

Metodi e Modelli Matematici

O.Caligaris - P.Oliva

Facoltà di Ingegneria - Università di Genova
Polo di Savona

Variabili Aleatorie e Distribuzioni di Probabilità.

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

Ad esempio, quando si considera il problema di scegliere un numero a caso compreso tra 0 ed 1.

La probabilità di estrarre 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.

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

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

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

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

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

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

mentre è naturale definire

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

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 (PDF) di una variabile aleatoria se

- $$\varphi(t) \geq 0 \quad \text{per ogni } t \in \mathbb{R}$$

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

In tal caso si ha

$$\mathcal{P}(\xi \leq x) = \int_{-\infty}^x \varphi(t) dt \quad , \quad \mathcal{P}(x_0 \leq \xi \leq x_1) = \int_{x_0}^{x_1} \varphi(t) dt$$

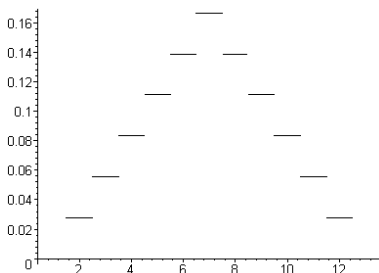
La funzione

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

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

Ad ogni variabile aleatoria discreta 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



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

Definizione

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

Definizione

- 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

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

Se ξ è una variabile aleatoria continua e se $f : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione derivabile ed invertibile che supponiamo crescente per comodità, possiamo considerare la variabile aleatoria $f(\xi)$ e possiamo calcolare che

$$\mathcal{P}(x_0 \leq f(\xi) \leq x_1) = \mathcal{P}(f^{-1}(x_0) \leq \xi \leq f^{-1}(x_1)) =$$

$$\int_{f^{-1}(x_0)}^{f^{-1}(x_1)} \varphi(t) dt = \int_{x_0}^{x_1} \frac{\varphi(f^{-1}(s))}{f'(f^{-1}(s))} ds$$

per cui la sua funzione distribuzione di probabilità risulta definita da

$$\psi(s) = \frac{\varphi(f^{-1}(s))}{f'(f^{-1}(s))}$$

In tal modo si ha

$$\begin{aligned} E(f(\xi)) &= \int_{-\infty}^{+\infty} s \frac{\varphi(f^{-1}(s))}{f'(f^{-1}(s))} ds = \\ &= \int_{-\infty}^{+\infty} f(t) \frac{\varphi(t)}{f'(t)} f'(t) dt = \int_{-\infty}^{+\infty} f(t) \varphi(t) dt = \mu \end{aligned}$$

ed inoltre

$$\begin{aligned} \sigma^2(f(\xi)) &= \int_{-\infty}^{+\infty} (s - \mu)^2 \frac{\varphi(f^{-1}(s))}{f'(f^{-1}(s))} ds = \\ &= \int_{-\infty}^{+\infty} (f(t) - \mu)^2 \frac{\varphi(t)}{f'(t)} f'(t) dt = \int_{-\infty}^{+\infty} (f(t) - \mu)^2 \varphi(t) dt \end{aligned}$$

Definizione

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

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

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

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

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

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

Ed è utile ricordare ancora che

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

Si ha inoltre:

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

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

$$\mu_k = \sum_{i=0}^k \binom{k}{i} \mu'_i \mu^{k-i} \quad (1)$$

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 1 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 $\mu_1 = \mu'_1 = \mu$).

La funzione generatrice dei momenti si rivela molto comoda per il calcolo dei momenti di una variabile aleatoria.

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

$$M_\xi(t) = \int_{-\infty}^{+\infty} \sum_{k=0}^{+\infty} \frac{(st)^k}{k!} \varphi(s) ds = \sum_{k=0}^{+\infty} \int_{-\infty}^{+\infty} s^k \varphi(s) ds \frac{t^k}{k!}$$

per cui

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

e quindi

$$\mu'_k = \frac{d^k}{dt^k} M_\xi(0)$$

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

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

se ne ricava pertanto che

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

La 2 è nota come **disuguaglianza di Tchebichev** e ne possiamo trarre una interessante conseguenza: per $\varepsilon = k\sigma$ otteniamo che

$$\mathcal{P}(|\xi - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad (3)$$

Pertanto

$$\mathcal{P}(|\xi - \mu| < k\sigma) = 1 - \mathcal{P}(|\xi - \mu| \geq k\sigma) \geq 1 - \frac{1}{k^2} \quad (4)$$

Se ora consideriamo la seguente tabella

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

Si vede 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.

Una delle conseguenze della disuguaglianza di Tchebichev prende il nome di "Legge dei Grandi Numeri" e si ricava come segue.

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

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

ha media

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

e varianza

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

inoltre vale il seguente teorema

Teorema

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

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

Allora

$$\mathcal{P}(|S_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0 \quad (5)$$

per $n \rightarrow +\infty$

La 6 è nota con il nome di "**Legge Debole dei Grandi Numeri**" ed esprime un concetto in base al quale la media di n uscite di una variabile aleatoria differisce dalla media della variabile aleatoria per una quantità infinitesima con n .

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

È possibile anche dimostrare, ma la dimostrazione è più complessa, la "**Legge Forte dei Grandi Numeri**" che asserisce che

Teorema

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

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

Allora

$$\mathcal{P}(\lim_n S_n = \mu) = 1 \quad (6)$$

In entrambi i casi il concetto espresso è che la media di S_n converge alla media μ la differenza risiede nel modo in cui tale convergenza avviene e nelle proprietà che tale convergenza consente di trasferire sul limite.

Piú precisamente la legge debole dei grandi numeri afferma che la successione di variabili aleatorie S_n converge a μ in probabilità, mentre la legge forte dei grandi numeri garantisce che S_n converge a μ quasi certamente.

Usando la terminologia derivante dalla teoria della misura, cui la teoria della probabilità astratta sostanzialmente si sovrappone, la formulazione debole parla di convergenza in misura, mentre la formulazione forte parla di convergenza puntuale quasi ovunque. È noto che, essendo lo spazio di probabilità di misura finita (uguale ad 1), una successione quasi ovunque convergente è anche convergente in misura, e che da una successione convergente in misura si può estrarre una sottosuccessione quasi ovunque convergente.

Consideriamo due variabili aleatorie discrete indipendenti ξ, η aventi PDF rispettivamente f e g , e la variabile aleatoria ζ che restituisce la somma delle due

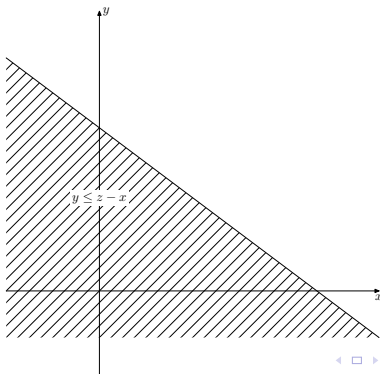
$$\zeta = \xi + \eta$$

Possiamo trovare la densità di probabilità della variabile ζ osservando che

$$\begin{aligned} h(\gamma) &= \mathcal{P}(\zeta = \gamma) = \\ &= \sum_{\alpha} \mathcal{P}(\xi + \eta = \gamma | \xi = \alpha) \mathcal{P}(\xi = \alpha) = \sum_{\alpha} \mathcal{P}(\eta = \gamma - \alpha) \mathcal{P}(\xi = \alpha) = \\ &\quad \sum_{\alpha} f(\gamma - \alpha) g(\alpha) \end{aligned}$$

Nel caso in cui ξ e η siano variabili aleatorie continue indipendenti avremo, come si vede dalla figura 1,

$$\begin{aligned}\mathcal{P}(\zeta \leq z) &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{z-x} g(y) dy \right) f(x) dx = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^z g(s-x) f(x) ds dx = \int_{-\infty}^z \left(\int_{-\infty}^{+\infty} g(s-x) f(x) dx \right) ds\end{aligned}$$



Possiamo allora calcolare che

$$\begin{aligned}
 E(\zeta) &= E(\xi + \eta) = \int_{-\infty}^{+\infty} s \left(\int_{-\infty}^{+\infty} g(s-x)f(x)dx \right) ds = \\
 &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} sg(s-x)f(x)ds \right) dx = \\
 &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} (x+t)g(t)f(x)dx \right) dt = \\
 &= \int_{-\infty}^{+\infty} g(t) \left(\int_{-\infty}^{+\infty} xf(x)dx \right) dt + \int_{-\infty}^{+\infty} f(x) \left(\int_{-\infty}^{+\infty} tg(t)dt \right) dx = \\
 &= E(\xi) \int_{-\infty}^{+\infty} g(t)dt + E(\eta) \int_{-\infty}^{+\infty} f(x)dx = E(\xi) + E(\eta)
 \end{aligned}$$

Siano ξ e η due variabili aleatorie indipendenti le cui PDF sono f ed g , rispettivamente, e sia

$$\zeta = \xi\eta$$

Avremo che

$$\mathcal{P}(\xi\eta \leq \alpha) = \int_A f(x)g(y)dxdy$$

dove

$$A = \{(x, y) \in \mathbb{R}^2 : xy \leq \alpha\}$$

è l'insieme tratteggiato nella figura 2;

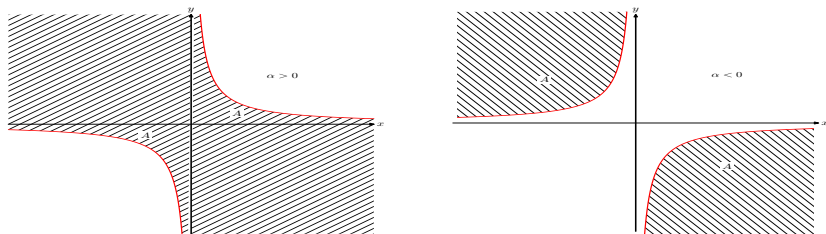


Figure :

Pertanto

$$\begin{aligned}\mathcal{P}(\xi\eta \leq \alpha) &= \int_A f(x)g(y)dx dy = \\ &= \int_{-\infty}^0 \int_{\frac{\alpha}{x}}^{+\infty} f(x)g(y)dy dx + \int_0^{+\infty} \int_{-\infty}^{\frac{\alpha}{x}} f(x)g(y)dy dx =\end{aligned}$$

posto $y = \frac{s}{x}$ da cui $dy = \frac{ds}{x}$

$$\begin{aligned}&= \int_{-\infty}^0 \int_{\alpha}^{-\infty} f(x)g\left(\frac{s}{x}\right) \frac{ds}{x} dx + \int_0^{+\infty} \int_{-\infty}^{\alpha} f(x)g\left(\frac{s}{x}\right) \frac{ds}{x} dx = \\ &= \int_{-\infty}^0 \left(- \int_{-\infty}^{\alpha} f(x)g\left(\frac{s}{x}\right) \frac{ds}{x} \right) dx + \int_0^{+\infty} \left(\int_{-\infty}^{\alpha} f(x)g\left(\frac{s}{x}\right) \frac{ds}{x} \right) dx = \\ &\hspace{25em} \int_{-\infty}^{\alpha} \left(\int_{-\infty}^{+\infty} f(x)g\left(\frac{s}{x}\right) \frac{dx}{|x|} \right) ds\end{aligned}$$

dal che si deduce che la PDF di ζ è data da

$$\varphi(s) = \int_{-\infty}^{+\infty} f(x)g\left(\frac{s}{x}\right) \frac{dx}{|x|}$$

Possiamo inoltre calcolare media e varianza di ζ come segue.

$$\begin{aligned}\mu_{\zeta} &= \int_{-\infty}^{+\infty} s \left(\int_{-\infty}^{+\infty} f(x) g\left(\frac{s}{x}\right) \frac{dx}{|x|} \right) ds = \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x) g\left(\frac{s}{x}\right) \frac{s}{|x|} ds \right) dx = \\ &= \int_0^{+\infty} \left(\int_{-\infty}^{+\infty} f(x) g\left(\frac{s}{x}\right) \frac{s}{x} ds \right) dx - \int_{-\infty}^0 \left(\int_{-\infty}^{+\infty} f(x) g\left(\frac{s}{x}\right) \frac{s}{x} ds \right) dx =\end{aligned}$$

posto $t = \frac{s}{x}$ da cui $dt = \frac{ds}{x}$

$$\begin{aligned}&= \int_0^{+\infty} \left(\int_{-\infty}^{+\infty} f(x) g(t) t x dt \right) dx + \int_{-\infty}^0 \left(\int_{-\infty}^{+\infty} f(x) g(t) t x dt \right) dx = \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x) g(t) t x dt \right) dx = \mu_{\xi} \mu_{\eta}\end{aligned}$$

Per quanto concerne la varianza avremo che

$$\sigma_{\zeta}^2 = \int_{-\infty}^{+\infty} s^2 \left(\int_{-\infty}^{+\infty} f(x)g\left(\frac{s}{x}\right) \frac{dx}{|x|} \right) ds - \mu_{\zeta}^2$$

e

$$\begin{aligned} \int_{-\infty}^{+\infty} s^2 \left(\int_{-\infty}^{+\infty} f(x)g\left(\frac{s}{x}\right) \frac{dx}{|x|} \right) ds &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x)g\left(\frac{s}{x}\right) \frac{s^2}{|x|} ds \right) dx = \\ &= \int_0^{+\infty} \left(\int_{-\infty}^{+\infty} f(x)g\left(\frac{s}{x}\right) \frac{s^2}{x} ds \right) dx - \int_{-\infty}^0 \left(\int_{-\infty}^{+\infty} f(x)g\left(\frac{s}{x}\right) \frac{s^2}{x} ds \right) dx = \end{aligned}$$

posto $t = \frac{s}{x}$ da cui $dt = \frac{ds}{x}$

$$\begin{aligned} &= \int_0^{+\infty} \left(\int_{-\infty}^{+\infty} f(x)g(t) \frac{x^2 t^2}{x} x dt \right) dx + \int_{-\infty}^0 \left(\int_{-\infty}^{+\infty} f(x)g(t) \frac{x^2 t^2}{x} x dt \right) dx = \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x)g(t) x^2 t^2 dt \right) dx = (\sigma_{\xi}^2 + \mu_{\xi}^2)(\sigma_{\eta}^2 + \mu_{\eta}^2) \end{aligned}$$

Possiamo allora concludere che

$$\sigma_{\zeta}^2 + \mu_{\zeta}^2 = (\sigma_{\xi}^2 + \mu_{\xi}^2)(\sigma_{\eta}^2 + \mu_{\eta}^2)$$

Se ξ è una variabile aleatoria e $\alpha \in \mathbb{R}_+$ allora

$$\mathcal{P}(\alpha\xi \leq x) = \mathcal{P}\left(\xi \leq \frac{x}{\alpha}\right) = \int_{-\infty}^{\frac{x}{\alpha}} f(t)dt = \int_{-\infty}^x \frac{1}{\alpha} f\left(\frac{s}{\alpha}\right) ds$$

D'altro canto se $\alpha \in \mathbb{R}_-$

$$\mathcal{P}(\alpha\xi \leq x) = \mathcal{P}\left(\xi \geq \frac{x}{\alpha}\right) = \int_{\frac{x}{\alpha}}^{+\infty} f(t)dt = - \int_{-\infty}^x \frac{1}{\alpha} f\left(\frac{s}{\alpha}\right) ds$$

per cui la PDF di $\alpha\xi$ è data da

$$g(s) = \frac{1}{|\alpha|} f\left(\frac{s}{\alpha}\right)$$

Siano η e ξ due variabili aleatorie indipendenti, ξ positiva, le cui PDF sono g ed f , rispettivamente, e sia

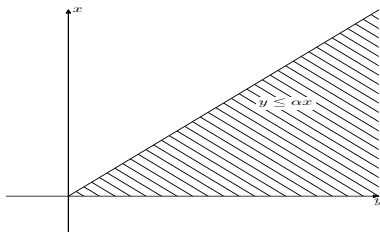
$$\zeta = \frac{\eta}{\xi}$$

Avremo che

$$\mathcal{P}\left(\frac{\eta}{\xi} \leq \alpha\right) = \mathcal{P}(\eta \leq \alpha\xi) = \int_A f(x)g(y)dx dy$$

dove

$$A = \{(x, y) \in \mathbb{R}^2 : y \leq \alpha x\}$$



Pertanto

$$\begin{aligned}\mathcal{P}\left(\frac{\eta}{\xi} \leq \alpha\right) &= \int_A f(x)g(y)dx dy = \int_0^{+\infty} \int_0^{\alpha x} f(x)g(y)dy dx = \\ &\text{posto } y = tx \text{ da cui } dy = xdt \\ &= \int_0^{+\infty} \int_0^{\alpha} f(x)g(tx)xdt dx = \int_0^{\alpha} \left(\int_0^{+\infty} xf(x)g(tx)dx \right) dt\end{aligned}$$

dal che si deduce che la PDF di ζ è data da

$$\varphi(t) = \int_0^{+\infty} xf(x)g(tx)dx$$

Se ξ è una variabile aleatoria e $n \in \mathbb{R}_+$ allora

$$\mathcal{P}\left(\frac{\xi}{n} \leq \alpha\right) = \mathcal{P}(\xi \leq \alpha n) = \int_{-\infty}^{\alpha n} f(x) dx = \int_{-\infty}^{\alpha} n f(nt) dt$$

per cui la PDF di $\frac{\xi}{n}$ è data da

$$g(t) = n f(nt)$$

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

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

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

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.

Possiamo giustificare la definizione osservando che:

$$\begin{aligned}
 \mathcal{P}(\xi \leq x, \eta = y) &\approx \mathcal{P}(\xi \leq x, y \leq \eta \leq y + k) = \\
 &= \frac{\int_{-\infty}^x \int_y^{y+k} f(t, s) ds dt}{\int_{-\infty}^{+\infty} \int_y^{y+k} f(t, s) ds dt} = \\
 &= \frac{\int_{-\infty}^x \int_y^{y+k} f(t, s) ds dt}{\int_y^{y+k} \int_{-\infty}^{+\infty} f(t, s) dt ds} = \frac{\int_{-\infty}^x \int_y^{y+k} f(t, s) ds dt}{\int_y^{y+k} \psi(s) ds} = \\
 &\approx \frac{\int_{-\infty}^x f(t, y) k dt}{\psi(y) k} = \int_{-\infty}^x \frac{f(t, y) k}{\psi(y) k} dt = \int_{-\infty}^x \frac{f(t, y)}{\psi(y)} dt
 \end{aligned}$$

È la distribuzione di una variabile aleatoria che restituisce un valore scelto in un intervallo $[a, b]$ con il criterio di equiprobabilità.

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

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

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

La distribuzione triangolare è utile per definire una variabile aleatoria che assuma valori compresi tra a e b ed abbia una moda c .

$$\phi(t) = \begin{cases} 0 & t < a \\ \frac{2(t-a)}{(b-a)(c-a)} & a \leq t < c \\ \frac{2(b-t)}{(b-a)(b-c)} & c < t \leq b \\ 0 & t > b \end{cases}$$

il valor medio e la varianza sono dati da

$$\mu = \frac{a + b + c}{3} \quad \sigma^2 = \frac{a^2 + b^2 + c^2 - bc - ab - ac}{18}$$

e la funzione generatrice dei momenti è

$$M_{\xi}(t) = \frac{e^{ta}(b-c) - e^{tc}(b-a) + e^{tb}(c-a)}{t^2(b-a)(c-a)(b-c)}$$

Consideriamo una prova con probabilità di successo p ;
ripetiamola indefinitamente sotto l'ipotesi che

- p rimane costante:
- l'esito della prova non dipende dalle prove precedentemente effettuate.

Ad esempio possiamo considerare un tiratore che ha la capacità di colpire il bersaglio con probabilità p ad ogni tiro o una lampada che può guastarsi con probabilità p ad ogni accensione.

Sia ξ la variabile aleatoria che restituisce il numero del primo tentativo in cui la prova ha successo. Avremo che

$$\mathcal{P}(\xi = k) = (1 - p)^{k-1}p$$

in quanto la prova ha avuto esito negativo (la probabilità di insuccesso è $1 - p$) per $k - 1$ volte ed ha avuto successo la k -esima volta.

Pertanto la distribuzione di probabilità della variabile aleatoria ξ è definita da

$$\varphi(k) = \mathcal{P}(\xi = k) = (1 - p)^{k-1}p$$

e si verifica subito che

$$\sum_1^{+\infty} p(1 - p)^{k-1} = p \sum_0^{+\infty} (1 - p)^k = p \frac{1}{1 - (1 - p)} = 1$$

Inoltre

$$\begin{aligned} \mu &= \sum_1^{+\infty} kp(1 - p)^{k-1} = p \sum_1^{+\infty} k(1 - p)^{k-1} = \\ &= p \sum_1^{+\infty} -\frac{d}{dp}(1 - p)^k = -p \frac{d}{dp} \sum_1^{+\infty} (1 - p)^k = -p \left(-\frac{1}{p^2} \right) = \frac{1}{p} \end{aligned}$$

mentre

$$\begin{aligned}
 \sum_1^{+\infty} k^2 p(1-p)^{k-1} &= \sum_1^{+\infty} (k^2 + k)p(1-p)^{k-1} - \sum_1^{+\infty} kp(1-p)^{k-1} = \\
 &= p \sum_1^{+\infty} \frac{d^2}{dp^2} (1-p)^{k+1} - \mu = p \frac{d^2}{dp^2} \sum_1^{+\infty} (1-p)^{k+1} - \mu = \\
 &= p \frac{d^2}{dp^2} \left(\frac{1}{p} - 2 - p \right) - \mu = \frac{2}{p^2} - \mu
 \end{aligned}$$

per cui

$$\sigma^2 = \sum_1^{+\infty} k^2 p(1-p)^{k-1} - \mu^2 = \frac{2}{p^2} - \mu - \mu^2 = \frac{1-p}{p^2}$$

Per quanto riguarda la funzione generatrice dei momenti

$$\begin{aligned} M_{\xi}(t) = E(e^{t\xi}) &= \sum_{k=1}^{+\infty} e^{tk} p(1-p)^{k-1} = \frac{p}{1-p} \sum_{k=1}^{+\infty} (e^t(1-p))^k = \\ &= \frac{p}{1-p} \left(\frac{e^t(1-p)}{1 - e^t(1-p)} \right) = \frac{pe^t}{1 - e^t(1-p)} \end{aligned}$$

Definizione

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

$$\mathcal{P}(\xi = k) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1 - p)^{n-k}$$

Per giustificare la formula precedente 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 mentre il valore 0 compare 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 possono ottenere, ad esempio, k successi su n prove 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	0	1	1	0	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

$$\binom{n}{k}$$

(Ciascuna combinazione è individuata dalla sequenza dei k numeri, compresi tra 1 ed n , che indicano la posizione dei successi.)

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

$$(1 - p)^2 \cdot p + (0 - p)^2 q = q^2 p + p^2 q = pq(p + q) = pq$$

e su n esperimenti per le proprietà della varianza avremo

$$\sigma^2 = E((\xi - \mu)^2) = npq$$

e

$$\sigma = \sqrt{npq}$$

Alternativamente possiamo calcolare la media e la varianza di una variabile aleatoria Bernoulliana ξ usando direttamente la definizione:

$$\begin{aligned}
 \mu = E(\xi) &= \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^n k \binom{n}{k} p^k q^{n-k} = \\
 &= \sum_{k=1}^n k \frac{n(n-1) \cdots (n-(k-1))}{k!} p^k q^{n-k} = \\
 &= np \sum_{k=1}^n \frac{(n-1)(n-2) \cdots (n-(k-1))}{(k-1)!} p^{k-1} q^{n-1-(k-1)} = \\
 &= np \sum_{k=0}^{n-1} \frac{(n-1)(n-2) \cdots (n-1-(k-1))}{k!} p^k q^{n-1-k} = \\
 &= np(p+q)^{n-1} = np
 \end{aligned}$$

$$\begin{aligned}
\sigma^2 &= E((\xi - \mu)^2) = \sum_{k=0}^n (k - np)^2 \binom{n}{k} p^k q^{n-k} = \\
&= \sum_{k=0}^n k^2 \binom{n}{k} p^k q^{n-k} - (np)^2 = \sum_{k=1}^n k^2 \binom{n}{k} p^k q^{n-k} - (np)^2 = \\
&= np \sum_{k=1}^n k \frac{(n-1)(n-2) \cdots (n-(k-1))}{(k-1)!} p^{k-1} q^{n-1-(k-1)} - (np)^2 = \\
&= np \sum_{k=0}^{n-1} (k+1) \frac{(n-1)(n-2) \cdots (n-1-(k-1))}{k!} p^k q^{n-1-k} - (np)^2 = \\
&= np \left(\sum_{k=0}^{n-1} k \binom{n-1}{k} p^k q^{n-1-k} + \sum_{k=0}^{n-1} \binom{n-1}{k} p^k q^{n-1-k} \right) - (np)^2 = \\
&= np((n-1)p + 1) - (np)^2 = np(np + q) - (np)^2 = npq
\end{aligned}$$

Per calcolare la funzione generatrice dei momenti possiamo procedere come segue

$$\begin{aligned}
 M_{\xi}(t) &= E(e^{t\xi}) = \sum_{k=0}^n e^{tk} \binom{n}{k} p^k q^{n-k} = \\
 &= \sum_{k=0}^n \binom{n}{k} (pe^t)^k q^{n-k} = (pe^t + q)^n
 \end{aligned}$$

la funzione densità di probabilità (PDF) di una variabile aleatoria di Bernoulli ξ è definita da

$$P(\xi = k) = \binom{n}{k} p^k q^{n-k}$$

mentre la funzione di distribuzione cumulativa (CDF) è

$$F(\nu) = \sum_{k=0}^{\nu} P(\xi = k) = \sum_{k=0}^{\nu} \binom{n}{k} p^k q^{n-k}$$

Consideriamo un esperimento bernoulliano, consideriamo cioè una serie di prove ripetute con due soli possibili esiti: successo con probabilità p ed insuccesso con probabilità q .

Consideriamo la variabile aleatoria che restituisce il minimo numero ξ di tentativi necessari per ottenere r successi.

Possiamo allora vedere che la probabilità $\mathcal{P}(\xi = k)$ che si ottengano r successi al tentativo k si può calcolare considerando che

- al tentativo k si è verificato un successo (che ha probabilità p)
- nelle precedenti $k - 1$ prove si sono verificati $r - 1$ successi e $k - 1 - (r - 1)$ insuccessi (con probabilità $\binom{k-1}{r-1} p^{r-1} q^{k-r}$)

Pertanto

$$\mathcal{P}(\xi = k) = p \binom{k-1}{r-1} p^{r-1} q^{k-r} = \binom{k-1}{r-1} p^r q^{k-r}$$

definisce la funzione densità di probabilità della distribuzione di Pascal.

Talvolta si considera, in luogo di ξ , la variabile η che restituisce il numero di fallimenti che precedono il successo r -esimo. In tal caso si ha $h = k - r$ e

$$\mathcal{P}(\eta = y) = \binom{h+r-1}{r-1} p^r q^h$$

Possiamo calcolare che

$$\mu_{\xi} = \frac{r}{p} \quad , \quad \sigma_{\xi}^2 = \frac{rq}{p^2} \quad , \quad M_{\xi}(t) = \frac{pe^t}{(1 - qe^t)^r}$$

e

$$\mu_{\eta} = \frac{r(1-p)}{p} \quad , \quad \sigma_{\eta}^2 = \frac{r(1-p)}{p^2}$$

A titolo di esempio vediamo come è possibile calcolare la media μ_ξ e la varianza σ_ξ ;

$$\begin{aligned}
 \mu_\xi &= \sum_{k=r}^{+\infty} kp \binom{k-1}{r-1} p^{r-1} q^{k-r} = \sum_{k=r}^{+\infty} p^r \frac{k(k-1)\dots(k-r+1)}{(r-1)!} q^{k-r} = \\
 &= \frac{p^r}{(r-1)!} \sum_{k=r}^{+\infty} k(k-1)\dots(k-r+1) q^{k-r} = \\
 &= \frac{p^r}{(r-1)!} \sum_{k=r}^{+\infty} \frac{d^r}{dq^r} q^k = \frac{p^r}{(r-1)!} \frac{d^r}{dq^r} \sum_{k=r}^{+\infty} q^k = \\
 &= \frac{p^r}{(r-1)!} \frac{d^r}{dq^r} \sum_{k=0}^{+\infty} q^k = \frac{p^r}{(r-1)!} \frac{d^r}{dq^r} \frac{1}{1-q} = \frac{p^r}{(r-1)!} r! \frac{1}{(1-q)^{r+1}} = \\
 &= \frac{p^r}{(r-1)!} \frac{r!}{p^{r+1}} = \frac{r}{p}
 \end{aligned}$$

mentre

$$\begin{aligned}
 \sum_{k=r}^{+\infty} k^2 p \binom{k-1}{r-1} p^{r-1} q^{k-r} &= \sum_{k=r}^{+\infty} k p^r \frac{k(k-1) \dots (k-r+1)}{(r-1)!} q^{k-r} = \\
 &= \frac{p^r}{(r-1)!} \sum_{k=r}^{+\infty} k k(k-1) \dots (k-r+1) q^{k-r} = \frac{p^r}{(r-1)!} \sum_{k=r}^{+\infty} k \frac{d^r}{dq^r} q^k = \\
 &= \frac{p^r}{(r-1)!} \frac{d^r}{dq^r} \sum_{k=r}^{+\infty} k q^k = \frac{p^r}{(r-1)!} \frac{d^r}{dq^r} \sum_{k=0}^{+\infty} k q^k = \\
 &= \frac{p^r}{(r-1)!} \frac{d^r}{dq^r} \frac{q}{(1-q)^2} = \frac{p^r}{(r-1)!} \left(\frac{1}{(1-q)^2} - \frac{1}{(1-q)} \right) = \\
 &= \frac{p^r}{(r-1)!} \left(\frac{(r+1)!}{p^{r+2}} - \frac{r!}{p^{r+1}} \right) = \\
 &= \frac{p^r}{(r-1)!} r! \frac{(r+q)}{p^{r+2}} = \frac{r(r+q)}{p^2}
 \end{aligned}$$

da cui

$$\sigma_{\xi}^2 = \sum_{k=r}^{+\infty} k^2 p \binom{k-1}{r-1} p^{r-1} q^{k-r} - \mu_{\xi}^2 = \frac{r(r+q)}{p^2} - \frac{r^2}{p^2} = \frac{rq}{p^2}$$

Consideriamo un centralino telefonico che in media riceve λ chiamate all'ora e supponiamo di voler determinare la probabilità che riceva k chiamate in un'ora.

Suddividiamo l'ora in n parti uguali ciascuna della durata di $\frac{1}{n}$; durante ciascuno degli n periodi di durata $\frac{1}{n}$ la probabilità che si riceva una chiamata è λ/n , pertanto la probabilità che si ricevano k chiamate si può ottenere considerando la probabilità che una variabile aleatoria binomiale relativa ad n prove ripetute con probabilità di successo $\frac{\lambda}{n}$ assuma valore k . Sia quindi ξ la variabile aleatoria che restituisce il numero di successi ottenuti.

Avremo che

$$\begin{aligned}\varphi(k) = \mathcal{P}(\xi = k) &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\ &= \lambda^k \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} =\end{aligned}$$

Se ora consideriamo di far tendere n a $+\infty$ avremo

$$\left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda} \quad \left(1 - \frac{\lambda}{n}\right)^{-k} \rightarrow 1$$

$$\begin{aligned}
\binom{n}{k} \left(\frac{1}{n}\right)^k &= \frac{n!}{k!(n-k)!} \frac{1}{n^k} = \frac{1}{k!} \frac{n!}{(n-k)! n^k} = \\
&= \frac{1}{k!} \frac{n^n e^{-n} \sqrt{2\pi n}}{(n-k)^{n-k} e^{-(n-k)} \sqrt{2\pi(n-k)} n^k} = \\
&= \frac{1}{k!} \frac{1}{e^k} \sqrt{\frac{n}{n-k}} \frac{n^n}{(n-k)^n} \frac{(n-k)^k}{n^k} = \\
&= \frac{1}{k!} \frac{1}{e^k} \sqrt{\frac{n}{n-k}} \left(1 - \frac{k}{n}\right)^{-n} \left(1 - \frac{k}{n}\right)^k \rightarrow \frac{1}{k!}
\end{aligned}$$

non appena si tenga conto che

$$\left(1 - \frac{k}{n}\right)^{-n} \rightarrow e^k \quad \text{e} \quad \sqrt{\frac{n}{n-k}} \rightarrow 1$$

Pertanto la funzione distribuzione di probabilità della variabile aleatoria considerata è data da

$$\varphi(k) = \frac{1}{k!} \lambda^k e^{-\lambda}$$

In generale chiamiamo variabile aleatoria di Poisson la variabile che restituisce il numero di eventi accaduti nell'unità di tempo, noto il fatto che il numero medio di eventi che accadono nell'unità di tempo è λ .

Per calcolare media, varianza ed i momenti della distribuzione di Poisson è utile calcolare la funzione generatrice dei momenti.

$$M_{\xi}(t) = \sum_{k=0}^{+\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{+\infty} \frac{(\lambda e^t)^k}{k!} e^{-\lambda} = e^{\lambda(e^t-1)}$$

e le sue derivate

$$\frac{d}{dt} M_{\xi}(t) = e^{\lambda(e^t-1)} \lambda e^t$$

$$\frac{d^2}{dt^2} M_{\xi}(t) = e^{\lambda(e^t-1)} \lambda^2 e^t + e^{\lambda(e^t-1)} \lambda e^t$$

Calcolando in $t = 0$ si ottiene

$$\frac{d}{dt} M_{\xi}(0) = \lambda = \mu'_1 = \mu \qquad \frac{d^2}{dt^2} M_{\xi}(0) = \lambda^2 + \lambda = \mu'_2$$

da cui

$$\mu = \mu'_1 = \lambda \qquad \sigma^2 = \mu'_2 - \mu^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Sui può verificare che la somma di due variabili aleatorie di Poisson di media λ e μ è ancora una variabile di Poisson di media $\lambda + \mu$, infatti se

$$\varphi(k) = \frac{1}{k!} \lambda^k e^{-\lambda} \quad \text{e} \quad \psi(k) = \frac{1}{k!} \mu^k e^{-\mu}$$

possiamo calcolare la densità di probabilità della variabile somma mediante la

$$\begin{aligned} \theta(k) &= \sum_{h=0}^k \varphi(h) \psi(k-h) = \sum_{h=0}^k \frac{1}{h!} \lambda^h e^{-\lambda} \frac{1}{(k-h)!} \mu^{k-h} e^{-\mu} = \\ &= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{h=0}^k \frac{k!}{h!(k-h)!} \lambda^h \mu^{k-h} = \\ &= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{h=0}^k \binom{k}{h} \lambda^h \mu^{k-h} = e^{-(\lambda+\mu)} \frac{1}{k!} (\lambda + \mu)^k \end{aligned}$$

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

$$\begin{aligned}\varphi(n_1, n_2, \dots, n_k) &= \mathcal{P}(\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_k = n_k) = \\ &= \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}\end{aligned}$$

Infatti su n tentativi si sono verificati n_1 successi con probabilità p_1 , sui restanti $n - n_1$ tentativi si sono verificati n_2 successi con probabilità p_2 e così via fino ad ottenere n_k successi su $(n - n_1 - n_2 - \dots - n_{k-1})$ tentativi per cui

$$\varphi(n_1, n_2, \dots, n_k) = \mathcal{P}(\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_k = n_k) =$$

$$= \frac{n!}{n_1!(n - n_1)!} p_1^{n_1} \frac{(n - n_1)!}{n_2!(n - n_1 - n_2)!} p_2^{n_2} \dots$$

$$\dots \frac{(n - n_1 - n_2 - \dots - n_{k-1})!}{(n_k)!(n - n_1 - n_2 - \dots - n_k)!} p_k^{n_k} =$$

(dal momento che $n - n_1 - n_2 - \dots - n_k = 0$)

$$= \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

Sia $\xi = (\xi_1, \xi_2, \dots, \xi_k)$ una variabile aleatoria distribuita multinomialmente relativa al caso in cui gli eventi $A_1, A_2, \dots, A_k$ hanno probabilità di accadimento $p_1, p_2, \dots, p_k$ e sia $\eta = \eta_1 + \eta_2 + \dots + \eta_k$ una variabile aleatoria le cui componenti η_j sono variabili indipendenti con densità di Poisson di media $\lambda_k = np_k$ Allora

$$\begin{aligned} \mathcal{P}(\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_k = n_k) &= \\ &= \mathcal{P}(\eta_1 = n_1, \eta_2 = n_2, \dots, \eta_n = n_k | \eta_1 + \eta_2 + \dots + \eta_k = n) \end{aligned}$$

Infatti, dal momento che le η_j son indipendenti si ha che

$$p_1 + p_2 + \cdots p_k = 1$$

e che

$$\begin{aligned} \mathcal{P}(\eta_1 = n_1, \eta_2 = n_2, \dots, \eta_k = n_k) &= \\ \left(\frac{(np_1)^{n_1} e^{-np_1}}{n_1!} \right) \left(\frac{(np_2)^{n_2} e^{-np_2}}{n_2!} \right) \cdots \left(\frac{(np_k)^{n_k} e^{-np_k}}{n_k!} \right) &= \\ = \left(\frac{(n)^{n_1+n_2+\cdots n_k} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}}{n_1! n_2! \cdots n_k!} \right) e^{-n(p_1+p_2+p_k)} &= \\ = \left(\frac{(n)^{n_1+n_2+\cdots n_k} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}}{n_1! n_2! \cdots n_k!} \right) e^{-n} \end{aligned}$$

D'altra parte

$$\begin{aligned}
 \mathcal{P}(\eta_1 = n_1, \eta_2 = n_2, \dots, \eta_k = n_k | \eta_1 + \eta_2 + \dots + \eta_k = n) &= \\
 &= \frac{\left(\frac{(n)^n p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}}{n_1! n_2! \dots n_k!} \right) e^{-n}}{\frac{n^n e^{-n}}{n!}} = \\
 &= \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k} = \\
 &= \mathcal{P}(\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_k = n_k)
 \end{aligned}$$

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

$$\varphi(k) = \mathcal{P}(\xi = k) = \binom{n}{k} \left(\frac{b}{b+w} \right)^k \left(\frac{w}{b+w} \right)^{n-k} = \binom{n}{k} \frac{b^k w^{n-k}}{b+w}$$

non appena si ricordi la distribuzione binomiale e si tenga presente che

$$p = \frac{b}{b+w} \quad , \quad q = \frac{w}{b+w}$$

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

$$\varphi(k) = \mathcal{P}(\xi = k) = \frac{\binom{b}{k} \binom{w}{n-k}}{\binom{b+w}{n}}$$

Infatti il denominatore conta quante $n - ple$ di palline si possono formare avendo a disposizione $b + w$ palline, mentre al numeratore c'è il numero delle $n - ple$ che contengono esattamente k palline nere che si possono ottenere combinando una $k - pla$ di palline nere, in numero di $\binom{b}{k}$, con una $(n - k) - pla$ di palline bianche, in numero di $\binom{w}{n-k}$.

Si calcola anche che

$$\mu = \frac{nb}{b+w} \text{ e } \sigma^2 = \frac{nbw(b+w-n)}{(b+w)^2(b+w-1)} \text{ infatti}$$

$$\begin{aligned} \mu &= \sum_{k=0}^n k \mathcal{P}(\xi = k) = \sum_{k=0}^n k \frac{\binom{b}{k} \binom{w}{n-k}}{\binom{b+w}{n}} = \\ &= \frac{1}{\binom{b+w}{n}} \sum_{k=1}^n b \binom{b-1}{k-1} \binom{w}{n-k} = \\ &= \frac{1}{\binom{b+w}{n}} \sum_{k=0}^{n-1} b \binom{b-1}{k} \binom{w}{n-k-1} = \\ &= b \frac{1}{\binom{b+w}{n}} \binom{b+w-1}{n-1} = \\ &= \frac{n!(b+w-n)!}{(b+w)!} \frac{(b+w-1)!}{(n-1)!(b+w-n)!} = \\ &= \frac{bn}{b+w} \end{aligned}$$

Lemma

- Identità di Vandermonde - Si ha

$$\binom{m+n}{k} = \sum_{h=0}^k \binom{m}{h} \binom{n}{k-h}$$

Dimostrazione. Osserviamo che

- $\binom{m+n}{k}$ è il numero dei modi con cui si possono scegliere k elementi tra $m+n$
- $\binom{m}{h}$ è il numero dei modi con cui si possono scegliere h elementi tra m
- $\binom{n}{k-h}$ è il numero dei modi con cui si possono scegliere $k-h$ elementi tra n

Si possono scegliere k elementi tra $m+n$ prendendone h tra i primi m e $k-h$ tra gli altri n , quindi, per h fissato, possiamo scegliere k elementi tra $m+n$ in $\binom{m}{h} \binom{n}{k-h}$ modi.

Sommando su h si ottiene la formula.

Inoltre

$$\begin{aligned}
\sum_{k=0}^n k^2 \mathcal{P}(\xi = k) &= \sum_{k=0}^n k^2 \frac{\binom{b}{k} \binom{w}{n-k}}{\binom{b+w}{n}} = \frac{1}{\binom{b+w}{n}} \sum_{k=1}^n k b \binom{b-1}{k-1} \binom{w}{n-k} = \\
&= \frac{1}{\binom{b+w}{n}} \sum_{k=0}^{n-1} (k+1) b \binom{b-1}{k} \binom{w}{n-k-1} = \frac{b}{\binom{b+w}{n}} \left(\sum_{k=1}^{n-1} k \binom{b-1}{k} \binom{w}{n-k-1} + \sum_{k=0}^{n-1} \binom{b-1}{k} \binom{w}{n-k-1} \right) = \\
&= \frac{b}{\binom{b+w}{n}} \left(\sum_{k=1}^{n-1} (b-1) \binom{b-2}{k-1} \binom{w}{n-k-1} + \binom{b+w-1}{n-1} \right) = \\
&= \frac{b}{\binom{b+w}{n}} \left(\sum_{k=0}^{n-2} (b-1) \binom{b-2}{k} \binom{w}{n-k-2} + \binom{b+w-1}{n-1} \right) = \\
&= \frac{b}{\binom{b+w}{n}} \left((b-1) \binom{b+w-2}{n-2} + \binom{b+w-1}{n-1} \right) = \\
&= \frac{b}{\binom{b+w}{n}} \left((b-1) \frac{(b+w-2)!}{(n-2)!(b+w-n)!} + \frac{(b+w-1)!}{(n-1)!(b+w-n)!} \right) = \\
&= \frac{bn!(b+w-n)!}{(b+w)!} \left(\frac{(b-1)(n-1)(b+w-2)! + (b+w-1)!}{(n-1)!(b+w-n)!} \right) = \\
&= \frac{bn((b+w-2)!(b-1)(n-1) + (b+w-1)!)}{(b+w)!} = \\
&= \frac{bn(nb - n - b + 1 + b + w - 1)}{(b+w)(b+w-1)} = \frac{bn(nb - n + w)}{(b+w)(b+w-1)}
\end{aligned}$$

da cui

$$\begin{aligned}
 \sigma^2 &= \frac{bn(nb - n + w)}{(b + w)(b + w - 1)} - \frac{b^2 n^2}{(b + w)^2} = \\
 &= \frac{nb(nb - n + w)(b + w) - b^2 n^2(b + w - 1)}{(b + w)^2(b + w - 1)} = \\
 &= \frac{nb(nb^2 - nb + bw - nbw - nw + w^2 - nb^2 - nbw - nb)}{(b + w)^2(b + w - 1)} = \\
 &= \frac{nbw(b + w - n)}{(b + w)^2(b + w - 1)}
 \end{aligned}$$

Possiamo anche osservare che si ha

$$\begin{aligned}
 \frac{\binom{b}{k} \binom{w}{n-k}}{\binom{b+w}{n}} &= \frac{b!}{k!(b-k)!} \frac{w!}{(n-k)!(w-(n-k))!} \frac{n!(b+w-n)!}{(b+w)!} = \\
 &= \binom{n}{k} \frac{b!}{(b-k)!} \frac{w!}{(w+k-n)!} \frac{(b+w-n)!}{(b+w)!} = \\
 &= \binom{n}{k} \frac{b(b-1) \cdots (b-k+1) w(w-1) \cdots (w+k-n+1)}{(b+w)(b+w-1) \cdots (b+w-n+1)} = \\
 &= \binom{n}{k} \frac{b}{b+w} \frac{b-1}{b+w-1} \cdots \frac{(b-k+1)}{b+w-k+1} \cdot \\
 &\quad \frac{w}{b+w-k} \frac{w-1}{b+w-k-1} \cdots \frac{w-n+k+1}{b+w-n+1}
 \end{aligned}$$

Ora se $b + w = N$, $\frac{b}{b+w} = p$ e $\frac{w}{b+w} = q$ si ha, dividendo numeratori e denominatori per $(b + w)$,

$$\frac{\binom{b}{k} \binom{w}{n-k}}{\binom{b+w}{n}} =$$

$$= \binom{n}{k} \frac{p}{1} \frac{p - \frac{1}{N}}{1 - \frac{1}{N}} \dots \frac{p - \frac{k-1}{N}}{1 - \frac{k-1}{N}} \frac{q}{1 - \frac{k}{N}} \frac{q - \frac{1}{N}}{1 - \frac{k+1}{N}} \dots \frac{q - \frac{n-k-1}{N}}{1 - \frac{n-1}{N}}$$

e

$$\lim_{N \rightarrow +\infty} \frac{\binom{b}{k} \binom{w}{n-k}}{\binom{b+w}{n}} = \binom{n}{k} p^k q^{n-k}$$

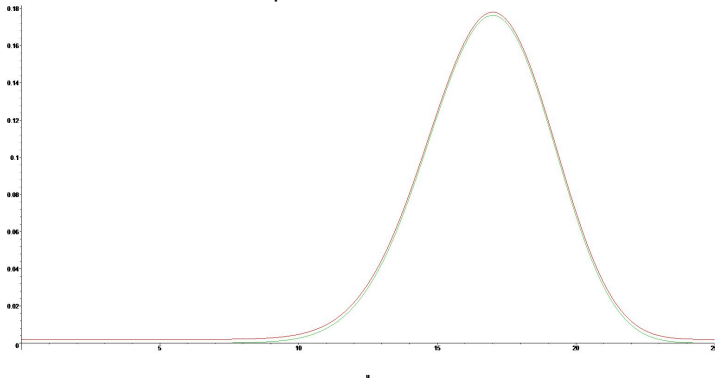
Osserviamo anche che, con queste notazioni, la media e la varianza si esprimono come

$$\mu = np \quad \sigma^2 = npq \frac{N-n}{N-1}$$

e che pertanto la media della distribuzione ipergeometrica è uguale alla media della distribuzione binomiale; per quanto concerne la varianza, possiamo vedere che il rapporto tra la varianza della ipergeometrica e la varianza della binomiale è dato da

$$\frac{N-n}{N-1} \rightarrow 1 \quad \text{per} \quad N \rightarrow +\infty$$

È pertanto evidente che per N grande la distribuzione ipergeometrica si riduce a quella binomiale.



In figura è riportata la PDF di una variabile aleatoria Binomiale in rosso e di una variabile aleatoria Ipergeometrica in verde nel caso in cui $p = 0.7$, $q = 0.3$, $N = 7200$, $n = 24$. Il grafico della PDF binomiale e' stato disegnato aumentato di 0.002 dal momento che senza innalzamento i due grafici si sovrappongono completamente.

Consideriamo ancora un centralino telefonico che in media riceve λ chiamate nell'unità di tempo; abbiamo già visto che la variabile aleatoria che restituisce il numero di chiamate in un'unità di tempo ha una distribuzione di Poisson di media e varianza λ .

Consideriamo ora la variabile aleatoria che restituisce il tempo che intercorre tra una chiamata e l'altra. A questo scopo definiamo $P_n(h)$ come la probabilità che ci siano n chiamate in un intervallo di tempo h ;

Dal momento che la media delle chiamate è, in questo caso, λh avremo che

$$P_0(h) = e^{-\lambda h}$$

Consideriamo un intervallo di tempo $[0, t]$; sia T la variabile aleatoria che restituisce il tempo in cui avviene la prima chiamata a partire da 0.

La probabilità che $T > t$, si ottiene calcolando la probabilità che in $[0, t]$ non ci siano state chiamate e, se dividiamo $[0, t]$ in n parti uguali, possiamo calcolare $\mathcal{P}(T > t)$ osservando che in ciascuno dei sottointervalli ottenuti ci sono state 0 chiamate.

Avremo pertanto che

$$\mathcal{P}(T > t) = \left(P_0 \left(\frac{t}{n} \right) \right)^n = e^{-\lambda \frac{t}{n} n} = e^{-\lambda t} = \int_t^{+\infty} \varphi(t) dt$$

per cui

$$\int_{-\infty}^t \varphi(t) dt = \mathcal{P}(T < t) = 1 - \int_t^{+\infty} \varphi(t) dt = 1 - e^{-\lambda t}$$

e possiamo trovare la PDF φ della distribuzione esponenziale semplicemente derivando rispetto a t

$$\varphi(t) = \lambda e^{-\lambda t}$$

Si verifica subito che media e varianza sono date da:

$$\mu = \int_0^{+\infty} \lambda t e^{-\lambda t} dt = -te^{-\lambda t} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda t} dt = -\frac{e^{-\lambda t}}{\lambda} \Big|_0^{+\infty} = \frac{1}{\lambda}$$

mentre

$$\begin{aligned} \int_0^{+\infty} \lambda t^2 e^{-\lambda t} dt &= -t^2 e^{-\lambda t} \Big|_0^{+\infty} + \int_0^{+\infty} 2te^{-\lambda t} dt = \\ &= 2 \left(-\frac{te^{-\lambda t}}{\lambda} \Big|_0^{+\infty} - \int_0^{+\infty} \frac{e^{-\lambda t}}{\lambda} dt \right) = 2 \left(-\frac{e^{-\lambda t}}{\lambda^2} \Big|_0^{+\infty} \right) = \frac{2}{\lambda^2} \end{aligned}$$

per cui

$$\sigma = \frac{2}{\lambda^2} - \mu^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

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.

In precedenza, nel 1733, De Moivre aveva pubblicato una derivazione della legge normale come limite di una distribuzione binomiale ed anche Laplace la conosceva già almeno dal 1774. Gauss invece arrivò a considerare la distribuzione normale studiando il problema di stimare un parametro noto un certo numero di sue osservazioni. Per questo scopo applicò quello che si chiama oggi principio di massima verosimiglianza.

Una derivazione della legge normale molto interessante è dovuta ad Herschel , 1850, che la dedusse studiando la distribuzione a due dimensioni degli errori di misurazione della posizione di una stella. In seguito, nel 1860 James Clerk Maxwell estese le argomentazioni di Herschel al caso tridimensionale, studiando la distribuzione di probabilità della velocità delle molecole in un gas.

Infine nel 1941 un ingegnere elettrico, Vernon D.Landon, studiando il rumore associato al voltaggio in circuito elettrico osservò che era distribuito con densità invarianti rispetto all'intensità del disturbo e provò che questa caratteristica identifica la distribuzione normale.

La varietà e la diversità dei problemi che conducono alla distribuzione normale giustificano quindi pienamente il ruolo centrale che questa distribuzione occupa.

