed f è differenziabile in P_0 . Allora

- $\nabla f(x) = 0$;

se inoltre f ammette derivate seconde continue in P_0 ,

- $Hf(x)$ è semidefinita positiva (negativa).

Osservazione. Se $\nabla f(x) = 0$ e se $Hf(x)$ non è definito, allora P_0 non è né punto di massimo relativo, né punto di minimo relativo per f ; un punto siffatto viene solitamente indicato con il nome di 'punto sella'. $\square$

TEOREMA 1.5. Se $f \in C^2(A)$; e se P_0 è interno al suo dominio e se

- $\nabla f(P_0) = 0$
- $Hf(P_0)$ è definita positiva (negativa)

allora P_0 è punto di minimo (massimo) relativo per f .

Anche per le funzioni di due variabili si può definire e studiare la convessità:

**Sia $f : A \rightarrow \mathbb{R}$, e sia $A \subset \mathbb{R}^2$ convesso, cioè supponiamo che A contenga ogni segmento di retta i cui estremi siano contenuti in A ;
diciamo che f è convessa se**

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad \forall x, y \in A, \quad \forall \lambda \in (0, 1)$$

Inoltre f si dice strettamente convessa se vale la disuguaglianza stretta.

Osservazione. si può dimostrare che se f è convessa allora i suoi insiemi di livello L_c sono a loro volta convessi un insieme convesso $\square$

Inoltre possiamo anche dimostrare che

TEOREMA 1.6. Sia $f : A \rightarrow \mathbb{R}$ convessa, A aperto; allora

- f è continua in A
- $f'(P, Q)$ esiste $\forall P \in A, \forall Q \in \mathbb{R}^2$.

Come per le funzioni di una variabile la convessità si può caratterizzare utilizzando le derivate come si vede dall'enunciato del teorema seguente.

TEOREMA 1.7. *Sia $f : A \rightarrow \mathbb{R}$, $A \subset \mathbb{R}^2$ convesso, aperto; supponiamo inoltre $f \in C^2(A)$, allora sono condizioni equivalenti:*

- f è convessa

-

$$f(y) \geq f(P_0) + \langle \nabla f(P_0), P - P_0 \rangle \quad \forall P, P_0 \in A$$

- $Hf(P)$ è semidefinita positiva.

Inoltre

Ciascuna delle seguenti condizioni è sufficiente per la successiva:

- $Hf(P)$ è definita positiva $\forall P \in A$;
- $f(P) > f(P_0) + \langle \nabla f(P_0), P - P_0 \rangle \quad \forall P, P_0 \in A, P \neq P_0$;
- f è strettamente convessa.

Si può inoltre vedere che se f è strettamente convessa e se $f(P) \rightarrow +\infty$ per $P \rightarrow \infty$; allora esiste uno ed un solo punto $P_0 \in \mathbb{R}^n$ tale che

$$f(P_0) = \min\{f(P) : P \in \mathbb{R}^2\}$$

6. Massimi e minimi vincolati.

Le condizioni fin qui trovate per caratterizzare i punti di massimo e di minimo relativo sono utilizzabili soltanto nel caso in cui si cerchino massimi e minimi di f all'interno di un determinato insieme; nel caso in cui si vogliano cercare massimi e minimi su insiemi che contengano anche punti non interni, questi ultimi andranno considerati a parte esattamente come a parte debbono essere considerati gli estremi di un intervallo se si considerano funzioni di una variabile.

Questo scopo si può raggiungere considerando le restrizioni di f ai punti non interni; tali restrizioni sono funzioni che dipendono da una sola variabile e si può cercare di trattarle con i risultati noti per tal caso.

Ovviamente lo scopo è individuare eventuali massimi o minimi per mezzo di condizioni necessarie e, se si è certi della loro esistenza, tra essi scegliere massimi e minimi assoluti.

A questo scopo è utile considerare il problema di trovare massimi e minimi di una funzione $f(x, y)$ sull'insieme dei punti del piano che soddisfano l'equazione $g(x, y) = 0$

In questo modo, infatti, è possibile identificare in molti casi l'insieme dei punti di frontiera (e quindi non interni) di un insieme.

Più precisamente ci riferiremo a questo problema come al problema di

Cercare massimi e minimi relativi di f vincolati a $g = 0$

6.1. funzioni definite implicitamente. Per studiare il problema è necessario conoscere qualche cosa in più sull'insieme

$$G = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$$

Più precisamente è necessario rendersi conto che G può essere rappresentato localmente mediante il grafico di una funzione φ .

Per chiarire il concetto consideriamo un semplice esempio.

Sia

$$g(x, y) = x^2 + y^2 - 1$$

ovviamente $g \in \mathcal{C}^1$ ed inoltre

$$\nabla g(x, y) = (2x, 2y) \neq (0, 0)$$

per ognuno dei punti tali che

$$g(x, y) = 0$$

È ben noto che l'equazione

$$g(x, y) = x^2 + y^2 - 1 = 0$$

identifica una circonferenza di raggio unitario centrata in $(0, 0)$.

Per illustrare la possibilità di rappresentare la circonferenza localmente in un punto P_0 mediante una funzione φ possiamo considerare i seguenti casi

- se $P_0 = (0, 1)$ possiamo rappresentare la circonferenza mediante la funzione

$$y = \sqrt{1 - x^2}$$

- se $P_0 = (0, -1)$ possiamo rappresentare la circonferenza mediante la funzione

$$y = -\sqrt{1 - x^2}$$

- se $P_0 = (1, 0)$ possiamo rappresentare la circonferenza mediante la funzione

$$x = \sqrt{1 - y^2}$$

- se $P_0 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ possiamo rappresentare la circonferenza sia mediante la funzione

$$y = \sqrt{1 - x^2}$$

sia mediante la funzione

$$x = \sqrt{1 - y^2}$$

- se $P_0 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ possiamo rappresentare la circonferenza sia mediante la funzione

$$y = \sqrt{1 - x^2}$$

sia mediante la funzione

$$x = -\sqrt{1 - y^2}$$

In generale non è, tuttavia, possibile trovare esplicitamente la funzione φ , come abbiamo fatto nell'esempio appena visto, tuttavia è per taluni scopi sufficiente sapere che questa funzione esiste.

A questo proposito si può dimostrare che

TEOREMA 1.8. - *delle funzioni implicite di U. Dini* - Se g è sufficientemente regolare ($g \in \mathcal{C}^1$, $\nabla g(x, y) \neq (0, 0)$) l'insieme

$$\{(x, y) \in \mathbb{R}^2, : g(x, y) = 0\}$$

può essere rappresentato localmente, (cioè in un intorno di ogni suo punto), come grafico di una funzione

$$y = \varphi(x)$$

6.2. il principio dei moltiplicatori di Lagrange. Si può trovare una condizione necessaria affinché un punto P_0 sia di minimo o di massimo per f vincolato a $g = 0$; possiamo enunciare tale condizione come segue

TEOREMA 1.9. - *dei moltiplicatori di Lagrange*- Se $f, g \in \mathcal{C}^1$, $\nabla f(P_0) \neq (0, 0)$ e $P_0 = (x_0, y_0)$ è un punto di minimo o di massimo per f vincolato a $g(x, y) = 0$, allora

$$\nabla f(P_0) \parallel \nabla g(P_0)$$

o equivalentemente esiste $\lambda \in \mathbb{R}$ tale che

$$\nabla f(P_0) = \lambda \nabla g(P_0)$$

Se infatti $\nabla f(P_0)$ e $\nabla g(P_0)$ non fossero paralleli, tenendo conto del fatto che $\nabla f(P_0)$ è ortogonale alla curva definita da $f(x, y) = f(x_0, y_0)$ mentre $\nabla g(P_0)$ è ortogonale alla curva definita da $g(x, y) = 0$ avremmo una situazione simile a quella illustrata nella figura 1.9.

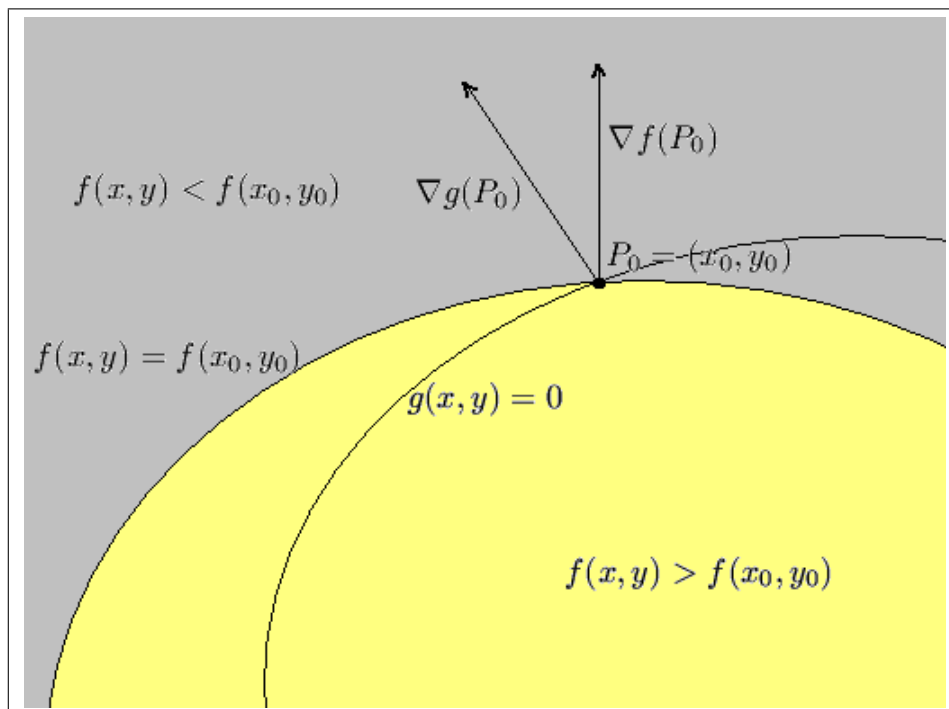


FIGURA 1.9. Principio dei moltiplicatori di Lagrange

Dalla figura si vede che ci sarebbero punti soddisfacenti l'equazione $g(x, y) = 0$ tali che $f(x, y) > f(x_0, y_0)$ ed anche punti tali che $f(x, y) < f(x_0, y_0)$.

Ciò escluderebbe che P_0 sia un punto di minimo o di massimo di f vincolato a $g = 0$

Possiamo dimostrare con maggior precisione il risultato come segue.

Siano $f, g : A \rightarrow \mathbb{R}$, $P_0 = (x_0, y_0) \in A \subset \mathbb{R}^2$, A aperto, $f, g \in C^1(A)$, e supponiamo che $g(x_0, y_0) = 0$. Supponiamo inoltre che $\nabla g(x_0, y_0) \neq 0$, il che significa, a meno di cambiare il nome delle variabili, che si può supporre $g_y(x_0, y_0) \neq 0$; allora si può dimostrare che esiste una funzione φ definita in un intorno di x_0 che assume valori in un intorno di y_0 e per la quale si ha

$$g(x, \varphi(x)) = 0$$

Pertanto la funzione $f(x, \varphi(x))$ ammette in x_0 un punto di minimo relativo se e solo se $P_0 = (x_0, y_0)$ è un punto di minimo per f vincolato a $g = 0$.

Di conseguenza, se $P_0 = (x_0, y_0)$ è un minimo relativo per f vincolato a $g = 0$ si ha

$$\frac{d}{dx} f(x, \varphi(x)) = f_x(x_0, y_0) + f_y(x_0, y_0) \varphi'(x_0) = 0$$

ed anche

$$g_x(x_0, y_0) + g_y(x_0, y_0) \varphi'(x_0) = 0$$

e la coppia $(1, \varphi'(x_0))$ è soluzione non banale del sistema algebrico lineare omogeneo la cui matrice dei coefficienti è data da

$$\begin{pmatrix} \nabla f(x_0, y_0) \\ \nabla g(x_0, y_0) \end{pmatrix}$$

Ne segue che esistono $\alpha, \beta \in \mathbb{R}$ non entrambi nulli, tali che

$$\alpha \nabla f(x_0, y_0) + \beta \nabla g(x_0, y_0) = 0$$

e, dal momento che $\nabla g(x_0, y_0) \neq 0$, ne viene che deve essere $\alpha \neq 0$.

Possiamo pertanto affermare, a meno di dividere per α , che esiste λ tale che

$$\nabla f(x_0, y_0) + \lambda \nabla g(x_0, y_0) = 0$$

Viceversa, posto

$$h(x) = f(x, \varphi(x))$$

se $h'(x_0) = 0$ e $h''(x_0) > 0$, (x_0, y_0) è un punto di minimo relativo per f vincolato a $g = 0$.

Concludiamo osservando un semplice fatto, spesso utile quando si trattano problemi di programmazione lineare.

TEOREMA 1.10. *Sia $f : A \rightarrow \mathbb{R}$, $A \subset \mathbb{R}^2$ convesso, chiuso e limitato, f convessa e continua; allora il massimo di f in A è assunto anche in punti che sono sulla frontiera di A*

DIMOSTRAZIONE. Sia

$$f(P) = \max\{f(Q) : Q \in A\}$$

allora, se P è interno ad A , detti $Q, R \in A$ gli estremi del segmento ottenuto intersecando A con una qualunque retta passante per P , si ha

$$P = \lambda Q + (1 - \lambda)R$$

e

$$f(P) \leq \lambda f(Q) + (1 - \lambda)f(R) \leq \max\{f(Q), f(R)\}$$

□

Osservazione. Nel caso in cui A sia poliedrale, cioè se

$$A = \{P \in \mathbb{R}^2 : g_i(P) \leq 0, g_i \text{ lineare}, i = 1, \dots, m\}$$

il massimo si può cercare solo tra i vertici della frontiera.

□

CAPITOLO 2

INTEGRAZIONE PER LE FUNZIONI DI DUE VARIABILI

1. Definizione di integrale doppio.

Se f è una funzione di 2 variabili positiva e se $R = [a, b] \times [c, d]$ è un rettangolo contenuto nel suo dominio, possiamo considerare il problema di calcolare il volume V delimitato dal piano (x, y) dal grafico di f e dal cilindro generato da R con generatrici parallele all'asse z (si veda la figura 2.1).

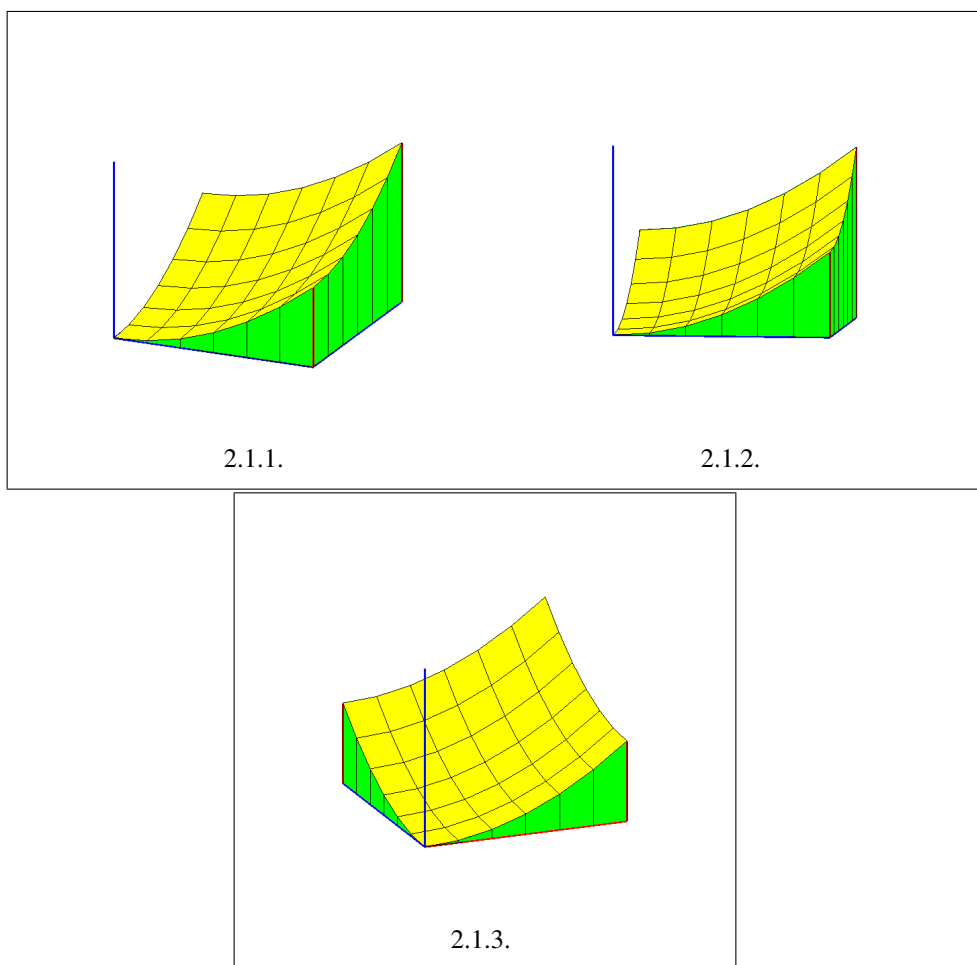


FIGURA 2.1.

Il volume può essere definito

- considerando una partizione di R ,
- definendo in corrispondenza le somme superiori e le somme inferiori di f relative alla partizione scelta,
- dichiarando una funzione integrabile se, al variare delle partizioni, l'estremo inferiore delle somme superiori e l'estremo superiore delle somme inferiori coincidono,
- in tal caso chiamiamo il loro valore comune

$$\iint_R f(x, y) dx dy$$

L'esistenza dell'integrale è assicurata, similmente a quanto accade per le funzioni di una variabile, dalla continuità della funzione integranda, e si può anche dimostrare che è sufficiente che f sia continua su R a meno di un insieme di area 0. Possiamo in altre parole dimostrare il seguente risultato.

Se f è una funzione di due variabili limitata su un insieme chiuso e limitato D ed è continua a meno di un sottoinsieme di misura 0, allora f è integrabile su D

(Pur non entrando nei particolari della definizione di area, possiamo ricordare che è possibile calcolare l'area di insiemi piani significativi usando la teoria dell'integrazione per le funzioni di una variabile)

Quando una funzione è integrabile possiamo approssimare il suo integrale su R anche usando le somme di Riemann; tali somme possono anche essere usate per dare la definizione di integrabilità e si calcolano come segue:

- Si suddivide il rettangolo R in rettangoli più piccoli R_j , ad esempio ottenuti suddividendo i lati di R in parti uguali di ampiezza δ_x e δ_y , rispettivamente (figura 2.3);
- si sceglie in maniera arbitraria un punto (ξ_j, η_j) in ognuno dei rettangoli R_j e si calcola la quota $f(\xi_j, \eta_j)$ (figura 2.2.1)
- si sostituisce al volume delimitato su R_j dalla funzione f il volume del parallelepipedo di base R_j ed altezza $f(\xi_j, \eta_j)$ (figura 2.2.2)
- si calcola la somma di tutti i contributi così ottenuti da ciascuno dei rettangoli della partizione.

Le somme di Riemann sono così definite da

$$(2.1) \quad \mathcal{R}(f) = \sum_j f(\xi_j, \eta_j) \delta_x \delta_y$$

e, quando la partizione è abbastanza fine, cioè suddivide R in rettangoli R_j abbastanza piccoli, $\mathcal{R}(f)$ approssima il valore di $\iint_R f(x, y) dx dy$

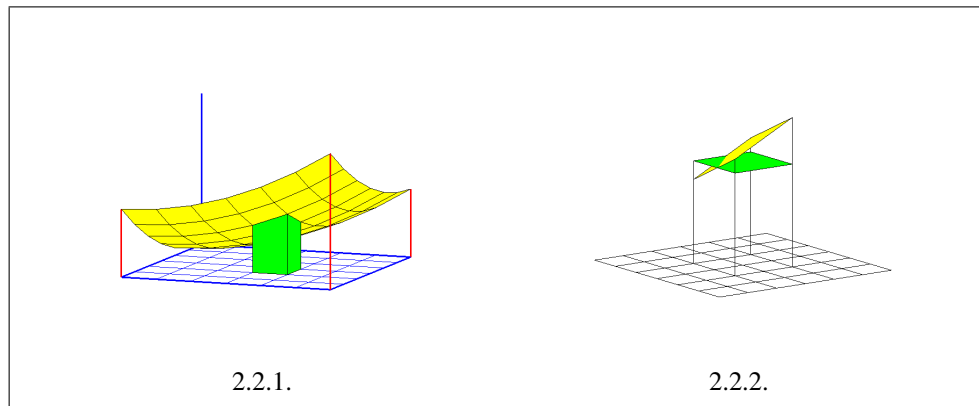


FIGURA 2.2.

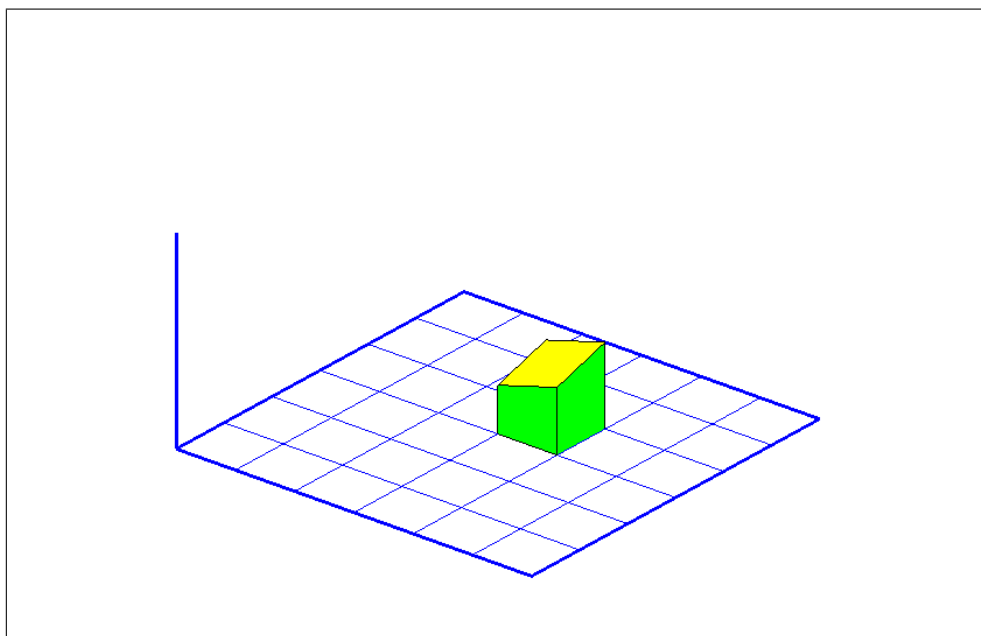


FIGURA 2.3. Singolo elemento di volume

2. Formule di riduzione per gli integrali doppi.

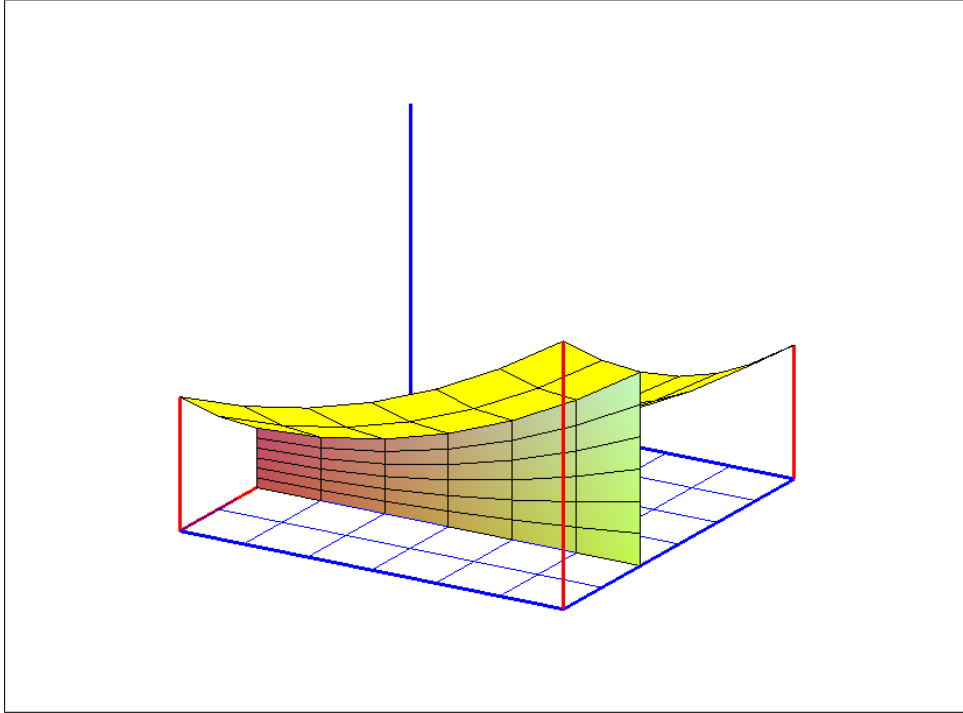
Purtroppo non disponiamo, per il calcolo di un integrale doppio, di uno strumento tanto potente quanto il teorema fondamentale del calcolo integrale; questo risultato si può infatti estendere anche al calcolo delle funzioni di più variabili, ma si colloca in un contesto più generale: quello delle forme differenziali e del teorema di Stokes.

Occorre quindi cercare altre vie per il calcolo degli integrali doppi.

Se definiamo

$$S(x) = \{(x, y, z) \in \mathbb{R}^3 : a \leq x \leq b, 0 \leq z \leq f(x, y)\}$$

$S(x)$ rappresenta una sezione del volume V , si veda figura 2.4.

FIGURA 2.4. $S(x)$

ed il calcolo integrale per le funzioni di una variabile consente di calcolare la sua area $A(x)$ mediante la

$$A(x) = \int_a^b f(x, y) dy$$

Possiamo considerare il volume V come la somma (infinita) dei volumi elementari $A(x)$ (che sono nulli) per $x \in [a, b]$; naturalmente la somma infinita si calcola integrando $A(x)$ su $[a, b]$ e quindi

$$\iint_R f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx$$

In maniera del tutto simile possiamo calcolare

$$\iint_R f(x, y) dx dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

Il calcolo di un integrale doppio può quindi essere ridotto al calcolo di due integrali semplici (formule di riduzione).

Il vincolo fin qui posto sul dominio di integrazione, (R è un rettangolo) non può tuttavia essere mantenuto e quindi è necessario definire

$$\iint_A f(x, y) dx dy$$

per una classe di sottoinsiemi del piano un po' più generale.

È naturale considerare per questo scopo la classe dei **domini normali**

Diciamo che un insieme D è un dominio normale rispetto all'asse x se

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \alpha(x) \leq y \leq \beta(x)\}$$

**dove $[a, b]$ è un intervallo reale e α e β sono funzioni continue su $[a, b]$.
(Si veda la figura 2.5).**

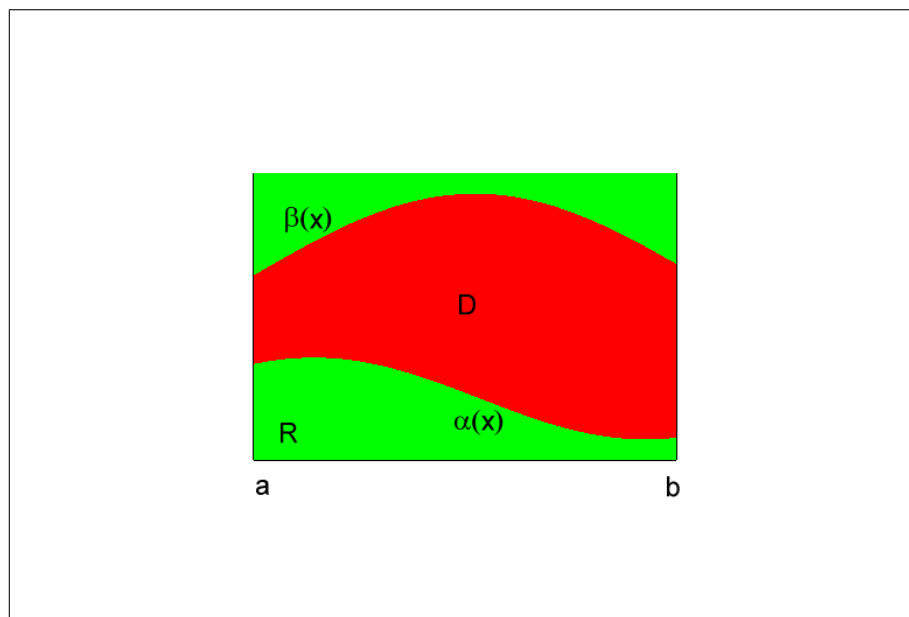


FIGURA 2.5. Dominio normale rispetto all'asse x

Diciamo che un insieme E è un dominio normale rispetto all'asse y se

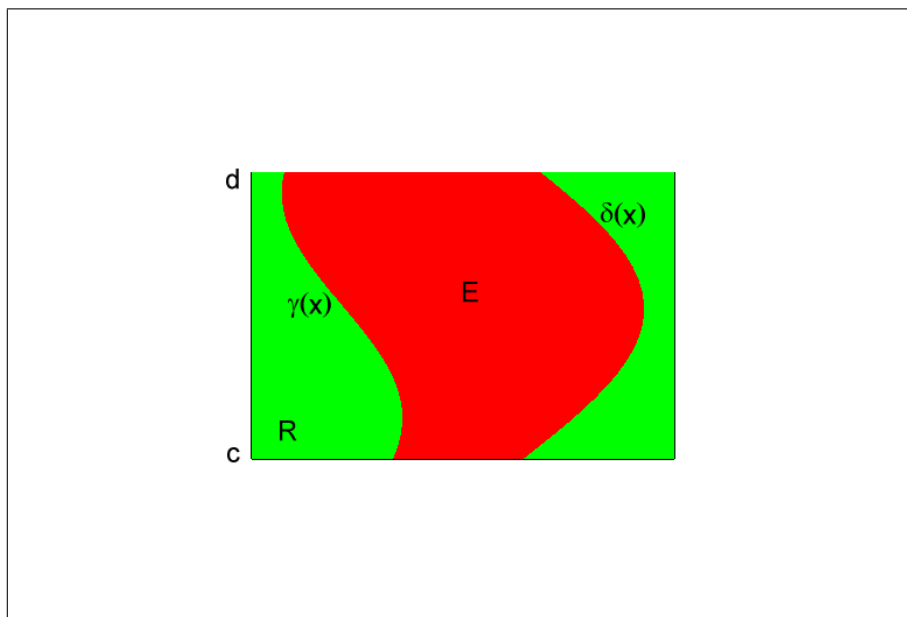
$$E = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \gamma(y) \leq x \leq \delta(y)\}$$

**dove $[c, d]$ è un intervallo reale e γ e δ sono funzioni continue su $[c, d]$.
(Si veda la figura 2.6).**

Per definire, ad esempio,

$$\iint_D f(x, y) dx dy$$

possiamo

FIGURA 2.6. Dominio normale rispetto all'asse x

- definire una funzione

$$\tilde{f}(x, y) = \begin{cases} f(x, y) & \text{se } (x, y) \in D \\ 0 & \text{se } (x, y) \notin D \end{cases}$$

- considerare un rettangolo $R \supset D$
- definire

$$\iint_D f(x, y) dx dy = \iint_R \tilde{f}(x, y) dx dy$$

Nel compiere questa procedura, possiamo osservare che, poichè si può dimostrare che il grafico di una funzione continua ha area nulla e poichè la definizione di $\tilde{f}$ può generare discontinuità solo nei punti dei grafici di α e di β , se f è continua a meno di insiemi di area nulla tale risulta anche $\tilde{f}$ e pertanto

Una funzione f continua a meno di un insieme di area nulla è integrabile su un dominio normale D .

A completamento occorre poi osservare che

$$\begin{aligned}\iint_D f(x, y) dx dy &= \iint_R \tilde{f}(x, y) dx dy \\ &= \int_a^b \left(\int_c^d \tilde{f}(x, y) dy \right) dx = \int_a^b \left(\int_{\alpha(x)}^{\beta(x)} f(x, y) dy \right) dx\end{aligned}$$

ed in maniera del tutto simile

$$\begin{aligned}\iint_E f(x, y) dx dy &= \iint_R \tilde{f}(x, y) dx dy \\ &= \int_c^d \left(\int_a^b \tilde{f}(x, y) dx \right) dy = \int_c^d \left(\int_{\gamma(y)}^{\delta(y)} f(x, y) dx \right) dy\end{aligned}$$

Infine osserviamo che quanto abbiamo visto è applicabile ad insiemi che siano unione finita di domini normali. Questo ci permette di considerare la maggior parte degli insiemi che si incontrano nella pratica del calcolo.

3. Cambiamento di variabili negli integrali doppi

3.1. Cambiamento di variabili lineari. Consideriamo ora il problema di calcolare l'area di un parallelogrammo A che abbia come lati i vettori (a, b) e (c, d) .

Semplici considerazioni di geometria permettono di stabilire che

$$\text{Area}(A) = ad - bc = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ed inoltre, se teniamo conto del fatto che il volume del cilindro di altezza 1 (figura 2.7) che ha per base il parallelogrammo è uguale ad $\text{Area}(A)$, possiamo anche affermare che

$$\text{Area}(A) = \iint_A 1 dx dy = \iint_B \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} du dv$$

dove

$$B = \{(u, v) : 0 \leq u \leq 1, 0 \leq v \leq 1\} = [0, 1] \times [0, 1]$$

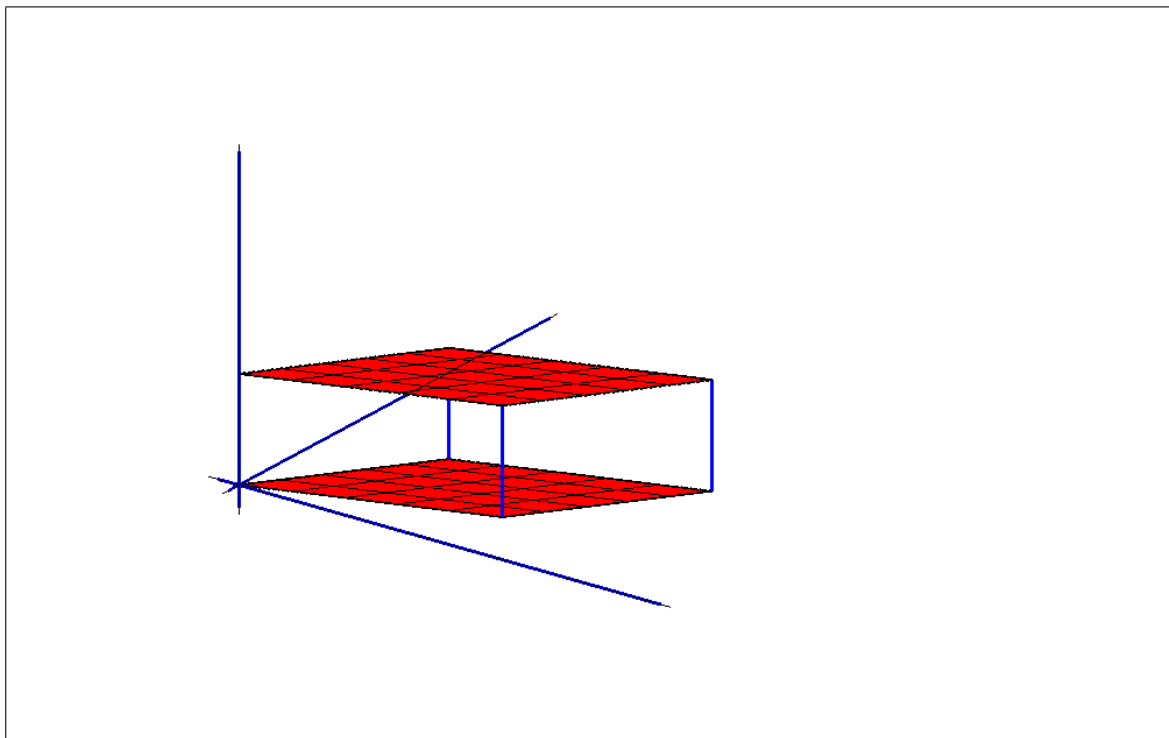


FIGURA 2.7. Volume ed area

ed osservare che il quadrato $B = [0, 1] \times [0, 1]$ si trasforma nel parallelogramma A mediante le corrispondenze

$$(2.2) \quad \begin{cases} x = au + bv \\ y = cu + dv \end{cases} \quad \text{cioè} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

Se supponiamo che

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0$$

la corrispondenza è biunivoca e può essere invertita; sia

$$(2.3) \quad \begin{cases} u = \alpha x + \beta y \\ v = \gamma x + \delta y \end{cases} \quad \text{cioè} \quad \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

la corrispondenza inversa.

Con riferimento alla definizione di integrale possiamo anche osservare che una partizione del quadrato B in quadrati più piccoli B_j corrisponde ad una suddivisione del parallelogramma A in parallelogrammi A_j , simili, più piccoli (si veda la figura 2.8).

Pertanto se f è una funzione definita su A , per calcolare

$$\iint_A f(x, y) dx dy$$

possiamo calcolare le somme di Riemann usando la partizione di A in parallelogrammi, che risulta più naturale di una partizione in rettangoli; Le somme di Riemann in questo caso risultano essere

$$\mathcal{R}(f) = \sum_j f(x_j, y_j) \text{Area}(A_j) = \sum_j f(x_j, y_j) \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{Area}(B_j)$$

Ma esiste un unico punto $(u_j, v_j) \in B$ tale che

$$(x_j, y_j) = (au_j + bv_j, cu_j + dv_j)$$

per cui

$$\mathcal{R}(f) = \sum_j f(au_j + bv_j, cu_j + dv_j) \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{Area}(B_j)$$

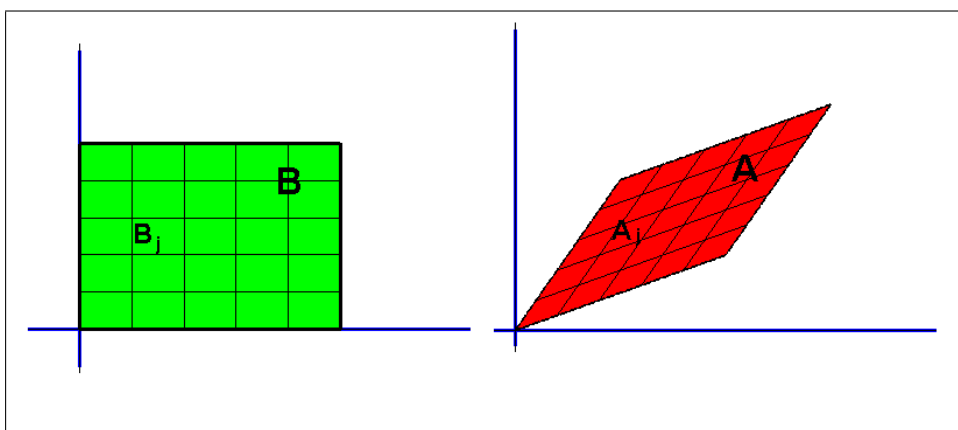


FIGURA 2.8. Cambiamento di variabili lineare

Tali somme al raffinarsi della partizione si approssimano a

$$\iint_B f(au + bv, cu + dv) \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} du dv$$

e quindi possiamo concludere che

$$\iint_A f(x, y) dx dy = \iint_B f(au + bv, cu + dv) \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} du dv$$

3.2. Coordinate Polari nel piano. Possiamo usare in luogo di 2.2 anche altre trasformazioni; ad esempio possiamo usare la trasformazione in coordinate polari che è definita da:

$$(2.4) \quad \begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases} \quad \rho \geq 0, \quad \theta \in [0, 2\pi]$$

La 2.4 trasforma

- le rette $\rho = R$ in circonferenze centrate nell'origine di raggio R
- le rette $\theta = \alpha$ in semirette passanti per l'origine inclinate di un angolo α rispetto al semiasse positivo dell'asse x .
- i settori di corona circolare nel piano (x, y)

$$A = \{(x, y) : r \leq \sqrt{x^2 + y^2} \leq R, \alpha \leq \tan y/x \leq \beta\} = \{(x, y) : r \leq \rho \leq R, \alpha \leq \theta \leq \beta\}$$

in rettangoli

$$B = \{(\rho, \theta) : \rho \leq R, \alpha \leq \theta \leq \beta\} = [r, R] \times [\alpha, \beta]$$

nel piano (ρ, θ) . (Si veda la figura 2.9).

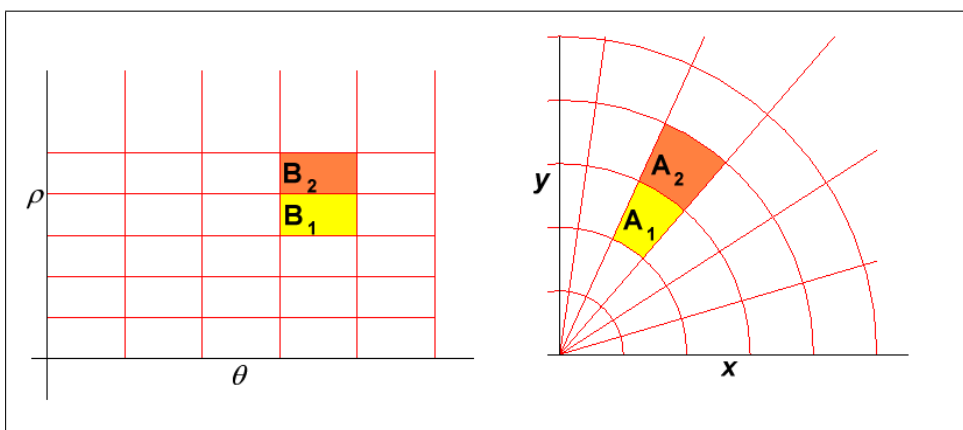


FIGURA 2.9. Cambiamento di variabili in coordinate polari

Purtroppo, la trasformazione definita dalle 2.4 non è biunivoca nè invertibile ed inoltre se

$$B_1 = [R - \delta_r, R] \times [\alpha, \beta]$$

e

$$B_2 = [R, R + \delta_r] \times [\alpha, \beta]$$

l'immagine A_1 di B_1 ed A_2 di B_2 hanno aree diverse anche se B_1 e B_2 hanno aree uguali.

Più precisamente si vede che se l'area di A_1 è più piccola di quella di A_2 poichè A_1 è più vicino all'origine di A_2 .

Possiamo calcolare che:

$$\begin{aligned} \text{Area}(B) &= (R - r)(\beta - \alpha) \\ \text{Area}(A) &= \frac{1}{2}(R^2 - r^2)(\beta - \alpha) = \frac{1}{2}(R + r)\text{Area}(B) \end{aligned}$$

Pertanto non possiamo procedere, come nel caso di 2.2 in quanto il fattore di conversione per ottenere Area (A) da Area (B) non è costante.

Possiamo tuttavia affermare che

$$(2.5) \quad \text{Area}(A) = \iint_A 1 dx dy$$

e la 2.5 si può ottenere come somma di settori circolari più piccoli A_j delimitati da circonferenze di raggio ρ e $\rho + \delta_\rho$ e aventi ampiezza δ_θ .

Se

$$B_j = [\rho, \rho + \delta_\rho] \times [\theta, \theta + \delta_\theta]$$

l'area di ciascuno dei settori A_j è data da

$$\text{Area}(A_j) = \frac{1}{2}(2\rho + \delta_\rho)\delta_\rho\delta_\theta = \frac{1}{2}(2\rho + \delta_\rho)\text{Area}(B_j)$$

ed inoltre se δ_ρ è piccolo e trascurabile avremo che

$$\text{Area}(A_j) \approx \frac{1}{2}2\rho\text{Area}(B_j)$$

Poichè

$$\text{Area}(A) = \sum_j \text{Area}(A_j) \approx \sum_j \rho \text{Area}(B_j)$$

Possiamo affermare che

$$\text{Area}(A) = \iint_B \rho d\rho d\theta$$

Se poi f è una funzione definita su A , possiamo affermare che

$$\iint_A f(x, y) dx dy = \iint_A f(\rho \cos \theta, \rho \sin \theta) \rho d\rho d\theta$$

INTEGRAZIONE DI FUNZIONI DI TRE VARIABILI

1. Definizione di integrale triplo.

Le formule di riduzione che abbiamo usato per il calcolo di un integrale doppio consentono di ricondurre il problema al calcolo di due integrali semplici.

In modo del tutto simile possiamo trovare il modo di calcolare un integrale triplo, cioè l'integrale di una funzione f di tre variabili (x, y, z) su un dominio V contenuto in $\mathbb{R}^3$

$$\iiint_V f(x, y, z) dx dy dz$$

Il concetto di area che è naturalmente collegato al concetto di integrale semplice e quello di volume che è caratteristico dell'integrale doppio si estende al concetto di ipervolume a quattro dimensioni per gli integrali tripli.

Inoltre, come nel caso di due variabili in cui abbiamo osservato che

$$\text{Area}(A) = \iint_A 1 dx dy = \text{Volume}(C_A)$$

se C_A è il cilindro di base A e di altezza 1, possiamo dire che

$$\text{Volume}(V) = \iiint_V 1 dx dy dz = \text{IperVolume}(C_V)$$

dove C_V è il cilindro di base V e di altezza 1.

2. Formule di riduzione per gli integrali tripli.

Per gli integrali tripli sono possibili diverse scomposizioni che danno origine a diverse formule di riduzione che riteniamo utile illustrare mediante qualche esempio.

Ci occuperemo allo scopo di calcolare

$$\iiint_V f(x, y, z) dx dy dz$$

dove

$$(3.1) \quad V = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq z \leq 2 - \sqrt{x^2 + y^2}\}$$

La parte di $\mathbb{R}^3$ definita dalla 3.1 è quella indicata nella figura 3.1.2 3.1.1

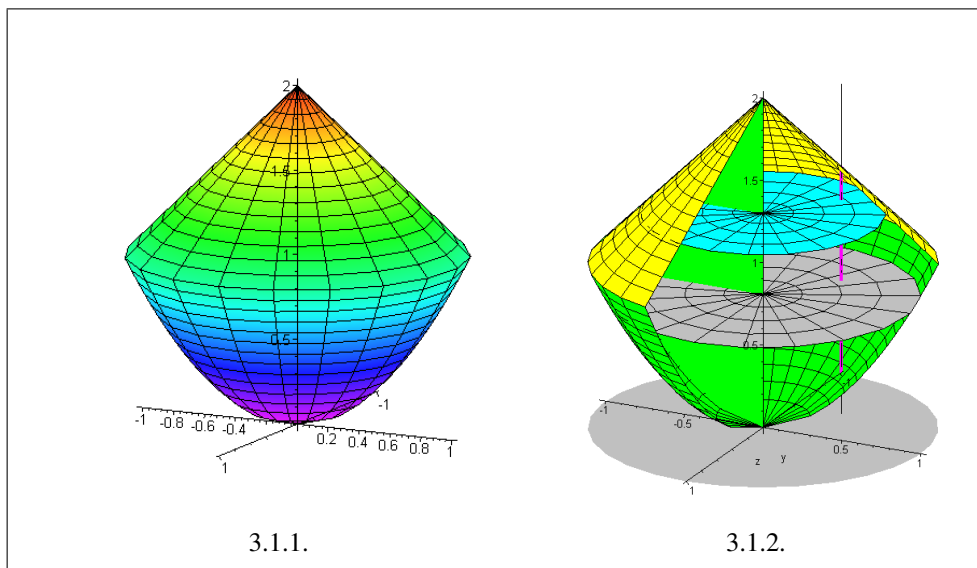


FIGURA 3.1.

Possiamo integrare su V sommando, cioè integrando rispetto a z , i valori ottenuti mediante il calcolo dell'integrale doppio sulle sezioni di V definite da

$$S(z) = \{(x, y) : (x, y, z) \in V\}$$

Avremo pertanto che

$$\iiint_V f(x, y, z) dx dy dz = \int_0^2 \left(\iint_{S(z)} f(x, y, z) dx dy \right) dz$$

e gli integrali indicati si calcolano come già sappiamo.

Possiamo anche calcolare l'integrale triplo considerando la proiezione D del solido V e calcolando

$$\iiint_D \left(\int_{F(x,y)} f(x, y, z) dx, dy \right) dz = \iint_D \left(\int_{x^2+y^2}^{2-\sqrt{x^2+y^2}} f(x, y, z) dx, dy \right) dz$$

3. Cambiamento di variabili per gli integrali tripli.

Anche per gli integrali tripli è utile considerare qualche cambiamento di variabile allo scopo di semplificare i calcoli nel caso di solidi con particolari simmetrie.

I cambiamenti di variabile più comuni sono quello lineare, quello in coordinate cilindriche e quello in coordinate sferiche, che ora illustriamo brevemente.

