

LABORATORIO DI MECCANICA E TERMODINAMICA

A.A. 2025-2026

ANDREA BRESSAN

ESPERIENZE DI MECCANICA

Misura della costante elastica di una molla elicooidale

Deformazioni elastiche

Una deformazione è detta **elastica** se il corpo torna allo **stato originario** quando vengono meno **le forze che ne hanno causato la deformazione**.

La deformazione è **elastica** se le forze applicate sono inferiori ad un limite che dipende dal materiale, dalla temperatura, dal tipo di deformazione considerata, etc.

Per le deformazioni elastiche si osserva una relazione di **proporzionalità tra sollecitazione e deformazione**.

$$\Delta L \propto F$$

Questo comportamento è noto come **legge di Hooke**.

La **legge di Hooke** è valida per la maggior parte dei minerali, per il vetro, per i materiali ceramici e per i metalli. Per i metalli duttili è vera per carichi modesti.

Nell'esperienza si studierà la deformazione di una molla elicoidale. Useremo nel fare ciò sia:

- Un metodo statico
- Un metodo dinamico

Deformazioni elastiche

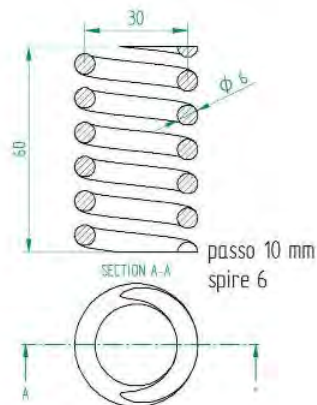
Quando all'estremità libera di una molla è applicata una forza, la molla esercita una forza di richiamo proporzionale allo spostamento di tale estremo rispetto alla posizione di equilibrio.

Il modulo della forza di richiamo è proporzionale alla deformazione (*legge di Hooke*).

$$\vec{F} = -k\vec{x}$$

Dove \vec{x} è la deformazione della molla rispetto la sua lunghezza normale $\vec{x} = (\ell - \ell_0)\hat{x}$ con ℓ_0 la lunghezza della molla a riposo. La legge di Hooke vale per piccole deformazioni fino al limite di elasticità della molla, oltre il quale le deformazioni divengono anelastiche, ovvero permanenti.

La costante di proporzionalità k è detta *costante elastica* della molla e dipende dal materiale di cui è costituita, dal diametro del filo e della molla e dal numero di spire



Molle ad elica 5 N/m

Dati prodotto

Weight	0.091 kg
Brand	3B Scientific
Art. No.	U8401010 [1000741]



Descrizione

Per prove di allungamento e oscillazione, dotate di due occhielli di sospensione.

Indice di rigidità: 5 N/m
 Max. capacità di carico: 2 N
 Lunghezza: 60 mm
 Diametro: 20 mm

$$T_{max} \sim 6 \sqrt{\frac{0.25}{5}} = 3s$$

Molle ad elica 1,5 N/m

Dati prodotto

Weight	0.03 kg
Brand	3B Scientific
Art. No.	U8405840 [1000786]



Descrizione

Per prove di allungamento e oscillazione, dotate di due occhielli di sospensione.

Indice di rigidità: 1,5 N/m
 Max. capacità di carico: 0,75 N
 Lunghezza: 120 mm
 Diametro: 20 mm

$$T_{max} \sim 6 \sqrt{\frac{0.8}{1.5}} = 4s$$

Molle ad elica 16 N/m

Dati prodotto

Weight	0.005 kg
Art. No.	U11025 [1002702]



Descrizione

Per prove di allungamento e oscillazione, dotate di due occhielli di sospensione.

Indice di rigidezza: 16 N/m
 Max. capacità di carico: 2,5 N
 Lunghezza: 115 mm
 Diametro: 6 mm

$$T_{max} \sim 6 \sqrt{\frac{0.25}{16}} = 0.8s$$

Molle ad elica 3,9 N/m

Dati prodotto

Weight	0.05 kg
Art. No.	U15027 [1002945]



Descrizione

Per prove di allungamento e oscillazione, dotate di due occhielli di sospensione.

Indice di rigidezza: 3,9 N/m
 Lunghezza: 30 mm
 Diametro: 34 mm

Misura della costante elastica – Metodo statico

La costante elastica di una molla può essere determinata sperimentalmente misurando le elongazioni $\Delta\ell$ al variare delle forze applicate.