Sia ξ_n la variabile aleatoria binomiale definita dalla somma di n variabili Bernoulliane ξ relative ad una prova con probabilità di successo p e probabilità di insuccesso $q = 1 - p$.

$$\xi_n = \xi + \xi + \xi + \dots + \xi = n\xi$$

La media di ξ è $\mu = p$, la sua varianza è $\sigma = \sqrt{pq}$ per cui la media di ξ_n è np la sua varianza è $\sqrt{npq}$ mentre la sua densità di probabilità è definita da

$$\mathcal{B}_n(k) = \mathcal{P}(\xi_n) = \binom{n}{k} p^k q^{n-k}$$

Utilizzando la Formula di Stirling possiamo approssimare il valore della distribuzione binomiale mediante la

$$\begin{aligned}
 \mathcal{B}_n(h) &= \binom{n}{h} p^h q^{n-h} = \frac{n!}{h!(n-h)!} p^h q^{n-h} = \\
 &\approx \frac{n^n e^{-n} \sqrt{2\pi n}}{h^h e^{-h} \sqrt{2\pi h} (n-h)^{(n-h)} e^{-(n-h)} \sqrt{2\pi(n-h)}} p^h q^{(n-h)} = \\
 &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{h(n-h)}} \frac{n^n}{h^h (n-h)^{n-h}} p^h q^{(n-h)} \quad (7)
 \end{aligned}$$

Se normalizziamo la variabile aleatoria ξ_n mediante la

$$\zeta_n = \frac{\xi_n - np}{\sqrt{npq}}$$

la funzione $G_n(k)$ densità di probabilità di ζ_n è data da

$$G_n(k) = \sqrt{npq} \mathcal{B}_n(np + k\sqrt{npq})$$

Infatti

$$\mathcal{B}_n(k) = 1 \cdot \mathcal{B}_n(k) = \mathcal{P}(\xi_n = k)$$

per cui, detta $G_n(k)$ la PDF di ζ_n , avremo che

$$\begin{aligned} \frac{1}{\sqrt{npq}} G_n(k) &= \mathcal{P}\left(k \leq \zeta_n \leq k + \frac{1}{\sqrt{npq}}\right) = \\ &= \mathcal{P}(k\sqrt{npq} \leq \xi_n - np \leq k\sqrt{npq} + 1) = \\ &= \mathcal{B}_n(np + k\sqrt{npq}) \end{aligned}$$

Possiamo allora mostrare che

$$G_n(k) \rightarrow \frac{1}{\sqrt{2\pi}} e^{-\frac{k^2}{2}}$$

Definiamo

$$h = np + k\sqrt{npq} = np + \delta\sqrt{n}, \quad \text{con} \quad \delta = k\sqrt{pq}$$

avremo

$$n - h = n - np - k\sqrt{npq} = nq - \delta\sqrt{n}$$

per cui sia h che $n - h$ sono grandi per n grande e possiamo usare la 7 per approssimare $\mathcal{B}_n(h)$ e calcolare

$$\begin{aligned}
 \sqrt{npq}\mathcal{B}_n(np + \delta\sqrt{n}) &\approx \\
 &= \alpha_n\beta_n = \left(\sqrt{npq} \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{(np + \delta\sqrt{n})(nq - \delta\sqrt{n})}} \right) \\
 &\quad \left(\frac{n^n}{(np + \delta\sqrt{n})^{(np+\delta\sqrt{n})} (nq - \delta\sqrt{n})^{(nq-\delta\sqrt{n})}} p^{(np+\delta\sqrt{n})} q^{(nq-\delta\sqrt{n})} \right)
 \end{aligned}$$

Osserviamo subito che

$$\alpha_n = \sqrt{npq} \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{(np + \delta\sqrt{n})(nq - \delta\sqrt{n})}} \rightarrow \frac{1}{\sqrt{2\pi}}$$

e che

$$\begin{aligned}
 \beta_n &= \frac{n^n}{(np + \delta\sqrt{n})^{(np+\delta\sqrt{n})} (nq - \delta\sqrt{n})^{(nq-\delta\sqrt{n})}} p^{(np+\delta\sqrt{n})} q^{(nq-\delta\sqrt{n})} = \\
 &= \frac{n^{np+\delta\sqrt{n}+nq-\delta\sqrt{n}}}{(np + \delta\sqrt{n})^{(np+\delta\sqrt{n})} (nq - \delta\sqrt{n})^{(nq-\delta\sqrt{n})}} p^{(np+\delta\sqrt{n})} q^{(nq-\delta\sqrt{n})} = \\
 &\quad \frac{n^{np+\delta\sqrt{n}}}{(np + \delta\sqrt{n})^{(np+\delta\sqrt{n})}} \frac{n^{nq-\delta\sqrt{n}}}{(nq - \delta\sqrt{n})^{(nq-\delta\sqrt{n})}}
 \end{aligned}$$

Si ha

$$\begin{aligned}
 \beta_n &= \left(1 + \frac{\delta}{p\sqrt{n}}\right)^{-np-\delta\sqrt{n}} \left(1 - \frac{\delta}{q\sqrt{n}}\right)^{-nq+\delta\sqrt{n}} = \\
 &= e^{(-np-\delta\sqrt{n}) \ln\left(1 + \frac{\delta}{p\sqrt{n}}\right) + (-nq+\delta\sqrt{n}) \ln\left(1 - \frac{\delta}{q\sqrt{n}}\right)}
 \end{aligned}$$

Ma

$$\begin{aligned}
 & (-np - \delta\sqrt{n}) \ln \left(1 + \frac{\delta}{p\sqrt{n}} \right) + (-nq + \delta\sqrt{n}) \ln \left(1 - \frac{\delta}{q\sqrt{n}} \right) \approx \\
 & \approx (-np - \delta\sqrt{n}) \left(\frac{\delta}{p\sqrt{n}} - \frac{\delta^2}{2p^2n} \right) + (-nq + \delta\sqrt{n}) \left(-\frac{\delta}{q\sqrt{n}} - \frac{\delta^2}{2q^2n} \right) \approx \\
 & -\frac{np\delta}{p\sqrt{n}} + \frac{np\delta^2}{2p^2n} - \frac{\delta^2}{p} + \frac{\delta^3}{2p^2\sqrt{n}} + \frac{nq\delta}{q\sqrt{n}} + \frac{nq\delta^2}{2q^2n} - \frac{\delta^2}{q} - \frac{\delta^3}{2q^2\sqrt{n}} \rightarrow \\
 & \rightarrow \frac{\delta^2}{2p} + \frac{\delta^2}{2q} - \frac{\delta^2}{p} - \frac{\delta^2}{q} = -\left(\frac{\delta^2}{2p} + \frac{\delta^2}{2q} \right) = \\
 & = -\frac{k^2pq}{2p} - \frac{k^2pq}{2q} = -\frac{k^2}{2}
 \end{aligned}$$

Pertanto

$$\beta_n \rightarrow e^{\frac{-k^2}{2}}$$

ed infine

$$\sqrt{npq}\mathcal{B}_n(np + \delta\sqrt{n}) \rightarrow \frac{1}{2\sqrt{\pi}}e^{-\frac{k^2}{2}}$$

Quindi possiamo affermare che $\zeta_n \rightarrow \zeta$ essendo ζ la variabile aleatoria la cui PDF è

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

Ne segue che

$$\frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

è la funzione distribuzione di probabilità della variabile aleatoria cui tende ξ_n per n grande.

Una stella si individua mediante la sua longitudine, misurata con un errore ξ , e la sua declinazione, misurata con un errore η ; Herschel (1850) postulò che ξ e η sono variabili aleatorie indipendenti ed hanno la stessa PDF f . Allora la distribuzione di probabilità della variabile aleatoria (ξ, η) è della forma

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

Inoltre postulò che ϕ , espressa usando le coordinate polari (ρ, θ) , fosse indipendente da θ , per cui

$$\phi(x, y) = f(x)f(y) = g(x^2 + y^2)$$

Ne dedusse infine che ϕ doveva allora essere quella che ora chiamiamo densità di probabilità gaussiana.

Maxwell (1860) usò argomentazioni sostanzialmente identiche nello studio della cinetica dei gas estendendo l'idea di Herschel al caso tridimensionale.

Infatti, ponendo $x = 0$ e $f(0) = \alpha$ si ricava $f(x)\alpha = g(x^2)$ da cui

$$\frac{g(x^2)}{\alpha} \frac{g(y^2)}{\alpha} = g(x^2 + y^2)$$

dividendo per α^2

$$\frac{g(x^2)}{\alpha^2} \frac{g(y^2)}{\alpha^2} = \frac{g(x^2 + y^2)}{\alpha^2}$$

e

$$\ln \left(\frac{g(x^2)}{\alpha^2} \right) + \ln \left(\frac{g(y^2)}{\alpha^2} \right) = \ln \left(\frac{g(x^2 + y^2)}{\alpha^2} \right) \text{ da cui}$$

Pertanto la funzione $z \mapsto \ln \left(\frac{g(z)}{\alpha^2} \right)$ è lineare e quindi

$$\ln \left(\frac{g(z)}{\alpha^2} \right) = kz \quad \text{e} \quad \ln \left(\frac{g(x^2)}{\alpha^2} \right) = kx^2$$

È opportuno ricordare che una funzione f additiva e continua è lineare, infatti in tal caso si ha

$$f(nx) = nf(x)$$

e di conseguenza

$$f\left(\frac{x}{m}\right) = \frac{1}{m}f(x)$$

per ogni n intero, e per ogni m naturale. Se ne deduce che

$$f\left(\frac{n}{m}x\right) = \frac{n}{m}f(x)$$

$$f(qx) = qf(x)$$

per ogni q razionale. La continuità garantisce poi che

$$f(\alpha x) = \alpha f(x)$$

per ogni x reale e l'omogeneità di f che essendo già additiva risulta lineare.

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

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Gauss considerò il problema di stimare un parametro θ note $n + 1$ sue osservazioni $x_0, x_1, \dots, x_n$ ed usò allo scopo il principio di massima verosimiglianza.

Se indichiamo con $f(x_i, \theta)$ la probabilità di ottenere la misura x_i condizionata al fatto che il parametro cercato è θ , la probabilità di avere ottenuto le osservazioni x_i , che supponiamo indipendenti, è data da

$$P(\theta) = \prod f(x_i, \theta)$$

Cerchiamo di determinare f in modo che $P(\theta)$ sia massima in corrispondenza del valor medio $\hat{\theta}$ delle osservazioni. L'argomento di massimo non cambia se consideriamo $\ln(P(\theta))$ in luogo di $P(\theta)$ e quindi θ deve soddisfare la condizione

$$\sum_0^n \frac{\partial}{\partial \theta} \ln(f(x_i, \theta)) = 0$$

Se poniamo

$$\ln(f(x_i, \theta)) = g(\theta - x_i) = g(u)$$

se supponiamo cioè che $f(x_i, \theta)$ dipenda solo dall'errore commesso nel considerare θ invece di x_i , dovrà risultare

$$\sum_0^n g'(\theta - x_i) = 0 \quad (8)$$

Se vogliamo che il massimo sia assunto per

$$\theta = \hat{\theta} = \frac{1}{n+1} \sum_0^n x_i \quad (9)$$

dovrà aversi

$$\sum_0^n g'(\hat{\theta} - x_i) = 0 \quad (10)$$

Le equazioni 10, 9 sono, in generale, incompatibili, anzi possiamo subito osservare che qualora la distribuzione degli errori di misurazione sia uniforme si avrebbe $f(x_i, \theta)$ costante per cui tale risulterebbe anche $P(\theta)$ e la 10 perderebbe di significato attribuendo a qualunque valore di θ la stessa affidabilità. Ora, se consideriamo il caso in cui una sola delle osservazioni differisce dalla media, cioè se poniamo

$$x_0 = (n+1)u, \quad x_1 = x_2 = \cdots = x_n = 0$$

per modo che

$$\hat{\theta} = u, \quad \hat{\theta} - x_0 = -nu, \quad \hat{\theta} - x_i = u - 0, i = 1, \cdots n$$

sostituendo in 10 otteniamo

$$g'(-nu) + ng'(u) = 0 \tag{11}$$

Per $n = 1$ la 11 fornisce

$$g'(-u) = -g'(u)$$

e ne viene che g' deve essere antisimmetrica inoltre la 11, garantisce che la g' , supposta continua, è lineare.

La linearità permette di concludere che

$$g'(u) = au, \quad g(u) = \frac{1}{2}au^2 + b$$

Infine usando la definizione di g e tenendo conto che la distribuzione di probabilità deve essere normalizzata, otteniamo che $a < 0$ ed anche

$$f(x, \theta) = \sqrt{\frac{\alpha}{2\pi}} e^{-\frac{1}{2}\alpha(x-\theta)^2} \quad (12)$$

Nel caso in cui la 12 valga la funzione che definisce la verosimiglianza diventa

$$\begin{aligned}\ln(P(\theta)) &= \sum_0^n \ln \left(\sqrt{\frac{\alpha}{2\pi}} e^{-\frac{1}{2}\alpha(x_i - \theta)^2} \right) = \\ &= (n + 1) \ln \left(\sqrt{\frac{\alpha}{2\pi}} \right) + \sum_0^n -\frac{1}{2}\alpha(x_i - \theta)^2\end{aligned}$$

Ne segue che

$$(\ln(P(\theta)))' = \sum_0^n \alpha(x_i - \theta) = \sum_0^n \alpha x_i - (n + 1)\theta$$

Che evidentemente assume massimo proprio per $\theta = \hat{\theta}$ ed inoltre il punto di massimo è unico (si tratta di un paraboloide concavo). Resta libero il parametro α che definisce diverse forme della distribuzione.

Landon studiò la distribuzione di probabilità della variabile aleatoria ξ che rappresenta il rumore da cui è affetto il voltaggio osservato in un circuito elettrico. Egli osservò che tale distribuzione sembrava non cambiare forma al variare della deviazione standard σ del rumore.

Sia quindi $p(x, \sigma)$ la probabilità che il rumore valga x nel caso in cui la sua deviazione standard sia σ .

Supponiamo che alla variabile aleatoria ξ si sommi una quantità aleatoria $\Delta\xi$ indipendente da ξ la cui distribuzione si chiami q .

Avremo che

$$f(x) = \int_{-\infty}^{+\infty} p(x - \epsilon, \sigma^2) q(\epsilon) d\epsilon$$

è la distribuzione di probabilità della variabile aleatoria $\xi + \Delta\xi$

Usando lo sviluppo di Taylor di ordine 2 di $p(x, \sigma)$ otteniamo

$$\begin{aligned} f(x) &= \\ &= \int_{-\infty}^{+\infty} q(\epsilon) \left(p(x, \sigma^2) - \epsilon \frac{\partial p(x, \sigma^2)}{\partial x} + \epsilon^2 \frac{\partial^2 p(x, \sigma^2)}{2 \partial x^2} + R \right) d\epsilon \end{aligned}$$

e

$$f(x) = p(x, \sigma^2) - \mu(q) \frac{\partial p(x, \sigma^2)}{\partial x} + \text{Var}(q) \frac{1}{2} \frac{\partial^2 p(x, \sigma^2)}{\partial x^2} + R$$

Pertanto supponendo che la media $\mu(q)$ sia nulla e trascurando i momenti di ordine superiore al secondo, si ha

$$f(x) = p(x, \sigma^2) + \text{Var}(q) \frac{1}{2} \frac{\partial^2 p(x, \sigma^2)}{\partial x^2} + R$$

D'altro canto, poichè si suppone che la forma della distribuzione di probabilità del rumore non cambi con la varianza e dal momento che $\text{Var}(\xi + \Delta\xi) = \sigma^2 + \text{Var}(\Delta\xi)$, deve essere

$$f(x) = p(x, \sigma^2 + \text{Var}(q)) = p(x, \sigma^2) + \text{Var}(q) \frac{\partial p(x, \sigma^2)}{\partial \sigma} + R$$

Ove si sia usato lo sviluppo di Taylor di $p(x, \cdot)$.

Possiamo quindi dedurre confrontando i due sviluppi che

$$\frac{\partial p(x, \sigma^2)}{\partial \sigma} = \frac{1}{2} \frac{\partial^2 p(x, \sigma^2)}{\partial x^2}$$

che è una equazione di diffusione che , con condizioni iniziali $p(x, 0) = \delta(x)$ fornisce come soluzione

$$p(x, \sigma^2) = N(x, 0, \sigma)$$

Si consideri una sbarra omogenea infinita, adiabatica, avente inizialmente una distribuzione di temperatura $u_0(x)$ (x denota l'ascissa sulla sbarra).

L'andamento della temperatura nello spazio e nel tempo $u(x, t)$ sarà regolato nella seguente maniera: se Q indica la quantità di calore e T la temperatura, si ha, in una porzione della sbarra, di lunghezza Δx e sezione S ,

$$\Delta Q = c\sigma S\Delta x\Delta T.$$

D'altro canto la quantità di calore che fluisce attraverso la sezione ad ascissa x è

$$\Delta Q = KST'(x)\Delta t$$

per cui, a meno di costanti,

$$\frac{\Delta T}{\Delta t} = \frac{T'(x + \Delta x) - T'(x)}{\Delta x}$$

e, al limite per Δx e $\Delta t \rightarrow 0$, si ha che la temperatura $u(x, t)$ deve soddisfare il seguente problema alle derivate parziali

$$\begin{cases} u_t = u_{xx} \\ u(x, 0) = u_0(x) \end{cases}$$

Cerchiamo una soluzione $u : \mathbb{R} \times \bar{R}_+ \rightarrow \mathbb{R}$ tale che u , u_x , u_{xx} sono continue, assolutamente integrabili in x

$$|u_t(x, t)| \leq \phi_a(x) \quad , \quad \int_{-\infty}^{+\infty} \phi_a(x) dx < +\infty \quad , \quad \forall t \geq a > 0 \quad .$$

Supponiamo inoltre che u_0 , u'_0 , u''_0 siano continue e assolutamente integrabili in $\mathbb{R}$.

Poniamo

$$\begin{aligned}v(\omega, t) &= \mathcal{F}(u(\cdot, t))(\omega) \\v_0(\omega) &= \mathcal{F}(u_0)(\omega) = v(\omega, 0).\end{aligned}$$

Avremo

$$\mathcal{F}(u_{xx}(\cdot, t))(\omega) = -\omega^2 v(\omega, t)$$

e

$$\begin{aligned}\frac{\partial}{\partial t} \mathcal{F}(u(\cdot, t))(\omega) &= \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} u(s, t) e^{-i\omega s} ds = \\&= \int_{-\infty}^{+\infty} u_t(s, t) e^{-i\omega s} ds = \mathcal{F}(u_t(\cdot, t))(\omega).\end{aligned}$$

Applicando la trasformata di Fourier all'equazione, si ha

$$\mathcal{F}(u_t(\cdot, t) - u_{xx}(\cdot, t)) = 0$$

e

$$v_t(\omega, t) + \omega^2 v(\omega, t) = 0 \quad , \quad v(\omega, 0) = v_0(\omega).$$

Ne viene

$$v(\omega, t) = e^{-\omega^2 t} v_0(\omega) = \mathcal{F}(k(\cdot, t))(\omega) \mathcal{F}(u_0)(\omega)$$

essendo

$$k(x, t) = \frac{e^{-x^2/(4t)}}{2(\pi t)} \quad , \quad t > 0$$

e perciò

$$v(\omega, t) = \mathcal{F}(k(\cdot, t) * u_0)(\omega).$$

Ne segue che

$$\mathcal{F}(u(\cdot, t)) = \mathcal{F}(k(\cdot, t) * u_0)$$

e per l'iniettività della trasformata di Fourier, se $t > 0$,

$$\begin{aligned} u(x, t) &= (k(\cdot, t) * u_0)(x) = \frac{1}{2(\pi t)} \int_{-\infty}^{+\infty} e^{-\frac{s^2}{4t}} u_0(x - s) ds = \\ &= \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{-z^2} u_0(x - 2zt) dz = \\ &= \frac{1}{2(\pi t)} \int_{-\infty}^{+\infty} e^{-\frac{(x-u)^2}{4t}} u_0(u) du. \end{aligned}$$

Dall'ultimo membro e dalle ipotesi su u_0 segue immediatamente la continuità di $u(\cdot, t)$ per $t > 0$. In modo analogo si prova la continuità di u_x e u_{xx} .

Poiché u'_0 è assolutamente integrabile su $\mathbb{R}$, u_0 è limitata e dal penultimo membro si ha

$$\lim_{t \rightarrow 0^+} u(x, t) = u_0(x).$$

Usando il teorema 27.13 si prova la assoluta integrabilità di u , u_x , u_{xx} .

Analogamente si verificano le ipotesi su u_t .

Siano ξ_1 e ξ_2 variabili aleatorie gaussiane di media 0 e varianza 1 e consideriamo $\xi = \alpha\xi_1 + \beta\xi_2$ con $\alpha^2 + \beta^2 = 1$, allora ξ è una variabile aleatoria normale standard.

Avremo che la PDF di ξ_1 e di ξ_2 è data da

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

e si può affermare che $\alpha\xi_1$ e $\beta\xi_2$ hanno rispettivamente, le seguenti PDF:

$$\frac{1}{\sqrt{2\pi}|\alpha|} e^{-\frac{u^2}{2\alpha^2}}, \quad \frac{1}{\sqrt{2\pi}|\beta|} e^{-\frac{u^2}{2\beta^2}}$$

Ne viene che la PDF ϕ della somma ξ sarà data da

$$\phi(v) = \frac{1}{\sqrt{2\pi}|\alpha|} \frac{1}{\sqrt{2\pi}|\beta|} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2\alpha^2}} e^{-\frac{(v-u)^2}{2\beta^2}} du$$

e si calcola

$$\begin{aligned}
 \phi(v) &= \frac{1}{2\pi|\alpha\beta|} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2\alpha^2} - \frac{v^2}{2\beta^2} - \frac{uv}{\beta^2} + \frac{uv}{\beta^2}} du = \\
 &= \frac{1}{2\pi|\alpha\beta|} e^{-\frac{v^2}{2\beta^2}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2\alpha^2} - \frac{uv}{\beta^2} + \frac{uv}{\beta^2}} du = \\
 &= \frac{1}{2\pi|\alpha\beta|} e^{-\frac{v^2}{2\beta^2}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2\alpha^2\beta^2} + \frac{uv}{\beta^2}} du = \\
 &= \frac{1}{2\pi|\alpha\beta|} e^{-\frac{v^2}{2\beta^2}} \int_{-\infty}^{+\infty} e^{-\left(\frac{u}{\sqrt{2}|\alpha\beta|} - \frac{v}{\sqrt{2}\beta^2}|\alpha\beta|\right)^2 + \frac{v^2}{2\beta^4}\alpha^2\beta^2} du = \\
 &= \frac{1}{2\pi|\alpha\beta|} e^{-\frac{v^2}{2\beta^2} + \frac{v^2}{2\beta^4}\alpha^2\beta^2} \int_{-\infty}^{+\infty} e^{-\left(\frac{u}{\sqrt{2}|\alpha\beta|} - \frac{v}{\sqrt{2}\beta^2}|\alpha\beta|\right)^2} du
 \end{aligned}$$

Ora, tenendo conto che

$$-\frac{1}{2\beta^2} + \frac{1}{2\beta^4}\alpha^2\beta^2 = -\frac{1}{2}$$

e calcolato l'integrale ponendo

$$t = \frac{u}{\sqrt{2}|\alpha\beta|} - \frac{v}{\sqrt{2}\beta^2}|\alpha\beta|$$

si conclude

$$\phi(v) = \frac{1}{2\pi|\alpha\beta|} e^{-\frac{v^2}{2}} \sqrt{2\pi}|\alpha\beta| = \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}$$

e ξ è a sua volta normale e standard.

Chiaramente qualora $\alpha^2 + \beta^2 \neq 1$ possiamo considerare che

$$\alpha\xi_1 + \beta\xi_2 = \sqrt{\alpha^2 + \beta^2} \left(\frac{\alpha}{\sqrt{\alpha^2 + \beta^2}}\xi_1 + \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}\xi_2 \right)$$

inoltre, se η_1 e η_2 sono gaussiane di media μ_i e varianza σ_i^2 possiamo scrivere che

$$\begin{aligned} \alpha\eta_1 + \beta\eta_2 &= \alpha\mu_1 + \alpha\sigma_1\xi_1 + \beta\mu_2 + \beta\sigma_2\xi_2 = \\ &= \alpha\mu_1 + \beta\mu_2 + (\alpha\sigma_1\xi_1 + \beta\sigma_2\xi_2) = \\ \alpha\mu_1 + \beta\mu_2 + \sqrt{\alpha^2\sigma_1^2 + \beta^2\sigma_2^2} &\left(\frac{\alpha\sigma_1}{\sqrt{\alpha^2\sigma_1^2 + \beta^2\sigma_2^2}}\xi_1 + \frac{\beta\sigma_2}{\sqrt{\alpha^2\sigma_1^2 + \beta^2\sigma_2^2}}\xi_2 \right) = \\ &\alpha\mu_1 + \beta\mu_2 + \sqrt{\alpha^2\sigma_1^2 + \beta^2\sigma_2^2}\xi \end{aligned}$$

essendo ξ una gaussiana normale standard.

Ne segue che $\alpha\eta_1 + \beta\eta_2$ ha una distribuzione gaussiana di media $\alpha\mu_1 + \beta\mu_2$ e varianza $\alpha^2\sigma_1^2 + \beta^2\sigma_2^2$ in quanto

$$\left(\frac{\alpha\sigma_1}{\sqrt{\alpha^2\sigma_1^2 + \beta^2\sigma_2^2}}\xi_1 + \frac{\beta\sigma_2}{\sqrt{\alpha^2\sigma_1^2 + \beta^2\sigma_2^2}}\xi_2 \right)$$

è una gaussiana standard.

Le distribuzioni legate ai test statistici.

La variabile aleatoria χ^2 a ν gradi di libertà restituisce la somma dei quadrati di ν variabili aleatorie ξ_i indipendenti, aventi distribuzione gaussiana con media 0 e varianza 1 (distribuzioni normali standardizzate).

$$\chi^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_\nu^2$$

Per ricavare la PDF della distribuzione χ^2 a ν gradi di libertà è opportuno procedere come segue.

Se ξ è una variabile aleatoria gaussiana standard e se $\eta = \xi^2$, la PDF φ di η è data da

$$\varphi(t) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{s}{2}}}{\sqrt{s}} & s \geq 0 \\ 0 & s < 0 \end{cases}$$

Infatti si ha

$$\begin{aligned} \mathcal{P}(\eta \leq \alpha) &= \mathcal{P}(\xi^2 \leq \alpha) = \mathcal{P}(-\sqrt{\alpha} \leq \xi \leq \sqrt{\alpha}) = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{\alpha}}^{\sqrt{\alpha}} e^{-\frac{t^2}{2}} dt = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{\alpha}} e^{-\frac{t^2}{2}} dt = \end{aligned}$$

posto $s = t^2$, per cui $ds = 2t dt$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\alpha} \frac{e^{-\frac{s}{2}}}{2\sqrt{s}} ds = \frac{1}{\sqrt{2\pi}} \int_0^{\alpha} \frac{e^{-\frac{s}{2}}}{\sqrt{s}} ds$$

Se η_1, η_2 sono variabili aleatorie $\eta_i = \xi_i^2$ con ξ_i gaussiana standard e se $\eta = \eta_1 + \eta_2$, la PDF φ di η è data da

$$\varphi(t) = \begin{cases} \frac{1}{2} e^{-\frac{s}{2}} & s \geq 0 \\ 0 & s < 0 \end{cases}$$

Infatti, se chiamiamo φ_i la PDF di η_i , si ha

$$\begin{aligned} \varphi(s) &= \int_{-\infty}^{+\infty} \varphi_1(t) \varphi_2(s-t) dt = \\ &= \int_0^s \varphi_1(t) \varphi_2(s-t) dt = \frac{1}{2\pi} \int_0^s \frac{e^{-\frac{t}{2}} e^{-\frac{s-t}{2}}}{\sqrt{t} \sqrt{s-t}} dt = \\ &= \frac{1}{2\pi} e^{-\frac{s}{2}} \int_0^s \frac{1}{\sqrt{t} \sqrt{s-t}} dt \end{aligned}$$

Posto $\sqrt{t(s-t)} = ut$ si ha

$$t(s-t) = u^2 t^2, \quad s-t = u^2 t, \quad s = (1+u^2)t$$

$$t = \frac{s}{1+u^2}, \quad dt = -\frac{2us}{(1+u^2)^2} du$$

per cui

$$\begin{aligned} \int \frac{1}{\sqrt{t(s-t)}} dt &= \int -\frac{2us}{ut(1+u^2)^2} du = \\ &= \int -\frac{2us}{u \frac{s}{1+u^2} (1+u^2)^2} du = -2 \int \frac{1}{1+u^2} du = \\ &= -2 \arctan(u) = -2 \arctan \sqrt{\frac{s-t}{t}} \end{aligned}$$

Ne viene

$$\int_0^s \frac{1}{\sqrt{t}\sqrt{s-t}} dt = -2 \arctan \sqrt{\frac{s-t}{t}} \Big|_0^s = 2 \frac{\pi}{2} = \pi$$

e si può dedurre che

$$\varphi(s) = \frac{1}{2\pi} e^{-\frac{s}{2}} \pi = \frac{1}{2} e^{-\frac{s}{2}}$$

è la distribuzione di η .

Possiamo ora provare che la variabile aleatoria

$$\eta = \eta_1 + \eta_2 + \dots + \eta_\nu = \xi_1^2 + \xi_2^2 + \dots + \xi_\nu^2$$

dove ξ_i sono variabili aleatorie gaussiane normalizzate, ha una PDF definita da

$$\varphi_\nu(u) = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} u^{\frac{\nu}{2}-1} e^{-\frac{u}{2}}$$

Le verifiche fatte precedentemente consentono di affermare che la precedente affermazione è vera per $\nu = 1$ e per $\nu = 2$. Pertanto per verificare la medesima è sempre vera sarà sufficiente provare che, supposta vera per ν è vera anche per $\nu + 2$.

Si ha

$$\begin{aligned}
 \varphi_{\nu+2}(u) &= \varphi_{\nu}(u) * \varphi_2(u) = \int_0^u \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} (u-t)^{\frac{\nu}{2}-1} e^{-\frac{u-t}{2}} \frac{1}{2} e^{-\frac{t}{2}} dt = \\
 &= \frac{1}{2^{\frac{\nu}{2}+1} \Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{u}{2}} \int_0^u (u-t)^{\frac{\nu}{2}-1} du = \frac{1}{2^{\frac{\nu}{2}+1} \Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{u}{2}} \left[-\frac{(u-t)^{\frac{\nu}{2}}}{\frac{\nu}{2}} \right]_0^u = \\
 &= \frac{1}{2^{\frac{\nu}{2}+1} \Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{u}{2}} \frac{(u)^{\frac{\nu}{2}}}{\frac{\nu}{2}} = \frac{1}{2^{\frac{\nu}{2}+1} \Gamma\left(\frac{\nu}{2} + 1\right)} u^{\frac{\nu}{2}} e^{-\frac{u}{2}}
 \end{aligned}$$

Media, varianza della variabile aleatoria χ^2 sono date da

$$\begin{aligned}\mu &= \nu \\ \sigma^2 &= 2\nu\end{aligned}$$

Infatti

$$\begin{aligned}\mu &= \int_0^{+\infty} \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} u u^{\frac{\nu}{2}-1} e^{-\frac{u}{2}} du = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} u^{\frac{\nu}{2}} e^{-\frac{u}{2}} du = \\ &= \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} (2t)^{\frac{\nu}{2}} e^{-t} 2dt \frac{2^{\frac{\nu}{2}+1} \Gamma\left(\frac{\nu}{2} + 1\right)}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} = \\ &= \frac{2^{\frac{\nu}{2}+1} \frac{\nu}{2} \Gamma\left(\frac{\nu}{2}\right)}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} = \nu\end{aligned}$$

mentre

$$\begin{aligned}
\int_0^{+\infty} \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} u^2 u^{\frac{\nu}{2}-1} e^{-\frac{u}{2}} du &= \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} u^{\frac{\nu}{2}+1} e^{-\frac{u}{2}} du = \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} (2t)^{\frac{\nu}{2}+1} e^{-t} 2 dt \frac{2^{\frac{\nu}{2}+2} \Gamma\left(\frac{\nu}{2}+2\right)}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} = \\
&= \frac{2^{\frac{\nu}{2}+2} \left(\frac{\nu}{2}+1\right) \frac{\nu}{2} \Gamma\left(\frac{\nu}{2}\right)}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} = 4\left(\frac{\nu}{2}+1\right) \frac{\nu}{2}
\end{aligned}$$

e

$$\sigma^2 = 4\left(\frac{\nu}{2}+1\right) \frac{\nu}{2} - \nu^2 = 2\nu$$

Inoltre la generatrice dei momenti della variabile aleatoria χ^2 è

$$M_{\chi^2}(t) = (1 - 2t)^{-\nu/2}$$

È la distribuzione di una variabile aleatoria T che restituisce il rapporto

$$T = \frac{\xi}{\sqrt{\frac{\eta}{\nu}}}$$

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
Per ricavare la PDF di student cominciamo con l'osservare che:

Se ξ è una variabile aleatoria con densità χ^2 a ν gradi di libertà, la sua PDF sarà data da

$$\varphi_{\nu}(u) = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} u^{\frac{\nu}{2}-1} e^{-\frac{u}{2}}$$

Sia η la variabile aleatoria definita da

$$\eta = \sqrt{\frac{\xi}{\nu}}$$

Avremo che

$$\begin{aligned} \mathcal{P}(\eta \leq \alpha) &= \mathcal{P}\left(\sqrt{\frac{\xi}{\nu}} \leq \alpha\right) = \mathcal{P}(\xi \leq \nu\alpha^2) = \\ &= \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{\nu\alpha^2} u^{\frac{\nu}{2}-1} e^{-\frac{u}{2}} du = \end{aligned}$$

posto $t = \sqrt{\frac{u}{\nu}}$ da cui $u = \nu t^2$ e $du = 2\nu t dt$

$$= \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{\alpha} t^{\nu-2} \nu^{\frac{\nu}{2}-1} e^{-\frac{\nu t^2}{2}} 2\nu t dt = \frac{1}{2^{\frac{\nu}{2}-1} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{\alpha} t^{\nu-1} \nu^{\frac{\nu}{2}} e^{-\frac{\nu t^2}{2}} dt$$

e possiamo concludere che la PDF di η è data da

$$\varphi(t) = \begin{cases} \frac{1}{2^{\frac{\nu}{2}-1}\Gamma(\frac{\nu}{2})} t^{\nu-1} \nu^{\frac{\nu}{2}} e^{-\frac{\nu t^2}{2}} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Sia ora η_1 una variabile aleatoria gaussiana standard, la cui PDF è ovviamente data da

$$g_1(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

e sia $\eta_2 = \sqrt{\frac{\xi}{\nu}}$ dove ξ è una variabile aleatoria con densità χ^2 a ν gradi di libertà; per quanto detto in precedenza la PDF di η_2 è nulla prima di 0 ed è data da

$$g_2(t) = \frac{1}{2^{\frac{\nu}{2}-1} \Gamma\left(\frac{\nu}{2}\right)} t^{\nu-1} \nu^{\frac{\nu}{2}} e^{-\frac{\nu t^2}{2}}$$

per $t \geq 0$.

La variabile aleatoria T di student è definita mediante la

$$T = \frac{\eta_1}{\eta_2}$$

e possiamo ricavarne la PDF φ ricordando che si tratta del quoziente di due variabili aleatorie di cui conosciamo la densità di probabilità.

Avremo:

$$\begin{aligned}
 \varphi(t) &= \int_0^{+\infty} x \frac{1}{2^{\frac{\nu}{2}-1} \Gamma\left(\frac{\nu}{2}\right)} x^{\nu-1} \nu^{\frac{\nu}{2}} e^{-\frac{\nu x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2 x^2}{2}} dx = \\
 &= \frac{\nu^{\frac{\nu}{2}}}{\sqrt{2\pi} 2^{\frac{\nu}{2}-1} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} x^{\nu} e^{-\frac{x^2}{2}(\nu+t^2)} dx = \\
 \text{posto } \frac{x^2}{2}(\nu+t^2) &= s \text{ si ha } x = \sqrt{\frac{2s}{\nu+t^2}} \text{ e } dx = \frac{\sqrt{2} ds}{\sqrt{\nu+t^2} \sqrt{s}} \\
 &= \frac{\nu^{\frac{\nu}{2}}}{\sqrt{2\pi} 2^{\frac{\nu}{2}-1} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} \frac{2^{\frac{\nu}{2}} s^{\frac{\nu}{2}}}{(\nu+t^2)^{\frac{\nu}{2}}} e^{-s} \frac{\sqrt{2}}{\sqrt{\nu+t^2}} \frac{1}{2\sqrt{s}} ds = \\
 &= \frac{\nu^{\frac{\nu}{2}} 2^{\frac{\nu}{2}-\frac{1}{2}}}{\sqrt{2\pi} 2^{\frac{\nu}{2}-1} \Gamma\left(\frac{\nu}{2}\right) (\nu+t^2)^{\frac{\nu}{2}}} \int_0^{+\infty} s^{\frac{\nu}{2}-\frac{1}{2}} e^{-s} ds = \\
 &= \frac{\nu^{\frac{\nu}{2}}}{\sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right)} \frac{1}{(\nu)^{\frac{\nu+1}{2}} (1+\frac{t^2}{\nu})^{\frac{\nu+1}{2}}} \int_0^{+\infty} s^{\frac{\nu+1}{2}-1} e^{-s} ds = \\
 &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \Gamma\left(\frac{\nu}{2}\right) \left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}}
 \end{aligned}$$

Pertanto la funzione di distribuzione della variabile aleatoria T a ν gradi di libertà è data da:

$$\varphi(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

La media e la varianza di T risultano essere

$$\mu = 0$$

$$\sigma^2 = \frac{\nu}{\nu - 2}$$

Infatti

$$\mu = \int_{-\infty}^{+\infty} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} t \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt = 0$$

per la simmetria dell'integranda;

Inoltre

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{+\infty} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} t^2 \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt = \\ &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \int_{-\infty}^{+\infty} \frac{\nu t}{2} \left(-\frac{2}{\nu-1}\right) \frac{d}{dt} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu-1}{2}} dt =\end{aligned}$$

integrando per parti

$$\begin{aligned}&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(-\left(\frac{\nu t}{\nu-1}\right) \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu-1}{2}} \right) \Big|_{-\infty}^{+\infty} + \frac{\nu}{\nu-1} \int_{-\infty}^{+\infty} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu-1}{2}} dt = \\ &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \frac{\nu}{\nu-1} \int_{-\infty}^{+\infty} \left(1 + \frac{t^2}{\nu-2} \frac{\nu-2}{\nu}\right)^{-\frac{\nu-1}{2}} dt =\end{aligned}$$

posto $s = t\sqrt{\frac{\nu-2}{\nu}}$

$$\begin{aligned}&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \frac{\nu}{\nu-1} \int_{-\infty}^{+\infty} \left(1 + \frac{s^2}{\nu-2}\right)^{-\frac{\nu-1}{2}} \sqrt{\frac{\nu}{\nu-2}} ds = \\ &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \frac{\nu}{\nu-1} \sqrt{\frac{\nu}{\nu-2}} \frac{\sqrt{(\nu-2)\pi}\Gamma\left(\frac{\nu-2}{2}\right)}{\Gamma\left(\frac{\nu-1}{2}\right)} = \\ &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\nu}{\nu-1} \frac{\Gamma\left(\frac{\nu-2}{2}\right)}{\Gamma\left(\frac{\nu-1}{2}\right)} = \frac{\frac{\nu}{2} - \frac{1}{2}}{\frac{\nu}{2} - 1} = \frac{\nu}{\nu-2}\end{aligned}$$

È la distribuzione di una variabile aleatoria F che restituisce il rapporto

$$F = \frac{\eta/\mu}{\xi/\nu}$$

dove η ed ξ sono variabili aleatorie con distribuzione χ^2 a μ e ν gradi di libertà, rispettivamente.

Per ricavare la PDF di F ricordiamo che la PDF g di η/μ è data da

$$g(u) = \frac{\mu}{2^{\frac{\mu}{2}} \Gamma\left(\frac{\mu}{2}\right)} (\mu u)^{\frac{\mu}{2}-1} e^{-\frac{\mu u}{2}}$$

mentre la PDF f di ξ/ν è data da

$$f(u) = \frac{\nu}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} (\nu u)^{\frac{\nu}{2}-1} e^{-\frac{\nu u}{2}}$$

Possiamo quindi ricavare la PDF φ di $F = \frac{\eta/\mu}{\xi/\nu}$ usando quanto conosciamo sul rapporto di due variabili aleatorie.

Avremo

$$\begin{aligned}
 \varphi(t) &= \int_0^{+\infty} x \frac{\nu}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \nu^{\frac{\nu}{2}-1} x^{\frac{\nu}{2}-1} e^{-\frac{\nu x}{2}} \frac{\mu}{2^{\frac{\mu}{2}} \Gamma\left(\frac{\mu}{2}\right)} \mu^{\frac{\mu}{2}-1} (tx)^{\frac{\mu}{2}-1} e^{-\frac{\mu tx}{2}} dx = \\
 &= \frac{\nu^{\frac{\nu}{2}} \mu^{\frac{\mu}{2}}}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\mu}{2}} \Gamma\left(\frac{\mu}{2}\right)} \int_0^{+\infty} x^{\frac{\nu}{2}} e^{-\frac{\nu x}{2}} t^{\frac{\mu}{2}-1} x^{\frac{\mu}{2}-1} e^{-\frac{\mu tx}{2}} dx = \\
 &= t^{\frac{\mu}{2}-1} C_{\nu,\mu} \int_0^{+\infty} x^{\frac{\nu+\mu}{2}-1} e^{-\frac{x}{2}(\nu+t\mu)} dx =
 \end{aligned}$$

posto $\frac{x}{2}(\nu + t\mu) = s$ si ha $dx = \frac{2}{\nu+t\mu} ds$

$$\begin{aligned}
 &= t^{\frac{\mu}{2}-1} C_{\nu,\mu} \int_0^{+\infty} \left(\frac{2s}{\nu+t\mu}\right)^{\frac{\nu+\mu}{2}-1} e^{-s} \frac{2}{\nu+t\mu} ds = \\
 &= \frac{t^{\frac{\mu}{2}-1} C_{\nu,\mu} 2^{\frac{\nu+\mu}{2}}}{(\nu+t\mu)^{\frac{\nu+\mu}{2}}} \int_0^{+\infty} s^{\frac{\nu+\mu}{2}-1} e^{-s} ds = \\
 &= \frac{\nu^{\frac{\nu}{2}} \mu^{\frac{\mu}{2}} \Gamma\left(\frac{\nu+\mu}{2}\right) 2^{\frac{\nu+\mu}{2}}}{2^{\frac{\nu}{2}} 2^{\frac{\mu}{2}} \Gamma\left(\frac{\nu}{2}\right) \Gamma\left(\frac{\mu}{2}\right)} \frac{t^{\frac{\mu}{2}-1}}{(\nu+t\mu)^{\frac{\nu+\mu}{2}}} = \\
 &= \frac{\nu^{\frac{\nu}{2}} \mu^{\frac{\mu}{2}} \Gamma\left(\frac{\nu+\mu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \Gamma\left(\frac{\mu}{2}\right)} t^{\frac{\mu}{2}-1} (\nu+t\mu)^{-\frac{\nu+\mu}{2}}
 \end{aligned}$$

Pertanto la funzione di distribuzione della variabile aleatoria F è data da

$$\varphi(t) = \begin{cases} \frac{\nu^{\frac{\nu}{2}} \mu^{\frac{\mu}{2}} \Gamma(\frac{\nu+\mu}{2})}{\Gamma(\frac{\nu}{2}) \Gamma(\frac{\mu}{2})} t^{\frac{\mu}{2}-1} (\nu + t\mu)^{-\frac{\nu+\mu}{2}} & t > 0 \\ 0 & \text{altrimenti} \end{cases}$$

La media e varianza risultano essere

$$\mu = \frac{\nu_2}{\nu_2 - 2}$$

$$\sigma^2 = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)^2}$$

Infatti posto

$$C(\mu, \nu) = \frac{\nu^{\frac{\nu}{2}} \mu^{\frac{\mu}{2}} \Gamma\left(\frac{\nu+\mu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \Gamma\left(\frac{\mu}{2}\right)}$$

$$\begin{aligned} \int_0^{+\infty} t^{k+1} \varphi(t) dt &= C(\mu, \nu) \int_0^{+\infty} t^{\frac{\mu}{2}+k} (\nu + t\mu)^{-\frac{\nu+\mu}{2}} dt = C(\mu, \nu) \int_0^{+\infty} t^{\frac{\mu}{2}+k} \left(\nu \left(1 + t \frac{\mu}{\nu} \right) \right)^{-\frac{\nu+\mu}{2}} dt = \\ &= C(\mu, \nu) \nu^{-\frac{\nu+\mu}{2}} \int_0^{+\infty} t^{\frac{\mu}{2}+k} \left(1 + t \frac{\mu+2k+2}{\nu-2k-2} \frac{\nu-2k-2}{\mu+2k+2} \frac{\mu}{\nu} \right)^{-\frac{\nu+\mu}{2}} dt = \end{aligned}$$

posto $s = \frac{\nu-2k-2}{\mu+2k+2} \frac{\mu}{\nu} t$

$$= C(\mu, \nu) \nu^{-\frac{\nu+\mu}{2}} \int_0^{+\infty} s^{\frac{\mu}{2}+k} \left(\frac{\nu}{\mu} \frac{\mu+2k+2}{\nu-2k-2} \right)^{\frac{\mu}{2}+k+1} \left(1 + s \frac{\mu+2k+2}{\nu-2k-2} \right)^{-\frac{\nu+\mu}{2}} dt =$$

$$= C(\mu, \nu) \nu^{-\frac{\nu+\mu}{2}} \left(\frac{\nu}{\mu} \frac{\mu+2k+2}{\nu-2k-2} \right)^{\frac{\mu}{2}+k+1} \int_0^{+\infty} s^{\frac{\mu}{2}+k} \left(1 + s \frac{\mu+2k+2}{\nu-2k-2} \right)^{-\frac{\nu+\mu}{2}} dt =$$

$$= \nu^{-\frac{\nu+\mu}{2}} \left(\frac{\nu}{\mu} \frac{\mu+2k+2}{\nu-2k-2} \right)^{\frac{\mu}{2}+k+1} \frac{C(\mu, \nu)}{C(\mu+2k+2, \nu-2k-2)} =$$

$$\begin{aligned} &= \frac{\nu^{\frac{\nu}{2}} \mu^{\frac{\mu}{2}} \nu^{-\frac{\nu+\mu}{2}} \nu^{\frac{\mu}{2}+k+1} (\mu+2k+2)^{\frac{\mu}{2}+k+1}}{\mu^{\frac{\mu}{2}+k+1} (\nu-2k-2)^{\frac{\mu}{2}+k+1} (\nu-2k-2)^{\frac{\nu}{2}-k-1} (\mu+2k+2)^{\frac{\mu}{2}+k+1}} \\ &= \frac{\Gamma\left(\frac{\nu+\mu}{2}\right) \Gamma\left(\frac{\nu}{2} - k - 1\right) \Gamma\left(\frac{\mu}{2} + k + 1\right)}{\Gamma\left(\frac{\nu}{2}\right) \Gamma\left(\frac{\mu}{2}\right) \Gamma\left(\frac{\nu+\mu}{2}\right)} = \frac{\nu^{k+1} \Gamma\left(\frac{\nu}{2} - k - 1\right) \Gamma\left(\frac{\mu}{2} + k + 1\right)}{\mu^{k+1} \Gamma\left(\frac{\nu}{2}\right) \Gamma\left(\frac{\mu}{2}\right)} \end{aligned}$$

Siamo così in grado di calcolare i momenti rispetto all'origine di ogni ordine

Per $k = 0$ otteniamo il momento di ordine 1, Cioè la media,

$$\int_0^{+\infty} t\varphi(t)dt = \frac{\nu}{\mu} \frac{\Gamma\left(\frac{\nu}{2} - 1\right) \frac{\mu}{2} \Gamma\left(\frac{\mu}{2}\right)}{\Gamma\left(\frac{\nu}{2} - 1\right) \Gamma\left(\frac{\nu}{2} - 1\right) \Gamma\left(\frac{\mu}{2}\right)} = \frac{\nu}{\nu - 2}$$

mentre per $k = 1$ possiamo calcolare la varianza come segue

$$\begin{aligned} \int_0^{+\infty} t^2\varphi(t)dt &= \frac{\nu^2}{\mu^2} \frac{\Gamma\left(\frac{\nu}{2} - 2\right) \frac{\mu}{2} \left(\frac{\mu}{2} + 1\right) \Gamma\left(\frac{\mu}{2}\right)}{\Gamma\left(\frac{\nu}{2} - 1\right) \left(\frac{\nu}{2} - 2\right) \Gamma\left(\frac{\nu}{2} - 2\right) \Gamma\left(\frac{\mu}{2}\right)} = \\ &= \frac{\nu^2}{\mu^2} \frac{\mu(\mu + 2)}{(\nu - 2)(\nu - 4)} \end{aligned}$$

$$\sigma^2 = \frac{\nu^2}{\mu^2} \frac{\mu(\mu + 2)}{(\nu - 2)(\nu - 4)} - \left(\frac{\nu}{\nu - 2}\right)^2 = \frac{2\nu^2(\mu + \nu - 2)}{\mu(\nu - 2)(\nu - 4)}$$

Il quadrato T^2 di una variabile aleatoria di Student a ν gradi di libertà è una variabile aleatoria di Fisher con $(1, \nu)$ gradi di libertà, infatti T è una variabile aleatoria che restituisce il rapporto

$$T = \frac{\xi}{\sqrt{\frac{\eta}{\nu}}}$$

dove ξ è una gaussiana normale ed η ha distribuzione χ^2 a ν gradi di libertà, per cui

$$\mathcal{P}(T^2 \leq \alpha) = \mathcal{P}\left(\frac{\xi^2}{\frac{\eta}{\nu}} \leq \alpha\right)$$

Osservando che ξ^2 , in quanto quadrato di una gaussiana normale, ha una distribuzione χ^2 a un grado di libertà, possiamo concludere che T^2 è una distribuzione di Fisher a $(1, \nu)$ gradi di libertà.

La distribuzione γ è definita, per $\alpha, \beta > 0$ da

$$\varphi(t) = \begin{cases} \frac{t^{\alpha-1} e^{-t/\beta}}{\beta^{\alpha} \Gamma(\alpha)} & t > 0 \\ 0 & \text{altrimenti} \end{cases}$$

La presenza dei due parametri α e β consente di adattare la distribuzione ai dati che si desidera rappresentare. In particolare il parametro α descrive la forma della distribuzione mentre β è semplicemente un fattore di scala.

Media, varianza e generatrice dei momenti sono date da

$$\mu = \alpha\beta$$

$$\sigma^2 = \alpha\beta^2$$

$$M_\gamma(t) = (1 - \beta t)^{-\alpha} \text{ per } t < \frac{1}{\beta}$$

Infatti

$$\begin{aligned} \mu &= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{+\infty} t^\alpha e^{-t/\beta} dt = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{+\infty} (s\beta)^\alpha e^{-s} \beta ds = \\ &= \frac{\beta^{\alpha+1} \Gamma(\alpha+1)}{\beta^\alpha \Gamma(\alpha)} = \alpha\beta \end{aligned}$$

mentre

$$\begin{aligned} \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{+\infty} t^{\alpha+1} e^{-t/\beta} dt &= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{+\infty} (s\beta)^{\alpha+1} e^{-s} \beta ds = \\ &= \frac{\beta^{\alpha+2} \Gamma(\alpha+2)}{\beta^\alpha \Gamma(\alpha)} = \alpha\beta^2(\alpha+1) \end{aligned}$$

Inoltre

$$\begin{aligned}
 M_{\gamma}(t) &= \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^{+\infty} e^{tx} x^{\alpha-1} e^{-x/\beta} dx = \\
 &= \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^{+\infty} \left(\frac{1}{\left(\frac{1}{\beta} - t\right)} \right)^{\alpha-1} e^{-s} \frac{1}{\frac{1}{\beta} - t} ds = \\
 &= \frac{\beta^{\alpha}\Gamma(\alpha)}{\beta^{\alpha}\Gamma(\alpha) \left(\frac{1}{\beta} - t\right)^{\alpha}} = \frac{1}{\left(\frac{1}{\beta} - t\right)^{\alpha}}
 \end{aligned}$$

Ovviamente il precedente integrale converge solo per $t < \frac{1}{\beta}$.
 La moda infine è $(\alpha - 1)\beta$ per $\alpha > 1$, come si deduce immediatamente calcolando il punto di massimo della funzione $t^{\alpha-1}e^{-t/\beta}$ la cui derivata

$$(\alpha - 1)t^{\alpha-2}e^{-t/\beta} - \frac{1}{\beta}t^{\alpha-1}e^{-t/\beta} = t^{\alpha-2}e^{-t/\beta} \left((\alpha - 1) - \frac{1}{\beta} \right) = 0$$

per

$$t = (\alpha - 1)\beta$$

nel caso in cui $\alpha > 1$

Nel caso in cui $\alpha = k \in \mathbb{N}$ e $\beta = 1$ la distribuzione γ definisce una variabile aleatoria che è somma di k variabili aleatorie esponenziali di media $\lambda = \frac{1}{\beta}$.

Infatti se consideriamo due variabili aleatorie con densità esponenziale di media λ la loro somma avrà una densità di probabilità definita dalla convoluzione

$$\begin{aligned}\psi(t) &= \int_{-\infty}^{+\infty} \lambda e^{-\lambda s} \lambda e^{-\lambda(t-s)} ds = \int_0^t \lambda e^{-\lambda s} \lambda e^{-\lambda(t-s)} ds = \\ & \lambda^2 e^{-\lambda t} \int_0^t ds = t \lambda^2 e^{-\lambda t}\end{aligned}$$

Inoltre se sommiamo due variabili aleatorie con densità nulla prima di 0 e uguale a

$$\frac{t^{k-1} \lambda^k e^{-\lambda t}}{(k-1)!} \quad \text{e} \quad \lambda e^{-\lambda t}$$

altrove, otterremo una nuova variabile aleatoria la cui PDF si ottiene per convoluzione nella forma

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{1}{(k-1)!} s^{k-1} \lambda^k e^{-\lambda s} \lambda e^{-\lambda(t-s)} ds &= \\ &= \frac{1}{(k-1)!} \lambda^{k+1} e^{-\lambda t} \int_0^t s^{k-1} ds = \\ &= \frac{1}{k!} \lambda^{k+1} t^k e^{-\lambda t} \end{aligned}$$

La distribuzione β è definita da

$$\varphi(t) = \begin{cases} \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)} & 0 < t < 1 \\ 0 & \text{altrimenti} \end{cases} = \quad (13)$$

$$= \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1}(1-t)^{\beta-1} & 0 < t < 1 \\ 0 & \text{altrimenti} \end{cases} \quad (14)$$

Anche in questo caso si tratta di una distribuzione che, in virtù dei due parametri da cui dipende, ben si presta a descrivere una variabile aleatoria di cui si conoscano le caratteristiche.