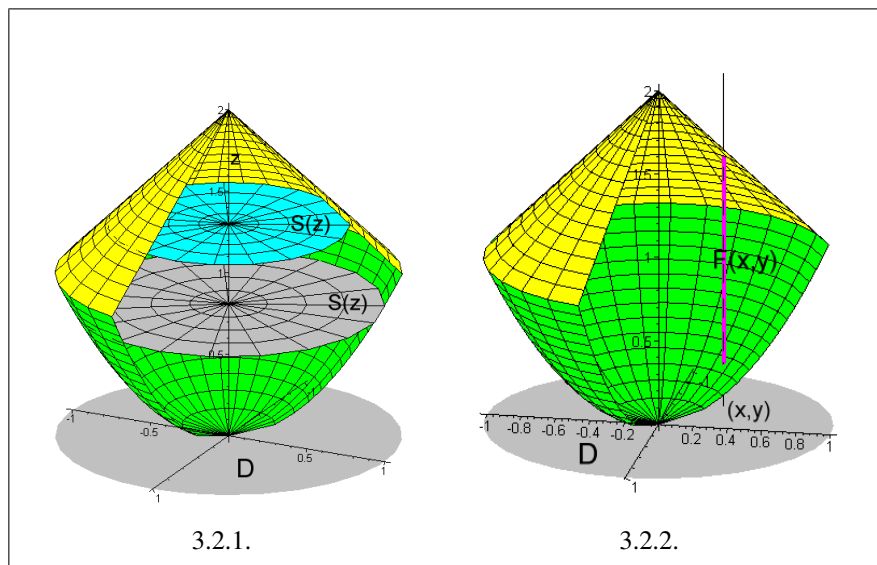


FIGURA 3.2.

3.1. Cambio di variabili lineare. Si tratta del cambiamento di variabili definito dalle

$$\begin{cases} x = a_1 r + b_1 s + c_1 t \\ y = a_2 r + b_2 s + c_2 t \\ z = a_3 r + b_3 s + c_3 t \end{cases}$$

Se $A \subset \mathbb{R}^3$ e se B è il trasformato di A mediante il cambiamento di variabili lineari si ha

$$\begin{aligned} \iiint_A f(x, y, z) dx dy dz = \\ \iiint_B f(a_1 r + b_1 s + c_1 t, a_2 r + b_2 s + c_2 t, a_3 r + b_3 s + c_3 t) \left| \frac{\partial(x, y, z)}{\partial(r, s, t)} \right| dr ds dt \end{aligned}$$

dove

$$\left| \frac{\partial(x, y, z)}{\partial(r, s, t)} \right| = \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

3.2. Coordinate cilindriche. Si tratta del cambiamento di variabili definito dalle

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \\ z = z \end{cases} \quad \rho \geq 0 \quad 0 \leq \theta \leq 2\pi \quad , z \in \mathbb{R}$$

Se $A \subset \mathbb{R}^3$ e se B è il trasformato di A mediante il cambiamento di variabili in coordinate cilindriche si ha

$$\iiint_A f(x, y, z) dx dy dz = \iiint_B f(\rho \cos \theta, \rho \sin \theta, z) \rho d\rho d\theta dz$$

3.3. Coordinate sferiche. Si tratta del cambiamento di variabili definito dalle

$$\begin{cases} x = \rho \cos \theta \cos \varphi \\ y = \rho \sin \theta \cos \varphi \\ z = \rho \sin \varphi \end{cases} \quad \rho \geq 0 \quad 0 \leq \theta \leq 2\pi \quad -\pi/2 \leq \varphi \leq \pi/2$$

Se $A \subset \mathbb{R}^3$ e se B è il trasformato di A mediante il cambiamento di variabili in coordinate sferiche si ha

$$\begin{aligned} \iiint_A f(x, y, z) dx dy dz = \\ \iiint_B f(\rho \cos \theta \cos \varphi, \rho \sin \theta \cos \varphi, \rho \sin \varphi) \rho \cos \varphi d\rho d\theta d\varphi \end{aligned}$$

CAPITOLO 4

INTEGRALI MULTIPLI IMPROPRI

Come nel caso degli integrali semplici, possiamo considerare il problema di calcolare l'integrale di una funzione di due o più variabili che non siano limitate o su domini di integrazione non limitati.

Qui illustriamo l'argomento con qualche esempio che è significativo anche per il seguito e che fornisce un utile strumento per affrontare, se necessario anche gli altri casi.

Consideriamo pertanto una funzione f definita su $\mathbb{R}^2$ limitata ed integrabile su ogni insieme limitato e chiuso di $\mathbb{R}^2$ (ad esempio continua) e sia $D \subset \mathbb{R}^2$ un sottoinsieme non limitato di $\mathbb{R}^2$.

In tali condizioni non è lecito definire

$$\iint_D f(x, y) dx dy$$

in senso proprio, tuttavia possiamo procedere come segue:

Innanzitutto assicuriamoci di poter lavorare con una funzione sempre positiva; se $f \geq 0$ nulla è da fare ma se così non è basta definire

$$f_+(x, y) = \max\{f(x, y), 0\} \quad \text{e} \quad f_-(x, y) = \min\{f(x, y), 0\}$$

osservare che

$$f = f_+ + f_-$$

e calcolare

$$\iint_D f(x, y) dx dy = \iint_{D_+} f_+(x, y) dx dy + \iint_{D_-} f_-(x, y) dx dy$$

dove

$$D_+ = \{(x, y) \in D : f(x, y) \geq 0\}, \quad D_- = \{(x, y) \in D : f(x, y) \leq 0\}$$

e chiedere che entrambi gli integrali a secondo membro esistano e non diano luogo ad una forma indeterminata.

Supponiamo quindi che $f \geq 0$ e consideriamo una successione di insiemi D_n soddisfacente le seguenti condizioni:

- D_n è chiuso e limitato
- $D_{n+1} \supset D_n$
- per ogni insieme limitato e chiuso K contenuto in D si può trovare un $D_{\bar{n}}$ tale che $D_{\bar{n}} \supset K$

È evidente che le condizioni sopra elencate esprimono il concetto che la successione di domini D_n riempie, invade, l'insieme D ed infatti una successione che soddisfa tali condizioni si chiama successione di domini invadenti D .

DEFINIZIONE 4.1. Se $f \geq 0$ e se D_n è una successione di domini invadenti D allora definiamo

$$\iint_D f(x, y) dx dy = \lim_n \iint_{D_n} f(x, y) dx dy$$

Si può dimostrare che, qualora il limite esista, è indipendente dalla successione di domini invadenti usata.

1. Qualche esempio

Consideriamo il problema di calcolare

$$\iint_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dx dy$$

Definiamo

$$D_n = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq n^2\}$$

allora D_n è una successione di domini invadenti $\mathbb{R}^2$ e quindi

$$\begin{aligned} \iint_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dx dy &= \lim_n \iint_{D_n} e^{-(x^2+y^2)/2} dx dy = \\ &= \lim_n \int_0^{2\pi} \left(\int_0^n \rho e^{-(\rho^2)/2} d\rho \right) d\theta = \\ &= \lim_n 2\pi \int_0^n \rho e^{-(\rho^2)/2} d\rho = 2\pi \lim_n \left. -e^{-(\rho^2)/2} \right|_0^n = 2\pi \lim_n (1 - e^{-(n^2)/2}) = 2\pi \end{aligned}$$

Il risultato appena ricavato ha una conseguenza interessante, infatti, poiché il valore dell'integrale non dipende dalla successione di domini invadenti usata, possiamo rifare il calcolo anche usando la successione definita da

$$Q_n = [-n, n] \times [-n, n]$$

ed otterremo lo stesso risultato.

Avremo

$$\begin{aligned}
2\pi &= \iint_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dx dy = \lim_n \iint_{Q_n} e^{-(x^2+y^2)/2} dx dy = \\
&= \lim_n \int_{-n}^n \left(\int_{-n}^n e^{-(x^2)/2} e^{-(y^2)/2} dx \right) dy = \\
&= \lim_n \left(\int_{-n}^n e^{-(y^2)/2} dy \right) \left(\int_{-n}^n e^{-(x^2)/2} dx \right) = \lim_n \left(\int_{-n}^n e^{-(t^2)/2} dt \right)^2 = \left(\int_{-\infty}^{+\infty} e^{-(t^2)/2} dt \right)^2
\end{aligned}$$

e possiamo affermare che

$$\int_{-\infty}^{+\infty} e^{-(t^2)/2} dt = \sqrt{2\pi}$$

e che

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(t^2)/2} dt = 1$$

CAPITOLO 5

LA SOMMA DI INFINITI TERMINI: LE SERIE.

Il problema di sommare un numero non finito di quantità numeriche è stato per lungo tempo considerato privo di senso, ma è giustificabile facilmente, anche dal punto di vista intuitivo, non appena si consideri il seguente esempio.

Sia $I = [0, 1]$ e consideriamo una successione di intervalli così definita: poniamo

$$\begin{aligned}I_1 &= [0, 1/2] \\I_2 &= [1/2, 1/4] \\I_3 &= [1/4, 1/8] \\I_4 &= [1/8, 3/4] \\&\dots\dots\dots\end{aligned}$$

E' ovvio che

$$\cup I_k = [0, 1]$$

ed inoltre la lunghezza del segmento I_k è data da

$$\ell(I_k) = 1/2^k$$

Pertanto

$$1 = \ell([0, 1]) = \sum_{k=1}^{+\infty} \frac{1}{2^k}$$

Possiamo cercare di puntualizzare il concetto di somma infinita mediante la seguente definizione

DEFINIZIONE 5.1. Sia a_k una successione di numeri reali e definiamo

$$S_n = \sum_{k=1}^n a_k$$

Se $\lim S_n$ esiste finito, diciamo che

$$\sum_{k=1}^{+\infty} a_k = S = \lim S_n$$

In tal caso si dice che $\sum_{k=1}^{+\infty} a_k$ è una serie convergente che ha per somma S .

Se $\lim S_n = +\infty (-\infty)$ diciamo che $\sum_{k=1}^{+\infty} a_k$ è una serie positivamente (negativamente) divergente.

Se $\lim S_n$ non esiste diciamo che la serie non è determinata.

Consideriamo ora qualche esempio importante di serie

Sia $x \in \mathbb{R}$ possiamo considerare $a_n = x^n$ e avremo

$$S_n = \sum_{k=0}^n a_k = \sum_{k=0}^n x^k = 1 + x + x^2 + x^3 + \dots + x^n$$

Se osserviamo che

$$xS_n = x + x^2 + x^3 + x^4 + \dots + x^{n+1}$$

si ottiene

$$(1 - x)S_n = 1 - x^{n+1}$$

e, per $x \neq 1$,

$$S_n = \frac{1 - x^{n+1}}{1 - x}$$

Di qui si vede che

- se $|x| < 1$ $\lim S_n = \frac{1}{1-x}$
- se $x \geq 1$ $\lim S_n = +\infty$
- se $x \leq -1$ $\lim S_n$ non esiste.

Pertanto

$$\boxed{\sum_{k=0}^{+\infty} x^k = \frac{1}{1-x} \quad \text{se } |x| < 1}$$

mentre per i restanti valori di x la serie è divergente o indeterminata.

$\sum x^k$ si chiama **serie geometrica di ragione x** .

Possiamo ottenere facilmente altri esempi di serie convergenti. usando la formula di Taylor.

consideriamo lo sviluppo di McLaurin della funzione e^x

$$e^x - 1 - x - \frac{x^2}{2!} - \frac{x^3}{3!} - \dots - \frac{x^n}{n!} = R_{n+1}(x)$$

dove il resto R_{n+1} si può esprimere nella forma di Lagrange mediante la

$$|R_{n+1}(x)| \leq e^{|x|} \frac{|x|^{n+1}}{(n+1)!}$$

Pertanto, se definiamo

$$S_n = \sum_{k=0}^n \frac{x^k}{k!}$$

si ha

$$|e^x - S_n| \leq |R_{n+1}(x)| \leq e^{|x|} \frac{|x|^{n+1}}{(n+1)!}$$

e tenendo conto che

$$\lim \frac{|x|^{n+1}}{(n+1)!} = 0 \text{ per ogni } x \in \mathbb{R}$$

si ha

$$e^x = \lim S_n = \sum_{k=0}^{+\infty} \frac{x^k}{k!}$$

In maniera del tutto analoga si prova che

$$\sin(x) = \sum_{k=0}^{+\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \forall x \in \mathbb{R}$$

$$\cos(x) = \sum_{k=0}^{+\infty} (-1)^k \frac{x^{2k}}{(2k)!} \forall x \in \mathbb{R}$$

$$\ln(1+x) = \sum_{k=1}^{+\infty} (-1)^{k-1} \frac{x^k}{k} \forall x \in [-1/2, 1]$$

CAPITOLO 6

ELEMENTI DI PROBABILITÀ E STATISTICA.

Il calcolo delle probabilità nasce agli inizi del '700 per rispondere alle istanze prodotte dal gioco d'azzardo.

Il suo scopo era inizialmente introdurre una stima numerica dei rischi ad esso connessi e successivamente lo stesso metodo fu esteso per studiare i fenomeni caratterizzati da elementi di incertezza.

Vediamo innanzi tutto di stabilire quali sono gli elementi necessari per parlare di probabilità

**Occorre innanzi tutto considerare una famiglia di eventi che costituiscono lo spazio su cui definiremo la probabilità.
Possiamo ad esempio associare all'idea di evento quella di sottoinsieme di un insieme dato che chiameremo evento certo.**

1. Spazi di probabilità

Consideriamo, ad esempio, il più semplice tra i giochi d'azzardo, e cioè il lancio di una moneta, possiamo dire che gli eventi possibili sono

- L'uscita di Testa $\{T\}$
- L'uscita di Croce $\{C\}$

mentre l'insieme $\mathcal{U} = \{T, C\}$ costituisce l'evento certo.

In questo caso è naturale definire la probabilità degli eventi che entrano in gioco:

$$(6.1) \quad \mathcal{P}(T) = \frac{1}{2}$$

$$(6.2) \quad \mathcal{P}(C) = \frac{1}{2}$$

$$(6.3) \quad \mathcal{P}(\mathcal{U}) = \mathcal{P}(T) + \mathcal{P}(C) = 1$$

$$(6.4) \quad \mathcal{P}(\emptyset) = 0$$

giustificando la definizione con il fatto che su due possibili uscite una sola è favorevole nel caso si consideri T o C , mentre entrambe vanno bene nel caso si consideri l'unione di T e C . È ovvio che con ciò supponiamo che T e C si presentino con ugual frequenza, cioè che la moneta sia non truccata.

Un secondo esempio di spazio di probabilità si può costruire considerando il caso del lancio di due dadi.

Se le facce sono numerate, come al solito, da 1 a 6 possiamo identificare l'esito del lancio con la coppia di numeri (i, j) (punteggio) che si leggono sulla faccia superiore del primo e del secondo dado.

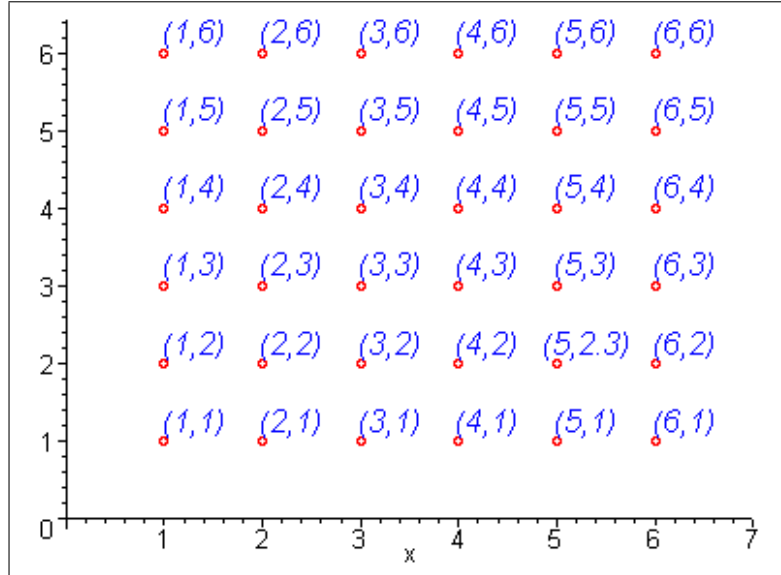


FIGURA 6.1. Lo spazio $\mathcal{U}$ degli eventi nel caso del lancio di due dadi

In tal modo possiamo identificare ciascuna delle 36 possibili uscite (eventi) con il punto del piano cartesiano di coordinate (i, j) ; indicheremo tale evento con il simbolo $A_{i,j}$, (si veda la figura 6.1).

Poichè nel caso di dadi non truccati ogni evento è equiprobabile possiamo affermare che la probabilità di $A_{i,j}$ è data da

$$\mathcal{P}(A_{i,j}) = \frac{1}{36}$$

Ovviamente possiamo considerare anche altri eventi, ad esempio possiamo cercare di stimare la probabilità che si presenti l'evento $A_{1,4}$ oppure l'evento $A_{5,6}$; possiamo indicare il nuovo evento come

$$B = A_{1,4} \cup A_{5,6} = \{A_{1,4}, A_{5,6}\}$$

ed è ragionevole stimare che, poichè accettiamo 2 eventi su 36 possibili,

$$\mathcal{P}(B) = \frac{2}{36} = \frac{1}{36} + \frac{1}{36}$$

Se vogliamo stimare la probabilità che si presenti uno qualunque degli eventi $A_{i,j}$, cioè se vogliamo stimare la probabilità che si presenti l'evento

$$\mathcal{U} = \{A_{i,j} : i, j = 1..6\}$$

poichè accettiamo 36 possibilità su 36 possiamo dire che

$$\mathcal{P}(\mathcal{U}) = 1$$

e che $\mathcal{U}$ è l'evento certo.

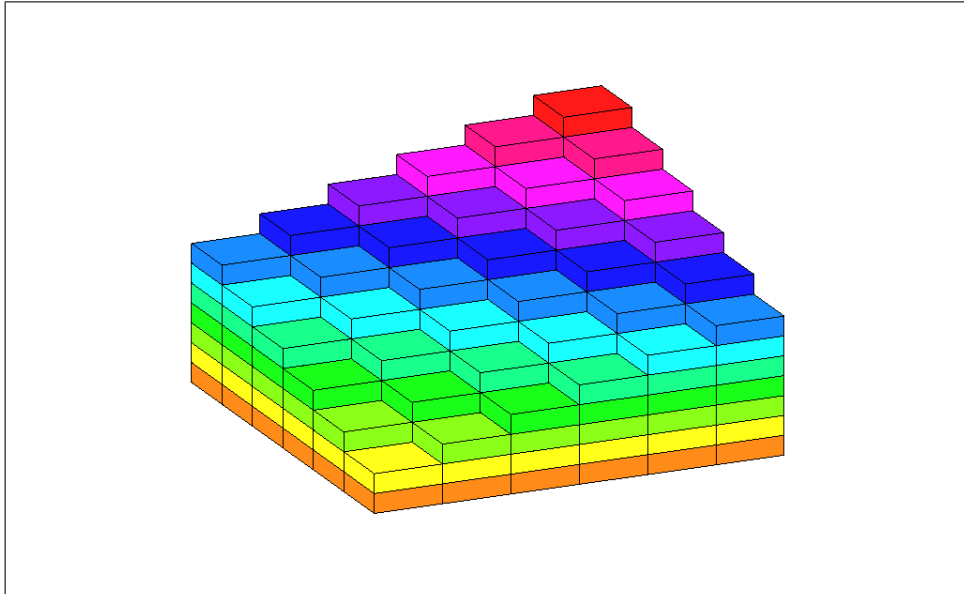


FIGURA 6.2. Istogramma relativo al lancio di due dadi, (supposti distinti)

In generale possiamo parlare di spazio di probabilità finito se è assegnata una famiglia di eventi

$$\mathcal{F} = \{A_i : i = 1..N\}$$

soddisfacente le seguenti condizioni

- $\emptyset \in \mathcal{F}$
- $A_1, A_2 \in \mathcal{F} \implies A_1 \cup A_2 \in \mathcal{F}$

ed è assegnata una funzione

$$\mathcal{P} : \mathcal{F} \rightarrow \mathbb{R}$$

che associa ad ogni $A \in \mathcal{F}$ un valore reale $\mathcal{P}(A)$, che chiamiamo probabilità che l'evento accada, soddisfacente le seguenti proprietà:

- $\mathcal{P}(A) \geq 0$
- **se $\mathcal{U} = \bigcup_{A \in \mathcal{F}} A \in \mathcal{F}$, si ha $\mathcal{P}(\mathcal{U}) = 1$**
- **se $A_1, A_2 \in \mathcal{F}$ sono tali che $A_1 \cap A_2 = \emptyset$ allora**

$$\mathcal{P}(A_1 \cup A_2) = \mathcal{P}(A_1) + \mathcal{P}(A_2)$$

Osserviamo che, dal momento che la famiglia dei possibili eventi è finita, avremo che

$$\mathcal{U} = \bigcup_{A \in \mathcal{F}} A \in \mathcal{F}$$

$\mathcal{U}$ rappresenta l'evento certo ed anche l'ambiente in cui si individuano gli eventi.

Per assegnare uno spazio di probabilità, quindi, è necessario assegnare un insieme $\mathcal{U}$ una famiglia $\mathcal{F}$ di sottoinsiemi di $\mathcal{U}$ ed una funzione P definita su $\mathcal{F}$ a valori in $\mathbb{R}$.

Ci riferiremo quindi ad uno spazio di probabilità come ad un terna $(\mathcal{U}, \mathcal{F}, \mathcal{P})$

Nel caso dell'esempio precedente la famiglia $\mathcal{F}$ è costituita da ciascuno degli eventi (i, j) che abbiamo rappresentato nel piano con i punti $A_{i,j}$ e dalle unioni finite di essi.

Ad esempio

$$B = \{A_{1,3}, A_{4,1}\} = A_{1,3} \cup A_{4,1}$$

è l'evento che accade se almeno una tra le uscite $A_{1,3}$ e $A_{4,1}$ accade.

Poichè le uscite del lancio dei dadi sono ritenute equiprobabili possiamo affermare che $\mathcal{P}(A_{i,j}) = \frac{1}{36}$ per ogni possibile uscita $A_{i,j}$ e quindi, ad esempio,

$$\mathcal{P}(B) = \mathcal{P}(A_{1,3} \cup A_{4,1}) = \mathcal{P}(A_{1,3}) + \mathcal{P}(A_{4,1}) = \frac{1}{36} + \frac{1}{36} = \frac{2}{36}$$

Ovviamente

$$\mathcal{P}(U) = \mathcal{P}\left(\bigcup_{A \in \mathcal{F}} A\right) = 36 \frac{1}{36} = 1$$

Come abbiamo già osservato, ci riferiamo ad uno spazio di probabilità come ad una terna $(\mathcal{U}, \mathcal{F}, \mathcal{P})$ in cui $\mathcal{U}$ è lo spazio di probabilità, $\mathcal{F}$ è la famiglia di tutti gli eventi e $\mathcal{P}$ è la misura di probabilità definita su $\mathcal{U}$.

Si possono provare i seguenti risultati:

- $0 \leq \mathcal{P}(A) \leq 1$ per ogni $A \in \mathcal{F}$
- $\mathcal{P}(\emptyset) = 0$
- Se $A_1, A_2 \in \mathcal{F}$ $A_1 \subset A_2$ allora

$$\mathcal{P}(A_1) \subset \mathcal{P}(A_2)$$

ed inoltre

$$\mathcal{P}(A_2 \subset A_1) = \mathcal{P}(A_2) - \mathcal{P}(A_1)$$

- $\mathcal{P}(A^c) = 1 - \mathcal{P}(A)$ per ogni $A \in \mathcal{F}$
- $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$
- Se $B = \bigcup_{i=1}^N A_i$ e se $A_i \cap A_j = \emptyset$ allora

$$\mathcal{P}(B) = \sum_{i=1}^N \mathcal{P}(A_i)$$

- $\mathcal{P}(A) = \mathcal{P}(A \cap B) + \mathcal{P}(A \cap B^c)$
- se $A_i \cap A_j = \emptyset$, ed esiste i_0 tale che $A \subset A_{j_0}$, allora

$$(6.5) \quad \boxed{\mathcal{P}(A) = \mathcal{P}(A \cap A_1) + \mathcal{P}(A \cap A_2) + \dots \mathcal{P}(A \cap A_N)}$$

DEFINIZIONE 6.1. Se $A, B \in \mathcal{F}$ definiamo probabilità di A condizionata a B e la denotiamo con $\mathcal{P}(A|B)$ il valore

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

La probabilità di A condizionata a B , $\mathcal{P}(A|B)$, definisce la probabilità di accadimento di A nel caso sia accaduto B

Naturalmente si ha

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

Nel caso in cui

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

diciamo che A e B sono eventi indipendenti, (la probabilità di accadimento di A non è cambiata dal fatto che B è accaduto).

In tal caso si ha

$$\mathcal{P}(A \cap B) = \mathcal{P}(A)\mathcal{P}(B)$$

Si può dimostrare che

Se $A_i \cap A_j = \emptyset$, ed esiste i_0 tale che $A \subset A_{j_0}$, allora

$$\mathcal{P}(A) = \mathcal{P}(A_1)\mathcal{P}(A|A_1) + \mathcal{P}(A_2)\mathcal{P}(A|A_2) + \dots \mathcal{P}(A_N)\mathcal{P}(A|A_N)$$

e ne segue il seguente teorema di Bayes

TEOREMA 6.1. - di Bayes - Se $A_1, A_2, \dots, A_N \in \mathcal{F}$ sono eventi tali che $A_i \cap A_j = \emptyset$ e $\bigcup_{i=1}^N A_i$ allora

$$\mathcal{P}(A_k|B) = \frac{\mathcal{P}(A_k)\mathcal{P}(A|A_k)}{\sum_{i=1}^N \mathcal{P}(A_i)\mathcal{P}(A|A_i)}$$

Per capire perchè il teorema sia vero, consideriamo il caso in cui $N = 2$; allora

- $A_1 \cup A_2 \supset A$
- $A_1 \cap A_2 = \emptyset$
- $\mathcal{P}(B) = \mathcal{P}(B \cap (A_1 \cup A_2)) = \mathcal{P}(B \cap A_1) + \mathcal{P}(B \cap A_2)$

e

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

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

da cui

$$\mathcal{P}(A_1|B)\mathcal{P}(B) = \mathcal{P}(B \cap A_1) = \mathcal{P}(B|A_1)\mathcal{P}(A_1)$$

Ne segue che

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

2. Qualche applicazione della regola di Bayes.

2.1. Si consideri una popolazione in cui è diffusa una malattia e si supponga di voler sottoporre l'intera popolazione ad un test con lo scopo di determinare per ciascuna persona se è affetta o no dalla malattia.

Una situazione di questo genere si presenterebbe se si considerasse l'opportunità di procedere ad uno screening di massa per rivelare, ad esempio, la diffusione dell'AIDS.

In tal caso sarebbe ragionevole supporre che:

- Nella popolazione 1 persona su 1000 è affetta dalla malattia (**incidenza della malattia**).
- Il test usato per rivelare la malattia è positivo (e quindi indica la presenza della malattia) in 95 casi per ogni 100 persone malate esaminate (**sensitività del test**).
- Il test usato per rivelare la malattia è positivo (e quindi indica la presenza della malattia) in 1 caso per ogni 100 persone sane esaminate (**specificità del test**).

Indichiamo con

- A l'evento infetto
- A' l'evento sano
- T_p l'evento test positivo
- T_n l'evento test negativo

Indichiamo inoltre con

- a l'incidenza della malattia (nel caso prima citato $a = 1/1000 = 0.001$)
- p la sensitività del test (nel caso prima citato $p = 95/100 = 0.95$)
- q la specificità del test (nel caso prima citato $q = 1/100 = 0.01$)

Possiamo facilmente calcolare che

$$\begin{aligned}
\mathcal{P}(A) &= a & \mathcal{P}(A') &= 1 - a \\
\mathcal{P}(T_p|A) &= p & \mathcal{P}(T_n|A) &= 1 - p \\
\mathcal{P}(T_p|A') &= q & \mathcal{P}(T_n|A') &= 1 - q
\end{aligned}$$

da cui

$$\begin{aligned}
\mathcal{P}(T_p) &= \mathcal{P}(T_p \cap A) + \mathcal{P}(T_p \cap A') = \\
&= \mathcal{P}(T_p|A)\mathcal{P}(A) + \mathcal{P}(T_p|A')\mathcal{P}(A') = pa + q(1 - a)
\end{aligned}$$

Pertanto, applicando il teorema di Bayes

$$\mathcal{P}(A|T_p) = \frac{\mathcal{P}(T_p|A)\mathcal{P}(A)}{\mathcal{P}(T_p|A)\mathcal{P}(A) + \mathcal{P}(T_p|A')\mathcal{P}(A')} = \frac{pa}{pa + q(1 - a)}$$

Se utilizziamo i valori di incidenza sensitività e specificità prima introdotti, possiamo ricavare che:

$$\mathcal{P}(A|T_p) = \frac{pa}{pa + q(1 - a)} = 0.0874$$

e questo significa che la probabilità che una persona in cui il test ha dato esito positivo sia effettivamente infetto è inferiore al 9%.

Se ne deduce la poca convenienza ad effettuare uno screening di massa.

Vale la pena osservare che neppure test più precisi consentono di raggiungere risultati più significativi.

Infatti se supponiamo di migliorare il test in modo da avere la certezza di individuare la malattia nelle persone infette, cioè se supponiamo la sensitività $p = 100/100 = 1$, mantenendo $a = 0.001$ e $q = 0.01$, otteniamo

$$\mathcal{P}(A|T_p) = \frac{pa}{pa + q(1 - a)} = 0.091$$

In realtà aumentando la sensitività, la specificità necessariamente diminuisce; se supponiamo che

$$a = 0.001 \quad , \quad p = 1 \quad , \quad q = 0.002$$

la situazione peggiora e si ha addirittura

$$\mathcal{P}(A|T_p) = \frac{pa}{pa + q(1 - a)} = 0.05$$

Solo migliorando la specificità senza peggiorare la sensitività si ottiene qualcosa di meglio; infatti se

$$a = 0.001 \quad , \quad p = 0.95 \quad , \quad q = 1/300 = 0.0034$$

si ha

$$\mathcal{P}(A|T_p) = \frac{pa}{pa + q(1 - a)} = 0.22$$

Un miglioramento decisivo nell'affidabilità dei risultati si ottiene invece nel caso in cui l'incidenza della malattia sia alta; infatti per

$$a = 0.01 \quad , \quad p = 0.95 \quad , \quad q = 0.01$$

si ha

$$\mathcal{P}(A|T_p) = \frac{pa}{pa + q(1-a)} = 0.5$$

e addirittura

$$\mathcal{P}(A|T_p) = \frac{pa}{pa + q(1-a)} = 0.9$$

nel caso in cui l'incidenza salga ad $a = .1$.

2.2. Un secondo interessante esempio si ha considerando la seguente situazione.

Si supponga che un capo di governo debba prendere una decisione di politica economica e disponga di tre consiglieri che identifichiamo con

$$A \quad B \quad C$$

Si supponga di avere una stima della loro affidabilità descritta come segue:

- la probabilità che il consigliere A abbia un'opinione corretta è $1/6$,
- la probabilità che il consigliere B abbia un'opinione corretta è $1/3$,
- la probabilità che il consigliere C abbia un'opinione corretta è $1/2$.

Indichiamo con

A_r l'evento il consigliere A ha espresso un'opinione corretta,

B_r l'evento il consigliere B ha espresso un'opinione corretta,

C_r l'evento il consigliere C ha espresso un'opinione corretta.

Esprimiamo l'affidabilità dei consiglieri assegnando agli eventi A_r, B_r e C_r opportuni valori di probabilità

$$\mathcal{P}(A_r) = 1/6 \quad \mathcal{P}(B_r) = 1/3 \quad \mathcal{P}(C_r) = 1/2$$

Il capo di governo chiede ai suoi consiglieri una previsione sulle conseguenze che la sua decisione avrà sul tasso di disoccupazione, a distanza di un anno.

Al quesito i tre consiglieri rispondono secondo i valori descritti nella tabella che segue, dove nella colonna D sono riportate le probabilità che la disoccupazione diminuisca, nella colonna S che rimanga stabile e nella colonna I che il tasso di disoccupazione aumenti.

	D	S	I
A	1/10	1/10	8/10
B	6/10	2/10	2/10
C	2/10	6/10	2/10

TABELLA 6.1. Previsioni dei consiglieri sul tasso di disoccupazione.

Trascorso un anno si rileva che il tasso di disoccupazione è aumentato; alla luce di questo fatto le percentuali di affidabilità dei tre consiglieri vanno riviste e ciò può essere fatto ridefinendo la probabilità di A_r , B_r e C_r alla luce del fatto che l'evento I si è verificato.

Dovremo, in altre parole valutare

$$\begin{aligned}\mathcal{P}(A_r|I) &= \frac{\mathcal{P}(I|A_r)\mathcal{P}(A_r)}{\mathcal{P}(I)} \\ \mathcal{P}(B_r|I) &= \frac{\mathcal{P}(I|B_r)\mathcal{P}(B_r)}{\mathcal{P}(I)} \\ \mathcal{P}(C_r|I) &= \frac{\mathcal{P}(I|C_r)\mathcal{P}(C_r)}{\mathcal{P}(I)}\end{aligned}$$

ma

$$\begin{aligned}\mathcal{P}(I) &= \mathcal{P}(I \cap A_r) + \mathcal{P}(I \cap B_r) + \mathcal{P}(I \cap C_r) = \\ &= \mathcal{P}(I|A_r)\mathcal{P}(A_r) + \mathcal{P}(I|B_r)\mathcal{P}(B_r) + \mathcal{P}(I|C_r)\mathcal{P}(C_r)\end{aligned}$$

Ora, la probabilità che avvenga I ammesso che un consigliere abbia ragione si legge nell'ultima colonna della tabella, per cui

$$\begin{aligned}\mathcal{P}(I|A_r) &= 8/10 \\ \mathcal{P}(I|B_r) &= 2/10 \\ \mathcal{P}(I|C_r) &= 2/10\end{aligned}$$

per cui

$$P(I) = \frac{8}{10} \frac{1}{6} + \frac{2}{10} \frac{1}{3} + \frac{2}{10} \frac{1}{2} = \frac{3}{10}$$

Ne viene che le nuove percentuali di affidabilità sono

$$\begin{aligned}\mathcal{P}(A_r|I) &= \frac{\mathcal{P}(I|A_r)\mathcal{P}(A_r)}{\mathcal{P}(I)} = \frac{\frac{8}{10} \frac{1}{6}}{\frac{3}{10}} = \frac{4}{9} \\ \mathcal{P}(B_r|I) &= \frac{\mathcal{P}(I|B_r)\mathcal{P}(B_r)}{\mathcal{P}(I)} = \frac{\frac{2}{10} \frac{1}{3}}{\frac{3}{10}} = \frac{2}{9} \\ \mathcal{P}(C_r|I) &= \frac{\mathcal{P}(I|C_r)\mathcal{P}(C_r)}{\mathcal{P}(I)} = \frac{\frac{2}{10} \frac{1}{2}}{\frac{3}{10}} = \frac{1}{3}\end{aligned}$$

3. Qualche richiamo di calcolo combinatorio.

Per studiare un po' di probabilità discreta è utile conoscere qualche elemento di calcolo combinatorio.

Il calcolo combinatorio si occupa di stabilire il numero delle possibili uscite di semplici esperimenti; si fonda essenzialmente sul principio seguente:

Se un esperimento ha n_1 possibili esiti, un secondo esperimento ha n_2 possibili esiti, un terzo esperimento ha n_3 possibili esiti, allora il numero dei possibili esiti della sequenza dei tre esperimenti è

$$n_1 n_2 n_3$$

Le più comuni conseguenze di questo principio portano a un certo numero di definizioni che descriviamo brevemente.

3.1. Disposizioni di n elementi a k a k . Parliamo di **disposizioni** (o anche, se $k = n$, di **permutazioni**) di n elementi a k a k quando consideriamo i gruppi di k elementi scelti tra gli n dati.

Riteniamo due gruppi distinti se differiscono per un elemento o per l'ordine con cui gli elementi sono scelti.

Indichiamo con

$${}_n D_k$$

il numero delle disposizioni di n elementi a k a k .

Poichè per il primo elemento di ciascun gruppo abbiamo n scelte, per il secondo ne abbiamo $(n - 1)$ per il terzo ne abbiamo $(n - 2)$ e così via, possiamo calcolare che

$${}_n D_k = n(n - 1)(n - 2)(n - 3) \dots (n - (k - 1)) = \frac{n!}{(n - k)!}$$

Il numero delle disposizioni di n elementi ad n ad n , cioè delle permutazioni, risulta

$$P_n = {}_n D_n = n!$$

Qualora gli n elementi da cui si sceglie presentino sottogruppi di elementi uguali, siano $n_1, n_2, n_3, \dots, n_k$ le permutazioni risultano in numero di

$$\frac{n!}{n_1! n_2! n_3! \dots n_k!}$$

3.2. Combinazioni di n elementi a k a k . Parliamo di **combinazioni** di n elementi a k a k quando consideriamo gruppi di k elementi scelti tra gli n dati senza distinguerli in base all'ordine degli elementi.

Indichiamo con

$${}_nC_k$$

il numero delle combinazioni di n elementi a k a k .

Poichè un gruppo di k elementi ammette $k!$ diversi modi possiamo calcolare che

$${}_nC_k = \frac{{}_nD_k}{k!} = \frac{n(n-1)(n-2)(n-3)\dots(n-(k-1))}{k!} = \frac{n!}{k!(n-k)!}$$

Il numero ${}_nC_k$ si chiama coefficiente binomiale e si indica con

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

Ricordiamo che i coefficienti binomiali possono essere ricavati dal triangolo di Tartaglia e che trovano una importante applicazione nella formula del binomio di Newton.

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

3.3. Campioni ordinati. È anche utile ricordare qualche formula per stimare il numero di possibili campioni estratti da una popolazione.

Per aiutarci assimiliamo la popolazione ad un'un'urna piena di palline e l'estrazione degli elementi del campione all'estrazione delle palline dall'urna.

Possiamo operare un campionamento con ripetizione estraendo una pallina, osservandola e rimettendola nell'urna dopo aver annotato l'informazione relativa.

In tal caso, se operiamo k estrazioni, avremo

$$\overbrace{nnn\dots n}^{k \text{ volte}} = n^k$$

possibili uscite in quanto per ogni elemento estratto avremo sempre n possibili scelte.

Possiamo anche operare un campionamento senza ripetizione, estraendo, osservando e non rimettendo la pallina nell'urna; in tal caso per la prima estrazione avremo n possibilità, per la seconda $n-1$, per la terza $n-2$ e così via. Pertanto in questo caso avremo

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

possibili uscite.

4. Le variabili aleatorie

Se è assegnato uno spazio di probabilità $(\mathcal{U}, \mathcal{F}, \mathcal{P})$ diciamo che ξ è una variabile aleatoria definita su $(\mathcal{U}, \mathcal{F}, \mathcal{P})$ se è data una funzione

$$\xi : \mathcal{U} \rightarrow \mathbb{R}$$

Nel caso del lancio di due dadi la terna $(\mathcal{U}, \mathcal{F}, \mathcal{P})$ è definita da

- $\mathcal{U}$ è l'insieme delle 36 coppie di valori (i, j)
- $\mathcal{F}$ è la famiglia di tutti i sottoinsiemi di $\mathcal{U}$
- $\mathcal{P}$ è definita da $\mathcal{P}(A_{i,j}) = \frac{1}{36}$ ed inoltre se $A \in \mathcal{F}$ è un sottoinsieme costituito da k elementi possiamo definire $\mathcal{P}(A) = k \frac{1}{36}$

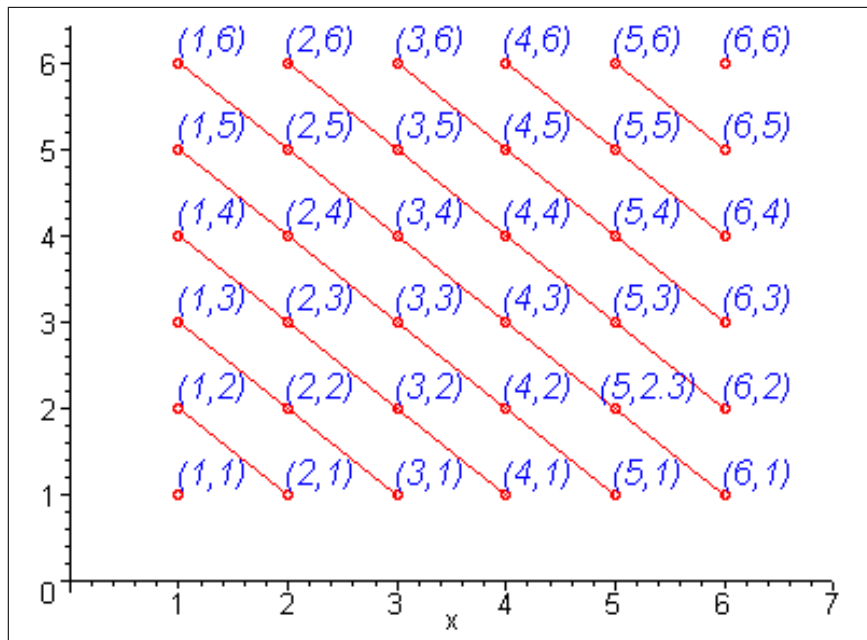


FIGURA 6.3. Eventi che forniscono lo stesso valore per ξ , (congiunti da un segmento)

Possiamo costruire un esempio di variabile aleatoria sullo spazio di probabilità associato al lancio di due dadi assegnando ad ogni uscita $(i, j) \in \mathcal{U}$ il valore ottenuto sommando i punteggi; in altre parole

$$\xi(A_{i,j}) = i + j$$

Possiamo osservare che i punti (eventi) che nella figura 6.3 sono congiunti da un segmento di retta forniscono lo stesso valore per la variabile aleatoria ξ e possiamo riassumere i valori assunti dalla variabile aleatoria ξ nella seguente tabella, dove all'incrocio della riga i -esima e della colonna j -esima possiamo leggere il valore $i + j$ assunto da ξ .

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

TABELLA 6.2. I valori assunti dalla Variabile aleatoria ξ

Possiamo ora definire alcuni elementi caratteristici di una variabile aleatoria:

DEFINIZIONE 6.2. Se ξ è una variabile aleatoria discreta definita su uno spazio di probabilità $(\mathcal{U}, \mathcal{F}, \mathcal{P})$ e se

$$\mathcal{U} = \{A_1, A_2, \dots, A_n\}$$

- la **media** μ di ξ è definita da

$$\mu = \sum_i \xi(A_i) \mathcal{P}(A_i)$$

- la **varianza** σ^2 di ξ è definita da

$$\sigma^2 = \text{Var}(\xi) = \sum_i (\xi(A_i) - \mu)^2 \mathcal{P}(A_i)$$

- lo **scarto quadratico medio** di ξ è definito da

$$\sigma = \sqrt{\sigma^2}$$

- la **moda** M di ξ è definita da

$$M = \xi(A_i)$$

dove A_i è tale che

$$\mathcal{P}(A_i) = \max_j \mathcal{P}(A_j)$$

- la **mediana** m di ξ è definita da

$$\mathcal{P}(\xi \leq m) = \mathcal{P}(\xi \geq m)$$

- il **momento di ordine** k μ_k di ξ è definito da

$$\mu_k = \sum_i (\xi(A_i) - \mu)^k \mathcal{P}(A_i)$$

- il **momento di ordine k , rispetto all'origine** μ'_k di ξ è definito da

$$\mu'_k = \sum_i (\xi(A_i))^k \mathcal{P}(A_i)$$

DEFINIZIONE 6.3. Chiamiamo **funzione di distribuzione di probabilità della variabile aleatoria ξ** la funzione

$$\varphi : \mathcal{U} \rightarrow \mathbb{R}$$

definita da

$$\varphi(i) = \mathcal{P}(\xi = A_i)$$

Se φ è la funzione distribuzione di probabilità di una variabile aleatoria ξ avremo che

$$\begin{aligned} \mathcal{P}(n \leq \xi \leq m) &= \sum_{i=n}^m \varphi(i) \\ \mathcal{P}(\xi \leq n) &= \sum_{i \leq n} \varphi(i) \end{aligned}$$

DEFINIZIONE 6.4. Se ξ è una variabile aleatoria discreta la cui densità di probabilità è φ , definiamo **speranza matematica di ξ**

$$E(\xi) = \sum_i \xi(A_i) \varphi(i) = \sum_i \xi(A_i) \mathcal{P}(\xi = A_i) = \mu$$

DEFINIZIONE 6.5. Se ξ è una variabile aleatoria discreta la cui densità di probabilità è φ e se $g : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione, definiamo una nuova variabile aleatoria che indichiamo con $g(\xi)$ mediante la funzione distribuzione di probabilità

$$\psi(i) = g(\xi(A_i)) \varphi(i)$$

Definiamo inoltre

$$E(g(\xi)) = \sum_i g(\xi(A_i)) \varphi(i) =$$

Osserviamo che si ha

$$\begin{aligned} \sigma &= E((\xi - \mu)^2) \\ \mu_k &= E((\xi - \mu)^k) \\ \mu'_k &= E(\xi^k) \end{aligned}$$

DEFINIZIONE 6.6. Se ξ è una variabile aleatoria discreta la cui densità di probabilità è φ , definiamo **funzione generatrice dei momenti di ξ** la

$$M_\xi(t) = E(e^{t\xi}) = \sum_i e^{t\xi(A_i)} \varphi(i) =$$

4.1. Un esempio. Consideriamo il solito esempio dei dadi e la solita variabile ξ che associa ad ogni uscita in $\mathcal{U}$ la somma dei punteggi, possiamo vedere che

$$\begin{aligned}\varphi(2) &= \mathcal{P}(\xi = 2) = \frac{1}{36} = \mathcal{P}(\xi = 12) = \varphi(12) \\ \varphi(3) &= \mathcal{P}(\xi = 3) = \frac{2}{36} = \mathcal{P}(\xi = 11) = \varphi(11) \\ \varphi(4) &= \mathcal{P}(\xi = 4) = \frac{3}{36} = \mathcal{P}(\xi = 10) = \varphi(10) \\ \varphi(5) &= \mathcal{P}(\xi = 5) = \frac{4}{36} = \mathcal{P}(\xi = 9) = \varphi(9) \\ \varphi(6) &= \mathcal{P}(\xi = 6) = \frac{5}{36} = \mathcal{P}(\xi = 8) = \varphi(8) \\ \varphi(7) &= \mathcal{P}(\xi = 7) = \frac{6}{36}\end{aligned}$$

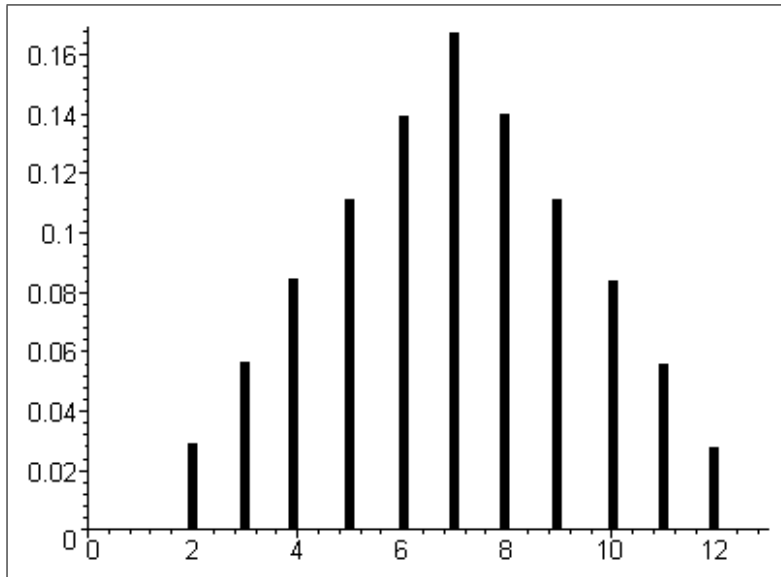


FIGURA 6.4. Distribuzione di probabilità della variabile aleatoria che restituisce il punteggio ottenuto lanciando due dadi.

Possiamo rappresentare questa distribuzione di probabilità mediante il grafico in figura 6.4 (scatter plot) oppure mediante il grafico in figura 6.5.

Quest'ultimo grafico si chiama istogramma ed è quello che più frequentemente viene adottato per rappresentare la funzione di distribuzione di una variabile aleatoria discreta.

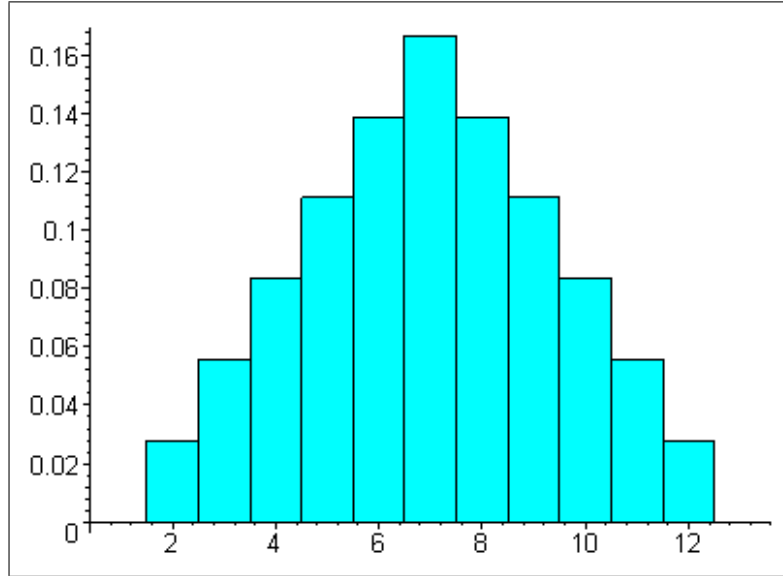


FIGURA 6.5. Istogramma relativo alla variabile aleatoria che restituisce il punteggio ottenuto lanciando due dadi.