Se all'estremità inferiore di una molla sospesa verticalmente è appesa una massa m , la forza applicata coincide con la forza peso della massa e la *legge di Hooke* si riscrive nel seguente modo:

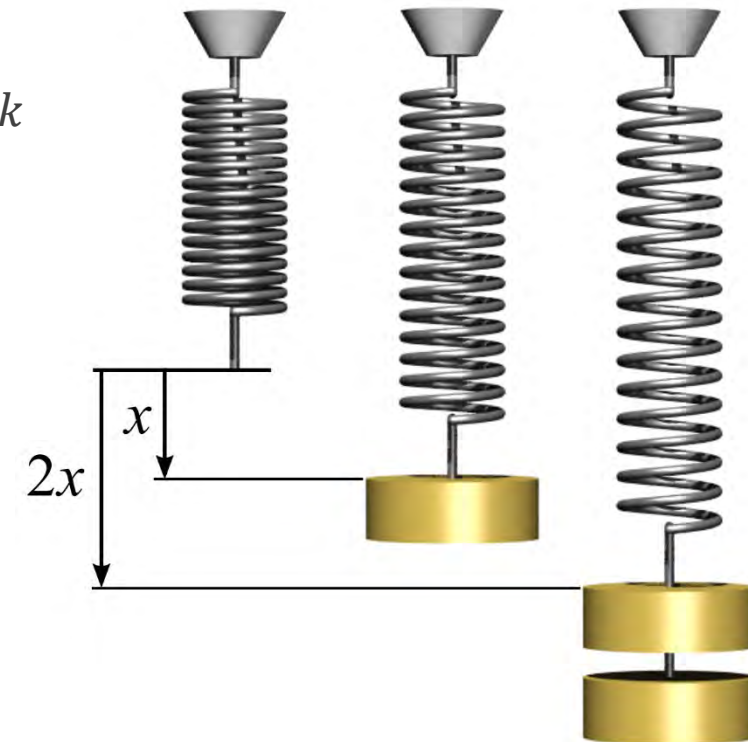
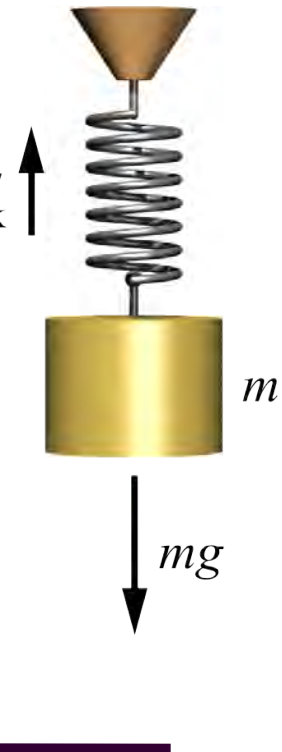
$$mg = k\Delta\ell = k(\ell - \ell_0)$$

Usando diverse masse e misurando i rispettivi $\Delta\ell$ ottengo diverse misure di k

$$k_i = \frac{m_i g}{\Delta\ell_i}$$

Nel caso reale alla molla è già appeso un piattello (di una certa massa m_p), inoltre la molla ha una sua massa ed è quindi anch'essa soggetta alla forza peso. Subisce già un allungamento ℓ_s anche a piattello (per le masse) scarico.

$$(m_p + m_{\text{eff. statica}})g = k(\ell_s - \ell_0)$$



Misura della costante elastica – Metodo statico

Misurando la differenza tra l'allungamento a piatto scarico

$$(m_p + m_{\text{eff. statica}})g = k(\ell_s - \ell_0)$$

e l'allungamento in seguito all'aggiunta di una massa m_i

$$(m_i + m_p + m_{\text{eff. statica}})g = k(\ell_i - \ell_0)$$

E facendo la differenza membro a membro tra le due equazioni otteniamo:

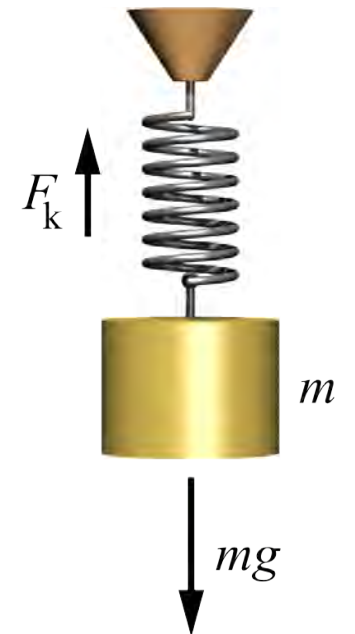
$$m_i g = k_i (\ell_i - \ell_s)$$

Da cui il valore della costante elastica della molla:

$$k_i = \frac{m_i g}{(\ell_i - \ell_s)}$$

Gli errori sono tutti errori di sensibilità per cui possiamo scrivere:

$$\frac{\Delta k_i}{k_i} = \frac{\Delta m_i}{m_i} + 2 \frac{\Delta \ell}{(\ell_i - \ell_s)}$$