I momenti della distribuzione β rispetto all'origine si calcolano facilmente non appena si ricordi che

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

Si ha infatti

$$\begin{aligned}\mu_k &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^{+\infty} t^{\alpha-1+k}(1-t)^{\beta-1} dt = \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + k)\Gamma(\beta)}{\Gamma(\alpha + \beta + k)} = \prod_{n=0}^{k-1} \frac{\alpha + n}{\alpha + \beta + n}\end{aligned}$$

$\mu_1 = \mu$ fornisce la media e $\mu_2 - \mu_1^2 = \sigma^2$ fornisce la varianza che sono quindi date da:

$$\mu = \frac{\alpha}{\alpha + \beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)(\alpha + \beta + 1)}$$

La funzione generatrice dei momenti si può poi ricavare ricordando che i momenti sono i coefficienti del suo sviluppo di McLaurin.

$$M_\beta(t) = 1 + \sum_{k=1}^{+\infty} \left(\prod_{n=0}^{k-1} \frac{\alpha + n}{\alpha + \beta + n} \right) \frac{t^k}{k!}$$

Cercando il punto di massimo di $t^{\alpha-1}(1-t)^{\beta-1}$ si vede poi che la moda è

$$\frac{\alpha - 1}{\alpha + \beta - 2}$$

La distribuzione β , per $\alpha = n + 1, \beta = m + 1 \in \mathbb{N}$ diventa

$$\varphi(t) = \begin{cases} \frac{(n+m+1)!}{n! m!} t^n (1-t)^m & 0 < t < 1 \\ 0 & \text{altrimenti} \end{cases}$$

e rappresenta la probabilità di aver ottenuto in una serie di $n + m$ prove bernoulliane ripetute, n successi ed m insuccessi, posto che sia t la probabilità di successo della prova bernoulliana stessa. Possiamo dedurre da questa osservazione che il valore di t per cui questa probabilità è massima è $t = \frac{n}{n+m}$; infatti

$$\begin{aligned} (t^n(1-t)^m)' &= nt^{n-1}(1-t)^m + mt^n(1-t)^{m-1} = \\ &= (mt + (1-t)n) (t^{n-1}(1-t)^{m-1}) = 0 \iff t = \frac{n}{n+m} \end{aligned}$$

Possiamo quindi dire che in presenza di una sequenza di esiti bernoulliani con n successi ed m insuccessi è ragionevole supporre uguale a $\frac{n}{n+m}$ la probabilità di successo del singolo evento.

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

$$F(t) = \int_0^t f(s)ds = \frac{t}{2\pi} \quad , \quad G(t) = \int_0^t g(s)ds = 1 - e^{-t^2/2}$$

Si ha allora

$$F^{-1}(s) = 2\pi s \quad , \quad G^{-1}(s) = \sqrt{-2 \ln(1 - s)}$$

e le variabili

$$\xi = \sqrt{-2 \ln(1 - t)} \cos(2\pi s) \quad , \quad \eta = \sqrt{-2 \ln(1 - t)} \sin(2\pi s)$$

dove s e t sono distribuite uniformemente, risultano distribuite con densità gaussiana di media 0 e di varianza 1.

Diciamo che è assegnata una popolazione se è assegnato un insieme $\mathcal{U}$ ed una variabile aleatoria ξ definita su $\mathcal{U}$.

Se ad esempio siamo interessati a stimare il diametro di sfere di acciaio per cuscinetti, possiamo considerare la popolazione i cui elementi sono le sfere prodotte e la variabile aleatoria ξ che associa ad ogni sfera s_k il suo diametro $\xi(k) = d_k$. La variabile ξ che restituisce il diametro di ciascuna sfera descrive la popolazione che stiamo esaminando. ξ potrebbe ad esempio avere una funzione di distribuzione, ad esempio gaussiana di media μ e varianza σ^2 .

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_k come la variabile aleatoria che restituisce il diametro della k -esima sferetta scelta in $\mathcal{U}$.

Diciamo che è assegnato uno stimatore, o riassunto campionario, se è assegnata una funzione $\phi = \phi(x_1, x_2, \dots, x_n)$.

Possiamo ad esempio considerare uno stimatore del diametro delle sferette considerando la media dei diametri delle n sferette che abbiamo estratto per costruire il campione.

In tal caso

$$\phi(x_1, x_2, \dots, x_n) = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

Si consideri una popolazione individuata da una variabile aleatoria ξ su un insieme $\mathcal{U}$ di media μ e varianza σ ed un campione $x = (x_1, x_2, \dots, x_n)$ di taglia n .

Sia

$$\mu = E(\xi)$$

la media della popolazione e

$$\bar{x}_n = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

la media campionaria

Sia $\mu_{\bar{x}_n}$ la media della variabile aleatoria $\bar{x}_n$, cioè il valor medio della media campionaria; si ha



$$\mu_{\bar{x}_n} = E(\bar{x}_n) = \mu$$

infatti

$$E(\bar{x}_n) = \frac{1}{n} (E(x_1) + E(x_2) + \cdots E(x_n)) = \frac{n\mu}{n} = \mu$$



$$\sigma_{\bar{x}_n}^2 = E((\bar{x}_n - \mu_{\bar{x}_n})^2) = E((\bar{x}_n - \mu)^2) = \frac{\sigma^2}{n}$$

infatti, tenendo conto che le variabili x_i sono indipendenti,

$$\begin{aligned} E((\bar{x}_n - \mu_{\bar{x}_n})^2) &= E((\bar{x}_n - \mu)^2) = \\ &= \frac{E((x_1 - \mu)^2) + E((x_2 - \mu)^2) + \cdots E((x_n - \mu)^2)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

$\bar{x}_n$ è distribuito normalmente con media μ e varianza σ^2/n . Infatti, poichè ξ è distribuita normalmente con media μ e varianza σ^2 allora anche ciascuna delle x_i segue la stessa distribuzione e, dal momento che la somma di distribuzioni gaussiane è ancora gaussiana, anche $\bar{x}_n$ ha una PDF Gaussiana la cui media e varianza sono, in accordo con il punto precedente, μ e $\frac{\sigma^2}{n}$, rispettivamente. Inoltre, per il teorema del limite centrale, si ottiene che

$$\frac{\bar{x}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

è asintoticamente ($n \geq 30$) distribuito normalmente con media 0 e varianza 1.

Chiamiamo varianza campionaria la variabile aleatoria definita da

$$S_n^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

Si ha che

$$E(S_n^2) = \frac{n-1}{n} \sigma^2$$

infatti possiamo facilmente calcolare che

$$\begin{aligned} (x_i - \bar{x}_n)^2 &= (x_i - \mu + \mu - \bar{x}_n)^2 = \\ &= (x_i - \mu)^2 - 2(x_i - \mu)(\bar{x}_n - \mu) + (\bar{x}_n - \mu)^2 \end{aligned}$$

perciò

$$\begin{aligned}
 \sum_{i=1}^n (x_i - \bar{x}_n)^2 &= \sum_{i=1}^n (x_i - \mu)^2 - 2(\bar{x}_n - \mu) \sum_{i=1}^n (x_i - \mu) + n(\bar{x}_n - \mu)^2 = \\
 &= \sum_{i=1}^n (x_i - \mu)^2 - 2n(\bar{x}_n - \mu)^2 + n(\bar{x}_n - \mu)^2 = \sum_{i=1}^n (x_i - \mu)^2 - n(\bar{x}_n - \mu)^2
 \end{aligned}
 \tag{15}$$

e ne segue

$$\begin{aligned}
 E(S_n^2) &= \\
 &= \frac{1}{n} E \left(\sum_{i=1}^n (x_i - \bar{x}_n)^2 \right) = \frac{1}{n} \left(E \left(\sum_{i=1}^n (x_i - \mu)^2 \right) - nE(\bar{x}_n - \mu)^2 \right) = \\
 &= \frac{1}{n} \left(n\sigma^2 - n\frac{\sigma^2}{n} \right) = \frac{n-1}{n} \sigma^2
 \end{aligned}$$

Definiamo inoltre

$$\hat{S}_n^2 = \frac{(x_1 - \bar{x}_n)^2 + (x_2 - \bar{x}_n)^2 + \dots + (x_n - \bar{x}_n)^2}{n - 1}$$

$\hat{S}_n^2$ è una definizione alternativa di varianza campionaria per la quale si ha

$$nS_n^2 = (n - 1)\hat{S}_n^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

Possiamo dimostrare che, nel caso in cui le variabili aleatorie x_i siano Gaussiane di media μ e varianza σ^2 , la media campionaria $\bar{x}_n$ e la varianza campionaria $\hat{S}_n^2$ sono tra loro indipendenti.

Infatti, supponiamo, senza perdere di generalità, che $\mu = 0$ e $\sigma = 1$, avremo che

$$\hat{S}_n^2 = \frac{1}{n-1} \left((x_1 - \bar{x}_n)^2 + \sum_{i=2}^n (x_i - \bar{x}_n)^2 \right) =$$

essendo $\sum_{i=1}^n (x_i - \bar{x}_n) = 0$

$$\frac{1}{n-1} \left(\left(\sum_{i=2}^n (x_i - \bar{x}_n) \right)^2 + \sum_{i=2}^n (x_i - \bar{x}_n)^2 \right) =$$

Ora sia $\varphi(t) = \varphi(t_1, t_2, \dots, t_n)$ la PDF congiunta delle variabili aleatorie indipendenti gaussiane $(x_1, x_2, \dots, x_n)$

$$\varphi(t) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} \sum_{i=1}^n t_i^2}$$

e, posto $\bar{t} = \frac{1}{n} \sum t_i$, consideriamo la trasformazione di coordinate definita da:

$$\begin{cases} s_1 = \bar{t} \\ s_2 = t_2 - \bar{t} \\ \dots \\ s_n = t_n - \bar{t} \end{cases}$$

La trasformazione è lineare ed il suo Jacobiano è $\frac{1}{n}$ inoltre si ha $t_i = s_i + s_1$ per $i = 2, \dots, n$ e

$$\begin{aligned} t_1 &= \sum_{i=1}^n t_i - \sum_{i=2}^n t_i = n\bar{t} - \sum_{i=2}^n t_i = \bar{t} - \sum_{i=2}^n t_i + (n-1)\bar{t} = \\ &= \bar{t} - \sum_{i=2}^n (t_i - \bar{t}) = \bar{t} - \sum_{i=2}^n s_i = s_1 - \sum_{i=2}^n s_i \end{aligned}$$

per cui possiamo esprimere φ in funzione di s mediante la

$$\varphi(s) = \frac{n}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}((s_1 - \sum_{i=2}^n s_i)^2 + \sum_{i=2}^n (s_i + s_1)^2)}$$

e dal momento che

$$\begin{aligned}
 \left(s_1 - \sum_{i=2}^n s_i \right)^2 + \sum_{i=2}^n (s_i + s_1)^2 &= \\
 &= s_1^2 + \left(\sum_{i=2}^n s_i \right)^2 - 2s_1 \sum_{i=2}^n s_i + \sum_{i=2}^n s_i^2 + \sum_{i=2}^n s_1^2 + 2s_1 \sum_{i=2}^n s_i = \\
 &= ns_1^2 + \left(\sum_{i=2}^n s_i \right)^2 + \sum_{i=2}^n s_i^2
 \end{aligned}$$

si ha

$$\varphi(s) = \frac{n}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}ns_1^2} e^{-\frac{1}{2}\left(\left(\sum_{i=2}^n s_i\right)^2 + \sum_{i=2}^n s_i^2\right)}$$

Quindi, se $(y_1, y_2, \dots, y_n)$ è la variabile aleatoria ottenuta da $(x_1, x_2, \dots, x_n)$ mediante la trasformazione lineare indicata, si ha che y_1 è indipendente da $y_2, \dots, y_n$ e $\bar{x}$ è indipendente da $\hat{S}^2$

È un risultato notevole che la variabile aleatoria definita da $\hat{S}_n^2$ ha una distribuzione di tipo χ_{n-1}^2 , cioè ha una distribuzione di tipo χ^2 ad $(n - 1)$ gradi di libertà. Infatti, dal momento che

$$\bar{x}_{n+1} = \frac{x_{n+1} + n\bar{x}_n}{n+1} = \bar{x}_n + \frac{1}{n+1}(x_{n+1} - \bar{x}_n)$$

si ha

$$\begin{aligned} n\hat{S}_{n+1}^2 &= \sum_{i=1}^{n+1} (x_i - \bar{x}_{n+1})^2 = \sum_{i=1}^{n+1} \left((x_i - \bar{x}_n) - \frac{1}{n+1}(x_{n+1} - \bar{x}_n) \right)^2 = \\ &= \sum_{i=1}^{n+1} \left((x_i - \bar{x}_n)^2 - 2 \frac{(x_{n+1} - \bar{x}_n)(x_i - \bar{x}_n)}{n+1} + \frac{(x_{n+1} - \bar{x}_n)^2}{(n+1)^2} \right) = \\ &= \sum_{i=1}^n (x_i - \bar{x}_n)^2 + (x_{n+1} - \bar{x}_n)^2 - 2 \frac{1}{n+1} (x_{n+1} - \bar{x}_n)^2 + \frac{n+1}{(n+1)^2} (x_{n+1} - \bar{x}_n)^2 = \\ &= (n-1)\hat{S}_n^2 + \frac{n}{n+1} (x_{n+1} - \bar{x}_n)^2 \end{aligned}$$

Ora

$$\hat{S}_2^2 = \frac{1}{2}(x_1 - x_2)^2$$

ha una distribuzione χ_1^2 in quanto quadrato di una gaussiana; inoltre la formula precedente consente di verificare che $\hat{S}_{n+1}^2$ ha una distribuzione χ_n^2 se $\hat{S}_n^2$ ha una distribuzione χ_{n-1}^2 ricordando che $\frac{n}{n+1}(x_{n+1} - \bar{x}_n)^2$ è gaussiana.

Pertanto per il principio di induzione $\hat{S}_n^2$ ha una distribuzione di tipo χ_{n-1}^2 per ogni $n \in \mathbb{N}$

Possiamo quindi affermare che la variabile aleatoria

$$\xi_2 = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2$$

ha una distribuzione di tipo χ^2 ad $n - 1$ gradi di libertà.

Abbiamo quindi verificato che

$$\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

è una variabile aleatoria gaussiana standard (media 0 e varianza 1).

$$\sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2 = \frac{n}{\sigma^2} S^2 = \frac{n-1}{\sigma^2} \hat{S}^2$$

è una variabile aleatoria con distribuzione di tipo χ^2 ad $n - 1$ gradi di libertà.

Quindi

$$\sqrt{\frac{\sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2}{n-1}} = \sqrt{\frac{nS^2}{\sigma(n-1)}} = \sqrt{\frac{\hat{S}^2}{\sigma}}$$

è la radice di una variabile aleatoria di tipo χ^2 ad $n - 1$ gradi di libertà divisa per i suoi gradi di libertà

e pertanto

$$\frac{\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{\hat{S}^2}{\sigma^2}}} = \sqrt{n} \frac{\bar{x} - \mu}{\hat{S}} = \sqrt{n-1} \frac{\bar{x} - \mu}{S}$$

ha una distribuzione di Student ad $n - 1$ gradi di libertà.

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 n_1 ed n_2 rispettivamente estratti dalle due popolazioni, le variabili aleatorie

$$\frac{n_i}{\sigma_i^2} S_i^2 = \frac{n_i - 1}{\sigma_i^2} \hat{S}_i^2$$

hanno distribuzioni di tipo χ^2 ad $n_i - 1$ gradi di libertà. Poichè

$$\frac{\frac{n_1}{(n_1-1)\sigma_1^2} S_1^2}{\frac{n_2}{(n_2-1)\sigma_2^2} S_2^2} = \frac{\frac{1}{\sigma_1^2} \hat{S}_1^2}{\frac{1}{\sigma_2^2} \hat{S}_2^2}$$

possiamo affermare che

$$\frac{1}{\sigma_2^2} \sigma_1^2 \frac{1}{\hat{S}_1^2} \hat{S}_2^2$$

ha una distribuzione di probabilità di Fisher con $n_1 - 1, n_2 - 1$ gradi di libertà.

Si consideri un esperimento in cui sono possibili 2 uscite

$$A_1, A_2$$

di cui si ipotizzano le probabilità

$$p_1, p_2$$

Si supponga di eseguire n volte l'esperimento e di ottenere delle frequenze di accadimento

$$x_1, x_2$$

relative alla variabile aleatoria $\xi = (\xi_1, \xi_2)$ che ha una distribuzione binomiale

Uno stimatore delle differenza tra frequenze osservate e frequenze ipotizzate si può definire come

$$\chi^2 = \frac{(x_1 - np_1)^2}{np_1} + \frac{(x_2 - np_2)^2}{np_2} \quad (16)$$

Chiaramente x_1 ed x_2 non sono indipendenti e si ha

$$\begin{aligned} \chi^2 &= \frac{(x_1 - np_1)^2}{np_1} + \frac{(x_2 - np_2)^2}{np_2} = \frac{(x_1 - np_1)^2}{np_1 p_2} p_2 + \frac{(x_2 - np_2)^2}{np_2 p_1} p_1 = \\ &= \frac{(x_1 - np_1)^2 p_2 + ((n - x_1) - n(1 - p_1))^2 p_1}{np_1 p_2} = \\ &= \frac{(x_1 - np_1)^2 p_2 + (-x_1 + np_1)^2 p_1}{np_1 p_2} = \frac{(x_1 - np_1)^2 (p_2 + p_1)}{np_1 p_2} = \\ &= \left(\frac{x_1 - np_1}{\sqrt{np_1 p_2}} \right)^2 \end{aligned}$$

Dal momento che

$$\frac{x_1 - np_1}{\sqrt{np_1 p_2}}$$

tende ad una distribuzione normale standard, per il teorema del limite centrale, si ha che

$$\left(\frac{x_1 - np_1}{\sqrt{np_1 p_2}} \right)^2$$

si può approssimare mediante una distribuzione con densità χ^2 ad un grado di libertà.

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

$$x_j \quad j = 1..k$$

relative alla variabile aleatoria $\xi = (\xi_1, \xi_2, \dots, \xi_k)$ che ha una distribuzione multinomiale

Uno stimatore delle differenza tra frequenze osservate e frequenze ipotizzate si può definire come

$$\chi^2 = \sum_{j=1}^k \frac{(x_j - np_j)^2}{np_j} \quad (17)$$

tenendo conto che $n_j = np_j$ è il valor medio di ξ_j

Consideriamo ora la variabile aleatoria η che ha k componenti indipendenti con distribuzione di Poisson di media $\lambda_j = np_j$ condizionate da

$$\eta_1 + \eta_2 + \cdots + \eta_k = n$$

Abbiamo visto in precedenza che η ha la stessa distribuzione di ξ e quindi possiamo utilizzare in luogo della 17 la

$$\chi^2 = \sum_{j=1}^k \frac{(y_j - \lambda_j)^2}{\lambda_j} = \sum_{j=1}^k \left(\frac{y_j - \lambda_j}{\sqrt{\lambda_j}} \right)^2 \quad (18)$$

Se ora facciamo $n \rightarrow +\infty$, per il teorema del limite centrale, possiamo affermare che

$$\frac{y_j - \lambda_j}{\sqrt{\lambda_j}} \rightarrow \zeta_j$$

dove ζ_j ha distribuzione normale standard, pertanto avremo che

$$\chi^2 = \sum_{j=1}^k \zeta_j^2 = \zeta_1^2 + \zeta_2^2 + \cdots + \zeta_k^2$$

è la somma di k variabili aleatorie normali standard condizionate da

$$0 = \eta_1 + \eta_2 + \cdots + \eta_k - n =$$

$$\begin{aligned} \sqrt{n} \left(\sqrt{p_1} \frac{\eta_1 - np_1}{\sqrt{np_1}} + \sqrt{p_2} \frac{\eta_2 - np_2}{\sqrt{np_2}} + \cdots + \sqrt{p_k} \frac{\eta_k - np_k}{\sqrt{np_k}} \right) = \\ \sqrt{n} \left(\sqrt{p_1} \frac{\eta_1 - \lambda_1}{\sqrt{\lambda_1}} + \sqrt{p_2} \frac{\eta_2 - \lambda_2}{\sqrt{\lambda_2}} + \cdots + \sqrt{p_k} \frac{\eta_k - \lambda_k}{\sqrt{\lambda_k}} \right) \end{aligned}$$

da cui

$$0 = \left(\sqrt{p_1} \frac{\eta_1 - \lambda_1}{\sqrt{\lambda_1}} + \sqrt{p_2} \frac{\eta_2 - \lambda_2}{\sqrt{\lambda_2}} + \cdots + \sqrt{p_k} \frac{\eta_k - \lambda_k}{\sqrt{\lambda_k}} \right)$$

ed, al limite,

$$0 = \sqrt{p_1} \zeta_1 + \sqrt{p_2} \zeta_2 + \cdots + \sqrt{p_k} \zeta_k$$

Sia φ la PDF della variabile aleatoria ottenuta; avremo che

$$\begin{aligned} \varphi(x) &= \\ &= \mathcal{P}(\zeta_1^2 + \zeta_2^2 + \cdots + \zeta_k^2 \leq x \mid \sqrt{p_1} \zeta_1 + \sqrt{p_2} \zeta_2 + \cdots + \sqrt{p_k} \zeta_k = 0) = \\ &= \lim_{\epsilon \rightarrow 0} \mathcal{P}(\zeta_1^2 + \zeta_2^2 + \cdots + \zeta_k^2 \leq x \mid |\sqrt{p_1} \zeta_1 + \sqrt{p_2} \zeta_2 + \cdots + \sqrt{p_k} \zeta_k| < \epsilon) \end{aligned}$$

Dal momento che ζ_j è una variabile normale standard, la sua PDF è

$$g_j(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

per cui

$$\mathcal{P}(\zeta_1^2 + \zeta_2^2 + \cdots + \zeta_k^2 \leq x | \sqrt{p_1}\zeta_1 + \sqrt{p_2}\zeta_2 + \cdots + \sqrt{p_k}\zeta_k | < \epsilon) = \frac{\frac{1}{\sqrt{2\pi}^k} \int_V e^{-\frac{u_1^2 + u_2^2 + \cdots + u_k^2}{2}} du_1 du_2 \dots, du_k}{\mathcal{P}(|\sqrt{p_1}\zeta_1 + \sqrt{p_2}\zeta_2 + \cdots + \sqrt{p_k}\zeta_k| < \epsilon)}$$

essendo V la parte della sfera n -dimensionale compresa tra i piani

$$\sqrt{p_1}\zeta_1 + \sqrt{p_2}\zeta_2 + \cdots + \sqrt{p_k}\zeta_k = \pm \epsilon$$

Poichè l'integranda dipende solo dalla distanza dall'origine e $|\sqrt{p_1}\zeta_1 + \sqrt{p_2}\zeta_2 + \cdots + \sqrt{p_k}\zeta_k| < \epsilon$ rappresenta la parte di spazio delimitata da due piani paralleli ed equidistanti da un piano per l'origine, il calcolo non dipende da come è inclinato il piano e possiamo quindi sostituire la condizione data con la

$$|\sqrt{p_k}\zeta_k| < \epsilon$$

Ne segue che

$$\begin{aligned} \frac{1}{\sqrt{2\pi}^k} \int_V e^{-\frac{u_1^2 + u_2^2 + \cdots + u_k^2}{2}} du_1 du_2 \dots, du_k &\approx \\ &\approx \frac{1}{\sqrt{2\pi}^k} 2\epsilon \int_{V_{k-1}} e^{-\frac{u_1^2 + u_2^2 + \cdots + u_{k-1}^2}{2}} du_1 du_2 \dots, du_{k-1} = \\ &= \frac{1}{\sqrt{2\pi}} 2\epsilon \chi_{k-1}^2(x) \end{aligned}$$

essendo V_{k-1} la sfera $(n-1)$ -dimensionale di raggio x e, conseguentemente, χ_{k-1}^2 la CDF di una variabile aleatoria di tipo χ^2 a $k-1$ gradi di libertà

D'altro canto $\sqrt{p_1}\zeta_1 + \sqrt{p_2}\zeta_2 + \cdots + \sqrt{p_k}\zeta_k$ è una variabile aleatoria normale standard in quanto combinazione lineare di variabili aleatorie normali standard per cui la sua PDF è

$$g(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

e quindi

$$\begin{aligned}\mathcal{P}(|\sqrt{p_1}\zeta_1 + \sqrt{p_2}\zeta_2 + \cdots + \sqrt{p_k}\zeta_k| < \epsilon) &= \frac{1}{\sqrt{2\pi}} \int_{-\epsilon}^{+\epsilon} e^{-\frac{u^2}{2}} \approx \\ &\approx \frac{1}{\sqrt{2\pi}} 2\epsilon\end{aligned}$$

Ne segue che

$$\mathcal{P}(\zeta_1^2 + \zeta_2^2 + \cdots + \zeta_k^2 \leq x | |\sqrt{p_1}\zeta_1 + \sqrt{p_2}\zeta_2 + \cdots + \sqrt{p_k}\zeta_k| < \epsilon) = \frac{\frac{1}{\sqrt{2\pi}} 2\epsilon \chi_{k-1}^2(x)}{\frac{1}{\sqrt{2\pi}} 2\epsilon} = \chi_{k-1}^2(x)$$

(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 .

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 rugoso, 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, in base ai caratteri appena descritti, nella seguente tabella

Caratteri	Frequenza osservata
Rotondo Giallo	315
Rotondo verde	108
rugoso Giallo	101
rugoso verde	32

I tipi di piselli possibili sono pertanto 4 ed i possibili incroci sono 16 e possono essere elencati nella tabella

A	LG	LG	LG	LG	Lv	Lv	Lv	Lv	rG	rG	rG	rG	rv	rv	rv	rv
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
B	LG	Lv	rG	rv	LG	Lv	rG	rv	LG	Lv	rG	rv	LG	Lv	rG	rv
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
C	LG	LG	LG	LG	LG	Lv	Lv	Lv	LG	LG	rG	rG	LG	Lv	rG	rv

dove accanto alle caratteristiche delle piante *A* e *B* incrociate sono riportate le caratteristiche *C* dell'incrocio.

Dalla tabella si evince 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 le frequenze previste per ciascun caso e la relativa differenza:

Caratteri	Freq. osservata	Freq. 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

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.

A questo scopo possiamo usare il test del χ^2 e calcolare la quantità

$$\frac{f_o - f_p}{\sqrt{f_p}}, \quad k = 1, 2, 3, 4$$

per ognuna delle frequenze.

Sommando i quadrati delle 4 variabili indicate in tabella ?? otterremo una nuova variabile che, se gli errori sono distribuiti secondo una gaussiana, è distribuita con una densità 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

$$\mathcal{P}(\chi^2 < .352) = .05$$

$$\mathcal{P}(\chi^2 < .584) = .10$$

$$\mathcal{P}(\chi^2 < 6.25) = .90$$

$$\mathcal{P}(\chi^2 < 7.81) = .95$$

per cui

$$.352 < .47 < .584 < 6.25 < 7.81$$

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%

Modelli di crescita di una Popolazione.

Thomas Robert Malthus pubblicò tra il 1798 ed il 1826 sei edizioni del suo libro "An Essay on the Principle of Population" in cui sosteneva che la crescita della popolazione sarebbe stata insostenibile per le risorse presenti sulla terra e argomentava che la crescita di una popolazione poteva essere controllata da fattori quali malattie, carestie e guerre da un lato e aborto, controllo delle nascite, prostituzione, ritardo del matrimonio e celibato dall'altro.

Il Modello Esponenziale di Malthus prevede che la crescita di una popolazione sia proporzionale al numero di individui;

Sia

$$n(t)$$

il numero di individui di una popolazione al tempo t Se c è il tasso di crescita n dovrà soddisfare il problema di cauchy

$$\begin{cases} \dot{n}(t) = cn(t) \\ n(t_0) = n_0 \end{cases}$$

dove c è il tasso di crescita

Possiamo risolvere l'equazione ed ottenere

$$n(t) = n_0 e^{c(t-t_0)} \quad (19)$$

Per applicare il modello occorre conoscere c
 Possiamo determinarne il valore usando i dati sperimentali.
 Allo scopo, sono sufficienti due valori.

$$n_0 = n(t_0) \quad n_1 = n(t_1)$$

Infatti, dalla 19 per $t = t_1$ si ricava

$$n_1 = n_0 e^{c(t_1 - t_0)} \quad \text{e} \quad c = \frac{\ln\left(\frac{n_1}{n_0}\right)}{t_1 - t_0}$$

Per $n_1 = 2n_0$ si parla di **tempo di raddoppio**

Per $n_1 = \frac{1}{2}n_0$ si parla di **tempo di dimezzamento**

c , il tasso di crescita istantaneo, è dato da

$$c = \frac{\dot{n}(t)}{n(t)}$$

La seguente tabella riporta i dati reali della popolazione mondiale rilevate a metà anno come sono riportati in <http://www.census.gov>

Anno	Valore vero
1950	2,556,505,579
1955	2,781,184,998
1960	3,042,389,609
1970	3,712,813,318
1980	4,452,686,744
1985	4,858,097,523
1990	5,288,828,246
2000	6,088,683,554
2005	6,469,129,736
2010	6,853,019,414
2011	6,929,098,151

Per $t_0 = 1960$ $n(t_0) = n_0 = 3042389609$

Per $t_1 = 1970$ $n(t_1) = n_1 = 3712813318$

Pertanto

$$c = \frac{\ln(\frac{n_1}{n_0})}{t_1 - t_0} = \frac{1}{10} \ln \left(\frac{3712813318}{3042389609} \right) \approx .1991466339$$

e

$$n(t) = 3042389609 e^{.1991466339(t-1960)}$$

Se calcoliamo la soluzione e mettiamo a confronto i valori previsti dal Modello di Malthus con i valori veri, otteniamo la seguente tabella.

Anno	Malthus	Valore vero
1950	2,493,024,500	2,556,505,579
1955	2,754,042,830,	2,781,184,998
1960	3,042,389,609	3,042,389,609
1970	3,712,813,343	3,712,813,318
1980	4,530,972,209	4,452,686,744
1985	5,005,362,574	4,858,097,523
1990	5,529,421,352	5,288,828,246
2000	6,747,889,586	6,088,683,554
2005	7,454,390,012	6,469,129,736
2010	8,234,860,571	6,853,019,414
2011	8,400,498,946	6,929,098,151

Il modello Malthusiano sottostima, anche se di poco il dato per il 1950 mentre sovrastima di parecchio i valori degli anni 2005 - 2011.

Il modello esponenziale è adatto a rappresentare un gran numero di fenomeni; citiamo ad esempio il decadimento di materiale radioattivo.

Ogni elemento radioattivo decade in maniera proporzionale alla propria massa.

Ad esempio il torio ha un tempo di dimezzamento di 1.65×10^{10} anni;

Pertanto il torio decade seguendo la legge

$$q = q_0 e^{ct}$$

dove c si può ricavare dalla

$$c = \frac{\ln(\frac{1}{2})}{1.65 \times 10^{10}} = -1.33 \times 10^{-18}$$

Il modello di Malthus non prevede condizionamenti dell'ambiente e considera il tasso di crescita costante, mentre si vede subito, esaminando i dati sperimentali, che il tasso di crescita diminuisce. Possiamo tentare di interpretare questa decrescita osservando che una popolazione per svilupparsi consuma le risorse dall'ambiente che, non essendo illimitate, ne condizionano lo sviluppo.

Un modello che tenga conto anche di queste condizioni fu proposto da Verhulst (1846)

Si assume che il tasso di accrescimento c diminuisca linearmente con il numero degli individui.

$$c(n) = a - bn \quad a, b > 0$$

Poichè

$$c(\bar{n}) = a - b\bar{n} = 0 \quad \Longleftrightarrow \quad \bar{n} = \frac{a}{b}$$

è evidente che la presenza di un numero $\bar{n}$ di individui annulla il tasso di crescita ed impedisce una ulteriore crescita

$\bar{n}$ assume pertanto il significato di massimo numero di individui sostenibili dall'ambiente.

Possiamo definire un modello mediante il problema di Cauchy

$$\begin{cases} \dot{n}(t) = an(t) - bn^2(t) = n(t)(a - bn(t)) \\ n(t_0) = n_0 \end{cases}$$

L'equazione differenziale è nota come equazione logistica

La sua soluzione si chiama funzione logistica.

L'equazione Logistica

- 1 ha due punti di equilibrio (soluzioni costanti) per

$$n(t) = 0 \quad n(t) = \frac{a}{b} = n_{max}$$

corrispondono, rispettivamente, al caso in cui

- la popolazione si è estinta
- la popolazione ha saturato l'ambiente

- 2 poichè

$$an - bn^2 > 0 \quad \text{per} \quad 0 < n < \frac{a}{b}$$

la soluzione è

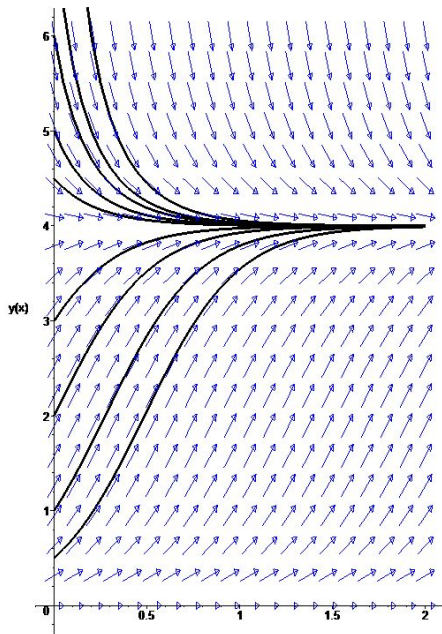
- crescente se $n_0 < \frac{a}{b}$
- decrescente se $n_0 > \frac{a}{b}$

- 3 poichè

$$\ddot{n}(t) = (a - 2bn(t))\dot{n}(t)$$

- la soluzione è convessa ($\dot{n}$ crescente), se $n(t) < \frac{a}{2b}$
- concava ($\dot{n}$ decrescente), per $\frac{a}{b} > n(t) > \frac{a}{2b}$.

$$y'(x) = 4y(x) - y^2(x)$$



Si può congetturare che n tende ad un valore asintotico n_{max}

- crescendo se $n_0 < n_{max}$
- decrescendo se $n_0 > n_{max}$
- (rimane costante se $n_0 = n_{max}$)

per

$$n(t) = \frac{a}{2b} = \frac{n_{max}}{2}$$

c'è un punto di flesso

La crescita è molto rapida fino a che $n(t) = \frac{a}{2b} = \frac{n_{max}}{2}$ rallenta avvicinandosi all'asintoto.

Possiamo integrare l'equazione

$$\begin{cases} \dot{n}(t) = an(t) - bn^2(t) = n(t)(a - bn(t)) \\ n(t_0) = n_0 \end{cases}$$

Vale il teorema di esistenza ed unicità e ci sono due soluzioni costanti

$$n(t) = 0 \quad , \quad n(t) = \frac{a}{b}$$

Che sono le soluzioni di equilibrio del modello). Si ha:

$$\frac{\dot{n}(t)}{n(t)(a - bn(t))} = 1$$

e

$$\int_{t_0}^t \frac{\dot{n}(s)}{n(s)(a - bn(s))} ds = t - t_0 \qquad \int_{n(t_0)}^{n(t)} \frac{dx}{x(a - bx)} = t - t_0$$

Si ha:

$$\frac{1}{x(a - bx)} = \frac{1}{a} \left(\frac{1}{x} + \frac{b}{a - bx} \right)$$

da cui

$$\begin{aligned} \int_{n_0}^{n(t)} \frac{dx}{x(a - bx)} &= \frac{1}{a} \left(\ln |x| - \ln |a - bx| \right) \Big|_{n_0}^{n(t)} = \\ &= \frac{1}{a} \ln \left(\frac{n(t)}{a - bn(t)} \right) \left(\frac{a - bn_0}{n_0} \right) \end{aligned}$$

Ne viene

$$t - t_0 = \frac{1}{a} \ln c \left(\frac{n(t)}{a - bn(t)} \right)$$

dove $\frac{n(t)}{a - bn(t)}$ ha lo stesso segno di $c = \frac{n_0}{a - bn_0}$ in ciascuno degli intervalli in cui è definita

Si ricava

$$n(t) = \frac{a}{b + ce^{-a(t-t_0)}} = \frac{1}{\beta + \gamma e^{-a(t-t_0)}} \quad (20)$$

dove $\beta = \frac{b}{a}$ e $\gamma = \frac{c}{a}$

Le costanti a, β e γ possono essere determinate se si conoscono

$$n(t_0 - h) = p_0$$

$$n(t_0) = p_1$$

$$n(t_0 + h) = p_2$$

infatti si ha

$$\frac{1}{n(t)} = \beta + \gamma e^{-a(t-t_0)}$$

ed è sufficiente risolvere il sistema

$$\begin{cases} \frac{1}{p_0} = \beta + \gamma e^{ah} \\ \frac{1}{p_1} = \beta + \gamma \\ \frac{1}{p_2} = \beta + \gamma e^{-ah} \end{cases}$$

Posto $H = e^{-ah}$ otteniamo

$$\begin{cases} \frac{1}{p_0} = \beta + \gamma \frac{1}{H} \\ \frac{1}{p_1} = \beta + \gamma \\ \frac{1}{p_2} = \beta + \gamma H \end{cases} \quad (21)$$

e pertanto

$$\begin{cases} \frac{1}{p_0} - \frac{1}{p_1} = \gamma \left(\frac{1}{H} - 1 \right) = \gamma \frac{1-H}{H} \\ \frac{1}{p_1} - \frac{1}{p_2} = \gamma(1-H) \end{cases} \quad (22)$$

$$\frac{\frac{1}{p_0} - \frac{1}{p_1}}{\frac{1}{p_1} - \frac{1}{p_2}} = \frac{1}{H}$$

Se ne deduce che

$$\frac{1}{H} = \frac{p_1 p_2 (p_1 - p_0)}{p_0 p_1 (p_2 - p_1)} \qquad H = \frac{p_0 p_1 (p_2 - p_1)}{p_1 p_2 (p_1 - p_0)}$$

ed è infine facile ricavare γ e β dalle 21,22.

Per $t_0 = 1960$ $h = 10$

$n(t_0 - h) = p_0 = 2,555,360,972$

$n(t_0) = p_1 = 3,039,669,330$

$n(t_0 + h) = p_2 = 3,708,067,105$

Pertanto

$$n(t) = \frac{10^9}{-.88 + 1.21 * e^{-.005t+9.8}}$$

Tabella dei valori previsti dal Modello di Verhulst a fronte dei valori effettivi.

Anno	Verhulst 1	Valore vero
1950	2,556,505,606	2,556,505,579
1955	2,781,454,590	2,781,184,998
1960	3,042,389,614	3,042,389,609
1970	3,712,813,351	3,712,813,318
1980	4,695,962,886	4,452,686,744
1985	5,381,578,768	4,858,097,523
1990	6,274,505,666	5,288,828,246
2000	9,218,652,877	6,088,683,554
2005	11,906,613,960	6,469,129,736
2010	16,634,169,740	6,853,019,414
2011	18,045,274,540	6,929,098,151

È immediato osservare che i dati ottenuti sono nettamente peggiori di quelli relativi al modello esponenziale, tuttavia se cambiamo la scelta dei parametri la situazione migliora.

Per $t_0 = 1970$ $h = 15$

$$n(t_0 - h) = p_0 = 2,781,184,998$$

$$n(t_0) = p_1 = 3,712,813,618$$

$$n(t_0 + h) = p_2 = 4,858,097,523$$

Pertanto

$$n(t) = \frac{10^{10}}{0.5499 + 2.14 * e^{-.0023t+46.1}}$$

Tabella dei valori previsti dal Modello di Verhulst a fronte dei valori effettivi.

Anno	Verhulst 2	Valore vero
1950	2,516,381,991	2,556,505,579
1955	2,781,185,002	2,781,184,998
1960	3,068,396,960	3,042,389,609
1970	3,712,813,617	3,712,813,318
1980	4,452,705,967	4,452,686,744
1985	4,858,097,515	4,858,097,523
1990	5,286,190,494	5,288,828,246
2000	6,205,196,592	6,088,683,554
2005	6,692,388,106	6,469,129,736
2010	7,194,859,309	6,853,019,414
2011	7,296,941,395	6,929,098,151

Con la nuova scelta di dati iniziali i valori previsti sono nettamente piu' vicini ai valori effettivamente osservati. Ciò evidenzia come sia importante la taratura dei parametri.

Anno	Malthus	Verhulst 1	Verhulst 2	Valore vero	M	V-1	V-2
1950	2,493,024,500	2,556,505,606	2,516,381,991	2,556,505,579		$p_0 \checkmark$	
1955	2,754,042,830	2,781,454,590	2,781,185,002	2,781,184,998			$p_0 \checkmark$
1960	3,042,389,609	3,042,389,614	3,068,396,960	3,042,389,609	$p_0 \checkmark$	$p_1 \checkmark$	
1970	3,712,813,343	3,712,813,351	3,712,813,617	3,712,813,318	$p_1 \checkmark$	$p_1 \checkmark$	$p_1 \checkmark$
1980	4,530,972,209	4,695,962,886	4,452,705,967	4,452,686,744			
1985	5,005,362,574	5,381,578,768	4,858,097,515	4,858,097,523			$p_2 \checkmark$
1990	5,529,421,352	6,274,505,666	5,286,190,494	5,288,828,246			
2000	6,747,889,586	9,218,652,877	6,205,196,592	6,088,683,554			
2005	7,454,390,012	11,906,613,960	6,692,388,106	6,469,129,736			
2010	8,234,860,571	16,634,169,740	7,194,859,309	6,853,019,414			
2011	8,400,498,946	18,045,274,540	7,296,941,395	6,929,098,151			

Consideriamo ora il caso di una popolazione la cui crescita è influenzata da un prelievo ad opera di agenti esterni;
Ad esempio

- un allevamento ittico, da cui viene pescata una quantità fissa di pesce
- una specie in ambiente protetto e senza antagonisti il cui sviluppo deve essere controllato mediante un prelievo periodico per mantenere il numero di esemplari stabile.

Se la popolazione si sviluppa con tasso di crescita $a - bn$ ed è sottoposta ad prelievo costante pari ad una quantità c avremo

$$\begin{cases} \dot{n}(t) = an(t) - bn^2(t) - c \\ n(t_0) = n_0 \end{cases}$$

$a, b, c > 0$

Il valore di c è critico e può portare all'estinzione della popolazione.

È utile determinare il massimo prelievo che non causa l'estinzione della popolazione.

Vale il teorema di esistenza ed unicità

L'equazione ha soluzioni costanti se

$$-bx^2 + ax - c = 0$$

ha soluzioni reali.

Possiamo pertanto distinguere tre casi

$$a^2 - 4bc < 0$$

Non ci sono soluzioni costanti e si ha

$$\frac{\dot{n}(t)}{an(t) - bn^2(t) - c} = 1$$

$$\int_{t_0}^t \frac{\dot{n}(s)}{an(s) - bn^2(s) - c} ds = t - t_0$$

$$\int_{n(t_0)}^{n(t)} \frac{dx}{ax - bx^2 - c} = t - t_0$$

L'integranda può essere decomposta in fratti semplici mediante la

$$\begin{aligned}\frac{1}{ax - bx^2 - c} &= -\frac{1}{b} \frac{1}{\left(x - \frac{a}{2b}\right)^2 - \frac{a^2}{4b^2} + \frac{c}{b}} = \\ &= -\frac{1}{b} \frac{1}{\left(x - \frac{a}{2b}\right)^2 + \delta} = -\frac{1}{b\delta} \frac{1}{\left(\frac{1}{\sqrt{\delta}} \left(x - \frac{a}{2b}\right)\right)^2 + 1} =\end{aligned}$$

dove si è posto

$$\delta = \frac{4bc - a^2}{4b^2}$$

Pertanto

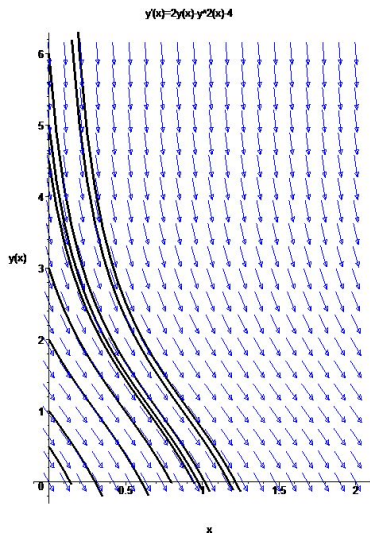
$$\int \frac{dx}{ax - bx^2 - c} = -\frac{1}{b\sqrt{\delta}} \arctan \frac{1}{\sqrt{\delta}} \left(x - \frac{a}{2b}\right) + \text{cost.}$$

e

$$\begin{aligned}-\frac{1}{b\sqrt{\delta}} \arctan \frac{1}{\sqrt{\delta}} \left(n(t) - \frac{a}{2b}\right) + \\ + \frac{1}{b\sqrt{\delta}} \arctan \frac{1}{\sqrt{\delta}} \left(n_0 - \frac{a}{2b}\right) = t - t_0\end{aligned}$$

Inoltre

$$n(t) = \frac{a}{2b} + \sqrt{\delta} \tan \left(t_0 + \frac{1}{b\sqrt{\delta}} \arctan \frac{1}{\sqrt{\delta}} \left(n_0 - \frac{a}{2b} \right) - t \right)$$



Popolazione nel caso di
crescita logistica con prelievo
costante per

$$a = 2$$

$$b = 1$$

$$c = 4$$

$$a^2 - 4bc = -12 < 0$$

$$a^2 - 4bc = 0$$

$$-bx^2 + ax - c = 0 \text{ per } x = \frac{a}{2b}$$

$$n(t) = \alpha = \frac{a}{2b}$$

è una soluzione costante dell'equazione differenziale
Si ha inoltre

$$\frac{1}{ax - bx^2 - c} = -\frac{1}{b} \frac{1}{\left(x - \frac{a}{2b}\right)^2}$$

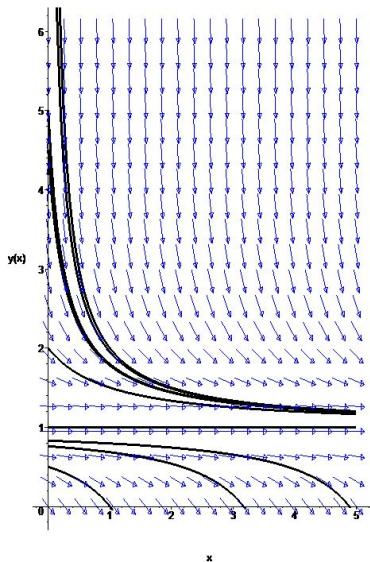
da cui

$$\int_{n_0}^{n(t)} \frac{dx}{ax - bx^2 - c} = \frac{1}{b} \frac{1}{\left(n(t) - \frac{a}{2b}\right)} - \frac{1}{b} \frac{1}{\left(n_0 - \frac{a}{2b}\right)} = t - t_0$$

e

$$n(t) = \frac{1}{2b} \left(a + \frac{2bn_0 - a}{1 + \left(n_0b - \frac{a}{2}\right)(t - t_0)} \right)$$

$$y'(x) = 2y(x) - y^2(x) - 1$$



Popolazione nel caso di crescita logistica con prelievo costante per

$$a = 2$$

$$b = 1$$

$$c = 1$$

$$a^2 - 4bc = 0$$

$$a^2 - 4bc > 0$$

In questo caso l'equazione

$$-bx^2 + ax - c = 0$$

ammette due soluzioni reali e distinte

$$\alpha = \frac{a}{2b} + \sqrt{\frac{a^2 - 4bc}{4b^2}}, \quad \beta = \frac{a}{2b} - \sqrt{\frac{a^2 - 4bc}{4b^2}}$$

e quindi possiamo trovare due soluzioni costanti di equilibrio per il sistema.

Si ha

$$-bx^2 + ax - c = -b(x - \alpha)(x - \beta)$$

e

$$\frac{1}{-bx^2 + ax - c} = \frac{1}{b(\beta - \alpha)} \left(\frac{1}{x - \alpha} - \frac{1}{x - \beta} \right)$$

da cui

$$\int \frac{1}{-bx^2 + ax - c} dx = \frac{1}{b(\beta - \alpha)} (\ln |x - \alpha| - \ln |x - \beta|) + \varphi(x)$$

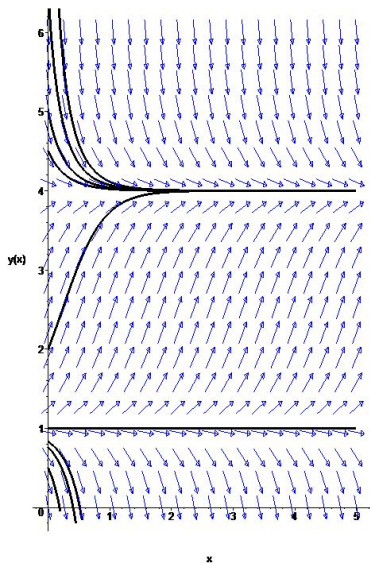
e φ è una funzione costante a tratti

$$\begin{aligned} \int_{n_0}^{n(t)} \frac{1}{-bx^2 + ax - c} dx &= \\ &= \frac{1}{b(\beta - \alpha)} \ln \left[\left(\frac{n(t) - \alpha}{n(t) - \beta} \right) \left(\frac{n_0 - \beta}{n_0 - \alpha} \right) \right] = t - t_0 \end{aligned}$$

da cui

$$n(t) = \frac{\alpha - \beta \frac{n_0 - \alpha}{n_0 - \beta} e^{-b(\beta - \alpha)(t - t_0)}}{1 - \frac{n_0 - \alpha}{n_0 - \beta} e^{-b(\beta - \alpha)(t - t_0)}}$$

$$y'(x) = 5y(x) - y^2(x) - 4$$



Popolazione nel caso di crescita logistica con prelievo costante per

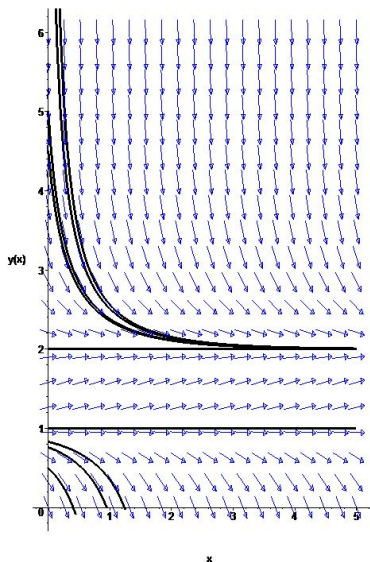
$$a = 5$$

$$b = 1$$

$$c = 4$$

$$a^2 - 4bc = 9 > 0$$

$$y'(x) = 3y(x) - y^2(x) - 2$$



Andamento della popolazione
con prelievo costante nei casi in
cui

$$a = 3$$

$$b = 1$$

$$c = 2$$

$$a^2 - 4bc = 1 > 0$$

Come si può facilmente osservare dai grafici della soluzione $n(t)$ il modello rende ragione dei seguenti fatti

- ① se $a^2 - 4bc < 0$ la popolazione si estingue in un tempo finito
- ② se $a^2 - 4bc = 0$
 - la popolazione si estingue in un tempo finito se $n_0 < \frac{a}{2b}$
 - mentre tende a stabilizzarsi sul valore $\frac{a}{2b}$ asintoticamente se $n_0 > \frac{a}{2b}$
- ③ se $a^2 - 4bc > 0$
 - la popolazione si estingue in un tempo finito se $n_0 < \beta$
 - mentre tende a stabilizzarsi sul valore β asintoticamente se $n_0 > \beta$

Pertanto, pur di partire con un sufficiente numero di individui, non si ha estinzione della popolazione per

$$c \leq \frac{a^2}{4b}$$

$c = \frac{a^2}{4b}$ è la più alta quantità che si può prelevare senza causare l'estinzione della popolazione in un tempo finito

È la rendita massima dell'allevamento.

Nel caso la popolazione non si estingua, essa si assesta su un valore di regime pari a

$$\alpha = \frac{a}{2b} + \sqrt{\frac{a^2 - 4bc}{4b^2}} < \frac{a}{b}$$

dove $\frac{a}{b}$ è il valore a regime della popolazione in assenza di prelievo.

Nel modello differenziale di Malthus c rappresenta il tasso di crescita istantaneo

Pertanto, nel periodo T , avremo una crescita pari a

$$\gamma = cT$$

e γ rappresenta il tasso di crescita nel periodo T

Possiamo allora scrivere che

$$\begin{cases} n(t_0 + T) = n(t_0) + cTn(t_0) = n(t_0) + \gamma n(t_0) \\ n(t_0) = n_0 \end{cases}$$

Sia dunque γ il tasso di accrescimento di una popolazione relativo ad un periodo di tempo T . Si ha

$$\begin{cases} n(t_0 + T) = n(t_0) + \gamma n(t_0) \\ n(t_0) = n_0 \end{cases}$$

L'equazione che definisce l'evoluzione del modello si può considerare corretta solo se T è abbastanza piccolo; nel caso in cui T sia grande, la popolazione, dopo un tempo T , sarà cresciuta anche per opera degli individui nati nei periodi

$$\left[t_0 + h\frac{T}{k}, t_0 + (h+1)\frac{T}{k} \right], \quad h = 0, 1, \dots, k-1$$

Il tasso di accrescimento relativo all'intervallo $[t_0 + h\frac{T}{k}, t_0 + (h+1)\frac{T}{k}]$ sarà:

$$\frac{\gamma}{k} = \frac{cT}{k}$$

e, posto

$$n_h = n\left(t_0 + h\frac{T}{k}\right) \quad h = 0..k$$

avremo

$$\begin{cases} n_0 = n(t_0) \\ n_h = n_{h-1} + \frac{cT}{k} n_{h-1} \end{cases} \quad 1 \leq h \leq k$$

Si ha

$$n_h = \left(1 + \frac{cT}{k}\right) n_{h-1} = \quad (23)$$

$$= \left(1 + \frac{cT}{k}\right)^2 n_{h-2} = \quad (24)$$

$$= \left(1 + \frac{cT}{k}\right)^h n(t_0) = \quad (25)$$

$$= \left(1 + \frac{cT}{k}\right)^h n_0 \quad (26)$$

Confrontando le soluzioni della 19 e della 26, avremo, dopo un tempo T

$$n(t_0 + T) = n_0 e^{cT} \text{ oppure } n(t_0 + T) = n_0 \left(1 + \frac{cT}{k}\right)^k$$

Si rileva che per $k \rightarrow +\infty$

$$\left(1 + \frac{cT}{k}\right)^k \rightarrow e^{cT}$$

La soluzione del modello discreto tende alla soluzione del modello continuo.

$$n(t_0 + T) = n_0 e^{cT} \text{ oppure } n(t_0 + T) = n_0 \left(1 + \frac{cT}{k}\right)^k$$

Con lo stesso procedimento si può calcolare il rendimento sul periodo T di un capitale al tasso c se si contabilizzano gli interessi su ognuno dei sottoperiodi di lunghezza $\frac{T}{k}$ (interesse composto),

$$n(t_0 + T) = n_0 \left(1 + \frac{cT}{k}\right)^k$$

È il valore maturato al tempo T partendo da un capitale n_0 con un tasso di interesse c sul periodo T contabilizzato su $k - 1$ sottoperiodi.