Osservando il grafico possiamo osservare che il valore che compare con maggior frequenza è il 7, che pertanto è la moda della variabile ξ .

Possiamo calcolare la probabilità $\mathcal{P}(\xi \leq 7)$ sommando le aree dei rettangoli mostrati in figura (si osservi che la base dei rettangoli è lunga 1) ed otteniamo che $\mathcal{P}(\xi \leq 7) = \frac{21}{36}$.

In modo simile possiamo calcolare $\mathcal{P}(\xi \geq 7) = \frac{21}{36}$ e possiamo concludere che la mediana della variabile aleatoria ξ è $m = 7$.

Un semplice calcolo mostra che la varianza di ξ è $\sigma^2 = \frac{35}{6}$.

Abbiamo già notato che i rettangoli che costituiscono l'istogramma hanno base lunga 1; ciò permette di definire la funzione di distribuzione di ξ come

$$(6.6) \quad \varphi(x) = \begin{cases} \mathcal{P}(\xi = \eta) & \text{per } \eta - 0.5 \leq x \leq \eta + 0.5 \text{ } \eta = 2, 3, 4, \dots, 12 \\ 0 & \text{altrove} \end{cases}$$

Con tale definizione si vede che φ è definita su tutto $\mathbb{R}$ ed inoltre si vede che

$$\int_{-\infty}^{+\infty} \varphi(t) dt = \sum_i \varphi(i) = 1$$

essendo la somma estesa a tutti i valori η che variabile discreta ξ può assumere.

5. Variabili aleatorie continue

Talvolta tuttavia non è possibile considerare uno spazio di probabilità discreto, finito o numerabile.

Ciò accade ad esempio quando si considera il problema di scegliere un numero a caso compreso tra 0 ed 1.

Infatti la probabilità di estrarre, ad esempio, il valore 0.3 non si può calcolare considerando il rapporto tra casi favorevoli, uno solo, e casi possibili, infiniti non numerabili.

Anche la definizione di media e varianza presentano qualche problema in quanto occorre definire come si intende procedere per calcolare la somma di un numero infinito, non numerabile, di addendi.

Per chiarire la questione possiamo osservare che, se è difficile definire la probabilità che la variabile aleatoria ξ il cui valore è il numero scelto a caso in $[0, 1]$ assuma il valore x , è invece naturale definire la probabilità che $\xi \in [x, x + h]$.

In tal caso infatti possiamo identificare i casi favorevoli con un segmento di lunghezza h e la totalità dei casi con l'intero intervallo $[0, 1]$ che risulta ovviamente di lunghezza 1.

Pertanto

$$\mathcal{P}(x \leq \xi \leq x + h) = \frac{h}{1}$$

Ricordando il significato di somma dell'integrale, possiamo definire la funzione distribuzione di probabilità della variabile aleatoria ξ come la funzione continua φ tale che

$$(6.7) \quad \mathcal{P}(x \leq \xi \leq x + h) = h = \int_x^{x+h} \varphi(t) dt$$

per ogni $x \in [0, 1]$ e per ogni h abbastanza piccolo.

Ne deduciamo che

$$(6.8) \quad \frac{1}{h} \int_x^{x+h} \varphi(t) dt = 1$$

e, passando al limite per $h \rightarrow 0$, poichè abbiamo supposto φ continua,

$$\varphi(x) = 1$$

Da quanto abbiamo detto appare ragionevole che, nel caso di una variabile aleatoria continua ξ , non è significativo definire

$$(6.9) \quad \mathcal{P}(\xi = x)$$

mentre è naturale definire

$$(6.10) \quad \mathcal{P}(x_0 \leq \xi \leq x_1) = \int_{x_0}^{x_1} \varphi(t) dt$$

dove φ è la funzione di distribuzione di probabilità di ξ .

Pertanto supporremo nota una variabile aleatoria continua ξ se è nota la sua funzione di distribuzione di probabilità φ .

Una funzione $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, continua, è la funzione di distribuzione di probabilità di una variabile aleatoria se

- $\varphi(t) \geq 0$
- **per ogni $t \in \mathbb{R}$**
- $\int_{-\infty}^{+\infty} \varphi(t) dt = 1$

In tal caso si ha

$$\mathcal{P}(\xi \leq x) = \int_{-\infty}^x \varphi(t) dt$$

$$\mathcal{P}(x_0 \leq \xi \leq x_1) = \int_{x_0}^{x_1} \varphi(t) dt$$

La funzione

$$F(x) = \mathcal{P}(\xi \leq x) = \int_{-\infty}^x \varphi(t) dt$$

si chiama distribuzione cumulativa di probabilità della variabile aleatoria ξ di densità di probabilità φ .

Osserviamo che ad ogni variabile aleatoria discreta (finita) si può associare una variabile aleatoria continua la cui densità è una funzione costante a tratti, nulla al di fuori di un insieme limitato (nel caso in cui la variabile sia discreta e finita).

Per chiarire il concetto consideriamo la variabile aleatoria ξ che fornisce il punteggio ottenuto nel lancio di due dadi; in tal caso la funzione distribuzione di probabilità φ può essere definita come segue

$$\varphi(x) = \begin{cases} 0 & x < 1.5 \\ k/36 & k + .5 \leq x < k + 1.5 & k = 1..5 \\ 6/36 & 6.5 \leq x < 7.5 \\ (12 - k)/36 & k + .5 \leq x < k + 1.5 & k = 7..11 \\ 0 & x \geq 12.5 \end{cases}$$

Come nel caso delle variabili aleatorie discrete possiamo definire

DEFINIZIONE 6.7. Se ξ è una variabile aleatoria continua che ha densità di probabilità φ ,

- la **media** μ di ξ è definita da

$$\mu = E(\xi) = \int_{-\infty}^{+\infty} x\varphi(x)dx$$

- la **varianza** σ^2 di ξ è definita da

$$\sigma^2 = \text{Var}(\xi) = E((\xi - \mu)^2) = \int_{-\infty}^{+\infty} (x - \mu)^2 \varphi(x)dx$$

- lo **scarto quadratico medio** di ξ è definito da

$$\sigma = \sqrt{\sigma^2}$$

- la **moda** M di ξ è definita da

$$M = \sup_{x \in \mathbb{R}} \varphi(x)$$

- la **mediana** m di ξ è definita da

$$\mathcal{P}(\xi \leq m) = \int_{-\infty}^m \varphi(x)dx = \int_m^{+\infty} \varphi(x)dx = \mathcal{P}(\xi \geq m)$$

- il **momento di ordine** k μ_k di ξ è definito da

$$\mu_k = E((\xi - \mu)^k) = \int_{-\infty}^{+\infty} (x - \mu)^k \varphi(x)dx$$

- il **momento di ordine** k , **rispetto all'origine** μ'_k di ξ è definito da

$$\mu_k = E(\xi^k) = \int_{-\infty}^{+\infty} x^k \varphi(x)dx$$

DEFINIZIONE 6.8. Se ξ è una variabile aleatoria continua e se $f : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione possiamo definire la variabile aleatoria $f(\xi)$ assegnandone la funzione distribuzione di probabilità definita da

$$\psi(t) = f(t)\varphi(t)$$

In tal modo

$$\mathcal{P}(x_0 \leq f(\xi) \leq x_1) = \int_{x_0}^{x_1} f(t)\varphi(t)dt$$

$$E(f(\xi)) = \int_{-\infty}^{+\infty} f(t)\varphi(t)dt$$

DEFINIZIONE 6.9. Se ξ è una variabile aleatoria continua la cui densità di probabilità è φ , definiamo **funzione generatrice dei momenti** di ξ la

$$M_\xi(t) = E(e^{t\xi}) = \int_{-\infty}^{+\infty} e^{tx} \varphi(x)dx$$

5.1. Le variabili aleatorie discrete infinite. Un caso che si colloca a metà tra quello delle variabili aleatorie discrete finite ed il caso delle variabili aleatorie continue è il caso delle variabili aleatorie discrete che assumono una quantità infinita numerabile di valori.

Il caso delle variabili aleatorie infinite ci suggerisce come comportarci in questo caso.

Se lo spazio $\mathcal{U}$ degli eventi è numerabile, allora potremo scrivere che

$$\mathcal{U} = \{A_i, i \in \mathbb{N}\}$$

e sarà sufficiente definire la probabilità di ciascun evento

$$\mathcal{P}(A_i) = p_i$$

con la condizione

$$\sum_{i=1}^{+\infty} p_i = 1$$

Media, varianza, funzione di distribuzione possono essere definiti mediante le

DEFINIZIONE 6.10. Se ξ è una variabile aleatoria discreta numerabile definiamo

- la **media** μ di ξ come

$$(6.11) \quad \mu = E(\xi) = \sum_{i=1}^{+\infty} i p_i$$

- la **varianza** σ^2 di ξ come

$$(6.12) \quad \sigma^2 = \text{Var}(\xi) = E((\xi - \mu)^2) = \sum_{i=1}^{+\infty} (i - \mu)^2 p_i$$

- lo **scarto quadratico medio** o deviazione standard di ξ

$$\sigma = \sqrt{\sigma^2}$$

- il **momento k -esimo** μ_k di ξ come

$$\mu_k = \sum_{i=1}^{+\infty} (i - \mu)^k p_i$$

- il **momento k -esimo rispetto all'origine** μ'_k di ξ come

$$\mu'_k = \sum_{i=1}^{+\infty} i^k p_i$$

- la **funzione di distribuzione** φ di ξ come

$$(6.13) \quad \varphi(n) = \mathcal{P}(\xi = A_i)$$

e si ha

$$(6.14) \quad \mathcal{P}(\xi \leq n) = \sum_{i=1}^n \mathcal{P}(\xi = A_i) \quad n \in \mathbb{N}$$

Naturalmente possiamo estendere la definizione di φ ad $\mathbb{R}$ definendola costante sugli intervalli del tipo $[k, k+1]$ e ponendola uguale a 0 prima di 1.

Inoltre

DEFINIZIONE 6.11. Se ξ è una variabile aleatoria discreta infinita la cui densità di probabilità è φ , e se $f : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione possiamo definire la variabile aleatoria $f(\xi)$ assegnandone la funzione distribuzione di probabilità definita da

$$\psi(k) = f(k)\varphi(k) = f(k)p_k$$

Quindi

$$E(f(\xi)) = \sum_{i=1}^{+\infty} f(i)\varphi(i)$$

Definiamo anche definiamo **funzione generatrice dei momenti** di ξ la

$$M_\xi(t) = E(e^{t\xi}) = \sum_{k=1}^{+\infty} e^{kt}\varphi(k) =$$

6. Distribuzioni di probabilità doppie

Siano $(\mathcal{U}_1, \mathcal{F}_1, \mathcal{P}_1)$ e $(\mathcal{U}_2, \mathcal{F}_2, \mathcal{P}_2)$ due spazi di probabilità consideriamo la variabile aleatoria che indichiamo con (ξ, η) definita sullo spazio $\mathcal{U}_1 \times \mathcal{U}_2$ mediante la

$$F(x, y) = \mathcal{P}(\xi \leq x, \eta \leq y) = \int_{-\infty}^x \left(\int_{-\infty}^y f(t, s) ds \right) dt$$

F è la distribuzione cumulativa di probabilità della variabile (ξ, η) ed f è la sua funzione distribuzione di probabilità

Se f è continua possiamo affermare che

$$\frac{\partial^2 F}{\partial x \partial y} = f(x, y)$$

Naturalmente devono essere verificate le seguenti condizioni:

•

$$f(x, y) \geq 0$$

•

$$\int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(t, s) ds \right) dt = 1$$

Inoltre se

$$F_1(x) = \mathcal{P}(\xi \leq x) = \int_{-\infty}^x \left(\int_{-\infty}^{+\infty} f(t, s) ds \right) dt$$

$$F_2(y) = \mathcal{P}(\eta \leq y) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^y f(t, s) dt \right) ds = \int_{-\infty}^y \left(\int_{-\infty}^{+\infty} f(t, s) dt \right) ds$$

F_1 ed F_2 sono le distribuzioni cumulative delle variabili aleatorie ξ e η , rispettivamente le cui funzioni di distribuzione sono date da

$$\varphi(t) = \int_{-\infty}^{+\infty} f(t, s) ds$$

$$\psi(s) = \int_{-\infty}^{+\infty} f(t, s) dt$$

Nel caso in cui le variabili aleatorie ξ e η siano indipendenti, allora (ξ, η) ha una distribuzione di probabilità

$$f(t, s) = \varphi(t)\psi(s)$$

dove φ e ψ sono le funzioni di distribuzione di ξ e η , rispettivamente.

È utile ricordare che la probabilità della variabile aleatoria ξ condizionata alla variabile aleatoria η si può definire mediante la

$$\mathcal{P}(\xi \leq x | \eta \leq y) = \int_{-\infty}^x \left(\int_{-\infty}^y \frac{f(t, s)}{\psi(s)} ds \right) dt$$

per cui $\frac{f(t, s)}{\psi(s)}$ è la sua funzione di distribuzione di probabilità.

7. Qualche proprietà di valor medio e varianza

Abbiamo già definito cosa intendiamo per

- media o valor medio
- varianza
- momento
- momento rispetto all'origine
- speranza matematica
- funzione generatrice dei momenti

di una variabile aleatoria ξ sia nel caso discreto finito sia nel caso continuo sia nel caso discreto infinito.

Tra le proprietà del valor medio ricordiamo che

- $E(\alpha\xi + \beta\eta) = \alpha E(\xi) + \beta E(\eta)$
- **se ξ e η sono variabili aleatorie indipendenti**

$$E(\xi\eta) = E(\xi)E(\eta)$$

In altre parole la speranza matematica è una applicazione lineare; La verifica di tale fatto non è banale e si può ad esempio ottenere provando il risultato nel caso discreto finito ed utilizzando un passaggio al limite per gli altri casi.

Si può provare che anche la varianza gode di proprietà di linearità, infatti

- $\text{Var}(\alpha\xi) = \alpha^2 \text{Var}(\xi)$
- **se ξ e η sono variabili aleatorie indipendenti**

$$\text{Var}(\xi \pm \eta) = \text{Var}(\xi) + \text{Var}(\eta)$$

È anche utile ricordare che

$$\begin{aligned}\sigma^2 &= E((\xi - \mu)^2) = E(\xi^2 - 2\mu\xi + \mu^2) = \\ &= E(\xi^2) - 2\mu E(\xi) + \mu^2 = \\ &= E(\xi^2) - 2\mu^2 + \mu^2 = E(\xi^2) - \mu^2 = E(\xi^2) - (E(\xi))^2\end{aligned}$$

Vediamo ora come la funzione generatrice dei momenti si riveli molto comoda per il calcolo dei momenti di una variabile aleatoria.

Cominciamo con l'osservare che

$$(t - \mu)^k = \sum_{i=0}^k \binom{k}{i} t^i \mu^{k-i}$$

moltiplicando per $\varphi(t)$ ed integrando, otteniamo

$$(6.15) \quad \mu_k = \sum_{i=0}^k \binom{k}{i} \mu^i \mu^{k-i}$$

e se ne ricava che per trovare i momenti rispetto al valor medio μ_k è sufficiente conoscere i momenti rispetto all'origine μ'_k .

Casi particolari della 6.15 sono

$$\begin{aligned}\mu_2 &= \sigma_2 = \mu'_2 - \mu^2 \\ \mu_3 &= \mu'_3 - 3\mu'_2\mu + 2\mu^3\end{aligned}$$

(ricordiamo che $\mu_0 = \mu'_0 = 1$ e $\mu_1 = \mu'_1 = \mu$).

D'altro canto i momenti di una variabile aleatoria si possono trovare a partire dalla funzione generatrice dei momenti che è definita da:

$$M_\xi(t) = E(e^{t\xi}) = \int_{-\infty}^{+\infty} e^{ts} \varphi(s) ds$$

Infatti si può verificare che la funzione M_ξ è sviluppabile in serie di McLaurin ed il suo sviluppo è dato da

$$M_\xi(t) = \sum_{k=0}^{+\infty} \mu'_k \frac{t^k}{k!}$$

e quindi

$$\mu'_k = \frac{d^k}{dt^k} M_\xi(t)$$

8. La disuguaglianza di Tchebichev e la legge dei grandi numeri

In questa sezione ci occupiamo di due risultati fondamentali: la disuguaglianza di Tchebichev e la legge dei grandi numeri, cominciando a parlare della prima.

Sia ξ una variabile aleatoria con media μ e varianza σ^2 , allora si avrà che

$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{+\infty} (t - \mu)^2 \varphi(t) dt = \\ &= \int_{\{t : |t - \mu| \leq \varepsilon\}} (t - \mu)^2 \varphi(t) dt + \int_{\{t : |t - \mu| > \varepsilon\}} (t - \mu)^2 \varphi(t) dt \geq \\ &\geq \int_{\{t : |t - \mu| \geq \varepsilon\}} (t - \mu)^2 \varphi(t) dt \geq \int_{\{t : |t - \mu| \geq \varepsilon\}} \varepsilon^2 \varphi(t) dt = \\ &= \varepsilon^2 \mathcal{P}(|\xi - \mu| \geq \varepsilon) \end{aligned}$$

se ne ricava pertanto che

$$(6.16) \quad \mathcal{P}(|\xi - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

La 6.16 è nota come disuguaglianza di Tchebichev e ne possiamo trarre una interessante conseguenza: per $\varepsilon = k\sigma$ otteniamo che

$$(6.17) \quad \mathcal{P}(|\xi - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Pertanto

$$(6.18) \quad \mathcal{P}(|\xi - \mu| < k\sigma) = 1 - \mathcal{P}(|\xi - \mu| \geq k\sigma) \geq 1 - \frac{1}{k^2}$$

Se ora consideriamo la seguente tabella

k	$1 - \frac{1}{k^2}$
1	0
2	.75
3	.88
4	.93
5	.95
6	.97

TABELLA 6.3. Valori approssimati di $1 - \frac{1}{k^2}$

Si vede pertanto che se ξ è una variabile aleatoria di media μ e di varianza σ^2 , allora la probabilità che il valore assunto da ξ sia vicino alla media μ per meno di 2 volte la varianza è del 75% e sale all'88% se ci accontentiamo di un errore inferiore a 3 volte la varianza.

Va osservato che, nonostante fornisca risultati soddisfacenti, la disuguaglianza di Tchebichev non è molto precisa.

8.1. La legge dei grandi numeri. Un'altra delle conseguenze della disuguaglianza di Tchebichev prende il nome di Legge dei grandi numeri e si ricava come segue.

Siano $\xi_1, \xi_2, \dots, \xi_n$ variabili aleatorie tutte con media μ e varianza σ^2 , e consideriamo la variabile aleatoria

$$S_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}$$

Avremo che

$$E(S_n) = \frac{1}{n}(E(\xi_1) + E(\xi_2) + \dots + E(\xi_n)) = \mu$$

ed inoltre

$$\text{Var}(S_n) = \frac{1}{n^2}(\text{Var}(\xi_1) + \text{Var}(\xi_2) + \dots + \text{Var}(\xi_n)) = \frac{\sigma^2}{n}$$

Pertanto

$$(6.19) \quad \mathcal{P}(|S_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0$$

per $n \rightarrow +\infty$

La 6.19 è nota con il nome di **legge dei grandi numeri** ed esprime un concetto in base al quale la media delle uscite di una variabile aleatoria differisce dalla media della variabile aleatoria di una quantità infinitesima.

Va tuttavia sottolineato che la legge dei grandi numeri fornisce informazioni di carattere qualitativo e quindi non può essere usata per stime di tipo quantitativo.

9. Normalizzazione di una variabile aleatoria.

Sia ξ una variabile aleatoria di media μ e di varianza σ^2 con distribuzione di probabilità φ .

Sia cioè

$$\begin{aligned}\int_{-\infty}^{+\infty} \varphi(t) dt &= 1 \\ \mathcal{P}(a \leq \xi \leq b) &= \int_a^b \varphi(t) dt \\ E(\xi) &= \int_{-\infty}^{+\infty} t \varphi(t) dt = \mu \\ E(\xi^2) &= \int_{-\infty}^{+\infty} (t - \mu)^2 \varphi(t) dt = \sigma^2\end{aligned}$$

e consideriamo la variabile aleatoria

$$\xi^* = \frac{\xi - \mu}{\sigma}$$

Per le proprietà di media e varianza possiamo affermare che ξ^* è una variabile normalizzata (o standardizzata), intendendo con ciò che ξ^* ha media 0 e varianza 1.

Allo scopo di determinare la funzione di distribuzione di ξ^* consideriamo la variabile aleatoria η la cui funzione di distribuzione è definita da

$$\sigma \varphi(\mu + \sigma t)$$

Possiamo allora verificare che

$$\begin{aligned}\int_{-\infty}^{+\infty} \sigma \varphi(\mu + \sigma t) dt &= \int_{-\infty}^{+\infty} \varphi(s) ds = 1 \\ \int_{-\infty}^{+\infty} t \sigma \varphi(\mu + \sigma t) dt &= \int_{-\infty}^{+\infty} \left(\frac{s - \mu}{\sigma} \right) \varphi(s) ds = \\ &= \frac{1}{\sigma} \left(\int_{-\infty}^{+\infty} s \varphi(s) ds - \mu \int_{-\infty}^{+\infty} \varphi(s) ds \right) = 0 \\ \int_{-\infty}^{+\infty} t^2 \sigma \varphi(\mu + \sigma t) dt &= \int_{-\infty}^{+\infty} \left(\frac{s - \mu}{\sigma} \right)^2 \varphi(s) ds = \\ &= \frac{1}{\sigma} \int_{-\infty}^{+\infty} (s - \mu)^2 \varphi(s) ds = 1\end{aligned}$$

$$\begin{aligned}
\mathcal{P}(a \leq \eta \leq b) &= \int_a^b \sigma \varphi(\mu + \sigma t) dt = \\
&= \int_{\mu + \sigma a}^{\mu + \sigma b} \sigma \varphi(s) \frac{1}{\sigma} ds = \mathcal{P}(\mu + \sigma a \leq \xi \leq \mu + \sigma b) = \\
&= \mathcal{P}\left(a \leq \frac{\xi - \mu}{\sigma} \leq b\right)
\end{aligned}$$

Ne deduciamo che η e $\xi^* = \frac{\xi - \mu}{\sigma}$ hanno la stessa distribuzione di probabilità e quindi sono la stessa variabile aleatoria.

In altre parole la funzione

$$\sigma \varphi(\mu + \sigma t)$$

è la distribuzione di probabilità di

$$\xi^* = \frac{\xi - \mu}{\sigma}$$

CAPITOLO 7

QUALCHE DISTRIBUZIONE DI PROBABILITÀ

Le funzioni di distribuzione di probabilità sono fondamentali per descrivere il comportamento delle variabili aleatorie che ci interessano.

Ogni variabile aleatoria ha una sua distribuzione e per definirne la proprietà è utile fare riferimento ad alcune distribuzioni note che sono in grado di descrivere la maggior parte delle variabili aleatorie con cui normalmente si lavora.

1. La distribuzione uniforme

La più semplice funzione di distribuzione di probabilità è quella di una variabile aleatoria che restituisce un valore scelto in un intervallo $[a, b]$ con il criterio di equiprobabilità.

Abbiamo già visto che in tal caso

$$\mathcal{P}(x \leq \xi \leq x + h) = \frac{h}{b - a}$$

e che la sua distribuzione di densità è

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

La funzione generatrice dei momenti si calcola mediante la

$$M_{\xi}(t) = \frac{1}{b-a} \int_a^b e^{tx} dx = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

Se ne ricava subito che

$$\mu = \frac{b+a}{2}$$
$$\sigma^2 = \frac{(b-a)^2}{12}$$

Descriviamo ora tre importanti distribuzioni di probabilità che sono, come vedremo, tra loro legate:

- La distribuzione binomiale di Bernoulli
- La distribuzione di Poisson
- La distribuzione normale di Gauss

Si tratta di una distribuzione discreta finita (Bernoulli), una distribuzione discreta numerabile (Poisson) ed una distribuzione continua (Gauss) che possono essere derivate a partire dal concetto elementare di prova bernoulliana.

2. La distribuzione binomiale di Bernoulli

DEFINIZIONE 7.1. *Chiamiamo prova bernoulliana un esperimento che ha due soli possibili esiti:*

- *Successo, cui associamo il valore 1 con probabilità p*
- *Insuccesso, cui associamo il valore 0 con probabilità q*

essendo ovviamente $p + q = 1$.

Chiamiamo variabile aleatoria bernoulliana la variabile aleatoria ξ che restituisce il numero di successi che si sono verificati su n prove ripetute (lanci) dell'esperimento.

Possiamo calcolare la probabilità che la variabile aleatoria ξ assuma il valore k mediante la

$$\varphi(k) = \mathcal{P}(\xi = k) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Possiamo giustificare la formula precedente se descriviamo la successione di n prove ripetute con una stringa di elementi che assumono il valore 1 oppure 0 a seconda che la corrispondente prova abbia avuto o no successo.

0	1	1	0	0	1	0	0	1	0	0	1	0	1	0	0	0	1	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

affinchè ci siano k successi la stringa dovrà contenere esattamente k volte il valore 1 (ed $n - k$ volte il valore 0) e quindi, poichè in ogni elemento 1 si presenta con probabilità p e 0 con probabilità q , una stringa con k successi avrà una probabilità di comparire uguale a

$$p^k q^{n-k}$$

d'altro canto, poichè siamo unicamente interessati a contare il numero di successi, e non l'ordine con cui si verificano, dovremo tener conto che si verificano, ad esempio, 2 successi in tanti modi diversi

0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

il cui numero è dato dalle combinazioni di n oggetti a k a k e cioè

$$\binom{n}{k}$$

(Infatti ciascuna stringa può essere individuata dalla sequenza dei k numeri, compresi tra 1 ed n , che indicano la posizione dei successi nella stringa.)

Possiamo calcolare la media della variabile bernoulliana ξ osservando che la media in ciascuna prova è

$$1 \cdot p + 0 \cdot q = p$$

e su n esperimenti, essendo la media lineare, avremo

$$\mu = E(\xi) = np$$

La varianza della variabile bernoulliana ξ in ciascuna prova è

$$(1-p)^2 \cdot p + 0 \cdot (0-q)^2 q = q^2 p + p^2 q = pq(p+q) = pq$$

e su n esperimenti essendo la varianza lineare avremo

$$\sigma^2 = E((\xi - \mu)^2) = npq$$

e

$$\sigma = \sqrt{npq}$$

Per calcolare la funzione generatrice dei momenti possiamo procedere come segue

$$\begin{aligned} M_\xi(t) = E(e^{t\xi}) &= \sum_{k=0}^n e^{tk} \binom{n}{k} p^k q^{n-k} = \\ &= \sum_{k=0}^n \binom{n}{k} (pe^t)^k q^{n-k} = (pe^t + q)^n \end{aligned}$$

Un esempio notevole di prove bernoulliane si costruisce considerando il lancio di una moneta ripetuto per n volte; in tal caso $p = q = \frac{1}{2}$.

3. La distribuzione di Poisson

Si consideri il caso di m prove ripetute bernoulliane in cui la probabilità di successo della singola prova sia $\frac{1}{n}$

Un esempio di tale situazione di si può realizzare considerando un pagliaio con un numero $n-1$ di pagliuzze dal quale si voglia estrarre 1 ago; la probabilità di estrarre l'ago sarà ovviamente $\frac{1}{n}$ e la prova sarà ripetuta m volte.

Sia ξ la variabile aleatoria che restituisce il numero di successi ottenuti nelle m prove fatte. Avremo che

$$\begin{aligned} \varphi(k) = \mathcal{P}(\xi = k) &= \binom{m}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{m-k} = \\ &= \frac{m(m-1)(m-2) \cdots (m-(k-1))}{k!} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^m \left(1 - \frac{1}{n}\right)^{-k} = \\ &= \frac{1}{k!} \frac{m}{n} \left(\frac{m}{n} - \frac{1}{n}\right) \cdots \left(\frac{m}{n} - \frac{k-1}{n}\right) \left(1 - \frac{1}{n}\right)^m \left(1 - \frac{1}{n}\right)^{-k} \end{aligned}$$

Se ora consideriamo di far tendere n a $+\infty$ mantenendo

$$\frac{m}{n} = \lambda$$

(oppure chiedendo semplicemente che $\frac{m}{n} \rightarrow \lambda$), otteniamo che

$$\varphi(k) = \frac{1}{k!} \lambda \left(\lambda - \frac{1}{n} \right) \left(\lambda - \frac{2}{n} \right) \cdots \left(\lambda - \frac{k-1}{n} \right) \left(1 - \frac{1}{n} \right)^{n\lambda} \left(1 - \frac{1}{n} \right)^{-k}$$

e, per $n \rightarrow +\infty$

$$\varphi(k) = \frac{1}{k!} \lambda^k e^{-\lambda}$$

La funzione distribuzione di probabilità di Poisson è quindi data da

$$\boxed{\varphi(k) = \frac{1}{k!} \lambda^k e^{-\lambda}}$$

Per calcolare media, varianza ed i momenti della distribuzione di Poisson è utile calcolare la funzione generatrice dei momenti.

$$M_{\xi}(t) = \sum_{k=0}^{+\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{+\infty} \frac{(\lambda e^t)^k}{k!} e^{-\lambda} = e^{-\lambda} e^{\lambda(e^t-1)}$$

e le sue derivate

$$\begin{aligned} \frac{d}{dt} M_{\xi}(t) &= e^{\lambda(e^t-1)} \lambda e^t \\ \frac{d^2}{dt^2} M_{\xi}(t) &= e^{\lambda(e^t-1)} \lambda^2 e^t + e^{\lambda(e^t-1)} \lambda e^t \end{aligned}$$

e calcolarle in $t = 0$.

$$\begin{aligned} \frac{d}{dt} M_{\xi}(0) &= \lambda = \mu'_1 = \mu \\ \frac{d^2}{dt^2} M_{\xi}(0) &= \lambda^2 + \lambda = \mu'_2 \end{aligned}$$

da cui

$$\begin{aligned} \mu &= \mu'_1 = \lambda \\ \sigma^2 &= \mu'_2 - \mu^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \end{aligned}$$

4. La distribuzione gaussiana

Sia $\xi_n = k$ una variabile aleatoria bernoulliana relativa ad n prove ripetute, allora

$$\mathcal{P}(\xi_n = k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

Sia ha

$$\begin{aligned} E(\xi_n) &= \mu = np \\ Var(\xi_n) &= \sigma^2 = npq \end{aligned}$$

Consideriamo la variabile normalizzata (di media 0 e varianza 1)

$$\xi_n^* = \frac{\xi_n - \mu}{\sigma}$$

ξ_n^* assume i valori

$$x_k = \frac{k - \mu}{\sigma} \quad k = 1, \dots, n$$

e

$$\mathcal{P}\left(\xi_n^* = \frac{k - \mu}{\sigma}\right) = \mathcal{P}(\xi_n = k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

Possiamo ricavare inoltre che

$$\begin{aligned} k &= \mu + \sigma x_k = np + \sigma x_k \\ n - k &= n - np - \sigma x_k = nq - \sigma x_k \end{aligned}$$

dove $\sigma = \sqrt{npq}$ e si vede che, per $n \rightarrow +\infty$ tanto k quanto $n - k$ tendono a $+\infty$, inoltre

$$\begin{aligned} \frac{k}{np} &= 1 + \sqrt{\frac{q}{np}} x_k \\ \frac{n-k}{nq} &= 1 + \sqrt{\frac{p}{nq}} x_k \\ x_{k+1} - x_k &= \Delta x_k = \frac{1}{\sigma} = \frac{1}{\sqrt{npq}} \end{aligned}$$

Possiamo allora scrivere che

$$\begin{aligned} \mathcal{P}(x' \leq \xi_n^* \leq x'') &= \sum_{x_k \in [x', x'']} \mathcal{P}(\xi_n^* = x_k) = \\ &= \sum_{\{k: x_k \in [x', x'']\}} \mathcal{P}(\xi_n = k) = \\ &= \sum_{\{k: x_k \in [x', x'']\}} \frac{n!}{k!(n-k)!} p^k q^{n-k} \end{aligned}$$

Poichè, per n grande, tanto k quanto $n - k$ sono grandi, possiamo usare la formula di Stirling per approssimare il fattoriale e quindi

$$\begin{aligned} \frac{n!}{k!(n-k)!} p^k q^{n-k} &\approx \frac{n^n e^{-n} \sqrt{2\pi n}}{k^k e^{-k} \sqrt{2\pi k} (n-k)^{n-k} e^{-(n-k)} \sqrt{2\pi(n-k)}} p^k q^{n-k} = \\ &= \frac{1}{\sqrt{2\pi}} \frac{n^k}{k^k} \frac{n^{n-k}}{(n-k)^{n-k}} \sqrt{\frac{n}{k(n-k)}} p^k q^{n-k} = \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \sqrt{\frac{n}{k(n-k)}} \end{aligned}$$

Ma, sempre per n grande,

$$\sqrt{\frac{n}{k(n-k)}} = \sqrt{\frac{n}{(np + \sigma x_k)(nq - \sigma x_k)}} \approx \sqrt{\frac{n}{n^2 pq}} = \frac{1}{\sqrt{npq}}$$

D'altro canto

$$\begin{aligned}
\ln \left(\frac{np}{k} \right)^k \left(\frac{nq}{n-k} \right)^{n-k} &= -k \ln \left(\frac{np}{k} \right) - (n-k) \ln \left(\frac{nq}{n-k} \right) = \\
&= -(np + \sigma x_k) \ln \left(1 + \sqrt{\frac{q}{np}} \right) - \\
&\quad - (nq - \sigma x_k) \ln \left(1 - \sqrt{\frac{p}{nq}} x_k \right) + \omega \left(\frac{1}{\sqrt{n}} \right) = \\
&= -(np + \sigma x_k) \left(\sqrt{\frac{q}{np}} x_k - \frac{1}{2} \frac{q}{np} x_k^2 \right) - \\
&\quad - (nq - \sigma x_k) \left(-\sqrt{\frac{p}{nq}} x_k - \frac{1}{2} \frac{p}{nq} x_k^2 \right) + \omega \left(\frac{1}{\sqrt{n}} \right) = \\
&= -x_k \left[np \sqrt{\frac{q}{np}} - nq \sqrt{\frac{p}{nq}} \right] - x_k^2 \left[\sqrt{npq} \sqrt{\frac{q}{np}} + \sqrt{npq} \sqrt{\frac{p}{nq}} \right] - \\
&\quad - x_k^2 \left[\frac{1}{2} np \frac{q}{np} + \frac{1}{2} nq \frac{p}{nq} \right] - x_k^3 \left[\frac{q \sqrt{npq}}{2np} + \frac{p \sqrt{npq}}{nq} \right] = \\
&= -x_k^2 + \frac{1}{2} x_k^2 + \omega \left(\frac{1}{\sqrt{n}} \right)
\end{aligned}$$

dove, come al solito, si indica con ω una funzione infinitesima.

Ne concludiamo pertanto che

$$\begin{aligned}
\mathcal{P}(x' \leq \xi_n^* \leq x'') &\approx \\
&\approx \sum_{x_k \in [x', x'']} \frac{1}{\sqrt{2\pi}} \left(\frac{np}{k} \right)^k \left(\frac{nq}{n-k} \right)^{n-k} \sqrt{\frac{n}{k(n-k)}} \approx \\
&\approx \sum_{x_k \in [x', x'']} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \Delta x_k \rightarrow \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-\frac{1}{2}t^2} dt
\end{aligned}$$

È pertanto naturale considerare la funzione distribuzione di probabilità definita da

$$\boxed{\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}}$$

La φ si chiama funzione di distribuzione di probabilità Gaussiana di media 0 e varianza 1. (Si ricordi che è ottenuta come limite di distribuzioni normalizzate).

Utilizzando il teorema di integrazione per sostituzione (poichè l'integrale è improprio deve essere calcolato come limite di integrali propri ed a questi si applica il teorema di integrazione per sostituzione) possiamo ricavare che la funzione

$$\varphi(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

è una funzione di distribuzione di probabilità, che chiameremo ancora Gaussiana, di media μ e di varianza σ^2 .

Possiamo calcolare la funzione generatrice dei momenti della funzione di distribuzione gaussiana, infatti

$$M_\xi(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

e, operando il cambio di variabili $\frac{x-\mu}{\sigma} = u$, da cui $x = \sigma u + \mu$,

$$\begin{aligned} M_\xi(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{t\sigma u + t\mu} e^{-\frac{u^2}{2}} \sigma du = \\ &= \sigma \frac{1}{\sigma\sqrt{2\pi}} e^{t\mu} \int_{-\infty}^{+\infty} e^{t\sigma u} e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} e^{t\mu} \int_{-\infty}^{+\infty} e^{\frac{1}{2}(2t\sigma u - u^2)} du = \\ &= \frac{1}{\sqrt{2\pi}} e^{t\mu} \int_{-\infty}^{+\infty} e^{\frac{1}{2}(\sigma^2 t^2 - \sigma^2 t^2 + 2t\sigma u - u^2)} du = \\ &= \frac{1}{\sqrt{2\pi}} e^{t\mu + \frac{1}{2}\sigma^2 t^2} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\sigma t - u)^2} du = \\ &= \frac{1}{\sqrt{2\pi}} e^{t\mu + \frac{1}{2}\sigma^2 t^2} \sqrt{2\pi} = e^{t\mu + \frac{1}{2}\sigma^2 t^2} \end{aligned}$$

La legge di distribuzione di probabilità normale è nota come distribuzione Gaussiana anche se Gauss si riferisce ad essa in una sua pubblicazione solo nel 1809, mentre in precedenza, nel 1733 De Moivre aveva pubblicato una derivazione della legge normale come limite di una distribuzione binomiale ed anche se era già nota a Laplace almeno dal 1774.

Una derivazione della legge normale molto interessante è dovuta ad Herschel, 1850, che la dedusse studiando la distribuzione $\phi(x, y)$ a due dimensioni degli errori di misurazione della posizione di una stella.

Una stella si individua mediante la sua longitudine, misurata con un errore x , e la sua declinazione, misurata con un errore y ; Herschel postulò che

x ed y sono indipendenti

per cui se f è la densità di probabilità di x e di y , si ha

$$\phi(x, y) = f(x)f(y)$$

inoltre postulò che la distribuzione ϕ , espressa usando le coordinate polari (ρ, θ) , fosse indipendente da θ , per cui

$$\phi(x, y) = f(x)f(y) = g(x^2 + y^2)$$

Ponendo $x = 0$ si ricava

$$f(x)f(0) = g(x^2)$$

da cui

$$\frac{g(x^2)}{f(0)} \frac{g(y^2)}{f(0)} = g(x^2 + y^2)$$

dividendo per $(f(0))^2$

$$\frac{g(x^2)}{(f(0))^2} \frac{g(y^2)}{(f(0))^2} = \frac{g(x^2 + y^2)}{(f(0))^2}$$

e

$$\ln \left(\frac{g(x^2)}{(f(0))^2} \right) + \ln \left(\frac{g(y^2)}{(f(0))^2} \right) = \ln \left(\frac{g(x^2 + y^2)}{(f(0))^2} \right)$$

Pertanto la funzione

$$z \mapsto \ln \left(\frac{g(z)}{(f(0))^2} \right)$$

è lineare e quindi

$$\ln \left(\frac{g(z)}{(f(0))^2} \right) = kz$$

e

$$\ln \left(\frac{g(x^2)}{(f(0))^2} \right) = kx^2$$

Ne deduciamo che

$$g(x^2) = h^2 e^{kx^2}$$

e le costanti h e k possono essere determinate in modo da aversi una distribuzione di probabilità. Innanzi tutto si vede che deve essere $h > 0$ e $k < 0$, imponendo poi che la distribuzione abbia media μ e varianza σ otteniamo

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

5. Due distribuzioni discrete

5.1. La distribuzione multinomiale. Consideriamo un esperimento che possa avere k possibili esiti, che indichiamo con

$$A_1, A_2, \dots, A_k$$

con probabilità

$$p_1, p_2, \dots, p_k \quad , \quad p_1 + p_2 + \dots + p_k = 1$$

e supponiamo di replicarlo per n volte; consideriamo la variabile aleatoria ξ che restituisce la n -pla di valori

$$n_1, n_2, \dots, n_k \quad , \quad n_1 + n_2 + \dots + n_k = n$$

dove n_i è il numero di volte in cui si è verificato l'evento A_i .

La funzione distribuzione di probabilità di ξ è data da

$$\begin{aligned}\varphi(n_1, n_2, \dots, n_k) &= \mathcal{P}(\xi_1 = p_1, \xi_2 = p_2, \dots, \xi_k = p_k) = \\ &= \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}\end{aligned}$$

5.2. La distribuzione ipergeometrica. Consideriamo un'urna contenente b palline nere e w palline bianche e supponiamo di estrarre per n volte una pallina rimettendola. dopo ogni estrazione, nell'urna.

Consideriamo la variabile aleatoria ξ che restituisce il numero di volte in cui si è estratta una pallina nera; allora la densità di probabilità di ξ si può calcolare mediante la

$$\begin{aligned}\varphi(k) = \mathcal{P}(\xi = k) &= \binom{n}{k} \left(\frac{b}{b+w} \right)^k \left(\frac{w}{b+w} \right)^{n-k} = \\ &= \binom{n}{k} \frac{b^k w^{n-k}}{b+w}\end{aligned}$$

non appena si ricordi la distribuzione binomiale e si tenga presente che

$$p = \frac{b}{b+w}, \quad q = \frac{w}{b+w}$$

Qualora l'esperimento si ripeta senza rimettere la pallina estratta nell'urna, (campionamento senza ripetizione), si può vedere che la densità di probabilità della nuova variabile aleatoria ξ che conta il numero delle palline nere estratte è

$$\varphi(k) = \mathcal{P}(\xi = k) = \frac{\binom{b}{k} \binom{w}{n-k}}{\binom{b+w}{n}}$$

Si calcola anche che

$$\begin{aligned}\mu &= \frac{nb}{b+w} \\ \sigma^2 &= \frac{nbw(b+w-n)}{(b+w)^2(b+w-1)}\end{aligned}$$

6. Le distribuzioni legate ai test statistici.

6.1. La distribuzione χ^2 . Si tratta della distribuzione di probabilità di una variabile aleatoria χ^2 che restituisce la somma di ν variabili aleatorie ξ_i indipendenti, aventi distribuzione gaussiana con media 0 e varianza 1 (distribuzioni normali standardizzate).

$$\chi^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_\nu^2$$

La funzione di distribuzione della variabile aleatoria χ^2 è data da

$$\varphi(t) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} t^{\nu/2-1} e^{-t/2} & t \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

ν rappresenta i gradi di libertà, media, varianza e generatrice dei momenti sono date da

$$\begin{aligned} \mu &= \nu \\ \sigma^2 &= 2\nu \\ M_{\chi}^2(t) &= (1 - 2t)^{-\nu/2} \end{aligned}$$

6.2. La distribuzione T di Student. È la distribuzione di una variabile aleatoria T che restituisce il rapporto

$$T = \frac{\xi}{\sqrt{\eta/\nu}}$$

dove ξ è una variabile aleatoria con densità di probabilità gaussiana normale (media 0 e varianza 1) ed η è una variabile aleatoria con distribuzione χ^2 a ν gradi di libertà

La funzione di distribuzione della variabile aleatoria T è data da

$$\varphi(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma(\nu/2)} \left(1 + \frac{t^2}{\nu}\right)$$

ν rappresenta i gradi di libertà, media e varianza sono date da

$$\begin{aligned} \mu &= 0 \\ \sigma^2 &= \frac{\nu}{\nu - 2} \end{aligned}$$

6.3. La distribuzione F di Fisher. È la distribuzione di una variabile aleatoria F che restituisce il rapporto

$$F = \frac{\xi_1/\nu_1}{\xi_2/\nu_2}$$

dove ξ ed ξ_2 sono variabili aleatorie con distribuzione χ^2 a ν_1 e ν_2 gradi di libertà, rispettivamente.

La funzione di distribuzione della variabile aleatoria F è data da

$$\varphi(t) = \begin{cases} \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \nu_1^{\nu_1/2} \nu_2^{\nu_2/2} t^{\nu_2/2-1} (\nu_2 + \nu_1 t)^{-(\nu_1+\nu_2)/2} & t > 0 \\ 0 & \text{altrimenti} \end{cases}$$

media e varianza sono date da

$$\mu = \frac{\nu_2}{\nu_2 - 2}$$

$$\sigma^2 = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)^2}$$

7. La distribuzione γ .

La distribuzione γ è definita, per $\alpha, \beta > 0$ da

$$\varphi(t) = \begin{cases} t^{\alpha-1} e^{-t/\beta} & t > 0 \\ 0 & \text{altrimenti} \end{cases}$$

media, varianza e generatrice dei momenti sono date da

$$\mu = \alpha\beta$$

$$\sigma^2 = \alpha\beta^2$$

$$M_\gamma(t) = (1 - \beta t)^{-\alpha}$$

8. La distribuzione β .

La distribuzione β è definita da

$$\varphi(t) = \begin{cases} \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha, \beta)} & 0 < t < 1 \\ 0 & \text{altrimenti} \end{cases} = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1}(1-t)^{\beta-1} & 0 < t < 1 \\ 0 & \text{altrimenti} \end{cases}$$

media e varianza sono date da

$$\mu = \frac{\alpha}{\alpha + \beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)(\alpha + \beta + 1)}$$

Si vede anche la sua moda è $\frac{\alpha-1}{\alpha+\beta-2}$.

9. Variabili casuali con distribuzione assegnata.

Qualora sia necessario utilizzare dati generati casualmente con funzione distribuzione di probabilità fissata, possiamo procedere come segue.

Sia ϕ la distribuzione che si vuole considerare e sia

$$F(x) = \int_{-\infty}^x \phi(t) dt$$

la funzione di distribuzione cumulativa; allora si ha

$$\mathcal{P}(a \leq \xi \leq b) = \int_a^b \phi(t) dt = F(b) - F(a)$$

e quindi

$$\mathcal{P}(F^{-1}(a) \leq \xi \leq F^{-1}(b)) = b - a$$

e

$$\mathcal{P}(a \leq F(\xi) \leq b) = b - a$$

Pertanto $F(\xi)$ ha una distribuzione uniforme e quindi poichè

$$\xi = F^{-1}(F(\xi))$$

possiamo generare valori distribuiti con densità di probabilità ϕ , considerando valori generati con densità uniforme ed applicando a tali valori F^{-1} .

Tale procedimento non è tuttavia applicabile, ad esempio, per determinare valori distribuiti con densità gaussiana in quanto non è possibile determinare esplicitamente F^{-1} nel caso in cui

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

In tal caso, che peraltro è di rilevante importanza possiamo osservare che se ξ e η sono variabili aleatorie indipendenti con distribuzione gaussiana normale, allora (ξ, η) è una variabile aleatoria la cui funzione distribuzione di probabilità è

$$\frac{1}{2\pi} e^{-(t^2+s^2)/2}$$

Pertanto

$$\mathcal{P}((\xi, \eta) \in A) = \frac{1}{2\pi} \iint_A e^{-(t^2+s^2)/2} dt ds = \frac{1}{2\pi} \iint_B \rho e^{-(\rho^2)/2} d\rho d\theta$$

dove B ed A sono l'uno il trasformato dell'altro rispetto al cambio di variabili in coordinate polari.

Ne viene che possiamo identificare due nuove variabili (R, Θ) la cui densità di probabilità è data da

$$\frac{1}{2\pi} \rho e^{-(\rho^2)/2} = \left(\frac{1}{2\pi}\right) (\rho e^{-(\rho^2)/2}) = f(\theta)g(\rho)$$

dove

$$f(\theta) = \frac{1}{2\pi}$$

$$g(\rho) = \rho e^{-(\rho^2)/2}$$

Quindi per quanto visto in precedenza, per generare valori casuali di Θ e di R possiamo utilizzare valori uniformemente distribuiti θ ed r e applicare

a tali valori le funzioni F^{-1} e G^{-1} , rispettivamente, dove

$$F(t) = \int_0^t f(s)ds = 2\pi t \quad , \quad G(t) = \int_0^t g(s)ds = 1 - e^{-t^2/2}$$

Si ha allora

$$F^{-1}(s) = \frac{s}{2\pi} \quad , \quad G^{-1}(s) = \sqrt{-2 \ln(1-s)}$$

e le variabili

$$\xi = \sqrt{-2 \ln(1-t)} \cos\left(\frac{s}{2\pi}\right) \quad , \quad \eta = \sqrt{-2 \ln(1-t)} \sin\left(\frac{s}{2\pi}\right)$$

dove s e t sono distribuite uniformemente, risultano distribuite con densità gaussiana di media 0 e di varianza 1.

10. I test statistici

Consideriamo il seguente esempio.

Supponiamo di voler determinare se una moneta è truccata lanciandola un numero N di volte ed osservando il risultato dei lanci.

Possiamo definire truccata una moneta in cui la probabilità di uscita di T , che chiameremo p è diversa dalla probabilità di uscita di C che è $q = (1 - p)$. Sarà quindi truccata una moneta per la quale $p \neq 0.5$.

Osserviamo comunque che allo scopo di ottenere benefici da una scommessa potremmo anche essere interessati a conoscere se $p > 0.5$ o equivalentemente se $p < 0.5$.

I dati di cui vogliamo individuare qualche caratteristica, la popolazione (il nome deriva dalle origini dello studio statistico), sono costituiti da tutti i possibili lanci della moneta in questione.

I dati di cui disponiamo sono costituiti dalle uscite degli N lanci osservati, campione.