Il Modello preda Predatore.

Il modello preda-predatore descrive le interazioni di due specie una delle quali si nutre dell'altra. Fu introdotto attorno al 1925 indipendentemente da Vito Volterra ed Alfred Lotka e la sua genesi è degna di menzione.

Tra il 1920 ed il 1930 il biologo Umberto D'Ancona si trovò ad analizzare i dati relativi al pescato del porto di Fiume; tra il 1914 ed il 1919 si era osservato un aumento sensibile della percentuale di pescato costituita da selaci, cioè da pesci predatori quali squali e razze, di scarso interesse commerciale.

I dati di cui egli disponeva sono raccolti nella seguente tabella.

Anno	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923
% di Selaci	11.9	21.4	22.1	21.1	36.4	27.3	16.0	15.9	14.8	10.7

Dai dati risulta evidente una crescita della parte di pescato costituita da predatori ed una corrispondente decrescita della parte di pesci non predatori.

Tale crescita era ovviamente conseguenza di una crescita della popolazione di selaci presente nel mare Adriatico ed era ragionevole che fosse stata provocata da una diminuzione dell'attività di pesca dovuta alle operazioni belliche in quel periodo, tuttavia questa osservazione non spiega come mai invece la popolazione delle prede non era cresciuta nello stesso modo, anzi percentualmente era diminuita.

D'Ancona, dopo aver tentato di spiegare il fenomeno con considerazioni di carattere biologico, si rivolse a Vito Volterra che per studiare il fenomeno formulò il modello che ora porta il nome e che fornisce una spiegazione dei dati osservati, aprendosi inoltre a numerosissime applicazioni e miglioramenti.

In maniera del tutto indipendente lo stesso tipo di modello fu introdotto da Alfred James Lotka nel corso del suo lavoro di biofisico e pubblicato nel libro "Elements of Mathematical Biology" del 1924.

- $x(t)$ numero delle prede al tempo t
- $y(t)$ numero dei predatori al tempo t
- In assenza di predatori le prede crescono secondo un modello malthusiano

$$\dot{x}(t) = ax(t)$$

- Il tasso di crescita a delle prede diminuisce in proporzione al numero di predatori

$$a - \alpha y(t)$$

- In assenza di prede i predatori si estinguono seguendo un modello malthusiano

$$\dot{y}(t) = -by(t)$$

- Il tasso di accrescimento b dei predatori aumenta in maniera proporzionale al numero di prede

$$-b + \beta x(t)$$

La dinamica delle due popolazioni può essere descritta dal seguente sistema differenziale (**Equazioni di Lotka-Volterra**)

$$\begin{cases} \dot{x}(t) = ax(t) - \alpha y(t)x(t) \\ \dot{y}(t) = -by(t) + \beta x(t)y(t) \end{cases} \quad (27)$$

con i dati iniziali

$$\begin{cases} x(t_0) = x_0 \\ y(t_0) = y_0 \end{cases}$$

Il sistema differenziale da studiare è semplice ma non si può integrare esplicitamente

Si può dimostrare che:

- 1 Il sistema ammette due punti stazionari $(\frac{b}{\beta}, \frac{a}{\alpha}), (0, 0)$
- 2 Le orbite che giacciono nel primo quadrante sono curve chiuse
- 3 Le soluzioni del sistema sono periodiche
- 4 Vale per le soluzioni del sistema un principio di conservazione delle medie in quanto si vede che

$$\frac{1}{T} \int_0^T x(s) ds = \frac{b}{\beta}, \quad \frac{1}{T} \int_0^T y(s) ds = \frac{a}{\alpha}$$

Il sistema 27 ha soluzioni costanti

Infatti

$$(x(t), y(t)) = (\xi, \eta)$$

sono soluzioni del sistema se

$$\begin{cases} 0 = a\xi - \alpha\eta\xi \\ 0 = -b\eta + \beta\eta\xi \end{cases}$$

e cioè quando

$$\xi = 0, \eta = 0$$

oppure quando

$$\xi = \frac{b}{\beta}, \eta = \frac{a}{\alpha}$$

Le orbite del sistema sono curve chiuse nel I quadrante del piano delle fasi.

Le orbite del sistema, sono descritte da

$$\begin{cases} \dot{x} = x(t) \\ \dot{y} = y(t) \end{cases}$$

Il sistema si può porre nella forma

$$\begin{cases} \dot{x}(t) = ax(t) - \alpha y(t)x(t) = \phi(x(t), y(t)) \\ \dot{y}(t) = -by(t) + \beta x(t)y(t) = \psi(x(t), y(t)) \end{cases} \quad (28)$$

Se l'orbita del sistema può essere rappresentata localmente da una funzione $y = y(x)$, si avrà

$$\begin{cases} x(t) = x \\ y(t) = y(x(t)) = y(x) \end{cases}$$

e si potrà ricavare

$$\dot{y}(t) = y'(x(t))\dot{x}(t)$$

per cui, sostituendo nella 28, si ottiene

$$y'(x) = \frac{\psi(x, y(x))}{\phi(x, y(x))} = \frac{-by(x) + \beta xy(x)}{ax - \alpha xy(x)}$$

Separando le variabili avremo

$$y'(x) \left(\frac{a}{y(x)} - \alpha \right) = \left(\frac{-b}{x} + \beta \right)$$

ed integrando con i dati iniziali

$$\begin{cases} x(t_0) = \bar{x} \\ y(t_0) = \bar{y} \end{cases}$$

cioè per $y(\bar{x}) = \bar{y}$

Otteniamo

$$a \ln y(x) - \alpha y(x) - a \ln \bar{y} + \alpha \bar{y} = -b \ln x + \beta x + b \ln \bar{x} - \beta \bar{x}$$

Ora se poniamo

$$g(y) = a \ln y - \alpha y$$

$$f(x) = -b \ln x + \beta x$$

$$G(y) = g(y) - g(\bar{y})$$

$$F(x) = f(x) - f(\bar{x})$$

possiamo riscrivere la 240 come

$$g(y) = f(x) + g(\bar{y}) - f(\bar{x}) \quad (29)$$

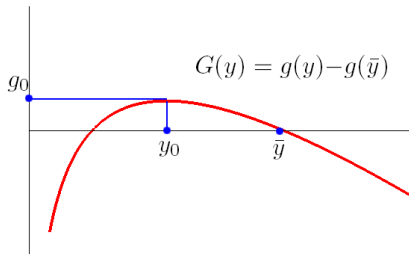
ovvero come

$$G(y) = F(x)$$

Nella figura è rappresentato il grafico della funzione

$$G(y) = g(y) - g(\bar{y})$$

$$y_0 = \frac{a}{\alpha}$$

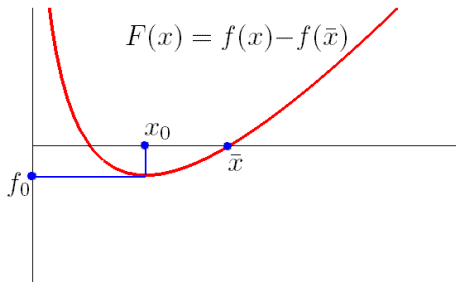


$$\begin{aligned} g_0 = G(y_0) &= a \ln \frac{a}{\alpha} - \alpha \frac{a}{\alpha} = \\ &= a \left(\ln \frac{a}{\alpha} - 1 \right) > 0 \end{aligned}$$

Nella figura è rappresentato il grafico della funzione

$$F(x) = f(x) - f(\bar{x})$$

$$x_0 = \frac{b}{\beta}$$



$$\begin{aligned} f_0 = f(x_0) &= -b \ln \frac{b}{\beta} + \beta \frac{b}{\beta} = \\ &= -b \left(\ln \frac{b}{\beta} - 1 \right) > 0 \end{aligned}$$

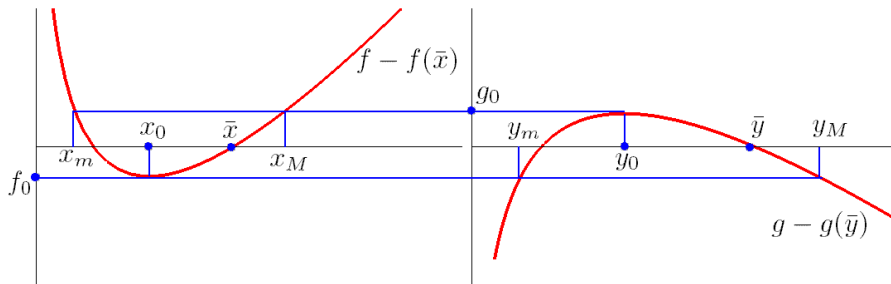
Se G^{-1} è l'inversa di una opportuna restrizione di G dalla 29 allora

$$y(x) = G^{-1}(F(x))$$

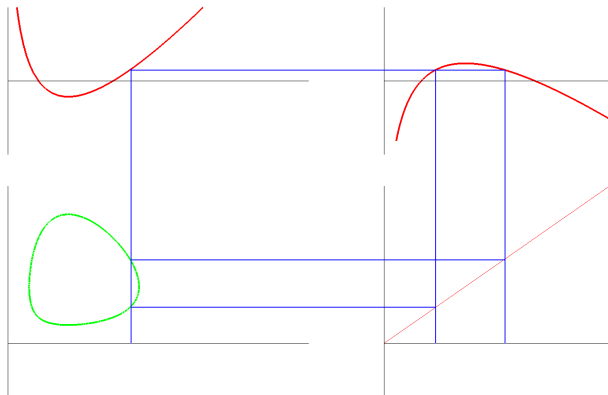
e possiamo osservare y è definita se e solo se

$$[f_0, g_0] = (-\infty, g_0] \cap [f_0, +\infty) \neq \emptyset$$

$$\text{e che } F(x_m) = F(x_M) = g_0 \quad G(y_m) = G(y_M) = f_0$$



Possiamo disegnare un'orbita del sistema $G^{-1}(F(x))$ componendo graficamente G^{-1} ed F ,



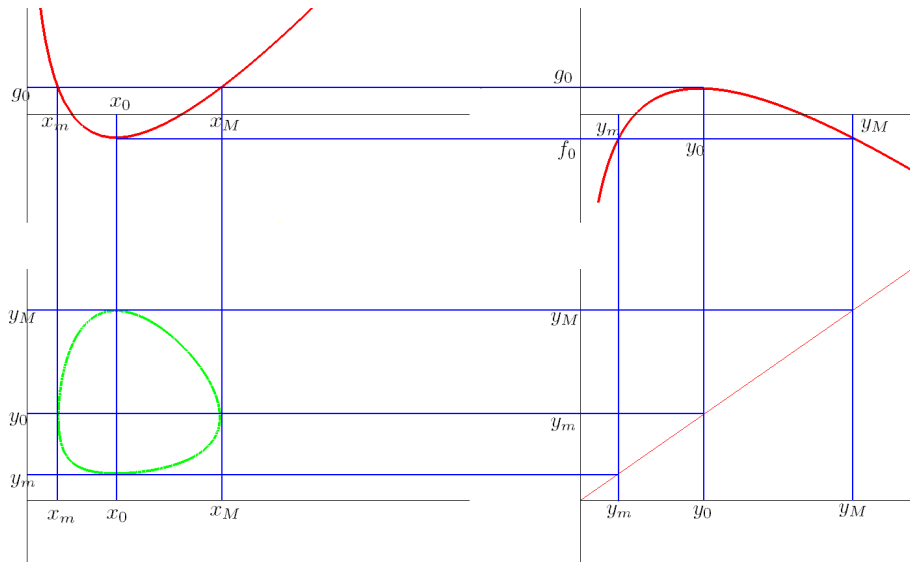
Osserviamo che

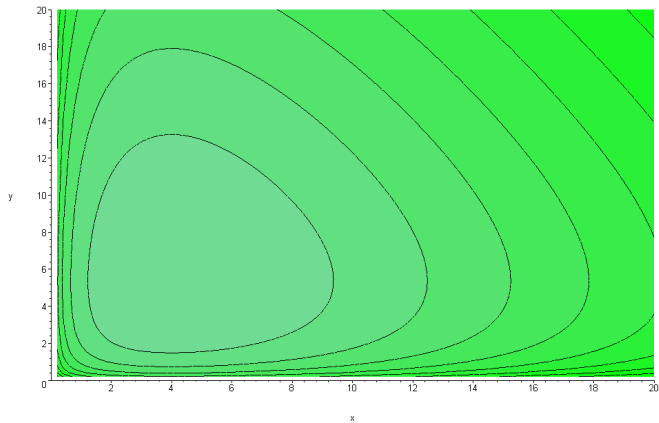
- F è decrescente per $x \in [x_m, x_0]$
- F è crescente per $x \in [x_0, x_M]$
- F assume tutti i valori in $[f_0, g_0]$,
- G è crescente ed invertibile su $[y_m, y_0]$
- G assume tutti i valori in $[f_0, g_0]$
- G è crescente ed invertibile su $[y_0, y_M]$
- G assume tutti i valori in $[f_0, g_0]$

Pertanto

- G^{-1} è definita da $[f_0, g_0]$ a valori in $[y_m, y_0]$ oppure in $[y_0, y_M]$
- $G^{-1}(F(\cdot))$ è definita e decrescente su $[x_m, x_0]$

L'andamento della curva definita dalla 29 si può osservare nella figura che segue





La figura mostra le orbite del sistema.

La funzione

$$U(x, y) = a \ln y - \alpha y + b \ln x - \beta x$$

nella 240, definisce implicitamente le traiettorie del sistema nel piano delle fasi. Si chiama integrale primo del sistema

Dalla 240 si ha,

$$U(x, y(x)) = c = f(\bar{x}) - g(\bar{y})$$

ed inoltre se viceversa

$$U(x(t), y(t)) = c$$

allora

$$U_x(x(t), y(t))\dot{x}(t) + U_y(x(t), y(t))\dot{y}(t) = 0$$

da cui

$$\frac{\psi(x(t), y(t))}{\phi(x(t), y(t))} = \frac{\dot{y}(t)}{\dot{x}(t)} = -\frac{U_x(x(t), y(t))}{U_y(x(t), y(t))} \quad (30)$$

Dalla 30 il gradiente di U è proporzionale al secondo membro del sistema differenziale

Il campo di direzioni definito dal sistema ed il campo di direzioni definito dal gradiente di un integrale primo coincidono.

Più precisamente coincidono le direzioni, non i vettori: il fattore di proporzionalità non solo può non essere 1 ma può non essere costante.

Lo studio delle orbite del sistema ci consente anche di affermare che

- Le orbite del sistema sono chiuse
- Le orbite contengono al loro interno il punto stazionario (x_0, y_0)
- Le orbite sono contenute nel primo quadrante cioè

$$x_m, x_M, y_m, y_M > 0$$

- le orbite non passano per nessun punto stazionario (gli unici punti stazionari diversi da (x_0, y_0) sono sugli assi)

Il sistema è autonomo quindi le traslate di una soluzione sono a loro volta soluzioni

Se

$(x(t), y(t))$ risolve il sistema differenziale

allora

$(x(t + T), y(t + T))$ è soluzione.

infatti

$$\begin{cases} \dot{x}(t + T) = ax(t + T) - \alpha y(t + T)x(t + T) \\ \dot{y}(t + T) = -by(t + T) + \beta x(t + T)y(t + T) \end{cases}$$

Per ogni punto del piano delle fasi passa una ed una sola orbita. Se $(x(t), y(t))$ e $(\xi(t), \eta(t))$ sono orbite del sistema tali che

$$\begin{cases} x(t_0) = \bar{x} \\ y(t_0) = \bar{y} \end{cases} \quad \begin{cases} \xi(\tau_0) = \bar{x} \\ \eta(\tau_0) = \bar{y} \end{cases} \quad \text{e se} \quad \begin{cases} \xi_1(t) = x(t - \tau_0 + t_0) \\ \eta_1(t) = y(t - \tau_0 + t_0) \end{cases}$$

allora (ξ_1, η_1) è soluzione del sistema e soddisfa la condizione

$$\begin{cases} \xi_1(\tau_0) = \bar{x} \\ \eta_1(\tau_0) = \bar{y} \end{cases}$$

Ne deduciamo che

$$\begin{cases} \xi(t) = \xi_1(t) = x(t - \tau_0 + t_0) \\ \eta(t) = \eta_1(t) = y(t - \tau_0 + t_0) \end{cases}$$

e quindi (ξ, η) ed (x, y) percorrono la stessa orbita a meno di una traslazione nel tempo.

Possiamo quindi anche provare che

Le soluzioni del sistema sono periodiche;

La curva è costituita di due tratti che possono essere rappresentati come funzioni su un intervallo limitato quindi la sua lunghezza ℓ è finita,

$$\begin{aligned} \dot{x}^2(t) + \dot{y}^2(t) &= \phi^2(\hat{x}, \hat{y}) + \psi^2(\hat{x}, \hat{y}) \geq \\ &\geq \min_{\substack{(x-x_0)^2 + (y-y_0)^2 > \epsilon \\ x_m \leq x \leq x_M, \quad y_m \leq y \leq y_M}} \{\phi^2(x, y) + \psi^2(x, y)\} \geq m > 0 \end{aligned}$$

$(\phi, \psi$ si annullano contemporaneamente solo nei punti $(0, 0)$ e $(\tilde{x}, \tilde{y})$ che abbiamo già visto non appartenere all'orbita.)

Ne segue che

$$\int_0^{+\infty} \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt = +\infty$$

e pertanto esiste $T > 0$ tale che

$$\int_0^T \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt = \ell$$

Ma allora trascorso il tempo T la soluzione deve ripassare per il punto iniziale $(\bar{x}, \bar{y})$ e, dal momento che per ogni punto passa una sola orbita, da li' è costretta a ripercorrere quell'orbita stessa. Ne viene che la soluzione è periodica.

I valori medi si mantengono costanti, cioè

$$\frac{1}{T} \int_0^T x(s) ds = \frac{b}{\beta} , \quad \frac{1}{T} \int_0^T y(s) ds = \frac{a}{\alpha} \quad (31)$$

Dal sistema, dividendo la prima equazione per x e la seconda per y , otteniamo

$$\begin{cases} \frac{\dot{x}(t)}{x(t)} = a - \alpha y(t) \\ \frac{\dot{y}(t)}{y(t)} = -b + \beta x(t) \end{cases} \quad (32)$$

e per la periodicità delle soluzioni, integrando sul periodo,

$$\begin{cases} 0 = \ln \frac{x(T)}{x(0)} = \int_0^T \frac{\dot{x}(s)}{x(s)} ds = \int_0^T (a - \alpha y(s)) ds = aT - \alpha \int_0^T y(s) ds \\ 0 = \ln \frac{y(T)}{y(0)} = \int_0^T \frac{\dot{y}(s)}{y(s)} ds = \int_0^T (-b + \beta x(s)) ds = -bT + \beta \int_0^T x(s) ds \end{cases}$$

da cui si può ricavare la 31

Per studiare un sistema preda predatore nel quale si intervenga con un prelievo costante h sulle prede e k sui predatori possiamo scrivere

$$\begin{cases} \dot{x}(t) = ax(t) - \alpha y(t)x(t) - hx(t) \\ \dot{y}(t) = -by(t) + \beta x(t)y(t) - ky(t) \\ x(t_0) = x_0 \\ y(t_0) = y_0 \end{cases} \quad (33)$$

Se $a - h$ e $b + k$ sono positivi, ci riconduce immediatamente al caso senza prelievo.

Viene modificata la sola condizione di equilibrio ed il valore medio

$$\bar{x} = \frac{b + k}{\beta} \qquad \bar{y} = \frac{a - h}{\alpha}$$

I risultati trovati consentono di interpretare i dati di D'Ancona sull'aumento della percentuale di Selaci nell'Adriatico.

L'effetto della pesca infatti può essere introdotto nel modello ipotizzando un prelievo costante proporzionale alla popolazione con coefficiente ϵ sia per le prede che per i predatori. ($h = k = \epsilon$). Il punto di equilibrio diviene in tal caso

$$\bar{x} = \frac{b + \epsilon}{\beta} \qquad \bar{y} = \frac{a - \epsilon}{\alpha}$$

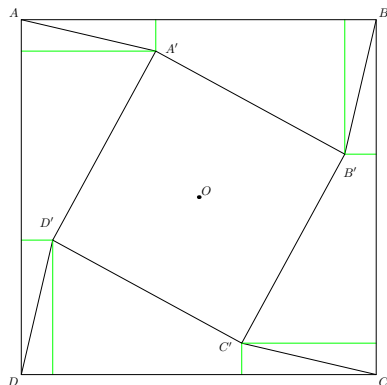
e quindi si vede che se ϵ diminuisce $\bar{x}$ diminuisce e $\bar{y}$ aumenta causando così un aumento della percentuale di y cioè dei predatori.

Il seguente problema è dovuto a Hugo Steinhaus:

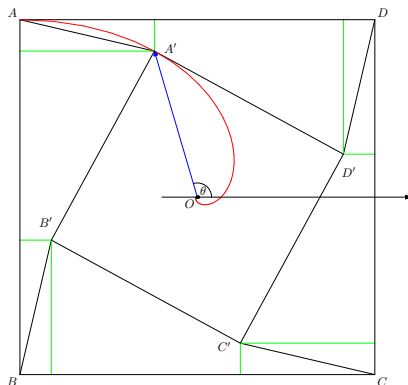
Quattro cani si trovano ai quattro vertici A, B, C, D di un prato quadrato. I cani cominciano a correre con velocità costante puntando ciascuno il cane che si trova nel vertice successivo. Dopo quanto tempo i quattro cani si incontrano e quanta strada hanno percorso? Quale percorso hanno compiuto?

In un istante successivo a quello iniziale i cani saranno, rispettivamente nelle posizioni A' , B' , C' , D' e, per simmetria, i vettori AA' , BB' , CC' , DD' , avranno la stessa lunghezza ed inclinazione rispetto ai lati.

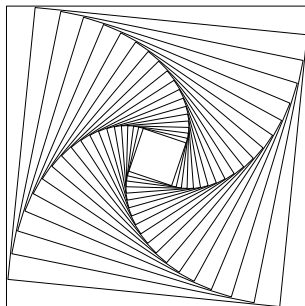
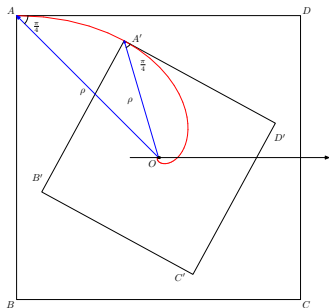
Pertanto $A'B'C'D'$ è un quadrato rimpicciolito e ruotato rispetto al centro comune O .



I cani si muovono seguendo una traiettoria che si mantiene tangente ad uno dei lati dei quadrati
 Cioè il vettore tangente alla curva percorsa (vettore velocità) giace su un lato del quadrato.

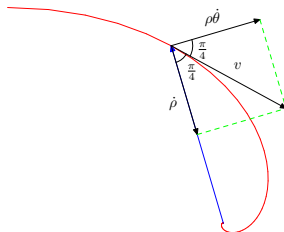


Pertanto il vettore tangente ed il raggio vettore rispetto al centro O formano un angolo costante ed uguale a $\frac{\pi}{4}$.



Poichè l'angolo tra raggio vettore e vettore tangente si mantiene costante, la traiettoria percorsa è una spirale logaritmica centrata in O .

Le componenti radiale $\dot{\rho}(t)$ e tangenziale $\rho(t)\dot{\theta}(t)$ della velocità v formeranno quindi un angolo costante pari a $\frac{\pi}{4}$ e quindi



$$\dot{\rho}(t) = -\frac{\sqrt{2}}{2}|\nu| \quad , \quad \rho(t)\dot{\theta}(t) = -\frac{\sqrt{2}}{2}|\nu|$$

Ne deduciamo che

$$\rho(t) = \frac{\sqrt{2}}{2}(\ell - |\nu|t) \quad , \quad \dot{\theta}(t) = -\frac{|\nu|}{\ell - |\nu|t} = \frac{|\nu|}{|\nu|t - \ell}$$

e

$$\theta(t) = \ln(\ell - |\nu|t) + \frac{3}{4}\pi - \ln(\ell) = \ln\left(\frac{\ell - |\nu|t}{\ell}\right) + \frac{3}{4}\pi = \ln\left(1 - \frac{|\nu|t}{\ell}\right) + \frac{3}{4}\pi$$

e possiamo ricavare

$$(\ell - |\nu|t) = e^{\theta - \frac{3}{4}\pi}$$

da cui

$$\rho(\theta) = \frac{\sqrt{2}}{2}e^{\theta - \frac{3}{4}\pi}$$

Consideriamo un riferimento polare (ρ, θ) con origine nel centro dei quadrati O .

Siano

$(x_a(t), y_a(t))$ $(x_b(t), y_b(t))$ $(x_c(t), y_c(t))$ $(x_d(t), y_d(t))$

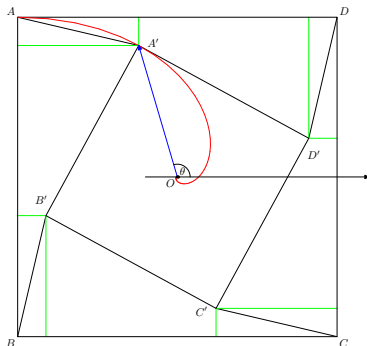
le equazioni parametriche delle curve descritte dai cani in A, B, C, D

Sia

$$\begin{cases} x_a(t) = \rho(t) \cos \theta(t) \\ y_a(t) = \rho(t) \sin \theta(t) \end{cases}$$

dove

$$\rho(t) = \rho(\theta(t))$$



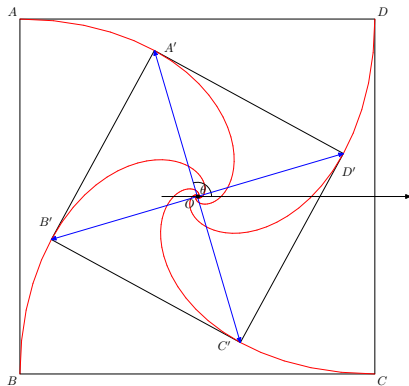
Per simmetria

$$\begin{cases} x_a(t) = \rho(t) \cos \theta(t) \\ y_a(t) = \rho(t) \sin \theta(t) \end{cases}$$

$$\begin{cases} x_d(t) = \rho(t) \sin \theta(t) \\ y_d(t) = -\rho(t) \cos \theta(t) \end{cases}$$

$$\begin{cases} x_c(t) = -\rho(t) \cos \theta(t) \\ y_c(t) = -\rho(t) \sin \theta(t) \end{cases}$$

$$\begin{cases} x_b(t) = -\rho(t) \sin \theta(t) \\ y_b(t) = \rho(t) \cos \theta(t) \end{cases}$$



La traiettoria dei cani sarà quindi definita da

$$(\rho, \theta)$$

Possiamo considerare ρ e θ come funzione del tempo t ed avremo una descrizione completa (geometrica e cinematica) delle traiettorie.

Possiamo anche considerare ρ come funzione di θ , nel qual caso avremo soltanto una descrizione geometrica della traiettoria.

Ovviamente

$$\rho(t) = \rho(\theta(t))$$

dove ρ a primo membro è funzione di t mentre a secondo membro è funzione di θ . Denotiamo

$$\dot{\rho} = \frac{d\rho}{dt} \quad \text{e} \quad \rho' = \frac{d\rho}{d\theta}$$

Il vettore tangente alla curva descritta dal cane in un generico punto $A(t)$ è dato da

$$T(t) = \begin{cases} \dot{x}_a(t) = (\rho'(\theta(t)) \cos \theta(t) - \rho(t) \sin \theta(t)) \dot{\theta}(t) \\ \dot{y}_a(t) = (\rho'(\theta(t)) \sin \theta(t) + \rho(t) \cos \theta(t)) \dot{\theta}(t) \end{cases}$$

Il vettore che indica la direzione che il cane deve seguire è

$$D(t) - A(t) = \begin{cases} \rho(\theta(t)) \sin \theta(t) - \rho(\theta(t)) \cos \theta(t) \\ -\rho(\theta(t)) \cos \theta(t) - \rho(\theta(t)) \sin \theta(t) \end{cases}$$

Dovrà aversi

$$T(t) = k[D(t) - A(t)]$$

dove k è una costante di proporzionalità.

Avremo

$$\begin{cases} (\rho'(\theta(t)) \cos \theta(t) - \rho(t) \sin \theta(t))\dot{\theta}(t) = \\ \quad = k(\rho(\theta(t)) \sin \theta(t) - \rho(\theta(t)) \cos \theta(t)) \\ (\rho'(\theta(t)) \sin \theta(t) + \rho(t) \cos \theta(t))\dot{\theta}(t) = \\ \quad = k(-\rho(\theta(t)) \cos \theta(t) - \rho(\theta(t)) \sin \theta(t)) \end{cases}$$

ed dividendo membro a membro

$$\frac{\rho' \cos \theta - \rho \sin \theta}{\rho' \sin \theta + \rho \cos \theta} = \frac{\rho \sin \theta - \rho \cos \theta}{-\rho \cos \theta - \rho \sin \theta}$$

da cui

$$\begin{aligned} -\dot{\rho} \rho \cos^2 \theta + \rho^2 \sin \theta \cos \theta - \rho \dot{\rho} \cos \theta \sin \theta + \rho^2 \sin^2 \theta = \\ \rho \dot{\rho} \sin^2 \theta - \rho \dot{\rho} \sin \theta \cos \theta + \rho^2 \sin \theta \cos \theta - \rho^2 \cos^2 \theta \end{aligned}$$

ed infine

$$\rho(\theta(t))\rho'(\theta(t)) - \rho^2(\theta(t)) = 0 \quad \text{e} \quad \rho'(\theta(t)) = \rho(\theta(t))$$

Pertanto la geometria della traiettoria è definita dall'equazione

$$\rho'(\theta) = \rho(\theta)$$

e quindi da

$$\rho(\theta) = ce^{\theta}$$

Se teniamo conto che, detto ℓ il lato del quadrato,

$$\rho\left(\frac{3}{4}\pi\right) = \frac{\sqrt{2}}{2}\ell$$

avremo

$$\rho(\theta) = \ell \frac{\sqrt{2}}{2} e^{\theta - \frac{3}{4}\pi}$$

Per ottenerne la legge oraria occorre tenere anche conto del fatto che la curva è percorsa con velocità costante in modulo.

In altre parole

$$\sqrt{\dot{x}^2(t) + \dot{y}^2(t)} = |v|$$

da cui

$$\dot{\rho}^2(t) + \rho^2(t)\dot{\theta}^2(t) = |v|^2$$

e

$$((\rho')^2(t) + \rho^2(t))\dot{\theta}^2(t) = |v|^2 \quad (34)$$

Se ne deduce che

$$\ell^2 e^{2\theta(t)} \dot{\theta}^2(t) = |v|^2$$

$$e^{\theta(t)} \dot{\theta}(t) = \pm \frac{|v|}{\ell}$$

e quindi, dal momento che all'istante iniziale $t = 0$, il cane si trova ad esempio in A , e che θ diminuisce col tempo.

$$e^{\theta(t)} = \frac{3}{4}\pi - \frac{|v|}{\ell}t$$

Infine

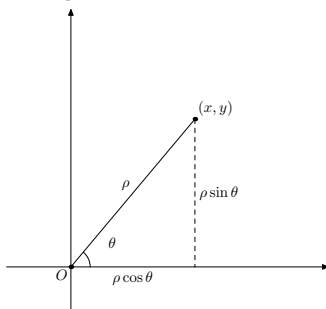
$$\begin{cases} \rho(t) = \rho(\theta(t)) = e^{\theta(t)} = e^{\frac{3}{4}\pi - \frac{|v|}{\ell}t} \\ \theta(t) = \ln \left(e^{\frac{3}{4}\pi - \frac{|v|}{\ell}t} \right) \end{cases}$$

Si definisce spirale logaritmica una curva nel piano descritta da $\gamma(t) = (x(t), y(t))$ tale che il raggio vettore $\gamma(t)$ ed il vettore tangente alla curva $\dot{\gamma}(t)$ formino un angolo fissato α per ogni t .

Usando un riferimento polare possiamo individuare, una curva mediante una funzione $\rho = \rho(\theta)$ essendo

$$\gamma(\theta) = \begin{cases} x = \rho(\theta) \cos \theta \\ y = \rho(\theta) \sin \theta \end{cases}$$

Le sue equazioni parametriche



L'angolo formato da γ e $\dot{\gamma}$ si può individuare mediante la

$$\begin{aligned} \langle \gamma(\theta), \dot{\gamma}(\theta) \rangle &= x(\theta)\dot{x}(\theta) + y(\theta)\dot{y}(\theta) = \\ &= \|(x(\theta), y(\theta))\| \|\dot{x}(\theta), \dot{y}(\theta)\| (\cos \alpha) \end{aligned}$$

e poichè

$$\begin{cases} x(\theta)\dot{x}(\theta) + y(\theta)\dot{y}(\theta) = \frac{d}{d\theta} \frac{1}{2} \rho^2(\theta) = \rho(\theta)\dot{\rho}(\theta) \\ \|(x(\theta), y(\theta))\|^2 = \rho^2(\theta) \\ \|\dot{x}(\theta), \dot{y}(\theta)\|^2 = \dot{\rho}^2(\theta) + \rho^2(\theta) \end{cases}$$

Otteniamo

$$\rho^2(\theta)\dot{\rho}^2(\theta) = (\cos^2 \alpha)\rho^2(\theta)(\dot{\rho}^2(\theta) + \rho^2(\theta))$$

Pertanto

$$\dot{\rho}^2(\theta) = (\cos^2 \alpha)(\dot{\rho}^2(\theta) + \rho^2(\theta))$$

$$\dot{\rho}^2(\theta) = \frac{\cos^2 \alpha}{1 - \cos^2 \alpha} \rho^2(\theta)$$

ed infine

$$\dot{\rho}(\theta) = (\cot \alpha) \rho(\theta)$$

e

$$\rho(\theta) = ke^{(\theta \cot \alpha)}$$

L'ultima equazione descrive la forma della spirale logaritmica ma non dice nulla riguardo al modo in cui essa viene percorsa.

Descrive le proprietà geometriche della spirale logaritmica ma non le sue proprietà cinematiche.

Se supponiamo che θ sia funzione di un variabile t avremo

$$\begin{cases} x(t) = \rho(\theta(t)) \cos \theta(t) \\ y(t) = \rho(\theta(t)) \sin \theta(t) \end{cases}$$

da cui

$$\begin{cases} x(t)\dot{x}(t) + y(t)\dot{y}(t) = \frac{d}{dt} \frac{1}{2} \rho^2(t) = \rho(t)\dot{\rho}(t)\dot{\theta}(t) \\ \|(x(t), y(t))\|^2 = \rho^2(t) \\ \|(\dot{x}(t), \dot{y}(t))\|^2 = (\dot{\rho}^2(t) + \rho^2(t))\dot{\theta}^2(t) \end{cases}$$

e

$$\rho^2(t)\dot{\rho}^2(t)\dot{\theta}^2(t) = (\cos^2 \alpha)\rho^2(t)\dot{\theta}^2(t)(\dot{\rho}^2(t)\rho^2(t))$$

Pertanto

$$\dot{\rho}^2(t) = (\cos^2 \alpha)(\dot{\rho}^2(t) + \rho^2(t))$$

$$\dot{\rho}^2(t) = \frac{\cos^2 \alpha}{1 - \cos^2 \alpha} \rho^2(t)$$

ed infine

$$\dot{\rho}(t) = (\cot \alpha)\rho(t)$$

da cui

$$\rho(t) = ke^{-t \cot \alpha}$$

La velocità è descritta dal vettore $(\dot{x}(t), \dot{y}(t))$ la cui norma è

$$\|(\dot{x}(t), \dot{y}(t))\|^2 = \dot{x}^2(t) + \dot{y}^2(t) = (\dot{\rho}^2(t) + \rho^2(t))\dot{\theta}^2(t)$$

possiamo imporre che la spirale venga percorsa con velocità costante in modulo chiedendo che

$$(\dot{\rho}^2(t) + \rho^2(t))\dot{\theta}^2(t) = |v|$$

da cui si può ricavare $\theta(t)$.

Due specie x e y condividono le risorse di uno stesso territorio
Ciascuna specie avrebbe un tasso di crescita di tipo logistico se visse in un ambiente isolato,

$$\begin{cases} \dot{x}(t) = (a - Ax(t))x(t) \\ \dot{y}(t) = (b - By(t))y(t) \end{cases} \quad \begin{cases} x(t_0) = x_0 \\ y(t_0) = y_0 \end{cases}$$

A causa dei mutui effetti il tasso di crescita; della popolazione x è

$$a - Ax - \alpha y$$

(diminuisce proporzionalmente a y)

Analogamente il tasso di crescita della popolazione y è

$$b - By - \beta x$$

(diminuisce proporzionalmente a x)

Ciascuna specie si sviluppa sottraendo risorse all'altra.
Il sistema che descrive lo sviluppo delle due popolazioni è

$$\begin{cases} \dot{x}(t) = (a - Ax(t) - \alpha y(t))x(t) \\ \dot{y}(t) = (b - By(t) - \beta x(t))y(t) \end{cases} \quad (35)$$

ove $a, b, A, B, \alpha, \beta > 0$, con i dati iniziali

$$\begin{cases} x(t_0) = x_0 \\ y(t_0) = y_0 \end{cases}$$

Dalla prima equazione, dividendo per x ed integrando otteniamo che

$$\frac{\dot{x}(t)}{x(t)} = (a - Ax(t) - \alpha y(t))$$

$$\int_{x_0}^{x(t)} \frac{ds}{s} = \int_{t_0}^t (a - Ax(s) - \alpha y(s)) ds$$

$$\ln \frac{x(t)}{x_0} = \int_{t_0}^t (a - Ax(s) - \alpha y(s)) ds$$

e

$$x(t) = x_0 e^{\int_{t_0}^t (a - Ax(s) - \alpha y(s)) ds}$$

In maniera simile otteniamo che

$$y(t) = y_0 e^{\int_{t_0}^t (b - By(s) - \beta x(s)) ds}$$

Da cui

$$x(t), y(t) > 0$$

Dal sistema 35, per $k = \max\{a, b\}$ si ha

$$\begin{aligned} (x + y)' &= (a - Ax(t) - \alpha y(t))x(t) + \\ &\quad + (b - By(t) - \beta x(t))y(t) \leq k(x + y) \end{aligned}$$

Quindi x ed y sono limitate, in quanto

$$0 \leq (x + y) \leq e^{kt} \leq e^{kT}$$

e quindi esistono in ogni intervallo $[0, T]$.

Il sistema 35 ammette soluzioni costanti che si ottengono risolvendo il sistema algebrico

$$\begin{cases} 0 = (a - Ax - \alpha y)x \\ 0 = (b - By - \beta x)y \end{cases}$$

Le soluzioni costanti sono individuate dai punti di intersezione degli assi con le rette

$$R_1 : (a - Ax - \alpha y) = 0$$

$$R_2 : (b - By - \beta x) = 0$$

La retta R_1 interseca gli assi nei punti

$$(\hat{\xi}, 0) \quad , \quad (0, \eta^{\#})$$

dove

$$\hat{\xi} = \frac{a}{A} \quad \eta^{\#} = \frac{a}{\alpha}$$

La retta R_2 interseca gli assi nei punti

$$(\xi^{\#}, 0) \quad , \quad (0, \hat{\eta})$$

dove

$$\xi^{\#} = \frac{b}{\beta} \quad \hat{\eta} = \frac{b}{B}$$

Le rette R_1 ed R_2 si intersecano nel punto

$$(\tilde{x}, \tilde{y})$$

dove

$$\tilde{x} = \frac{aB - \alpha b}{AB - \alpha\beta} \quad , \quad \tilde{y} = \frac{Ab - a\beta}{AB - \alpha\beta}$$

Pertanto le soluzioni costanti sono individuate dai punti

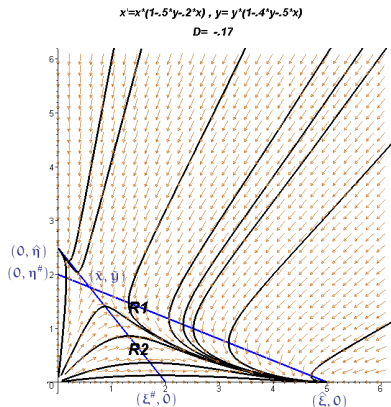
$$(0, 0) \quad , \quad (0, \hat{\eta}) \quad , \quad (\hat{\xi}, 0) \quad , \quad (\tilde{x}, \tilde{y})$$

L'ultimo punto si considera solo nel caso in cui

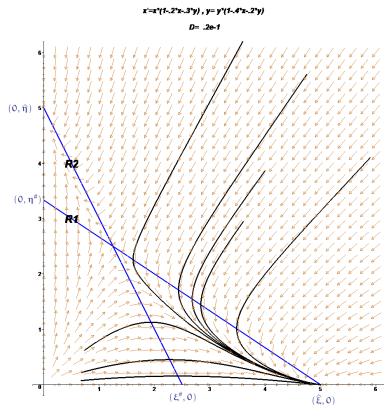
- ① R_1 ed R_2 non siano parallele ($AB - \alpha\beta \neq 0$)
- ② si trovi nel primo quadrante ($\tilde{x} > 0$, $\tilde{y} > 0$).

Possiamo disegnare il campo di vettori associato al sistema e studiare le orbite.

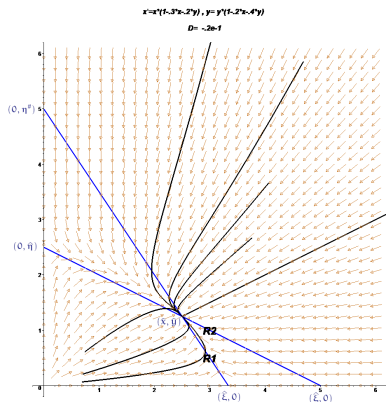
Si presentano quattro casi nei quali



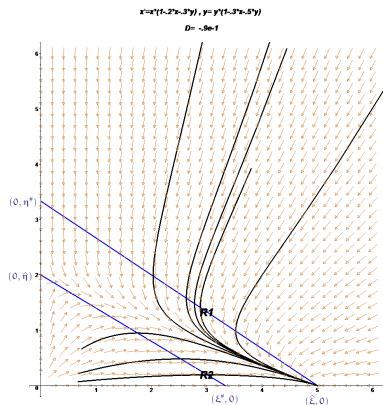
Andamento del numero di individui di due popolazioni in competizione nel caso in cui la popolazione prevalente dipenda dai dati iniziali



Andamento del numero di individui di due popolazioni in competizione nel caso in cui la popolazione prevalente sia la seconda



Andamento del numero di individui di due popolazioni in competizione nel caso in cui le due popolazioni tendano a raggiungere un equilibrio stabile



Andamento del numero di individui di due popolazioni in competizione nel caso in cui la popolazione prevalente sia la prima

Consideriamo la seguente situazione

- 1 La popolazione x , isolata, decrescerebbe secondo la legge

$$\dot{x}(t) = -ax(t)$$

- 2 La popolazione y , isolata, decrescerebbe secondo la legge

$$\dot{y}(t) = -by(t)$$

- 3 Il tasso di crescita della x aumenta proporzionalmente ad y , così che

$$\dot{x}(t) = (-a + \beta y(t))x(t)$$

- 4 Il tasso di crescita della y aumenta proporzionalmente ad x , così che

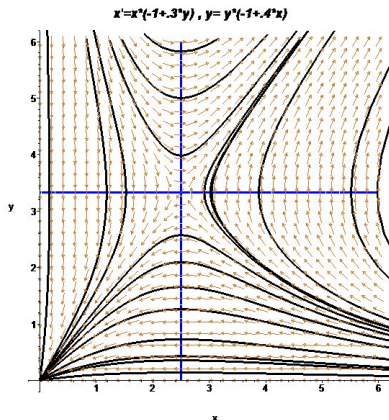
$$\dot{y}(t) = (-b + \alpha x(t))y(t)$$

ne consegue che

$$\begin{cases} \dot{x}(t) = (-a + \beta y(t))x(t) \\ \dot{y}(t) = (-b + \alpha x(t))y(t) \end{cases}$$

Il sistema ammette come soluzioni costanti (punti critici)

$$(0, 0) \quad , \quad \left(\frac{a}{\beta}, \frac{b}{\alpha}\right)$$



ne consegue che

$$\begin{cases} \dot{x}(t) = (-a + \beta y(t))x(t) \\ \dot{y}(t) = (-b + \alpha x(t))y(t) \end{cases}$$

Il sistema ammette come soluzioni costanti (punti critici)

$$(0, 0) \quad , \quad \left(\frac{a}{\beta}, \frac{b}{\alpha}\right)$$

- 1 La x , in assenza della y , crescerebbe secondo la legge logistica

$$\dot{x}(t) = (a - b(x(t)))x(t)$$

- 2 La y , in assenza della x , crescerebbe secondo la legge logistica

$$\dot{y}(t) = (c - dy(t))y(t)$$

- 3 La presenza della y aumenta il tasso di crescita della x proporzionalmente ad x , così che

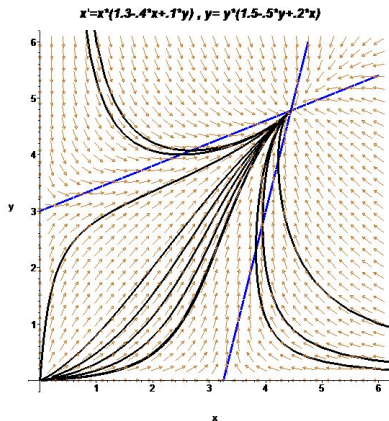
$$\dot{x}(t) = (a - b(x(t) + \gamma y(t)))x(t)$$

- 4 La presenza della x aumenta il tasso di crescita della y proporzionalmente ad x , così che

$$\dot{y}(t) = (c - dy(t) + \delta x(t))y(t)$$

ne consegue che

$$\begin{cases} \dot{x}(t) = (a - b(x(t) + \gamma y(t)))x(t) \\ \dot{y}(t) = (c - d y(t) + \delta x(t))y(t) \end{cases}$$



Che ammette come punti critici

$$E_1 = (0, 0) \quad , \quad E_4 = \left(\frac{A}{D}, \frac{B}{D}\right)$$

$$E_2 = \left(\frac{a}{b}, 0\right) \quad , \quad E_3 = \left(0, \frac{c}{d}\right)$$

dove

$$A = c\gamma + ad$$

$$B = cb + a\delta$$

$$D = bd - \gamma\delta$$

Consideriamo un sistema lineare 2×2 a coefficienti costanti. Cominciamo ad esaminare la stabilità dei sistemi differenziali lineari omogenei a coefficienti costanti di due equazioni in due incognite.

$$\begin{cases} \dot{x}(t) = ax(t) + by(t) \\ \dot{y}(t) = cx(t) + dy(t) \end{cases} \quad a, b, c, d \in \mathbb{R} \quad (36)$$

Possiamo scrivere il sistema anche usando notazioni vettoriali come

$$\dot{u}(t) = Au(t) \quad (37)$$

con

$$u(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \quad \text{ed} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Se $u(t) = (x(t), y(t))$ è la soluzione di un sistema,
Chiamiamo orbita del sistema la curva descritta parametricamente dalle equazioni

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

Le orbite (o traiettorie) si rappresentano nel piano (x, y) (Piano delle Fasi)

La definizione resta valida anche nel caso in cui il sistema non sia lineare.

Come visto precedentemente,

Se il sistema è autonomo per ogni punto del piano delle fasi passa una ed una sola orbita.

Ogni matrice 2×2 può essere trasformata mediante una matrice di passaggio P , non singolare in una matrice canonica C (forma canonica di Jordan)

In altre parole

è possibile trovare P matrice non singolare tale che

$$C = P^{-1}AP$$

sia una delle seguenti matrici, dove λ_1, λ_2 sono gli autovalori di A .

- ① Nel caso di autovalori reali e distinti λ_1, λ_2 ,

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

- ② Nel caso di autovalori complessi e coniugati $\lambda_1 = \alpha + i\beta$,
 $\lambda_2 = \alpha - i\beta$,

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

- ③ Nel caso di autovalori reali e coincidenti $\lambda_1 = \lambda_2 = \lambda$,

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad \text{oppure} \quad \begin{pmatrix} \lambda & 0 \\ \gamma & \lambda \end{pmatrix}$$

Il sistema dato può essere trasformato in un sistema più semplice la cui matrice dei coefficienti rientra in uno dei casi appena elencati, mediante la trasformazione

$$u = Pv \text{ o equivalentemente } v = P^{-1}u$$

Infatti si ottiene

$$P\dot{v} = \dot{u} = Au = APv$$

da cui

$$\dot{v} = P^{-1}APv = Cv$$

Se siamo in grado di studiare le orbite del sistema $\dot{v} = Cv$, otteniamo anche informazioni sul sistema di partenza.

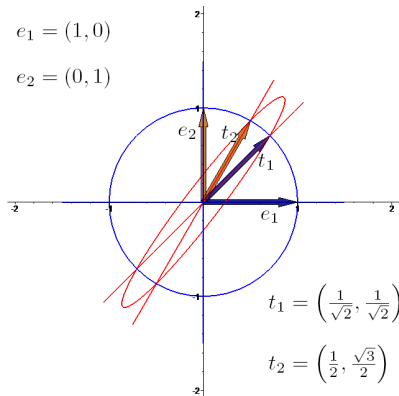
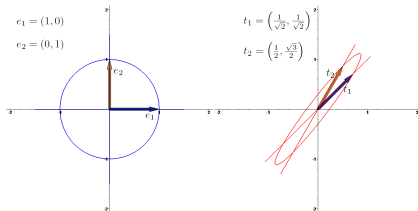
La trasformazione $u = Pv$ è lineare e non degenera e provoca quindi solo una deformazione delle orbite.

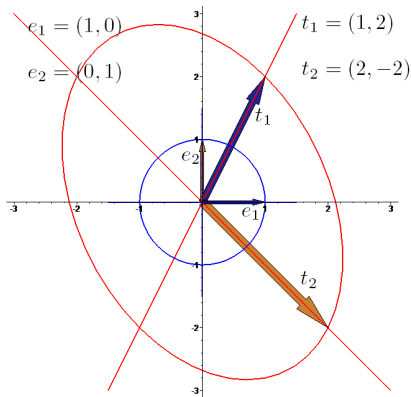
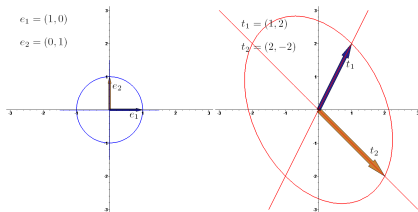
Diventa importante studiare il comportamento delle orbite dei sistemi associati alle matrici canoniche che abbiamo appena elencato.