Del campione siamo in grado di conoscere la distribuzione di probabilità che è una distribuzione discreta binomiale la cui media è $\mu = Np$ e la cui varianza è $\sigma^2 = \sqrt{Npq}$.

Calcoliamo la probabilità che la variabile aleatoria ξ che restituisce il numero di teste uscite negli N lanci sia compresa tra n ed m .

Avremo che

$$\mathcal{P}(n \leq \xi \leq m) = \sum_{k=n}^m \binom{N}{k} p^k q^{N-k}$$

e quindi, se ad esempio $N = 100$, $n = 45, 40, 30$, $m = 55, 60, 70$,

$$\mathcal{P}(45 \leq \xi \leq 55) \approx 0.73$$

$$\mathcal{P}(40 \leq \xi \leq 60) \approx 0.96$$

$$\mathcal{P}(30 \leq \xi \leq 70) \approx 0.99$$

Possiamo quindi affermare che

- nel caso il numero di teste uscite sia superiore a 55 o inferiore a 45, si è verificato un evento che, supponendo la moneta non truccata, ha la probabilità di verificarsi del $100 - 73 = 27\%$
- nel caso il numero di teste uscite sia superiore a 60 o inferiore a 40, si è verificato un evento che, supponendo la moneta non truccata, ha la probabilità di verificarsi del $100 - 96 = 4\%$
- nel caso il numero di teste uscite sia superiore a 70 o inferiore a 30, si è verificato un evento che, supponendo la moneta non truccata, ha la probabilità di verificarsi del $100 - 99 = 1\%$

e potremmo in base a queste considerazioni concludere che

- nel caso il numero di teste uscite sia superiore a 55 o inferiore a 45, siamo fiduciosi che la moneta sia truccata con un livello di significatività del 73%
- nel caso il numero di teste uscite sia sia superiore a 60 o inferiore a 40, siamo fiduciosi che la moneta sia truccata con un livello di significatività del 96%
- nel caso il numero di teste uscite sia superiore a 70 o inferiore a 30, siamo fiduciosi che la moneta sia truccata con un livello di significatività del 99%

Abbiamo in altre parole usato la variabile ξ che restituisce il numero di teste uscite per stimare la veridicità dell'affermazione:

La moneta non è truccata

Chiamiamo la variabile ξ stimatore e l'affermazione *La moneta non è truccata* ipotesi H_0

Osserviamo che accettare che la moneta sia truccata, cioè rigettare l'ipotesi H_0 che *La moneta non è truccata* con un livello di significatività del 99% non vuol dire essere certi che la moneta è truccata con il 99% di possibilità in quanto non è tra l'altro affatto chiaro rispetto a quale spazio tale probabilità sia calcolata.

Potremmo anche calcolare la probabilità che il numero di teste uscito sia compreso tra n ed m utilizzando il fatto che la distribuzione binomiale si può approssimare con la distribuzione normale.

In tal caso dobbiamo normalizzare i dati osservando che

$$\begin{aligned}\mathcal{P}(n \leq \xi \leq m) &= \mathcal{P}\left(\frac{n - Np}{\sqrt{Npq}} \leq \frac{\xi - Np}{\sqrt{Npq}} \leq \frac{m - Np}{\sqrt{Npq}}\right) = \\ &= G\left(\frac{m - Np}{\sqrt{Npq}}\right) - G\left(\frac{n - Np}{\sqrt{Npq}}\right)\end{aligned}$$

dove

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

è la funzione distribuzione cumulativa normale standardizzata (gaussiana).

Otterremo i seguenti risultati che differiscono di poco dai precedenti

$$\mathcal{P}(45 \leq \xi \leq 55) \approx 0.68$$

$$\mathcal{P}(40 \leq \xi \leq 60) \approx 0.95$$

$$\mathcal{P}(30 \leq \xi \leq 70) \approx 0.99$$

e potremmo trarre conclusioni simili alle precedenti.

Qualora fossimo interessati a verificare l'ipotesi H_0 a fronte dell'ipotesi H_1 che la probabilità di successo p sia inferiore a 0.5, potremmo calcolare

$$\mathcal{P}(\xi \leq 55) \approx 0.86$$

$$\mathcal{P}(\xi \leq 60) \approx 0.98$$

$$\mathcal{P}(\xi \leq 70) \approx 0.99$$

e concludere

- nel caso il numero di teste uscite sia superiore a 55, siamo fiduciosi che la moneta sia truccata con $p > .05$ con un livello di significatività del 86%
- nel caso il numero di teste uscite sia superiore a 60, siamo fiduciosi che la moneta sia truccata con $p > .05$ con un livello di significatività del 98%
- nel caso il numero di teste uscite sia superiore a 70, siamo fiduciosi che la moneta sia truccata con $p > .05$ con un livello di significatività del 99%

Questo esempio mostra come sia possibile trarre conclusioni sulle caratteristiche di una popolazione (l'esito di tutti i possibili lanci di una moneta) esaminandone soltanto un campione (l'esito di un numero finito di lanci della moneta) e stimando le caratteristiche mediante la somma degli esiti (1 nel caso di T , successo, 0 in caso contrario).

Per poter procedere occorre precisare meglio i concetti di popolazione, campione e stimatore.

10.1. Popolazioni. Diciamo che è assegnata una popolazione se è assegnato un insieme $\mathcal{U}$ ed una variabile aleatoria ξ definita su $\mathcal{U}$.

Chiamiamo popolazione l'insieme $\xi(\mathcal{U})$.

Se ad esempio siamo interessati a stimare il diametro di 9 sferette di acciaio contenute in un cuscinetto, possiamo indicare ciascuna sferetta con un indice e considerare la popolazione i cui elementi sono

$$\{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9\}$$

In questo caso $\mathcal{U}$ è costituito dalle 9 sferette

$$\mathcal{U} = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9\}$$

e possiamo considerare la variabile aleatoria ξ che associa ad ogni sferetta il suo diametro $\xi(k) = d_k$ in modo che

$$\xi(\mathcal{U}) = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9\}$$

La variabile ξ descrive la popolazione che stiamo esaminando.

Se fossimo interessati a studiare il diametro delle sferette che costituiscono la produzione mensile di una macchina, sarebbe molto scomodo elencarle tutte e quindi potremmo descrivere la popolazione assegnando per la variabile ξ una funzione di distribuzione, ad esempio gaussiana di medio μ e varianza σ^2 .

Nel caso del lancio della moneta la popolazione è costituita da tutti i lanci.

10.2. Campioni. Data una popolazione definita dalla variabile aleatoria ξ sullo spazio $\mathcal{U}$ diciamo che è assegnato un campione di taglia n se sono assegnate n variabili aleatorie $X_1, X_2, \dots, X_n$ sullo spazio $\mathcal{U}$. Indichiamo con X la n -pla delle variabili aleatorie $X_1, X_2, \dots, X_n$ che costituiscono il campione.

Ad esempio possiamo considerare

- X_1 come la variabile aleatoria che restituisce il diametro di una sferetta scelta in $\mathcal{U}$
- X_2 come la variabile aleatoria che restituisce il diametro di una seconda sferetta scelta in $\mathcal{U}$
- X_k come la variabile aleatoria che restituisce il diametro di una k -esima sferetta scelta in $\mathcal{U}$

Nel caso del lancio di una moneta il campione è dato dalle N variabili aleatorie che restituiscono gli esiti degli N lanci effettuati.

10.3. Stimatore. Diciamo che è assegnato uno stimatore, o riassunto campionario, se è assegnata una funzione $\phi = \phi(X_1, X_2, \dots, X_n)$ delle n variabili aleatorie $X_1, X_2, \dots, X_n$ che costituiscono il campione.

Possiamo ad esempio considerare uno stimatore del diametro delle sferette considerando la media dei diametri delle n sferette di cui abbiamo esaminato il diametro per costruire il campione.

In tal caso

$$\bar{X} = \phi(X_1, X_2, \dots, X_n) = \frac{1}{n}(X_1 + X_2 + \dots + X_n) = \frac{1}{n}(d_1 + d_2 + \dots + d_n)$$

Nel caso del lancio di una moneta lo stimatore è costituito dalla somma dei risultati di ciascun lancio, per cui ϕ è il numero di successi ottenuti su n lanci.

Uno stimatore è a sua volta una variabile aleatoria.

Uno stimatore spesso usato è costituito dalla media dei campioni, come visto nell'esempio delle sferette); indicheremo tale stimatore con

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

11. Risultati sulle distribuzioni campionarie

Si consideri una popolazione individuata da una variabile aleatoria ξ su un insieme $\mathcal{U}$

un campione $X = (X_1, X_2, \dots, X_n)$ di taglia n estratto dalla popolazione individuata da ξ ed $\mathcal{U}$

Si possono provare alcuni fatti:

11.1. Se $\mu = E(\xi)$ è la media della popolazione ed $\bar{X}$ è la media campionaria

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

allora

- $E(\bar{X}) = \mu_{\bar{X}} = \mu$
- $E((X - \bar{X})^2) = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

11.2. Se ξ è distribuita normalmente con media μ e varianza σ^2 allora $\bar{X}$ è distribuito normalmente con media μ e varianza σ^2/n .

11.3. Se ξ è distribuita con media μ e varianza σ^2 (anche non normalmente) allora

$$\frac{\bar{X} - \mu}{\sigma\sqrt{n}}$$

è asintoticamente ($n \geq 30$) distribuito normalmente con media μ e varianza σ^2/n .

11.4. Se ξ è distribuita normalmente con media μ e se

$$S = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n}$$

è la varianza campionaria, allora la variabile aleatoria

$$\frac{\bar{X} - \mu}{S\sqrt{n-1}}$$

ha una distribuzione di probabilità di Student con $n - 1$ gradi di libertà.

11.5. Se ξ_1, ξ_2 sono due popolazioni normalmente distribuite con varianza σ_1^2, σ_2^2 e se X_1, X_2 sono due campioni di taglia m ed n rispettivamente estratti dalle due popolazioni, allora la variabile aleatoria

$$\frac{(mS_1^2)/((m-1)\sigma_1^2)}{nS_2^2/((n-1)\sigma_2^2)}$$

ha una distribuzione di probabilità di Fisher con $n-1, m-1$ gradi di libertà.

11.6. Supponiamo di disporre di un certo numero di dati in base ai quali la probabilità di accadimento di un evento A_k è x_k e supponiamo di voler verificare se la distribuzione di probabilità di tali eventi ha una funzione densità di probabilità ξ in base alla quale la probabilità attesa degli eventi A_k è ξ_k .

Possiamo considerare che la variabile aleatoria ξ descrive la popolazione degli eventi possibili di cui i valori x_k costituiscono un campione e possiamo considerare come stimatore la variabile aleatoria definita da

$$\chi^2 = \frac{(\xi_1 - x_1)^2}{x_1} + \frac{(\xi_2 - x_2)^2}{x_2} + \dots + \frac{(\xi_n - x_n)^2}{x_n}$$

a proposito della quale possiamo affermare che ha una distribuzione di probabilità χ^2

- $n-1$ gradi di libertà, nel caso in cui della distribuzione ξ siano noti i parametri
- $n-1-m$ gradi di libertà, nel caso in cui della distribuzione ξ si debbano ricavare i parametri per mezzo di m dati dei riassunti campionari.

Per avere informazioni su una variabile aleatoria è necessario conoscere la distribuzione di probabilità.

Possiamo, per analogia con quanto visto in precedenti casi, congetturare che la variabile aleatoria in esame abbia una particolare distribuzione di probabilità, tuttavia avremo successivamente la necessità di verificare se tale congettura è ragionevole.

In altri casi può accadere di aver bisogno di verificare la ragionevolezza di una ipotesi H_0 mediante affermazioni del tipo se H_0 è vera allora la variabile aleatoria ξ ha una distribuzione di probabilità di un certo tipo (Gaussiana, χ^2, T, F)

In ogni caso, stabilita la funzione di distribuzione di probabilità che si vuole sottoporre a test, si procede come segue:

- si individua la precisione ϵ con cui si vuole procedere: se ad esempio si fissa $\epsilon = .05$ significa che si vuole essere certi al 95%
- si individuano α, β in modo che

$$\mathcal{P}(\alpha \leq \xi \leq \beta) = .95$$

- si estrae un valore per ξ e se tale valore cade in $[\alpha, \beta]$ si accetta l'ipotesi, in caso contrario si rifiuta.

11.7. Il Test χ^2 . supponiamo di disporre di una serie di dati e di voler verificare se essi sono distribuiti in accordo con una certa ipotesi che chiamiamo H_1 .

Come esempio possiamo adottare il famoso esperimento di Mendel:

Mendel incrociò tra di loro due tipi di piselli ciascuno dei quali aveva due caratteri

- la forma del seme: (Liscio o angoloso, L oppure r)
- il colore del seme: (Giallo o verde, G oppure v)

Egli supponeva che

- Liscio e Giallo fossero caratteri dominanti (distinti dall'iniziale Maiuscola)
- ruvido e verde fossero caratteri recessivi (iniziale minuscola)

egli supponeva cioè che dall'incrocio di due piante si ottenesse una terza pianta con i caratteri scelti tra i dominanti delle due che l'avevano generata.

Mendel ottenne 556 piselli che classificò secondo i caratteri appena descritti nella seguente tabella

Caratteri	Frequenza osservata
Rotondo Giallo	315
Rotondo verde	108
angoloso Giallo	101
angoloso verde	32

TABELLA 7.1. Risultati dell'esperimento di Mendel.

I tipi di piselli possibili sono pertanto 4 ed i possibili incroci sono 16 e possono essere elencati nella tabella 7.1 dove accanto alle caratteristiche delle piante A e B incrociate sono riportate le caratteristiche dell'incrocio.

A	+	B	→	C
LG	+	LG	→	LG
LG	+	Lv	→	LG
LG	+	rG	→	LG
LG	+	rv	→	LG
Lv	+	LG	→	LG
Lv	+	Lv	→	Lv
Lv	+	rG	→	LG
Lv	+	rv	→	Lv
rG	+	LG	→	LG
rG	+	Lv	→	LG

rG	+	rG	→	rG
rG	+	rv	→	rG
rv	+	LG	→	LG
rv	+	Lv	→	Lv
rv	+	rG	→	rG
rv	+	rv	→	rv

TABELLA 7.2. Possibili incroci.

Dalla tabella si evince che

- un pisello Liscio e Giallo si presenta in 9 casi su 16
- un pisello Liscio e verde si presenta in 3 casi su 16
- un pisello rugoso e Giallo si presenta in 3 casi su 16
- un pisello rugoso e verde si presenta in 1 caso su 16

per cui possiamo dire che

$$\mathcal{P}(LG) = \frac{9}{16} 556 = 321.75$$

$$\mathcal{P}(Lv) = \frac{3}{16} 556 = 104.75$$

$$\mathcal{P}(rG) = \frac{3}{16} 556 = 104.75$$

$$\mathcal{P}(rv) = \frac{1}{16} 556 = 34.75$$

e possiamo affiancare alle frequenze osservate nella tabella 7.1 le frequenze previste per ciascun caso:

Caratteri	Frequenza osservata	frequenza prevista	Differenza
	f_o	f_p	$f_o - f_p$
Liscio Giallo	315	312.75	2.25
Liscio verde	108	104.25	3.75
rugoso Giallo	101	104.25	-3.25
rugoso verde	32	34.75	-2.75

TABELLA 7.3. Risultati dell'esperimento di Mendel e frequenze previste.

Possiamo ora chiederci se i dati sperimentali confermano l'ipotesi che abbiamo fatto sui caratteri delle due piante:

Liscio dominante e rugoso recessivo,

Giallo dominante e verde recessivo.

Si può vedere che è lecito supporre che la variabile aleatoria

$$\xi_k = \frac{f_o - f_p}{\sqrt{f_p}}, \quad k = 1, 2, 3, 4$$

segue una distribuzione gaussiana con media 0 e varianza 1.

Pertanto la somma dei quadrati delle 4 variabili aleatorie indicate in tabella 7.3 segue una distribuzione χ^2 con $3 = 4 - 1$ gradi di libertà.

Calcoliamo quindi

$$\chi^2 = \frac{(315 - 312.75)^2}{312.75} + \frac{(108 - 104.25)^2}{104.25} + \frac{(101 - 104.25)^2}{104.25} + \frac{(32 - 34.75)^2}{34.75} \approx 0.47$$

e consideriamo la tavola che fornisce i valori di χ^2 con 3 gradi di libertà.

Possiamo osservare che la tavola fornisce l'indicazione che

$$\mathcal{P}(\chi^2 < .352) = .05$$

$$\mathcal{P}(\chi^2 < .584) = .10$$

$$\mathcal{P}(\chi^2 < .95) = 7.81$$

$$\mathcal{P}(\chi^2 < .584) = 11.3$$

per cui

$$.352 < .47 < .584 < 7.81 < 11.3$$

e pertanto possiamo concludere che se gli eventi fossero casuali avremmo trovato un evento che accade con probabilità maggiore del 5% ma minore del 10%, per cui possiamo affermare che l'ipotesi fatta può essere accettata al livello del 90% ma deve essere rifiutata al livello del 95%

Un secondo esempio di applicazione del test χ^2 si trova considerando il seguente problema:

Si supponga di lanciare un dado a forma di tetraedro per 100 volte e di ottenere le seguenti uscite:

1	2	3	4
23	26	24	27

TABELLA 7.4

Ci si pone il problema di stabilire se il dado è non truccato.

La frequenza attesa per ciascuno dei punteggi è $\frac{100}{4} = 25$ e possiamo quindi calcolare

$$\chi^2 = \frac{(23 - 25)^2}{25} + \frac{(26 - 25)^2}{25} + \frac{(24 - 25)^2}{25} + \frac{(27 - 25)^2}{25} \approx .4$$

dalla tabella dei valori di χ^2 con $4 - 1 = 3$ gradi di libertà si ottiene che

$$\mathcal{P}(\chi^2 < .352) = .05$$

$$\mathcal{P}(\chi^2 < .584) = .10$$

per cui

$$\mathcal{P}(\chi^2 < .352) < \mathcal{P}(\chi^2 < .4) < \mathcal{P}(\chi^2 < .584)$$

per cui, come prima, se il dado fosse truccato, avremmo sperimentato un evento la cui probabilità di accadimento è compresa tra il 5% ed il 10%,

per cui possiamo affermare che l'ipotesi fatta può essere accettata al livello del 90% ma deve essere rifiutata al livello del 95%

CAPITOLO 8

IL TEOREMA DEL LIMITE CENTRALE

Il teorema del limite centrale è un risultato di grande importanza in quanto sancisce il fatto che la sovrapposizione di un gran numero di variabili aleatorie aventi media e varianza comune conduce ad una variabile con distribuzione normale (gaussiana).

Più precisamente possiamo dire che

Siano

$$\xi_1, \xi_2, \dots, \xi_n$$

variabili aleatorie indipendenti aventi la stessa distribuzione di probabilità con media μ e varianza σ^2 .

Siano

$$\eta_1, \eta_2, \dots, \eta_n$$

le corrispondenti variabili normalizzate

$$\eta_k = \frac{\xi_k - \mu}{\sigma}$$

e consideriamo la variabile aleatoria ζ definita da

$$\zeta_n = \frac{\eta_1 + \eta_2 + \dots + \eta_n}{\sqrt{n}}$$

(Poichè $E(\eta_i) = 0$ avremo che $E(\zeta_n) = nE(\eta_1 + \eta_2 + \dots + \eta_n) = 0$ e inoltre, poichè $\text{Var}(\eta_i) = 1$ avremo che $\text{Var}(\eta_1 + \eta_2 + \dots + \eta_n) = n$ e quindi $\text{Var}(\zeta_n) = 1$)

Si ha

$$\zeta_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n - n\mu}{\sqrt{n}\sigma}$$

e si può dimostrare che

$$\lim_{n \rightarrow +\infty} \mathcal{P}(\alpha \leq \zeta_n \leq \beta) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-x^2/2} dx$$

o equivalentemente che

$$\lim_{n \rightarrow +\infty} \mathcal{P}(a \leq \xi_1 + \xi_2 + \dots + \xi_n \leq b) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-n\mu}{\sqrt{n}\sigma}}^{\frac{b-n\mu}{\sqrt{n}\sigma}} e^{-x^2/2} dx$$

1. Un esempio di applicazione

Dal teorema del limite centrale risulta che moltissime variabili aleatorie possono essere approssimate mediante una variabile che abbia distribuzione, gaussiana.

Se ad esempio consideriamo la variabile aleatoria ξ che restituisce il numero di successi in una serie di N prove bernoulliane con probabilità di successo p e probabilità di insuccesso $q = 1 - p$, possiamo calcolare la probabilità di ottenere un numero di successi compreso tra n ed m utilizzando la distribuzione di Bernoulli mediante la seguente formula

$$\mathcal{P}(n \leq \xi \leq m) = \sum_{k=n}^m \binom{N}{k} p^k q^{N-k}$$

È subito evidente che, se $m - n$ è grande non è facile portare a termine il calcolo; per ovviare a questo inconveniente possiamo allora utilizzare il teorema del limite centrale e sostituire alla distribuzione di Bernoulli una distribuzione Gaussiana.

Procederemo considerando ξ come la somma di N variabili Bernoulliane ξ_i

$$\xi = \xi_1 + \xi_2 + \xi_3 + \dots + \xi_N$$

che assumono ciascuna valori 0 con probabilità q ed 1 con probabilità p .

La media di ciascuna delle ξ_i sarà $\mu = p$ e la loro varianza sarà $\sigma = \sqrt{pq}$, pertanto

$$\begin{aligned} \mathcal{P}(n \leq \xi \leq m) &= \mathcal{P}(n \leq \xi_1 + \xi_2 + \xi_3 + \dots + \xi_N \leq m) \approx \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{\frac{n-Np}{\sqrt{N}\sqrt{pq}}}^{\frac{m-Np}{\sqrt{N}\sqrt{pq}}} e^{-x^2/2} dx = G\left(\frac{m-Np}{\sqrt{N}\sqrt{pq}}\right) - G\left(\frac{n-Np}{\sqrt{N}\sqrt{pq}}\right) \end{aligned}$$

dove

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2/2} dx$$

è la funzione di distribuzione cumulativa Gaussiana standardizzata i cui valori si trovano tabulati.

Se ad esempio supponiamo di effettuare $N = 100$ lanci di una moneta non truccata per cui $p = q = 1/2$, possiamo valutare la probabilità che si abbiano un numero di Teste compreso tra $n = 45$ e $m = 55$ mediante la

$$\begin{aligned} \mathcal{P}(45 \leq \xi \leq 55) &\approx G\left(\frac{55 - 50}{\sqrt{100}\sqrt{1/4}}\right) - G\left(\frac{45 - 50}{\sqrt{100}\sqrt{1/4}}\right) = \\ &G(1) - G(-1) \approx 0.68 \end{aligned}$$

Abbiamo visto che, se la moneta non è truccata, su 100 lanci possiamo aspettarci un numero di successi compreso tra 45 e 55 con probabilità

del 68% e quindi potremo dire che la probabilità che i successi non siano compresi tra 45 e 55 è del 32%.

Per verificare se la moneta è truccata con un livello di significatività del 32% potremmo allora effettuare una serie di lanci e rifiutare l'ipotesi della moneta non truccata se otteniamo un numero di successi maggiore di 55 o minore di 45.

Qualora volessimo una maggiore precisione, ad esempio del 60% dovremmo avere

$$G\left(\frac{m - Np}{\sqrt{N}\sqrt{pq}}\right) - G\left(\frac{n - Np}{\sqrt{N}\sqrt{pq}}\right) = .40$$

e quindi dovrebbe risultare

$$G\left(\frac{m - Np}{\sqrt{N}\sqrt{pq}}\right) = .5 + .2 = .7 \quad \text{e} \quad G\left(\frac{n - Np}{\sqrt{N}\sqrt{pq}}\right) = .5 - .2 = .3$$

da cui

$$\frac{m - Np}{\sqrt{N}\sqrt{pq}} = .53 \quad \text{e} \quad \frac{n - Np}{\sqrt{N}\sqrt{pq}} = -.53$$

Se ne ricava che

$$m = 50 + 5 * (.53) \approx 52 \quad \quad n = 50 - 5 * (.53) \approx 48$$

CAPITOLO 9

REGRESSIONE LINEARE: LA RETTA DEI MINIMI QUADRATI

Siano assegnate n coppie di dati (punti di $\mathbb{R}^2$)

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

e si consideri il problema di determinare l'equazione di una retta

$$y = ax + b$$

in corrispondenza della quale risulti minima la quantità

$$\epsilon(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$$

$\epsilon(a, b)$ è una funzione convessa delle variabili (a, b) che tende a $+\infty$ per $(a, b) \rightarrow \infty$ e pertanto ammette uno ed un solo punto di minimo assoluto che si può trovare annullando $\nabla \epsilon$.

Per risolvere il problema dovremo pertanto risolvere il sistema definito dalle equazioni

$$\begin{cases} \frac{\partial \epsilon}{\partial a} = \sum_{i=1}^n -2(y_i - ax_i - b)x_i = 0 \\ \frac{\partial \epsilon}{\partial b} = \sum_{i=1}^n -2(y_i - ax_i - b) = 0 \end{cases}$$

Ne viene che

$$\begin{cases} \sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0 \\ \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - b \sum_{i=1}^n 1 = 0 \end{cases}$$

ovvero

$$(9.1) \quad \begin{cases} \sum_{i=1}^n x_i y_i = a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i = a \sum_{i=1}^n x_i + nb \end{cases}$$

Dalla seconda delle 9.1 si può vedere che

$$nb = \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i$$

ed anche

$$b = \frac{\sum_{i=1}^n y_i}{n} - a \frac{\sum_{i=1}^n x_i}{n} = \bar{y} - a\bar{x}$$

dove

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

indicano la media dei valori x_i ed y_i , rispettivamente.

Dalla prima delle 9.1 si può invece ottenere che

$$\begin{aligned} \sum_{i=1}^n x_i y_i &= a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \\ &= a \sum_{i=1}^n x_i^2 + \sum_{i=1}^n x_i \left(\frac{\sum_{i=1}^n y_i}{n} - a \frac{\sum_{i=1}^n x_i}{n} \right) \end{aligned}$$

e

$$\begin{aligned} \sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n} &= a \left(\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right) \\ n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i &= a \left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right) \end{aligned}$$

ed infine

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

Inoltre

$$\begin{aligned} nb &= \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \left(\frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \right) = \\ &= \frac{1}{n (\sum_{i=1}^n x_i)^2 - (\sum_{i=1}^n x_i)^2} \left[n \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \left(\sum_{i=1}^n x_i \right)^2 \sum_{i=1}^n y_i \right. \\ &\quad \left. - n \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i + \left(\sum_{i=1}^n x_i \right)^2 \sum_{i=1}^n y_i \right] \end{aligned}$$

e se ne conclude che

$$b = \frac{n \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - n \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

Ora, tenendo conto che

$$\begin{aligned}
 \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= \\
 &= \sum_{i=1}^n x_i y_i - \sum_{i=1}^n \bar{x} y_i - \sum_{i=1}^n x_i \bar{y} + \sum_{i=1}^n \bar{x} \bar{y} = \\
 &= \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} - n\bar{x}\bar{y} + n\bar{x}\bar{y} = \\
 &= \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} = \sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}
 \end{aligned}$$

e che

$$\begin{aligned}
 \sum_{i=1}^n (x_i - \bar{x})^2 &= \sum_{i=1}^n x_i^2 - 2 \sum_{i=1}^n \bar{x} x_i + \sum_{i=1}^n \bar{x}^2 = \\
 &= \sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2
 \end{aligned}$$

si ricava che

$$\begin{aligned}
 \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} &= \\
 &= \frac{1}{n} \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \\
 &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - n^2 \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2} = a
 \end{aligned}$$

Pertanto possiamo esprimere a e b mediante le seguenti formule

$$(9.2) \quad \begin{cases} a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ \bar{y} = a\bar{x} + b \end{cases}$$

La prima delle due uguaglianze premette di concludere che a è invariante rispetto alla traslazione degli assi: cioè usando $x - x_0$ ed $y - y_0$ in luogo di x ed y il valore di a non cambia.

La stessa trasformazione cambia invece il valore di b , come si vede dalla seconda uguaglianza da cui si vede anche che la retta di regressione passa per il punto di coordinate $(\bar{x}, \bar{y})$ che è il baricentro dei dati.

Possiamo anche osservare che a meno di operare una traslazione dei dati riportando l'origine degli assi nel baricentro $(\bar{x}, \bar{y})$, possiamo supporre che

$$(9.3) \quad \begin{cases} a = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \\ b = 0 \end{cases}$$

Ora, siano

- s_x^2 la varianza dei dati x_i

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

- s_y^2 la varianza dei dati y_i

$$s_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$$

- s_{xy} la covarianza dei dati (x_i, y_i)

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

Possiamo scrivere la retta di regressione nella forma

$$y - \bar{y} = \frac{s_{xy}}{s_x^2} (x - \bar{x})$$

e se invertiamo il ruolo di x e di y l'equazione diventa

$$x - \bar{x} = \frac{s_{xy}}{s_y^2} (y - \bar{y})$$

Possiamo misurare la correlazione tra i dati utilizzando il coefficiente definito da

$$(9.4) \quad r = \frac{s_{xy}}{s_x s_y}$$

mediante il quale le equazioni delle due rette prima introdotte diventano

$$\frac{y - \bar{y}}{s_y} = r \frac{x - \bar{x}}{s_x}$$

e

$$r \frac{y - \bar{y}}{s_y} = \frac{x - \bar{x}}{s_x}$$

Chiaramente le due rette coincidono soltanto nel caso in cui

$$r^2 = 1 \quad \text{cioè} \quad r = \pm 1$$

e il fatto che questo accada è indice della correlazione dei dati cioè del fatto che i dati si trovano su una retta.

È ragionevole quindi stimare la maggiore o minore correlazione tra i dati confrontando r^2 con 1: più r^2 è vicino ad 1 e più i dati sono da considerarsi linearmente correlati.

Possiamo inoltre misurare la dispersione dei dati attorno alla retta di regressione mediante la

$$S_{y,x} = \sqrt{\frac{\sum_{i=1}^n (y_i - y_i^s)^2}{n}}$$

dove

$$y_i^s = ax_i + b$$

Pertanto

$$\begin{aligned} S_{y,x}^2 &= \frac{\sum_{i=1}^n (y_i - ax_i - b)^2}{n} = \sum_{i=1}^n (y_i^2 + a^2 x_i^2 + b^2 - 2ax_i y_i - 2by_i + 2abx_i) = \\ &= \sum_{i=1}^n y_i^2 + a^2 \sum_{i=1}^n x_i^2 + nb^2 - 2a \sum_{i=1}^n x_i y_i - 2b \sum_{i=1}^n y_i + 2ab \sum_{i=1}^n x_i = \end{aligned}$$

e per la definizione di a e b

$$\begin{aligned} &= \sum_{i=1}^n y_i^2 + a^2 \left[\frac{1}{a} \left(\sum_{i=1}^n x_i y_i - b \sum_{i=1}^n x_i \right) \right] + nb^2 - \\ &\quad - 2a \sum_{i=1}^n x_i y_i - 2b \sum_{i=1}^n y_i + 2ab \left[\frac{1}{a} \left(\sum_{i=1}^n y_i - nb \right) \right] = \\ &= \sum_{i=1}^n y_i^2 + a \sum_{i=1}^n x_i y_i - ab \sum_{i=1}^n x_i + nb^2 - \\ &\quad - 2a \sum_{i=1}^n x_i y_i - 2b \sum_{i=1}^n y_i + 2b \sum_{i=1}^n y_i - 2nb^2 = \\ &= \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i - ab \sum_{i=1}^n x_i - nb^2 = \\ &= \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i - b \left(a \sum_{i=1}^n x_i + nb \right) = \end{aligned}$$

e ancora per la definizione di a e b

$$= \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i - b \sum_{i=1}^n y_i$$

Quindi

$$S_{y,x}^2 = \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i - b \sum_{i=1}^n y_i$$

Si può d'altro canto verificare che

$$\sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i - b \sum_{i=1}^n y_i = \sum_{i=1}^n (y_i - \bar{y})^2 - a \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

infatti

$$\begin{aligned} & \sum_{i=1}^n (y_i - \bar{y})^2 - a \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \\ &= \sum_{i=1}^n y_i^2 - 2\bar{y} \sum_{i=1}^n y_i + n\bar{y}^2 - a \sum_{i=1}^n x_i y_i + a \sum_{i=1}^n x_i \bar{y} + \\ & \quad + a \sum_{i=1}^n \bar{x} y_i - a \sum_{i=1}^n \bar{x} \bar{y} = \\ &= \sum_{i=1}^n y_i^2 - 2n\bar{y}^2 + n\bar{y}^2 - a \sum_{i=1}^n x_i y_i + an\bar{x}\bar{y} + an\bar{x}\bar{y} - an\bar{x}\bar{y} = \\ &= \sum_{i=1}^n y_i^2 - n\bar{y}^2 - a \sum_{i=1}^n x_i y_i + an\bar{x}\bar{y} = \\ &= \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i + n\bar{y}(\bar{y} - a\bar{x}) = \\ &= \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i + n\bar{y}b = \\ &= \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i + b \sum_{i=1}^n y_i \end{aligned}$$

Le precedenti considerazioni permettono quindi di affermare che

$$\begin{aligned} S_{y,x}^2 &= \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i - b \sum_{i=1}^n y_i = \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - a \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \\ &= \frac{ns_y^2 - ans_{xy}}{n} = s_y^2 - as_{xy} \end{aligned}$$

e dal momento che $a = \frac{s_{xy}}{s_x}$

$$\begin{aligned} &= s_y^2 \left(1 - \frac{s_{xy}}{s_y^2} \right) = s_y^2 \left(1 - \frac{s_{xy}^2}{s_x^2 s_y^2} \right) = \\ & \quad s_y^2 (1 - r^2) \end{aligned}$$

Ne viene quindi che

$$\frac{S_{y,x}}{s_y^2} = (1 - r^2)$$

e

$$r^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_i^s)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

D'altro canto

$$\begin{aligned} \sum_{i=1}^n (y_i - \bar{y})^2 &= \sum_{i=1}^n (y_i - y_i^s + y_i^s - \bar{y})^2 = \\ &= \sum_{i=1}^n (y_i - y_i^s)^2 + \sum_{i=1}^n (y_i^s - \bar{y})^2 + 2 \sum_{i=1}^n (y_i - y_i^s)(y_i^s - \bar{y}) \end{aligned}$$

Poichè valgono le equazioni normali 9.1 che definiscono a e b

$$\begin{aligned} \sum_{i=1}^n (y_i - y_i^s)(y_i^s - \bar{y}) &= \sum_{i=1}^n (y_i - ax_i - b)(ax_i + b - \bar{y}) = \\ &= (b - \bar{y}) \sum_{i=1}^n (y_i - ax_i - b) + a \sum_{i=1}^n x_i (y_i - ax_i - b) \\ &= (b - \bar{y}) \left(\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - nb \right) + a \left[\sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i \right] = 0 \end{aligned}$$

avremo

$$\begin{aligned} r^2 &= 1 - \frac{\sum_{i=1}^n (y_i - y_i^s)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - y_i^s)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \\ &= \frac{\sum_{i=1}^n (y_i^s - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\text{Variazione spiegata}}{\text{Variazione totale}} \end{aligned}$$

Possiamo anche calcolare, dalla 9.4, che

$$\begin{aligned}
 r &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2\right) \left(n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i\right)^2\right)}} = \\
 &= \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{(\overline{x^2} - \bar{x}^2)(\overline{y^2} - \bar{y}^2)}} = \\
 &= \frac{n \sum_{i=1}^n x'_i y'_i - \sum_{i=1}^n x'_i \sum_{i=1}^n y'_i}{\sqrt{\left(n \sum_{i=1}^n x'^2_i - \left(\sum_{i=1}^n x'_i\right)^2\right) \left(n \sum_{i=1}^n y'^2_i - \left(\sum_{i=1}^n y'_i\right)^2\right)}} = \\
 &= \frac{\sum_{i=1}^n x'_i y'_i}{\sqrt{(\sum_{i=1}^n x'^2_i)(\sum_{i=1}^n y'^2_i)}}
 \end{aligned}$$

CAPITOLO 10

ANALISI DEI COMPONENTI PRINCIPALI.

Quando si esaminano i dati statistici relativi ad un elevato numero di caratteristiche di una popolazione è opportuno considerare i dati in un sistema di riferimento che ne metta in evidenza le caratteristiche salienti.

Supponiamo ad esempio di campionare una funzione f su un intervallo $[a, b]$; dividiamo cioè l'intervallo in $2n - 1$ parti uguali ed indichiamo con

$$a = t_1, t_2, \dots, t_{2n} = b$$

i punti mediante i quali tale suddivisione è operata.

Consideriamo poi n coppie di dati ottenute come segue

$$P_1 = (f(t_1), f(t_2)), P_2 = (f(t_3), f(t_4)), \dots, P_n = (f(t_{2n-1}), f(t_{2n}))$$

È ragionevole che, se la funzione è continua, ci siano piccole differenze tra ascissa e ordinata di ognuno dei punti P_k e pertanto essi tenderanno a disporsi lungo la bisettrice primo-terzo quadrante $y = x$.

Quindi potremo ottenere una rappresentazione più significativa dei dati se useremo come assi coordinati le bisettrici $y = x$ ed $y = -x$, cioè se opereremo una rotazione di 45° degli assi.

Ricordiamo che se $P = (x, y)$ sono le coordinate originali, si opera una rotazione degli assi applicando una trasformazione lineare la cui matrice di rappresentazione è

$$A = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

Pertanto

$$AP_1 = \left(\frac{\sqrt{2}}{2} f(t_1) + \frac{\sqrt{2}}{2} f(t_2), -\frac{\sqrt{2}}{2} f(t_1) + \frac{\sqrt{2}}{2} f(t_2) \right)$$

ed è evidente che delle due coordinate è rilevante solo la prima, mentre la seconda è piccola.

Ne viene che si possono ricostruire i valori originali con sufficiente precisione anche se le seconde coordinate dei punti trasformati sono trascurate.

Per illustrare il metodo occorre premettere qualche considerazione sulle forme quadratiche,

1. Forme quadratiche ed autovalori.

Sia A una matrice $n \times n$ e consideriamo la funzione

$$f(u) = \langle Au, u \rangle \quad , u \in \mathbb{R}^n$$

f si chiama forma quadratica su $\mathbb{R}^n$ e si può vedere che è sempre possibile supporre che la matrice A che la individua sia simmetrica.

Infatti, ad esempio per $n = 2$, se

$$A_1 = \begin{pmatrix} a & d_1 \\ d_2 & c \end{pmatrix}$$

ed $u = \begin{pmatrix} x \\ y \end{pmatrix}$ avremo che

$$\begin{aligned} f(u) = \langle A_1 u, u \rangle &= ax^2 + d_1 xy + d_2 yx + cy^2 = ax^2 + (d_1 + d_2)xy + cy^2 = \\ &= ax^2 + 2bxy + cy^2 = \langle Au, u \rangle \end{aligned}$$

per $b = (d_1 + d_2)/2$, da cui

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

D'altro canto, se A è una matrice simmetrica possiamo verificare che

$$\langle Au, v \rangle = \langle u, Av \rangle$$

ed inoltre

$$f(u+h) = \langle A(u+h), (u+h) \rangle = \langle Au, u \rangle + 2\langle Au, h \rangle - \langle Ah, h \rangle$$

per cui

$$f(u+h) - f(u) = 2\langle Au, h \rangle + \langle Ah, h \rangle = 2\langle Au, u \rangle + \|h\|\omega(h)$$

con ω funzione infinitesima per $h \rightarrow 0$ (si ricordi che $|\langle Ah, h \rangle| \leq \|Ah\|\|h\|$ ed $Ah \rightarrow 0$ se $h \rightarrow 0$).

Dalla definizione di differenziale si ottiene allora che f è differenziabile e che

$$\nabla f(u) = 2Au$$

Come caso particolare, se $A = I$ (la matrice identica) si ha

$$g(u) = \langle u, u \rangle = \|u\|^2 \quad , \quad \nabla g(u) = 2u$$

Consideriamo ora il problema di trovare

$$\max_{g(u)-1=0} f(u) = \max_{\|u\|^2-1=0} \langle Au, u \rangle$$

Per il teorema di Weierstraß, dal momento che f è continua e che $\|u\|^2 - 1 = 0$ definisce la superficie della sfera di centro l'origine e raggio 1, che è

chiusa e limitata, possiamo affermare che il massimo esiste. Sia u_1 il punto in cui tale massimo è assunto

$$\langle Au_1, u_1 \rangle = \max_{\|u\|^2=1} \langle Au, u \rangle$$

$$\|u_1\|^2 = 1$$

D'altro canto, il teorema dei moltiplicatori di Lagrange consente di affermare che esiste λ_1 tale che

$$\nabla f(u_1) = \lambda_1 \nabla g(u_1)$$

per cui deve essere

$$Au_1 = \lambda_1 u_1$$

Dal momento che la precedente equazione è soddisfatta

- λ_1 è un autovalore di A
- u_1 è un autovettore di A corrispondente all'autovalore λ_1

Possiamo inoltre osservare che

$$\max_{\|u\|^2=1} \langle Au, u \rangle = \langle Au_1, u_1 \rangle = \langle \lambda_1 u_1, u_1 \rangle = \lambda_1 \|u_1\|^2 = \lambda_1$$

per cui λ_1 è il valore massimo assunto da $\langle Au, u \rangle$ sulla sfera $\|u\|^2 = 1$

Consideriamo ora lo spazio vettoriale V_1 generato da u_1

$$V_1 = \{\lambda u_1 : \lambda \in \mathbb{R}\}$$

e lo spazio $V_1^\perp$ ortogonale a V_1

$$V_1^\perp = \{v \in \mathbb{R}^n : \langle v, u_1 \rangle = 0\}$$

$V_1^\perp$ è uno spazio vettoriale $(n-1)$ -dimensionale e la matrice A definisce una forma quadratica f_1 su $V_1^\perp$ in quanto

per ogni $v \in V_1^\perp$ si ha

$$\langle Av, u_1 \rangle = \langle v, Au_1 \rangle = \langle v, \lambda_1 u_1 \rangle = 0$$

e $Av \in V_1^\perp$

Se ora ripetiamo il procedimento per f_1 su $V_1^\perp$ possiamo provare un nuovo autovalore λ_2 ed un nuovo autovettore u_2 con le stesse caratteristiche di λ_1 e u_1 .

Chiaramente si può iterare fino a trovare:

- n autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$ decrescenti in valore
- n autovettori $u_1, u_2, \dots, u_n$, uno per ogni autovalore, che risultano ortogonali tra loro.

Gli autovettori $u_1, u_2, \dots, u_n$ formano quindi una base ortonormale con la caratteristica che lungo il primo asse si trova il punto di massimo della forma quadratica $\langle Au, u \rangle$ sulla sfera unitaria in $\mathbb{R}^2$, lungo il secondo asse si trova il massimo della forma quadratica $\langle Au, u \rangle$ sulla sfera unitaria in $V_1^\perp$ e così via fino all' n -esimo asse.

2. L'applicazione all'analisi delle componenti principali.

Torniamo ora ai nostri dati $P_k = (u_k^1, u_k^2, \dots, u_k^n) \in \mathbb{R}^n$ e cerchiamo di individuare una combinazione lineare delle componenti

$$Q_k = \alpha_1 u_k^1 + \alpha_2 u_k^2 + \dots + \alpha_n u_k^n$$

in modo che la varianza di Q_k sia massima (massima significatività della variabile).

Definiamo

$$v_i = \text{Var}(u_k^i)$$

$$c_{ij} = \text{Cov}(u_k^i, u_k^j)$$

la varianza delle singole componenti e la covarianza delle componenti a due a due e sia R la **matrice di covarianza** dei dati definita mediante la

$$R = \begin{pmatrix} v_1 & c_{12} & c_{13} & \cdots & c_{1n} \\ c_{21} & v_2 & c_{23} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ c_{n1} & c_{n2} & c_{n3} & \cdots & v_n \end{pmatrix}$$

Possiamo allora verificare che

$$\text{Var}(Q_k) = \langle Ra, a \rangle$$

dove

$$a = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

$\text{Var}(Q_k)$ è quindi una forma quadratica cui possiamo applicare il metodo visto nella precedente sezione e mediante tale metodo possiamo individuare in ordine decrescente di significatività le componenti dei dati.

2.1. Un esempio. Per illustrare gli effetti del metodo consideriamo i punti del grafico di $\sin(t)$ nell'intervallo $[0, 2\pi]$; suddividiamo l'intervallo in 199 parti uguali, in modo da individuare 200 punti in $[0, 2\pi]$ e calcoliamo i valori assunti da $\sin(t)$ in tali punti.

Le seguenti istruzioni di Matlab producono come risultato un vettore t che contiene i 200 punti in $[0, 2\pi]$, ed i vettori x ed y che contengono i valori assunti da $\sin(6t)$ nei punti pari e dispari rispettivamente

```
clear all;
step=2*pi/199;
t=0:step:2*pi;
xx=0:2*step: 2*pi;
yy=step:2*step: 2*pi;
x=sin(6*xx);
y=sin(6*yy);
```

Possiamo esaminare nel piano la distribuzione dei punti di coordinate $(x(k), y(k))$ con $k = 1, \dots, 100$ mediante le seguenti istruzioni

```
figure(1)
plot(x, y, ' + ' );
```

`axis([-2 2 -2 2]);`
 che producono la seguente figura 10.1

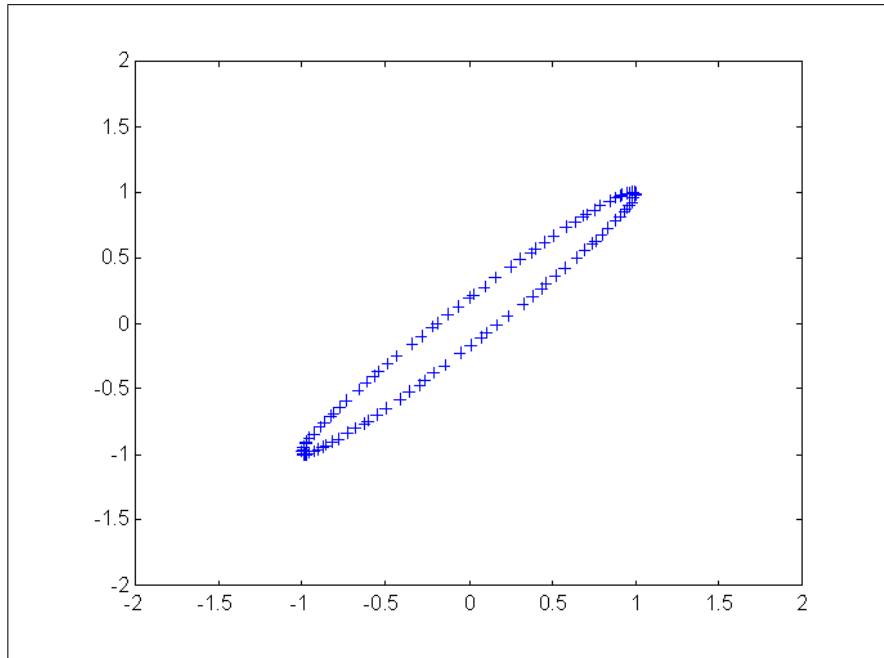


FIGURA 10.1.

Le istruzioni

```
R=cov(x,y);
[V,D]=eig(R)
```

calcolano la matrice di covarianza R e le matrici V e D , dove V è la matrice le cui colonne sono gli autovettori di R e D è una matrice diagonale con gli autovalori sulla diagonale principale; in altre parole V è la matrice tale che

$$VR = RD \quad \text{cioè tale che} \quad V^{-1}RV = D$$

La matrice V pertanto è la matrice di passaggio dal sistema di coordinate originale a quello individuato dagli autovalori di R .

Poichè ci interessa tener conto della componente relativa al massimo autovalore, di assicuriamo anche che l'autovalore più grande sia in posizione (1,1) mediante le istruzioni

```
if D(1,1)>D(2,2)
vv=V(:,1);
V(:,1)=V(:,2);
V(:,2)=vv;
end
```

La matrice V , quindi, può essere usata per effettuare un cambio di base che metta in evidenza le componenti principali.

Le seguenti istruzioni

```
tr=[x(:),y(:)]*V;
```

```

tr1=tr(:,1);
tr2=tr(:,2);
figure(2)
plot(tr1,tr2,'r+')
axis([-2 2 -2 2]);

```

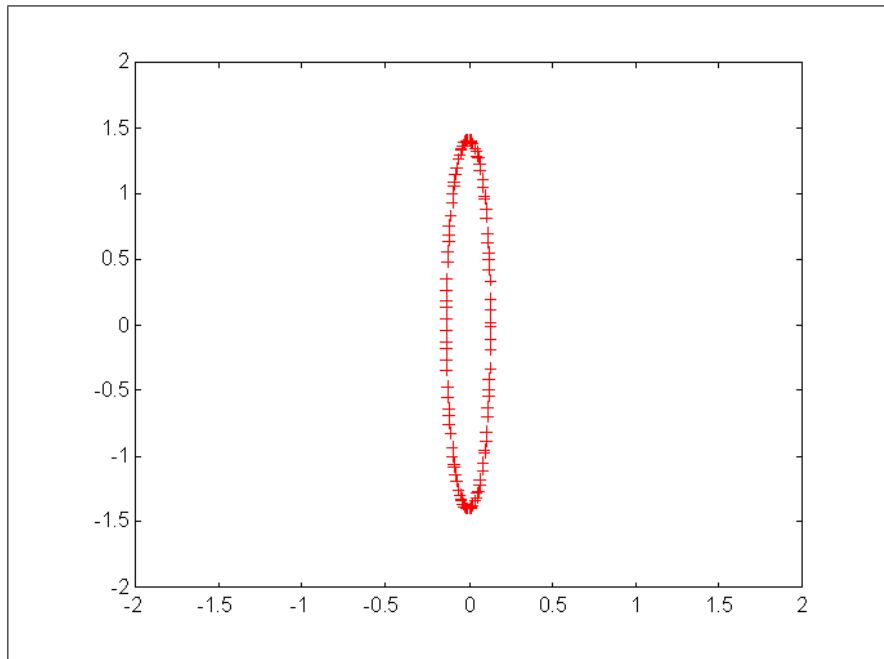


FIGURA 10.2.