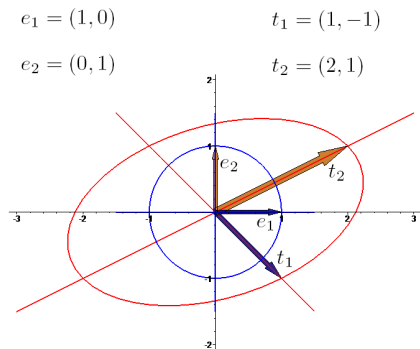
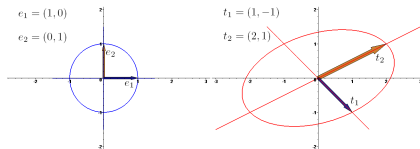
Di seguito sono riportate le immagini degli assi coordinati e del cerchio unitario mediante la trasformazione lineare

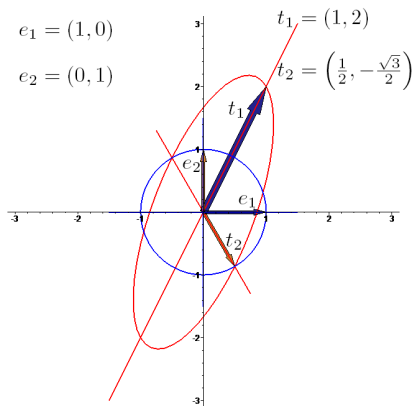
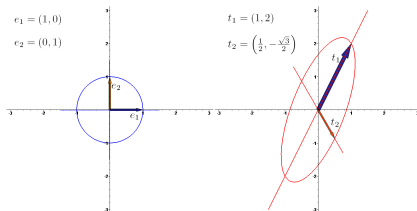
$$\begin{cases} \xi = ax + by \\ \eta = cx + dy \end{cases} \quad \text{ovvero} \quad \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

per diversi casi dei coefficienti









In tal caso il sistema diventa

$$\begin{cases} \dot{\xi} = \lambda_1 \xi \\ \dot{\eta} = \lambda_2 \eta \end{cases}$$

possiamo allora determinare le soluzioni nella forma

$$\begin{cases} \xi(t) = c_1 e^{\lambda_1 t} \\ \eta(t) = c_2 e^{\lambda_2 t} \end{cases}$$

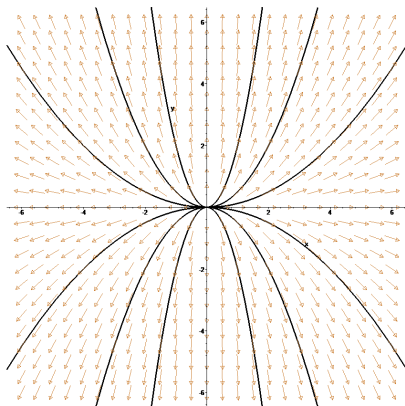
Entrambi gli autovalori sono non nulli

Le traiettorie sono date da $y = x^{\frac{\lambda_2}{\lambda_1}}$ e a seconda del segno di λ_1 e λ_2 e del fatto che $\frac{\lambda_2}{\lambda_1} \gtrless 1$ si possono avere i seguenti casi:

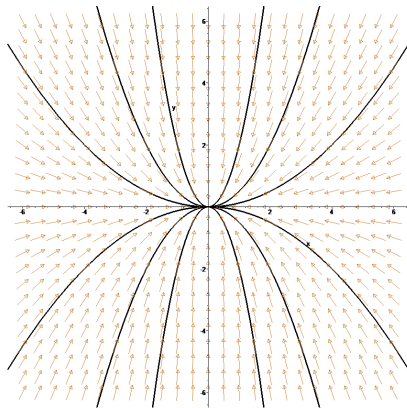
Autovalori reali e distinti $|\lambda_2| > |\lambda_1|$

$$x' = x, \quad y' = 2y$$

$$x' = -x, \quad y' = -2y$$



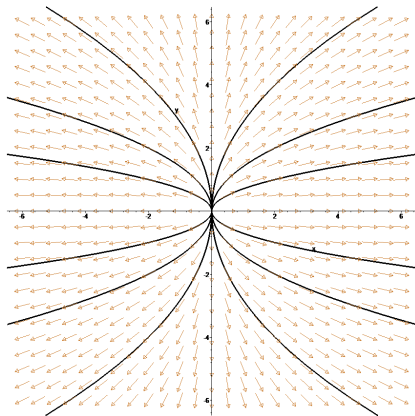
$$\lambda_1, \lambda_2 > 0$$



$$\lambda_1, \lambda_2 < 0$$

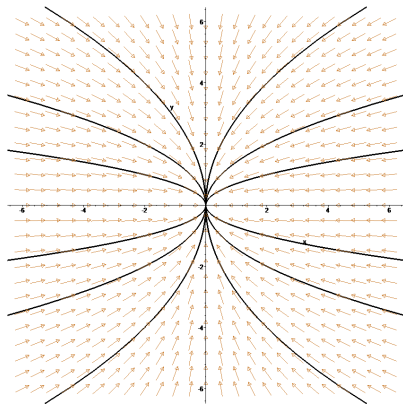
Autovalori reali e distinti $|\lambda_2| < |\lambda_1|$

$$x' = 2x, \quad y' = y$$



$$\lambda_1, \lambda_2 > 0$$

$$x' = -2x, \quad y' = -y$$

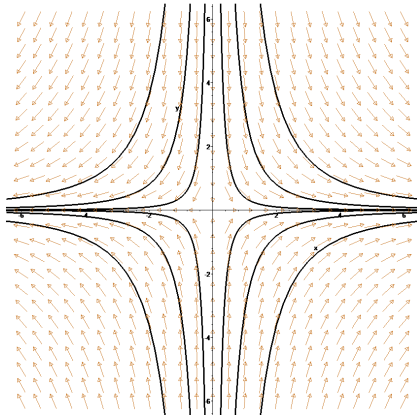


$$\lambda_1, \lambda_2 < 0$$

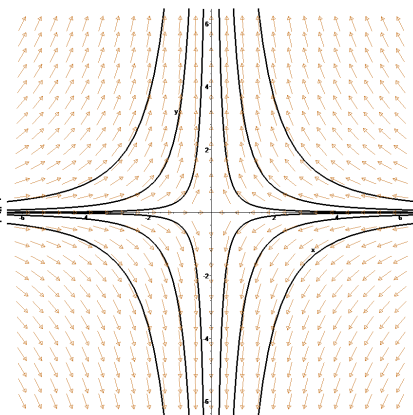
Autovalori reali e distinti $|\lambda_2| > |\lambda_1|$

$$x' = x, \quad y' = -2y$$

$$x' = -x, \quad y' = 2y$$



$$\lambda_1 > 0, \lambda_2 < 0$$



$$\lambda_1 < 0, \lambda_2 > 0$$

Se uno degli autovalori è nullo allora il sistema diventa

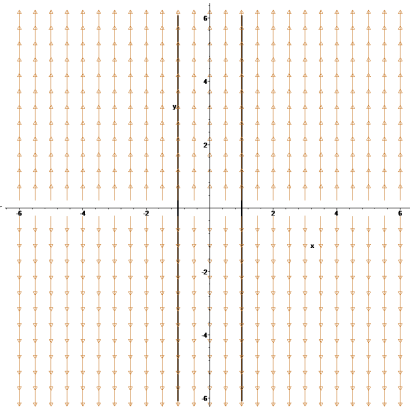
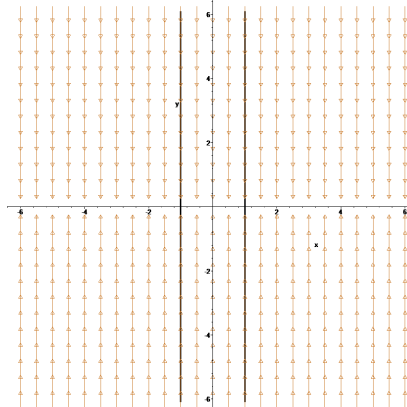
$$\begin{cases} \dot{\xi} = \lambda_1 \xi \\ \dot{\eta} = 0 \end{cases} \quad \text{oppure} \quad \begin{cases} \dot{\xi} = 0 \\ \dot{\eta} = \lambda_2 \eta \end{cases}$$

e le traiettorie sono semirette parallele ad uno degli assi il cui verso di percorrenza dipende dal segno dell'autovalore non nullo.

Autovalori reali e distinti

$$x' = 0, \quad y' = -y$$

$$x' = 0, \quad y' = y$$

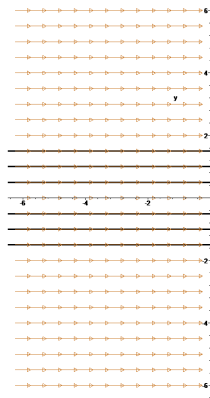


$$\lambda_1 = 0, \lambda_2 < 0$$

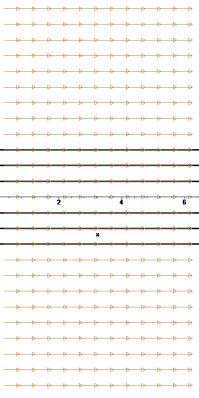
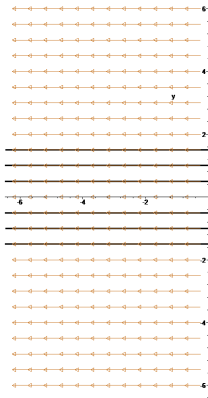
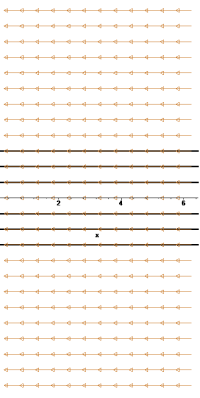
$$\lambda_1 = 0, \lambda_2 > 0$$

Autovalori reali e distinti

$$x' = -x, \quad y' = 0$$



$$x' = x, \quad y' = 0$$



$$\lambda_1 < 0, \lambda_2 = 0$$

$$\lambda_1 > 0, \lambda_2 = 0$$

In tal caso il sistema diventa uno dei due seguenti

$$\begin{cases} \dot{\xi} = \lambda \xi \\ \dot{\eta} = \lambda \eta \end{cases} \quad \text{oppure} \quad \begin{cases} \dot{\xi} = \lambda \xi \\ \dot{\eta} = \gamma \xi + \lambda \eta \end{cases}$$

Supponiamo che $\lambda \neq 0$. Nel primo caso la soluzione del sistema è

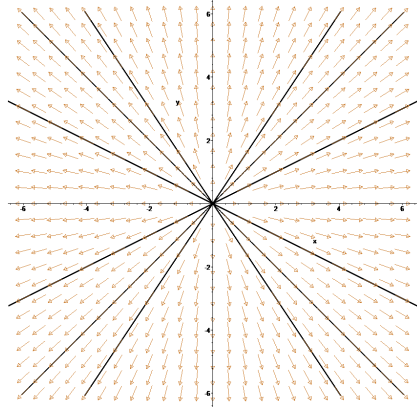
$$\begin{cases} \xi(t) = c_1 e^{\lambda t} \\ \eta(t) = c_2 e^{\lambda t} \end{cases}$$

e le orbite sono semirette che passano per l'origine, hanno coefficiente angolare $\frac{c_2}{c_1}$ e sono percorse con verso che si avvicina o si allontana dall'origine a seconda del segno di λ .

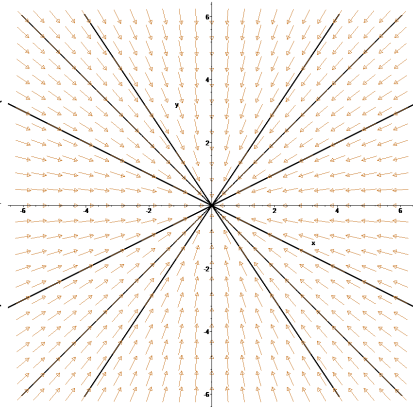
Autovalori reali e coincidenti non nulli

$$x' = x, \quad y' = y$$

$$x' = -x, \quad y' = -y$$



$$\lambda_1, \lambda_2 > 0$$



$$\lambda_1, \lambda_2 < 0$$

Nel secondo caso la soluzione del sistema è

$$\begin{cases} \xi(t) = c_1 e^{\lambda t} \\ \eta(t) = (\gamma c_1 t + c_2) e^{\lambda t} \end{cases}$$

da cui si ricava

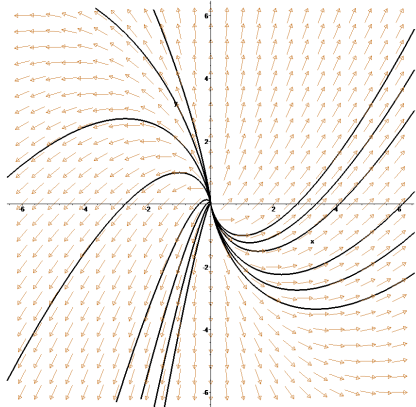
$$\begin{cases} t = \frac{1}{\lambda} \ln \left(\frac{\xi}{c_1} \right) \\ \eta = \eta(t) = \left(\gamma \frac{1}{\lambda} \ln \left(\frac{\xi}{c_1} \right) + \frac{c_2}{c_1} \right) \xi \end{cases}$$

e si possono disegnare le curve che rappresentano le orbite del sistema.

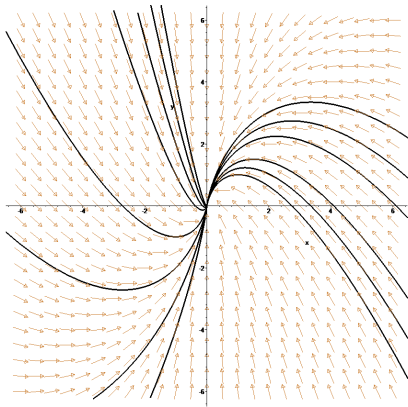
Autovalori reali e coincidenti non nulli

$$x' = x, \quad y' = x + y$$

$$x' = -x, \quad y' = x - y$$



$$\lambda_1, \lambda_2 > 0$$



$$\lambda_1, \lambda_2 < 0$$

Se infine $\lambda = 0$ otteniamo che, nel primo caso, ogni punto è un punto singolare, mentre nel secondo caso le soluzioni sono date da

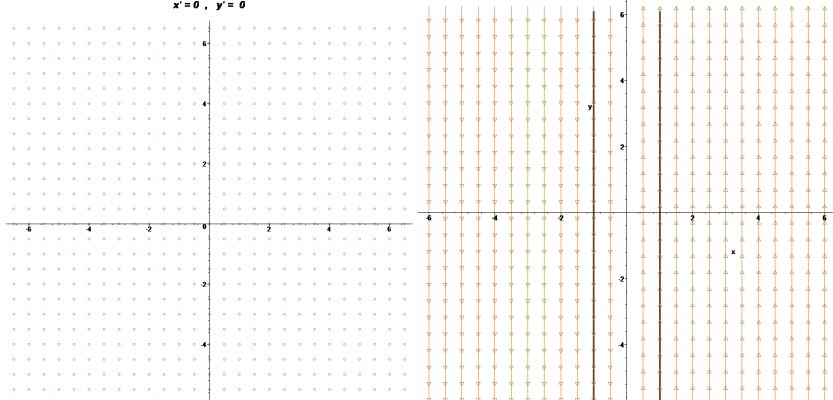
$$\begin{cases} \xi(t) = c_1 \\ \eta(t) = \gamma c_1 t + c_2 \end{cases}$$

e le orbite sono ancora rette parallele ad un'asse.

Autovalori reali e coincidenti nulli

$x' = 0, y' = x$

$x' = 0, y' = 0$



In tal caso il sistema diventa

$$\begin{cases} \dot{\xi} = \alpha\xi - \beta\eta \\ \dot{\eta} = \beta\xi + \alpha\eta \end{cases}$$

e può essere studiato mediante un ulteriore cambio di variabili. Infatti se poniamo

$$\begin{cases} \xi = \rho \cos \theta \\ \eta = \rho \sin \theta \end{cases}$$

possiamo ricavare che

$$\rho^2(t) = \xi^2(t) + \eta^2(t) \quad \tan \theta(t) = \frac{\eta(t)}{\xi(t)}$$

da cui, derivando si ottiene

$$2\rho(t)\dot{\rho}(t) = 2\xi(t)\dot{\xi}(t) + 2\eta(t)\dot{\eta}(t)$$

e

$$(1 + \tan^2 \theta(t))\dot{\theta}(t) = \frac{\xi(t)\dot{\eta}(t) - \eta(t)\dot{\xi}(t)}{\xi^2(t)}$$

e tenendo conto del sistema si deduce che

$$\begin{cases} \dot{\rho} = \alpha\rho \\ \dot{\theta} = \beta \end{cases}$$

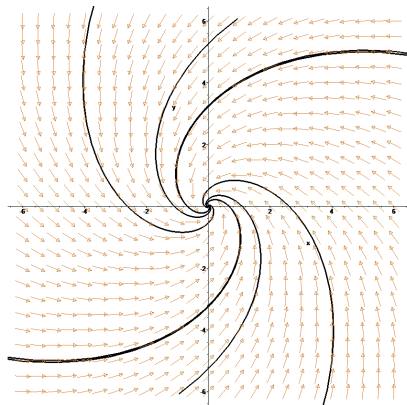
da cui si ricava che

$$\begin{cases} \rho(t) = C_1 e^{\alpha t} \\ \theta(t) = \beta t + C_2 \end{cases}$$

Le traiettorie sono spirali o, nel caso in cui $\alpha = 0$, cerchi. Il verso di percorrenza è determinato dal segno di β , mentre le spirali che tendono all'origine si distinguono dalle spirali che se ne allontanano mediante il segno di α (negativo o positivo rispettivamente)

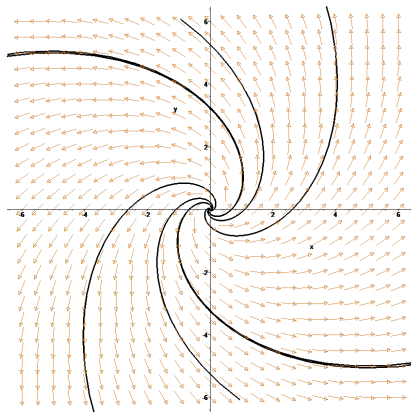
Autovalori complessi e coniugati con parte reale non nulla

$$x' = -x - y, \quad y' = x - y$$



Parte reale negativa

$$x' = x - y, \quad y' = x + y$$

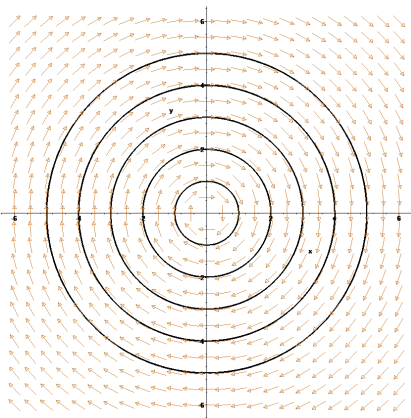
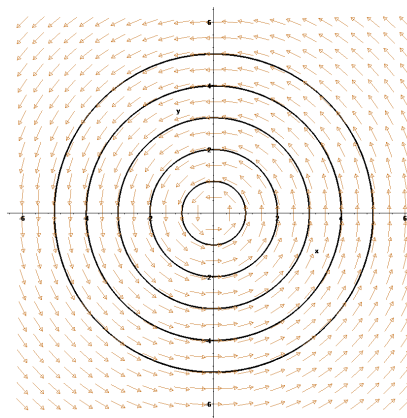


Parte reale positiva

Autovalori complessi e coniugati con parte reale nulla

$$x' = -y, \quad y' = x$$

$$x' = y, \quad y' = -x$$



$$\lambda_1, \lambda_2 > 0$$

$$\lambda_1, \lambda_2 < 0$$

Sia $A(t)$ una funzione a valori matrici $n \times n$, continua su $[t_0, +\infty)$ e consideriamo il sistema lineare

$$\dot{x}(t) = A(t)x(t)$$

Sia G una matrice fondamentale del sistema; allora la soluzione identicamente nulla è stabile se e solo se

$$\|G(t)\| \leq K \quad , \quad \forall t \geq t_0$$

Inoltre la soluzione nulla è asintoticamente stabile se e solo se

$$\lim_{t \rightarrow +\infty} \|G(t)\| = 0$$

Infatti la soluzione $x(t)$ del problema di Cauchy

$$\begin{cases} \dot{x}(t) = A(t)x(t) \\ x(t_0) = x_0 \end{cases}$$

si può esprimere nella forma

$$x(t) = G(t)G^{-1}(t_0)x_0 \quad (38)$$

e

$$\|x(t)\| \leq \|G(t)\| \|G^{-1}(t_0)\| \|x_0\|.$$

Ciò permette di concludere che le condizioni proposte per la stabilità e l'asintotica stabilità della soluzione nulla sono sufficienti.

Viceversa sia la soluzione nulla stabile; per $\|x_0\| < \delta$ si ha $\|x(t)\| < 1$

Sia $D = \{\delta e_i, i = 1..n\}$; D è una base di $\mathbb{R}^n$,

Ogni soluzione $x_{\delta e_i}$ che corrisponde al dato iniziale $x_0 = \delta e_i$, è limitata in norma;

$x_{\delta e_i}$ sono le colonne di una matrice fondamentale che risulta pertanto limitata.

Analogamente, se la soluzione nulla è asintoticamente stabile

$$\lim_{t \rightarrow +\infty} \|x_{\delta e_i}(t)\| = 0.$$

e quindi

$$\lim_{t \rightarrow +\infty} \|G(t)\| = 0$$

Se la matrice dei coefficienti è costante le soluzioni sono combinazioni lineari di funzioni del tipo

$$t^\gamma e^{\alpha t} \cos(\beta t) \quad , \quad t^\gamma e^{\alpha t} \sin(\beta t)$$

in corrispondenza degli autovalori $\alpha \pm i\beta$, con molteplicità γ della matrice dei coefficienti.

Esaminando il comportamento di tali funzioni si deduce subito che

- ① se A ha tutti gli autovalori con parte reale negativa, allora la soluzione nulla è asintoticamente stabile
- ② se A ha qualche autovalore con parte reale positiva, allora la soluzione nulla non è stabile (ne' asintoticamente stabile, ovviamente)
- ③ se A ha tutti gli autovalori con parte reale negativa o nulla, e gli autovalori con parte reale nulla sono semplici, allora la soluzione è stabile ma non asintoticamente stabile
- ④ se A ha autovalori con parte reale negativa o nulla, e gli autovalori con parte reale nulla hanno molteplicità superiore 1, allora la soluzione non è asintoticamente stabile ma può essere sia stabile che instabile a seconda della forma che assumono le soluzioni.

Se il sistema non è lineare, la stabilità è una questione più delicata

Si può, ad esempio, risolvere utilizzando il metodo delle funzioni di Lyapunov;

Talvolta è possibile studiare la stabilità di un sistema non lineare utilizzando informazioni sulla stabilità del sistema lineare ottenuto usando lo sviluppo di Taylor del primo ordine della funzione che compare a secondo membro. (Sistema Linearizzato) .
È interessante conoscere sotto quali condizioni la stabilità del sistema linearizzato è sufficiente per la stabilità del sistema originario.

(stabilità in prima approssimazione)

A questo proposito possiamo provare che se

$$\dot{x}(t) = Ax(t) + f(x(t)) \quad (39)$$

dove A è una matrice $n \times n$, $S = \{y \in \mathbb{R}^n : \|y\| < a\}$ ed $f : S \rightarrow \mathbb{R}$, è continua e tale che $f(0) = 0$ e

$$\lim_{x \rightarrow 0} \frac{f(x)}{\|x\|} = 0$$

Se tutti gli autovalori di A hanno parte reale negativa, allora la soluzione nulla è asintoticamente stabile per il sistema 39 se invece la soluzione nulla non è stabile per il sistema lineare

$$\dot{x}(t) = Ax(t) \quad (40)$$

allora non è stabile neanche per il sistema 39

Dimostrazione. Sia G la matrice fondamentale principale del sistema linearizzato $\dot{x} = Ax$ e sia x la soluzione di 39 definita in un intervallo massimale $[t_0, b)$; poiché gli autovalori di A hanno parte reale negativa esiste $\alpha > 0$ tale che

$$\|G(t)\| \leq Ke^{-\alpha t}$$

Se $\|x\| < \sigma$ si ha

$$\|f(x)\| \leq \frac{\alpha}{2K} \|x\|;$$

e possiamo supporre che $a < \sigma$.

Poniamo

$$z(t) = G(t - t_0)x_0 + \int_{t_0}^t G(t - s)f(x(s))ds$$

si ha

$$\begin{aligned} z'(t) &= G'(t - t_0)x_0 + \int_{t_0}^t G'(t - s)f(x(s))ds + G(0)f(x(t)) = \\ &= A \left(G(t - t_0)x_0 + \int_{t_0}^t G(t - s)f(x(s))ds \right) + f(x(t)) = \\ &= Az(t) + f(x(t)) \end{aligned}$$

e

$$z(t_0) = x_0.$$

Tenuto conto del fatto che

$$\dot{x}(t) = Ax(t) + f(x(t)) \quad , \quad x(t_0) = x_0$$

si ha

$$(z - x)'(t) = A(z - x)(t) \quad , \quad (z - x)(t_0) = 0$$

da cui

Pertanto

$$x(t) = G(t - t_0)x_0 + \int_{t_0}^t G(t - s)f(x(s))ds$$

Allora

$$\|x(t)\| \leq \kappa e^{-\alpha(t-t_0)}\|x_0\| + \int_{t_0}^t \frac{\alpha}{2} e^{-\alpha(t-s)}\|x(s)\|ds$$

da cui

$$e^{\alpha(t-t_0)}\|x(t)\| \leq \kappa\|x_0\| + \int_{t_0}^t \frac{\alpha}{2} e^{\alpha(s-t_0)}\|x(s)\|ds$$

e per il lemma di Gronwall

$$e^{\alpha(t-t_0)} \|x(t)\| \leq K \|x_0\| e^{\alpha(t-t_0)/2}$$

e

$$\|x(t)\| \leq K \|x_0\| e^{-\alpha(t-t_0)} \leq K \|x_0\| \quad , \quad \forall t \geq t_0 \quad (41)$$

Pertanto si ha, se $0 < \varepsilon < \sigma$ e $\|x_0\| < \varepsilon/K$,

$$\|x(t)\| < \varepsilon \quad , \quad \forall t \in [t_0, b)$$

e poichè la soluzione è prolungabile, deve essere $b = +\infty$.
Passando al limite per $t \rightarrow +\infty$ nella 41 si conclude anche l'asintotica stabilità.

Per verificare la seconda affermazione si può provare che se G ha qualche autovalore positivo si può scegliere una soluzione x in modo che $x(t) \rightarrow +\infty$ quando $t \rightarrow +\infty$, scegliendo opportunamente i dati iniziali.

Consideriamo una matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Gli autovalori sono le soluzioni dell'equazione

$$\det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc) \quad (42)$$

Se $T = a + d$ è la traccia e $D = ad - bc$ è il determinante della matrice possiamo riscrivere la 42 come

$$\lambda^2 - T\lambda + D$$

e gli autovalori sono

$$\lambda = \frac{T}{2} \pm \sqrt{\frac{T^2}{4} - D} \quad (43)$$

- Se $D > 0$

$$\frac{T^2}{4} > \frac{T^2}{4} - D$$

e

- Nel caso in cui $\frac{T^2}{4} - D \geq 0$

$$\left| \frac{T}{2} \right| > \sqrt{\frac{T^2}{4} - D}$$

Gli autovalori λ_1, λ_2 hanno il segno di T

- Nel caso in cui $\frac{T^2}{4} - D < 0$

Gli autovalori sono complessi e coniugati la loro parte reale $\Re \lambda_1, \Re \lambda_2$ ha lo stesso segno di T

- Se $D = 0$

$$\lambda_1 = 0, \quad \lambda_2 = T$$

- Se $D < 0$

Si ha $\frac{T^2}{4} - D > \frac{T^2}{4}$

Gli autovalori hanno uno segno positivo e l'altro negativo.

Serve a determinare il segno della parte reale delle radici di un polinomio.

Si può quindi usare per determinare il segno delle parti reali degli autovalori di una matrice.

Consideriamo un polinomio a coefficienti reali

$$P(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n$$

e la corrispondente matrice di Hurwitz definita da

$$H = \begin{pmatrix} a_1 & 1 & 0 & 0 & \cdots & 0 \\ a_3 & a_2 & a_1 & 1 & \cdots & 0 \\ a_5 & a_4 & a_3 & a_2 & \cdots & 0 \\ a_7 & a_6 & a_5 & a_4 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & a_n \end{pmatrix}$$

Condizione necessaria e sufficiente affinché $P(z)$ abbia radici tutte con parte reale negativa è che H abbia tutti i minori principali con determinante positivo.

Per costruire la matrice di Hurwitz si può osservare che sulla diagonale sono presenti i coefficienti del polinomio.

Se chiamiamo Δ_k il minore principale di ordine k cioè se

$$\Delta_1 = (a_1) \quad \Delta_2 = \begin{pmatrix} a_1 & 1 \\ a_3 & a_2 \end{pmatrix} \quad \Delta_3 = \begin{pmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{pmatrix}$$

$$\Delta_n = \begin{pmatrix} a_1 & 1 & 0 & 0 & \cdots & 0 \\ a_3 & a_2 & a_1 & 1 & \cdots & 0 \\ a_5 & a_4 & a_3 & a_2 & \cdots & 0 \\ a_7 & a_6 & a_5 & a_4 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & a_n \end{pmatrix}$$

$$P(z) = z^2 + a_1 z + a_2$$

ha radici tutte con parte reale negativa se

$$a_1 > 0 \quad , \quad a_2 > 0$$

$$P(z) = z^3 + a_1 z^2 + a_2 z + a_3$$

ha radici tutte con parte reale negativa se

$$a_1 > 0 \quad , \quad a_1 a_2 - a_3 > 0 \quad , \quad a_3 > 0$$

$$P(z) = z^4 + a_1 z^3 + a_2 z^2 + a_3 z + a_4$$

ha radici tutte con parte reale negativa se
i

$$a_1 > 0 \quad , \quad a_1 a_2 - a_3 > 0 \quad , \quad (a_1 a_2 - a_3) a_3 - a_1^2 a_4 > 0 \quad , \quad a_3 > 0$$

Lo sviluppo di due specie x e y che condividono le risorse di uno stesso ambiente è descritto dal sistema

$$\begin{cases} \dot{x}(t) = (a - Ax(t) - \alpha y(t))x(t) \\ \dot{y}(t) = (b - By(t) - \beta x(t))y(t) \end{cases} \quad (44)$$

con

$$\begin{cases} x(t_0) = x_0 \\ y(t_0) = y_0 \end{cases}$$

Ove $a, b, A, B, \alpha, \beta$ sono positivi.

x ed y sono limitate,

$$0 \leq (x + y) \leq e^{kt} \leq e^{kT}$$

e quindi esistono in ogni intervallo $[0, T]$.

Le soluzioni costanti sono individuate dai punti di intersezione degli assi con le rette

$$R_1 : (a - Ax - \alpha y) = 0$$

$$R_2 : (b - By - \beta x) = 0$$

La retta R_1 interseca gli assi nei punti

$$(\hat{\xi}, 0) \quad , \quad (0, \eta^{\#})$$

dove

$$\hat{\xi} = \frac{a}{A} \quad \eta^{\#} = \frac{a}{\alpha}$$

La retta R_2 interseca gli assi nei punti

$$(\xi^{\#}, 0) \quad , \quad (0, \hat{\eta})$$

dove

$$\xi^{\#} = \frac{b}{\beta} \quad \hat{\eta} = \frac{b}{B}$$

Per studiare la stabilità delle soluzioni costanti possiamo considerare il sistema linearizzato relativamente ad ognuna delle soluzioni costanti $(\bar{x}, \bar{y})$.

Il sistema linearizzato in $(\bar{x}, \bar{y})$ è associato alla matrice

$$M = \begin{pmatrix} a - 2A\bar{x} - \alpha\bar{y} & -\alpha\bar{x} \\ -\beta\bar{y} & b - 2B\bar{y} - \beta\bar{x} \end{pmatrix}$$

Per $(\bar{x}, \bar{y}) = (0, 0)$

$$M = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

M ha autovalori positivi e quindi la relativa soluzione è instabile.

Per $(\bar{x}, \bar{y}) = (\hat{\xi}, 0)$

$$M = \begin{pmatrix} a - 2A\frac{a}{A} & -\alpha\frac{a}{A} \\ 0 & b - \beta\frac{a}{A} \end{pmatrix} = \begin{pmatrix} -a & -\alpha\frac{a}{A} \\ 0 & \beta\left(\frac{b}{\beta} - \frac{a}{A}\right) \end{pmatrix} =$$

$$= \begin{pmatrix} -a & -\alpha\frac{a}{A} \\ 0 & \beta(\hat{\xi} - \xi^{\#}) \end{pmatrix}$$

M ha autovalori $-a$ e $\beta(\hat{\xi} - \xi^{\#})$, la stabilità dipende dalla mutua posizione di $\xi^{\#}$ e $\hat{\xi}$.

Per $(\bar{x}, \bar{y}) = (0, \hat{\eta})$

$$M = \begin{pmatrix} a - \alpha\frac{b}{B} & 0 \\ -\beta\frac{b}{B} & b - 2B\frac{b}{B} \end{pmatrix} = \begin{pmatrix} \alpha\left(\frac{a}{\alpha} - \frac{b}{B}\right) & 0 \\ -\beta\frac{b}{B} & -b \end{pmatrix} =$$

$$= \begin{pmatrix} \alpha(\hat{\eta} - \eta^{\#}) & 0 \\ -\beta\frac{b}{B} & -b \end{pmatrix}$$

M ha autovalori $-b$ e $\alpha(\hat{\eta} - \eta^{\#})$, la stabilità dipende dalla mutua posizione di $\eta^{\#}$ e $\hat{\eta}$.

Per $(\bar{x}, \bar{y}) = (\tilde{x}, \tilde{y})$

$$M = \begin{pmatrix} -A\tilde{x} & -\alpha\tilde{x} \\ -\beta\tilde{y} & -B\tilde{y} \end{pmatrix}$$

se indichiamo con D il determinante di M e con T la traccia di M , avremo che

$$D = \tilde{x}\tilde{y}(AB - \alpha\beta) \neq 0$$

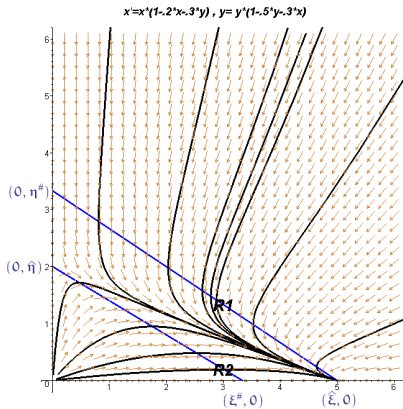
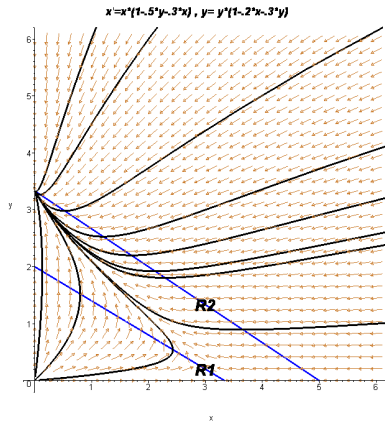
$$T = -(A\tilde{x} + B\tilde{y}) (< 0)$$

Se $D > 0$ si ha

$$\frac{T^2}{4} - D = (\tilde{x}A - \tilde{y}B)^2 + 4\tilde{x}\tilde{y}\alpha\beta > 0$$

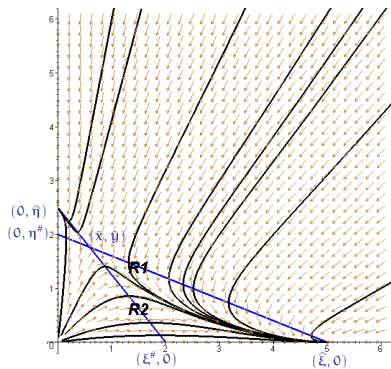
dal momento che $\tilde{x} > 0$, $\tilde{y} > 0$, gli autovalori di M sono sempre reali e distinti. hanno lo stesso segno di T (soluzione stabile)

Se $D < 0$ M ha due autovalori reali di segno opposto. (soluzione instabile)



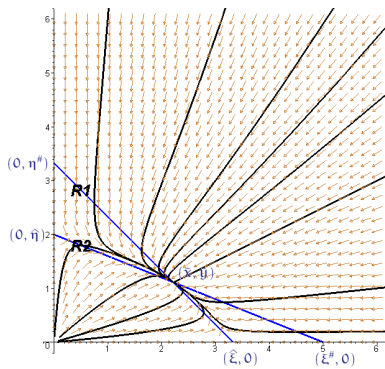
$$x' = x'(1 - 5y - 2x), \quad y' = y'(1 - 4y - 5x)$$

$$D = -.17$$



$$x' = x'(1 - 3x - 3y), \quad y' = y'(1 - 5y - 2x)$$

$$D = .9e-1$$



Il sistema che descrive il fenomeno è

$$\begin{cases} \dot{x}(t) = (-a + \beta y(t))x(t) \\ \dot{y}(t) = (-b + \alpha x(t))y(t) \end{cases}$$

Ammette come soluzioni costanti (punti critici)

$$(0, 0) \quad , \quad \left(\frac{b}{\alpha}, \frac{a}{\beta} \right)$$

La stabilità può essere studiata linearizzando

$$\nabla \begin{pmatrix} (-a + \beta y)x \\ (-b + \alpha x)y \end{pmatrix} = \begin{pmatrix} -a + \beta y & \beta x \\ \alpha y & -b + \beta x \end{pmatrix}$$

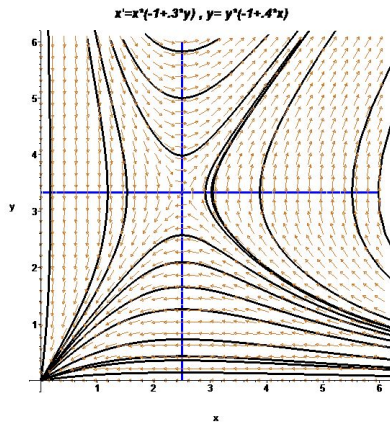
Per $x = 0, y = 0$ la matrice del sistema linearizzato è

$$\begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix}$$

La matrice ha autovalori entrambi negativi La soluzione nulla è asintoticamente stabile.

Per $x = \frac{a}{\beta}, y = \frac{b}{\alpha}$ La matrice del sistema linearizzato è

$$\begin{pmatrix} 0 & b\frac{\beta}{\alpha} \\ b\frac{\alpha}{\beta} & 0 \end{pmatrix}$$



La matrice ha autovalori

$$\sqrt{ab} \quad , \quad -\sqrt{ab}$$

La soluzione nulla è instabile.

Il sistema che descrive la situazione è

$$\begin{cases} \dot{x}(t) = (a - b(x(t) + \gamma y(t)))x(t) \\ \dot{y}(t) = (c - d y(t) + \delta x(t))y(t) \end{cases}$$

che ammette come punti critici

$$\begin{aligned} E_1 &= (0, 0) \quad , \quad E_4 = \left(\frac{A}{D}, \frac{B}{D} \right) \\ E_2 &= \left(\frac{a}{b}, 0 \right) \quad , \quad E_3 = \left(0, \frac{c}{d} \right) \end{aligned}$$

dove

$$\begin{aligned} A &= c\gamma + ad \\ B &= cb + a\delta \\ D &= bd - \gamma\delta \end{aligned}$$

La matrice Jacobiana per $\bar{E} = (\bar{x}, \bar{y})$ è

$$M = \begin{pmatrix} a - 2b\bar{x} + \gamma\bar{y} & \gamma\bar{x} \\ \delta\bar{y} & c - 2d\bar{y} + \delta\bar{x} \end{pmatrix}$$

Per $(\bar{x}, \bar{y}) = E_1 = (0, 0)$

$$M = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}$$

Gli autovalori sono $a, b > 0$

Il punto $E_1 = (0, 0)$ è instabile

Per $(\bar{x}, \bar{y}) = E_2 = (\frac{a}{b}, 0)$ la matrice del sistema linearizzato è

$$M = \begin{pmatrix} -a & \gamma \frac{a}{b} \\ 0 & c + \delta \frac{a}{b} \end{pmatrix} \quad -a < 0 \quad c + \delta \frac{a}{b} > 0$$

Gli autovalori sono $-a < 0$, $c + \delta \frac{a}{b} > 0$

Il punto $E_2 = E_2 = (\frac{a}{b}, 0)$ è un punto sella .

Per $(\bar{x}, \bar{y}) = E_3 = (0, \frac{c}{d})$ la matrice del sistema linearizzato è

$$M = \begin{pmatrix} a + \gamma \frac{c}{d} & 0 \\ \gamma \frac{c}{d} & -c \end{pmatrix}$$

Gli autovalori sono $a + \gamma \frac{c}{d} > 0$, $-c < 0$

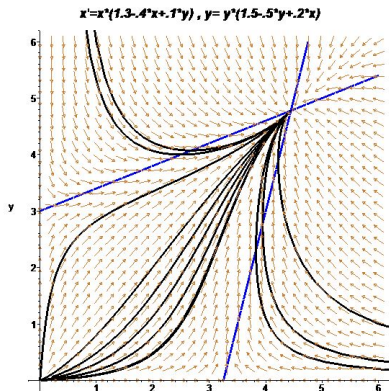
Il punto $E_3 = (0, \frac{c}{d})$ è un punto sella

Per $(\bar{x}, \bar{y}) = E_4 = (\frac{A}{D}, \frac{B}{D})$

Il punto $E_4 = (\frac{A}{D}, \frac{B}{D})$ appartiene al primo quadrante solo se $D > 0$ perchè il modello abbia senso questa condizione deve essere soddisfatta.

In tal caso la matrice del sistema linearizzato è

$$M = \begin{pmatrix} -b\frac{A}{D} & -\gamma\frac{A}{D} \\ \delta\frac{B}{D} & -d\frac{B}{D} \end{pmatrix}$$



$$\det M = (bd - \gamma\delta)AB = DAB > 0$$

mentre la traccia di M è negativa. Quindi il punto è stabile

Il Calcolo delle Variazioni.

Il problema classico di calcolo delle variazioni consiste nel minimizzare un funzionale integrale che dipende da una funzione e dalla sua derivata.

Più precisamente se

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}$$

è abbastanza regolare e se $\mathcal{C}^1$ è lo spazio delle funzioni

$$x : [a, b] \rightarrow \mathbb{R}$$

continue con la loro derivata, possiamo porci il problema di trovare

$$\min_{x \in \mathcal{C}^1, x(a)=\alpha, x(b)=\beta} \left\{ \int_a^b f(x(t), \dot{x}(t)) dt \right\}$$

La curva di lunghezza minima per due punti fissati

Determinare la curva $y(x)$, nel piano (x, y) , che passa per due punti

$$P_0 = (a, y_0) \quad \text{e} \quad P_1 = (b, y_1)$$

mediante la quale si passa dal punto P_0 al punto P_1 percorrendo la minima distanza.

La lunghezza di una curva si calcola mediante la

$$\int_a^b \sqrt{1 + (y'(x))^2} dx$$

Il problema si riduce a trovare

$$\min_{y \in C^1, y(a)=y_0, y(b)=y_1} \left\{ \int_a^b \sqrt{1 + (y'(x))^2} dx \right\}$$

La brachistocrona

Trovare la curva da seguire per minimizzare il tempo di percorrenza necessario perchè un punto materiale soggetto alla sola forza di gravità si sposti da un punto $A = (a, h)$, $h > 0$ ad un punto $B = (b, 0)$.

Indichiamo con $y(x)$ la funzione il cui grafico è percorso per passare dal punto A al punto B .

Il punto materiale descriverà una traiettoria rispetto alla quale la lunghezza d'arco è data da

$$s(x) = \int_a^x \sqrt{1 + (y'(x))^2} dx$$

Se consideriamo una partizione dell'intervallo $[a, b]$

$$\{a = x_0 < x_1 < x_2 < \dots < x_n = b\}$$

la lunghezza percorsa relativa all'intervallo $[x_k, x_{k+1}]$ è

$$\int_{x_k}^{x_{k+1}} \sqrt{1 + (y'(x))^2} dx = \sqrt{1 + (y'(c_k))^2} (x_{k+1} - x_k)$$

dove $c_k \in [x_k, x_{k+1}]$

La velocità di caduta dipende solo dalla quota per il principio di conservazione dell'energia e si ha

$$\frac{1}{2}mv^2 = mg(h - y) \quad \text{e} \quad v = \sqrt{2g(h - y(x))}$$

Nell'intervallo $[x_k, x_{k+1}]$ possiamo considerare che

$$y(x) = y(d_k) \quad \text{con} \quad d_k \in [x_k, x_{k+1}]$$

per cui

$$v = \sqrt{2g(h - y(d_k))}$$

Il tempo t_k necessario per percorrere la parte di curva relativa all'intervallo $[x_k, x_{k+1}]$ sarà

$$t_k = \frac{\sqrt{1 + (y'(c_k))^2}}{\sqrt{2g(h - y(d_k))}}(x_{k+1} - x_k) \quad , \quad c_k, d_k \in [x_k, x_{k+1}]$$

Sommando su k e passando al limite, si ottiene

$$t = \int_a^b \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{2g(h - y(x))}} dx \quad (45)$$

Il problema è ridotto a minimizzare il funzionale definito nella 45 sulle funzioni y di classe $\mathcal{C}^1$ tali che $y(a) = h$ e $y(b) = 0$.

$$\min_{y \in \mathcal{C}^1, y(a)=h, y(b)=0} \int_a^b \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{2g(h - y(x))}} dx$$

Il solido di rotazione avente minima superficie

Determinare il solido di rotazione che ha minima superficie.
In un riferimento cartesiano ortogonale (x, y, z) consideriamo una curva $y(x)$, $x \in [a, b]$ tale che

$$y(a) = \alpha \quad \text{e} \quad y(b) = \beta$$

e la superficie da essa generata dalla rotazione attorno all'asse x .
Possiamo parametrizzare la superficie nella forma

$$\begin{cases} x = x \\ y = y(x) \cos \theta \\ z = y(x) \sin \theta \end{cases} \quad x \in [a, b]$$

si ha

$$\frac{\partial x, y, z}{\partial x, \rho, \theta} = \begin{pmatrix} 1 & y'(x) \cos \theta & y'(x) \sin \theta \\ 0 & -y(x) \sin \theta & y(x) \cos \theta \end{pmatrix}$$

per cui l'area della superficie si può calcolare mediante la

$$\int_a^b y(x) \sqrt{1 + (y'(x))^2} dx$$

Il problema si riduce a quello di trovare il minimo del funzionale

$$2\pi \int_a^b y(x) \sqrt{1 + (y'(x))^2} dx$$

sotto le condizioni

$$y(a) = \alpha \quad y(b) = \beta$$

In altre parole occorre trovare

$$\min_{y \in C^1, y(a)=\alpha, y(b)=\beta} \int_a^b y(x) \sqrt{1 + (y'(x))^2} dx$$

Il solido di rotazione avente minima resistenza all'avanzamento

Determinare il solido di rotazione che offre minima resistenza all'avanzamento nell'aria.

La resistenza all'avanzamento in un gas, di densità tale da consentire l'ipotesi che le molecole si riflettano liberamente, può essere calcolata considerando la forza generata dagli urti delle molecole sulla superficie stessa.

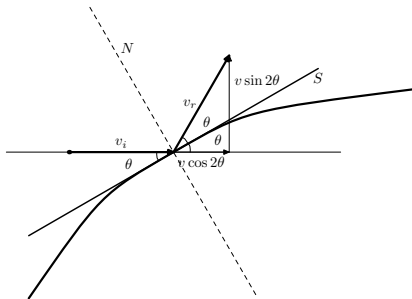
S è una porzione di superficie piana e forma un angolo θ con la direzione di avanzamento.

Su S agisce una forza F tale che $F\Delta t$ è uguale alla variazione di quantità di moto delle molecole che urtano la superficie.

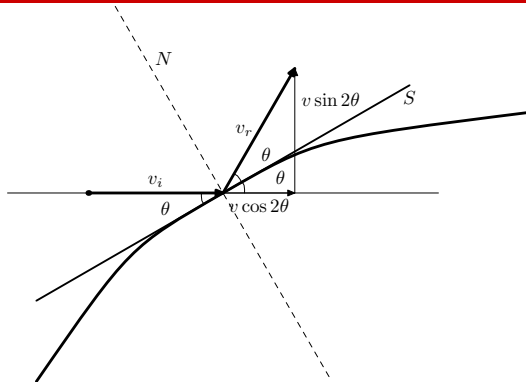
$$F\Delta t = \Delta Q$$

Una particella di massa μ che urta la superficie S con velocità

$$v_i = (v, 0, 0)$$



uguale alla velocità di avanzamento si allontana con una velocità v_r uguale in modulo a v , nel piano individuato da v e dalla normale N alla superficie e forma con essa un angolo uguale all'angolo di incidenza.



Pertanto

$$v_r = (v \cos 2\theta, v \sin 2\theta, 0)$$

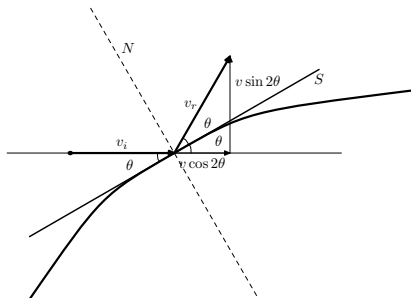
mentre la differenza tra le quantità di moto iniziale e finale è

$$\Delta Q = \mu v_i - \mu v_r = \mu(v(1 - \cos 2\theta), -v \sin 2\theta, 0)$$

Nel tempo Δt sulla superficie S agisce una forza F tale che

$$F \Delta t = \Delta Q$$

Osserviamo che F ha una componente opposta alla direzione del moto ed una componente ad essa perpendicolare. Quest'ultima non sortisce effetto in quanto il solido è di rotazione e la somma di tali componenti è pertanto nulla.

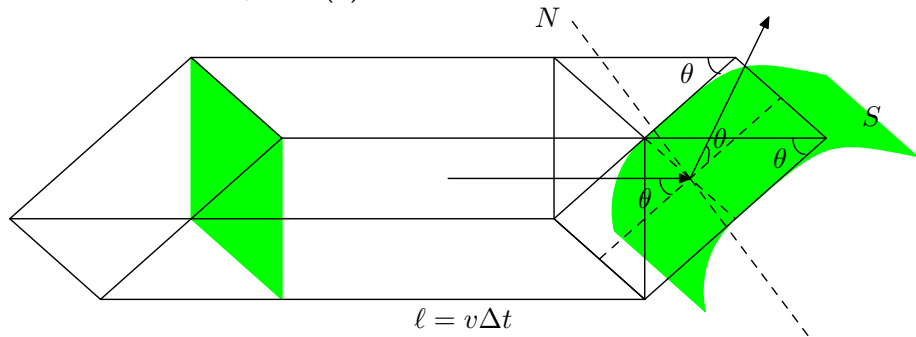


Sulla superficie S agisce quindi complessivamente una forza frenante

$$F = \frac{1}{\Delta t} m(v(1 - \cos 2\theta), 0, 0)$$

dove m è la massa totale delle particelle di gas che collidono con la superficie S nel tempo Δt .

Avremo che $m = \rho S \sin(\theta) v \Delta t$ in quanto



- nel tempo Δt collidono con la superficie S le particelle contenute nel volume di un cilindro di altezza $\ell = v\Delta t$ e di base S
- se ρ è la densità del gas e V è il volume considerato

$$m = \rho V$$

- Il fattore $\sin \theta$ rende conto dell'inclinazione di S

Pertanto su una superficie S inclinata di un angolo θ agisce una forza pari a

$$\frac{1}{\Delta t} v(1 - \cos(2\theta)) \rho S \sin(\theta) v \Delta t = 2\rho v^2 S \sin^3(\theta)$$

Supponiamo che il solido di rotazione sia definito in un riferimento cartesiano ortogonale (x, y, z) mediante la rotazione attorno all'asse x della curva

$$\begin{cases} x = x & x \in [a, b] \\ y = y(x) \end{cases}$$

con $y(0) = 0$ e $y(a) = R$.

Una parametrizzazione s di S è data da

$$s(x, t) = \begin{cases} x = x & x \in [0, a] \\ y = y(x) \cos t & t \in [0, 2\pi] \\ z = y(x) \sin t \end{cases}$$

Si ha

$$\nabla s(x, t) = \begin{pmatrix} 1 & y'(x) \cos t & y'(x) \sin t \\ 0 & -y(x) \sin t & y(x) \cos t \end{pmatrix}$$

e

$$N = (y(x)y'(x), -y(x) \cos t, -y(x) \sin t)$$

da cui

$$d\sigma = y(x) \sqrt{(1 + y'(x))^2}$$

La forza agente su un elemento ΔS di superficie sarà

$$2\rho v^2 \Delta S \sin^3(\theta)$$

per cui la resistenza si può calcolare mediante la

$$\begin{aligned} \int_S 2\rho v^2 \sin^3(\theta) d\sigma &= \\ &= \int_0^a \int_0^{2\pi} 2\rho v^2 \sin^3(\theta) y(x) \sqrt{1 + (y'(x))^2} dx dt = \\ &= 4\pi \rho v^2 \int_0^a \sin^3(\theta) y(x) \sqrt{1 + (y'(x))^2} dx \end{aligned}$$

Poichè

$$\sin \theta = \frac{y'(x)}{\sqrt{1 + (y'(x))^2}}$$

La forza resistente sarà

$$4\pi\rho v^2 \int_0^a \frac{(y'(x))^3 y(x)}{1 + (y'(x))^2} dx$$

Il problema si riduce a trovare il minimo del funzionale

$$4\pi\rho v^2 \int_0^a \frac{(y'(x))^3 y(x)}{1 + (y'(x))^2} dx$$

sulle funzioni $y \in C^1$ tali che , $y(0) = 0$ e $y(a) = R$ cioè a trovare

$$\min_{y \in C^1, y(0)=0, y(a)=R} 4\pi\rho v^2 \int_0^a \frac{(y'(x))^3 y(x)}{1 + (y'(x))^2} dx$$

Le condizioni necessarie il minimo di un funzionale

Sia

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}$$

continua con la sua derivata e $\mathcal{C}^1$ lo spazio delle funzioni

$$x : [a, b] \rightarrow \mathbb{R}$$

continue con la loro derivata.