Calcolano i trasformati $tr1$, $tr2$ dei punti (x, y) e li mostrano rispetto ad una coppia di assi ortogonali coincidenti con gli autovettori di R (si veda la figura 10.2).

Le successive istruzioni:

```

rtr=[tr1(:),tr2(:)]*inv(V);
rtr1=rtr(:,1);
rtr2=rtr(:,2);
figure(3)
plot(rtr1,rtr2,'g+')
axis([-2 2 -2 2]);

```

mostrano come applicando la trasformazione inversa i dati possano essere recuperati (si veda la figura 10.4.1).

Ora, possiamo osservare che la variazione della seconda componente dei dati trasformati è trascurabile rispetto alla prima, per cui, se la trascuriamo e applichiamo la trasformazione inversa ai dati privati di tale componente, otteniamo nuovi punti che differiscono di poco da quelli originali; possiamo operare usando le seguenti istruzioni:

```

nu=zeros(size(tr2));
ntr=[nu(:),tr2(:)]*inv(V);

```

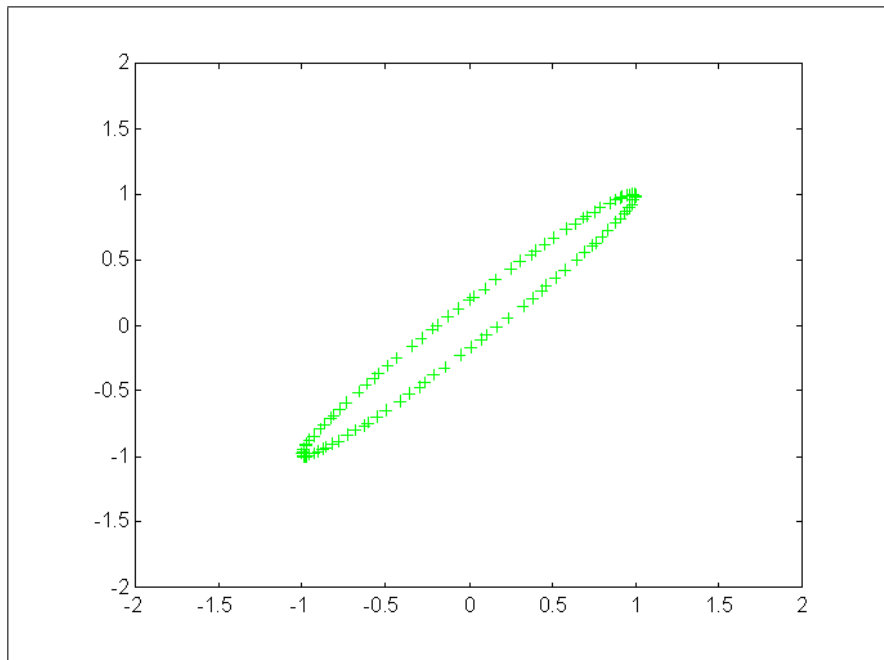


FIGURA 10.3.

```

ntr1=ntr(:,1);
ntr2=ntr(:,2);
figure(4)
plot(ntr1,ntr2,'rx',rtr1,rtr2,'g+')
axis([-2 2 -2 2]);

```

che forniscono anche una immagine dei nuovi punti come indicato in figura 10.4.1 ed un confronto con i punti originali 10.4.2.

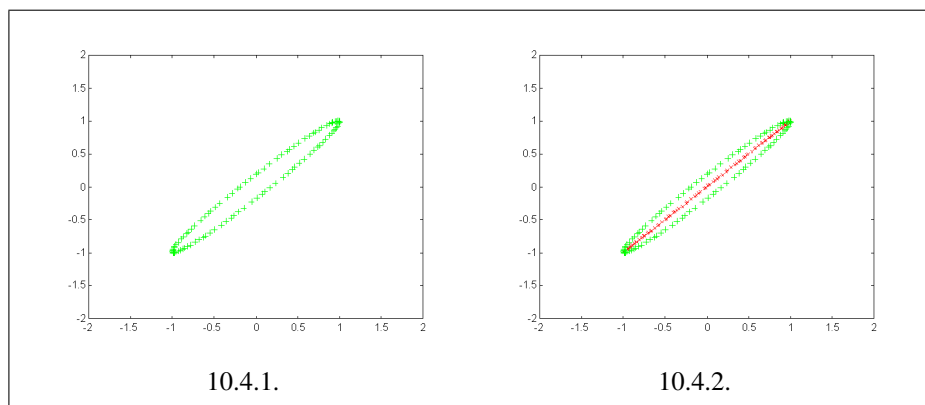


FIGURA 10.4.

È interessante ora osservare come dai punti originali e da quelli privati della componente meno significativa si possa ricostruire la funzione $\sin(6t)$

Le seguenti istruzioni

```
z=zeros(1,200);
```

```

zt=zeros(1,200);
for k=0:99
    zt(2*k+1)=x(k+1);
end
for k=1:100
    zt(2*k)=y(k);
end
for k=0:99
    z(2*k+1)=ntr2(k+1);
end
for k=1:100
    z(2*k)=ntr1(k);
end
figure(5)
plot(t,z)
figure(6)
plot(t,zt)

```

producono i due grafici riportati in figura 10.5, il primo dei quali riporta la funzione $\sin(t)$ ricostruita congiungendo con segmenti di retta i 200 punti originali, mentre la seconda riporta il grafico di $\sin(6t)$ ricostruito a partire dai punti ottenuti applicando la trasformazione inversa ai punti prima trasformati e poi privati della seconda componente.

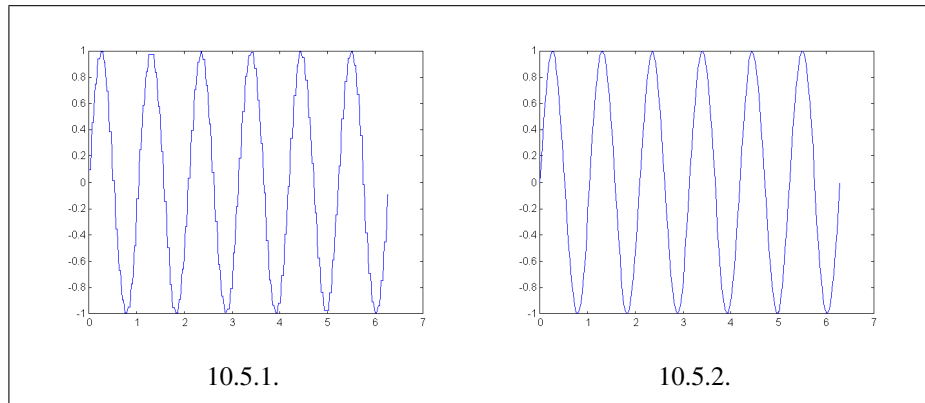


FIGURA 10.5.

Come si vede è evidente che la componente trascurata non ha peggiorato di molto il grafico, mentre la quantità di dati necessari a ricostruire l'immagine si è dimezzata.

Questo indica come può essere sviluppato un procedimento che consenta di immagazzinare dati (i punti del grafico della funzione) utilizzando al meglio le informazioni che contengono.

CAPITOLO 11

Appendice

1. LE FORMULE DI WALLIS E DI STIRLING-DE MOIVRE.

Intendiamo qui provare due formule che sono di grande utilità ed interesse.

La prima, nota come formula del prodotto di Wallis, consente di individuare una successione di razionali che ha per limite π , la seconda, dovuta a Stirling e De Moivre, consente di valutare, con notevole precisione, $n!$.

cominciamo con l'osservare che la definizione di fattoriale consente si affermare che

Si ha

$$(2n)!! = 2^n n! \quad , \quad (2n-1)!! = (2n)! \frac{q^2}{2^n n!}$$

Possiamo provare il seguente teorema

TEOREMA 11.1. -Wallis - Posto

$$w_n = ((2n)!!)^2 \frac{q^2}{((2n-1)!!)^2 (2n+1)}$$

si ha

$$\lim w_n = \pi \frac{q^2}{2}.$$

Infatti sia

$$s_n = \int_0^{\pi/2} (\sin x)^n dx$$

si ha

$$s_0 = \pi/2 \quad , \quad s_1 = 1$$

ed anche

$$\begin{aligned} s_n &= \int_0^{\pi/2} (\sin x)^{n-1} (\sin x) dx = -(\sin x)^{n-1} \cos x \Big|_0^{\pi/2} + \\ &\quad + \int_0^{\pi/2} (n-1)(\sin x)^{n-2} (\cos x)^2 dx = \\ &= (n-1) \int_0^{\pi/2} ((\sin x)^{n-2} - (\sin x)^n) dx = (n-1)(s_{n-2} - s_n) \end{aligned}$$

onde

$$s_n = n-1 \frac{q^2}{n} s_{n-2} \quad , \quad n \geq 2$$

Pertanto

$$\begin{aligned} s_{2n} &= 2n-1 \frac{q^2}{2n} s_{2n-2} = 2n-1 \frac{q^2}{2n} 2n-3 \frac{q^2}{2n-2} s_{2n-4} = \\ &= 2n-1 \frac{q^2}{2n} 2n-3 \frac{q^2}{2n-2} \dots 1 \frac{q^2}{2} s_0 = (2n-1)!! \frac{q^2}{(2n)!!} \pi \frac{q^2}{2} \end{aligned}$$

Analogamente

$$s_{2n+1} = 2n \frac{q^2}{2n+1} 2n-2 \frac{q^2}{2n-1} \dots 2 \frac{q^2}{3} s_1 = (2n)!! \frac{q^2}{(2n+1)!!}.$$

Ora, poiché

$$0 \leq s_n = \int_0^{\pi/2} (\sin x)^n dx \leq \pi \frac{q^2}{2},$$

mentre

$$s_{n+1} = \int_0^{\pi/2} (\sin x)^{n+1} dx \leq \int_0^{\pi/2} (\sin x)^n dx = s_n,$$

si ha che

$$\lim s_n = s \in \mathbb{R}.$$

Ma, come si verifica facilmente

$$w_n = s_{2n+1} \frac{q^2}{s_{2n}} \pi \frac{q^2}{2}$$

e pertanto

$$\lim w_n = s \frac{q^2}{s} \pi \frac{q^2}{2} = \pi \frac{q^2}{2}$$

Se ora osserviamo che si ha

$$w_n = (2^n n!)^2 (2^n n!)^2 \frac{q^2}{((2n)!)^2 (2n+1)}$$

da cui

$$v_n = (n!)^2 2^{2n} \frac{q^2}{(2n)! \sqrt{n}} = \sqrt{w_n (2n+1) \frac{q^2}{n}} \longrightarrow \sqrt{\pi}$$

possiamo concludere che

Si ha

$$(11.1) \quad \lim (n!)^2 2^{2n} \frac{q^2}{(2n)! \sqrt{n}} = \sqrt{\pi}$$

Come conseguenza della precedente formula possiamo provare la Formula di Stirling che è di grande utilità per approssimare i valori del fattoriale.

Formula di Stirling-De Moivre**Si ha**

$$\lim n^n e^{-n} \sqrt{2\pi n} \frac{q^2}{n!} = 1$$

da cui

$$n! = n^n e^{-n} \sqrt{2\pi n} \theta_n, \quad \text{con} \quad \lim \theta_n = 1.$$

La verifica della formula di Stirling procede come di seguito illustrato.

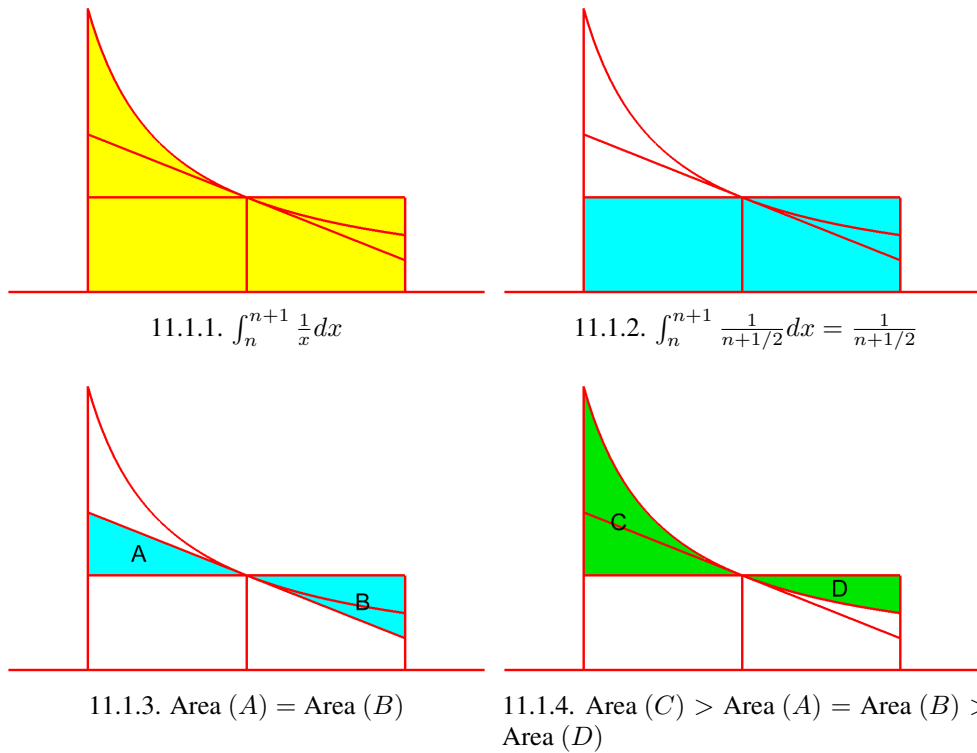


FIGURA 11.1. Grafici di funzioni di due variabili

consideriamo la funzione $1/x$ sull'intervallo $[n, n+1]$; la figura 11.1 mostra come sia vera la disuguaglianza

$$\frac{2}{2n+1} = \frac{1}{n+1/2} \leq \int_n^{n+1} \frac{dx}{x} = \ln \left(1 + \frac{1}{n} \right).$$

Ne segue che

$$1 \leq (n+1/2) \ln(1+1/n)$$

ed

$$e \leq \left(1 + \frac{1}{n} \right)^{n+1/2}.$$

Tenendo conto che $\ln$ è una funzione concava, si ottiene,

$$\begin{aligned} 1 + (\ln 2 + \ln 3 + \dots + \ln(n-1)) + \frac{1}{2} \ln n &\geq \\ &\geq \int_1^{3/2} \ln x dx + \int_{3/2}^{n-1/2} \ln x dx + \int_{n-1/2}^n \ln x dx = \\ &= \int_1^n \ln x dx = n \ln n - n + 1 = \\ &= \ln(n^n e^{-n}) + 1 \end{aligned}$$

da cui

$$\ln \left(\frac{n!}{\sqrt{n}} \right) \geq \ln(n^n e^{-n})$$

$$0 \leq u_n = \frac{n^n e^{-n} \sqrt{n}}{n!} \leq 1.$$

Pertanto u_n è limitata e possiamo di più provare che è crescente, verificando che $u_{n+1}/u_n \geq 1$. Si ha infatti

$$\frac{u_{n+1}}{u_n} = \frac{(n+1)^{n+1} e^{-n-1} \sqrt{n+1}}{(n+1)!} \frac{n!}{n^n e^{-n} \sqrt{n}} = \frac{1}{e} \left(1 + \frac{1}{n} \right)^{n+1/2} \geq 1.$$

Possiamo a questo punto affermare che

$$u_n \rightarrow u \in \mathbb{R}$$

e ne consegue che

$$\frac{u_n^2}{u_{2n}} \rightarrow u.$$

Ma

$$\frac{u_n^2}{u_{2n}} = \frac{n^{2n} e^{-2n} n}{(n!)^2} \frac{(2n)!}{(2n)^{2n} e^{-2n} \sqrt{2n}} = \frac{(2n)! \sqrt{n}}{(n!)^2 2^{2n}} \frac{1}{\sqrt{2}}.$$

Da cui, per la 11.1

$$u = \lim \frac{u_n^2}{u_{2n}} = \lim \frac{1}{v_n \sqrt{2}} = \frac{1}{\sqrt{2\pi}}$$

e si conclude che

$$\frac{n^n e^{-n} \sqrt{n}}{n!} \longrightarrow \frac{1}{\sqrt{2\pi}}$$

Possiamo anche osservare che,

$\forall n \in \mathbb{N}$

$$.367879 \leq \frac{1}{e} = u_1 \leq u_n \leq \lim u_n = \frac{1}{\sqrt{2\pi}} \leq .398943$$

CAPITOLO 12

QUALCHE ARGOMENTO DI PROBABILITÀ E STATISTICA.

1. Il Lancio di un Dado

Consideriamo un dado a forma di esaedro (cubo) con le facce numerate da 1 a 6; identifichiamo con x_k l'evento

è stato ottenuto il punteggio k

(la faccia superiore del dado mostra k), $k = 1..6$.

Evidentemente possiamo costruire una tabella in cui ad ogni evento si fa corrispondere la sua probabilità.

Evento	Probabilità
x_1	$1/6$
x_2	$1/6$
x_3	$1/6$
x_4	$1/6$
x_5	$1/6$
x_6	$1/6$

In altre parole

$$P(x_k) = \frac{1}{6} = p_k$$

Calcoliamo il **valor medio** μ e la **varianza** σ^2 della variabile aleatoria che restituisce il valore del punto ottenuto.

Avremo

$$\mu = \sum_{k=1}^6 \frac{1}{6} k = \frac{1}{6} \left(\frac{7 * 6}{2} \right) = \frac{1}{6} (1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2} = 3.5$$

$$\begin{aligned} \sigma^2 &= \sum_{k=1}^6 \frac{1}{6} \left(k - \frac{7}{2} \right)^2 = \frac{1}{6} \left(\sum_{k=1}^6 k^2 - 2 \frac{7}{2} \sum_{k=1}^6 k + \sum_{k=1}^6 \frac{49}{4} \right) = \\ &= \frac{91}{6} - 2 \frac{49}{4} + \frac{49}{4} = \frac{364 - 294}{24} = \frac{35}{12} \end{aligned}$$

In generale possiamo osservare che

$$\mu = \sum_{k=1}^6 x_k p_k$$

e

$$\begin{aligned}
\sigma^2 &= \sum_{k=1}^6 (x_k - \mu)^2 p_k = \sum_{k=1}^6 (x_k^2 - 2\mu x_k + \mu^2) p_k = \\
&= \sum_{k=1}^6 x_k^2 p_k - 2\mu \sum_{k=1}^6 x_k p_k + \mu^2 \sum_{k=1}^6 p_k = \sum_{k=1}^6 x_k^2 p_k - 2\mu^2 + \mu^2 = \\
&= \sum_{k=1}^6 x_k^2 p_k - \mu^2
\end{aligned}$$

Per cui

$$\sigma^2 = \sum_{k=1}^6 k^2 \frac{1}{6} - \left(\frac{7}{2}\right)^2 = \frac{6 * 7 * 13}{6} - \frac{49}{4} = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

2. Lancio di due Dadi

Supponiamo di lanciare due dadi le cui facce sono numerate come d'uso da 1 a 6 e cominciamo con l'individuare lo spazio di tutti i possibili eventi.

Possiamo allo scopo usare la seguente tabella.

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Gli eventi possibili sono 36 ciascuno dei quali ha la stessa probabilità di uscita, per cui, se indichiamo con

$$D_{i,j} = (i, j)$$

l'evento che si verifica quando sul primo dado esce i e sul secondo dado j , avremo

$$P(D_{i,j}) = \frac{1}{36}$$

L'insieme Ω di tutti gli eventi possibili definisce lo spazio di probabilità in cui operiamo.

Gli elementi di Ω sono gli eventi elementari $D_{i,j}$ e tutti gli eventi che si ottengono come unione di eventi elementari; ad esempio

$$P(\xi = 2) = P(\xi = 12) = \frac{1}{36}$$

$$P(\xi = 3) = P(\xi = 11) = \frac{2}{36}$$

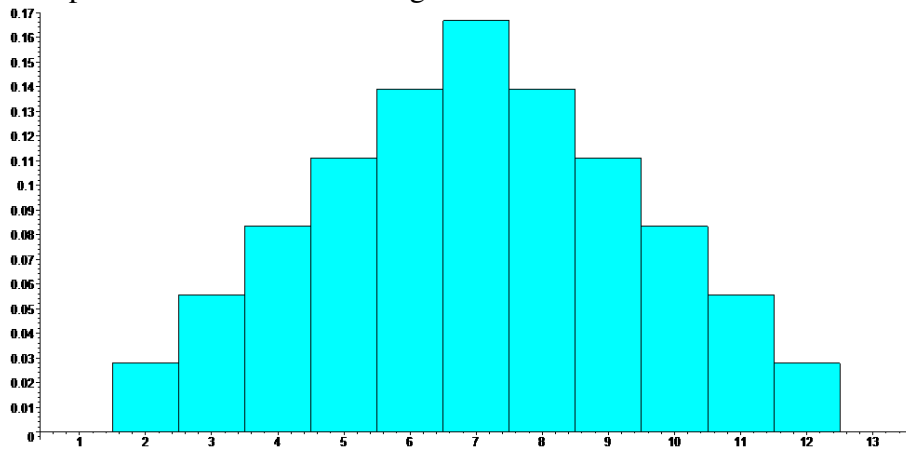
$$P(\xi = 4) = P(\xi = 10) = \frac{3}{36}$$

$$P(\xi = 5) = P(\xi = 9) = \frac{4}{36}$$

$$P(\xi = 6) = P(\xi = 8) = \frac{5}{36}$$

$$P(\xi = 7) = \frac{6}{36}$$

e riportate i risultati su un istogramma.

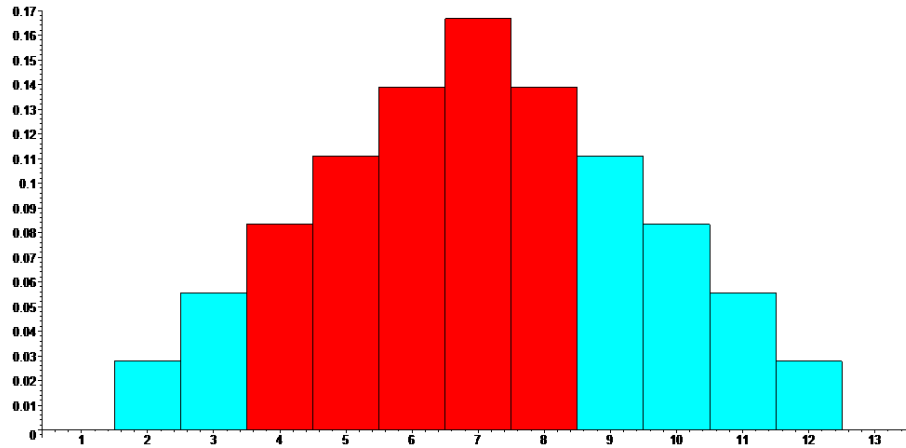


Ad esempio possiamo calcolare

$$P(\xi = 7) = \frac{6}{36} = \frac{1}{6}$$

$$P(\xi \neq 7) = 1 - P(\xi = 7) = 1 - \frac{1}{6} = \frac{5}{6}$$

$$P(4 \leq \xi \leq 8) = P(\xi = 4) + P(\xi = 5) + P(\xi = 6) + P(\xi = 7) + P(\xi = 8) = \frac{3 + 4 + 5 + 6 + 5}{36} = \frac{23}{36}$$



La media μ della variabile aleatoria ξ si può calcolare sommando la media del punteggio del primo dado con quella del punteggio del secondo dado:

$$\mu = \text{media del primo dado} + \text{media del secondo dado} = \frac{7}{2} + \frac{7}{2} = 7$$

oppure direttamente

$$\begin{aligned} \mu &= \frac{2 + 12 + 2(3 + 11) + 3(4 + 10) + 4(5 + 9) + 5(6 + 8) + 6 * 7}{36} = \\ &= 14 \frac{1 + 2 + 3 + 4 + 5 + 3}{36} = 14 \frac{18}{36} = 7 \end{aligned}$$

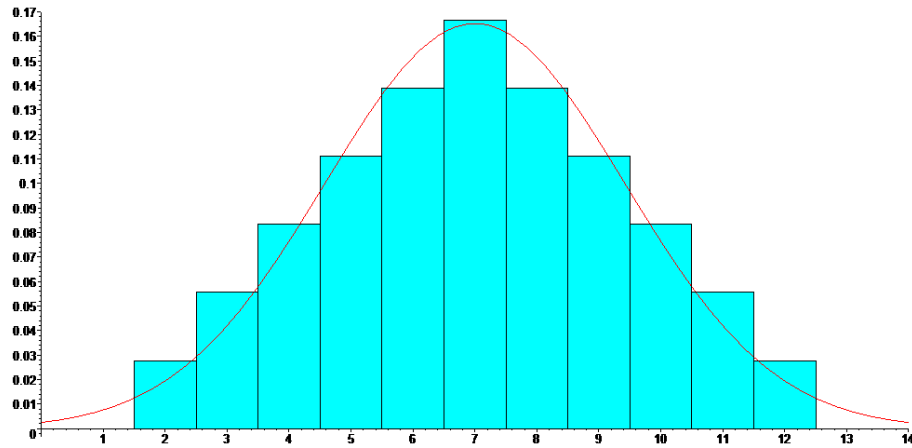
La moda è 7 e la mediana è ancora 7.

La varianza è

$$\begin{aligned} \sigma^2 &= \\ &= \frac{1}{36}(2-7)^2 + \frac{1}{36}(3-7)^2 + \frac{1}{36}(4-7)^2 + \frac{1}{36}(5-7)^2 + \frac{1}{36}(6-7)^2 + \frac{1}{36}(7-7)^2 + \\ &+ \frac{1}{36}(8-7)^2 + \frac{1}{36}(9-7)^2 + \frac{1}{36}(10-7)^2 + \frac{1}{36}(11-7)^2 + \frac{1}{36}(12-7)^2 = \\ &= \frac{1}{36}(25 + 2*16 + 3*9 + 4*4 + 5+5+4*4 + 3*9 + 2*16 + 25) = \frac{210}{36} = \frac{35}{6} \end{aligned}$$

Alternativamente

$$\begin{aligned} \sigma^2 &= \sum_{i=1}^6 \xi_i^2 p_i - \mu^2 = \\ &= \frac{1}{36}(4 + 2*9 + 3*16 + 4*25 + 5*36 + 6*49 + 5*64 + 4*81 + 3 + 100 + 2*121 + 144) - 49 = \\ &= \frac{1974}{36} - 49 = \frac{35}{6} \end{aligned}$$



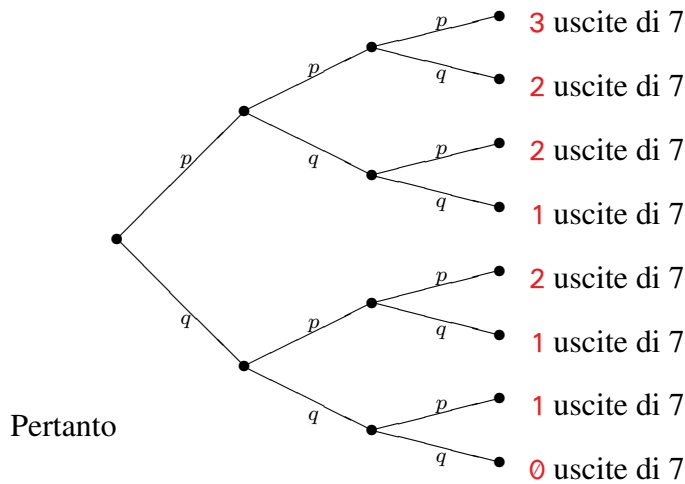
(in rosso è riportato il confronto con una gaussiana di media 7 e varianza $\frac{35}{6}$)

3. Lancio ripetuto di due dadi

Supponiamo di lanciare 2 dadi per 3 volte ed indichiamo con ξ la variabile aleatoria che restituisce il numero di volte in cui si è ottenuto il punteggio di 7 (nei tre lanci).

Come abbiamo visto prima

$$P(\text{uscita di } 7) = \frac{1}{6} = p, \quad P(\text{non uscita di } 7) = 1 - \frac{1}{6} = \frac{5}{6} = q$$



Pertanto

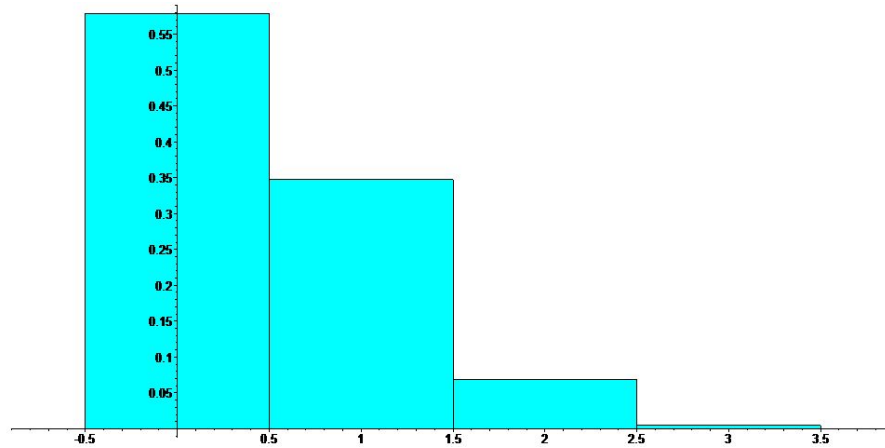
$$P(\xi = 0) = P(\text{nessun } 7) = \frac{5}{6} \frac{5}{6} \frac{5}{6} = \frac{125}{216} \approx 0.579$$

$$P(\xi = 1) = P(\text{un } 7) = \frac{1}{6} \frac{5}{6} \frac{5}{6} + \frac{5}{6} \frac{1}{6} \frac{5}{6} + \frac{5}{6} \frac{5}{6} \frac{1}{6} = 3 \frac{25}{216} \approx 0.347$$

$$P(\xi = 2) = P(\text{due } 7) = \frac{1}{6} \frac{1}{6} \frac{5}{6} + \frac{5}{6} \frac{1}{6} \frac{1}{6} + \frac{1}{6} \frac{5}{6} \frac{1}{6} = 3 \frac{5}{216} \approx 0.069$$

$$P(\xi = 3) = P(\text{tre } 7) = \frac{1}{6} \frac{1}{6} \frac{1}{6} = \frac{1}{216} \approx 0.005$$

L'istogramma che segue è relativo alla variabile aleatoria ξ e rappresenta la sua PDF.



Il problema consiste nel valutare il numero di successi ottenuti in 3 lanci considerando

successo: l'evento Il punteggio ottenuto è 7

che ha probabilità $p = \frac{1}{6}$,

insuccesso: l'evento Il punteggio ottenuto non è 7

che ha probabilità $p = \frac{5}{6}$.

Può essere generalizzato come segue:

4. Distribuzione Binomiale di Bernoulli

Consideriamo n ripetizioni di un evento a due stati

Successo	Insuccesso
Vincita	Perdita
Funzionante	Guasto
Testa	Croce
1	0

La probabilità di avere k Successi su n Lanci è

$$p^k q^{n-k}$$

e k successi su n lanci si possono ottenere in tanti modi quante sono le possibilità di disporre k segni $\star$ in n posizioni $\bigcirc$

$\bigcirc \star \bigcirc \bigcirc \bigcirc \bigcirc \star \bigcirc \star \star \bigcirc \bigcirc \star \star \bigcirc \bigcirc \star \star \bigcirc \bigcirc \bigcirc \bigcirc \star$

Abbiamo esemplificato nella figura il caso di 9 successi su 22 lanci. Naturalmente non è importante in quale delle 9 posizioni occupate da $\star$ si colloca il primo successo, in quale si colloca il secondo, in quale il terzo e così via; quindi ci sono molti modi per disporre sulle 9 posizioni dei successi i 9 successi che indichiamo con

$\textcircled{1} \textcircled{2} \textcircled{3} \textcircled{4} \textcircled{5} \textcircled{6} \textcircled{7} \textcircled{8} \textcircled{9}$

Pertanto

$\bigcirc \textcircled{1} \bigcirc \bigcirc \bigcirc \textcircled{2} \bigcirc \textcircled{3} \textcircled{4} \bigcirc \bigcirc \textcircled{5} \textcircled{6} \bigcirc \bigcirc \textcircled{7} \textcircled{8} \bigcirc \bigcirc \bigcirc \bigcirc \textcircled{9}$

è una possibile disposizione in cui il primo successo è collocato sulla prima posizione scelta il secondo sulla seconda e così fino al nono.

Ulteriori possibili disposizioni sono

○ ③ ○ ○ ○ ② ○ ① ⑨ ○ ○ ⑤ ⑧ ○ ○ ⑦ ⑥ ○ ○ ○ ○ ④

e

○ ⑨ ○ ○ ○ ② ○ ⑧ ④ ○ ○ ⑤ ⑥ ○ ○ ① ③ ○ ○ ○ ○ ⑤

Quindi per stimare il numero di possibilità di disporre k segni * in n posizioni ○ occorre

- contare in quanti modi si possono scegliere k posizioni su n disponibili; sia N_k il loro numero.
- contare in quanti modi si possono disporre i k segni * (successi) sulle k posizioni scelte; sia k_k il loro numero.
- Calcolare $\frac{N_k}{k_k}$ in quanto ogni possibilità di disporre k segni * in n posizioni è ottenuta k_k volte, se non si tiene conto dell'ordinamento.

Per contare in quanti modi si possono scegliere k posizioni su n disponibili possiamo osservare che:

- per il 1° dei k successi abbiamo a disposizione n posizioni
- per il 2° dei k successi abbiamo a disposizione $n - 1$ posizioni
- per il 3° dei k successi abbiamo a disposizione $n - 2$ posizioni
-
- per il k -esimo dei k successi abbiamo a disposizione $n - (k - 1)$ posizioni

Per cui avremo, in tutto,

$$N_k = n(n-1)(n-2)(n-3) \cdot \dots \cdot (n - (k-1)) = \frac{n!}{(n-k)!}$$

modi possibili

Per contare in quanti modi si possono disporre i k segni * (successi) sulle k posizioni scelte, possiamo osservare che

- per il 1° dei k successi abbiamo a disposizione k delle k posizioni scelte
- per il 2° dei k successi abbiamo a disposizione $k - 1$ delle k posizioni scelte
- per il 3° dei k successi abbiamo a disposizione $k - 2$ delle k posizioni scelte
-
- per il k -esimo dei k successi abbiamo a disposizione 1 delle k posizioni scelte

Per cui avremo, in tutto,

$$K_k = k(k-1)(k-2)(k-3) \cdot \dots \cdot 1 = k!$$

Concludendo, il numero di modi in cui si possono ottenere k successi su n lanci è

$$\frac{n(n-1)(n-2)(n-3) \cdots (n-(k+1))}{k!} = \frac{n!}{(n-k)!k!} = \binom{n}{k}$$

e la probabilità di ottenere k successi su n lanci è

$$\binom{n}{k} p^k q^{n-k}$$

Ricordiamo che i coefficienti binomiali si possono calcolare mediante il triangolo di Tartaglia. (Si veda appendice)

Se definiamo ξ la variabile aleatoria che restituisce il numero di successi ottenuti su n lanci, essendo p la probabilità di successo e $q = 1 - p$ la probabilità di insuccesso

ξ = numero di successi ottenuti in n lanci

p = probabilità di successo

q = probabilità di insuccesso

la funzione densità di probabilità (Probability Density Function, PDF) di ξ è definita da

$$P(\xi = k) = \binom{n}{k} p^k q^{n-k}$$

e la funzione di distribuzione cumulativa (Cumulative Distribution Function, CDF) è definita da

$$F(\nu) = \sum_{k=0}^{\nu} P(\xi = k) = \sum_{k=0}^{\nu} \binom{n}{k} p^k q^{n-k}$$

Ad esempio si ha

$$F(0) = \sum_{k=0}^0 P(\xi = k) = P(\xi = 0) = q^n$$

ed anche

$$F(n) = \sum_{k=0}^n P(\xi = k) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1$$

Con riferimento al caso precedentemente trattato dei 3 lanci di due dadi, definendo

Successo = esce 7, $p = \frac{1}{6}$

Insuccesso = non esce 7, $q = 1 - \frac{1}{6} = \frac{5}{6}$

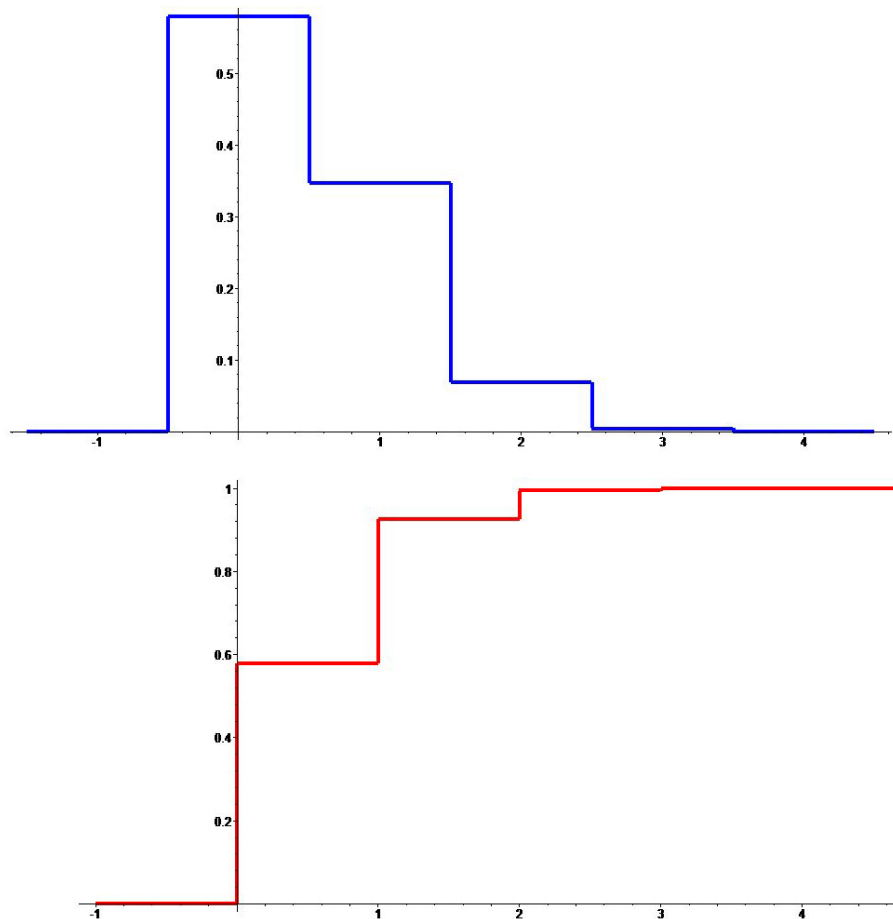
$$P(\xi = 0) = \binom{3}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^3 = 1 \frac{125}{216} = 0.579$$

$$P(\xi = 1) = \binom{3}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^2 = 3 \frac{25}{216} = 0.347$$

$$P(\xi = 2) = \binom{3}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^1 = 3 \frac{5}{216} = 0.069$$

$$P(\xi = 3) = \binom{3}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^0 = 1 \frac{1}{216} = 0.005$$

Di seguito sono riportate la PDF della variabile aleatoria ξ e la sua CDF



5. Applicazioni della distribuzione Binomiale

5.1.

Il 20% dei pezzi prodotti da una macchina in un giorno è difettoso.

Scegliamo a caso 4 pezzi della produzione giornaliera e consideriamo la variabile aleatoria ξ che restituisce il numero di pezzi difettosi tra i 4 scelti.

ξ = numero dei pezzi non difettosi su 4 pezzi scelti nella produzione giornaliera

Evidentemente possiamo ritenere

- **Successo** : che il pezzo non sia difettoso e si ha che la probabilità di successo è

$$p = \frac{80}{100} = \frac{4}{5} = 0.8$$

- **Insuccesso** : che il pezzo sia difettoso e si ha che la probabilità di insuccesso è

$$p = \frac{20}{100} = \frac{1}{5} = 0.2$$

ξ ha una distribuzione binomiale (bernoulliana) con $n = 4$ e $k = 5$.
Pertanto

$$P(\xi = 0) = \binom{4}{0} \left(\frac{2}{10}\right)^0 \left(\frac{8}{10}\right)^4 = 1 \frac{4096}{10^4} = 0.4096$$

$$P(\xi = 1) = \binom{4}{1} \left(\frac{2}{10}\right)^1 \left(\frac{8}{10}\right)^3 = 4 \frac{1024}{10^4} = 0.4096$$

$$P(\xi = 2) = \binom{4}{2} \left(\frac{2}{10}\right)^2 \left(\frac{8}{10}\right)^2 = 6 \frac{256}{10^4} = 0.1536$$

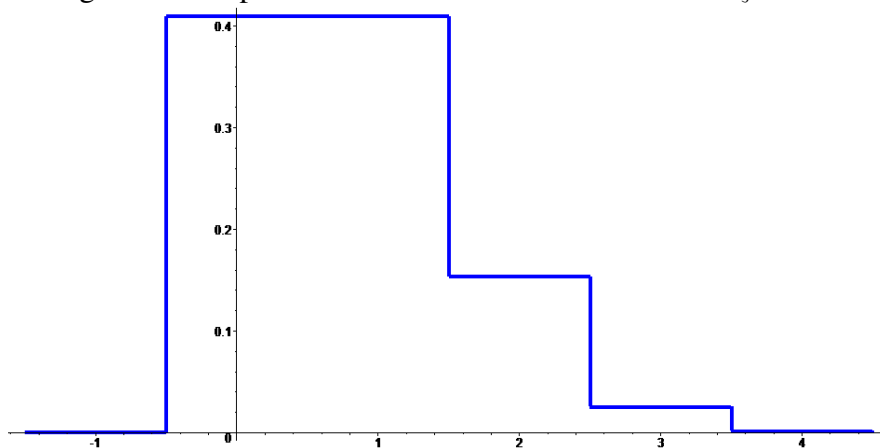
$$P(\xi = 3) = \binom{4}{3} \left(\frac{2}{10}\right)^3 \left(\frac{8}{10}\right)^1 = 4 \frac{64}{10^4} = 0.0256$$

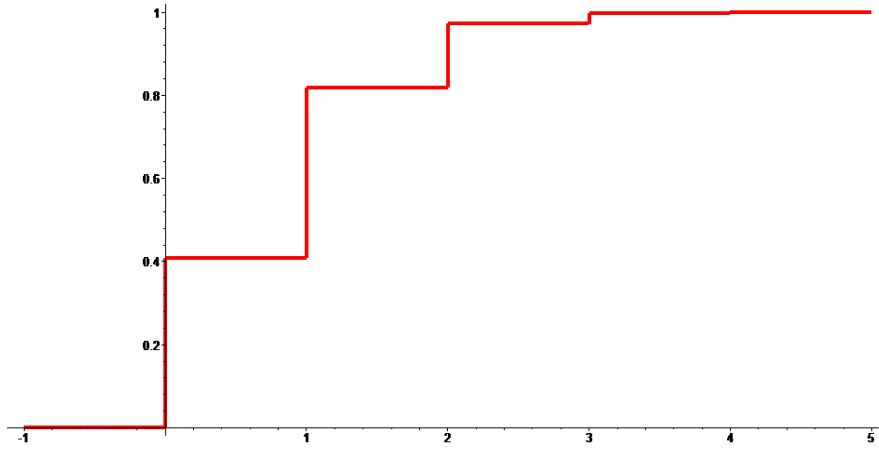
$$P(\xi = 4) = \binom{4}{4} \left(\frac{2}{10}\right)^4 \left(\frac{8}{10}\right)^0 = 1 \frac{16}{10^4} = 0.0016$$

Osserviamo che ovviamente si ha

$$0.4096 + 0.4096 + 0.1536 + 0.0256 + 0.0016 = 1$$

Di seguito sono riportate la PDF della variabile aleatoria ξ e la sua CDF





5.2.

Supponiamo di lanciare una moneta non truccata per $n = 10$ volte; definiamo successo l'uscita di Testa, che indichiamo con T ed insuccesso l'uscita di Croce che indichiamo con C .

Evidentemente

- $P(T) = \frac{1}{2} = p$ è la probabilità di successo.
- $P(C) = \frac{1}{2} = q = 1 - p$ è la probabilità di insuccesso.

Sia ξ il numero k di successi ottenuti in $n = 10$ lanci.

Evidentemente

$$P(\xi = 3) = \binom{10}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^7 = \frac{15}{128} \approx 0.1172$$

$$P(\xi = 4) = \binom{10}{4} \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^6 = \frac{105}{512} \approx 0.2051$$

$$P(\xi = 5) = \binom{10}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 = \frac{63}{256} \approx 0.2461$$

$$P(\xi = 6) = \binom{10}{6} \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^4 = \frac{105}{512} \approx 0.2051$$

e possiamo calcolare che si ottengano un numero di successi compreso tra 3 e 6 con probabilità

$$P(3 \leq \xi \leq 6) = \frac{60 + 105 + 126 + 105}{512} = \frac{99}{128} \approx 0.7734$$

Se F è la CDF della distribuzione di probabilità binomiale relativa ad $n = 10$ (possiamo ricavarne i valori dalla tabella in appendice), avremo che

$$P(3 \leq \xi \leq 6) = F(6) - F(2) \approx 0.82812 - 0.05469 \approx 0.77343$$

6. Probabilità condizionata. Teorema di Bayes

Ricordiamo che la probabilità di un evento A , condizionata all'accadimento dell'evento B si definisce come

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

Ne viene che

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

Nel caso in cui

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

diciamo che gli eventi A e B sono indipendenti; nel caso di eventi indipendenti si ha

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

6.1.

Consideriamo un'urna contenente 4 palline: una di colore Rosso, una di colore Bianco, una di colore Verde ed una di colore Nero.

Indichiamo con R, B, V, N , rispettivamente, l'evento

è estratta una pallina Rossa, Bianca, Verde, Nera

Consideriamo gli eventi

$$E_1 = B \cup R$$

$$E_2 = B \cup V$$

$$E_3 = B \cup N$$

e verifichiamo che gli eventi E_i sono a due a due indipendenti, ma non collettivamente indipendenti.

Si ha

$$P(E_1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(E_2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(E_3) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

ed inoltre

$$P(E_1)P(E_2) = \frac{1}{4} \quad P(E_1 \cap E_2) = P(B) = \frac{1}{4}$$

$$P(E_2)P(E_3) = \frac{1}{4} \quad P(E_2 \cap E_3) = P(B) = \frac{1}{4}$$

$$P(E_1)P(E_3) = \frac{1}{4} \quad P(E_1 \cap E_3) = P(B) = \frac{1}{4}$$

Possiamo anche verificare che E_1, E_2, E_3 non sono collettivamente indipendenti.

Infatti

$$P(E_1 \cap E_2 \cap E_3) = P(B) = \frac{1}{4}$$

mentre

$$P(E_1)P(E_2)P(E_3) = \frac{1}{8}$$

Se nell'urna viene aggiunta una quinta pallina Gialla, accade che

$$P(E_1) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

$$P(E_2) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

$$P(E_3) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

mentre

$$P(E_1)P(E_2) = \frac{4}{25} \quad P(E_1 \cap E_2) = P(B) = \frac{1}{5}$$

$$P(E_2)P(E_3) = \frac{4}{25} \quad P(E_2 \cap E_3) = P(B) = \frac{1}{5}$$

$$P(E_1)P(E_3) = \frac{4}{25} \quad P(E_1 \cap E_3) = P(B) = \frac{1}{5}$$

e gli eventi E_1 , E_2 , ed E_3 non sono più mutuamente indipendenti.