Consideriamo il problema di trovare

$$\min_{x \in \mathcal{C}^1, x(a)=\alpha, x(b)=\beta} \int_a^b f(x(t), \dot{x}(t)) dt$$

Supponiamo che

$$x_0 \in \mathcal{C}^1 \quad , \quad x_0(a) = \alpha \quad , \quad x_0(b) = \beta$$

sia tale che

$$\int_a^b f(x_0(t), \dot{x}_0(t)) dt = \min_{x \in \mathcal{C}^1, x(a)=\alpha, x(b)=\beta} \int_a^b f(x(t), \dot{x}(t)) dt$$

Sia $h \in \mathcal{C}^1$ tale che $h(a) = h(b) = 0$; la funzione

$$\phi(\lambda) = \int_a^b f(x_0(t) + \lambda h(t), \dot{x}_0(t) + \lambda \dot{h}(t)) dt$$

ammette minimo in $\lambda = 0$ per cui dovrà aversi

$$\phi'(0) = 0$$

D'altro canto

$$\begin{aligned}
\phi'(0) &= \\
\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \int_a^b (f(x_0(t) + \lambda h(t), \dot{x}_0(t) + \lambda \dot{h}(t)) - f(x_0(t), \dot{x}_0(t))) dt &= \\
&= \int_a^b (f_x(x_0(t), \dot{x}_0(t))h(t) + f_v(x_0(t), \dot{x}_0(t))\dot{h}(t)) dt = \\
&= \int_a^b f_x(x_0(t), \dot{x}_0(t))h(t)dt + f_v(x_0(t), \dot{x}_0(t))h(t) \Big|_a^b - \\
&\quad - \int_a^b \left(\frac{d}{dt} f_v(x_0(t), \dot{x}_0(t)) \right) h(t)dt = \\
&= \int_a^b \left(f_x(x_0(t), \dot{x}_0(t)) - \frac{d}{dt} f_v(x_0(t), \dot{x}_0(t)) \right) h(t)dt = 0
\end{aligned}$$

Si può dedurre che deve essere

$$f_x(x_0(t), \dot{x}_0(t)) - \frac{d}{dt} f_v(x_0(t), \dot{x}_0(t)) = 0 \quad (46)$$

La 46 è l'equazione di Eulero o euleriana del problema
Poichè l'integranda f è autonoma, si ha

$$\begin{aligned} f_x(x(t), \dot{x}(t)) - \frac{d}{dt} f_v(x(t), \dot{x}(t)) = \\ f_x(x(t), \dot{x}(t)) - f_{v,x}(x(t), \dot{x}(t))\dot{x}(t) - \\ - f_{v,v}(x(t), \dot{x}(t))\ddot{x}(t) \end{aligned}$$

e moltiplicando per $\dot{x}(t)$

$$\begin{aligned}
 & f_x(x(t), \dot{x}(t))\dot{x}(t) - f_{v,x}(x(t), \dot{x}(t))\dot{x}^2(t) - \\
 & \quad - f_{v,v}(x(t), \dot{x}(t))\ddot{x}(t)\dot{x}(t) = \\
 & = f_x(x(t), \dot{x}(t))\dot{x}(t) + f_v(x(t), \dot{x}(t))\ddot{x}(t) - f_v(x(t), \dot{x}(t))\ddot{x}(t) - \\
 & \quad - f_{v,x}(x(t), \dot{x}(t))\dot{x}^2(t) - f_{v,v}(x(t), \dot{x}(t))\ddot{x}(t)\dot{x}(t) = \\
 & \quad = \frac{d}{dt} \left(f(x(t), \dot{x}(t)) - \dot{x}(t)f_v(x(t), \dot{x}(t)) \right) = 0
 \end{aligned}$$

Perchè la 46 sia soddisfatta è quindi necessario e sufficiente che

$$f(x(t), \dot{x}(t)) - \dot{x}(t)f_v(x(t), \dot{x}(t)) = \text{costante} \quad (47)$$

L'ultima equazione si chiama integrale primo della euleriana e spesso è più facilmente risolvibile della euleriana originale.

La curva di lunghezza minima per due punti fissati

Avremo

$$f(y, v) = \sqrt{1 + v^2}$$

e pertanto

$$f_y(y, v) = 0 \quad f_v(y, v) = \frac{v}{\sqrt{1 + v^2}}$$

e la 46 può essere scritta nella forma

$$\frac{d}{dx} \left(\frac{y'(x)}{\sqrt{1 + (y'(x))^2}} \right) = 0$$

Ne viene

$$\frac{y''(x)\sqrt{1+(y'(x))^2} - y'(x)\frac{y'(x)y''(x)}{\sqrt{1+(y'(x))^2}}}{1+(y'(x))^2} =$$

$$= y''(x)\frac{1}{(1+(y'(x))^2)\sqrt{1+(y'(x))^2}} = 0$$

e pertanto se ne ricava

$$y''(x) = 0$$

e

$$y(x) = ax + b$$

Le costanti a, b si determinano imponendo il passaggio per i punti assegnati.

La brachistocrona

Avremo

$$f(y, v) = \frac{\sqrt{1 + v^2}}{\sqrt{(h - y)}}$$

e dalle 47 possiamo ricavare che deve essere

$$\frac{\sqrt{1 + (y'(x))^2}}{\sqrt{(h - y(x))}} - y'(x) \frac{y'(x)}{\sqrt{(h - y(x))} \sqrt{1 + (y'(x))^2}} = c$$

Se ne ricava che

$$\frac{1}{\sqrt{(h - y(x))} \sqrt{1 + (y'(x))^2}} = c$$

ed elevando al quadrato,

$$y(x) = h - \frac{k}{(1 + (y'(x))^2)}$$

Abbiamo ottenuto una equazione nella forma

$$y(x) = f(y'(x)) \quad (48)$$

ove f è una funzione sufficientemente regolare.
Cerchiamo soluzioni derivabili due volte e deriviamo ambo i membri dell'equazione. Avremo

$$y'(x) = f'(y'(x))y''(x)$$

cioè

$$1 = \frac{f'(y'(x))y''(x)}{y'(x)}$$

da cui, integrando,

$$x = x_0 + \int_{x_0}^x \frac{f'(y'(t))y''(t)}{y'(t)} dt = \int_{y'(x_0)}^{y'(x)} \frac{f'(s)}{s} ds$$

Posto $y'(x) = p$ possiamo allora concludere che

$$\begin{cases} x = \int_{p_0}^p \frac{f'(s)}{s} ds + c_0 \\ y = f(p) \end{cases}$$

Alternativamente possiamo cercare soluzioni parametriche dell'equazione nella forma

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \quad (49)$$

Se poniamo

$$\phi(t) = y'(x(t))$$

avremo dalla 48

$$y(t) = y(x(t)) = f(y'(x(t))) = f(\phi(t))$$

ed anche, poichè $y(t) = y(x(t))$

$$\dot{y}(t) = y'(x(t))\dot{x}(t)$$

da cui

$$\dot{x}(t) = \frac{\dot{y}(t)}{y'(x(t))} = \frac{f'(\phi(t))}{\phi(t)}\phi'(t)$$

Ne viene che la soluzione del problema può essere individuata parametricamente dalle

$$\begin{cases} x(t) = \int_{t_0}^t \frac{f'(\phi(s))}{\phi(s)} \phi'(s) ds + c_0 \\ y(t) = f(\phi(t)) \end{cases}$$

Pertanto tenendo conto che nel nostro caso

$$f(p) = \frac{k}{1 + p^2}$$

avremo una soluzione definita dalle seguenti equazioni parametriche

$$\begin{cases} x(t) = k \int_{t_0}^t \frac{2}{(1+\phi^2(s))^2} \phi'(s) ds + c_0 \\ y(t) = h - \frac{k}{1+\phi^2(t)} \end{cases}$$

essendo c_0 e k costanti da determinarsi per soddisfare le condizioni al contorno.

Si ottiene una notevole semplificazione dei calcoli se si sceglie

$$\phi(t) = \tan t$$

in tal caso infatti

$$\begin{cases} x(t) = k \int_{t_0}^t 2 \cos^2 s ds = x_0 + k \int_{t_0}^t (1 - \cos 2s) ds + c_0 = \\ \qquad \qquad \qquad = k(t - \frac{1}{2} \sin 2t + \tau_0) + c_0 \\ y(t) = h - k \cos^2 t = h - k \frac{1 - \cos(2t)}{2} \end{cases}$$

Il solido di rotazione avente minima superficie

Avremo

$$f(y, v) = y\sqrt{1 + v^2}$$

e dalle 47 possiamo ricavare che deve essere

$$y(x)\sqrt{1 + (y'(x))^2} - y'(x)\frac{y(x)y'(x)}{\sqrt{1 + (y'(x))^2}} = c$$

da cui

$$\frac{y(x)(1 + (y'(x))^2) - (y'(x))^2 y(x)}{\sqrt{1 + (y'(x))^2}} = c$$

e

$$\frac{y(x)}{\sqrt{1 + (y'(x))^2}} = c$$

da cui

$$y(x) = c\sqrt{1 + (y'(x))^2}$$

Operando come nel caso della brachistocrona, e tenendo conto che nel presente caso

$$f(p) = c\sqrt{1 + p^2}$$

avremo una soluzione definita dalle seguenti equazioni parametriche

$$\begin{cases} x(t) = \int_{t_0}^t \frac{c\phi'(s)}{\sqrt{1+\phi^2(s)}} ds + c_0 \\ y(t) = c\sqrt{1 + \phi^2(t)} \end{cases}$$

ed è evidente che si ottiene una notevole semplificazione dei calcoli se si sceglie

$$\phi(t) = \sinh t$$

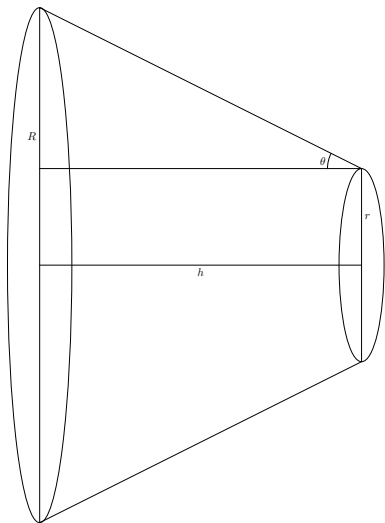
in tal caso infatti

$$\begin{cases} x(t) = \int_{t_0}^t \frac{c \cosh(s)}{\cosh(s)} ds = \int_{t_0}^t c ds = ct + c_0 \\ y(t) = c \cosh(t) \end{cases}$$

Il solido di rotazione avente minima resistenza all'avanzamento

Cominciamo ad esaminare il problema nel caso in cui il solido in esame è un tronco di cono di altezza h e raggi di base $R > r$ che avanza con la base minore in avanti.

In tal caso la resistenza opposta all'avanzamento dalla base minore è: $2\pi r^2 \rho v^2$ mentre la superficie laterale essendo generata dal segmento di retta di equazione $R - \frac{R-r}{h}x$ per $x \in [0, h]$ offre una resistenza che si può calcolare come abbiamo visto ed è, posto $\frac{R-r}{h} = \tan(\theta)$.



$$\begin{aligned}
 4\pi\rho v^2 \int_0^h \frac{\tan(\theta)^3}{1 + \tan(\theta)^2} (R - \tan(\theta)x) dx &= \\
 4\pi\rho v^2 \frac{\tan(\theta)^3}{1 + \tan(\theta)^2} \left(Rh - \tan(\theta) \frac{h^2}{2} \right) &= \\
 4\pi\rho v^2 h \tan(\theta) \sin(\theta)^2 \left(R - \tan(\theta) \frac{h}{2} \right) &= \\
 4\pi\rho v^2 (R - r) \sin(\theta) \left(R - \frac{R - r}{h} \frac{h}{2} \right) &= 2\pi\rho v^2 (R^2 - r^2) \sin^2(\theta)
 \end{aligned}$$

Pertanto la resistenza totale è

$$2\pi\rho v^2 (R^2 \sin(\theta)^2 + r^2 \cos(\theta)^2) = 2\pi\rho v^2 (R^2 + h^2 \sin(\theta)^2 - Rh \sin(2\theta))$$

che ammette minimo se

$$\theta = \frac{1}{2} \arctan\left(\frac{2}{h}\right)$$

e tenendo conto delle relazioni tra R , r e θ , si vede che se $h \rightarrow 0$
 $\theta \rightarrow \frac{\pi}{4}$.

In generale avremo, con riferimento a quanto detto in precedenza,

$$f(y, v) = \frac{v^3 y}{1 + v^2}$$

e dalle 47 possiamo ricavare che deve essere

$$\begin{aligned} \frac{(y'(x))^3 y(x)}{1 + (y'(x))^2} - y(x) \frac{(y'(x))^3 (3 + (y'(x))^2)}{(1 + (y'(x))^2)^2} &= \\ &= \frac{(y'(x))^3 y(x)}{(1 + (y'(x))^2)^2} (-2) = \textit{Costante} \end{aligned}$$

Se ne ricava che deve essere

$$\frac{(y'(x))^3 y(x)}{(1 + (y'(x))^2)^2} (-2) = C$$

ed anche

$$y(x) = c \frac{(1 + (y'(x))^2)^2}{(y'(x))^3}$$

Pertanto possiamo ottenere una soluzione in forma parametrica.
In questo caso si ha

$$f(p) = \frac{(1 + p^2)^2}{p^3}$$

e possiamo concludere che

$$\begin{cases} x(t) = c \int_{t_0}^t \left(\frac{1}{s} - \frac{2}{s^3} - \frac{3}{s^5} \right) ds + c_0 \\ y(t) = c \frac{(1+t^2)^2}{t^3} \end{cases}$$

da cui

$$\begin{cases} x(t) = c \left(\ln(t) + \frac{1}{t^2} + \frac{3}{4t^4} \right) + c_0 \\ y(t) = c \left(t + \frac{1}{t^3} + \frac{2}{t} \right) \end{cases}$$

Poichè si verifica subito che $y(t)$ non può mai assumere il valore nullo, si vede che il problema nella sua formulazione completa non ammette soluzione.

Per determinare una ragionevole approssimazione della soluzione del problema possiamo tenere conto che nel caso di un tronco di cono di altezza trascurabile la resistenza minima all'avanzamento nella direzione del suo asse si ha quando la sua apertura è di $\frac{\pi}{4}$. Possiamo allora imporre come condizione finale che la pendenza $p = y'(x) = 1$ e cioè che $t = 1$.

Le altre costanti si possono poi ricavare imponendo opportune condizioni al contorno.

Procediamo ad alcune approssimazioni.

Nel caso in cui y' sia molto grande, possiamo supporre che

$$f(y, v) = \frac{v^3 y}{1 + v^2} = vy \frac{v^2}{1 + v^2} \approx vy$$

ed in tal caso

$$\int_0^a \frac{(y'(x))^3 y(x)}{1 + (y'(x))^2} dx \approx \int_0^a y'(x) y(x) dx = \frac{(y'(a))^2}{2} = R^2$$

e quindi il funzionale è indipendente dalla scelta della funzione.

Possiamo anche procedere ad una approssimazione supponendo che y' sia piccolo per cui

$$f(y, v) = \frac{v^3 y}{1 + v^2} = y(v^3 - v^5 + v^7 + \dots) \approx v^3 y$$

In tal caso

$$\int_0^a \frac{(y'(x))^3 y(x)}{1 + (y'(x))^2} dx \approx \int_0^a (y'(x))^3 y(x) dx$$

Dalle 47 possiamo ricavare che deve essere

$$(y'(x))^3 y(x) - y'(x)(3(y'(x))^2) = -2(y'(x))^3 y(x) = C$$

Ne viene che

$$y'(x) = \frac{C}{\sqrt[3]{y(x)}}$$

la cui soluzione è data da $y(x) = (\alpha x + \beta)^{\frac{3}{4}}$

Imponendo le condizioni al contorno si ricava che $y(x) = R \left(\frac{x}{a}\right)^{\frac{3}{4}}$

Qualche risultato di Teoria del Controllo

Consideriamo una navicella, in discesa verso la superficie della luna da una quota iniziale h_0 con velocità iniziale v_0 .

La navicella

- è soggetta alla forza di gravità lunare g
- ha massa M e porta una quantità di carburante F , quindi la sua massa iniziale è $m_0 = M + F$
- può frenare utilizzando un motore che fornisce una forza frenante $u \leq \alpha$ consumando una quantità di carburante ku
- La quota della navicella all'istante t si indica con $h(t)$
- La velocità della navicella all'istante t si indica con $v(t)$
- La massa della navicella all'istante t si indica con $m(t)$

Il problema consiste nel far giungere la navicella a quota 0 con velocità nulla utilizzando i motori in modo da minimizzare il consumo di carburante.

Il moto del modulo sarà governato dal sistema

$$\begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g + \frac{u(t)}{m(t)} \\ \dot{m}(t) = -ku(t) \end{cases} \quad \begin{cases} h(0) = h_0 \\ v(0) = v_0 \\ m(0) = m_0 = M + F \end{cases}$$

e sarà nostro scopo trovare $T, (h, v, m)$ ed u in modo da

$$\text{Minimizzare} \quad \int_0^T u(t) dt$$

sotto i vincoli

$$\begin{cases} h(T) = 0 \\ v(T) = 0 \end{cases} \quad 0 \leq u(t) \leq \alpha$$

Per risolvere il problema occorre il principio del massimo di Pontryagin

Il problema di controllo

Sia

$$f_0 : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}$$

sufficientemente regolare (almeno di classe $\mathcal{C}^1$) e sia

$$F : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^n$$

una funzione vettoriale sufficientemente regolare.

Siano

$$x : [0, T] \rightarrow \mathbb{R}^n$$

di classe $\mathcal{C}^1$ a tratti e

$$u : [0, T] \rightarrow \mathbb{R}^k$$

di classe $\mathcal{C}^0$ a tratti.

Sia $U \subset \mathbb{R}^k$ un insieme chiuso

Consideriamo il problema di
Determinare il minimo del funzionale

$$\int_0^T f_0(x(t), u(t)) dt$$

al variare di T , tempo finale, di x ($\mathcal{C}^1$ a tratti) , e di u ($\mathcal{C}^0$ a tratti), tali che:

$$\begin{aligned}\dot{x}(t) &= F(x(t), u(t)) \\ x(0) &= x_0 \quad , \quad x(T) = x_1 \\ u(t) &\in U\end{aligned}$$

Il problema enunciato è un tipico problema di controllo ottimo.

x è la funzione di stato

u è la funzione di controllo

U è l'insieme dei possibili controlli.

Pontryagin ha dato condizioni necessarie per la soluzione del problema di controllo.

Tali risultati sono noti come

Principio del massimo di Pontryagin

Principio del massimo di Pontryagin

Se (x^*, u^*) è una soluzione del problema di controllo allora esistono

$\lambda_0 \in \mathbb{R}$, $\lambda : \mathbb{R} \rightarrow \mathbb{R}^n$ tali che, posto

$$H(x, u) = \lambda_0 f_0(x, u) + \langle \lambda, F(x, u) \rangle = \lambda_0 f_0(x, u) + \sum_{i=1}^n \lambda_i F_i(x, u)$$

si abbia

$$\dot{x}^*(t) = F(x^*(t), u^*(t)) \quad x^*(0) = x_0 \quad x^*(T) = x_1$$

$$\dot{\lambda}(t) = -\nabla_x H(x^*(t), u^*(t))$$

$$H(x^*(t), u^*(t)) = \max_{|u| \leq \alpha} H(x^*(t), u)$$

$$\lambda_0 = -1 \quad H(x^*(t), u^*(t)) = 0$$

Una Prova euristica del Principio del Massimo

Ci limitiamo al caso in cui si minimizza il funzionale

$$I(u) = \int_0^1 f_0(x(t), u(t)) dt$$

sotto le condizioni

$$\begin{cases} \dot{x}(t) = F(x(t), u(t)) \\ x(0) = x_0 \\ x(1) = x_1 \end{cases} \quad u \in U$$

essendo F una funzione a valori reali

($\dot{x}(t) = F(x(t), u(t))$ è un'equazione differenziale.)

Consideriamo un incremento per il controllo u^* in corrispondenza del quale si ottiene il minimo di I

$$u_\lambda = u^* + \lambda w$$

ed il corrispondente stato

$$\begin{cases} \dot{x}_\lambda(t) = F(x_\lambda(t), u_\lambda(t)) \\ x_\lambda(0) = x_0 \\ x_\lambda(1) = x_1 \end{cases}$$

Avremo che

$$\begin{aligned} 0 = \frac{d}{d\lambda} I(u_\lambda) &= \int_0^1 \left((f_0)_x \frac{d}{d\lambda} x_\lambda + (f_0)_u \frac{d}{d\lambda} u_\lambda \right) dt = \\ &= \int_0^1 ((f_0)_x \delta_\lambda + (f_0)_u w) dt \end{aligned}$$

dove

$$\delta_\lambda = \frac{d}{d\lambda} x_\lambda$$

Poichè

$$x_\lambda = x_0 + \int_0^t F(x_\lambda, u_\lambda) dt$$

si ha

$$\frac{d}{d\lambda} x_\lambda = \int_0^t \frac{d}{d\lambda} F(x_\lambda, u_\lambda) dt = \int_0^t \left(F_x \frac{dx_\lambda}{d\lambda} + F_u w \right) dt$$

Pertanto

$$\begin{cases} \dot{\delta}_\lambda = F_x \delta_\lambda + F_u w \\ \delta_\lambda(0) = 0 \\ \delta_\lambda(1) = 0 \end{cases}$$

$$(x(0) = x_\lambda(0) = x_0, x(1) = x_\lambda(1) = x_1)$$

Quindi

$$w = \frac{\dot{\delta}_\lambda - F_x \delta_\lambda}{F_u}$$

e

$$\begin{aligned}
0 &= \int_0^1 ((f_0)_x \delta_\lambda + (f_0)_u w) dt = \\
&= \int_0^1 \left((f_0)_x \delta_\lambda + (f_0)_u \left(\frac{\dot{\delta}_\lambda}{F_u} - \frac{F_x \delta_\lambda}{F_u} \right) \right) dt = \\
&= \int_0^1 \left(\left((f_0)_x - (f_0)_u \frac{F_x}{F_u} \right) \delta_\lambda + (f_0)_u \frac{\dot{\delta}_\lambda}{F_u} \right) dt = \\
&= \int_0^1 \left(\left((f_0)_x - (f_0)_u \frac{F_x}{F_u} \right) \delta_\lambda \right) dt + (f_0)_u \frac{\delta_\lambda}{F_u} \Big|_0^1 - \int_0^1 \left(\frac{d}{dt} \frac{(f_0)_u}{F_u} \delta_\lambda \right) dt = \\
&= \int_0^1 \left[\left((f_0)_x - (f_0)_u \frac{F_x}{F_u} - \frac{d}{dt} \frac{(f_0)_u}{F_u} \right) \right] \delta_\lambda dt = 0
\end{aligned}$$

Sfruttando l'arbitrarietà dell'incremento

$$\frac{d}{dt} \frac{(f_0)_u}{F_u} = (f_0)_x - (f_0)_u \frac{F_x}{F_u}$$

e, posto $\eta = \frac{(f_0)_u}{F_u}$ si ha

$$\begin{cases} \dot{\eta} = -(f_0)_x + \eta F_x \\ \eta F_u - (f_0)_u = 0 \end{cases}$$

Se infine poniamo

$$H(x, u) = -f_0(x, u) + \eta F(x, u)$$

abbiamo che

$$\begin{cases} \dot{\eta} = \nabla_x H(x, u) \\ \nabla_u H(x, u) = 0 \end{cases}$$

La seconda condizione afferma che u è un punto stazionario (gradiente nullo) in realtà si può dimostrare che u minimizza $H(x, \cdot)$

Qualche esempio di Problema di Controllo

Il problema del minimo tempo

Possiamo enunciare il problema come segue

Consideriamo una particella, che si muove su una retta, $x(t)$ è la distanza della particella da un punto fisso O , della retta, detto origine

Il moto della particella è governato dall'equazione differenziale

$$\ddot{x}(t) = u(t) \quad |u(t)| \leq 1 \quad (50)$$

dove $u(t)$ rappresenta una forza applicata alla particella.

Vogliamo trovare $u(t)$ in modo che la particella che parte con velocità e posizione iniziale nota, raggiunga l'origine con velocità nulla nel minimo tempo possibile.

Possiamo porre il problema in termini di controllo come segue

$$\text{Minimizzare } \int_0^T dt \quad (51)$$

Sugli archi che soddisfano l'equazione di controllo

$$\begin{cases} \dot{x}(t) = y(t) \\ \dot{y}(t) = u(t) \end{cases} \quad \begin{cases} x(0) = x_0 \\ y(0) = v_0 \end{cases}$$

ed i vincoli

$$\begin{cases} x(T) = 0 \\ y(T) = 0 \end{cases} \quad (52)$$

Con riferimento all'enunciato del principio del massimo avremo che

$$f_0((x, y), u) = 1 \quad F((x, y), u) = \begin{pmatrix} y \\ u \end{pmatrix}$$

per cui

$$H(x, u) = \lambda_0 + \lambda y + \mu u \quad (53)$$

Per il principio del massimo dovrà essere

$$\begin{pmatrix} \dot{\lambda} \\ \dot{\mu} \end{pmatrix} = - \begin{pmatrix} 0 \\ \lambda \end{pmatrix} = \begin{pmatrix} H_x \\ H_y \end{pmatrix}$$

da cui si ricava che

$$\begin{cases} \lambda(t) = h \\ \mu(t) = -ht + k \end{cases}$$

Poichè H è lineare in u , assumerà il suo massimo per $u = \pm 1$ e il segno di u deve essere scelto in dipendenza del segno di μ .

Più precisamente dovrà essere

- $u = 1$ nel caso in cui $\mu > 0$
- $u = -1$ nel caso in cui $\mu < 0$

Dal momento che

$$\mu(t) = -ht + k$$

si annulla una ed una sola volta, in $\bar{t}$, anche u potrà avere un solo cambiamento di segno.

Pertanto assumerà valore costante

$$\pm 1 \quad \text{per} \quad t \in [0, \bar{t}]$$

e

$$\mp 1 \quad \text{per} \quad t \in [\bar{t}, T]$$

Quando $u = 1$ avremo che

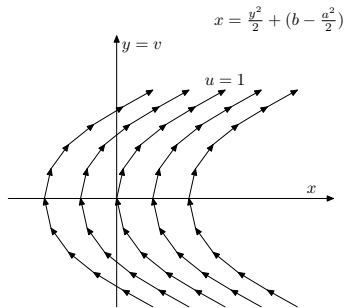
$$\begin{cases} \dot{x}(t) = y(t) \\ \dot{y}(t) = 1 \end{cases}$$

e

$$\begin{cases} x(t) = t^2/2 + at + b \\ y(t) = t + a \end{cases}$$

$$\begin{cases} x(t) = \frac{(t+a)^2}{2} + b - \frac{a^2}{2} \\ y(t) = t + a \end{cases}$$

Le traiettorie per $u = 1$ sono come in figura



Per $u = -1$ avremo che

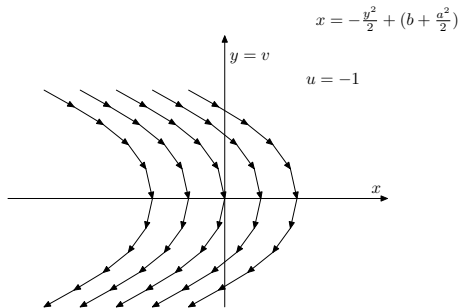
$$\begin{cases} \dot{x}(t) = y(t) \\ \dot{y}(t) = -1 \end{cases}$$

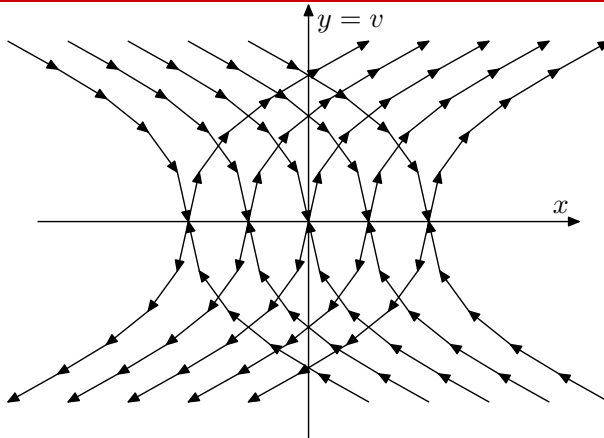
e

$$\begin{cases} x(t) = -t^2/2 + at + b \\ y(t) = -t + a = -(t - a) \end{cases}$$

$$\begin{cases} x(t) = -\frac{(t-a)^2}{2} + b + \frac{a^2}{2} \\ y(t) = -t + a = -(t - a) \end{cases}$$

Le traiettorie per $u = -1$ sono come in figura



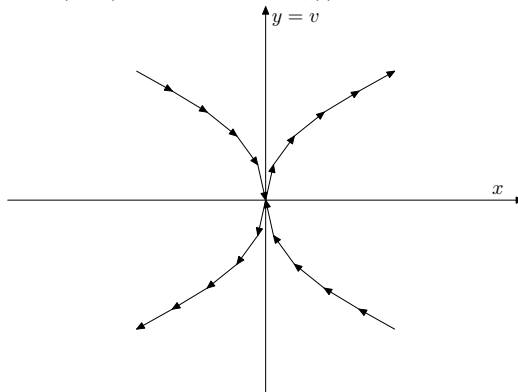


In ogni caso si avrà

$$x = \pm y^2 + (b \mp \frac{a^2}{2})$$

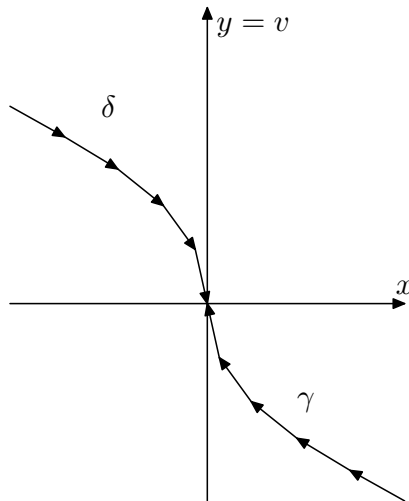
La figura mostra le traiettorie del sistema per $u = \pm 1$ nel piano (x, y) . In tale riferimento $x(t)$ individua posizione e $y(t)$ la velocità della particella.

$$\begin{cases} x(t) = -t^2/2 + at + b = \frac{-(t-a)^2}{2} + b + \frac{a^2}{2} \\ y(t) = -t + a = -(t-a) \end{cases} \quad \begin{cases} x(t) = t^2/2 + at + b = \frac{(t+a)^2}{2} + b - \frac{a^2}{2} \\ y(t) = t + a \end{cases}$$



Per raggiungere l'origine, occorrerà allora percorrere una delle due traiettorie che passano per $(0, 0)$ avendo cura di scegliere il verso diretto all'origine e non quello dall'origine.

La traiettoria γ che passa per l'origine e corrisponde a y positive è percorsa per $u = 1$ concordemente al verso positivo di y (su di essa $y = t + a$), la traiettoria δ che passa per l'origine e corrisponde a y negative è percorsa per $u = -1$ in senso opposto al verso positivo di y (su di essa $y = -t + a$). Quindi per raggiungere l'origine abbiamo a disposizione soltanto i due archi di traiettoria che abbiamo descritto e che sono illustrati nella figura seguente



Se $(x_0, v_0) = (x_0, y_0)$ giace su γ o su δ possiamo allora trovare una soluzione del problema, semplicemente tenendo fisso $u = \pm 1$.

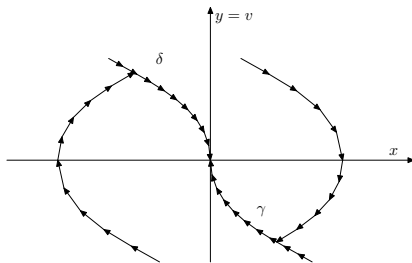
Altrimenti, partendo dal punto iniziale assegnato e percorrendo una delle altre traiettorie,

$$x = \pm y^2 + (b \mp \frac{a^2}{2})$$

dovremo raggiungere γ o δ e poi l'origine, imponendo che

$$x_0 = \pm y_0^2 + (b \mp \frac{a^2}{2})$$

Le traiettorie ottimali sono riportate nella figura.



Il Problema del modulo lunare

Consideriamo una navicella, in discesa verso la superficie della luna da una quota iniziale h_0 con velocità iniziale v_0 .

La navicella

- è soggetta alla forza di gravità lunare g
- ha massa M e porta una quantità di carburante F , quindi la sua massa iniziale è $m_0 = M + F$
- può frenare utilizzando un motore che fornisce una forza frenante $u \leq \alpha$ consumando una quantità di carburante ku

- La quota della navicella all'istante t si indica con $h(t)$
- La velocità della navicella all'istante t si indica con $v(t)$
- La massa della navicella all'istante t si indica con $m(t)$

La navicella è soggetta all'accelerazione di gravità g ed è frenata dai motori con forza u

I motori consumano carburante in quantità ku e la massa diminuisce del carburante consumato.

Quindi evidentemente

- $h(0) = h_0$
- $v(0) = v_0$
- $m(0) = m_0 = M + F$

$$m(t)\ddot{h}(t) = -m(t)g + u(t)$$

$$\dot{m}(t) = -ku(t)$$

$$0 \leq u(t) \leq \alpha$$

Il consumo di carburante nell'intervallo di tempo $[0, t]$ è

$$k \int_0^t u(s) ds$$

Il problema consiste nel far giungere la navicella a quota 0 con velocità nulla utilizzando i motori in modo da minimizzare il consumo di carburante.

Il moto del modulo sarà governato dal sistema

$$\begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g + \frac{u(t)}{m(t)} \\ \dot{m}(t) = -ku(t) \end{cases} \quad \begin{cases} h(0) = h_0 \\ v(0) = v_0 \\ m(0) = m_0 = M + F \end{cases} \quad (54)$$

e sarà nostro scopo trovare $T, (h, v, m)$ ed u in modo da

$$\text{Minimizzare} \quad \int_0^T u(t) dt \quad (55)$$

sotto i vincoli

$$\begin{cases} h(T) = 0 \\ v(T) = 0 \end{cases} \quad 0 \leq u(t) \leq \alpha$$

Con riferimento all'enunciato del principio del massimo avremo che

$$f_0((h, v, m), u) = u \quad F((h, v, m), u) = \begin{pmatrix} v \\ -g + \frac{u}{m} \\ -ku \end{pmatrix}$$

per cui

$$\begin{aligned} H((h, v, m), u) &= \lambda_0 u + \lambda v + \eta \left(-g + \frac{u}{m}\right) + \mu(-ku) = \\ &= \left(\lambda_0 + \frac{\eta}{m} - k\mu\right)u + \lambda v - \eta g \quad (56) \end{aligned}$$

e

$$H_h = 0 \quad H_v = \lambda \quad H_m = -\frac{\eta u}{m^2}$$

Per il principio del massimo dovrà essere

$$\begin{cases} \dot{\lambda}(t) = 0 \\ \dot{\eta}(t) = -\lambda(t) \\ \dot{\mu}(t) = \frac{\eta(t)u(t)}{m^2(t)} \end{cases}$$

Inoltre u deve essere scelto in modo da realizzare

$$\max_{u \in [0, \alpha]} H((h, v, m), u) = \max_{u \in [0, \alpha]} = (\lambda_0 + \frac{\eta}{m} - k\mu)u + \lambda v - \eta g$$

Poichè H è lineare in u , il suo massimo è assunto per $u = 0$ oppure $u = \alpha$; a seconda del segno di

$$\lambda_0 + \frac{\eta}{m} - k\mu$$

Più precisamente si avrà

- $u = 0$ nel caso in cui $\lambda_0 + \frac{\eta}{m} - k\mu < 0$
- $u = \alpha$ nel caso in cui $\lambda_0 + \frac{\eta}{m} - k\mu > 0$.

Per $u = 0$ la traiettoria è di caduta libera.

Se si applica il controllo $u = 0$ scegliendo come tempo iniziale $t = 0$

alla quota

$$\bar{h}$$

con velocità iniziale

$$\bar{v}$$

e massa iniziale

$$\bar{m}$$

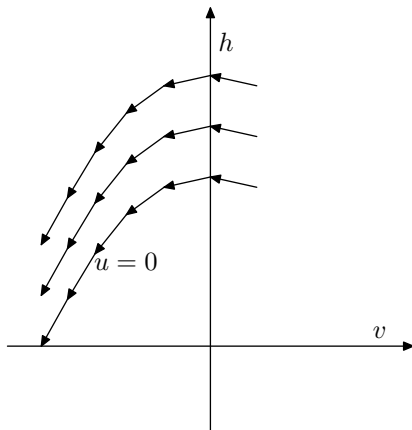
dovrà essere

$$\begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g \\ \dot{m}(t) = 0 \end{cases} \quad \text{da cui} \quad \begin{cases} h(t) = -\frac{1}{2}gt^2 + \bar{v}t + \bar{h} \\ v(t) = \bar{v} - gt \\ m(t) = \bar{m} \end{cases}$$

Sostituendo $t = \frac{\bar{v}-v}{\frac{g}{\bar{v}}}$ in
 $h = -\frac{1}{2}gt^2 + \bar{v}t + \bar{h}$ otteniamo
 la traiettoria nel piano (h, v) .

$$h = -\frac{1}{2g}v^2 + \left(\bar{h} + \frac{1}{2g}\bar{v}^2\right)$$

($v(t)$ è decrescente e
 negativa $\bar{h} > 0$, $\bar{v} < 0$ per cui
 la traiettoria giace nel
 quadrante con $h > 0$ e $v < 0$)

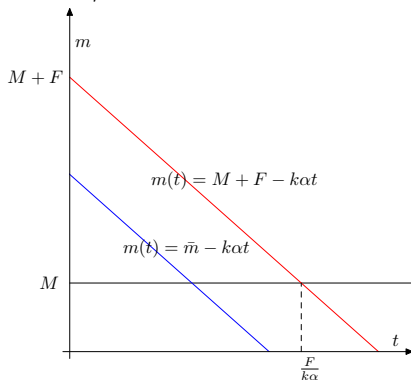


Se il controllo $u = \alpha$ è applicato a partire da $t = 0$ alla quota $\bar{h}$ con velocità iniziale $\bar{v}$ e massa iniziale $\bar{m}$, da

$$\begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g + \frac{\alpha}{m(t)} \\ \dot{m}(t) = -k\alpha \end{cases}$$

segue che

$$m(t) = \bar{m} - k\alpha t$$



La massa iniziale $\bar{m} \leq M + F$ si può ridurre al più ad M , per cui

$$M + F - k\alpha t \geq \bar{m} - k\alpha t \geq M \quad \Rightarrow \quad 0 \leq t \leq \frac{F}{k\alpha}$$

Inoltre

$$v(t) = \bar{v} + \int_0^t \left(-g + \frac{\alpha}{\bar{m} - k\alpha s} \right) ds$$

Affinchè il modulo possa frenare occorre che

$$\dot{v}(0) = -g + \frac{\alpha}{\bar{m}} \geq -g + \frac{\alpha}{M+F} \geq 0$$

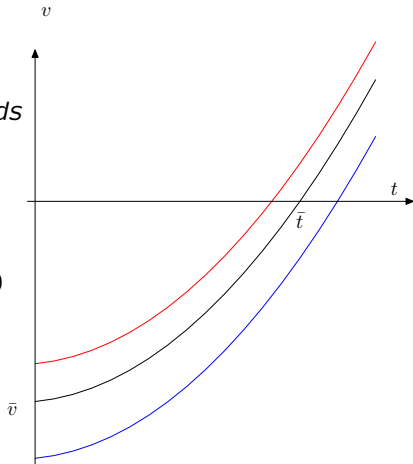
per cui

$$\frac{\alpha}{M+F} \geq g$$

Poichè $\bar{m} - k\alpha s$ è decrescente,

$$\dot{v}(s) = -g + \frac{\alpha}{\bar{m} - k\alpha s}$$

è crescente e , da $\dot{v}(0) > 0$ segue che $v(t)$ è crescente, convessa.



Infine

$$h(t) = \bar{h} + \int_0^t v(s) ds$$

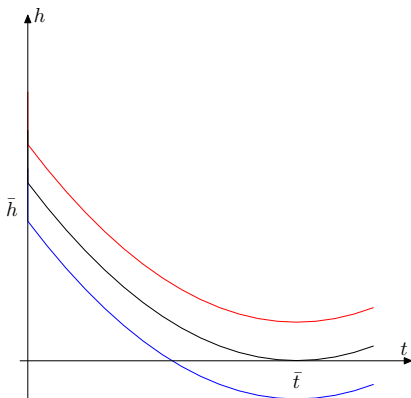
quindi sapendo che

$$v(t)$$

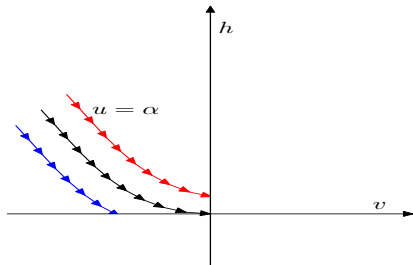
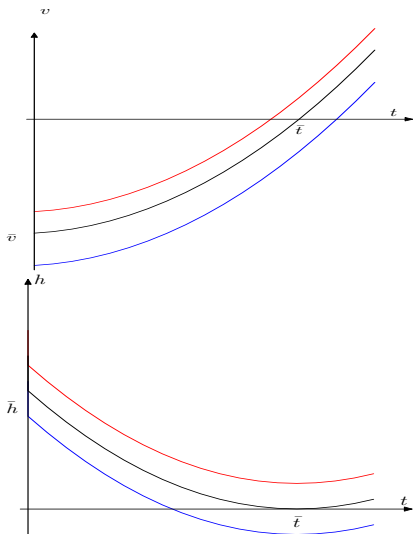
è negativa e crescente,

$$h(t)$$

è decrescente e convessa.



Il grafico della traiettoria nel piano (h, v) è simile ad un ramo di parabola convessa.



Si possono scegliere i dati iniziali in modo che la traiettoria corrispondente ad $u = \alpha$ passi, per qualche valore $t = T$, per il punto $(h, v) = (0, 0)$

Mentre non ci sono traiettorie che permettano di raggiungere il suolo con velocità nulla senza frenare ($u = 0$),

Pertanto per raggiungere l'origine abbiamo a disposizione soltanto la traiettoria γ corrispondente a $u = \alpha$ che passa per l'origine nel piano (h, v) .

Se le condizioni di partenza (h_0, v_0) giacciono sull'arco di curva γ la soluzione del problema, consiste nell'applicare il controllo $u = \alpha$, cioè frenare al massimo per l'intera discesa.

In caso contrario sarà necessario, partendo dal punto iniziale assegnato, raggiungere un punto dell'arco γ e da lì proseguire sull'arco γ .

Per raggiungere un punto di γ possiamo alternare controlli $u = 0$ ed $u = \alpha$,

Ma occorre ricordare che affinché il controllo sia ottimo la scelta dipende dal segno di

$$\lambda_0 + \frac{\eta}{m} - k\mu$$

Si ha

$$\left(\lambda_0 + \frac{\eta}{m} - k\mu \right)' = \frac{\dot{\eta}m - \eta\dot{m}}{m^2} - k\dot{\mu} = \frac{\dot{\eta}}{m} + \frac{\eta ku}{m^2} - \frac{k\eta u}{m^2} = \frac{\dot{\eta}}{m}$$

Ma

$$\dot{\eta} = -\lambda, \quad \dot{\lambda} = 0 \quad \Rightarrow \quad \dot{\eta} = \text{costante}$$

inoltre

$$m(t) = \bar{m} - k\alpha t > 0$$

quindi $\frac{\dot{\eta}}{m}$ ha segno costante, $(\lambda_0 + \frac{\eta}{m} - k\mu)$ è monotona e cambia segno al più una sola volta.

Si può concludere che il controllo ottimo si realizza utilizzando

- $u = 0$ per $t \in [0, \bar{t}]$
- $u = \alpha$ per $t \in [\bar{t}, T]$

con $T \leq \frac{F}{k\alpha}$

Una ulteriore condizione necessaria afferma che deve essere

$$H((h, v, m), u) = 0$$

per la scelta di u fissata.

Poichè u è costante a tratti e $\lambda = \bar{\lambda}$ è a sua volta costante si ha

$$\begin{aligned}\frac{dH}{dt} &= \left(\lambda_0 + \frac{\eta}{m} - k\mu \right)' u + \bar{\lambda} \dot{v} - \dot{\eta} g = \\ &= \left(\frac{\dot{\eta} m - \eta \dot{m}}{m^2} - k\dot{\mu} \right) u + \bar{\lambda} \left(-g + \frac{u}{m} \right) + \bar{\lambda} g = \\ &= \left(\frac{\dot{\eta}}{m} + \frac{\eta k u}{m^2} - \frac{k \eta u}{m^2} \right) + \frac{\bar{\lambda} u}{m} = \frac{\dot{\eta} u}{m} + \frac{\lambda u}{m} = 0\end{aligned}$$

quindi

$$H((h, v, m), u) \quad \text{è costante}$$

Si può infine verificare che si possono determinare λ , η , μ in modo che

$$H((h, v, m), u) = 0$$

Ricordiamo che

$$H((h, v, m), u) = \left(\lambda_0 + \frac{\eta}{m} - k\mu \right) u + \lambda v - \eta g$$

$$\begin{cases} \dot{\lambda} = 0 \\ \dot{\eta} = -\lambda \\ \dot{\mu} = \frac{\eta u}{m^2} \end{cases} \quad \begin{cases} \dot{h} = v \\ \dot{v} = -g + \frac{u}{m} \\ \dot{m} = -ku \end{cases} \quad u(t) = \begin{cases} 0 & t \in [0, \bar{t}] \\ \alpha & t \in [\bar{t}, T] \end{cases} \quad T \leq \frac{F}{k\alpha}$$

$$h(0) = h_0, \quad v(0) = v_0, \quad m(0) = m_0 = M + F, \quad v(T) = 0, \quad h(T) = 0$$

Dal momento che

$$H((h, v, m), u)$$

è costante basta verificare che è nulla in $t = 0$.

Per $t \in [0, \bar{t}]$

$$\begin{cases} \lambda = \bar{\lambda} \\ \eta = \bar{\eta} - \bar{\lambda}t \\ \mu = \bar{\mu} \end{cases}$$

$$\begin{aligned} H((h, v, m), u) &= \bar{\lambda}(v_0 - g\bar{t}) - (\bar{\eta} - \bar{\lambda}\bar{t}) = \\ &= \bar{\lambda}v_0 - \bar{\lambda}\bar{t}g - \bar{\eta} + \bar{\lambda}\bar{t}g = \bar{\lambda}v_0 - \bar{\eta}g = 0 \end{aligned}$$

per cui basta scegliere

$$\bar{\eta} = \frac{\bar{\lambda}v_0}{g}$$

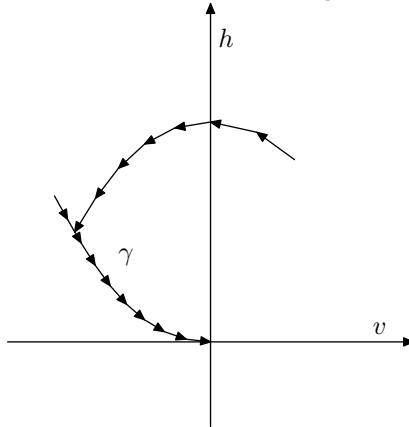
Per $t \in [\bar{t}, t_0]$

$$\begin{cases} \lambda = \bar{\lambda} \\ \eta = \bar{\lambda} \left(\frac{v_0}{g} t \right) \\ \mu = \bar{\mu} \end{cases}$$

$$\begin{aligned} H((h, v, m), u) &= \left(\lambda_0 + \bar{\lambda} \left(\frac{v_0}{g} \right) - k\bar{\mu} \right) \alpha + \bar{\lambda}(v_0 - gt) - \bar{\lambda}(v_0 - gt) = \\ &= \left(\lambda_0 + \bar{\lambda} \left(\frac{v_0}{g} \right) - k\bar{\mu} \right) \alpha = 0 \end{aligned}$$

per cui basta scegliere $\bar{\mu}$ di conseguenza

La traiettoria ottimale è riportata nella figura seguente



Le equazioni di discesa del Modulo Lunare

Il Modulo per allunare deve scendere in caduta libera fino ad un certo istante $\bar{t}$ e di lì innanzi frenare con la massima spinta possibile.

Esaminiamo la traiettoria che deve seguire

Troviamo innanzi tutto le traiettorie di caduta libera e di caduta frenata.

La traiettoria di caduta libera

Si ha per $u = 0$

Se

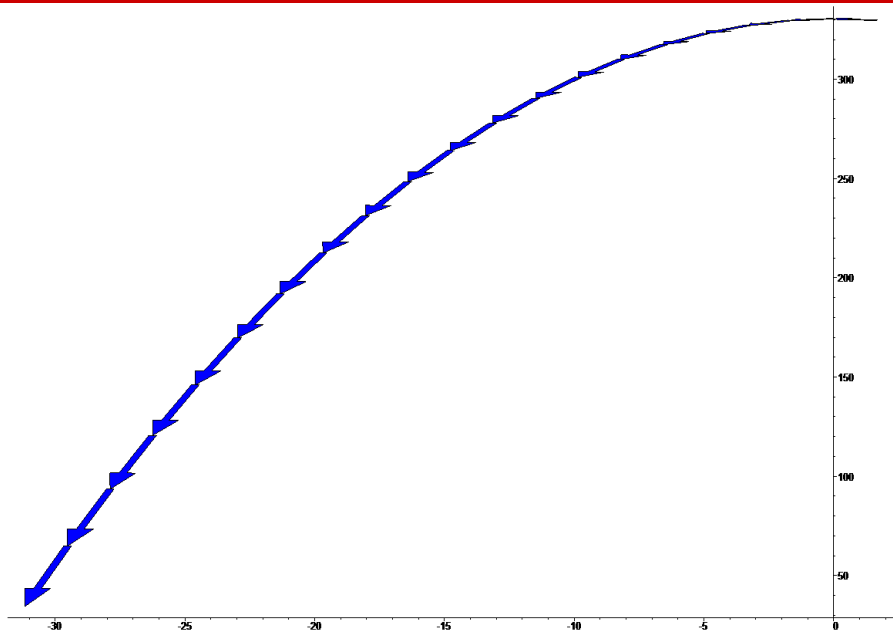
$$t = 0 \quad , \quad h(0) = h_0 \quad , \quad v(0) = v_0 \quad , \quad m(0) = m_0$$

avremo

$$\begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g \\ \dot{m}(t) = 0 \end{cases} \quad \text{da cui} \quad \begin{cases} h(t) = -\frac{1}{2}gt^2 + v_0t + h_0 \\ v(t) = v_0 - gt \\ m(t) = m_0 \end{cases}$$

e

$$h = -\frac{1}{2}gv^2 + \left(\bar{h} + \frac{1}{2}g\bar{v}^2 \right)$$



la traiettoria di caduta libera

La traiettoria di caduta frenata

Si ha per $u = \alpha$ e deve risultare

$$\begin{cases} \dot{h}(t) = v(t) \\ \dot{v}(t) = -g + \frac{\alpha}{m(t)} \\ \dot{m}(t) = -k\alpha \end{cases}$$

Sia $t = 0$ l'istante in cui il modulo tocca la superficie lunare. L'istante iniziale quindi sarà $\bar{t} < 0$.

Le condizioni finali saranno

$$\begin{cases} h(0) = 0 \\ v(0) = 0 \\ m(0) = M \end{cases}$$

ed integreremo per $t < 0$.

Avremo che

$$m(t) = M - k\alpha t \quad t < 0$$

mentre

$$v(0) - v(t) = \int_t^0 -g + \frac{\alpha}{M - k\alpha s} ds$$

da cui

$$\begin{aligned} v(t) &= \int_0^t -g + \frac{\alpha}{M - k\alpha s} ds = \\ &= -gt - \frac{\alpha}{k\alpha} \ln(M - k\alpha s) \Big|_0^t = \\ &= -gt - \frac{1}{k} \ln \left(\frac{M - k\alpha t}{M} \right) \end{aligned}$$

$$h(t) - h(0) = \int_0^t \left(-gt - \frac{1}{k} \ln \left(\frac{M - k\alpha s}{M} \right) \right) ds$$

$$\begin{aligned} h(t) &= -\frac{1}{2}gt^2 - \frac{1}{k} \int_0^t \ln \left(\frac{M - k\alpha s}{M} \right) ds \\ &= -\frac{1}{2}gt^2 - \frac{1}{k} \left[s \ln \left(\frac{M - k\alpha s}{M} \right) \Big|_0^t \right] - \int_0^t s \frac{M}{M - k\alpha s} \frac{-k\alpha}{M} ds \\ &= -\frac{1}{2}gt^2 - \frac{1}{k} t \ln \left(\frac{M - k\alpha t}{M} \right) - \frac{1}{k} \frac{k\alpha}{M} \int_0^t \frac{sM}{M - k\alpha s} ds \\ &= -\frac{1}{2}gt^2 - \frac{1}{k} t \ln \left(\frac{M - k\alpha t}{M} \right) + \frac{1}{k} \int_0^t \frac{-sk\alpha}{M - k\alpha s} ds \end{aligned}$$

Poichè

$$\frac{-sk\alpha}{M - k\alpha s} = \frac{M - sk\alpha}{M - k\alpha s} - \frac{M}{M - k\alpha s} = 1 - \frac{M}{M - k\alpha s}$$

si ha

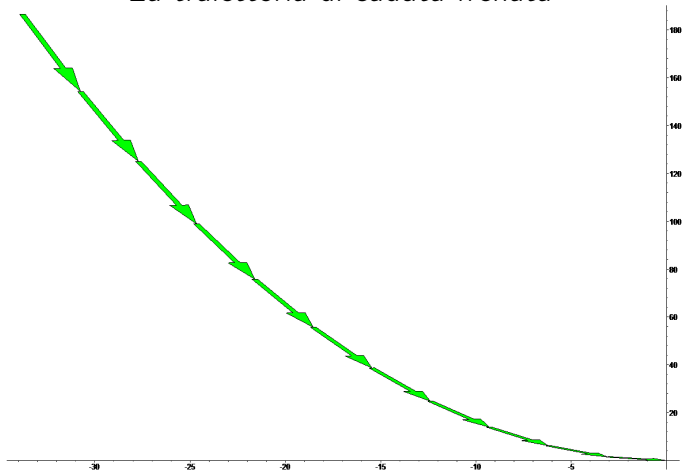
$$\begin{aligned} \int_0^t \frac{-sk\alpha}{M - k\alpha s} ds &= \int_0^t \left(1 - \frac{M}{M - k\alpha s} \right) ds \\ &= \left[s + \frac{M}{k\alpha} \ln \left(\frac{M - k\alpha s}{M} \right) \right]_0^t \end{aligned}$$

da cui

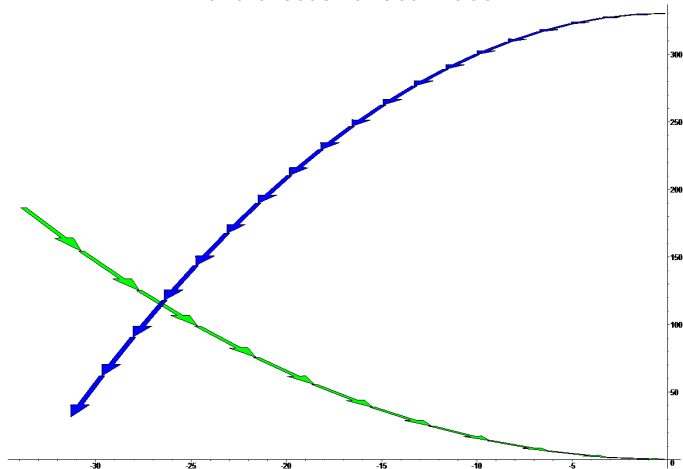
$$h(t) = -\frac{1}{2}gt^2 - \frac{t}{k} \ln \left(\frac{M - k\alpha t}{M} \right) + \frac{t}{k} + \frac{M}{k^2\alpha} \ln \left(\frac{M - k\alpha t}{M} \right)$$

$$h(t) = -\frac{1}{2}gt^2 + \frac{t}{k} + \frac{M - k\alpha t}{k^2\alpha} \ln \left(\frac{M - k\alpha t}{M} \right)$$

La traiettoria di caduta frenata



La traiettoria ottimale



Misura ed integrazione.