6.2. Il Teorema di Bayes. Siano $A_1, A_2 \supset B$, $A_1 \cap A_2 = \emptyset$ allora si ha

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} \quad , \quad P(B|A_i) = \frac{P(B \cap A_i)}{P(A_i)}$$

e quindi

$$P(A_i \cap B) = P(B)P(A_i|B) \quad , \quad P(B \cap A_i) = P(A_i)P(B|A_i)$$

ne segue che

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

e, dal momento che $B = (B \cap A_1) \cup (B \cap A_2)$

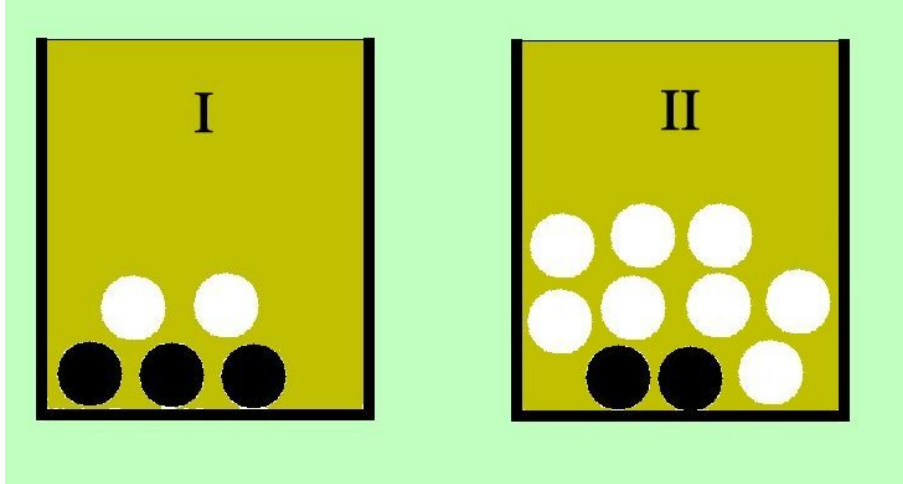
$$P(B) = P(B \cap A_1) + P(B \cap A_2) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$$

si può concludere che

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

6.3.

Si considerino due urne che indichiamo con I ed II e supponiamo che nell'urna I siano contenute 2 palline bianche (B) e 3 palline nere (N), mentre nell'urna II siano contenute 8 palline bianche 2 palline nere.



Si sceglie una delle due urne lanciando una moneta non truccata, per cui la probabilità di scegliere la prima urna ($P(I)$) e la probabilità di scegliere la seconda urna ($P(II)$) sono uguali a .5

$$P(I) = P(II) = \frac{1}{2}$$

Viene lanciata la moneta e viene estratta una pallina; conoscendo il colore della pallina estratta, calcolare la probabilità che sia stata estratta dalla prima o dalla seconda urna.

si ha

$$\begin{aligned} P(B|I) &= \frac{2}{5} & P(N|I) &= \frac{3}{5} \\ P(B|II) &= \frac{8}{10} & P(N|II) &= \frac{2}{10} \end{aligned}$$

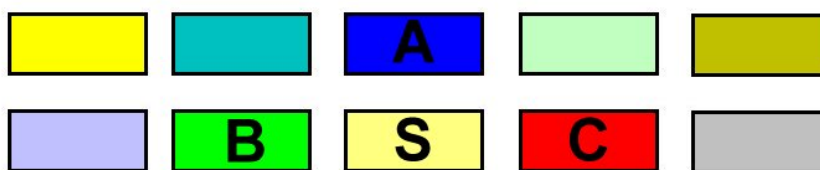
quindi

$$\begin{aligned} P(I|B) &= \frac{P(I)P(B|I)}{P(I)P(B|I) + P(II)P(B|II)} = \frac{2/5 \cdot 1/2}{2/5 \cdot 1/2 + 8/10 \cdot 1/2} = \frac{1}{3} \\ P(I|N) &= \frac{P(I)P(N|I)}{P(I)P(N|I) + P(II)P(N|II)} = \frac{3/5 \cdot 1/2}{3/5 \cdot 1/2 + 2/10 \cdot 1/2} = \frac{3}{4} \\ P(II|B) &= \frac{P(II)P(B|II)}{P(I)P(B|I) + P(II)P(B|II)} = \frac{8/10 \cdot 1/2}{2/5 \cdot 1/2 + 8/10 \cdot 1/2} = \frac{2}{3} \\ P(II|N) &= \frac{P(II)P(N|II)}{P(I)P(N|I) + P(II)P(N|II)} = \frac{2/10 \cdot 1/2}{3/5 \cdot 1/2 + 2/10 \cdot 1/2} = \frac{1}{4} \end{aligned}$$

6.4.

Durante una esercitazione scritta uno studente (S) si trova seduto nella prima fila dell'aula avendo a fianco ed immediatamente dietro tre altri studenti (A, B, C) con i quali riesce a consultarsi e dei quali valuta l'affidabilità secondo le seguenti stime.

A	B	C
60%	30%	10%



Una delle domande della prova è la seguente.

$$\int_{-\infty}^{+\infty} e^{-x^2/2} dx = \begin{cases} \textcircled{1} - & \sqrt{2\pi} \\ \textcircled{2} - & 2\pi \\ \textcircled{3} - & \sqrt{\pi} \end{cases}$$

(Ricordiamo che la risposta corretta è la **1** in quanto

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-x^2/2-y^2/2} dx dy &= \left(\int_{-\infty}^{+\infty} e^{-x^2/2} dx \right)^2 = \\ &= \int_0^{+\infty} \int_0^{2\pi} e^{-\rho^2/2} \rho d\rho d\theta = 2\pi e^{-\rho^2/2} \Big|_0^{+\infty} = 2\pi \end{aligned}$$

per cui $\int_{-\infty}^{+\infty} e^{-x^2/2} dx = \sqrt{2\pi}$.)

Ciascuna tra A,B,C fornisce la sua opinione in merito alla risposta corretta, come è riportato nella seguente tabella, dove su ogni riga è indicata con quale probabilità A, B o C ritengono corretta ciascuna delle risposte.

	①	②	③
A	60%	30%	10%
B	50%	30%	20%
C	30%	30%	40%

Vediamo ora come, tenendo conto del fatto che la risposta corretta è la **①** possiamo aggiornare la tabella della affidabilità di A,B,C.

Avremo che:

$$P(\textcircled{1} | A) = 60/100$$

$$P(\textcircled{1} | B) = 50/100$$

$$P(\textcircled{1} | C) = 30/100$$

da cui, indicando genericamente con X di volta in volta A,B,C , si ha:

$$P(X | \textcircled{1}) = \frac{P(\textcircled{1} | X)P(X)}{P(\textcircled{1})}$$

e

$$\begin{aligned} P(\textcircled{1}) &= P(\textcircled{1} | A)P(A) + P(\textcircled{1} | B)P(B) + P(\textcircled{1} | C)P(C) = \\ &= \frac{60}{100} \frac{60}{100} + \frac{50}{100} \frac{30}{100} + \frac{30}{100} \frac{10}{100} = \frac{54}{100} \end{aligned}$$

Infine

$$P(A | \textcircled{1}) = \frac{36/100}{54/100} = \frac{36}{54} \approx 67\%$$

$$P(B | \textcircled{1}) = \frac{15/100}{54/100} = \frac{15}{54} \approx 28\%$$

$$P(C | \textcircled{1}) = \frac{3/100}{54/100} = \frac{3}{54} \approx 5\%$$

6.5.

Negli anni '90 negli Stati Uniti fu proposto uno screening di massa per rilevare il numero di individui affetti da AIDS.

Una possibilità per rivelare la malattia consiste nel misurare la presenza di anticorpi che vengono prodotti dall'organismo in reazione all'AIDS. Il test, pur potendo essere molto accurato, può fornire risultati scorretti

- non rilevando la malattia in un individuo malato
- rilevando la malattia in un individuo sano

Chiamiamo

- **Sensitività** (p) del test la sua capacità di rilevare la malattia in un individuo malato; è ad esempio ragionevole supporre che

$$p = \frac{95}{100}$$

cioè che il test riveli la malattia nel 95% dei malati.

Ne consegue che il primo degli errori citati ha una probabilità del 5%.

- **Specificità** (q) del test la sua capacità di non rivelare la malattia in un individuo sano; è ad esempio ragionevole supporre che

$$q = \frac{99}{100}$$

cioè che il test riveli la malattia in un individuo sano nel 1% degli individui sani.

Supponiamo infine che nella popolazione in esame sia malato 1 individuo su 1000; supponiamo cioè che l'**incidenza** (a) della malattia sia

$$a = \frac{1}{1000}$$

Con questi dati vogliamo stimare l'efficacia di uno screening di massa. Indichiamo con

A	l'evento: l'individuo è malato di AIDS
A'	l'evento: l'individuo non è malato di AIDS
Tp	l'evento: l'individuo è risultato positivo al test
Tn	l'evento: l'individuo è risultato negativo al test

Avremo:

$$\begin{aligned} P(A) &= a = \frac{1}{1000} & P(A') &= 1 - a = \frac{999}{1000} \\ P(Tp|A) &= p = \frac{95}{100} & P(Tp|A') &= q = \frac{1}{100} \\ P(Tn|A) &= 1 - p = \frac{5}{100} & P(Tn|A') &= 1 - q = \frac{99}{100} \end{aligned}$$

Inoltre

$$P(Tp) = P(Tp|A)P(A) + P(Tp|A')P(A') = pa + q(1 - a)$$

per cui

$$\begin{aligned} P(A|Tp) &= \frac{P(Tp|A)P(A)}{P(Tp)} = \frac{pa}{pa + q(1 - a)} = \\ &= \frac{95/100 \cdot 1/1000}{95/100 \cdot 1/1000 + 1/100 \cdot 999/1000} = \frac{95}{95 + 999} = \frac{95}{1094} \approx 0.0847 \end{aligned}$$

In altre parole la probabilità che un individuo sia malato nel caso in cui il test abbia rivelato l'AIDS è inferiore al 9%.

Per comprendere come questo risultato sia influenzato dai dati che abbiamo usato calcoliamo la stessa probabilità usando set di parametri diversi.

p	q	a	$P(A Tp)$
$\frac{95}{100}$	$\frac{1}{100}$	$\frac{1}{1000}$	≈ 0.0847
$\frac{100}{100}$	$\frac{1}{10}$	$\frac{1}{1000}$	≈ 0.091
$\frac{100}{100}$	$\frac{2}{100}$	$\frac{1}{1000}$	≈ 0.05
$\frac{95}{100}$	$\frac{1}{300}$	$\frac{1}{1000}$	≈ 0.21
$\frac{90}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	≈ 0.55
$\frac{95}{100}$	$\frac{1}{100}$	$\frac{1}{10}$	≈ 0.9

6.6.

In una classe è svolta una prova scritta che consiste in 10 domande a risposta multipla ciascuna con 4 alternative.

Uno studente è sicuramente in grado di rispondere correttamente alla domanda nel caso conosca la materia, ma ha 1 probabilità su 4 di rispondere correttamente scegliendo una risposta a caso.

Se è noto, o è ragionevole supporre, che $2/3$ degli studenti conoscano la materia, con quale probabilità uno studente che ha risposto ad una certa domanda conosce la relativa materia?

Indichiamo con C l'evento lo studente conosce la materia e con C' in suo complementare (lo studente non conosce la materia);

indichiamo inoltre con B l'evento lo studente ha risposto correttamente e con B' il suo complementare (lo studente non ha risposto correttamente).

Avremo

$$P(C) = \frac{2}{3} \quad P(C') = \frac{1}{3}$$

$$P(B|C) = 1 \quad P(B|C') = \frac{1}{4}$$

Risulta

$$P(C|B) = \frac{P(B|C)P(C)}{P(B|C)P(C) + P(B|C')P(C')} =$$

$$= \frac{1 \cdot 2/3}{1 \cdot 2/3 + 1/4 \cdot 1/3} = \frac{8}{9}$$

Pertanto la probabilità che uno studente conosca la materia, avendo risposto correttamente alla domanda è $\frac{8}{9} \approx 0.89$

D'altro canto

$$P(C'|B) = \frac{P(B|C')P(C')}{P(B|C)P(C) + P(B|C')P(C')} = \frac{1/4 \cdot 1/3}{1 \cdot 2/3 + 1/4 \cdot 1/3} = \frac{1}{9}$$

Cioè la probabilità che uno studente non conosca la materia, avendo risposto correttamente alla domanda è $\frac{1}{9} \approx 0.11$

Evidentemente quindi la probabilità che uno studente che abbia risposto a k domande su 10 conosca veramente la materia è data da

$$P_k = \binom{n}{k} \frac{8^k}{9} \frac{1}{9^{n-k}}$$

Mentre la probabilità che, se ha risposto ad almeno 6 domande, abbia studiato è

$$P_6 + P_7 + P_8 + P_9 + P_{10} \approx 0.739$$

Se uno studente decide di scegliere a caso le risposte la probabilità di rispondere a k domande è

$$C_k = \binom{n}{k} \frac{1^k}{4} \frac{3^{n-k}}{4}$$

e quindi la probabilità che risponda correttamente ad almeno 6 domande su 10 è

$$C_6 + C_7 + C_8 + C_9 + C_{10} \approx 0.024$$

7. Distribuzioni di Probabilità

7.1.

Sia ξ una variabile aleatoria la cui distribuzione di probabilità è data da

$$\phi(t) = \begin{cases} \frac{1}{2}, & 0 < x < 1 \\ \frac{1}{4}, & 2 < x < 4 \\ 0 & \text{altrove} \end{cases}$$

(Ad esempio ξ può rappresentare il tempo di attesa per un autobus)

Calcolare la media μ e la varianza di ξ

$$\begin{aligned} \mu = E(\xi) &= \int_{-\infty}^{+\infty} x\phi(x)dx = \\ &= \int_0^1 x \frac{1}{2} dx + \int_2^4 x \frac{1}{4} dx = \\ &= \frac{x^2}{4} \Big|_0^1 + \frac{x^2}{8} \Big|_2^4 = \\ &= \frac{2 + 16 - 4}{8} = \frac{7}{4} = 2 - \frac{1}{4} = 2 - 0.25 = 1.75 \end{aligned}$$

$$\begin{aligned}
\sigma^2 &= E((\xi - \mu)^2) = \int_{-\infty}^{+\infty} \left(x - \frac{7}{4}\right)^2 \phi(x) dx = \\
&= \int_0^1 \left(x - \frac{7}{4}\right)^2 \frac{1}{2} dx + \int_2^4 \left(x - \frac{7}{4}\right)^2 \frac{1}{4} dx = \\
&= \frac{\left(x - \frac{7}{4}\right)^3}{6} \Big|_0^1 + \frac{\left(x - \frac{7}{4}\right)^3}{12} \Big|_2^4 = \\
&= \frac{\left(1 - \frac{7}{4}\right)^3}{6} + \frac{\left(4 - \frac{7}{4}\right)^3}{12} - \frac{\left(2 - \frac{7}{4}\right)^3}{12} = \\
&= \frac{1}{6} \left(\frac{4-7}{4}\right)^3 + \frac{1}{12} \left(\frac{16-7}{4}\right)^3 - \frac{1}{12} \left(\frac{8-7}{4}\right)^3 = \\
&= \frac{1}{4^3} \frac{1}{12} (-54 + 729 - 1) = \frac{674}{768}
\end{aligned}$$

7.2.

Stabilire per quali valori di $k \in \mathbb{R}$ la funzione

$$f(x) = \begin{cases} k & \text{se } |x| \leq 1 \\ \frac{k}{x^2} & \text{se } |x| > 1 \end{cases}$$

è una PDF per qualche variabile aleatoria ξ .

Affinchè f sia la PDF di una variabile aleatoria occorre che $f(x) \geq 0$ per ogni $x \in \mathbb{R}$ ed inoltre deve aversi che

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

Pertanto deve essere

$$\begin{aligned}
1 &= \int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{-1} \frac{k}{x^2} dx + \int_{-1}^1 k dx + \int_1^{+\infty} \frac{k}{x^2} dx = \\
&= -\frac{k}{x} \Big|_{-\infty}^{-1} + kx \Big|_{-1}^1 - \frac{k}{x} \Big|_1^{+\infty} = \\
&= k + 2k + k = 4k
\end{aligned}$$

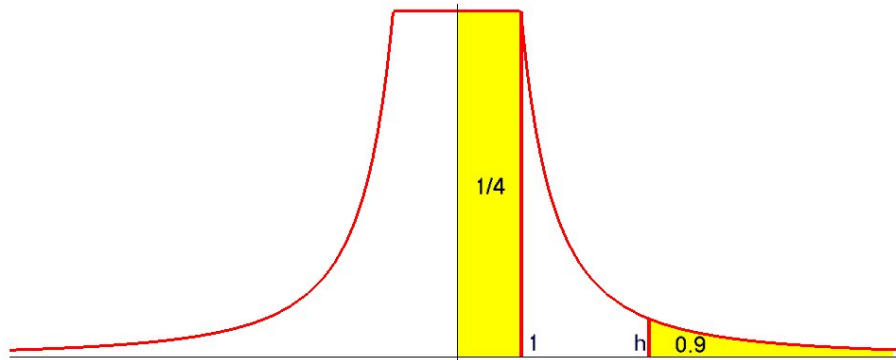
e

$$k = \frac{1}{4}$$

Si noti tuttavia che, poichè $xf(x)$ e, a maggior ragione, $(x - \mu)^2 f(x)$ non sono integrabili su $\mathbb{R}$ della variabile aleatoria individuata da ξ non è possibile calcolare nè media, nè la varianza.

Vogliamo ora, per $k = \frac{1}{4}$, determinare il valore di $h \in \mathbb{R}$ tale che

$$P(|\xi| \leq h) = 0.9$$



Osserviamo che, dal momento che

$$P(|\xi| \leq 1) = \int_{-1}^1 k dx = \frac{1}{2} < 0.9$$

dovrà risultare $h > 1$. ma

$$\int_1^h f(x) dx = -\frac{1}{4x} \Big|_1^h = \frac{k}{x} \Big|_1^{+\infty}$$

Per cui dovrà essere

$$1 - \frac{9}{10} - \frac{1}{4} = \frac{k}{x} \Big|_1^{+\infty}$$

e si ricava $h = \frac{20}{4}$

8. Uso delle tavole per i valori della CDF Normale Standard

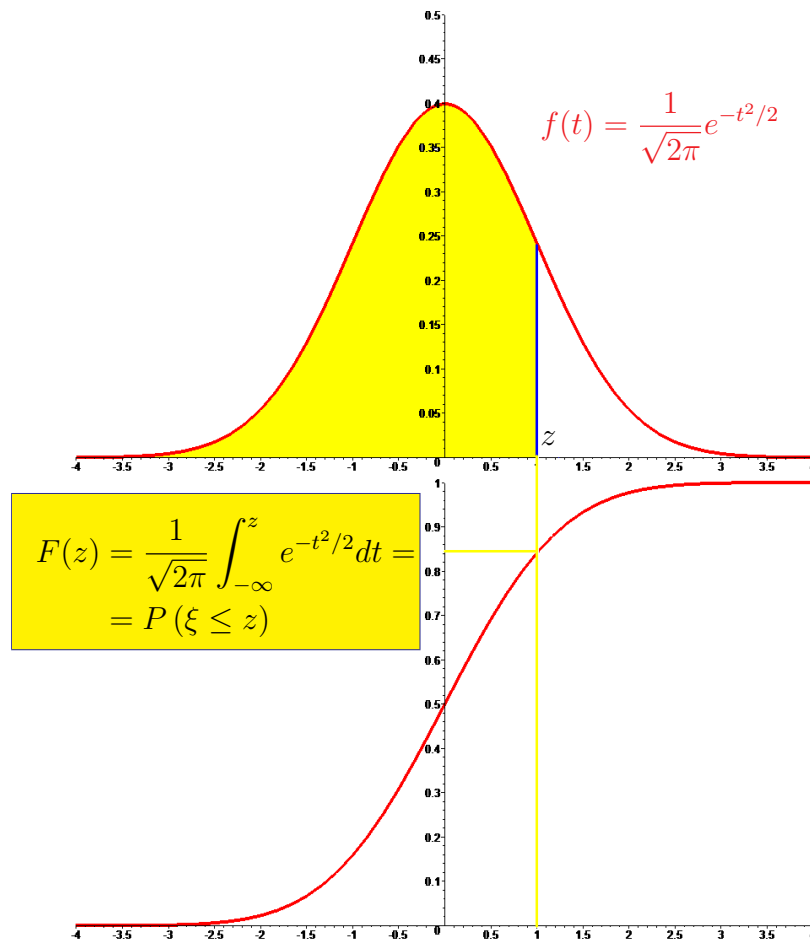
Sia ξ una variabile aleatoria Gaussiana di media μ e di varianza σ^2 ; allora se definiamo

$$z = \frac{\xi - \mu}{\sigma} \quad (\text{da cui } \xi = \mu + \sigma z)$$

z risulta una variabile aleatoria Gaussiana di media $\mu = 0$ e di varianza $\sigma^2 = 1$; una siffatta variabile aleatoria si dice normale standard.

La variabile aleatoria z ha una PDF ed una CDF che sono, rispettivamente date da

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \quad F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt$$



si ha

$$\begin{aligned}
F(z) = P(x \leq z) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt = \\
&= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^0 e^{-t^2/2} dt + \int_0^z e^{-t^2/2} dt \right) = \\
&= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z e^{-t^2/2} dt = \\
&= \frac{1}{2} + \frac{\sqrt{2}}{\sqrt{2\pi}} \int_0^{z/\sqrt{2}} e^{-s^2} ds = \\
&= \frac{1}{2} \left[1 + \frac{2}{\sqrt{\pi}} \int_0^{z/\sqrt{2}} e^{-s^2} ds \right] = \\
&= \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{z}{\sqrt{2}} \right) \right]
\end{aligned}$$

Poichè

$$P(\alpha \leq \xi \leq \beta) = P\left(\frac{\alpha - \mu}{\sigma} \leq z \leq \frac{\beta - \mu}{\sigma}\right) = F\left(\frac{\beta - \mu}{\sigma}\right) - F\left(\frac{\alpha - \mu}{\sigma}\right)$$

è evidentemente necessario disporre dei valori della funzione CDF della variabile normale standardizzata.

Tali valori possono essere facilmente calcolati mediante programmi di calcolo.

Ad esempio, usando EXCEL si calcola

$$\begin{aligned}
F(z) &= \text{NORMDIST}(z, 0, 1, \text{TRUE}) \\
F(z) &= (1 + \text{ERF}(z / \text{sqrt}(2))) / 2;
\end{aligned}$$

mentre usando MAPLE

$$F(z) = \text{statevalf}[\text{cdf}, \text{normald}](z);$$

oppure

$$F(z) = (1 + \text{evalf}(\text{erf}(z / \text{sqrt}(2)))) / 2;$$

ed usando MATLAB

$$F(z) = \text{normcdf}(z, 0, 1)$$

$$F(z) = (1 + \text{erf}(z / \text{sqrt}(2))) / 2;$$

I valori della funzione CDF della variabile normale standardizzata possono essere anche trovati facendo uso di tabelle simili a quella riportata in appendice; in essa sono indicati i valori di

$$\phi = \Phi\left(x + \frac{y}{100}\right)$$

dove

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-t^2/2} dt$$

e x, y, ϕ sono si trovano come segue

	*	*	*	*	*	y	*	*	*	*
+	-	-	-	-	-	-	-	-	-	-
+	-	-	-	-	-	-	-	-	-	-
+	-	-	-	-	-	-	-	-	-	-
x	-	-	-	-	-	ϕ	-	-	-	-
+	-	-	-	-	-	-	-	-	-	-
+	-	-	-	-	-	-	-	-	-	-
+	-	-	-	-	-	-	-	-	-	-

Per cui, ad esempio

$$\phi = \Phi(0.95) = \frac{1}{\sqrt{2\pi}} \int_0^{0.95} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_0^{0.9 + \frac{5}{100}} e^{-t^2/2} dt = 0.3289$$

	0	1	2	3	4	5	6	7	8	9
0.9	-	-	-	-	-	0.3289	-	-	-	-

E, similmente,

$$\phi = \Phi(1.58) = \frac{1}{\sqrt{2\pi}} \int_0^{1.58} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_0^{1.5 + \frac{8}{100}} e^{-t^2/2} dt = 0.4429$$

	0	1	2	3	4	5	6	7	8	9
1.5	-	-	-	-	-	-	-	-	0.4429	-

La tabella contiene valori di $\Phi(z)$ solo per valori di z compresi tra 0 e 3; per valori di $z > 3$ è ragionevole approssimare $\Phi(z)$ con 0.5, mentre per valori di z negativi si usa il fatto che la funzione densità di probabilità gaussiana f è una funzione pari e quindi Φ risulta dispari per cui

$$\Phi(z) = -\Phi(-z)$$

Va inoltre ricordato che talvolta nelle tabelle è riportato il valore di F per cui si leggono valori che differiscono da quelli di Φ per 0.5; precisamente si ha

$$F(z) = \Phi(z) + \frac{1}{2}$$

8.1.

Sia Z una variabile aleatoria normale standard ($\mu = 0$, $\sigma = 1$).

Calcoliamo

$$P(-1 \leq z \leq 1)$$

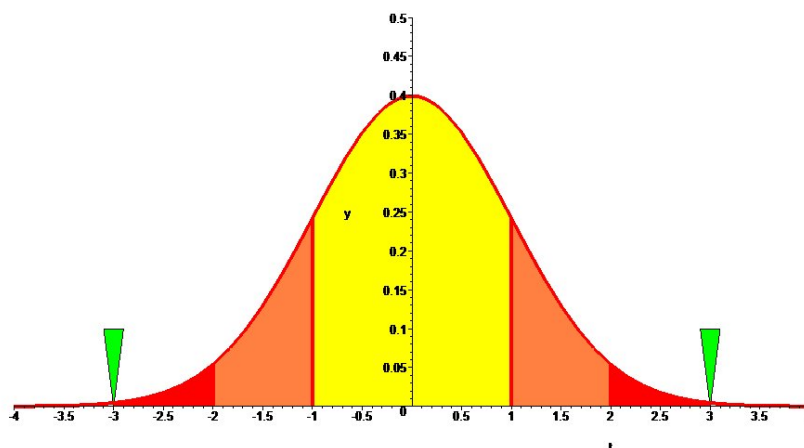
$$\begin{aligned} P(-1 \leq z \leq 1) &= \frac{1}{\sqrt{2\pi}} \int_{-1}^1 e^{-t^2/2} dt = \\ &= F(1) - F(-1) = \Phi(1) - \Phi(-1) = 2\Phi(1) \end{aligned}$$

Usando le tabelle si ha

$$\Phi(1) = -\Phi(-1) = 0.3413$$

e

$$P(-1 \leq z \leq 1) = 0.6826$$



Calcoliamo

$$P(-2 \leq z \leq 2)$$

$$\begin{aligned} P(-2 \leq z \leq 2) &= \frac{1}{\sqrt{2\pi}} \int_{-2}^2 e^{-t^2/2} dt = \\ &= F(2) - F(-2) = \Phi(2) - \Phi(-2) = 2\Phi(2) \end{aligned}$$

Usando le tabelle si ha

$$\Phi(2) = -\Phi(-2) = 0.4472$$

e

$$P(-2 \leq z \leq 2) = 0.8944$$

Calcoliamo

$$P(-3 \leq z \leq 3)$$

$$\begin{aligned}
 P(-3 \leq z \leq 3) &= \frac{1}{\sqrt{2\pi}} \int_{-3}^3 e^{-t^2/2} dt = \\
 &= F(3) - F(-3) = \Phi(3) - \Phi(-3) = 2\Phi(3)
 \end{aligned}$$

Usando le tabelle si ha

$$\Phi(3) = -\Phi(-3) = 0.4981$$

e

$$P(-3 \leq z \leq 3) = 0.9962$$

Se poi ξ è una variabile aleatoria gaussiana di media μ e varianza σ^2 allora

$$z = \frac{\xi - \mu}{\sigma}$$

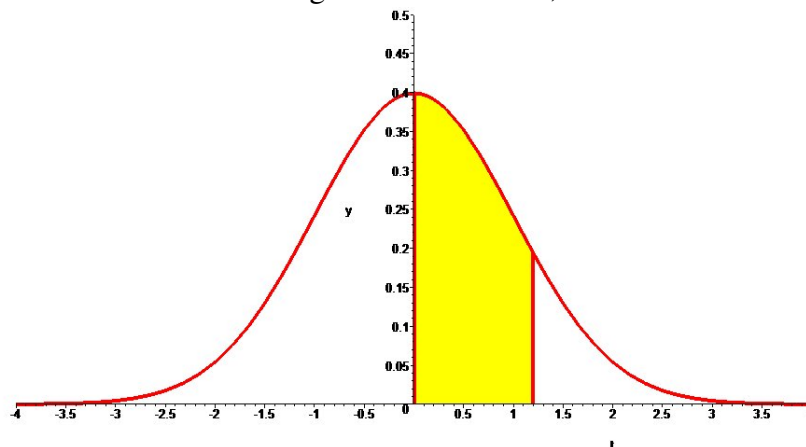
$$P(\mu - \sigma \leq \xi \leq \mu + \sigma) = P(-1 \leq z \leq 1) = 0.6826$$

$$P(\mu - 2\sigma \leq \xi \leq \mu + 2\sigma) = P(-2 \leq z \leq 2) = 0.8944$$

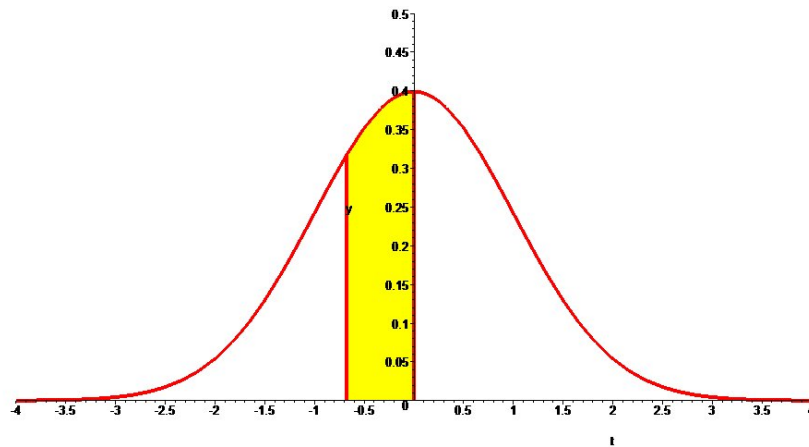
$$P(\mu - 3\sigma \leq \xi \leq \mu + 3\sigma) = P(-3 \leq z \leq 3) = 0.9962$$

8.2.

Sia z una variabile aleatoria gaussiana standard, calcolare



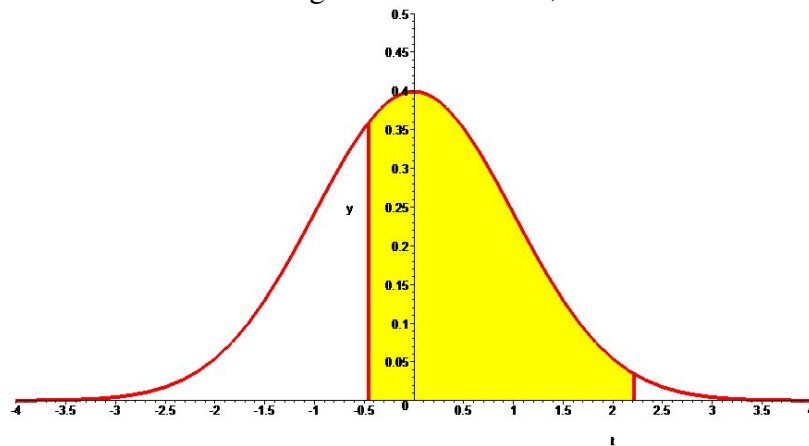
$$P(0 \leq z \leq 1.2) = \Phi(1.2) = .3849$$



$$P(-0.68 \leq z \leq 0) = -\Phi(-0.68) = \Phi(0.68) = .2517$$

8.3.

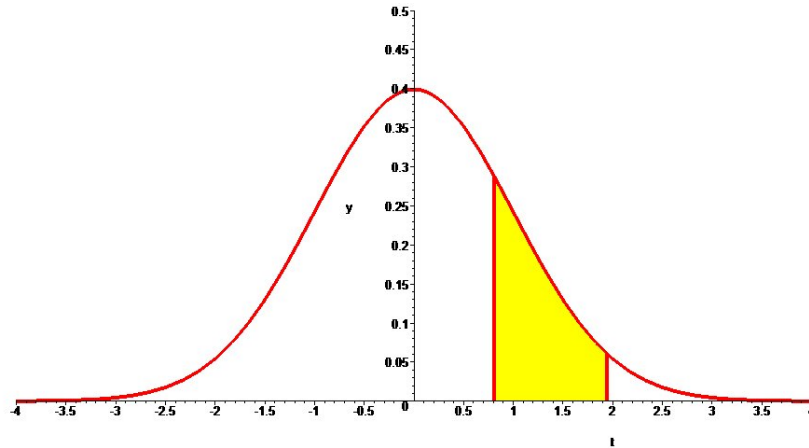
Sia z una variabile aleatoria gaussiana standard, calcolare



$$\begin{aligned} P(-0.46 \leq z \leq 2.21) &= F(2.21) - F(-0.46) = \\ &= \Phi(2.21) - \Phi(-0.46) = \Phi(2.21) + \Phi(0.46) = 0.4864 + 0.1772 = 0.6636 \end{aligned}$$

8.4.

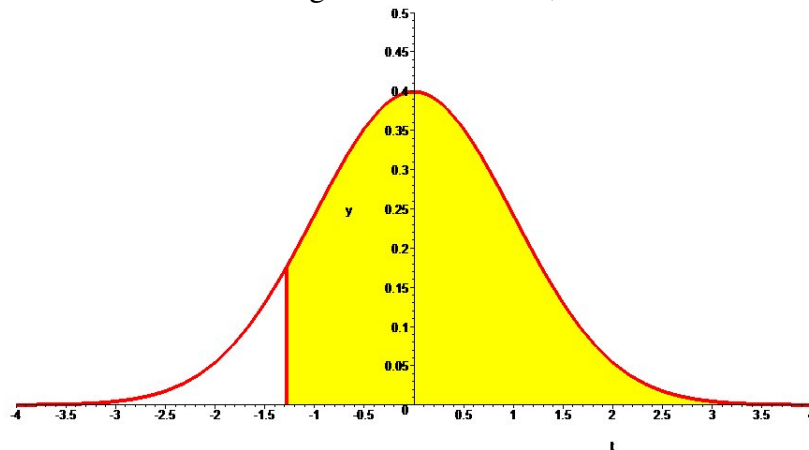
Sia z una variabile aleatoria gaussiana standard, calcolare



$$\begin{aligned}
 P(0.81 \leq z \leq 1.94) &= F(1.94) - F(0.81) = \\
 &= \Phi(1.94) - \Phi(0.81) = 0.4864 + 0.1772 = 0.6636
 \end{aligned}$$

8.5.

Sia z una variabile aleatoria gaussiana standard, calcolare



$$\begin{aligned}
 P(z \geq -1.28) &= 1 - F(-1.28) = 1 - 1/2 - \Phi(-1.28) = \\
 &= 1/2 + \Phi(1.28) = 0.5 + 0.3997 = 0.8997
 \end{aligned}$$

9. Calcolo dei valori della CDF Gaussiana Inversa

Sia ξ una variabile aleatoria Gaussiana di media μ e di varianza σ^2 ; e

$$z = \frac{\xi - \mu}{\sigma} \quad (\text{da cui } \xi = \mu + \sigma z)$$

la corrispondente variabile aleatoria normale standard.

Sia F la CDF di z , allora

$$F(z_a) = P(z \leq z_a) = A$$

se e solo se

$$z_a = F^{-1}(A)$$

Il valore z_a può essere ancora calcolato mediante programmi di calcolo.
Ad esempio, usando EXCEL si calcola

$$z_a = \text{NORMINV}(A, 0, 1)$$

mentre usando MAPLE

$$z_a = \text{statevalf}[\text{icdf}, \text{normald}](A);$$

ed usando MATLAB

$$z_a = \text{norminv}(A, 0, 1)$$

$$z_a = \sqrt{2} \operatorname{erfinv}(2A-1)$$

I valori della funzione CDF della variabile normale standardizzata possono essere anche trovati facendo uso di tabelle simili a quella riportata in appendice; in essa sono indicati i valori di

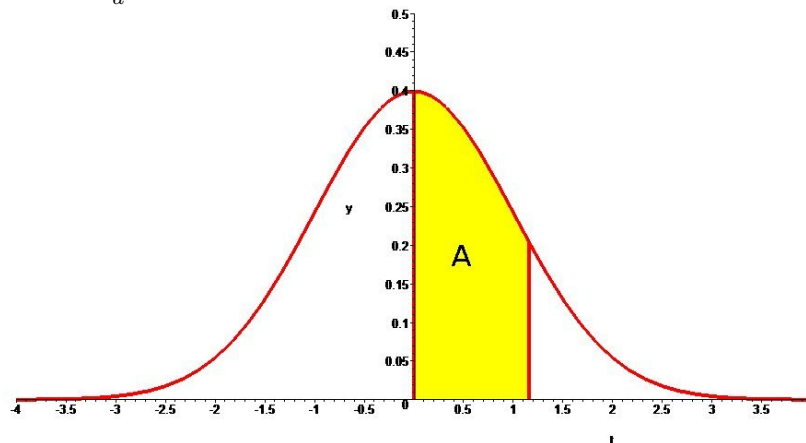
$$\phi = \Phi\left(x + \frac{y}{100}\right) = \frac{1}{\sqrt{2\pi}} \int_0^{x + \frac{y}{100}} e^{-t^2/2} dt$$

9.1.

Sia z una variabile aleatoria e sia z_a tale che

$$P(0 \leq z \leq z_a) = 0.3770$$

Determinare z_a



Dalle tavole della CDF Gaussiana Inversa si ricava che

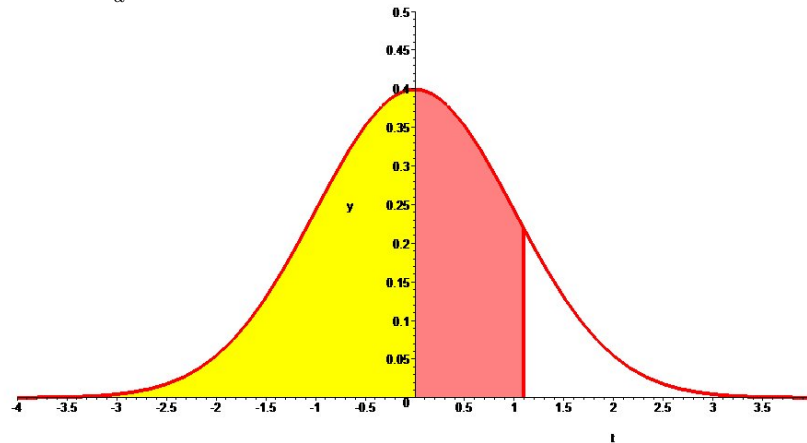
$$z_a = \Phi^{-1}(0.3770) = 1.16$$

9.2.

Sia z una variabile aleatoria e sia z_a tale che

$$P(z \leq z_a) = 0.8621$$

Determinare z_a



Dalle tavole della CDF Gaussiana Inversa si ricava che

$$z_a = F^{-1}(0.8621) = \Phi^{-1}(0.8621 - 0.5) = 1.09$$

9.3.

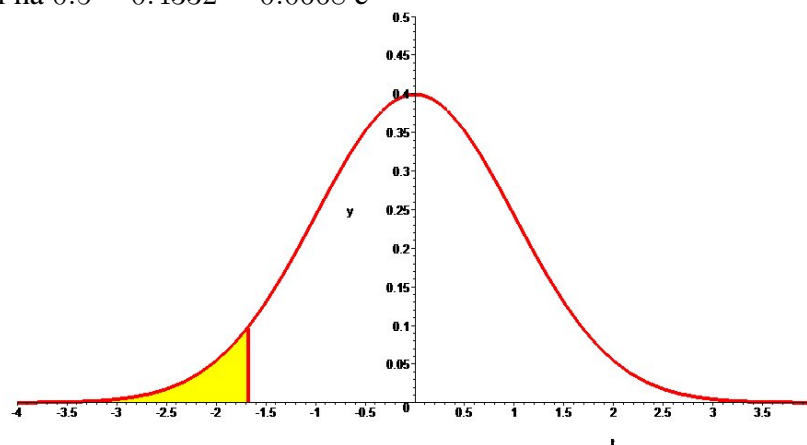
Sia z una variabile aleatoria e sia z_a tale che

$$P(z \leq z_a) = 0.4332$$

Determinare z_a

Poichè $0.4332 < 0.5$ z_a è negativo.

Si ha $0.5 - 0.4332 = 0.0668$ e



Dalle tavole della CDF Gaussiana Inversa si ricava che

$$z_a = F^{-1}(0.0668) = \Phi^{-1}(0.0668 - 0.5) = -\Phi^{-1}(0.4332) = -1.68$$

10. Approssimazione della Distribuzione Binomiale mediante la Distribuzione Normale

Qualora il numero di prove n sia grande e p e q siano troppo piccoli (in pratica si adotta la condizione $np, nq \geq 5$), la distribuzione binomiale può essere approssimata mediante la distribuzione normale standardizzata.

Se ξ_n è (la variabile aleatoria che restituisce) il numero di successi su n prove bernoulliane con probabilità di successo p , definiamo (una nuova variabile aleatoria) z_n mediante la

$$z_n = \frac{\xi_n - np}{\sqrt{npq}} \quad (\text{da cui } \xi_n = np + \sqrt{npq}z_n)$$

z_n è una variabile aleatoria con media $\mu = 0$ e varianza $\sigma^2 = 1$ che per $n \rightarrow +\infty$ può essere approssimata mediante una variabile gaussiana standard.

Più precisamente

$$\lim_n P(a \leq z_n \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

Pertanto

$$\begin{aligned} P(\alpha \leq \xi_n \leq \beta) &= P\left(\alpha - \frac{1}{2} \leq np + \sqrt{npq}z_n \leq \beta + \frac{1}{2}\right) = \\ &= P\left(\frac{\alpha - \frac{1}{2} - np}{\sqrt{npq}} \leq z_n \leq \frac{\beta + \frac{1}{2} - np}{\sqrt{npq}}\right) \end{aligned}$$

Alla luce di tutto ciò, la probabilità che su 10 lanci di una moneta non truccata si abbiano un numero di teste ξ compreso tra 3 e 6, si può calcolare tenendo conto che

$$\alpha = 3, \quad \beta = 6, \quad n = 10, \quad p = q = \frac{1}{2}$$

$$\begin{aligned} P(3 \leq \xi \leq 6) &= P\left(\frac{3 - \frac{1}{2} - 5}{\sqrt{2.5}} \leq z_n \leq \frac{6 + \frac{1}{2} - 5}{\sqrt{2.5}}\right) = \\ &= P(-1.58 \leq z_n \leq 0.95) \end{aligned}$$

Usando le tavole della CDF Gaussiana standardizzata riportate in appendice;

$$P(-1.58 \leq z_n \leq 0.95) = F(0.95) - F(1.58) = 0.4429 + 0.3288 = 0.7718$$

Possiamo confrontare il risultato con quello ottenuto direttamente mediante i valori della distribuzione cumulativa binomiale B_{10} relativa a 10 lanci; in questo modo si ottiene

$$B_{10}(6) - B_{10}(2) = 0.7734$$

ed osservare che l'errore di approssimazione commesso è dell'ordine di 0.0016

11. Stima di Probabilità

11.1.

Una macchina produce $n = 10000$ pezzi al giorno. La probabilità che un pezzo sia difettoso è del 20%

Vogliamo poter affermare che nella produzione giornaliera ci sono almeno n pezzi buoni.

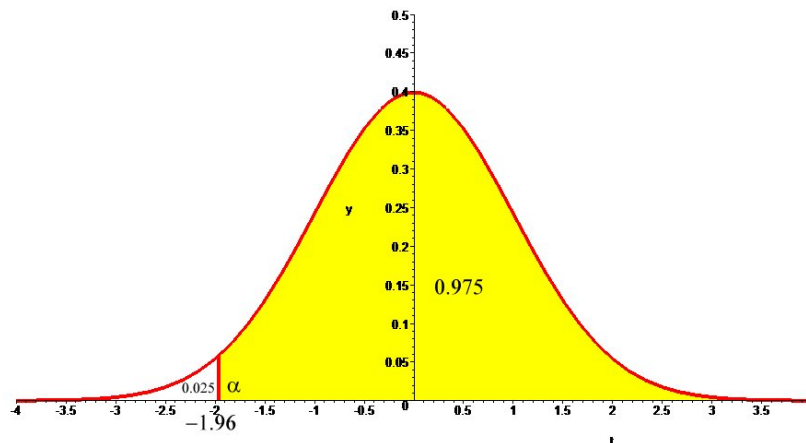
Determinare come dobbiamo scegliere n affinché la probabilità che l'affermazione sia corretta risulti superiore al 97.5%.

Sia ξ la variabile aleatoria che restituisce il numero di pezzi non difettosi prodotti in un giorno.

La probabilità che un pezzo sia difettoso è $q = P(D) = 0.2$ mentre la probabilità che sia buono è $p = P(B) = 0.8$; ξ ha una distribuzione binomiale di media $\mu = np = 10000 * 0.8 = 8000$ e di varianza $\sigma^2 = npq = 10000 * 0.16 = 1600$, ($\sigma = 40$).

Affinché l'affermazione sia corretta occorre che

$$P(\xi > x) > 0.975$$



Per calcolare il valore di x è conveniente osservare che deve essere

$$P(\xi > n) = P\left(\frac{\xi - \mu}{\sigma} > \frac{n - \mu}{\sigma}\right) > 0.975$$

$\frac{\xi - \mu}{\sigma}$ si può approssimare con una distribuzione normale standard z e se $\frac{x - \mu}{\sigma} = \alpha$ deve essere

$$1 - F(\alpha) = P(z > \alpha) > 0.975$$

da cui

$$F(\alpha) < 0.025$$

ed

$$\alpha < F^{-1}(0.025) = -1.96$$

Ne viene quindi che

$$\frac{n - \mu}{\sigma} = \alpha = -1.96$$

e

$$x \leq 8000 - 40 * 1.96 = 7921.6$$

Concludendo, l'affermazione:

La produzione giornaliera contiene almeno 7921 pezzi buoni
è corretta al 97.5%

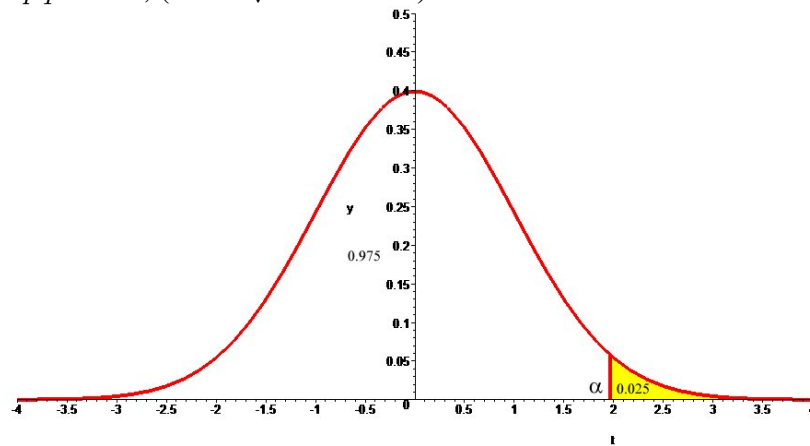
11.2.

Una moneta viene lanciata 1000 volte; La probabilità $p = P(T)$ di ottenere Testa è uguale alla probabilità $q = P(C)$ di ottenere Croce ed entrambe valgono 0.5

Vogliamo determinare m in modo che la probabilità di ottenere più di m volte Testa sia inferiore al 2.5%.

Sia ξ la variabile aleatoria che restituisce il numero di Teste (successi) ottenuti in $n = 1000$ lanci.

ξ ha una distribuzione binomiale di media $\mu = np = 500$ e di varianza $\sigma^2 = npq = 250$, ($\sigma = 5\sqrt{10} = 15.81$).



Per calcolare m osserviamo che

$$0.025 > P(\xi > m) = P\left(\frac{\xi - \mu}{\sigma} > \frac{m - \mu}{\sigma}\right) > 0$$

$\frac{\xi - \mu}{\sigma}$ si può approssimare con una distribuzione normale standard z e se

$$\alpha = \frac{m - \mu}{\sigma}$$

$$1 - F(z) = P(z > \alpha) < 0.025$$

per

$$\alpha > F^{-1}(0.025) = 1.96$$

Ne viene quindi che

$$\frac{m - \mu}{\sigma} > \alpha, m \leq \mu + \sigma\alpha$$

$$x \leq 500 + 15.81 * 1.96 = 500 + 30.98 = 530.98$$

11.3.

Una macchina produce sfere di acciaio di diametro medio uguale a 4cm.

Su un campione di 100 sfere se ne trovano

- 5 di diametro inferiore a 3.995cm.
- 12 di diametro superiore a 4.005cm.

Valutando la probabilità con le frequenze ottenute e supponendo il diametro delle sfere prodotte distribuito normalmente con media μ e varianza σ^2 , calcolare μ e σ .

Sia ξ la variabile aleatoria che restituisce il diametro delle sfere prodotte.