Il problema della misura di un insieme è di primaria importanza in innumerevoli problemi quali, ad esempio il calcolo di lunghezze, aree e volumi, la definizione di integrale, la teoria delle probabilità.

Sarebbe quindi molto comodo disporre di una definizione di misura soddisfacente requisiti intuitivamente indispensabili, in grado di attribuire un valore ad ogni insieme di $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$ ed eventualmente estendibile a spazi e situazioni più generali.

Banach e Tarski hanno però dimostrato che in $\mathbb{R}^3$ non c'è alcun modo di definire il volume senza che almeno una delle seguenti affermazioni sia vera:

- ① Il volume non è invariante per rotazione,
- ② il volume non è finitamente additivo,
- ③ Esistono insiemi di cui non si può definire il volume,
- ④ Non vale l'assioma della scelta.

Normalmente la scelta cade sull'accettare la terza scelta e dare

Procediamo alla costruzione della misura di Lebesgue in $\mathbb{R}^2$. Possiamo con argomenti analoghi a quelli mostrati anche costruire una misura in $\mathbb{R}$ (con qualche semplificazione) ed in $\mathbb{R}^3$ (con qualche complicazione).

Il procedimento usato può inoltre essere generalizzato a casi completamente astratti.

Consideriamo in $\mathbb{R}$ l'insieme $\mathcal{S}$ dei rettangoli.

Chiamiamo rettangolo il prodotto cartesiano R di due intervalli reali I e J , chiusi aperti, semichiusi o semiaperti. In altre parole diciamo che R è un rettangolo se

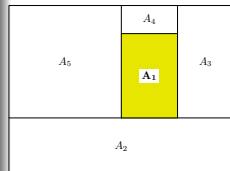
$$R = I \times J$$

con I, J intervalli reali.

Si verifica subito che $\mathcal{S}$ è una famiglia di insiemi soddisfacente le seguenti proprietà:

- $\emptyset \in \mathcal{S}$
- $A, B \in \mathcal{S} \implies A \cap B \in \mathcal{S}$
- se $A_1, A \in \mathcal{S}$, $A_1 \subset A$ allora esistono $\{A_k \in \mathcal{S} : k = 2, \dots, n\}$, a due a due disgiunti, tali che

$$A = \bigcup_{k=1}^n A_k$$



In generale

Definizione

Una famiglia $\mathcal{S}$ di insiemi si dice semianello se

- $\emptyset \in \mathcal{S}$
- $A, B \in \mathcal{S} \implies A \cap B \in \mathcal{S}$
- se $A_1, A \in \mathcal{S}$, $A_1 \subset A$ allora esistono $\{A_k \in \mathcal{S} : k = 2, \dots, n\}$, a due a due disgiunti, tali che

$$A = \bigcup_{k=1}^n A_k$$

Un semianello è quindi chiuso rispetto all'operazione di unione, tuttavia operando con intersezione e complementare si esce dalla famiglia dei rettangoli.

Definizione

Una famiglia di insiemi $\mathcal{A}$ si chiama **anello** se

$$A, B \in \mathcal{A} \implies A \Delta B = (A \setminus B) \cup (B \setminus A) \in \mathcal{A}, \quad A \cap B \in \mathcal{A}$$

Poichè

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

$$A \Delta (A \cap B) = A \setminus (A \cap B) \cup (A \cap B) \setminus A = A \setminus B \cup \emptyset = A \setminus B$$

Ogni anello risulta chiuso rispetto alle operazioni di differenza simmetrica, intersezione, unione, differenza, purchè applicate un numero finito di volte.

Poichè

$$A \setminus A = \emptyset$$

ogni anello contiene l'insieme vuoto;

- Il più piccolo anello è costituito dal solo insieme vuoto.
- L'insieme $\mathcal{P}(A)$ delle parti di un insieme A è un anello
- L'insieme $\{A, \emptyset\}$ è un anello
- la famiglia dei sottoinsiemi finiti di un insieme A è un anello di insiemi.
- La famiglia dei sottoinsiemi limitati della retta reale è un anello.

Definizione

Un anello di insiemi ammette un'unità $E \in \mathcal{A}$ se

$$E \cap A = A \quad \forall A \in \mathcal{A}$$

e si ha

$$E \supset \bigcup_{A \in \mathcal{A}} A$$

E è l'elemento massimale rispetto all'inclusione della famiglia di insiemi $\mathcal{A}$. Infatti

$$E \supset A \iff E \cap A = A$$

Definizione

Si chiama **algebra** di insiemi un anello $\mathcal{S}$ che ammetta un'unità

- L'insieme $\mathcal{P}(A)$ delle parti di un insieme A è un'algebra la cui unità è A
- L'insieme $\{A, \emptyset\}$ è un'algebra la cui unità è A
- la famiglia dei sottoinsiemi finiti di un insieme A è un anello di insiemi e risulta un'algebra se e solo se A è esso stesso un insieme finito.
- La famiglia dei sottoinsiemi limitati della retta reale è un anello di insiemi senza unità.

Teorema

Se $\mathcal{A}_\alpha$ è una famiglia di anelli allora

$$\mathcal{A} = \bigcap_{\alpha} \mathcal{A}_\alpha$$

è esso stesso un anello.

Teorema

Per ogni famiglia non vuota $\mathcal{S}$ di sottoinsiemi esiste uno ed un solo anello $\mathcal{A}(\mathcal{S})$ che contiene $\mathcal{S}$ ed è contenuto in ogni anello che contiene $\mathcal{S}$. $\mathcal{A}(\mathcal{S})$ si chiama anello generato dalla famiglia di insiemi $\mathcal{S}$.

Possiamo pertanto considerare l'anello generato dal semianello dei rettangoli

Definizione

Sia $\mathcal{E} = \mathcal{A}(\mathcal{S})$ l'anello generato dal semianello dei rettangoli. chiameremo gli elementi $A \in \mathcal{E}$ insiemi elementari

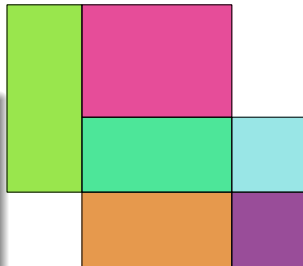
Pertanto $\mathcal{E}$ sarà identificato come la famiglia degli insiemi elementari.

Ogni famiglia di insiemi $\mathcal{S}$ genera un anello $\mathcal{A}(\mathcal{S})$,
 La costruzione di $\mathcal{A}(\mathcal{S})$ può non essere banale.
 È facile costruire $\mathcal{A}(\mathcal{S})$ nel caso in cui $\mathcal{S}$ sia un semianello.

Teorema

Sia $\mathcal{S}$ un semianello allora l'anello $\mathcal{A}(\mathcal{S})$ generato da $\mathcal{S}$ coincide con la famiglia $\mathcal{B}$ degli insiemi A che ammettono una decomposizione finita del tipo

$$A = \bigcup_{k=1}^n A_k \quad , \quad A_k \in \mathcal{S}$$



Definiamo la misura di un rettangolo in modo elementare per cui risulta definita una funzione

$$m : \mathcal{S} \rightarrow \mathbb{R}$$

tale che

- ① $m(\emptyset) = 0$
- ② $m(R) = (b - a)(d - c)$ essendo a, b e c, d gli estremi degli intervalli che definiscono R
- ③ La misura m assume solo valori positivi o nulli.
- ④ La misura m è additiva, cioè se $P = \bigcup_{k=1}^n P_k$ con $P_i \cap P_j = \emptyset$ per $i \neq j$ allora

$$m(P) = \sum_{k=1}^n m(P_k)$$

m è una misura additiva sul semianello $\mathcal{S}$.

Possiamo estendere m su $\mathcal{S}$ ad m' su $\mathcal{E}$ mediante la

Definizione

Se $A = \bigcup_{k=1}^n P_k \in \mathcal{E}$ definiamo

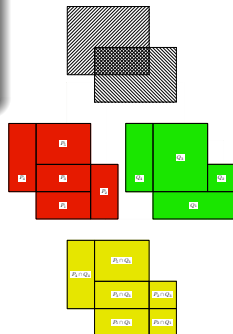
$$m'(A) = \sum_{k=1}^n m(P_k)$$

$m'(A)$ non dipende dalla maniera di rappresentare A come unione finita di rettangoli; infatti se

$$A = \bigcup_{k=1}^n P_k = \bigcup_{k=1}^m Q_k$$

si ha

$$m'(A) = \sum_{k=1}^n \sum_{h=1}^m m(P_k \cap Q_h)$$



si ha evidentemente che

- ① Se R è un rettangolo allora $m'(R) = m(R)$
- ② Se $A, B \in \mathcal{E}$ e $A \subset B$ avremo che

$$A \cup B \subset A \cup (B \setminus A) \quad \text{e} \quad A \cap (B \setminus A) = \emptyset$$

Quindi

$$m(A \cup B) \leq m(A) + m(B \setminus A) \leq m(A) + m(B)$$

per cui

m risulta finitamente subadditiva.

Se ci limitiamo a considerare solo la parte di $\mathbb{R}^2$ contenuta in un rettangolo $\mathcal{R}$ fissato allora $\mathcal{E}$ ammette un'unità e possiamo in tal caso provare che

Teorema

Siano $A, A_n \in \mathcal{E}$ per $n \in \mathbb{I}$ dove $\mathbb{I}$ è un insieme di indici, finito o numerabile, tali che

$$A \subset \bigcup_{n \in \mathbb{I}} A_n$$

allora

$$m'(A) \leq \sum_{n \in \mathbb{I}} m'(A_n)$$

In altre parole la misura m' è numerabilmente subadditiva su $\mathcal{E}$

Teorema

La subadditività numerabile della misura m' implica la sua additività numerabile cioè la sua σ -additività.

È interessante osservare che l'additività numerabile non si ottiene mediante un semplice passaggio al limite dalla additività finita. Occorre sfruttare le proprietà di compattezza di un insieme elementare chiuso e limitato.

Ricordiamo che stiamo lavorando in un rettangolo $\mathcal{R}$ fissato; possiamo allora dire che

Definizione

Chiamiamo misura esterna di un insieme A del piano il numero

$$\mu^*(A) = \inf_{A \subset \bigcup_k P_k} \sum m(P_k)$$

essendo l'estremo inferiore esteso a tutti i ricoprimenti finiti o numerabili dell'insieme A mediante rettangoli.

Osserviamo che:

- i rettangoli P_k possono essere supposti a due a due disgiunti;
- la soma (eventualmente infinita) che compare a secondo membro è la somma di termini positivi ed è maggiorata dalla misura di $\mathcal{R}$, quindi è convergente.
- la definizione di misura esterna non cambia se sostituiamo i rettangoli P_k con gli insiemi elementari;

Se A è un insieme elementare,

$$m'(A) = \sum_{k=1}^n m(P_k)$$

Pertanto

$$\mu^*(A) \leq \sum_{k=1}^n m(P_k) = m'(A)$$

d'altro canto se Q_j è una famiglia finita o numerabile di rettangoli che ricopre A , per la σ -subaddittività di m' si ha

$$m'(A) \leq \sum_{j=1}^n m'(Q_j) \leq \sum_{j=1}^n m(Q_j) \quad \text{e} \quad m'(A) \leq \mu^*(A)$$

Per cui, nel caso in cui A sia un insieme elementare,

$$\mu^*(A) = m'(A)$$

essendo l'estremo inferiore nella definizione di μ^* un minimo.

Teorema

Se

$$A \subset \bigcup_{n \in \mathbb{I}} A_n$$

dove $\{A_n, n \in \mathbb{I}\}$ è una famiglia finita o numerabile di insiemi, allora si ha

$$\mu^*(A) \leq \sum_{n \in \mathbb{I}} \mu^*(A_n)$$

In particolare $A \subset B \implies \mu^*(A) \leq \mu^*(B)$

Il teorema precedente dimostra come caso particolare la σ -subadditività della misura m' non appena si tenga conto che, sugli insiemi elementari la misura μ^* e la misura m' coincidono.

Tuttavia la misura esterna μ^* non risulta additiva. (esempio di Vitali)

Definizione

Diciamo che un insieme A è misurabile nel senso di Lebesgue, se per ogni $\epsilon > 0$ esiste un insieme elementare $B \in \mathcal{E}$ tale che

$$\mu^*(A \Delta B) < \epsilon$$

Indichiamo con $\mathcal{L}$ la famiglia degli insiemi misurabili secondo Lebesgue.

Se un insieme A è misurabile indichiamo semplicemente con $\mu(A)$ la sua misura esterna $\mu^(A)$ e la chiamiamo misura di Lebesgue dell'insieme A .*

Teorema

Se $A \in \mathcal{L}$ allora $\mathcal{R} \setminus A \in \mathcal{L}$; in altre parole se A è misurabile secondo Lebesgue anche il complementare di A è misurabile secondo Lebesgue.

Teorema

Se $A_1, A_2 \in \mathcal{L}$ allora

$$A_1 \cup A_2, A_1 \cap A_2 \in \mathcal{L}$$

In altre parole la classe degli insiemi misurabili secondo Lebesgue è chiusa rispetto all'unione ed all'intersezione finita.

Corollario

Se $A_1, A_2 \in \mathcal{L}$ allora

$$A_1 \setminus A_2, A_1 \Delta A_2 \in \mathcal{L}$$

In altre parole la classe degli insiemi misurabili secondo Lebesgue è chiusa rispetto alla differenza ed alla differenza simmetrica finita.

Quindi $\mathcal{L}$ è un anello di insiemi su cui è definita una misura $\mu = \mu^*$ numerabilmente subadditiva

Lemma

Siano A, B due insiemi allora

$$|\mu^*(A) - \mu^*(B)| \leq \mu^*(A \Delta B)$$

Teorema

Siano A_1, A_2 due insiemi misurabili disgiunti, allora

$$\mu \left(\bigcup_{k=1}^2 A_k \right) = \sum_{k=1}^2 \mu(A_k)$$

È chiaro che l'enunciato si può facilmente estendere ad un qualunque numero finito di insiemi a due a due disgiunti. Ciò permette di concludere che μ è una misura (finitamente) additiva.

Dal teorema appena dimostrato si deduce in particolare che

$$\mu(\mathcal{R}) = \mu(A) + \mu(\mathcal{R} \setminus A) \quad \text{e} \quad \mu(\mathcal{R} \setminus A) = 1 - \mu(A)$$

Teorema

Per ogni insieme misurabile $A \in \mathcal{L}$, e per ogni $\epsilon > 0$ esiste un insieme $B = \bigcup_n P_n$ dove gli insiemi P_n sono rettangoli aperti, tale che

$$A \subset B \quad 0 \leq \sum_n m(P_n) - \mu^*(A) = \sum_n m(P_n) - \mu(A) < \epsilon$$

Inoltre per ogni $\epsilon > 0$ esiste un insieme C chiuso tale che

$$C \subset A \quad 0 \leq \mu(A) - \mu(C) < \epsilon$$

Teorema

Se $\{A_k \in \mathcal{L} : k \in \mathbb{N}\}$ è una famiglia di insiemi misurabili numerabile, allora

$$\bigcup_{k=1}^{+\infty} A_k \quad \bigcap_{k=1}^{+\infty} A_k$$

sono insiemi misurabili.

Cioè $\mathcal{L}$ è un σ -anello e un δ -anello; poichè è evidente che $\mathcal{R} = [0, 1] \times [0, 1]$ è una unità per $\mathcal{L}$, possiamo affermare che $\mathcal{L}$ è una σ -algebra.

Teorema

Se $\{A_k \in \mathcal{L} : k \in \mathbb{N}\}$ è una famiglia numerabile di insiemi misurabili a due a due disgiunti, e se

$$A = \bigcup_{k=1}^{+\infty} A_k$$

allora

$$\mu(A) = \sum_{k=1}^{+\infty} \mu(A_k)$$

Il teorema precedente dimostra la σ -additività della misura di Lebesgue.

la σ —additività implica la continuità della misura μ , in altre parole si ha:

Teorema

Sia

$$A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \supset \dots$$

una successione di insiemi misurabili decrescente rispetto all'inclusione e sia

$$A = \bigcap_{k=1}^{+\infty} A_k$$

Allora

$$\mu(A) = \lim_k \mu(A_k)$$

Corollario

Sia

$$A_1 \subset A_2 \subset A_3 \subset \dots \subset A_n \subset \dots$$

una successione di insiemi misurabili crescente rispetto all'inclusione e sia

$$A = \bigcup_{k=1}^{+\infty} A_k$$

Allora

$$\mu(A) = \lim_k \mu(A_k)$$

Poichè $\emptyset$ è un insieme elementare, ogni insieme A la cui misura esterna $\mu^*(A)$ è nulla, è misurabile. Infatti

$$\mu^*(A \setminus \emptyset) = \mu^*(A) = 0 < \epsilon$$

Abbiamo con ciò completato la costruzione di una misura μ

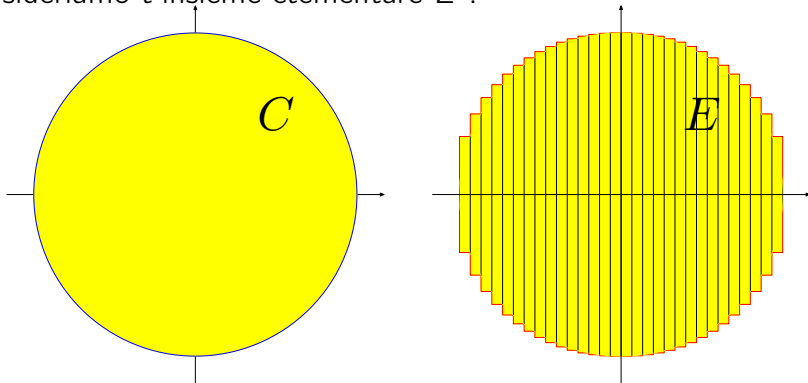
- definita su una σ –algebra di insiemi $\mathcal{L}$,
- σ –additiva,
- μ si chiama misura di Lebesgue nel piano
- $\mathcal{L}$ è la σ –algebra di Lebesgue nel piano;
- $\mathcal{L} \supset \mathcal{E}$,

Cioè la σ –algebra degli insiemi di Lebesgue contiene l'anello degli insiemi elementari

L'uso della definizione per provare la misurabilità di un insieme può essere facilmente implementato. Sia ad esempio

$$C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

Consideriamo l'insieme elementare E .



La figura mostra la differenza
simmetrica

$$C \Delta E = E \setminus C$$

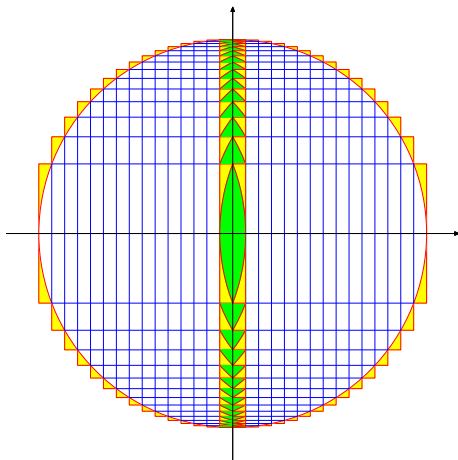
e mostra altresì che

$$C \setminus E$$

è contenuto in una striscia la cui altezza è 2 e la cui base può essere resa piccola quanto si vuole.

Ciò basta per assicurare la misurabilità di C .

Quanto fatto può essere adattato a molte e diverse situazioni.



Se

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

possiamo considerare l'insieme elementare E definito come l'unione numerabile dei quadrati di lato $\frac{\epsilon}{\sqrt{2}^n}$ aventi centro ciascuno in uno dei punti (numerabili) di A .

Evidentemente

$$E \Delta A \subset E$$

e si ha che

$$m(E) = \sum_1^{+\infty} \frac{\epsilon}{2^n} = \epsilon$$

Ne segue la misurabilità ed il fatto che

$$m(A) = 0$$

Una simile costruzione può essere ripetuta su ogni insieme limitato di uno spazio euclideo di dimensione finita.

La misura di Lebesgue nel piano è stata definita a partire dalla definizione di misura di un rettangolo;

il procedimento che abbiamo seguito si fonda unicamente su alcune proprietà astratte.

Più precisamente, ci siamo serviti delle seguenti proprietà:

- 1 L'area è una funzione definita sui rettangoli che assume valori positivi
- 2 L'area è una funzione (finitamente) additiva
- 3 i rettangoli nel piano formano un semi-anello

Fin qui abbiamo definito la misura di sottoinsiemi di un fissato rettangolo e quindi solo di insiemi limitati.

Possiamo estendere la definizione ad insiemi non limitati ma di misura finita

Possiamo considerare il semianello dei rettangoli in $\mathbb{R}^2$ e l'anello da essi generato tuttavia, in questo caso, l'anello generato non ammette unità.

Sia pertanto

$$m : \mathcal{S} \rightarrow \mathbb{R}$$

la misura σ -additiva definita sul semianello $\mathcal{S}$ dei rettangoli in $\mathbb{R}^2$ e sia $\mathcal{P}$ l'insieme delle parti di $\mathbb{R}^2$.
definiamo

$$\mu^* : \mathcal{P} \rightarrow \mathbb{R}$$

mediante la

$$\mu^*(A) = \inf \left\{ \sum_n m(B_n) \right\}$$

essendo l'estremo inferiore esteso a tutti i ricoprimenti finiti o numerabili di A mediante insiemi $B_n \in \mathcal{S}$ per tutti quegli insiemi A che ammettono almeno un ricoprimento $\bigcup_n A_n$, $A_n \in \mathcal{S}$ tale che $\sum_n m(A_n) < +\infty$.

Definizione

Diciamo che un insieme A è misurabile nel senso di Lebesgue, se per ogni $\epsilon > 0$ esiste un insieme $B \in \mathcal{A}(S)$ tale che

$$\mu^*(A \Delta B) < \epsilon$$

Indichiamo con $\mathcal{M}$ la famiglia degli insiemi misurabili secondo Lebesgue. La funzione μ^ , considerata solo su $\mathcal{M}$ si chiama misura di Lebesgue e si denota semplicemente con μ .*

È immediato verificare che se $A \in \mathcal{S}_m$ si ha

$$\mu(A) = m(A)$$

Inoltre, poichè μ^ è subadditiva, si ha*

$$\mu^*(A) \leq \mu^*(B) + \mu^*(A \Delta B) < +\infty$$

ogni volta che A è misurabile.

Teorema

La classe $\mathcal{M}$ degli insiemi misurabili è un anello.

Il teorema che asserisce che la classe degli insiemi misurabili è una σ -algebra va modificato come segue

Definizione

Se $A \in \mathcal{M}$ chiamiamo $\mathcal{M}_A$ la famiglia degli insiemi misurabili contenuti in A . In altre parole

$$\mathcal{M}_A = \{B \in \mathcal{M} : B \subset A\}$$

Teorema

Se $A \in \mathcal{M}$ allora $\mathcal{M}_A$ è una σ -algebra

Teorema

Se m è una misura su un semianello $\mathcal{S}$ senza unità, allora la classe $\mathcal{M}$ degli insiemi misurabili secondo Lebesgue è un δ -anello; se $\{A_n : n \in \mathbb{N}\}$ è una famiglia di insiemi misurabili secondo Lebesgue, allora

$$A = \bigcup_n A_n$$

è misurabile se e solo se esiste una costante K indipendente da N tale che

$$\mu \left(\bigcup_{n=1}^N A_n \right) \leq K \quad (57)$$

Si può dimostrare che la misura ottenuta è completa

Per concludere occorre anche definire la misurabilità di insiemi non limitati di misura infinita.

A questo fine osserviamo che possiamo trovare in $\mathbb{R}^2$ una famiglia di sottoinsiemi misurabili (ad esempio rettangoli) B_n con $n \in \mathbb{N}$ tale che

$$\mathbb{R}^2 = \bigcup_{n \in \mathbb{N}} B_n$$

Indentificheremo questa proprietà dicendo che la misura che abbiamo definito in $\mathbb{R}^2$ è σ -finita.

La proprietà potrà anche essere considerata in astratto.

Sia pertanto

$$m : \mathcal{S} \rightarrow \mathbb{R}$$

la misura σ -additiva e σ -finita che abbiamo definito sul semianello $\mathcal{S}$ dei rettangoli.

Passando all'anello $\mathcal{A}(\mathcal{S})$ generato da $\mathcal{S}$ e sostituendo B_n con

$$B_n \setminus \bigcup_{k=1}^{n-1} B_k$$

possiamo supporre che

$$\mathbb{R}^2 = \bigcup_{n \in \mathbb{N}} B_n \quad B_n \in \mathcal{A}(\mathcal{S}) \quad B_n \cap B_m = \emptyset \quad n \neq m$$

e che B_n sia misurabile.

Applicando ad m il procedimento di prolungamento della misura secondo Lebesgue otteniamo una misura μ definita su un δ -anello $\mathcal{M}$.

Se $B \in \mathcal{M}$, chiamiamo $\mathcal{M}_B$ la classe degli insiemi misurabili contenuti in B , cioè

$$\mathcal{M}_B = \{C \in \mathcal{M} : C \subset B\}$$

$\mathcal{M}_B$ è una σ -algebra la cui unità è B .

Consideriamo ora la famiglia

$$\mathcal{U} = \{A : A \cap B_n \in \mathcal{M}_{B_n}, n \in \mathbb{N}\} = \{A : A \cap B_n \in \mathcal{M}, n \in \mathbb{N}\}$$

In altre parole $\mathcal{U}$ è costituita dagli insiemi A tali che

$$A = \bigcup_{n \in \mathbb{N}} A_n \quad A_n \in \mathcal{M}_{B_n}$$

Si può verificare che $\mathcal{U}$ è una σ -algebra che chiameremo somma diretta delle σ -algebre $\mathcal{M}_{B_n}$.

Ad esempio se

$$A = \bigcup_k A_k$$

con $A_k \in \mathcal{U}$, allora esistono degli insiemi

$$C_n^k = A_k \cap B_n \in \mathcal{M}_{B_n}$$

tali che

$$A = \bigcup_k \bigcup_n C_n^k = \bigcup_n \bigcup_k C_n^k \quad \bigcup_k C_n^k \in \mathcal{M}_{B_n}$$

e quindi $A \in \mathcal{U}$.

Chiameremo misurabili gli insiemi della σ -algebra $\mathcal{U}$ e definiremo

$$\tilde{\mu} : \mathcal{U} \rightarrow \mathbb{R}$$

mediante la

$$\tilde{\mu}(A) = \sum_{n=1}^{+\infty} \mu(A_n)$$

essendo

$$A = \bigcup_{n \in \mathbb{N}} A_n \quad A_n \in \mathcal{M}_{B_n}$$

Poichè la serie a secondo membro è a termini positivi, possiamo affermare che la serie considerata è incondizionatamente convergente o positivamente divergente.

Per giustificare la definizione di $\tilde{\mu}$ occorre verificare che

Teorema

La σ -algebra $\mathcal{U}$ e la misura $\tilde{\mu}$ non dipendono dalla famiglia di insiemi disgiunti $B_n \in \mathcal{M}$ tali che $\bigcup_n B_n = X$.

Inoltre

$$\tilde{\mu}(A) = \mu(A) \quad \forall A \in \mathcal{M}$$

e $\tilde{\mu}$ è σ -additiva su $\mathcal{M}$.

Si può poi dimostrare che.

Teorema

La misura $\tilde{\mu}$ è σ -additiva su $\mathcal{U}$; inoltre

$$\{A \in \mathcal{U} : \tilde{\mu}(A) < +\infty\} = \mathcal{M}$$

è un δ -anello.

Funzioni misurabili

Sia in $\mathbb{R}$ che in $\mathbb{R}^2$ (e più in generale in $\mathbb{R}^n$) possiamo considerare una famiglia $\mathcal{T}$ di insiemi che chiamiamo insiemi aperti e che definisce una topologia.

Una topologia ci permette di definire gli intorni di un punto e di perseguire lo studio delle proprietà locali dello spazio e delle funzioni su esso definite.

Uno spazio dotato di una topologia si chiama spazio topologico. Se $A \subset X$, A è aperto e A^c è chiuso.

Essenzialmente $\mathcal{T}$, deve contenere $\mathbb{R}^n$ e $\emptyset$ e deve essere chiusa rispetto ad intersezione finita ed unione numerabile.

Il concetto di topologia è legato essenzialmente al concetto di limite e alla continuità.

Le funzioni continue definite tra spazi topologici conservano le proprietà topologiche.

La definizione stessa di continuità recita che una funzione è continua se la controimmagine di ogni aperto è un aperto. e si riduce alle ben note definizioni $\epsilon - \delta$ nel caso di spazi euclidei con la topologia usuale (indotta dalla metrica).

Tuttavia una topologia non ha la struttura di σ - algebra, pertanto

Definizione

Chiamiamo $\mathcal{B}$ la σ -algebra generata dagli aperti di $\mathbb{R}^n$. $\mathcal{B}$ si chiama σ -algebra degli insiemi di Borel, o Borelliani, in $\mathbb{R}^n$.

Per la definizione di misura esterna e di misurabilità si ha che

$$\mathcal{L} \supset \mathcal{B}$$

In altre parole ogni insieme borelliano è misurabile secondo Lebesgue.

Teorema

Sia X uno spazio topologico e sia $\mathcal{B}$ la σ -algebra degli insiemi di Borel in X ; allora se $A \in \mathcal{B}$

$$A = \bigcup_{n \in \mathbb{N}} G_n = \bigcap_{n \in \mathbb{N}} F_n$$

dove F_n sono insiemi chiusi e G_n sono aperti,

Teorema

La σ -algebra degli insiemi di Borel in $\mathbb{R}$ è generata da ciascuna delle seguenti famiglie di insiemi.

- ① *Gli intervalli del tipo $(-\infty, b)$*
- ② *Gli intervalli del tipo $[a, b)$*

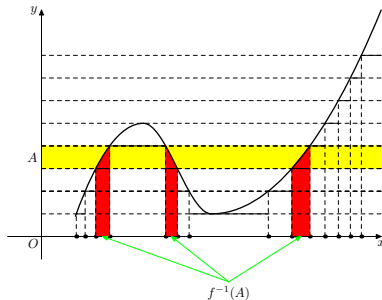
Definizione

Siano X, Y due insiemi, e siano $\mathcal{S}_X, \mathcal{S}_Y$ due famiglie di insiemi in X ed Y rispettivamente e sia

$$f : X \rightarrow Y$$

Diciamo che f è $(\mathcal{S}_X, \mathcal{S}_Y)$ -misurabile se

$$A \in \mathcal{S}_Y \implies f^{-1}(A) \in \mathcal{S}_X$$



Se ad esempio $\mathcal{S}_X$ ed $\mathcal{S}_Y$ sono le famiglie degli aperti di X ed Y rispettivamente una funzione f è $(\mathcal{S}_X, \mathcal{S}_Y)$ - misurabile se e solo se f è continua.

Definizione

Nel caso in cui $\mathcal{S}_X = \mathcal{B}_X$ ed $\mathcal{S}_Y = \mathcal{B}_Y$ siano la famiglia degli insiemi Borelliani in X e Y rispettivamente, una funzione $(\mathcal{S}_X, \mathcal{S}_Y)$ - misurabile si dice $\mathcal{B}$ -misurabile, o misurabile nel senso di Borel.

Definizione

Siano X un insieme, sia μ una misura σ -additiva su X e sia $\mathcal{B}$ la σ -algebra degli insiemi Borelliani in $\mathbb{R}$; diciamo che

$$f : X \rightarrow \mathbb{R}$$

è una funzione misurabile se

$$A \in \mathcal{B} \implies f^{-1}(A) \in \mathcal{S}_X$$

Teorema

Siano X, Y, Z tre insiemi e siano $\mathcal{S}_X, \mathcal{S}_Y, \mathcal{S}_Z$ tre famiglie di insiemi di X, Y, Z rispettivamente; siano

$$f : X \rightarrow Y \quad g : Y \rightarrow Z$$

due funzioni f $(\mathcal{S}_X, \mathcal{S}_Y)$ - misurabile e g $(\mathcal{S}_Y, \mathcal{S}_Z)$ - misurabile, allora $g(f(\cdot))$ è $(\mathcal{S}_X, \mathcal{S}_Z)$ - misurabile. In altre parole la composta di funzioni misurabili è misurabile.

Corollario

Siano

$$f : X \rightarrow Y \quad g : Y \rightarrow \mathbb{R}$$

Se g è borelliana e f è misurabile allora $g(f(\cdot))$ è misurabile; in particolare se f è continua e g è misurabile, allora $f(g(\cdot))$ è misurabile.

Teorema

Sia $f : X \rightarrow \mathbb{R}$ una funzione e sia $\mathcal{S}_X$ una σ -algebra su X ; allora f è misurabile se e solo se

$$f^{-1}(\{-\infty, c\}) \in \mathcal{S}_X \quad \forall c \in \mathbb{R}$$

Spesso, la condizione enunciata nel teorema precedente viene assunta come definizione di misurabilità.

Possiamo altresì provare che

Teorema

La somma, la differenza, il prodotto di due funzioni misurabili è misurabile. Il quoziente di due funzioni misurabili è misurabile a condizione che il denominatore non sia nullo.

Per quanto riguarda il limite di una successione di funzioni misurabili possiamo provare che

Teorema

Sia $f_n : X \rightarrow \mathbb{R}$ una successione di funzioni misurabili e supponiamo che $f_n(x) \rightarrow f(x)$ per ogni $x \in X$, allora f è una funzione misurabile.

La misura costruita in $\mathbb{R}^2$ è σ -additiva e completa quindi ogni sottoinsieme di misura nulla è esso stesso misurabile.

Pertanto è chiaro che la misurabilità di una funzione non è influenzata da insiemi di misura nulla ed è uso porre la seguente definizione.

Definizione

Siano $f, g : X \rightarrow \mathbb{R}$ due funzioni e supponiamo che μ sia una misura associata ad una σ -algebra $\mathcal{S}_\mu$, diciamo che f e g sono equivalenti se

$$\mu(\{x \in X : f(x) \neq g(x)\}) = 0$$

In altre parole diciamo che f e g sono equivalenti se differiscono su un insieme di misura nulla.

Inoltre

Definizione

Diciamo che una proprietà vale quasi ovunque in X rispetto alla misura μ se è vera su $X \setminus N$ con $\mu(N) = 0$.

Teorema

Siano $f, g : X \rightarrow \mathbb{R}$ due funzioni definite su uno spazio dotato di una misura μ e supponiamo che f e g siano equivalenti, cioè differiscano per un insieme di misura nulla, allora f è misurabile se e solo se g è misurabile.

Teorema

Siano f, g due funzioni reali di variabile reale continue, allora f e g sono equivalenti se e solo se $f(x) = g(x)$ per ogni $x \in X$.

Definizione

Sia $f_n : X \rightarrow \mathbb{R}$ una successione di funzioni definite su X insieme dotato di una misura μ . Diciamo che la successione f_n è convergente ad f quasi ovunque se

$$f_n(x) \rightarrow f(x) \quad \forall x \in X \setminus N \quad \mu(N) = 0$$

Teorema

Sia f_n una successione di funzioni misurabili quasi ovunque convergente ad una funzione f , allora f è misurabile.

Definizione

Sia $f_n : X \rightarrow \mathbb{R}$ una successione di funzioni definite su X insieme dotato di una misura μ . Diciamo che la successione f_n è convergente ad f in misura se per ogni $\delta > 0$

$$\lim_n \mu(\{x \in X : |f_n(x) - f(x)| \geq \delta\}) = 0$$

Teorema

-Severini - Egoroff - Sia X un insieme di misura finita e sia f_n una successione di funzioni misurabili quasi ovunque convergente ad f su X allora per ogni $\epsilon > 0$ esiste un insieme $X_\epsilon \subset X$ tale che

- $\mu(X_\epsilon) > \mu(X) - \epsilon$ cioè $0 \leq \mu(X) - \mu(X_\epsilon) < \epsilon$
- la successione f_n è uniformemente convergente ad f sull'insieme X_ϵ

Teorema

Sia X un insieme di misura finita e sia f_n una successione di funzioni misurabili, convergente quasi ovunque ad una funzione f . Allora f_n è convergente in misura ad f .

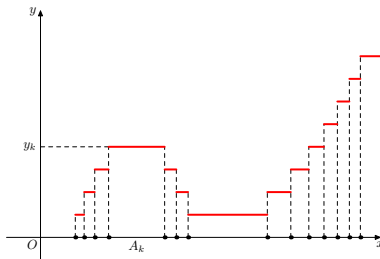
Teorema

Da ogni successione di funzioni f_n convergente in misura si può estrarre una successione convergente quasi ovunque.

Sia X è uno spazio dotato di una σ -algebra $\mathcal{S}_\mu$ su cui è definita una misura σ -additiva, completa.

Definizione

Diciamo che $s : X \rightarrow \mathbb{R}$ è una funzione semplice se è misurabile ed assume un numero finito o numerabile di valori.



Teorema

Una funzione s è semplice se e solo se

$$s(x) = \sum_{k \in \mathbb{I}} y_k \chi_{A_k}(x)$$

dove $y_k \in \mathbb{R}$, $A_k \in \mathcal{S}_\mu$ a due a due disgiunti ed $\mathbb{I}$ è un insieme finito o numerabile di indici e χ_A è la funzione caratteristica di A .

Teorema

Una funzione $f : X \rightarrow \mathbb{R}$ è misurabile se e solo se esiste una successione s_n di funzioni semplici uniformemente convergente ad f .

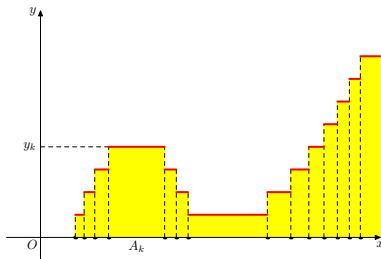
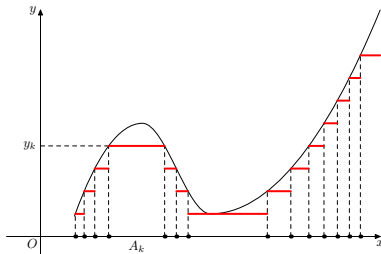
Possiamo sempre supporre che s_n sia crescente. e che, se f è positiva, le s_n siano funzioni positive.

Definizione

Se s è una funzione semplice definiamo

$$\int_X s(x) d\mu = \sum_{k \in \mathbb{I}} y_k \mu(A_k)$$

e diciamo che s è integrabile, o sommabile se la serie è assolutamente convergente.



Teorema

Sia s una funzione semplice, allora s è integrabile se e solo se $|s|$ è integrabile. Inoltre

$$\left| \int_X s d\mu \right| \leq \int_X |s| d\mu$$

Nella definizione di integrale di una funzione semplice è stato implicitamente usato il fatto che i valori assunti siano tra di loro distinti. Qualora ciò non fosse, si potrebbe, per ogni valore y_k assunto, ad esempio su A_k^1 e su A_k^2 , considerare l'insieme $A_k = A_k^1 \cup A_k^2$.

Ciò conduce alla necessità di eseguire la somma che definisce l'integrale della funzione semplice data usando un riordinamento dei termini. tuttavia non ne risultano ambiguità in quanto una funzione semplice è integrabile se la serie che definisce il suo integrale è assolutamente (e quindi anche incondizionatamente convergente).

Siano s e t due funzioni semplici integrabili su X e siano $\alpha, \beta \in \mathbb{R}$ allora $\alpha s + \beta t$ è una funzione semplice e risulta

$$\int_X \alpha s(x) + \beta t(x) d\mu = \alpha \int_X s(x) d\mu + \beta \int_X t(x) d\mu$$

Inoltre se s è una funzione semplice tale che $|s(x)| \leq M$ allora s è integrabile e si ha

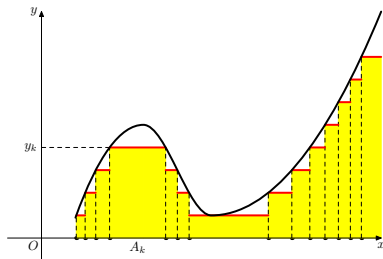
$$\left| \int_X s(x) d\mu \right| \leq M \mu(X)$$

Definizione

Sia f una funzione misurabile su X rispetto alla σ -algebra $\mathcal{S}_\mu$ associata alla misura μ . Diciamo che f è integrabile, o sommabile se esiste una successione di funzioni semplici s_n uniformemente convergenti ad f su X . In tal caso si definisce

$$\int_X f(x) d\mu = \lim_{n \rightarrow +\infty} \int_X s_n(x) d\mu$$

che esiste finito.



Osserviamo che se f è positiva, allora la successione s_n può essere scelta monotona crescente per cui

$$\int_X f(x) d\mu = \lim_{n \rightarrow +\infty} \int_X s_n(x) d\mu = \sup_n \int_X s_n(x) d\mu$$

Ovviamente la definizione precedente ha senso non appena si siano verificati i seguenti fatti:

- 1 Il limite che figura nella definizione esiste per ogni successione s_n di funzioni uniformemente convergenti ad f .
- 2 Il limite in questione non dipende dalla scelta della successione s_n .
- 3 Nel caso in cui f sia una funzione semplice la definizione precedente coincide con quella data per le funzioni semplici stesse.

La verifica dei fatti enumerati può essere fatta ricordando che

1

$$\left| \int_X s_n(x) d\mu - \int_X s_m(x) d\mu \right| \leq \mu(X) \sup_{x \in X} |s_n(x) - s_m(x)|$$

2

Se s_n e t_n sono due successioni di funzioni semplici uniformemente convergente ad f , allora anche la successione r_n definita da

$$r_n = \begin{cases} s_n & n = 2k \\ t_n & n = 2k + 1 \end{cases}$$

è una successione di funzioni semplici uniformemente convergente ad f .

Poichè per quanto abbiamo appena visto il limite di $\int_X s_n d\mu$, $\int_X t_n d\mu$ ed $\int_X r_n d\mu$ esistono, e poichè s_n , t_n dono successioni estratte da r_n i loro limiti devono essere tutti uguali.

- ③ Per dimostrare l'ultima affermazione possiamo osservare che se f è una funzione semplice possiamo scegliere la successione $s_n = f$

Usando le proprietà dei limiti e degli integrali delle funzioni semplici possiamo dimostrare che

Teorema

Sia f una funzione misurabile su X allora f è integrabile se e solo se per ogni $\epsilon > 0$ esiste una funzione semplice integrabile s tale che

$$|f(x) - s(x)| < \epsilon \quad \forall x \in X$$

.

Teorema

- ① Se $A \subset X$ è un insieme misurabile, allora $\int_A d\mu = \mu(A)$
- ② Se f è una funzione integrabile e se $\alpha \in \mathbb{R}$ allora
$$\int_X \alpha f d\mu = \alpha \int_X f d\mu$$
- ③ Se f, g sono funzioni integrabili allora
$$\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu$$
- ④ Ogni funzione limitata e misurabile su X risulta ivi integrabile.
- ⑤ Se f è una funzione misurabile $f \geq 0$ allora $\int_X f d\mu \geq 0$

Teorema

6 Se f, g sono funzioni integrabili, $f \geq g$ allora

$$\int_X f d\mu \geq \int_X g d\mu$$

7 Se f è una funzione misurabile

$$m \leq f(x) \leq M$$

$$\text{allora } m\mu(X) \leq \int_X f d\mu \leq M\mu(X)$$

8 Se f è misurabile, A è misurabile e $\mu(A) = 0$ allora $\int_A f d\mu = 0$

Teorema

- 9 *Se f, g sono due funzioni misurabili quasi ovunque uguali allora f è integrabile se e solo se g è integrabile e risulta $\int_X f d\mu = \int_X g d\mu$*
- 10 *Se f, φ sono funzioni misurabili tali che $|f(x)| \leq \varphi(x)$ quasi ovunque e se φ risulta integrabile allora anche f è integrabile su X*
- 11 *Se f è una funzione misurabile allora f è integrabile se e solo se $|f|$ è integrabile e risulta $|\int_X f d\mu| \leq \int_X |f| d\mu$*

Per quanto concerne la σ –additività dell'integrale. si ha:

Teorema

Sia f una funzione misurabile su X e sia $A \subset X$ un insieme misurabile. Se

$$A = \bigcup_{n \in \mathbb{N}} A_n \quad A_i \cap A_j = \emptyset \quad i \neq j$$

e se f è integrabile su A , allora anche

- $\int_{A_n} f(x) d\mu$ esiste,
- $\sum_{n \in \mathbb{N}} \int_{A_n} f(x) d\mu$ è assolutamente convergente

e risulta

$$\int_A f(x) d\mu = \sum_{n \in \mathbb{N}} \int_{A_n} f(x) d\mu$$

Corollario

Se f è una funzione misurabile, integrabile su X allora f è integrabile anche su ogni sottoinsieme Y misurabile $Y \subset X$.

Possiamo anche provare una sorta di teorema reciproco del precedente.

Teorema

Sia A un insieme misurabile tale che

$$A = \bigcup_{n \in \mathbb{N}} A_n \quad A_i \cap A_j = \emptyset \quad i \neq j$$

e supponiamo che $\sum_{n \in \mathbb{N}} \int_{A_n} |f(x)| d\mu$ sia convergente allora f è integrabile su A , e risulta

$$\int_A f(x) d\mu = \sum_{n \in \mathbb{N}} \int_{A_n} f(x) d\mu$$

Teorema

- **Tchebychev** - Sia φ una funzione integrabile su A , $\varphi(x) \geq 0$ e sia $c \in \mathbb{R}_+$ allora

$$\frac{1}{c} \int_A \varphi(x) d\mu \geq \mu(\{x \in A : \varphi(x) \geq c\})$$

Corollario

Sia f una funzione misurabile tale che

$$\int_A |f(x)| d\mu = 0$$

allora $f(x) = 0$ per quasi ogni $x \in A$.

Teorema

Sia f una funzione integrabile su $A \subset X$, allora

$$\forall \epsilon > 0 \exists \delta > 0 \text{ tale che } \left| \int_E f(x) d\mu \right| < \epsilon \quad \forall E \subset A \quad \mu(E) < \delta$$

Proviamo infine tre risultati di passaggio al limite sotto il segno di integrale.

Teorema

-Lebesgue - (Convergenza Dominata) *Sia f_n una successione di funzioni misurabili su A convergente ad f e supponiamo che esista φ integrabile tale che $|f_n(x)| \leq \varphi(x)$ Allora la funzione f è integrabile su A e si ha*

$$\int_A f_n(x) d\mu \rightarrow \int_A f(x) d\mu$$

Teorema

- **B. Levi - (Convergenza Monotona)** Sia f_n una successione monotona (crescente) di funzioni integrabili su A ; sia cioè

$$f_1(x) \leq f_2(x) \leq f_3(x) \leq \cdots \leq f_n(x) \leq \cdots$$

e supponiamo che

$$\int_A f_n(x) d\mu \leq K$$

Allora la successione f_n è convergente quasi ovunque ad una funzione integrabile

$$f(x) = \lim_n f_n(x)$$

ed inoltre

$$\int_A f_n(x) d\mu \rightarrow \int_A f(x) d\mu$$

Corollario

Sia f_n una successione di funzioni misurabili $f_n(x) \geq 0$ e sia

$$\sum_n \int_A f_n(x) d\mu < +\infty$$

sia inoltre

$$\sum_n f_n(x) = f(x)$$

quasi ovunque, allora

$$\int_A \sum_n f_n(x) d\mu = \sum_n \int_A f_n(x) d\mu$$

Teorema

- **Fatou** - Sia f_n una successione di funzioni misurabili, non negative e quasi ovunque convergenti ad f su A ; sia inoltre

$$\int_A f_n(x) d\mu \leq K$$

Allora la funzione f è integrabile su A e risulta

$$\int_A f(x) d\mu \leq K$$

Abbiamo fin qui considerato insiemi X su cui è assegnata una σ -algebra $\mathcal{S}$ ed una misura μ per i quali risulta $\mu(X) < +\infty$, abbiamo cioè considerato solo insiemi di misura finita.

Tuttavia è immediato osservare che tale restrizione non è accettabile in quanto preclude la possibilità di considerare integrazione su insiemi come rette, semirette o il piano intero.

Al fine di evitare questo inconveniente occorre estendere la definizione di integrale anche su insiemi di misura non limitata.

Allo scopo è opportuno dare una definizione di integrabilità un po' più generale; tale definizione consente anche di ammettere per la funzione integranda f valori infiniti, ma può essere data solo per funzioni positive. Siamo quindi condotti a considerare funzioni

$$f : X \rightarrow [0, +\infty]$$

È sufficiente definire

$$A_m^n = f^{-1} \left(\left[\frac{m}{2^n}, \frac{m+1}{2^n} \right) \right) \quad B_n = f^{-1}([n, +\infty])$$

e

$$s_n(x) = \sum_{m=0}^{n2^n-1} \frac{m}{2^n} \chi_{A_m^n}(x) + n \chi_{B_n}(x)$$

per dimostrare che

Teorema

Se

$$f : X \rightarrow [0, +\infty]$$

è misurabile allora esiste una successione s_n di funzioni semplici crescente e puntualmente convergente ad f .

Definizione

Sia $f : X \rightarrow [0, +\infty]$ misurabile, definiamo integrale di f mediante la

$$\int_X f d\mu = \sup_{s \in \mathcal{S}_f} \int_X s d\mu$$

ove $\mathcal{S}_f$ è l'insieme delle funzioni semplici s tali che $s \leq f$. Nel caso in cui f non abbia segno costante utilizziamo l'uguaglianza

$$f = f_+ + f_- \quad f_+ = \max\{f, 0\} \quad f_- = \min\{f, 0\}$$

e poniamo $\int_X f d\mu = \int_X f_+ d\mu + \int_X f_- d\mu$ non appena la somma a secondo membro ha senso.

Diciamo che f è integrabile tutte le volte che $\int_X f d\mu \in \mathbb{R}$. In tal caso è evidente che si ha anche $\int_X |f| d\mu \in \mathbb{R}$.

Osserviamo inoltre che nel caso in cui $f \geq 0$ ha sempre senso parlare di $\int_X f d\mu$ essendo il valore $+\infty$ un possibile valore.

Possiamo inoltre verificare che

Teorema

Se

$$f : X \rightarrow [0, +\infty)$$

è misurabile, $\mu(X)$ è finita e se $\int_X f d\mu < +\infty$ allora esiste una successione di funzioni semplici s_n uniformemente convergente ad f , integrabili tale che

$$\int_X f d\mu = \lim_n \int_X s_n d\mu$$

Teorema

- ① Se $A \subset X$ è un insieme misurabile, allora $\int_A d\mu = \mu(A)$
- ② Se f è una funzione integrabile e se $\alpha \in \mathbb{R}$ allora
$$\int_X \alpha f d\mu = \alpha \int_X f d\mu$$
- ③ Se f, g sono funzioni integrabili allora
$$\int_X (f + g)(x) d\mu = \int_X f d\mu + \int_X g d\mu$$
- ④ Se f è una funzione misurabile $f \geq 0$ allora $\int_X f d\mu \geq 0$
- ⑤ Se f, g sono funzioni integrabili, $f \geq g$ allora
$$\int_X f d\mu \geq \int_X g d\mu$$

Teorema

- ⑥ Se f è misurabile, A è misurabile e $\mu(A) = 0$ allora $\int_A f d\mu = 0$
- ⑦ Se f, g sono due funzioni misurabili quasi ovunque uguali allora f è integrabile se e solo se g è integrabile e risulta $\int_X f d\mu = \int_X g d\mu$
- ⑧ Se f, φ sono funzioni misurabili tali che $|f(x)| \leq \varphi(x)$ quasi ovunque e se φ risulta integrabile allora anche f è integrabile su X
- ⑨ Se f è una funzione misurabile allora f è integrabile se e solo se $|f|$ è integrabile e risulta $|\int_X f d\mu| \leq \int_X |f| d\mu$

Teorema

Sia f una funzione misurabile su X e sia $A \subset X$ un insieme misurabile. Se

$$A = \bigcup_{n \in \mathbb{N}} A_n \quad A_i \cap A_j = \emptyset \quad i \neq j$$

risulta

$$\int_A f(x) d\mu = \sum_{n \in \mathbb{N}} \int_{A_n} f(x) d\mu$$

Corollario

Sia f una funzione misurabile tale che

$$\int_A |f(x)| d\mu = 0$$

allora $f(x) = 0$ per quasi ogni $x \in A$.