Avremo, valutando la probabilità con le frequenze, che

$$P(\xi > 4.005) = \frac{12}{100}$$

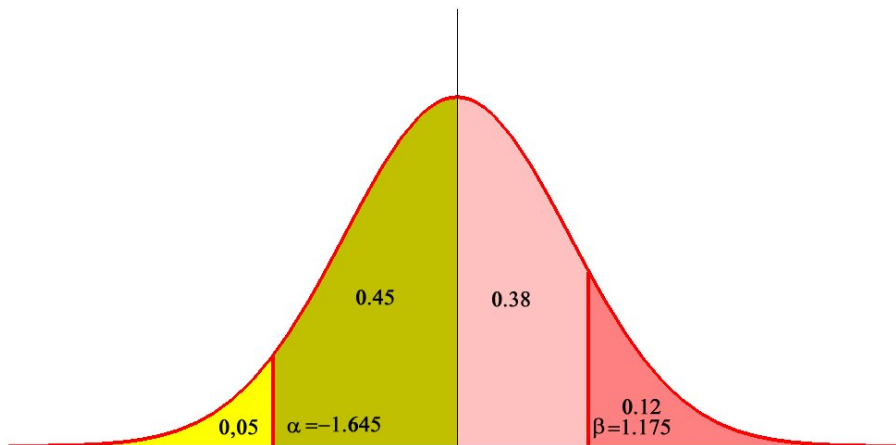
$$P(\xi < 3.995) = \frac{5}{100}$$

e

$$P\left(\frac{\xi - \mu}{\sigma} > \frac{4.005 - \mu}{\sigma}\right) = 0.12, \quad P\left(\frac{\xi - \mu}{\sigma} < \frac{3.995 - \mu}{\sigma}\right) = 0.05$$

Poichè la variabile aleatoria $z = \frac{\xi - \mu}{\sigma}$ è distribuita normalmente con $\mu = 0$ e $\sigma = 1$, posto

$$\alpha = \frac{3.995 - \mu}{\sigma}, \quad \beta = \frac{4.005 - \mu}{\sigma}$$



affinchè

$$F(z) = P(z < \alpha) = 0.05, \quad 1 - F(z) = P(z > \beta) = 0.12$$

deve risultare

$$\alpha = F^{-1}(0.05) = -1.645, \quad \beta = F^{-1}(0.88) = 1.175$$

Pertanto dovrà essere

$$\frac{3.995 - \mu}{\sigma} = \alpha = -1.645, \quad \text{e} \quad \frac{4.005 - \mu}{\sigma} = \beta = 1.175$$

Se ne deduce che

$$3.995 - \mu = -\sigma 1.645$$

$$4.005 - \mu = \sigma 1.175$$

e

$$0.010 = \sigma(1.175 + 1.645) = \sigma 2.820$$

Se ne conclude che

$$\sigma = \frac{1}{100} \frac{1}{2.82} = 0.0035$$

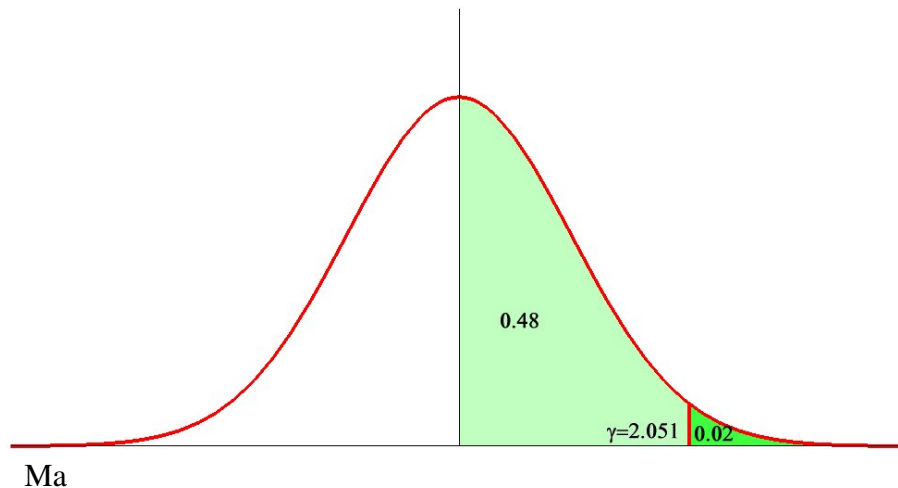
e

$$\mu = 3.995 + \frac{1}{100} \frac{1}{2.82} 1.645 = 3.995 + 0.0058 = 4.0009$$

$$(\mu = 3.995 - \frac{1}{100} \frac{1}{2.82} 1.175 = 4.005 - 0.0042 = 4.0009)$$

Assumiamo ora che $\mu = 4.0009$ e cerchiamo di determinare σ in modo che

$$P(\xi > 4.005) < \frac{2}{100}$$



$$z = \frac{\xi - \mu}{\sigma}$$

è una variabile aleatoria normalmente distribuita con media 0 e varianza 1 e quindi

$$1 - F\left(\frac{4.005 - \mu}{\sigma}\right) = P\left(z > \frac{4.005 - \mu}{\sigma}\right) < 0.02$$

Ne viene che

$$F\left(\frac{4.005 - \mu}{\sigma}\right) > 1 - 0.02 = 0.98$$

e

$$\frac{4.005 - \mu}{\sigma} > F^{-1}(0.98) = 2,053$$

$$4.005 - 4.0009 > \sigma 2.053$$

$$\sigma < \frac{4.005 - 4.0009}{2.053} = 0.0019$$

11.4.

Un parco è servito da due bagni. Si supponga che giornalmente 400 persone facciano uso dei bagni scegliendo a caso uno dei due con probabilità $\frac{1}{2}$

All'interno di ognuno dei bagni è collocato un dispenser che fornisce ogni utente di uno ed un solo asciugamani di carta.

Con quanti asciugamani devono essere riforniti i dispenser affinché la probabilità che restino vuoti sia inferiore al 2.5%?

Siano u_1 e u_2 gli utenti del primo e del secondo bagno, rispettivamente.

Sia N il numero di asciugamani caricati in ciascun dispenser ed indichiamo semplicemente con u il numero di utenti di uno dei bagni.

Ci interessa assicurare che

$$P(u > N) < 0.025$$

La distribuzione di probabilità di u è binomiale con

$$n = 400, p = q = \frac{1}{2}, \mu = np = 200, \sigma = \sqrt{npq} = \sqrt{100} = 10$$

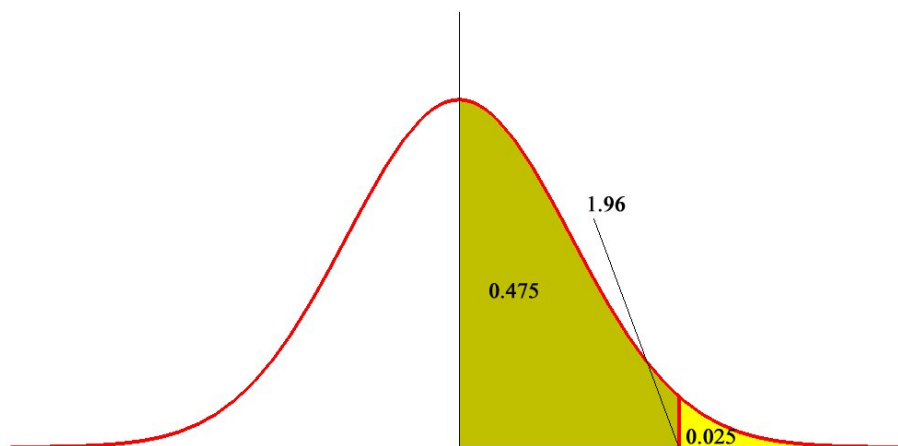
per cui possiamo normalizzare la variabile aleatoria u ed approssimarla mediante una distribuzione gaussiana standard.

Avremo che, posto

$$z = \frac{u - \mu}{\sigma}$$

$P(u > N) < 0.025$ è equivalente a

$$P\left(z > \frac{N - \mu}{\sigma}\right) < 0.025$$



quindi

$$1 - F\left(\frac{N - \mu}{\sigma}\right) = P\left(z > \frac{N - \mu}{\sigma}\right) < 0.025$$

e dovrà quindi essere

$$\frac{N - \mu}{\sigma} > F^{-1}(1 - 0.025) = 1.96$$

da cui

$$N > \mu + \sigma 1.96 = 200 + 10 \cdot 1.96 = 219.6$$

Dovranno essere quindi caricati in ogni dispenser almeno 220 asciugamani affinché la probabilità che non si vuotino sia inferiore a 0.025.

11.5.

Un addetto deve rifornire due macchine distributrici di lattine di una medesima bibita che sono collocate in punti differenti di uno stesso edificio.

Si stima che 100 clienti al giorno acquistino da una delle due macchine, scelta a caso con probabilità $\frac{1}{2}$, una lattina.

Quante lattine devono essere contenute in ciascuna macchina per essere certi al 97.5% che ogni cliente possa acquistare la propria lattina?

Sia u_k , con $k = 1, 2$, il numero degli utenti della prima e della seconda macchina. Affinchè ciascuno di essi possa acquistare la propria lattina occorre che, detto n_k , con $k = 1, 2$, il numero di lattine contenute nella prima e nella seconda macchina rispettivamente.

u_k è una variabile aleatoria binomiale di media μ e varianza σ^2 con

$$\mu = 50, \quad \sigma^2 = 25, \quad \sigma = 5$$

che può essere normalizzata ed approssimata da una variabile aleatoria gaussiana z normalizzata.

Dovrà risultare

$$\begin{aligned} P(u_k > n_k) &< 0.025 \\ P\left(\frac{u_k - \mu}{\sigma} > \frac{n_k - \mu}{\sigma}\right) &< 0.025 \end{aligned}$$

e posto $z_k = \frac{n_k - \mu}{\sigma}$

$$1 - F(z_k) = P(z > z_k) < 0.025$$

da cui

$$\begin{aligned} F(z_k) &> 0.975 \\ \frac{u_k - \mu}{\sigma} = z_k &> F^{-1}(0.975) = 1.96 \\ n_k &> 50 + 5 \cdot 1.96 = 59.8 \end{aligned}$$

11.6.

Consideriamo ora lo stesso problema, relativo però a due macchine che sono usate l'una con probabilità $\frac{1}{4}$ e $\frac{3}{4}$.

u_1 ed u_2 sono ancora due variabili aleatorie binomiali che hanno la stessa varianza

$$\sigma^2 = 100 \frac{3}{16}, \quad \sigma = 10 \frac{\sqrt{3}}{4} \approx 4.26$$

ed hanno media, rispettivamente,

$$\mu_1 = 100 \frac{1}{4} = 25, \quad \mu_2 = 100 \frac{3}{4} = 75$$

procedendo come sopra si ottiene quindi che deve essere

$$n_k > \mu_k + \sigma * 1.96 = 59.8$$

e quindi che

$$n_1 > 25 + 4.26 * 1.96 \approx 33.35$$

$$n_2 > 75 + 4.26 * 1.96 \approx 83.35$$

Osserviamo che, mentre nel primo caso per rifornire entrambe le macchine erano necessarie 120 lattine, in questo ne servono soltanto 118.

12. Test Statistici

12.1.

Determinare la probabilità di ottenere in $(n =)100$ lanci di una moneta non truccata ($p = q = \frac{1}{2}$) un numero di Teste (T) compreso tra 40 e 60.

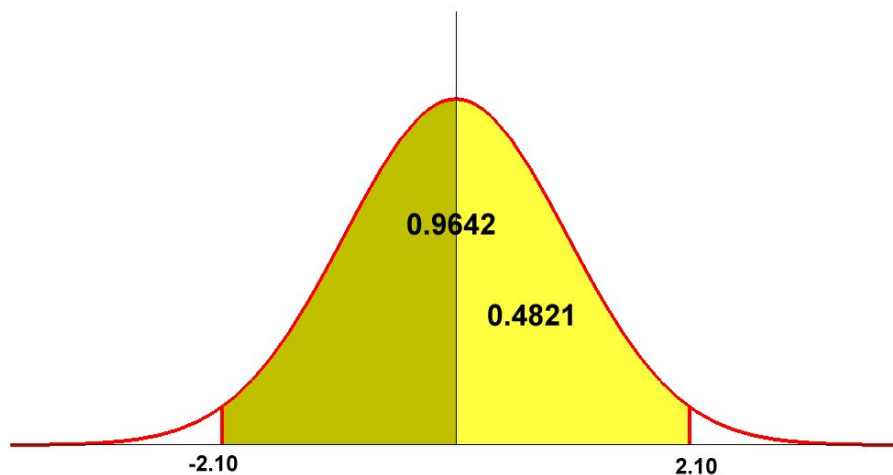
Si tratta di un processo bernoulliano con media μ e varianza σ^2 dati da

$$\mu = np = 100 \frac{1}{2}, \quad \sigma^2 = npq = 100 \frac{1}{2} \frac{1}{2} = 25$$

Poichè $np, nq > 5$ possiamo approssimare la distribuzione binomiale di Bernoulli utilizzando una distribuzione normale standard con la stessa media e varianza.

Avremo che

$$\begin{aligned} P(40 \leq T \leq 60) &= P\left(\frac{39.5 - \mu}{\sigma} \leq \frac{T - \mu}{\sigma} \leq \frac{60.5 - \mu}{\sigma}\right) = \\ &= P\left(-2.10 \leq \frac{T - \mu}{\sigma} \leq 2.10\right) = 2 \cdot 0.4821 = 0.9642 \end{aligned}$$



12.2.

Vogliamo ora testare l'ipotesi che una moneta non sia truccata.

Allo scopo possiamo procedere come segue:

- lanciamo la moneta 100 volte
- se otteniamo un numero di teste compreso tra 40 e 60 **accettiamo l'ipotesi che la moneta non sia truccata**
- se invece otteniamo più di 60 o meno di 40 teste **respingiamo l'ipotesi che la moneta non sia truccata**

Così facendo siamo guidati dalla convinzione che se si ottengono un numero troppo alto o troppo basso di teste, se si è verificato cioè un evento abbastanza raro, la moneta è truccata.

Tuttavia una moneta non truccata può fornire su 100 lanci più di 60 o meno di 40 teste per cui, se ciò fosse accaduto proprio in occasione della nostra serie di lanci sbagliaremmo a rigettare l'ipotesi che la moneta non sia truccata.

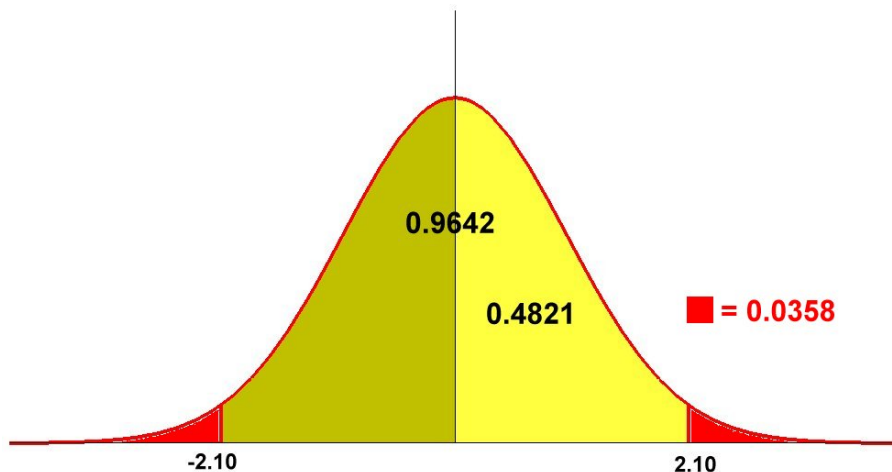
È allora importante stabilire quale è la probabilità di commettere un simile errore.

Se la moneta non è truccata otterremo un numero di teste compreso tra 40 e 60 con probabilità 0.9642, quindi la probabilità che le teste non siano comprese tra 40 e 60 è

$$P(T < 40 \text{ oppure } T > 60) = 1 - 0.9642 = 0.0358$$

Qualora ciò accada quindi noi rigetteremmo l'ipotesi che la moneta non sia truccata, erroneamente.

Ne segue che la probabilità di rigettare l'ipotesi nel caso in cui l'ipotesi sia corretta è 0.0358.

**12.3.**

Vogliamo programmare un test per verificare l'ipotesi che una moneta non

sia truccata disponendo dei risultati di 64 lanci e con un livello di significatività (probabilità di rigettare l'ipotesi la moneta non è truccata nel caso non sia effettivamente truccata) inferiore o uguale a 0,01

Rigetteremo l'ipotesi che la moneta non sia truccata nel caso in cui escano troppe o troppo poche teste; cioè se $T > m$ oppure $T < n$

Commetteremo un errore (di I specie) nel caso in cui l'ipotesi sia rigettata, essendo tuttavia vera; cioè nel caso in cui la moneta non truccata fornisce un numero di teste superiore ad m od inferiore ad n .

La probabilità di tale errore è quindi uguale alla probabilità che esca un numero di teste superiore ad m od inferiore ad n cioè alla

$$P(T < n \text{ oppure } T > m)$$

e deve essere inferiore o uguale a 0.01, per cui occorre trovare m ed n in modo che

$$P(T < n \text{ oppure } T > m) = 0.01$$

Per calcolare tale probabilità possiamo approssimare la distribuzione di Bernoulli con una distribuzione normale standard e poichè

$$\mu = np = 32 \quad , \quad \sigma = \sqrt{64 \frac{1}{4}} = 4$$

Possiamo normalizzare la variabile T che restituisce il numero di teste ponendo

$$\tau = \frac{T - \mu}{\sigma} = \frac{T - 32}{4}$$

Avremo quindi che

$$T \leq n \iff \tau \leq \frac{n - 32}{4}$$

$$T \geq m \iff \tau \geq \frac{m - 32}{4}$$

e deve essere

$$0.01 = P(T < n \text{ oppure } T > m) = P(\tau < \frac{n - 32}{4} \text{ oppure } \tau > \frac{m - 32}{4})$$

È ragionevole considerare

$$\frac{n - 32}{4} = -\frac{m - 32}{4} = -z_a$$

per cui essendo τ una variabile aleatoria normale standardizzata, per la simmetria della PDF Gaussiana, si ricava che deve essere

$$0.01 = F(-z_a) + 1 - F(z_a) = 2F(z_a)$$

e

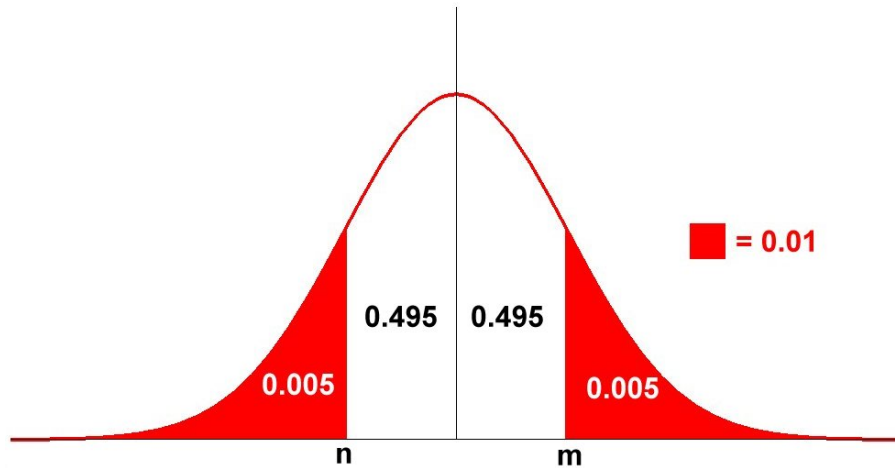
$$z_a = F^{-1}(0.05) = 2.575$$

Allora

$$\frac{n - 32}{4} = 2.575$$

e

$$32 + 2.575 \cdot 4 \approx 42, \quad 32 - 2.575 \cdot 4 \approx 22$$



Se ne conclude che, per operare con un livello di significatività di 0.01, si accetterà l'ipotesi che la moneta non sia truccata nel caso in cui si verifichi l'uscita di un numero di teste

$$22 \leq T \leq 42$$

12.4. Test statistico di ipotesi. È frequente la necessità di sottoporre a test statistico la validità di una ipotesi, tuttavia è essenziale ben definire cosa si intende con questo; infatti il linguaggio usato può essere frainteso se interpretato alla luce del senso comune.

Definiamo

H_0 l'affermazione che vogliamo sottoporre a test

e

H_a o H_1 una affermazione alternativa.

Va subito detto che H_a **non è necessariamente la negazione di H_0** ed anzi, in corrispondenza di una stessa ipotesi H_0 si possono programmare diversi test semplicemente scegliendo differenti ipotesi alternative H_a ciascuna delle quali in grado di mettere in luce differenti aspetti significativi.

Un esempio classico in grado di illustrare la situazione è il meccanismo di giudizio in un sistema giuridico nel quale:

Un individuo è considerato non colpevole fino a che non è provata la sua colpevolezza oltre ogni ragionevole dubbio.

Implicitamente una tale affermazione ritiene molto più grave giudicare colpevole un non colpevole piuttosto che giudicare non colpevole un colpevole.

Nei termini prima espressi avremo

H_0 non colpevole

e

H_a colpevole oltre ogni ragionevole dubbio.

Rigettare H_0 , nel caso in cui H_0 sia vera, significa giudicare colpevole un non colpevole ed è considerato più grave di accettare H_a nel caso in cui H_a sia falsa, cioè nel caso in cui si giudichi non colpevole oltre ragionevole dubbio un colpevole.

Diciamo che si commette un **errore di I specie** se si rigetta H_0 nel caso in cui H_0 è vera. (Si condanna un innocente).

Diciamo che si commette un **errore di II specie** se si rigetta H_a nel caso in cui H_a è vera. (Si assolve un colpevole, ma non oltre ogni ragionevole dubbio).

Definiamo inoltre

$P(\text{errore di I specie}) = \alpha$ livello di significatività del test.

$P(\text{errore di II specie}) = \beta$ potenza del test.

12.5. Uso dei Test statistici. Illustriamo il concetto di test statistico riprendendo il semplice ma significativo esempio che abbiamo già in precedenza considerato:

progettare un test per stabilire se una moneta è truccata; possiamo procedere in diverse maniere che esaminiamo qui di seguito.

Indichiamo con p la probabilità che esca Testa e con q la probabilità che esca Croce. e lanciamo la moneta 100 volte; chiamiamo T il numero di teste uscito e stabiliamo di adottare un livello di significatività $\alpha = 0.05$.

Se la moneta non è truccata T è una variabile aleatoria che ha una distribuzione binomiale (bernoulliana) di media μ e scarto quadratico σ dati da

$$\mu = np = 50 \quad , \quad \sigma = \sqrt{npq} = 5$$

e può essere approssimata con una variabile aleatoria normale standardizzata z definita dalla

$$z = \frac{n_t - \mu}{\sigma}$$

12.5.1. ---

Consideriamo l'ipotesi da testare

H_0 la moneta non è truccata cioè $p = q = \frac{1}{2}$.

a fronte della ipotesi alternativa

H_a la moneta è truccata cioè $p \neq \frac{1}{2}, q \neq \frac{1}{2}$.

In questo caso giudicheremo la moneta truccata se T è troppo grande $p > 0.5$ oppure se T è troppo piccolo $p < 0.5$ ed avremo un livello di significatività $\alpha = 0.05$ se

$$P(z < -z_a \text{ oppure } z > z_a) = 0.05$$

Dovrà quindi essere

$$0.05 = P(z > z_a) + P(z < -z_a) = 2P(z > z_a) = 2(1 - P(z < z_a)) = 2(1 - F(z_a))$$

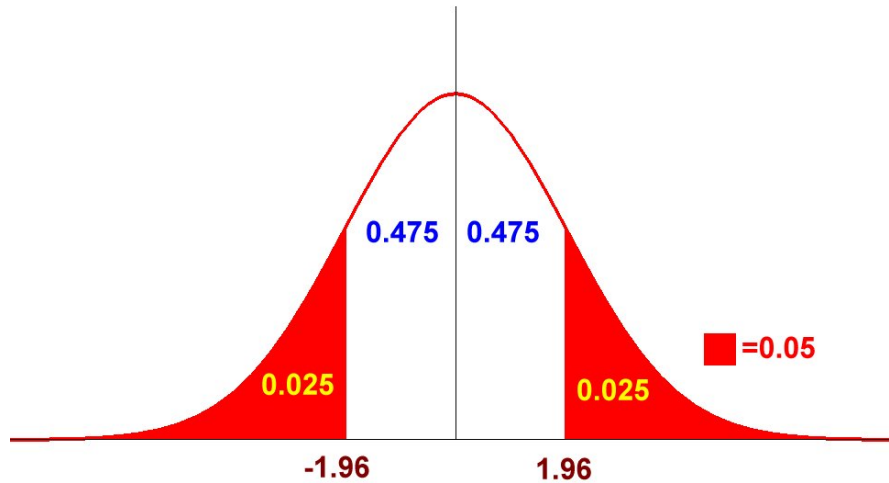
da cui

$$F(z_a) = 1 - 0.025 = 0.975 \quad , \quad z_a = F^{-1}(0.975) = 1.96$$

Rigetteremo cioè H_0 se $z < -z_a = -1.96$ oppure se $z > z_a = 1.96$ cioè se

$$\frac{T - \mu}{\sigma} < -1.96 \quad \text{oppure} \quad \frac{T - \mu}{\sigma} > 1.96$$

$$T < \mu - 1.96\sigma \approx 40 \quad \text{oppure} \quad T > \mu + 1.96\sigma \approx 60$$



Riassumendo possiamo affermare che decidendo di accettare l'ipotesi H_0 che la moneta non sia truccata a fronte dell'ipotesi alternativa H_a che $p \neq 0.5$ nel caso che il numero di teste uscite sia compreso tra 40 e 60, la probabilità di commettere un errore di prima specie e cioè di rigettare l'ipotesi la moneta non è truccata, quando realmente non è truccata, è 0.05

12.5.2.

Consideriamo l'ipotesi da testare

H_0 - la moneta non è truccata cioè $p = q = \frac{1}{2}$.
a fronte della ipotesi alternativa

H_a - la moneta è truccata cioè $p > \frac{1}{2}$.

In questo caso giudicheremo la moneta truccata se T è troppo grande ed se stabiliamo un livello di significatività di $\alpha = 0.05$ possiamo calcolare

$$P(z > z_a) = 0.05 = 1 - F(z_a)$$

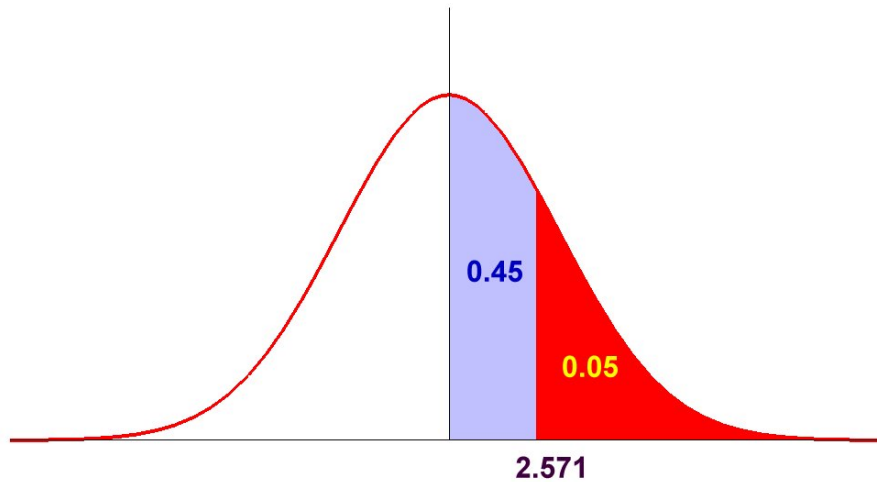
da cui

$$z_a = F^{-1}(0.95) = 2.571$$

Rigetteremo cioè H_0 se $z > z_a = 2.571$ da cui

$$\frac{T - \mu}{\sigma} > 2.571$$

$$T > \mu + 2.571\sigma \approx 63$$



Riassumendo possiamo affermare che decidendo di accettare l'ipotesi H_0 che la moneta non sia truccata a fronte dell'ipotesi alternativa H_a che $p > 0.5$ nel caso che il numero di teste uscite sia maggiore di 63, la probabilità di commettere un errore di prima specie e cioè di rigettare l'ipotesi la moneta non è truccata, quando realmente non è truccata, è 0.05

12.5.3.

Con riferimento ai due esempi precedenti osserviamo che se otteniamo un numero di teste $T = 62$, a parità di livello di significatività, nel primo caso rigettiamo l'ipotesi che la moneta sia truccata mentre nel secondo caso la accettiamo; ciò in conseguenza alla diversa formulazione dell'ipotesi H_a che deve essere scelta in modo da esprimere le esigenze del problema.

Se ad esempio il test è condotto allo scopo di stabilire se è equo giocare a Testa o Croce con quella moneta e si ha intenzione di puntare su Testa è chiaro che il secondo test meglio si adatta alla situazione.

12.6. Il Test χ^2 . Si consideri un esperimento in cui sono possibili k uscite

$$A_j \quad j = 1..k$$

di cui si ipotizzano le probabilità

$$p_j \quad j = 1..k$$

Si supponga di eseguire n volte l'esperimento e di ottenere delle frequenze di accadimento relative

$$x_j \quad j = 1..k$$

Uno stimatore della differenza tra frequenze osservate e frequenze ipotizzate si può definire come

$$\chi^2 = \sum_{j=1}^k \frac{(x_j - np_j)^2}{np_j}$$

e si può dimostrare che se $np_j \geq 5$ z^2 si può approssimare mediante una distribuzione χ^2 con $k-1$ gradi di libertà (supponendo che le frequenze teoriche possano essere stimate senza dover stimare statisticamente i parametri della popolazione).

Poichè i dati sono discreti e la variabile χ^2 è continua può essere opportuno apportare una correzione allo stimatore usando

$$\bar{\chi}^2 = \sum_{j=1}^k \frac{(|x_j - np_j| - .5)^2}{np_j}$$

$\bar{\chi}^2$ si chiama correzione di Yates.

Gli stimatori introdotti possono essere usati in test che tendano a stabilire se le frequenze teoriche p_j siano in accordo con i risultati ottenuti x_j .

12.7.

Si supponga di lanciare una moneta 200 volte e di ottenere 115 Tese e 85 Croci. e si supponga di voler sottoporre a test l'ipotesi H_0 che la moneta sia truccata.

Possiamo definire H_a in parecchi modi ed ottenere altrettanti risultati.

12.7.1.

Siano A_1 l'evento Testa ed A_2 l'evento Croce; con le notazioni precedenti avremo

$$p_1 = p_2 = 0.5 \quad x_1 = 115 \quad x_2 = 85 \quad n = 200$$

e possiamo calcolare

$$z^2 = \frac{(x_1 - np_1)^2}{np_1} + \frac{(x_2 - np_2)^2}{np_2} = \frac{(115 - 100)^2}{100} + \frac{(85 - 100)^2}{100} \approx 4.50$$

z^2 ha una distribuzione del tipo χ^2 con 1 grado di libertà.

Nel caso in cui le frequenze teoriche siano corrette z^2 assumerà valori molto piccoli e possiamo quindi decidere di sottoporre a test l'ipotesi

H_0 la moneta non è truccata, $z^2 \leq z_a$

a fronte dell'ipotesi alternativa

H_a la moneta è truccata, $z^2 > z_a$

Il valore χ_a^2 deve essere determinato dal valore di significatività α ; dovremo cioè scegliere χ_a^2 in modo che

$$P(\chi^2 > \chi_a^2) = \alpha$$

Se $\alpha = 0.05$, si ricava che $z_a = 3.84$ e quindi, poichè $z^2 = 4.50$ si conclude che la moneta è truccata.

12.7.2.

Usando la correzione di Yates otteniamo un valore

$$\begin{aligned}\bar{\chi}^2 &= \frac{(|x_1 - np_1| - 0.5)^2}{np_1} + \frac{(|x_2 - np_2| - 0.5)^2}{np_2} = \\ &= \frac{(|115 - 100| - 0.5)^2}{100} + \frac{(|85 - 100| - 0.5)^2}{100} \approx 4.205 > 3.84\end{aligned}$$

anche in questo caso si conclude che la moneta è truccata.

12.7.3.

Possiamo anche procedere usando il fatto che il lancio di una moneta non truccata per 200 volte segue una distribuzione binomiale di media $\mu = 100$ e scarto quadratico $\sigma = \sqrt{50} \approx 7.07$ ed usando un test a due code.

Detta z la variabile aleatoria normale standardizzata che approssima $\frac{x_1 - \mu}{\sigma}$ (ricordiamo che x_1 è il numero di teste), e scelto un livello di significatività $\alpha = 0.05$, possiamo sottoporre a test l'ipotesi

H_0 la moneta non è truccata,

a fronte di

H_a la moneta è truccata, $z < -z_\alpha$ oppure $z > z_\alpha$

Il valore z_α deve essere in modo che

$$P(z > z_\alpha) = \alpha/2$$

Se $\alpha = 0.05$, si ricava che $z_\alpha = 1.196$ e quindi, poichè

$$z = \frac{115 - 100}{7.07} \approx 2.12 > 1.196$$

si conclude che la moneta è truccata.

12.7.4.

Possiamo infine procedere usando il fatto che il lancio di una moneta non truccata per 200 volte segue una distribuzione binomiale di media $\mu = 100$ e scarto quadratico $\sigma = \sqrt{50} \approx 7.07$ ed usando un test a una sola coda.

Come prima, detta z la variabile aleatoria normale standardizzata che approssima $\frac{x_1 - \mu}{\sigma}$ e scelto un livello di significatività $\alpha = 0.05$, possiamo sottoporre a test l'ipotesi

H_0 la moneta non è truccata,

a fronte di

H_a la moneta è truccata, $z > z_\alpha$

Il valore z_α deve essere calcolato in modo che

$$P(z > z_\alpha) = \alpha$$

Se $\alpha = 0.05$, si ricava che $z_\alpha = 1.645$ e quindi, poichè

$$z = \frac{115 - 100}{7.07} \approx 2.12 > 1.645$$

si conclude che la moneta è truccata.

12.8. Test sulle Medie.

12.8.1.

Siano $\xi_1, \xi_2, \dots, \xi_n$ variabili aleatorie costituenti un campione di grandezza n . (Con ciò si intende che ξ_k è una variabile aleatoria che si ottiene estraendo un elemento dalla popolazione).

Definiamo una nuova variabile aleatoria $\bar{\xi}$, che chiamiamo media campionaria, mediante la

$$\bar{\xi} = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}$$

La distribuzione di probabilità di $\bar{\xi}$ è detta distribuzione campionaria della media.

Si può verificare che, se la popolazione ha media μ e varianza σ^2 allora la media e la varianza di $\bar{\xi}$ sono date da

$$\mu_{\bar{\xi}} = \mu \quad , \quad \sigma_{\bar{\xi}}^2 = \frac{\sigma^2}{n} \quad , \quad \left(\sigma_{\bar{\xi}} = \frac{\sigma}{\sqrt{n}} \right)$$

Si può inoltre dimostrare che la variabile aleatoria

$$\frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}}$$

ha, per n grande, una distribuzione normale standard.

$\frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}}$ si usa come stimatore della media.

12.8.2.

Un produttore afferma che la vita media delle lampade di sua produzione è $\mu = 1600$ ore. Vogliamo sottoporre a test l'affermazione con un livello di significatività $\alpha = 0.04$ e $\alpha = 0.01$.

Consideriamo un campione di 100 lampade e verifichiamone la vita media.

Supponiamo di ottenere vita media $\bar{\xi} = 1570$ ore e scarto quadratico $\sigma_{\bar{\xi}} = 120$ ore

Consideriamo l'ipotesi

$$H_0 \quad - \quad \mu = 1600$$

a fronte dell'ipotesi alternativa

$$H_a \quad - \quad \mu < 1600$$

Usando $\sigma_{\bar{\xi}}$ come stima di σ avremo che

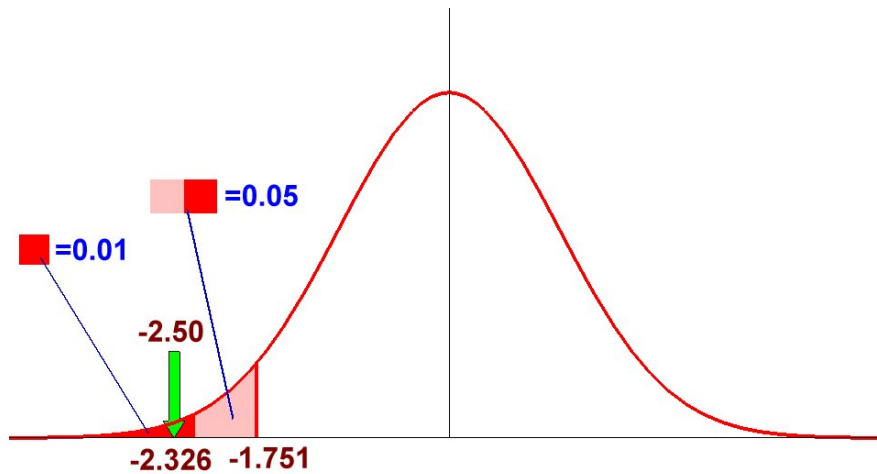
$$\frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}} = \frac{1570 - 1600}{120 / \sqrt{100}} \approx -2.50$$

Poichè $\frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}}$ è una variabile asintoticamente normale si ha che

$$P \left(\frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}} < z_a \right) = 0.04 \quad \text{se} \quad z_a = -1.751$$

mentre

$$P \left(\frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}} < z_a \right) = 0.01 \quad \text{se} \quad z_a = -2.326$$



Poichè

$$-2.50 < -2.326 < -1.752$$

Il valore che abbiamo ottenuto ha probabilità di uscire, nel caso in cui H_0 sia vera, inferiore a 0.01. Ne segue che H_0 va rigettata sia al livello di significatività 0.04 sia al livello 0.01.

12.8.3.

Il carico medio di rottura di un cavo è $\mu = 1800$ kg con una varianza $\sigma^2 = 100$ kg.

Il produttore afferma di aver migliorato la tecnica di produzione e di aver aumentato il carico medio di rottura.

Si provano 50 cavi e si riscontra una carico medio di rottura $\bar{\xi} = 1850$ kg.

È corretto sostenere l'affermazione con un livello di significatività $\alpha = 0.01$?

Possiamo sottoporre a test l'ipotesi

H_0 - $\mu = 1800$ kg.

a fronte dell'ipotesi alternativa

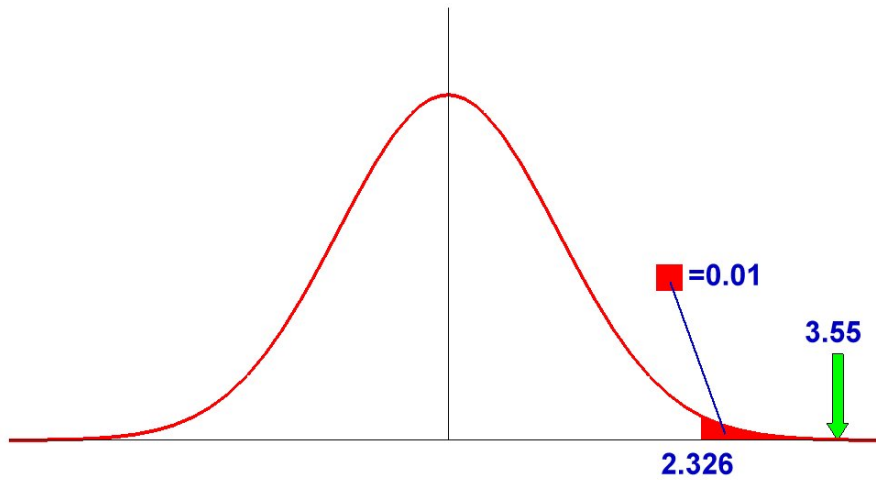
H_a - $\mu > 1800$ kg.

Calcoliamo

$$\frac{\bar{\xi} - \mu}{\sigma/\sqrt{n}} = \frac{1850 - 1800}{100/\sqrt{50}} \approx 3.55$$

Poichè $\frac{\bar{\xi} - \mu}{\sigma/\sqrt{n}}$ è una variabile asintoticamente normale si ha che

$$P\left(\frac{\bar{\xi} - \mu}{\sigma/\sqrt{n}} > z_a\right) = 0.01 \quad \text{se} \quad z_a = 2.326$$



Poichè

$$2.326 < 3.55$$

il valore che abbiamo ottenuto ha probabilità di uscire, nel caso in cui H_0 sia vera, inferiore a 0.01. Ne segue che H_0 va rigettata al livello di significatività 0.01 e quindi l'affermazione del produttore può essere accettata con lo stesso livello di significatività.

12.9. Test sulle Varianze.

12.9.1.

Siano $\xi_1, \xi_2, \dots, \xi_n$ variabili aleatorie costituenti un campione di grandezza n . (Con ciò si intende che ξ_k è una variabile aleatoria che si ottiene estraendo un elemento dalla popolazione).

Definiamo una nuova variabile aleatoria s^2 , che chiamiamo varianza campionaria, mediante la

$$s^2 = \frac{(\xi_1 - \bar{\xi})^2 + (\xi_2 - \bar{\xi})^2 + \dots + (\xi_n - \bar{\xi})^2}{n}$$

La distribuzione di probabilità di s^2 è detta distribuzione campionaria della varianza.

Si può verificare che, se la popolazione ha media μ e varianza σ^2 allora la media di s^2 è data da

$$\mu_{s^2} = \frac{n-1}{n} \sigma^2$$

Inoltre se la popolazione è distribuita normalmente, la variabile aleatoria definita da

$$\frac{ns^2}{\sigma^2}$$

ha una distribuzione di tipo χ^2 con $n - 1$ gradi di libertà.

$\frac{ns^2}{\sigma^2}$ si usa come stimatore della varianza.

12.9.2.

Nel passato la deviazione standard del peso di certe confezioni da 40. g. è stata $\sigma = 0.25$ g.

L'esame di un campione casuale di 20 confezioni ha accertato una deviazione standard di 0.32 g.

Stabilire se l'apparente aumento è significativo a livello $\alpha = 0.05$ e $\alpha = 0.01$.

Possiamo sottoporre a test l'ipotesi

$$H_0 \quad - \quad \sigma = 0.25 \text{ g.}$$

a fronte dell'ipotesi alternativa

$$H_a \quad - \quad \sigma > 0.25 \text{ g.}$$

Calcoliamo

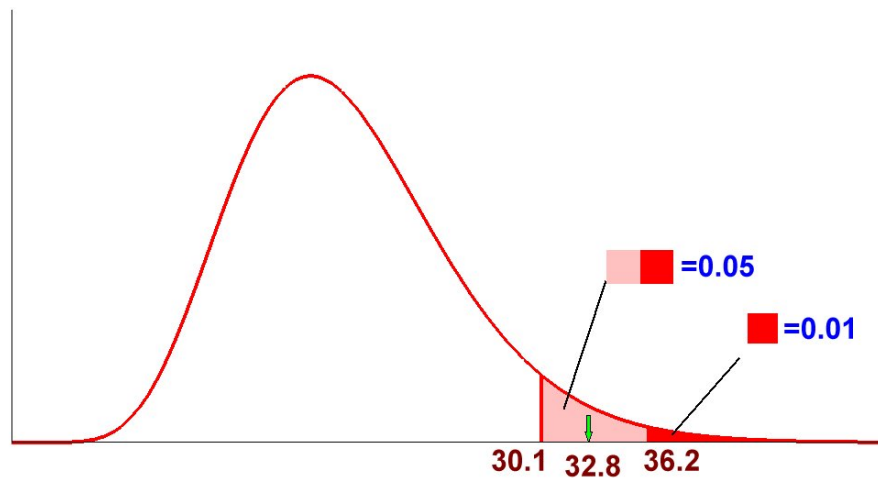
$$\frac{ns^2}{\sigma^2} = \frac{20 \cdot 0.32}{0.25} \approx 32.8$$

Poichè $\frac{ns^2}{\sigma^2}$ ha una distribuzione di tipo χ^2 con 19 gradi di libertà

$$P\left(\frac{ns^2}{\sigma^2} > z_a\right) = 0.05 \quad \text{se} \quad z_a = 30.1$$

mentre

$$P\left(\frac{ns^2}{\sigma^2} > z_a\right) = 0.01 \quad \text{se} \quad z_a = 36.2$$



Poichè

$$30.1 < 32.8 < 36.1$$

il valore che abbiamo ottenuto ha probabilità di uscire, nel caso in cui H_0 sia vera, inferiore a 0.05 ma superiore a 0.01.

Ne segue che H_0 va rigettata al livello di significatività 0.05 ma non può essere rifiutata al livello 0.01

CAPITOLO 13

Appendice

1. Il Problema del compleanno

Se in una stanza sono presenti N persone , qual'è la probabilità che almeno due tra loro abbiano lo stesso compleanno?

Esistono 366 compleanni diversi (tra i possibili compleanni c'è anche il 29 febbraio degli anni bisestili).

La probabilità che 2 persone abbiano compleanni diversi è :

$$P_2 = \frac{366}{366} \frac{365}{366}$$

La probabilità che 3 persone abbiano tutte compleanni diversi è :

$$P_3 = \frac{366}{366} \frac{365}{366} \frac{364}{366}$$

La probabilità che 4 persone abbiano tutte compleanni diversi è :

$$P_4 = \frac{366}{366} \frac{365}{366} \frac{364}{366} \frac{363}{366} = \frac{366!}{366!366^4}$$

La probabilità che n persone abbiano compleanni diversi è :

$$P_n = \frac{366!}{(366 - n)!366^n}$$

Si calcola che

$$P_{50} = \frac{366!}{(316)!366^{50}} \approx 0.03$$

$$P_{60} = \frac{366!}{(306)!366^{60}} \approx 0.006$$

Possiamo allora chiederci:

Quante persone ci devono essere in una stanza affinché almeno due tra esse abbiano lo stesso compleanno?

Se sono presenti n persone, la probabilità che tutte abbiano compleanni diversi è

$$P_n = \frac{366!}{(366 - n)!366^n}$$

Affinchè almeno due tra esse abbiano lo stesso compleanno ciò non deve accadere e quindi occorre calcolare la probabilità Q_n dell'evento complementare a

Tutte e n le persone hanno compleanni diversi.

la cui probabilità è P_n
Pertanto

$$Q_n = 1 - P_n = 1 - \frac{366!}{(366 - n)!366^n}$$

$$Q_{50} = 1 - 0.03 = 0.97$$

$$Q_{60} = 1 - 0.006 = 0.994$$

2. Il Paradosso di Bertrand

Data una circonferenza di raggio 1, calcolare la probabilità di tracciare una corda di lunghezza inferiore a $\sqrt{3}$ (il lato di un triangolo equilatero in essa inscritto).

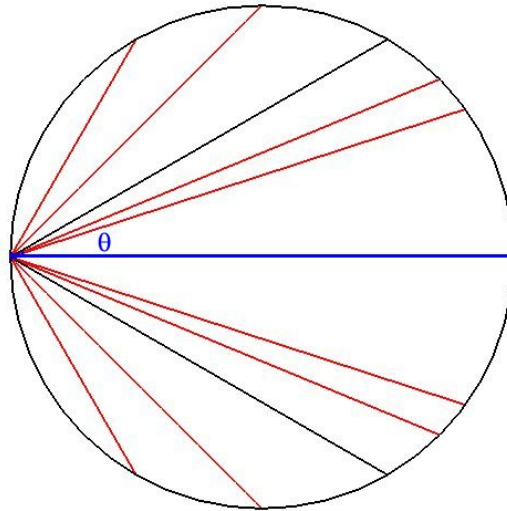
Possiamo procedere in diversi modi.

2.1.

- Consideriamo un punto sulla circonferenza e un valore

$$\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

- disegniamo una corda che passa per il punto fissato e forma un angolo θ con il diametro per il punto fissato.



Avremo una corda inferiore a $\sqrt{3}$ se

$$|\theta| \geq \frac{\pi}{6}$$

e quindi la probabilità di ottenere una corda inferiore a $\sqrt{3}$ è

$$\frac{\pi - \frac{\pi}{3}}{\pi} = \frac{2}{3}$$

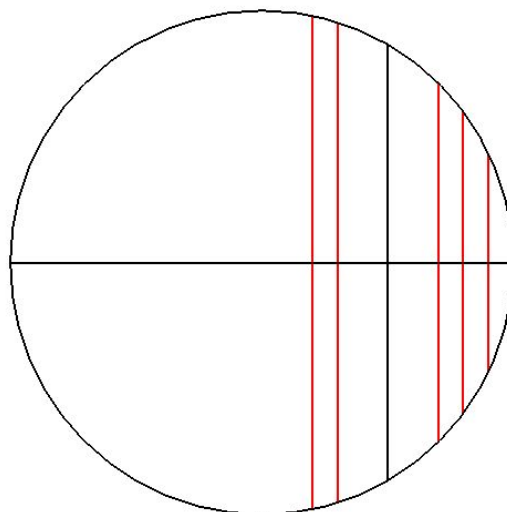
2.2.

Potremmo anche procedere come segue:

- consideriamo un punto su un raggio della circonferenza cioè un valore

$$t \in [0, 1]$$

- disegniamo una corda che passa per il punto fissato perpendicolare al raggio.



Avremo una corda inferiore a $\sqrt{3}$ se

$$t \in \left[\frac{1}{2}, 1\right]$$

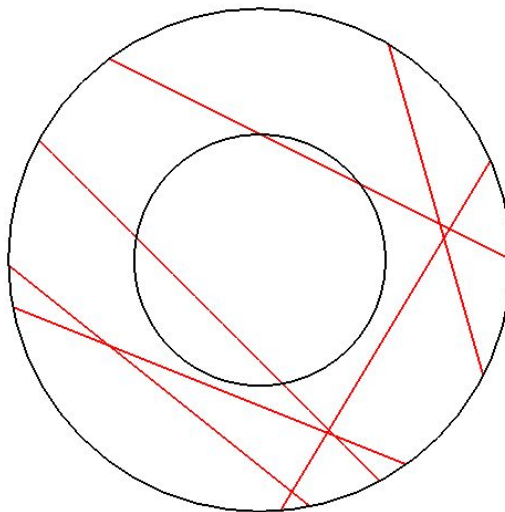
e quindi la probabilità di ottenere una corda inferiore a $\sqrt{3}$ è

$$\frac{1}{2}$$

2.3.