Teorema

Sia f una funzione integrabile su $A \subset X$, allora

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \text{tale che} \quad \left| \int_E f(x) d\mu \right| < \epsilon \quad \forall E \subset A$$

Teorema

- **Lebesgue - (Convergenza Dominata)** Sia $f_n : X \rightarrow [0, +\infty]$ una successione di funzioni misurabili su A convergente ad f e supponiamo che esista φ integrabile tale che

$$|f_n(x)| \leq \varphi(x)$$

Allora la funzione f è integrabile su A e si ha

$$\lim_n \int_A |f_n - f| d\mu \rightarrow 0$$

da cui

$$\int_A f_n(x) d\mu \rightarrow \int_A f(x) d\mu$$

Teorema

- **B. Levi - (Convergenza Monotona)** Sia $f_n : X \rightarrow [0, +\infty]$ una successione monotona (crescente) di funzioni misurabili positive su A ; sia cioè

$$0 \leq f_1(x) \leq f_2(x) \leq f_3(x) \leq \cdots \leq f_n(x) \leq \dots$$

Allora la successione f_n è convergente quasi ovunque ad una funzione misurabile

$$f : X \rightarrow [0, +\infty] \qquad f(x) = \lim_n f_n(x)$$

ed inoltre

$$\int_A f_n(x) d\mu \rightarrow \int_A f(x) d\mu$$

Corollario

Sia f_n una successione di funzioni misurabili $f_n(x) \geq 0$ e sia $\sum_n \int_A f_n(x) d\mu < +\infty$ sia inoltre

$$\sum_n f_n(x) = f(x)$$

quasi ovunque, allora $\int_A \sum_n f_n(x) d\mu = \sum_n \int_A f_n(x) d\mu$

Teorema

- **Fatou** - Sia $f_n : X \rightarrow [0, +\infty]$ una successione di funzioni misurabili, non negative allora

$$\int_X \liminf_n f_n(x) d\mu \leq \liminf_n \int_A f_n(x) d\mu$$

Per quanto riguarda i legami con la teoria dell'integrazione secondo Riemann, possiamo provare con sufficiente semplicità che

Teorema

Se f è integrabile secondo Riemann, allora f è anche integrabile secondo Lebesgue e il valore dell'integrale non cambia. Inoltre f è integrabile secondo Riemann se e solo se è continua a meno di un insieme di misura nulla

In analisi riveste una notevole importanza il teorema di riduzione di un integrale multiplo a più integrali semplici iterati. Nell'ambito della teoria dell'integrazione secondo Lebesgue, tale risultato prende il nome di Teorema di Fubini e può essere dimostrato come segue. Allo scopo sarà necessario tutta una serie di nozioni e risultati preliminari.

Definizione

Siano $\mathcal{S}_x, \mathcal{S}_y$ famiglie di insiemi contenuti, rispettivamente, negli spazi X, Y . Sia

$$Z = X \times Y$$

il prodotto cartesiano degli spazi considerati e sia

$$\mathcal{A} = \mathcal{S}_x \times \mathcal{S}_y$$

la famiglia di insiemi in X i cui elementi sono definiti da

$$A \in \mathcal{A} \iff A = A_x \times A_y \quad A_k \in \mathcal{S}_k$$

Osserviamo che nel caso in cui $X = Y = \mathbb{R}$, ed $\mathcal{S}_m$ è il semianello degli intervalli in $\mathbb{R}$, allora $X = \mathbb{R}^2$ ed $\mathcal{A}$ è il semianello dei rettangoli in $\mathbb{R}^2$.

Possiamo provare con argomenti standard che

Teorema

Se $\mathcal{S}_x, \mathcal{S}_y$ sono semianelli allora anche $\mathcal{A} = \mathcal{S}_x \times \mathcal{S}_y$ è un semianello .

È bene tuttavia osservare che nel caso in cui le famiglie di insiemi $\mathcal{S}_k$ siano degli anelli o delle σ -algebre, non è detto che tale risulti anche il loro prodotto.

Infatti se $\mathcal{S}_x = \mathcal{S}_y \in \mathcal{P}(\mathbb{R})$, e se

$$A = [0, 1] \times [0, 2] \in \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R}) \quad B = [1, 2] \times [0, 1] \in \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R})$$

non è vero che $A \cup B \in \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R})$.

Definizione

Siano $\mathcal{S}_x, \mathcal{S}_y$ semianelli su cui siano assegnate delle misure μ_x, μ_y . Definiamo su $\mathcal{A}$ una misura

$$\mu = \mu_x \times \mu_y$$

mediante la

$$\mu(A) = \mu_x(A_x) \times \mu_y(A_y) \quad \text{per} \quad A = A_x \times A_y$$

La misura μ si dice misura prodotto.

Teorema

La definizione precedente individua effettivamente una misura cioè una funzione di insieme additiva, che verrà indicata con il nome di prodotto delle misure μ_k .

Teorema

Se le misure μ_x, μ_y sono σ –additive, allora anche la misura prodotto μ è σ –additiva.

Definizione

Se μ_x, μ_y sono misure σ –additive sulle σ –algebre $\mathcal{S}_x, \mathcal{S}_y$, definiamo misura prodotto delle misure μ_x, μ_y il prolungamento secondo Lebesgue della misura

$$\mu_x \times \mu_y$$

Tale prolungamento sarà indicato come

$$\mu = \mu_x \otimes \mu_y$$

È opportuno osservare che la misura prodotto risulta automaticamente completa anche se tali non sono i fattori; essa inoltre risulta σ –additiva.

Teorema

Per ogni insieme misurabile $A \subset X \times Y$ definiamo

$$A_x = \{y : (x, y) \in A\}$$

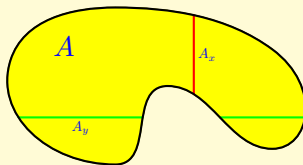
$$A_y = \{x : (x, y) \in A\}$$

Si ha

$$\mu(A) = \int_X \mu_y(A_x) d\mu_x = \int_Y \mu_x(A_y) d\mu_y$$

(Osserviamo che gli $\int_X \mu_y(A_x) d\mu_x$ e $\int_Y \mu_x(A_y) d\mu_y$ sono in realtà estesi soltanto agli insiemi

$U = \{x \in X : A_x \neq \emptyset\}$ e $V = \{y \in Y : A_y \neq \emptyset\}$, rispettivamente.)



Corollario

Sia f una funzione non negativa integrabile su X e sia $A = \{(x, y) \in D \times \mathbb{R} : 0 \leq y \leq f(x)\}$ $D \subset X$ allora

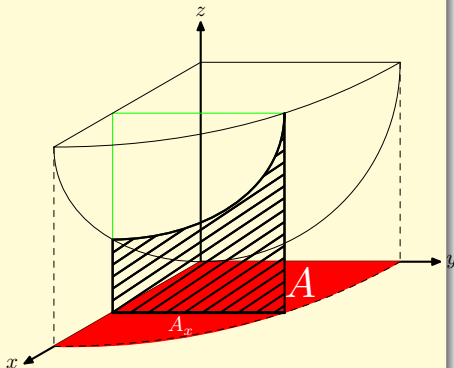
$$\mu(A) = \int_D f(x) d\mu_x$$

Teorema

- **Fubini** - Siano μ_x, μ_y misure definite su σ -algebre di sottoinsiemi degli spazi X ed Y , rispettivamente; supponiamo che μ_x e μ_y siano complete e σ -additive. Indichiamo con $\mu = \mu_x \otimes \mu_y$ la misura prodotto e supponiamo che $f(x, y)$ sia una funzione integrabile rispetto a μ sull'insieme $A \subset X \times Y$

Allora

$$\begin{aligned} \int_A f(x, y) d\mu &= \\ &= \int_X \left(\int_{A_x} f(x, y) d\mu_y \right) d\mu_x = \\ &= \int_Y \left(\int_{A_y} f(x, y) d\mu_x \right) d\mu_y \end{aligned}$$



Osserviamo che gli integrali estesi ad X e ad Y sono in realtà estesi soltanto agli insiemi $U = \{x \in X : A_x \neq \emptyset\}$ e $V = \{y \in Y : A_y \neq \emptyset\}$, rispettivamente.

Definizione

Sia $f : [a, b] \rightarrow \mathbb{R}$, diciamo che f è monotona in $[a, b]$ se

$$[f(x) - f(y)](x - y) \geq 0 \quad \forall x, y \in [a, b]$$

Possiamo facilmente provare che per le funzioni monotone valgono alcuni importanti risultati che riassumiamo qui di seguito.

Teorema

Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione monotona, allora per ogni $x_0 \in [a, b]$ f ammette limite da sinistra e da destra

$$f(x_0-) = \lim_{x \rightarrow 0^-} f(x) \quad f(x_0+) = \lim_{x \rightarrow 0^+} f(x)$$

e, se ad esempio supponiamo f crescente, si ha

$$f(x_0-) \leq f(x_0+)$$

Dal teorema precedente è evidente che

Teorema

Se $f : [a, b] \rightarrow \mathbb{R}$ è una funzione monotona allora f può essere discontinua in x_0 solo nel caso in cui

$$f(x_0-) \neq f(x_0+)$$

Diciamo, in questo caso, che f ammette in x_0 un “salto”. Diciamo altresì che f è continua da destra, o da sinistra in x_0 se accade che

$$f(x_0+) = f(x_0) \qquad f(x_0-) = f(x_0)$$

rispettivamente.

Teorema

Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione monotona allora f è continua su $[a, b]$ tranne un insieme al più numerabile di punti.

Teorema

Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione monotona (crescente), allora f è limitata e misurabile su $[a, b]$ e quindi risulta ivi integrabile.

Definizione

Diciamo che $h : [a, b] \rightarrow \mathbb{R}$ è una funzione di salto se esiste un sottoinsieme D al più numerabile di $[a, b]$

$$D = \{x_1, x_2, \dots, x_n, \dots\}$$

che supponiamo ordinato $x_1 < x_2 < \dots < x_n < \dots$ cui è associato un insieme di valori positivi $h_1, h_2, \dots, h_n, \dots$ in modo che

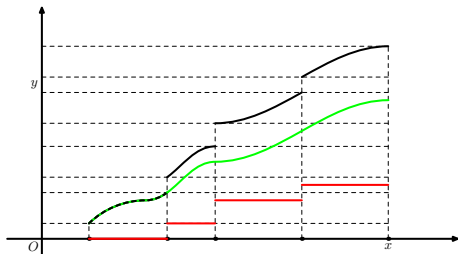
$$h(x) = \sum_{\{k: x_k < x\}} h_k$$

Teorema

Sia $h : [a, b] \rightarrow \mathbb{R}$ una funzione di salto, allora h è una funzione monotona crescente continua da sinistra in ogni punto ed è continua in ogni punto di $[a, b] \setminus D$.

Teorema

Se $f : [a, b] \rightarrow \mathbb{R}$ è una funzione monotona (crescente) continua da sinistra allora f può essere decomposta nella somma di una funzione φ monotona (crescente) e continua e di una funzione h di salto.



Usando un famoso risultato di F.Riesz si può dimostrare che

Teorema

Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione monotona, allora f è derivabile quasi ovunque su $[a, b]$. In altre parole è possibile trovare un insieme di misura nulla $N \subset [a, b]$ tale che f è derivabile per ogni $x \in [a, b] \setminus N$.

Corollario

Una funzione di salto ha quasi ovunque derivata nulla

Cominciamo ora a considerare lo studio della funzione integrale associata ad una funzione integrabile.

Definizione

Sia f una funzione integrabile su $[a, b]$ e definiamo

$$F(x) = \int_0^x f(t)dt$$

Se osserviamo che qualora $f \geq 0$ allora F è crescente, e se ricordiamo anche che

$$f = f_+ - f_-$$

è facile convincersi che F può essere espressa come differenza di funzioni crescenti.

Ad f quindi possono essere applicati tutti i risultati provati per la funzioni monotone ed in particolare si può affermare che

Teorema

Se f è una funzione integrabile su $[a, b]$ allora

$$\frac{d}{dx}F(x) = \frac{d}{dx} \int_0^x f(t)dt$$

esiste quasi ovunque su $[a, b]$.

Osserviamo esplicitamente che abbiamo fino a questo momento provato che F' esiste quasi ovunque e non ancora che $F'(x) = f(x)$. ciò sarà oggetto del lavoro che segue.

Definizione

Diciamo che una funzione f è di variazione limitata su $[a, b]$ se esiste una costante C tale che per ogni partizione dell'intervallo $[a, b]$

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

si abbia

$$\sum_{k=1}^n |f(x_k) - f(x_{k-1})| \leq C$$

Osserviamo che ogni funzione monotona è di variazione limitata in quanto la somma considerata è in tal caso sempre uguale a $|f(b) - f(a)|$,

Definizione

Sia f una funzione di variazione limitata su $[a, b]$, chiamiamo variazione totale di f su $[a, b]$

$$\mathcal{V}_a^b[f] = \sup \sum_{k=1}^n |f(x_k) - f(x_{k-1})|$$

l'estremo superiore essendo esteso a tutte le possibili partizioni di $[a, b]$

Se f è una funzione data su $\mathbb{R}$, diciamo che è di variazione limitata se

$$\mathcal{V}_{-\infty}^{+\infty}[f] = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \mathcal{V}_a^b[f]$$

è finita.

Teorema

Se f, g sono funzioni definite su $[a, b]$ e se $\alpha \in \mathbb{R}$ allora

①

$$\mathcal{V}_a^b[\alpha f] = |\alpha| \mathcal{V}_a^b[f]$$

②

$$\mathcal{V}_a^b[f + g] \leq \mathcal{V}_a^b[f] + \mathcal{V}_a^b[g]$$

③ se $a < b < c$ allora

$$\mathcal{V}_a^c[f] = \mathcal{V}_a^b[f] + \mathcal{V}_b^c[f]$$

Teorema

Sia f una funzione definita su $[a, b]$, allora la funzione

$$v(x) = \mathcal{V}_a^x[f] \tag{58}$$

è monotona crescente; inoltre se f è continua da sinistra o da destra in un punto x^* tale risulta anche v .

Definizione

Sia f una funzione definita su $[a, b]$ a variazione limitata allora possiamo considerare la funzione w definita da

$$w(x) = v(x) - f(x) \quad (59)$$

dove $v(x) = \mathcal{V}_a^x[f]$ è la variazione totale di f sull'intervallo $[a, x]$. È evidente che f può essere decomposta nella somma

$$f = v + w$$

Teorema

Se f è una funzione di variazione limitata su $[a, b]$ e se w è la funzione definita dalla 59, allora w è decrescente. Quindi f può essere rappresentata come differenza di funzioni monotone crescenti.

Corollario

Ogni funzione di variazione limitata ammette quasi ovunque derivata finita

Per le funzioni di variazione limitata è opportuno considerare una generalizzazione della nozione di funzione di salto.

Definizione

Sia $x_1 < x_2 < x_3 < \dots < x_n < \dots$ un insieme finito o numerabile di punti dell'intervallo $[a, b]$; e siano assegnati in corrispondenza di ogni x_n due numeri h_n e k_n tali che

$$\sum_n (|k_n| + |h_n|) < +\infty$$

Supponiamo che se $x_n = a$ allora $k_n = 0$ e che se $x_n = b$ allora $h_n = 0$. e chiamiamo funzione di salto la funzione σ definita da

$$\sigma(x) = \sum_{x_n \leq x} k_n + \sum_{x_n < x} h_n \quad (60)$$

Chiaramente la variazione totale della funzione σ è

$$\mathcal{V}_a^b[\sigma] = \sum_n (|k_n| + |h_n|) < +\infty$$

la funzione σ è continua in ogni punto tranne che per quegli x_n per i quali k_n oppure h_n risultano non nulli. In tal caso è facile convincersi che

$$k_n = \sigma(x_n) - \sigma(x_{n-}) \quad h_n = \sigma(x_{n+}) - \sigma(x_n)$$

Così come abbiamo fatto per le funzioni monotone possiamo anche per le funzioni a variazione limitata (che sono differenze di funzioni monotone) enunciare il seguente risultato.

Teorema

Se f è una funzione a variazione limitata tra $[a, b]$, allora f può essere decomposta in modo unico nella forma

$$f = \varphi + \sigma$$

dove σ è una funzione di salto e φ è una funzione continua.

Osserviamo infine che le funzioni costanti sono le uniche funzioni la cui variazione totale è nulla.

Osserviamo anche che se poniamo

$$\|f\| = \mathcal{V}_a^b[f]$$

definiamo sullo spazio delle funzioni di variazione limitata (che risulta spazio vettoriale) una quasi-norma.

Non risulta verificata che

$$\|f\| = 0 \iff f = 0$$

in quanto tutte le costanti hanno variazione totale nulla, tuttavia è possibile normare, per mezzo della variazione totale lo spazio delle funzioni di variazione limitata tali che $f(a) = 0$.

Lo spazio delle funzioni di variazione limitata può invece essere normato mediante la

$$\|f\| = |f(a)| + \mathcal{V}_a^b[f]$$

È già stato verificato che la funzione integrale

$$F(x) = \int_0^x f(t)dt$$

ammette quasi ovunque derivata finita. Non abbiamo tuttavia ancora provato che tale derivata coincide quasi ovunque con la funzione integranda. Questo sarà lo scopo del teorema che segue.

Teorema

Sia f una funzione sommabile su $[a, b]$ e definiamo

$$F(x) = \int_0^x f(t)dt$$

allora

$$f'(x) = f(x) \quad \text{quasi ovunque in } [a, b]$$

Abbiamo a questo punto stabilito che ogni funzione integrale è derivabile ed ha come derivata la sua integranda q meno di un insieme di misura nulla.

Vogliamo ora estendere al caso dell'integrale di Lebesgue la formula del teorema fondamentale del calcolo integrale di Newton-Leibnitz

$$f(b) - f(a) = \int_a^b f'(t) dt$$

Teorema

Sia f una funzione monotona crescente allora f' esiste quasi ovunque inoltre f' è sommabile e si ha

$$\int_a^b f'(x) dx \leq f(b) - f(a)$$

Definizione

Diciamo che una funzione f definita su un intervallo $[a, b]$ è assolutamente continua se comunque si scelga $\epsilon > 0$ esiste $\delta > 0$ tale che, comunque si scelga una famiglia finita di intervalli in $[a, b]$ a due a due disgiunti, che indicheremo con

$$\{(x_k, y_k) : k = 1, 2, \dots, n\}$$

tale che

$$\sum_{k=1}^n (y_k - x_k) < \delta$$

si ha

$$\sum_{k=1}^n [f(y_k) - f(x_k)] < \epsilon$$

È banale verificare che ogni funzione assolutamente continua è anche uniformemente continua, mentre si trovano esempi di funzioni uniformemente continue non assolutamente continue. Consideriamo ora le proprietà fondamentali delle funzioni assolutamente continue.

- ① f è assolutamente continua se e solo se comunque si scelga $\epsilon > 0$ esiste $\delta > 0$ tale che, comunque si scelga una famiglia **numerabile** di intervalli in $[a, b]$ a due a due disgiunti, che indicheremo con

$$\{(x_k, y_k) : k = 1, 2, \dots, n\} \quad \text{tale che} \quad \sum_{k=1}^{+\infty} (b_k - a_k) < \delta$$

si ha

$$\sum_{k=1}^{+\infty} [f(b_k) - f(a_k)] < \epsilon$$

- ② Ogni funzione assolutamente continua è anche di variazione limitata.
- ③ Se f, g sono assolutamente continue ed $\alpha \in \mathbb{R}$, allora anche $f + g$ ed αf sono assolutamente continue.
- ④ L'insieme delle funzioni assolutamente continue risulta un spazio vettoriale sottospazio dello spazio delle funzioni di variazione limitata.
- ⑤ se f è una funzione assolutamente continua allora

$$f = v - w$$

dove

$$v(x) = \mathcal{V}_a^x[f] \quad w(x) = f(x) - v(x)$$

e sia v che w risultano assolutamente continue e decrescenti.

Teorema

Sia f una funzione sommabile su $[a, b]$ allora la funzione

$$F(x) = \int_a^x f(t)dt$$

è una funzione assolutamente continua. Possiamo chiamare F integrale indefinito di f .

Lemma

Sia f una funzione monotona (crescente), assolutamente continua su $[a, b]$. Allora se $f'(x) = 0$ quasi ovunque si può affermare che f è costante.

Teorema

- **Lebesgue** - Sia f una funzione assolutamente continua su $[a, b]$, allora f è sommabile e si ha

$$f(x) - f(a) = \int_a^x f'(t) dt \quad x \in [a, b]$$

Possiamo a questo punto esaminare più precisamente al decomposizione di una funzione di variazione limitata.

Abbiamo già visto che ogni funzione di variazione limitata f può essere scritta come somma

$$f = \varphi + \sigma$$

dove σ è una funzione di salto e φ è una funzione continua (ovviamente a variazione limitata).

Ora poichè φ è a variazione limitata essa risulta derivabile quasi ovunque e possiamo considerare le funzioni

$$\psi(x) = \int_0^x \varphi'(t) dt \quad \chi(x) = \varphi(x) - \psi(x)$$

Teorema

Ogni funzione di variazione limitata f può essere decomposta nella somma di tre funzioni

$$f = \psi + \chi + \sigma \quad (61)$$

la prima assolutamente continua, la seconda continua e singolare, la terza di salto.

Ogni funzione che compare nella somma 61 risulta essere univocamente determinata a meno di una costante. possiamo normalizzare la rappresentazione rendendo univocamente determinate le tre componenti chiedendo che due di esse si annullino nell'origine.

Abbiamo finora considerato solo misure che siano positive, vogliamo ora prendere in esame anche il concetto di misura con segno. Sia X uno spazio di misura (finito) e sia μ una misura definita su una σ -algebra $\mathcal{S}_\mu$; sia f una funzione su X sommabile rispetto a μ .

Allora possiamo considerare una funzione Φ definita su $\mathcal{S}_\mu$ mediante la

$$\Phi(A) = \int_A f(x) d\mu$$

Si può verificare che la funzione Φ è una funzione σ -additiva sugli insiemi della σ -algebra $\mathcal{S}_\mu$.

Definizione

Chiamiamo misura con segno o più brevemente carica su X ogni funzione Φ definita su una σ -algebra $\mathcal{S}_\mu$ di sottoinsiemi di X che risulta σ -additiva.

Definizione

Sia Φ una carica su X , diciamo che E è negativo per Φ se comunque si consideri un insieme $F \in \mathcal{S}_\mu$ si ha

$$\Phi(E \cap F) \geq 0$$

È possibile dimostrare il seguente risultato

Teorema

Se Φ è una carica definita su X allora è possibile trovare due insiemi P ed N tali che

$$P = X \setminus N \quad P \text{ è negativo per } \Phi \quad N \text{ è positivo per } \Phi$$

Definizione

Sia X uno spazio di misura (finito) e sia μ una misura definita su una σ -algebra $\mathcal{S}_\mu$; Sia Φ una carica su X e siano P ed N due sottoinsiemi di X , rispettivamente positivo e negativo per Φ . Diciamo che la coppia di insiemi (P, N) costituisce una decomposizione di Hahn dello spazio X rispetto alla carica Φ .

La decomposizione di Hahn di uno spazio X non risulta univocamente determinata, tuttavia se (P, N) e (Q, M) sono due decomposizioni dello spazio X rispetto alla carica Φ , si ha, per ogni $E \in \mathcal{S}_\mu$

$$\Phi(E \cap P) = \Phi(E \cap Q) \quad \Phi(E \cap N) = \Phi(E \cap M) \quad (62)$$

Definizione

Se (P, N) è una decomposizione di Hahn dello spazio X , possiamo definire

$$\Phi_+(E) = \Phi(E \cap P) \quad \Phi_-(E) = -\Phi(E \cap N) \quad (63)$$

Φ_- , Φ_+ si dicono, rispettivamente, variazione inferiore e superiore della carica Φ e risulta evidente che

1

$$\Phi = \Phi_+ - \Phi_-$$

2 Φ_+ , Φ_- sono funzioni di insieme σ -additive e positive, cioè misure.

Inoltre

$$|\Phi| = \Phi_+ + \Phi_-$$

risulta essa pure una misura ed è detta variazione totale della carica Φ .

La coppia (Φ_+, Φ_-) è detta decomposizione di Jordan della carica Φ .

Proviamo infine il seguente

Teorema

-Radon - Nikodim Sia X uno spazio di misura su cui è definita una misura μ ed una σ -algebra $\mathcal{S}_\mu$. Sia Φ una carica assolutamente continua definita su $\mathcal{S}_\mu$, assolutamente continua rispetto a μ .

Allora esiste una funzione sommabile f definita su X tale che.

$$\Phi(A) = \int_A f(x) d\mu \quad \forall A \in \mathcal{S}_\mu$$

Definizione

Siano $f, g : [a, b] \longrightarrow \mathbb{R}$ due funzioni limitate; sia $P \in \mathcal{P}(a, b)$, $P = \{x_0, \dots, x_n\}$ e $\Xi = \{\xi_1, \dots, \xi_n\}$ una scelta di punti $x_{i-1} \leq \xi_i \leq x_i$ associata alla partizione P .

Definiamo la somma di Riemann – Stieltjes di f rispetto a g relativa alla partizione P ed alla scelta di punti Ξ .

$$R(f, g, P, \Xi) = \sum_{i=1}^n f(\xi_i)[g(x_i) - g(x_{i-1})]$$

f è integrabile secondo Riemann – Stieltjes in $[a, b]$ se esiste $I \in \mathbb{R}$ tale che $\forall \varepsilon > 0 \exists P_\varepsilon \in \mathcal{P}(a, b)$ tale che se $P < P_\varepsilon$ si ha

$$I = \int_a^b f(x) dg(x)$$

è l'integrale di Riemann-Stieltjes di f rispetto a g , f si dice integranda, g si dice integratrice.

Teorema

Siano $f, g : [a, b] \longrightarrow \mathbb{R}$ due funzioni limitate; allora f è integrabile rispetto a g se e solo se g è integrabile rispetto ad f . Si ha inoltre che

$$\int_a^b f(x)dg(x) + \int_a^b g(x)df(x) = f(b)g(b) - f(a)g(a).$$

Teorema

- integrazione per sostituzione - Supponiamo che $f, g : [a, b] \longrightarrow \mathbb{R}$ siano limitate e sia $\phi : [c, d] \longrightarrow [a, b]$ continua e strettamente crescente, $\phi(c) = a$, $\phi(d) = b$. Allora si ha che f è integrabile su $[a, b]$ rispetto a g se e solo se $f(\phi(\cdot))$ è integrabile rispetto a $g(\phi(\cdot))$. Inoltre

$$\int_a^b f(x)dg(x) = \int_c^d f(\phi(x))dg(\phi(x)).$$

Teorema

Siano $f, g : [a, b] \rightarrow \mathbb{R}$ limitate e supponiamo che $g \in C^1([a, b])$; allora, se f è integrabile rispetto a g , si ha che fg' è integrabile e

$$\int_a^b f(x) dg(x) = \int_a^b f(x)g'(x)dx.$$

Teorema

Siano $f, g : [a, b] \rightarrow \mathbb{R}$ limitate e sia g derivabile con g' limitata; supponiamo inoltre che f e g' siano integrabili su $[a, b]$; allora f è integrabile rispetto a g e si ha

$$\int_a^b f(x) dg(x) = \int_a^b f(x)g'(x)dx.$$

Teorema

Siano $f, h, g : [a, b] \rightarrow \mathbb{R}$ limitate e sia g monotona. Allora se f ed h sono integrabili rispetto a g , tale risulta anche $\alpha f + \beta h$, $\forall \alpha, \beta \in \mathbb{R}$ e si ha;

$$\int_a^b [\alpha f(x) + \beta h(x)] dg(x) = \alpha \int_a^b f(x) dg(x) + \beta \int_a^b h(x) dg(x).$$

Inoltre si ha che fh è integrabile rispetto a g .

Ricordando il teorema di integrazione per parti, si prova che:

Teorema

Siano $f, g, h : [a, b] \rightarrow \mathbb{R}$ limitate e siano g ed h monotone crescenti. Allora se f è integrabile rispetto a g e rispetto ad h si ha anche che f è integrabile rispetto ad $\alpha g + \beta h$ e

$$\int_a^b f(x) d[\alpha g(x) + \beta h(x)] = \alpha \int_a^b f(x) dg(x) + \beta \int_a^b f(x) dh(x).$$

Teorema

Siano $f, h, g : [a, b] \longrightarrow \mathbb{R}$ limitate e supponiamo che g sia monotona crescente; allora

- 1 Se f, h sono integrabili rispetto a g e se $f \geq h$ si ha

$$\int_a^b f(x)dg(x) \geq \int_a^b h(x)dg(x);$$

- 2 se definiamo $f_+ = \max\{f, 0\}$ ed $f_- = \min\{f, 0\}$, f è integrabile rispetto a g se e solo se f_+ ed f_- sono integrabili rispetto a g ;
- 3 f è integrabile rispetto a g se e solo se $|f|$ è integrabile rispetto a g e si ha

$$\left| \int_a^b f(x)dg(x) \right| \leq \int_a^b |f(x)|dg(x) ;$$

- 4 se f è integrabile rispetto a g , se $f \geq 0$ e se $[\alpha, \beta] \subset [a, b]$

$$0 \leq \int_{\alpha}^{\beta} f(x)dg(x) \leq \int_a^b f(x)dg(x)$$

- 5 se f è continua, g è strettamente crescente, $f \geq 0$ e $\int_a^b f(x)dg(x) = 0$ allora $f \equiv 0$.

Teorema

- Siano $f, g: [a, b] \longrightarrow \mathbb{R}$, f continua, g monotona crescente; allora f è integrabile rispetto a g e g è integrabile rispetto ad f .

Teorema

Consideriamo $f, g: [a, b] \longrightarrow \mathbb{R}$, f continua e g crescente, allora esiste $c \in [a, b]$ tale che

$$\int_a^b f(x) dg(x) = f(c) \int_a^b dg(x) = f(c)(g(b) - g(a))$$

Teorema

Siano $f, g: [a, b] \longrightarrow \mathbb{R}$, f continua, g crescente; supponiamo che g sia derivabile in $x_0 \in [a, b]$, allora

$$F(x) = \int_a^x f(t) dg(t)$$

è derivabile in x_0 e

$$F'(x_0) = f(x_0)g'(x_0).$$

Teorema

Consideriamo $f, g: [a, b] \longrightarrow \mathbb{R}$, f crescente, g continua; allora esiste $c \in [a, b]$ tale che

$$\int_a^b f(x) dg(x) = f(a) \int_a^c dg(x) + f(b) \int_c^b dg(x).$$

Teorema

Siano $f, h : [a, b] \longrightarrow \mathbb{R}$ continue, $h \geq 0$. Allora $\exists c \in [a, b]$:

$$\int_a^b f(x)h(x)dx = f(c) \int_a^b h(x)dx.$$

Teorema

Siano $f, h : [a, b] \longrightarrow \mathbb{R}$, f crescente, h continua; Allora $\exists c \in [a, b]$:

$$\int_a^b f(x)h(x)dx = f(a) \int_a^c h(x)dx + f(b) \int_c^b h(x)dx.$$

Teorema

Siano $f, h : [a, b] \longrightarrow \mathbb{R}$, f crescente $f \geq 0$, h continua, Allora $\exists c \in [a, b]$:

$$\int_a^b f(x)dh(x) = f(b) \int_c^b h(x)dx.$$

Consideriamo ora il problema di definire l'integrale di Riemann-Stieltjes anche su sottoinsiemi non limitati di $\mathbb{R}$.

Definizione

Siano $f, g : [a, +\infty) \longrightarrow \mathbb{R}$ e supponiamo che f sia integrabile rispetto a g in $[a, \delta] \forall \delta > a$.

Qualora

$$\lim_{\delta \rightarrow +\infty} \int_a^{\delta} f(x) dg(x)$$

esista, diciamo che f è integrabile rispetto a g su $[a, +\infty)$ e definiamo

$$\int_a^{+\infty} f(x) dg(x) = \lim_{\delta \rightarrow +\infty} \int_a^{\delta} f(x) dg(x).$$

Teorema

Siano $f, g : [a, +\infty) \longrightarrow \mathbb{R}$ e supponiamo che f sia integrabile rispetto a g in $[a, \delta] \forall \delta > a$; allora f è integrabile rispetto a g in $[a, +\infty)$ se e solo se
 $\forall \varepsilon > 0 \exists \delta_\varepsilon > 0$ tale che, se $x', x'' > \delta_\varepsilon$

$$\left| \int_{x'}^{x''} f(t) dg(t) \right| < \varepsilon$$

Corollario

Siano $f, g : [a, +\infty) \longrightarrow \mathbb{R}$ e supponiamo che g sia crescente. Allora, se $|f|$ è integrabile rispetto a g , anche f è integrabile rispetto a g su $[a, +\infty)$.

Teorema

Siano $f, g: [a, +\infty) \longrightarrow \mathbb{R}$, f limitata e g crescente e limitata su $\mathbb{R}$. Allora f è integrabile rispetto a g su $[a, +\infty)$.

Teorema

Siano $g: [a, b] \longrightarrow \mathbb{R}$ limitata e crescente, e sia $h: [a, b] \times I \longrightarrow \mathbb{R}$ continua, $I \subset \mathbb{R}$; allora

$$F(s) = \int_a^b h(t, s) dg(t)$$

è continua in I .

Se inoltre h è derivabile rispetto ad s e $\partial h(t, s)/\partial s$ è continua in $[a, b] \times I$, allora F è derivabile in I e si ha

$$F'(s) = \int_a^b \frac{\partial h}{\partial s}(t, s) dg(t).$$

Teorema

L'equazione di Black-Scholes

Problema

Stima del prezzo f di un'opzione di acquisto o di vendita di un bene il cui valore S cresce nel tempo con tasso μ ed è affetto da un fattore casuale gaussiano con media 0 e varianza σ con riferimento ad un tasso di investimento senza rischio r .

Il possesso di una opzione di acquisto o di vendita conferisce il diritto di acquistare o vendere una unita' del bene S al tempo stabilito T ad un prezzo stabilito K .

Il prezzo di una opzione deve essere tale che sia possibile trovare una combinazione di quote del bene e quote di investimento privo di rischio che garantiscano un rendimento pari a quello privo di rischio.

Il valore del bene S cresce seguendo la legge

$$dS = \mu S dt + \sigma S dW$$

Il valore dell'investimento privo di rischio B cresce seguendo la legge

$$dB = rB dt$$

Il prezzo dell'opzione è funzione del tempo t e del prezzo S del bene sottostante

$$f(t, S(t))$$

I è un portafoglio composto da beni S , opzioni f ed investimenti B a tasso r privo di rischio

$$I = -f + \alpha S + \beta B$$

In ipotesi di autofinanziamento ed usando la formula di Itô si ottiene

$$dI = \left((-f_t + \mu S f_S + \frac{1}{2} \sigma^2 S^2 f_{SS}) + (\alpha \mu S + \beta r B) \right) dt + (-f_S + \alpha) \sigma S dW$$

Per eliminare il rischio si richiede che

$$-f_S + \alpha = 0 \implies \alpha = f_S$$

Supponendo impossibile l'arbitraggio

$$dI = rI dt = r(-f + \alpha S + \beta B) dt = r(-f + f_S S + \beta B) dt$$

Combinando le due espressioni di dI si ottiene l'equazione di Black-Scholes

$$f_t + rSf_S + \frac{1}{2}\sigma^2 S^2 f_{SS} = rf \quad t < T$$

alla quale si associa (nel caso di una opzione di vendita, tipo PUT) la condizione

$$f(T, S) = \max(K - S, 0)$$

che si può usare per determinare $f(t, S(t))$.

Per le opzioni di tipo CALL di acquisto si usa la condizione

$$f(T, S) = \max(S - K, 0)$$

Variabili aleatorie e Processi Stocastici

Sia

$$(\Omega, \mathcal{F}, P)$$

uno spazio di probabilità.

- Ω , spazio dei campioni;
- $\mathcal{F}$, σ –algebra degli eventi in Ω
- P misura di probabilità su $(\Omega, \mathcal{F})$

Sia

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

una variabile aleatoria sia cioè X una funzione misurabile in $(\Omega, \mathcal{F}, P)$

X definisce su $\mathbb{R}$ una misura di probabilità ξ mediante la

$$\begin{aligned}\xi((\alpha, \beta)) &= P(\alpha \leq X \leq \beta) = \\ &= P(\{\omega \in \Omega : \alpha \leq X(\omega) \leq \beta\}) = \\ &= P(X^{-1}((\alpha, \beta)))\end{aligned}$$

si può verificare, usando il teorema di Radon-Nikodym che, sotto ipotesi non restrittive, esiste una funzione $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$P(\alpha \leq X \leq \beta) = \xi((\alpha, \beta)) = \int_{\alpha}^{\beta} \varphi(s) ds$$

φ è la densità di probabilità di X

Abbiamo anche che

$$\begin{aligned} E(X) &= \int_{\Omega} X(\omega) d\omega \approx \sum \alpha P(X^{-1}(\alpha, \beta)) \approx \\ &\approx \sum \int_{\alpha}^{\beta} s \varphi(s) ds = \int_{-\infty}^{+\infty} s \varphi(s) ds \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_{\Omega} X(\omega) d\omega \approx \sum \alpha^2 P(X^{-1}(\alpha, \beta)) \approx \\ &\approx \sum \int_{\alpha}^{\beta} s^2 \varphi(s) ds = \int_{-\infty}^{+\infty} s^2 \varphi(s) ds \end{aligned}$$

Chiamiamo Processo Stocastico una funzione

$$\mathbb{R} \ni t \mapsto X(t)$$

dove $X(t)$ è una variabile aleatoria su Ω

Avremo

$$X = X(t, \omega) \quad (t, \omega) \in \mathbb{R} \times \Omega$$

$$\mathbb{R} \times \Omega \ni (t, \omega) \mapsto X(t, \omega)$$

Per ogni t fissato $X(t, \omega)$ ha una densità di probabilità $\varphi(t, s)$

$$P(\alpha \leq X(t) \leq \beta) = \int_{\alpha}^{\beta} \varphi(t, s) ds$$

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

$$E(X^2(t)) = \int_{-\infty}^{+\infty} s^2 \varphi(t, s) ds = \sigma^2(t)$$

Il Processo di Wiener

È un processo stocastico W con le seguenti caratteristiche:

- $W(0) = 0$
- $W(t) - W(s)$ ha una densità di probabilità Gaussiana normale di media 0 e varianza $(t - s)$, $N(0, \sqrt{t - s})$.
- W ha incrementi indipendenti: cioè $W(t_1) - W(s_1)$ e $W(t_2) - W(s_2)$ sono variabili aleatorie indipendenti.

È possibile costruire un processo stocastico con le caratteristiche indicate.

Si ha

$$P(\alpha \leq W(t) - W(s) \leq \beta) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{\alpha}^{\beta} e^{-\frac{1}{2(t-s)}\tau^2} d\tau$$

La funzione $t \mapsto W(t, \omega)$ è continua per quasi ogni $\omega \in \Omega$ e non è derivabile con probabilità 1.

Modelli di Crescita

Sia x una quantità scalare che cresce nel tempo con tasso costante a

Allora

$$\frac{x(t+h) - x(t)}{h} = a$$

La precedente uguaglianza può essere espressa in diversi modi

$$\dot{x}(t) = a \quad , \quad \frac{dx}{dt} = a \quad dx = a dt \quad , \quad \int_0^t dx = \int_0^t a ds$$

$$x(t+h) = x(t) + \int_t^{t+h} a ds = x(t) + ah$$

Se x cresce con tasso proporzionale alla sua consistenza

Allora

$$\frac{x(t) - x(s)}{t - s} = \mu x(t)$$

La precedente uguaglianza può essere espressa da

$$\dot{x}(t) = \mu x(t), \quad \frac{dx}{dt} = \mu x, \quad dx = \mu x dt, \quad \int_0^t dx = \int_0^t \mu x(s) ds$$

Non si può procedere esplicitamente con l'integrazione,
Si può usare la definizione di integrale per ottenere una formula discreta che approssima x (Metodo di Eulero).

$$x(t + h) = x(t) + \int_t^{t+h} x(\bar{s}) ds = x(t) + x(\bar{s})h$$

Crescita con tasso costante alterata da un fattore di incertezza W .

Avremo

$$x(t) + \sigma W(t) - (x(s) + \sigma W(s)) = \mu(t - s)$$

da cui

$$x(t) - x(s) = \mu(t - s) + \sigma(W(t) - W(s))$$

e

$$dx = \mu dt + \sigma dW$$

Possiamo integrare numericamente

$$\int_t^{t+h} dx = \int_t^{t+h} \mu dt + \int_t^{t+h} \sigma dW$$

se siamo in grado di dare una definizione di

$$\int_t^{t+h} \sigma dW$$

Nel caso in cui σ sia costante possiamo usare l'idea di integrale di Riemann:

$$\int_0^t \sigma dW \approx \sum_{i=1}^{n-1} \sigma(W(t_{i+1}) - W(t_i)) = \sigma(W(t) - W(0)) = \sigma W(t)$$

Osserviamo che

$$E \left(\sum_{i=1}^{n-1} \sigma(W(t_{i+1}) - W(t_i)) \right) = 0$$

$$\text{Var} \left(\sum_{i=1}^{n-1} \sigma(W(t_{i+1}) - W(t_i)) \right) = \sum_{i=1}^{n-1} \sigma^2(t_{i+1} - t_i) = \sigma^2(t - 0) = \sigma^2 t$$

Ne viene che

$$\int_0^t \sigma dW$$

è una variabile aleatoria di media 0 e di varianza $\sigma^2 t$

Nel caso in cui σ sia funzione della sola t possiamo approssimare

$$\int_0^t \sigma(s) dW(s)$$

mediante le somme di Riemann - Stieltjes

$$\sum_{i=1}^{n-1} \sigma(s_i)(W(t_{i+1}) - W(t_i))$$

Paley e Wiener hanno dimostrato che le somme di Riemann, in questo caso, convergono ad un processo stocastico nel caso in cui

$$\sigma \in \mathcal{C}^1$$

Non è difficile estendere la definizione a funzioni della sola t meno regolari.

Il caso in cui $\sigma = \sigma(t, W)$ è necessario per poter considerare equazioni del tipo

$$dX = \mu(t, X(t))dt + \sigma(t, X(t))dW$$

Per questo scopo è necessario definire

$$\int_0^t \sigma(s, X(s))dW$$

dove X è un processo stocastico.

Calcoliamo ad esempio il più semplice di questi integrali:

$$\int_0^t WdW$$

Procedendo come per gli integrali di Riemann-Stieltjes
 Consideriamo una partizione dell'intervallo $[0, t]$

$$0 = t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = t$$

ed una scelta di punti s_i

$$t_i \leq s_i \leq t_{i-1}$$

Le corrispondenti somme di Riemann Stieltjes sono

$$RS_n = \sum_{i=1}^{n-1} W(s_i)(W(t_i) - W(t_{i-1}))$$

Si ha

$$\begin{aligned}
 & W(s_i)(W(t_i) - W(t_{i-1})) = \\
 = & W(t_{i-1})(W(t_i) - W(t_{i-1})) + (W(s_i) - W(t_{i-1}))(W(t_i) - W(t_{i-1})) = \\
 & \frac{1}{2} W(t_{i-1})(W(t_i) - W(t_{i-1})) + \frac{1}{2} W(t_{i-1})(W(t_i) - W(t_{i-1})) + \\
 & \quad + (W(s_i) - W(t_{i-1}))(W(t_i) - W(t_{i-1})) = \\
 & \quad \frac{1}{2} (W(t_i) + W(t_{i-1}))(W(t_i) - W(t_{i-1})) - \\
 & \quad - \frac{1}{2} (W(t_i) - W(t_{i-1}))(W(t_i) - W(t_{i-1})) + \\
 & \quad \quad + (W(s_i) - W(t_{i-1}))(W(t_i) - W(t_{i-1}))
 \end{aligned}$$

Inoltre

$$\begin{aligned}(W(s_i) - W(t_{i-1}))(W(t_i) - W(t_{i-1})) &= \\ &= (W(s_i) - W(t_{i-1}))(W(t_i) - W(s_i)) + \\ &\quad + (W(s_i) - W(t_{i-1}))(W(s_i) - W(t_{i-1}))\end{aligned}$$

per cui

$$\begin{aligned}W(s_i)(W(t_i) - W(t_{i-1})) &= \\ &= \frac{1}{2}((W^2(t_i) - W^2(t_{i-1})) - (W(t_i) - W(t_{i-1}))^2) + \\ &\quad + (W(s_i) - W(t_{i-1}))(W(t_i) - W(s_i)) + (W(s_i) - W(t_{i-1}))(W(s_i) - W(t_{i-1}))\end{aligned}$$

Ne deduciamo che

$$\begin{aligned}
 RS_n &= \sum_{i=1}^n W(s_i)(W(t_i) - W(t_{i-1})) = \\
 &\frac{1}{2} (W^2(t) - W^2(0)) - \frac{1}{2} \sum_{i=1}^n (W(t_i) - W(t_{i-1}))^2 + \\
 &\quad + \sum_{i=1}^n (W(s_i) - W(t_{i-1}))^2 + \\
 &\quad + \sum_{i=1}^n (W(s_i) - W(t_{i-1}))(W(t_i) - W(s_i))
 \end{aligned}$$

Se $s_i = (1 - \theta)t_{i-1} + \theta t_i$ si ha

$$E \left(\frac{1}{2} \sum_{i=1}^n (W(t_i) - W(t_{i-1}))^2 \right) = \sum_{i=1}^n (t_i - t_{i-1}) = \frac{1}{2}(t - 0)$$

$$E \left(\sum_{i=1}^n (W(s_i) - W(t_{i-1}))^2 \right) = \sum_{i=1}^n (s_i - t_{i-1}) = \theta(t - 0)$$

$$E \left(\sum_{i=1}^n (W(s_i) - W(t_{i-1}))(W(t_i) - W(s_i)) \right) = 0$$

Si dimostra che

$$RS_n \rightarrow \frac{1}{2} (W^2(t) - W^2(t_0)) + \left(\theta - \frac{1}{2} \right) (t - 0)$$

Il limite dipende dalla scelta di punti s_i attraverso θ .

Stratonovich propose di scegliere $\theta = \frac{1}{2}$: continua a valere la formula di integrazione per parti.

Itô propose di scegliere $\theta = 0$:

- non vale l'integrazione per parti
- ma la variabile aleatoria

$$X(t) = \int_0^t W dW$$

che si ottiene è non anticipativa.

cioè dipende solo dagli avvenimenti passati.

Con metodi simili si prova che

$$\sum_{i=1}^n (W(t_i) - W(t_{i-1}))^2 \rightarrow (t - 0) = t$$

più precisamente

$$\sum_{i=1}^n (W(t_i) - W(t_{i-1}))^2$$

tende ad una variabile aleatoria di **media t** e di **varianza 0** .
Si esprime questo fatto dicendo che

$$\int_{t_0}^t (dW)^2 = \int_{t_0}^t dt$$

ed introducendo la regola formale

$$(dW)^2 = dt$$

Analogamente da

$$E \left(\sum_{i=1}^n (W(t_i) - W(t_{i-1}))(t_i - t_{i-1}) \right) = 0$$

$$E \left(\sum_{i=1}^n (t_i - t_{i-1})(W(t_i) - W(t_{i-1})) \right) = 0$$

si conclude che $\sum_{i=1}^n (W(t_i) - W(t_{i-1}))(t_i - t_{i-1})$ e $\sum_{i=1}^n (t_i - t_{i-1})(W(t_i) - W(t_{i-1}))$ tendono a 0 per cui

$$\int_{t_0}^t dt dW = \int_{t_0}^t dW dt = 0$$

si prova che è ragionevole introdurre anche le seguenti regole formali

$$dt dW = dW dt = 0$$

La Formula di Itô

Sia

X un processo stocastico tale che

$$dX = a(t, X(t))dt + b(t, X(t))dW(t)$$

$$g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$$

Consideriamo

$$y(t) = g(t, X(t))$$

e studiamo il problema di derivare la funzione ottenuta

Itô ha dimostrato che

$$\begin{aligned} dy(t) &= \\ &= \left(g_t(t, X(t)) + a(t, X(t))g_x(t, X(t)) + \frac{b^2(t, X(t))}{2}g_{xx}(t, X(t)) \right) dt + \\ &\quad + b(t, X(t))g_x(t, X(t))dW(t) \end{aligned}$$

La formula si dimostra in forma integrale:

$$\begin{aligned} y(t) - y(t_0) &= g(t, X(t)) - g(t_0, X(t_0)) = \\ &= \int_{t_0}^t g_t(t, X(t)) + a(t, X(t))g_x(t, X(t)) + \frac{b^2(t, X(t))}{2}g_{xx}(t, X(t)) dt + \\ &\quad + \int_{t_0}^t b(t, X(t))g_x(t, X(t))dW(t) \end{aligned}$$

Si può ricavare formalmente usando la Formula di Taylor e le regole di moltiplicazione di dt e dW

Equazione di crescita per il prezzo di un bene

Il prezzo $S(t)$ di un bene si può descrivere mediante l'equazione

$$dS(t) = rS(t)dt + \sigma S(t)dW(t)$$

$S(t) = S(t, \omega)$ è un processo stocastico

Se consideriamo

$$g(t) = \ln(S(t))$$

usando la formula di Itô si ottiene

$$dg(t) = d \ln(S(t)) = \frac{dS(t)}{S(t)} - \frac{1}{2} \left(\frac{\sigma^2 S^2(t)}{S^2(t)} \right) dt = rdt - \frac{\sigma^2}{2} dt + \sigma dW(t)$$

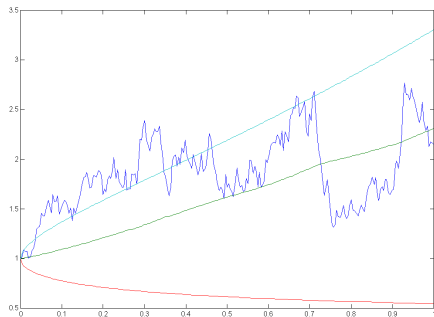
Si ha

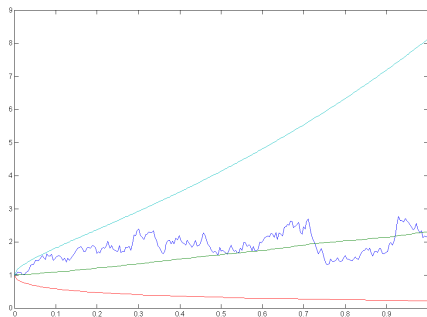
$$g(t) - g(0) = \ln \left(\frac{S(t)}{S(0)} \right) = rt - \frac{t\sigma^2}{2} + \sigma W(t)$$

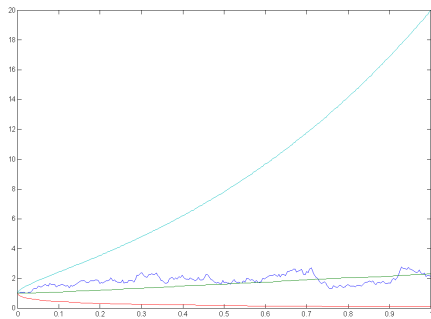
e di conseguenza

$$S(t) = S(0)e^{rt - \frac{t\sigma^2}{2} + \sigma W(t)}$$

Le figure che seguono riportano i grafici di una soluzione, il grafico della soluzione senza rumore e le bande di confidenza corrispondenti a σ , 2σ e 3σ .







L'equazione di Black-Scholes

Sia

$$f(t, S(t))$$

il prezzo dell'opzione di vendita di un bene S e sia B un investimento privo di rischio con rendimento r ;
consideriamo un portafoglio

$$I = -f + \alpha S + \beta B$$

La variazione di I può essere calcolata mediante la

$$dI = -df + \alpha dS + \beta dB$$

df può essere calcolato mediante la formula di Itô.
per cui

$$\begin{aligned} dI &= (-f_t + \mu S f_S + \frac{1}{2} \sigma^2 S^2 f_{SS}) dt + f_S \sigma dW + \alpha (\mu S dt + \sigma S dW) + \beta r B = \\ &= \left((-f_t + \mu S f_S + \frac{1}{2} \sigma^2 S^2 f_{SS}) + (\alpha \mu S + \beta r B) \right) dt + (-f_S + \alpha) \sigma S dW \end{aligned}$$

Per eliminare la componente di incertezza occorre imporre che

$$\alpha = f_S$$

ed in tal caso

$$\begin{aligned} dI &= \left((-f_t + \mu S f_S + \frac{1}{2} \sigma^2 S^2 f_{SS}) + \right. \\ &\quad \left. + (f_S \mu S + \beta r B) \right) dt + (-f_S + f_S) \sigma S dW = \\ &= \left((-f_t + \mu S f_S + \frac{1}{2} \sigma^2 S^2 f_{SS}) + \right. \\ &\quad \left. + (f_S \mu S + \beta r B) \right) dt = \end{aligned}$$

In condizioni di non arbitraggio si ha

$$dI = rIdt$$

per cui

$$dI = r(-f + \alpha S + \beta B)dt = r(-f + f_S S + \beta B)dt$$

e

$$\begin{aligned} r(-f + \alpha S + \beta B)dt &= r(-f + f_S S + \beta B)dt = dI = \\ &= rI = r(f + \alpha S + \beta B)dt = r(f - f_S S - \beta B)dt \end{aligned}$$

Si ottiene infine che

$$f_t + rsf_s + \frac{1}{2}\sigma^2 S^2 f_{SS} = rf \quad t < T$$

con la condizione

$$f(s, T) = \max(K - s, 0)$$

Il seguente è un programma scritto in Matlab che integra l'equazione stocastica

$$\begin{cases} dX = \lambda * Xdt + \sigma * (X)dW \\ X(0) = Xzero. \end{cases}$$


```
clf
randn('state',1)
T = 1; N = 28; Delta = T/N;
lambda = 0.05; sigma = 0.8; Xzero = 1;
Xem = zeros(1,N+1);
Xem(1) = Xzero;
for j = 1:N
Winc = sqrt(Delta)*randn;
Xem(j+1) = Xem(j) + Delta*lambda*Xem(j) +
sigma*Xem(j)*Winc;
end
plot([0:Delta:T],Xem,'r-')
```

Questo è il grafico che si ottiene