Potremmo infine procedere così

- consideriamo un cerchio di raggio $\frac{1}{2}$ e scegliamo un punto a caso nel cerchio
- disegniamo una corda che passa per il punto fissato perpendicolare al raggio che passa per il punto scelto.



Avremo una corda inferiore a $\sqrt{3}$ se il punto è stato scelto nella corona circolare che ha area

$$\pi - \frac{\pi}{4}$$

e quindi la probabilità di ottenere una corda inferiore a $\sqrt{3}$ è

$$\frac{\frac{3}{4}\pi}{\pi} = \frac{3}{4}$$

3. TRIANGOLO DI TARTAGLIA E BINOMIO DI NEWTON

Si ha

TEOREMA 13.1.

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$$

DIMOSTRAZIONE.

$$\begin{aligned} \binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n+1-k)!} = \\ &= \frac{n![(n+1-k) + k]}{k!(n+1-k)!} = \frac{(n+1)!}{k!(n+1-k)!} = \binom{n+1}{k} \end{aligned}$$

□

La precedente uguaglianza consente di costruire il così detto triangolo di Tartaglia:

$$\begin{array}{ccccccc} & & \binom{1}{0} & \binom{1}{1} & & & \\ & & \binom{2}{0} & \binom{2}{1} & \binom{2}{2} & & \\ & & \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & & \binom{3}{3} \\ & \dots & \dots & \dots & \dots & \dots & \dots \\ & \binom{n}{0} & \dots & \binom{n}{k-1} & & \binom{n}{k} & \dots & \binom{n}{n} \\ & \dots & \dots & \dots & \binom{n+1}{k} & \dots & \dots & \dots \\ & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{array}$$

È importante osservare che ogni elemento del triangolo si può ottenere dalla somma dei due elementi della riga precedente, che occupano la posizione sopra e a sinistra della posizione occupata dall'elemento considerato.

Vale inoltre il seguente risultato

TEOREMA 13.2. (*binomio di Newton*)

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

DIMOSTRAZIONE. E' immediato verificare che la formula vale per $n = 1$.

Proviamo ora che, se la formula è valida per n , allora è valida anche per $n + 1$. Si ha

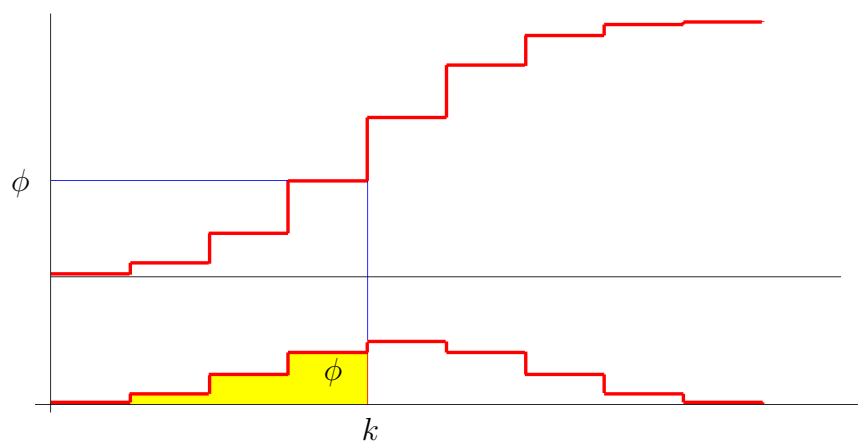
$$\begin{aligned}
 (13.1) \quad (a+b)^{n+1} &= (a+b)^n(a+b) = \\
 &= \left(\sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \right) (a+b) = \\
 &= a \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k + b \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k = \\
 &= \sum_{k=0}^n \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^n \binom{n}{k} a^{n-k} b^{k+1} = \\
 &= a^{n+1} + \sum_{k=1}^n \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^{n-1} \binom{n}{k} a^{n-k} b^{k+1} + b^{n+1} = \\
 &= a^{n+1} + \sum_{k=1}^n \binom{n}{k} a^{n+1-k} b^k + \sum_{k=1}^n \binom{n}{k-1} a^{n+1-k} b^k + b^{n+1} = \\
 &= a^{n+1} + \sum_{k=1}^n \left(\binom{n}{k} + \binom{n}{k-1} \right) a^{n+1-k} b^k + b^{n+1} = \\
 &= a^{n+1} + \sum_{k=1}^n \binom{n+1}{k} a^{n+1-k} b^k + b^{n+1} = \\
 &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k
 \end{aligned}$$

□

Il triangolo di Tartaglia per $n \leq 10$

n \ k	0	1	2	3	4	5	6	7	8	9	10
1	1										
1	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	9	1		
10	1	10	45	120	210	256	210	120	45	10	1

Tabelle
di alcune
distribuzioni cumulative binomiali



n=N	+	+	+	+	+	p	+	+	+	+
k	-	-	-	-	-	phi	-	-	-	-

N = numero delle prove
 k = numero dei successi
 p = probabilità di successo

Tabella della distribuzione cumulativa binomiale per $n = 3, 4, 5, 6$

[illegible]

Tabella della distribuzione cumulativa binomiale per $n = 7, 8, 9$

n=7	p	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
k														
0		0.932	0.698	0.478	0.210	0.082	0.028	0.008	0.002	0.000	0.000	0.000	0.000	0.000
1		0.998	0.956	0.850	0.577	0.329	0.159	0.063	0.019	0.004	0.000	0.000	0.000	0.000
2		1.000	0.996	0.974	0.852	0.647	0.420	0.227	0.096	0.029	0.005	0.000	0.000	0.000
3		1.000	1.000	0.997	0.967	0.874	0.710	0.500	0.290	0.126	0.033	0.003	0.000	0.000
4		1.000	1.000	1.000	0.995	0.971	0.904	0.773	0.580	0.353	0.148	0.026	0.004	0.000
5		1.000	1.000	1.000	1.000	0.996	0.981	0.938	0.841	0.671	0.423	0.150	0.044	0.002
6		1.000	1.000	1.000	1.000	1.000	0.998	0.992	0.972	0.918	0.790	0.522	0.302	0.068
7		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=8	p	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
k														
0		0.923	0.663	0.430	0.168	0.058	0.017	0.004	0.001	0.000	0.000	0.000	0.000	0.000
1		0.997	0.943	0.813	0.503	0.255	0.106	0.035	0.009	0.001	0.000	0.000	0.000	0.000
2		1.000	0.994	0.962	0.797	0.552	0.315	0.145	0.050	0.011	0.001	0.000	0.000	0.000
3		1.000	1.000	0.995	0.944	0.806	0.594	0.363	0.174	0.058	0.010	0.000	0.000	0.000
4		1.000	1.000	1.000	0.990	0.942	0.826	0.637	0.406	0.194	0.056	0.005	0.000	0.000
5		1.000	1.000	1.000	0.999	0.989	0.950	0.855	0.685	0.448	0.203	0.038	0.006	0.000
6		1.000	1.000	1.000	1.000	0.999	0.991	0.965	0.894	0.745	0.497	0.187	0.057	0.003
7		1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.983	0.942	0.832	0.570	0.337	0.077
8		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n=9	p	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
k														
0		0.914	0.630	0.387	0.134	0.040	0.010	0.002	0.000	0.000	0.000	0.000	0.000	0.000
1		0.997	0.929	0.775	0.436	0.196	0.071	0.020	0.004	0.000	0.000	0.000	0.000	0.000
2		1.000	0.992	0.947	0.738	0.463	0.232	0.090	0.025	0.004	0.000	0.000	0.000	0.000
3		1.000	0.999	0.992	0.914	0.730	0.483	0.254	0.099	0.025	0.003	0.000	0.000	0.000
4		1.000	1.000	0.999	0.980	0.901	0.733	0.500	0.267	0.099	0.020	0.001	0.000	0.000
5		1.000	1.000	1.000	0.997	0.975	0.901	0.746	0.517	0.270	0.086	0.008	0.001	0.000
6		1.000	1.000	1.000	1.000	0.996	0.975	0.910	0.768	0.537	0.262	0.053	0.008	0.000
7		1.000	1.000	1.000	1.000	1.000	0.996	0.980	0.929	0.804	0.564	0.225	0.071	0.003
8		1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.990	0.960	0.866	0.613	0.370	0.086
9		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabella della distribuzione cumulativa binomiale per $n = 10$

[illegible]

Tabella della distribuzione cumulativa binomiale per $n = 20$

[illegible]

Tabella della distribuzione cumulativa binomiale per $n = 30$

[illegible]

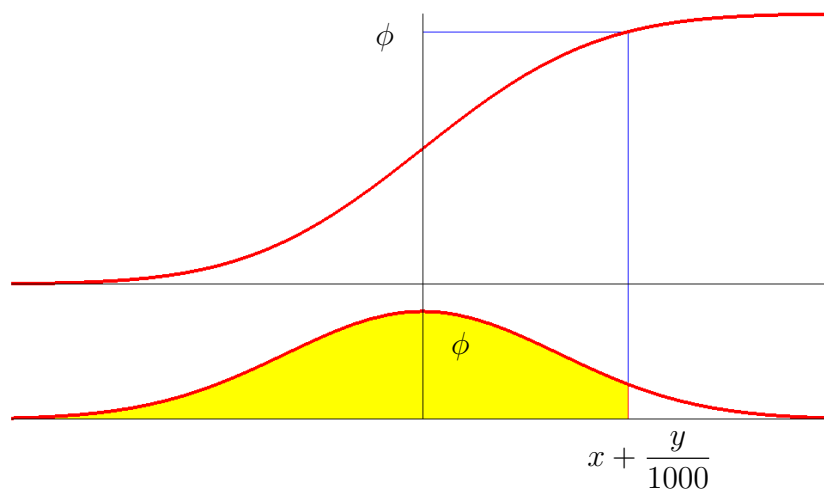
[illegible]

Tabella della distribuzione cumulativa normale

	0	1	2	3	4	5	6	7	8	9
0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Valori da 0 a .499

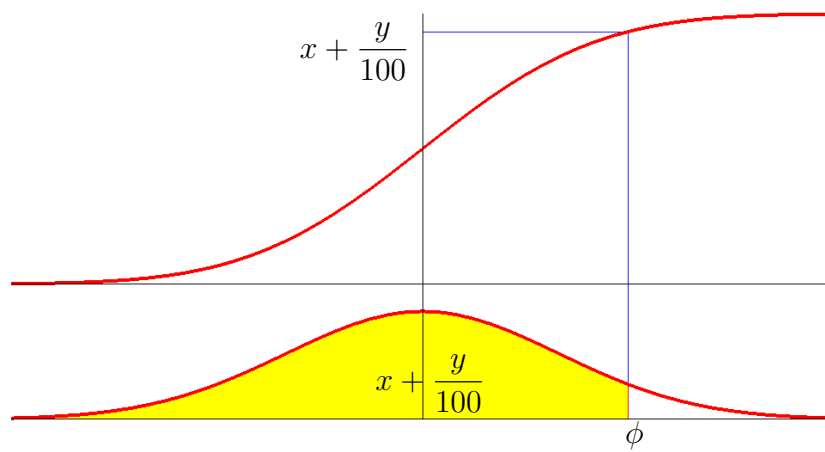
[illegible]

Tabella della inversa distribuzione cumulativa normale

	0	1	2	3	4	5	6	7	8	9
0	0.0000	0.0025	0.0050	0.0075	0.0100	0.0125	0.0150	0.0175	0.0201	0.0226
0.01	0.0251	0.0276	0.0301	0.0326	0.0351	0.0376	0.0401	0.0426	0.0451	0.0476
0.02	0.0502	0.0527	0.0552	0.0577	0.0602	0.0627	0.0652	0.0677	0.0702	0.0728
0.03	0.0753	0.0778	0.0803	0.0828	0.0853	0.0878	0.0904	0.0929	0.0954	0.0979
0.04	0.1004	0.1030	0.1055	0.1080	0.1105	0.1130	0.1156	0.1181	0.1206	0.1231
0.05	0.1257	0.1282	0.1307	0.1332	0.1358	0.1383	0.1408	0.1434	0.1459	0.1484
0.06	0.1510	0.1535	0.1560	0.1586	0.1611	0.1637	0.1662	0.1687	0.1713	0.1738
0.07	0.1764	0.1789	0.1815	0.1840	0.1866	0.1891	0.1917	0.1942	0.1968	0.1993
0.08	0.2019	0.2045	0.2070	0.2096	0.2121	0.2147	0.2173	0.2198	0.2224	0.2250
0.09	0.2275	0.2301	0.2327	0.2353	0.2378	0.2404	0.2430	0.2456	0.2482	0.2508
0.1	0.2533	0.2559	0.2585	0.2611	0.2637	0.2663	0.2689	0.2715	0.2741	0.2767
0.11	0.2793	0.2819	0.2845	0.2871	0.2898	0.2924	0.2950	0.2976	0.3002	0.3029
0.12	0.3055	0.3081	0.3107	0.3134	0.3160	0.3186	0.3213	0.3239	0.3266	0.3292
0.13	0.3319	0.3345	0.3372	0.3398	0.3425	0.3451	0.3478	0.3505	0.3531	0.3558
0.14	0.3585	0.3611	0.3638	0.3665	0.3692	0.3719	0.3745	0.3772	0.3799	0.3826
0.15	0.3853	0.3880	0.3907	0.3934	0.3961	0.3989	0.4016	0.4043	0.4070	0.4097
0.16	0.4125	0.4152	0.4179	0.4207	0.4234	0.4261	0.4289	0.4316	0.4344	0.4372
0.17	0.4399	0.4427	0.4454	0.4482	0.4510	0.4538	0.4565	0.4593	0.4621	0.4649
0.18	0.4677	0.4705	0.4733	0.4761	0.4789	0.4817	0.4845	0.4874	0.4902	0.4930
0.19	0.4958	0.4987	0.5015	0.5044	0.5072	0.5101	0.5129	0.5158	0.5187	0.5215
0.2	0.5244	0.5273	0.5302	0.5330	0.5359	0.5388	0.5417	0.5446	0.5476	0.5505
0.21	0.5534	0.5563	0.5592	0.5622	0.5651	0.5681	0.5710	0.5740	0.5769	0.5799
0.22	0.5828	0.5858	0.5888	0.5918	0.5948	0.5978	0.6008	0.6038	0.6068	0.6098
0.23	0.6128	0.6158	0.6189	0.6219	0.6250	0.6280	0.6311	0.6341	0.6372	0.6403
0.24	0.6433	0.6464	0.6495	0.6526	0.6557	0.6588	0.6620	0.6651	0.6682	0.6713
0.25	0.6745	0.6776	0.6808	0.6840	0.6871	0.6903	0.6935	0.6967	0.6999	0.7031
0.26	0.7063	0.7095	0.7128	0.7160	0.7192	0.7225	0.7257	0.7290	0.7323	0.7356
0.27	0.7388	0.7421	0.7454	0.7488	0.7521	0.7554	0.7588	0.7621	0.7655	0.7688
0.28	0.7722	0.7756	0.7790	0.7824	0.7858	0.7892	0.7926	0.7961	0.7995	0.8030
0.29	0.8064	0.8099	0.8134	0.8169	0.8204	0.8239	0.8274	0.8310	0.8345	0.8381

	0	1	2	3	4	5	6	7	8	9
0.3	0.8416	0.8452	0.8488	0.8524	0.8560	0.8596	0.8632	0.8669	0.8706	0.8742
0.31	0.8779	0.8816	0.8853	0.8890	0.8927	0.8965	0.9002	0.9040	0.9078	0.9116
0.32	0.9154	0.9192	0.9230	0.9269	0.9307	0.9346	0.9385	0.9424	0.9463	0.9502
0.33	0.9542	0.9581	0.9621	0.9661	0.9701	0.9741	0.9782	0.9822	0.9863	0.9904
0.34	0.9945	0.9986	1.0027	1.0069	1.0110	1.0152	1.0194	1.0237	1.0279	1.0322
0.35	1.0364	1.0407	1.0451	1.0494	1.0537	1.0581	1.0625	1.0669	1.0714	1.0758
0.36	1.0803	1.0848	1.0893	1.0939	1.0985	1.1031	1.1077	1.1123	1.1170	1.1217
0.37	1.1264	1.1311	1.1359	1.1407	1.1455	1.1503	1.1552	1.1601	1.1650	1.1700
0.38	1.1750	1.1800	1.1850	1.1901	1.1952	1.2004	1.2055	1.2107	1.2160	1.2212
0.39	1.2265	1.2319	1.2372	1.2426	1.2481	1.2536	1.2591	1.2646	1.2702	1.2759
0.4	1.2816	1.2873	1.2930	1.2988	1.3047	1.3106	1.3165	1.3225	1.3285	1.3346
0.41	1.3408	1.3469	1.3532	1.3595	1.3658	1.3722	1.3787	1.3852	1.3917	1.3984
0.42	1.4051	1.4118	1.4187	1.4255	1.4325	1.4395	1.4466	1.4538	1.4611	1.4684
0.43	1.4758	1.4833	1.4909	1.4985	1.5063	1.5141	1.5220	1.5301	1.5382	1.5464
0.44	1.5548	1.5632	1.5718	1.5805	1.5893	1.5982	1.6072	1.6164	1.6258	1.6352
0.45	1.6449	1.6546	1.6646	1.6747	1.6849	1.6954	1.7060	1.7169	1.7279	1.7392
0.46	1.7507	1.7624	1.7744	1.7866	1.7991	1.8119	1.8250	1.8384	1.8522	1.8663
0.47	1.8808	1.8957	1.9110	1.9268	1.9431	1.9600	1.9774	1.9954	2.0141	2.0335
0.48	2.0537	2.0748	2.0969	2.1201	2.1444	2.1701	2.1973	2.2262	2.2571	2.2904
0.49	2.3263	2.3656	2.4089	2.4573	2.5121	2.5758	2.6521	2.7478	2.8782	3.0902

Tabella della inversa distribuzione normale

Valori da .4000 a .4999

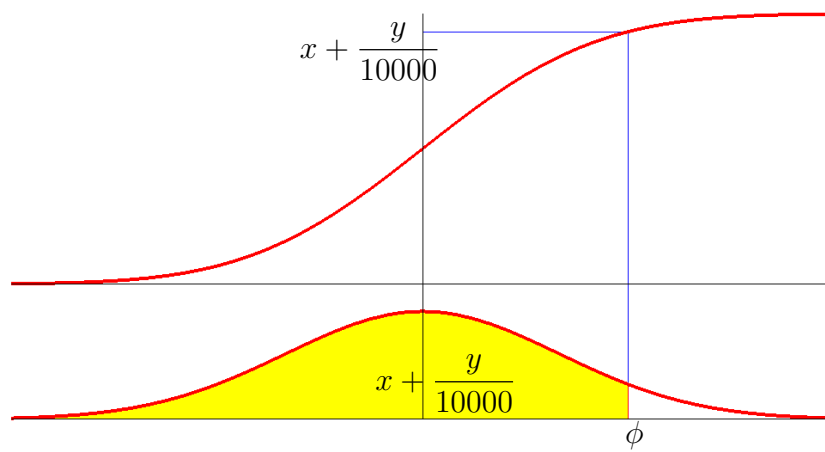
[illegible]

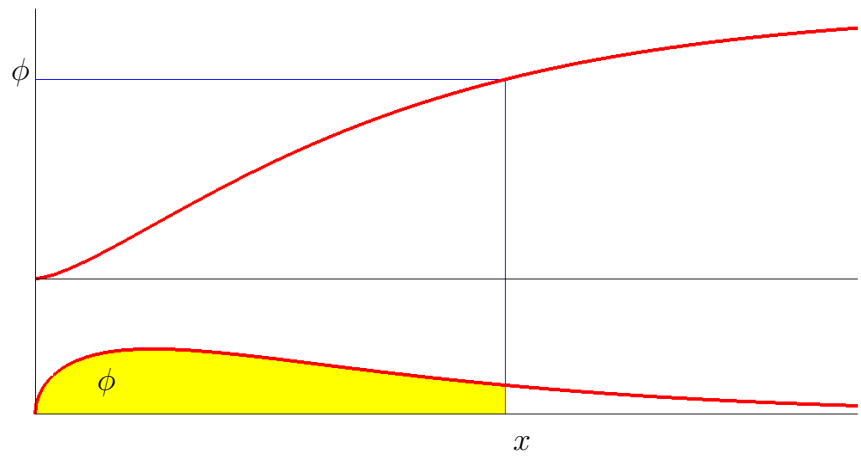
Tabella della distribuzione normale cumulativa inversa (code)

	0	1	2	3	4	5	6	7	8	9
0.4	1.2816	1.2821	1.2827	1.2833	1.2838	1.2844	1.2850	1.2856	1.2861	1.2867
0.401	1.2873	1.2878	1.2884	1.2890	1.2896	1.2901	1.2907	1.2913	1.2919	1.2925
0.402	1.2930	1.2936	1.2942	1.2948	1.2953	1.2959	1.2965	1.2971	1.2977	1.2983
0.403	1.2988	1.2994	1.3000	1.3006	1.3012	1.3018	1.3023	1.3029	1.3035	1.3041
0.404	1.3047	1.3053	1.3059	1.3064	1.3070	1.3076	1.3082	1.3088	1.3094	1.3100
0.405	1.3106	1.3112	1.3118	1.3124	1.3130	1.3135	1.3141	1.3147	1.3153	1.3159
0.406	1.3165	1.3171	1.3177	1.3183	1.3189	1.3195	1.3201	1.3207	1.3213	1.3219
0.407	1.3225	1.3231	1.3237	1.3243	1.3249	1.3255	1.3261	1.3267	1.3273	1.3279
0.408	1.3285	1.3291	1.3298	1.3304	1.3310	1.3316	1.3322	1.3328	1.3334	1.3340
0.409	1.3346	1.3352	1.3358	1.3365	1.3371	1.3377	1.3383	1.3389	1.3395	1.3401
0.41	1.3408	1.3414	1.3420	1.3426	1.3432	1.3438	1.3445	1.3451	1.3457	1.3463
0.411	1.3469	1.3476	1.3482	1.3488	1.3494	1.3500	1.3507	1.3513	1.3519	1.3525
0.412	1.3532	1.3538	1.3544	1.3551	1.3557	1.3563	1.3569	1.3576	1.3582	1.3588
0.413	1.3595	1.3601	1.3607	1.3614	1.3620	1.3626	1.3633	1.3639	1.3645	1.3652
0.414	1.3658	1.3664	1.3671	1.3677	1.3684	1.3690	1.3696	1.3703	1.3709	1.3716
0.415	1.3722	1.3728	1.3735	1.3741	1.3748	1.3754	1.3761	1.3767	1.3774	1.3780
0.416	1.3787	1.3793	1.3800	1.3806	1.3813	1.3819	1.3826	1.3832	1.3839	1.3845
0.417	1.3852	1.3858	1.3865	1.3871	1.3878	1.3885	1.3891	1.3898	1.3904	1.3911
0.418	1.3917	1.3924	1.3931	1.3937	1.3944	1.3951	1.3957	1.3964	1.3970	1.3977
0.419	1.3984	1.3990	1.3997	1.4004	1.4010	1.4017	1.4024	1.4031	1.4037	1.4044
0.42	1.4051	1.4057	1.4064	1.4071	1.4078	1.4084	1.4091	1.4098	1.4105	1.4112
0.421	1.4118	1.4125	1.4132	1.4139	1.4146	1.4152	1.4159	1.4166	1.4173	1.4180
0.422	1.4187	1.4193	1.4200	1.4207	1.4214	1.4221	1.4228	1.4235	1.4242	1.4249
0.423	1.4255	1.4262	1.4269	1.4276	1.4283	1.4290	1.4297	1.4304	1.4311	1.4318
0.424	1.4325	1.4332	1.4339	1.4346	1.4353	1.4360	1.4367	1.4374	1.4381	1.4388
0.425	1.4395	1.4402	1.4409	1.4417	1.4424	1.4431	1.4438	1.4445	1.4452	1.4459
0.426	1.4466	1.4473	1.4481	1.4488	1.4495	1.4502	1.4509	1.4516	1.4524	1.4531
0.427	1.4538	1.4545	1.4553	1.4560	1.4567	1.4574	1.4581	1.4589	1.4596	1.4603
0.428	1.4611	1.4618	1.4625	1.4632	1.4640	1.4647	1.4654	1.4662	1.4669	1.4676
0.429	1.4684	1.4691	1.4699	1.4706	1.4713	1.4721	1.4728	1.4736	1.4743	1.4750
0.43	1.4758	1.4765	1.4773	1.4780	1.4788	1.4795	1.4803	1.4810	1.4818	1.4825
0.431	1.4833	1.4840	1.4848	1.4855	1.4863	1.4871	1.4878	1.4886	1.4893	1.4901
0.432	1.4909	1.4916	1.4924	1.4931	1.4939	1.4947	1.4954	1.4962	1.4970	1.4977
0.433	1.4985	1.4993	1.5001	1.5008	1.5016	1.5024	1.5032	1.5039	1.5047	1.5055

	0	1	2	3	4	5	6	7	8	9
0.434	1.5063	1.5070	1.5078	1.5086	1.5094	1.5102	1.5110	1.5117	1.5125	1.5133
0.435	1.5141	1.5149	1.5157	1.5165	1.5173	1.5181	1.5189	1.5196	1.5204	1.5212
0.436	1.5220	1.5228	1.5236	1.5244	1.5252	1.5260	1.5268	1.5276	1.5285	1.5293
0.437	1.5301	1.5309	1.5317	1.5325	1.5333	1.5341	1.5349	1.5358	1.5366	1.5374
0.438	1.5382	1.5390	1.5398	1.5407	1.5415	1.5423	1.5431	1.5440	1.5448	1.5456
0.439	1.5464	1.5473	1.5481	1.5489	1.5498	1.5506	1.5514	1.5523	1.5531	1.5539
0.44	1.5548	1.5556	1.5565	1.5573	1.5581	1.5590	1.5598	1.5607	1.5615	1.5624
0.441	1.5632	1.5641	1.5649	1.5658	1.5666	1.5675	1.5683	1.5692	1.5701	1.5709
0.442	1.5718	1.5726	1.5735	1.5744	1.5752	1.5761	1.5770	1.5779	1.5787	1.5796
0.443	1.5805	1.5813	1.5822	1.5831	1.5840	1.5849	1.5857	1.5866	1.5875	1.5884
0.444	1.5893	1.5902	1.5910	1.5919	1.5928	1.5937	1.5946	1.5955	1.5964	1.5973
0.445	1.5982	1.5991	1.6000	1.6009	1.6018	1.6027	1.6036	1.6045	1.6054	1.6063
0.446	1.6072	1.6082	1.6091	1.6100	1.6109	1.6118	1.6127	1.6137	1.6146	1.6155
0.447	1.6164	1.6174	1.6183	1.6192	1.6202	1.6211	1.6220	1.6229	1.6239	1.6248
0.448	1.6258	1.6267	1.6276	1.6286	1.6295	1.6305	1.6314	1.6324	1.6333	1.6343
0.449	1.6352	1.6362	1.6371	1.6381	1.6391	1.6400	1.6410	1.6420	1.6429	1.6439
0.45	1.6449	1.6458	1.6468	1.6478	1.6487	1.6497	1.6507	1.6517	1.6527	1.6536
0.451	1.6546	1.6556	1.6566	1.6576	1.6586	1.6596	1.6606	1.6616	1.6626	1.6636
0.452	1.6646	1.6656	1.6666	1.6676	1.6686	1.6696	1.6706	1.6716	1.6726	1.6736
0.453	1.6747	1.6757	1.6767	1.6777	1.6788	1.6798	1.6808	1.6818	1.6829	1.6839
0.454	1.6849	1.6860	1.6870	1.6881	1.6891	1.6901	1.6912	1.6922	1.6933	1.6943
0.455	1.6954	1.6965	1.6975	1.6986	1.6996	1.7007	1.7018	1.7028	1.7039	1.7050
0.456	1.7060	1.7071	1.7082	1.7093	1.7104	1.7114	1.7125	1.7136	1.7147	1.7158
0.457	1.7169	1.7180	1.7191	1.7202	1.7213	1.7224	1.7235	1.7246	1.7257	1.7268
0.458	1.7279	1.7290	1.7302	1.7313	1.7324	1.7335	1.7347	1.7358	1.7369	1.7381
0.459	1.7392	1.7403	1.7415	1.7426	1.7438	1.7449	1.7461	1.7472	1.7484	1.7495
0.46	1.7507	1.7519	1.7530	1.7542	1.7553	1.7565	1.7577	1.7589	1.7600	1.7612
0.461	1.7624	1.7636	1.7648	1.7660	1.7672	1.7684	1.7696	1.7708	1.7720	1.7732
0.462	1.7744	1.7756	1.7768	1.7780	1.7792	1.7805	1.7817	1.7829	1.7841	1.7854
0.463	1.7866	1.7878	1.7891	1.7903	1.7916	1.7928	1.7941	1.7953	1.7966	1.7979
0.464	1.7991	1.8004	1.8017	1.8029	1.8042	1.8055	1.8068	1.8080	1.8093	1.8106
0.465	1.8119	1.8132	1.8145	1.8158	1.8171	1.8184	1.8197	1.8210	1.8224	1.8237

	0	1	2	3	4	5	6	7	8	9
0.466	1.8250	1.8263	1.8277	1.8290	1.8303	1.8317	1.8330	1.8344	1.8357	1.8371
0.467	1.8384	1.8398	1.8411	1.8425	1.8439	1.8453	1.8466	1.8480	1.8494	1.8508
0.468	1.8522	1.8536	1.8550	1.8564	1.8578	1.8592	1.8606	1.8620	1.8634	1.8649
0.469	1.8663	1.8677	1.8692	1.8706	1.8720	1.8735	1.8749	1.8764	1.8779	1.8793
0.47	1.8808	1.8823	1.8837	1.8852	1.8867	1.8882	1.8897	1.8912	1.8927	1.8942
0.471	1.8957	1.8972	1.8987	1.9003	1.9018	1.9033	1.9048	1.9064	1.9079	1.9095
0.472	1.9110	1.9126	1.9142	1.9157	1.9173	1.9189	1.9205	1.9220	1.9236	1.9252
0.473	1.9268	1.9284	1.9301	1.9317	1.9333	1.9349	1.9366	1.9382	1.9398	1.9415
0.474	1.9431	1.9448	1.9465	1.9481	1.9498	1.9515	1.9532	1.9549	1.9566	1.9583
0.475	1.9600	1.9617	1.9634	1.9651	1.9669	1.9686	1.9703	1.9721	1.9738	1.9756
0.476	1.9774	1.9791	1.9809	1.9827	1.9845	1.9863	1.9881	1.9899	1.9917	1.9936
0.477	1.9954	1.9972	1.9991	2.0009	2.0028	2.0047	2.0065	2.0084	2.0103	2.0122
0.478	2.0141	2.0160	2.0179	2.0198	2.0218	2.0237	2.0257	2.0276	2.0296	2.0315
0.479	2.0335	2.0355	2.0375	2.0395	2.0415	2.0435	2.0456	2.0476	2.0496	2.0517
0.48	2.0537	2.0558	2.0579	2.0600	2.0621	2.0642	2.0663	2.0684	2.0706	2.0727
0.481	2.0748	2.0770	2.0792	2.0814	2.0836	2.0858	2.0880	2.0902	2.0924	2.0947
0.482	2.0969	2.0992	2.1015	2.1038	2.1060	2.1084	2.1107	2.1130	2.1153	2.1177
0.483	2.1201	2.1225	2.1248	2.1272	2.1297	2.1321	2.1345	2.1370	2.1394	2.1419
0.484	2.1444	2.1469	2.1494	2.1520	2.1545	2.1571	2.1596	2.1622	2.1648	2.1675
0.485	2.1701	2.1727	2.1754	2.1781	2.1808	2.1835	2.1862	2.1890	2.1917	2.1945
0.486	2.1973	2.2001	2.2029	2.2058	2.2086	2.2115	2.2144	2.2173	2.2203	2.2232
0.487	2.2262	2.2292	2.2322	2.2353	2.2383	2.2414	2.2445	2.2476	2.2508	2.2539
0.488	2.2571	2.2603	2.2636	2.2668	2.2701	2.2734	2.2768	2.2801	2.2835	2.2869
0.489	2.2904	2.2938	2.2973	2.3009	2.3044	2.3080	2.3116	2.3152	2.3189	2.3226
0.49	2.3263	2.3301	2.3339	2.3377	2.3416	2.3455	2.3495	2.3534	2.3575	2.3615
0.491	2.3656	2.3697	2.3739	2.3781	2.3824	2.3867	2.3911	2.3954	2.3999	2.4044
0.492	2.4089	2.4135	2.4181	2.4228	2.4276	2.4324	2.4372	2.4421	2.4471	2.4522
0.493	2.4573	2.4624	2.4677	2.4730	2.4783	2.4838	2.4893	2.4949	2.5006	2.5063
0.494	2.5121	2.5181	2.5241	2.5302	2.5364	2.5427	2.5491	2.5556	2.5622	2.5690
0.495	2.5758	2.5828	2.5899	2.5972	2.6045	2.6121	2.6197	2.6276	2.6356	2.6437
0.496	2.6521	2.6606	2.6693	2.6783	2.6874	2.6968	2.7065	2.7164	2.7266	2.7370
0.497	2.7478	2.7589	2.7703	2.7822	2.7944	2.8071	2.8202	2.8338	2.8480	2.8627
0.498	2.8782	2.8943	2.9113	2.9290	2.9479	2.9677	2.9889	3.0115	3.0357	3.0619
0.499	3.0902	3.1214	3.1560	3.1947	3.2390	3.2905	3.3528	3.4319	3.5402	3.7195

Tabella
della
distribuzione cumulativa χ^2 inversa



	+	+	+	+	+	ϕ	+	+	+	+
d	-	-	-	-	-	x	-	-	-	-

d = gradi di libertà

Tabella della distribuzione χ^2 inversa

	0.005	0.01	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	0.995	0.999
1	0.00	0.00	0.00	0.00	0.02	0.10	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75	20.51
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55	22.46
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.95	26.12
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.73	26.76	31.26
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00	45.31
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56	51.18
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93	52.62
26	11.16	12.20	13.84	15.38	17.29	20.84	25.34	30.43	35.56	38.89	41.92	45.64	48.29	54.05
27	11.81	12.88	14.57	16.15	18.11	21.75	26.34	31.53	36.74	40.11	43.19	46.96	49.65	55.48
28	12.46	13.56	15.31	16.93	18.94	22.66	27.34	32.62	37.92	41.34	44.46	48.28	50.99	56.89
29	13.12	14.26	16.05	17.71	19.77	23.57	28.34	33.71	39.09	42.56	45.72	49.59	52.34	58.30
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67	59.70
31	14.46	15.66	17.54	19.28	21.43	25.39	30.34	35.89	41.42	44.99	48.23	52.19	55.00	61.10
32	15.13	16.36	18.29	20.07	22.27	26.30	31.34	36.97	42.58	46.19	49.48	53.49	56.33	62.49
33	15.82	17.07	19.05	20.87	23.11	27.22	32.34	38.06	43.75	47.40	50.73	54.78	57.65	63.87
34	16.50	17.79	19.81	21.66	23.95	28.14	33.34	39.14	44.90	48.60	51.97	56.06	58.96	65.25
35	17.19	18.51	20.57	22.47	24.80	29.05	34.34	40.22	46.06	49.80	53.20	57.34	60.27	66.62

	0.005	0.01	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	0.995	0.999
36	17.89	19.23	21.34	23.27	25.64	29.97	35.34	41.30	47.21	51.00	54.44	58.62	61.58	67.98
37	18.59	19.96	22.11	24.07	26.49	30.89	36.34	42.38	48.36	52.19	55.67	59.89	62.88	69.35
38	19.29	20.69	22.88	24.88	27.34	31.81	37.34	43.46	49.51	53.38	56.90	61.16	64.18	70.70
39	20.00	21.43	23.65	25.70	28.20	32.74	38.34	44.54	50.66	54.57	58.12	62.43	65.48	72.06
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.81	55.76	59.34	63.69	66.77	73.40
41	21.42	22.91	25.21	27.33	29.91	34.58	40.34	46.69	52.95	56.94	60.56	64.95	68.05	74.74
42	22.14	23.65	26.00	28.14	30.77	35.51	41.34	47.77	54.09	58.12	61.78	66.21	69.34	76.08
43	22.86	24.40	26.79	28.96	31.63	36.44	42.34	48.84	55.23	59.30	62.99	67.46	70.62	77.42
44	23.58	25.15	27.57	29.79	32.49	37.36	43.34	49.91	56.37	60.48	64.20	68.71	71.89	78.75
45	24.31	25.90	28.37	30.61	33.35	38.29	44.34	50.98	57.51	61.66	65.41	69.96	73.17	80.08
46	25.04	26.66	29.16	31.44	34.22	39.22	45.34	52.06	58.64	62.83	66.62	71.20	74.44	81.40
47	25.77	27.42	29.96	32.27	35.08	40.15	46.34	53.13	59.77	64.00	67.82	72.44	75.70	82.72
48	26.51	28.18	30.75	33.10	35.95	41.08	47.34	54.20	60.91	65.17	69.02	73.68	76.97	84.04
49	27.25	28.94	31.55	33.93	36.82	42.01	48.33	55.27	62.04	66.34	70.22	74.92	78.23	85.35
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49	86.66
51	28.73	30.48	33.16	35.60	38.56	43.87	50.33	57.40	64.30	68.67	72.62	77.39	80.75	87.97
52	29.48	31.25	33.97	36.44	39.43	44.81	51.33	58.47	65.42	69.83	73.81	78.62	82.00	89.27
53	30.23	32.02	34.78	37.28	40.31	45.74	52.33	59.53	66.55	70.99	75.00	79.84	83.25	90.57
54	30.98	32.79	35.59	38.12	41.18	46.68	53.33	60.60	67.67	72.15	76.19	81.07	84.50	91.87
55	31.73	33.57	36.40	38.96	42.06	47.61	54.33	61.67	68.80	73.31	77.38	82.29	85.75	93.17
56	32.49	34.35	37.21	39.80	42.94	48.55	55.33	62.73	69.92	74.47	78.57	83.51	86.99	94.46
57	33.25	35.13	38.03	40.65	43.82	49.48	56.33	63.79	71.04	75.62	79.75	84.73	88.24	95.75
58	34.01	35.91	38.84	41.49	44.70	50.42	57.33	64.86	72.16	76.78	80.94	85.95	89.48	97.04
59	34.77	36.70	39.66	42.34	45.58	51.36	58.33	65.92	73.28	77.93	82.12	87.17	90.72	98.32
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33	66.98	74.40	79.08	83.30	88.38	91.95	99.61
61	36.30	38.27	41.30	44.04	47.34	53.23	60.33	68.04	75.51	80.23	84.48	89.59	93.19	100.89
62	37.07	39.06	42.13	44.89	48.23	54.17	61.33	69.10	76.63	81.38	85.65	90.80	94.42	102.17
63	37.84	39.86	42.95	45.74	49.11	55.11	62.33	70.16	77.75	82.53	86.83	92.01	95.65	103.44
64	38.61	40.65	43.78	46.59	50.00	56.05	63.33	71.23	78.86	83.68	88.00	93.22	96.88	104.72
65	39.38	41.44	44.60	47.45	50.88	56.99	64.33	72.28	79.97	84.82	89.18	94.42	98.10	105.99
66	40.16	42.24	45.43	48.31	51.77	57.93	65.33	73.34	81.09	85.96	90.35	95.63	99.33	107.26
67	40.93	43.04	46.26	49.16	52.66	58.87	66.33	74.40	82.20	87.11	91.52	96.83	100.55	108.53
68	41.71	43.84	47.09	50.02	53.55	59.81	67.33	75.46	83.31	88.25	92.69	98.03	101.78	109.79
69	42.49	44.64	47.92	50.88	54.44	60.76	68.33	76.52	84.42	89.39	93.86	99.23	103.00	111.06
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33	77.58	85.53	90.53	95.02	100.43	104.21	112.32

	0.005	0.01	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	0.995	0.999
71	44.06	46.25	49.59	52.60	56.22	62.64	70.33	78.63	86.64	91.67	96.19	101.62	105.43	113.58
72	44.84	47.05	50.43	53.46	57.11	63.58	71.33	79.69	87.74	92.81	97.35	102.82	106.65	114.83
73	45.63	47.86	51.26	54.33	58.01	64.53	72.33	80.75	88.85	93.95	98.52	104.01	107.86	116.09
74	46.42	48.67	52.10	55.19	58.90	65.47	73.33	81.80	89.96	95.08	99.68	105.20	109.07	117.35
75	47.21	49.48	52.94	56.05	59.79	66.42	74.33	82.86	91.06	96.22	100.84	106.39	110.29	118.60
76	48.00	50.29	53.78	56.92	60.69	67.36	75.33	83.91	92.17	97.35	102.00	107.58	111.50	119.85
77	48.79	51.10	54.62	57.79	61.59	68.31	76.33	84.97	93.27	98.48	103.16	108.77	112.70	121.10
78	49.58	51.91	55.47	58.65	62.48	69.25	77.33	86.02	94.37	99.62	104.32	109.96	113.91	122.35
79	50.38	52.72	56.31	59.52	63.38	70.20	78.33	87.08	95.48	100.75	105.47	111.14	115.12	123.59
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33	88.13	96.58	101.88	106.63	112.33	116.32	124.84
81	51.97	54.36	58.00	61.26	65.18	72.09	80.33	89.18	97.68	103.01	107.78	113.51	117.52	126.08
82	52.77	55.17	58.84	62.13	66.08	73.04	81.33	90.24	98.78	104.14	108.94	114.69	118.73	127.32
83	53.57	55.99	59.69	63.00	66.98	73.99	82.33	91.29	99.88	105.27	110.09	115.88	119.93	128.56
84	54.37	56.81	60.54	63.88	67.88	74.93	83.33	92.34	100.98	106.39	111.24	117.06	121.13	129.80
85	55.17	57.63	61.39	64.75	68.78	75.88	84.33	93.39	102.08	107.52	112.39	118.24	122.32	131.04
86	55.97	58.46	62.24	65.62	69.68	76.83	85.33	94.45	103.18	108.65	113.54	119.41	123.52	132.28
87	56.78	59.28	63.09	66.50	70.58	77.78	86.33	95.50	104.28	109.77	114.69	120.59	124.72	133.51
88	57.58	60.10	63.94	67.37	71.48	78.73	87.33	96.55	105.37	110.90	115.84	121.77	125.91	134.75
89	58.39	60.93	64.79	68.25	72.39	79.68	88.33	97.60	106.47	112.02	116.99	122.94	127.11	135.98
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33	98.65	107.57	113.15	118.14	124.12	128.30	137.21
91	60.00	62.58	66.50	70.00	74.20	81.57	90.33	99.70	108.66	114.27	119.28	125.29	129.49	138.44
92	60.81	63.41	67.36	70.88	75.10	82.52	91.33	100.75	109.76	115.39	120.43	126.46	130.68	139.67
93	61.63	64.24	68.21	71.76	76.01	83.47	92.33	101.80	110.85	116.51	121.57	127.63	131.87	140.89
94	62.44	65.07	69.07	72.64	76.91	84.42	93.33	102.85	111.94	117.63	122.72	128.80	133.06	142.12
95	63.25	65.90	69.92	73.52	77.82	85.38	94.33	103.90	113.04	118.75	123.86	129.97	134.25	143.34
96	64.06	66.73	70.78	74.40	78.73	86.33	95.33	104.95	114.13	119.87	125.00	131.14	135.43	144.57
97	64.88	67.56	71.64	75.28	79.63	87.28	96.33	106.00	115.22	120.99	126.14	132.31	136.62	145.79
98	65.69	68.40	72.50	76.16	80.54	88.23	97.33	107.05	116.32	122.11	127.28	133.48	137.80	147.01
99	66.51	69.23	73.36	77.05	81.45	89.18	98.33	108.09	117.41	123.23	128.42	134.64	138.99	148.23
100	67.33	70.06	74.22	77.93	82.36	90.13	99.33	109.14	118.50	124.34	129.56	135.81	140.17	149.45

Elenco delle figure

1.1	Grafici di funzioni di due variabili	4
1.2	Curve di livello delle superfici in figura 1.1	5
1.3	Sezioni, per y fissato, dei grafici di figura 1.1	6
1.4	Sezioni, per x fissato, dei grafici di figura 1.1	7
1.5	Curve di livello e segno di f	11
1.6	Il grafico di una funzione derivabile parzialmente, non continua.	12
1.7	Definizione di derivata direzionale	14
1.8	Curve di livello e gradiente	17
1.9	Principio dei moltiplicatori di Lagrange	24
2.1		27
2.2		29
2.3	Singolo elemento di volume	29
2.4	$S(x)$	30
2.5	Dominio normale rispetto all'asse x	31
2.6	Dominio normale rispetto all'asse x	32
2.7	Volume ed area	34
2.8	Cambiamento di variabili lineare	35
2.9	Cambiamento di variabili in coordinate polari	36
3.1		40
3.2		41
6.1	Lo spazio $\mathcal{U}$ degli eventi nel caso del lancio di due dadi	52
6.2	Istogramma relativo al lancio di due dadi, (supposti distinti)	53
6.3	Eventi che forniscono lo stesso valore per ξ , (congiunti da un segmento)	62
6.4	Distribuzione di probabilità della variabile aleatoria che restituisce il punteggio ottenuto lanciando due dadi.	65
6.5	Istogramma relativo alla variabile aleatoria che restituisce il punteggio ottenuto lanciando due dadi.	66

10.1		117
10.2		118
10.3		119
10.4		119
10.5		120
11.1	Grafici di funzioni di due variabili	123

Indice

Capitolo 1. FUNZIONI DI DUE VARIABILI	3
1. La struttura di $\mathbb{R}^2$.	6
2. Limiti e continuità per le funzioni di 2 variabili.	9
3. Derivabilità e differenziabilità per funzioni di 2 variabili.	11
4. Derivate del secondo ordine: forma quadratica Hessiana.	18
5. Massimi e minimi per le funzioni di 2 variabili.	20
6. Massimi e minimi vincolati.	21
Capitolo 2. INTEGRAZIONE PER LE FUNZIONI DI DUE VARIABILI	27
1. Definizione di integrale doppio.	27
2. Formule di riduzione per gli integrali doppi.	29
3. Cambiamento di variabili negli integrali doppi	33
Capitolo 3. INTEGRAZIONE DI FUNZIONI DI TRE VARIABILI	39
1. Definizione di integrale triplo.	39
2. Formule di riduzione per gli integrali tripli.	39
3. Cambiamento di variabili per gli integrali tripli.	40
Capitolo 4. INTEGRALI MULTIPLI IMPROPRI	43
1. Qualche esempio	44
Capitolo 5. LA SOMMA DI INFINITI TERMINI: LE SERIE.	47
Capitolo 6. ELEMENTI DI PROBABILITÀ E STATISTICA.	51
1. Spazi di probabilità	51
2. Qualche applicazione della regola di Bayes.	56
3. Qualche richiamo di calcolo combinatorio.	59
4. Le variabili aleatorie	62
5. Variabili aleatorie continue	66
6. Distribuzioni di probabilità doppie	71
7. Qualche proprietà di valor medio e varianza	72
8. La disuguaglianza di Tchebichev e la legge dei grandi numeri	74
9. Normalizzazione di una variabile aleatoria.	76
Capitolo 7. QUALCHE DISTRIBUZIONE DI PROBABILITÀ	79
1. La distribuzione uniforme	79
2. La distribuzione binomiale di Bernoulli	80
3. La distribuzione di Poisson	81

4. La distribuzione gaussiana	82
5. Due distribuzioni discrete	86
6. Le distribuzioni legate ai test statistici.	87
7. La distribuzione γ .	89
8. La distribuzione β .	89
9. Variabili casuali con distribuzione assegnata.	89
10. I test statistici	91
11. Risultati sulle distribuzioni campionarie	95
Capitolo 8. IL TEOREMA DEL LIMITE CENTRALE	101
1. Un esempio di applicazione	102
Capitolo 9. REGRESSIONE LINEARE: LA RETTA DEI MINIMI QUADRATI	105
Capitolo 10. ANALISI DEI COMPONENTI PRINCIPALI.	113
1. Forme quadratiche ed autovalori.	114
2. L'applicazione all'analisi delle componenti principali.	116
Capitolo 11. Appendice	121
1. LE FORMULE DI WALLIS E DI STIRLING-DE MOIVRE.	121
Capitolo 12. QUALCHE ARGOMENTO DI PROBABILITÀ E STATISTICA.	127
1. Il Lancio di un Dado	127
2. Lancio di due Dadi	128
3. Lancio ripetuto di due dadi	132
4. Distribuzione Binomiale di Bernoulli	133
5. Applicazioni della distribuzione Binomiale	136
6. Probabilità condizionata. Teorema di Bayes	139
7. Distribuzioni di Probabilità	146
8. Uso delle tavole per i valori della CDF Normale Standard	148
9. Calcolo dei valori della CDF Gaussiana Inversa	155
10. Approssimazione della Distribuzione Binomiale mediante la Distribuzione Normale	158
11. Stima di Probabilità	159
12. Test Statistici	165
Capitolo 13. Appendice	179
1. Il Problema del compleanno	179
2. Il Paradosso di Bertrand	180
3. TRIANGOLO DI TARTAGLIA E BINOMIO DI NEWTON	182
Elenco delle figure	205
Indice analitico	209

Indice analitico

binomio di Newton, 183

triangolo di Tartaglia, 183