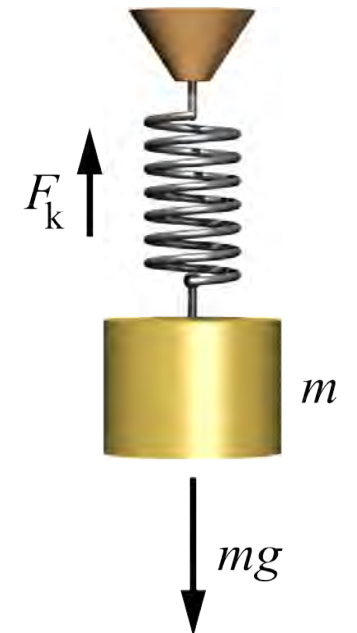
Misura della costante elastica – Metodo statico

Usando due molle con costanti elastiche diverse:

- Misurare con 5-6 valori di massa diversi (senza superare i limiti di carico delle molle)
- Fare un grafico di ℓ_i in funzione di m_i
- Fare un grafico di k_i in funzione di m_i
- Scegliere il valore più preciso di k_i (quale sarà ?)

Qualche commento/domanda:

- Gli errori da misura a misura sono indipendenti?
- Alcune molle potrebbero essere pretensionate (spire a contatto, lunghezza a riposo non raggiunta) e quindi non abbiamo allungamenti se non a partire da una massa m_x
- Alcune molle (con costanti elastiche piccole) si allungano vistosamente per il solo effetto della propria massa, con le spire vicino al punto di aggancio all'asta visibilmente più discostate di quelle vicine al punto di aggancio del piattello.



Quanto vale $m_{\text{eff. statica}}$

Anche se nella misura statica $m_{\text{eff. statica}}$ si cancella prendendo la differenza tra posizione di equilibrio iniziale e finale, è comunque istruttivo cercare un modello per calcolare quale sia la massa equivalente che agisce sulla molla dovuta al suo stesso peso.

Immaginiamo di dividere la molla in N parti di uguale lunghezza. La lunghezza di tali sezioni della molla sarà quindi $a = \ell_0/N$ e la massa pari a $\mu = m_m/N$. La costante elastica della molla completa è k , e quindi, considerando tale molla come formate da N molle in serie, la costante elastica κ di una ognuna di queste parti della molla sarà data da:

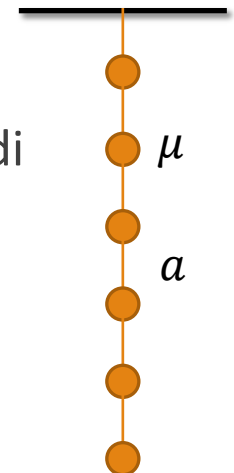
$$k = \frac{1}{\sum_{i=1}^N \frac{1}{\kappa}} = \frac{\kappa}{N}$$

Il primo elemento della molla sentirà il peso di tutte le $N - 1$ molle e si allungherà di

$$\Delta x_1 = \frac{(N - 1)\mu g}{\kappa}$$

La seconda sentirà il peso delle $N - 2$ molle e si allungherà di

$$\Delta x_2 = \frac{(N - 2)\mu g}{\kappa}$$



Quanto vale $m_{\text{eff. statica}}$

L'allungamento totale sarà dato dalla somma di tutti gli allungamenti:

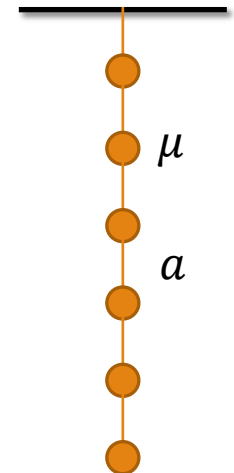
$$\Delta x = \sum_{i=1}^{N-1} \Delta x_i = \sum_{i=1}^{N-1} \frac{(N-i)\mu g}{\kappa}$$

$$\Delta x = \frac{\mu g}{\kappa} \sum_{i=1}^{N-1} (N-i) = \frac{\mu g}{\kappa} \left(N^2 - \frac{N(N-1)}{2} \right)$$

$$\Delta x = \frac{\mu N(N-1)g}{\kappa} \cong \frac{\mu g N^2}{\kappa} = \frac{\mu N g N}{\kappa} = \frac{m_m g}{2 \kappa}$$

Ovvero la molla si è allungata come se fosse stata appesa una massa:

$$m_{\text{eff. statica}} = \frac{m_m}{2}$$



Quanto vale $m_{\text{eff. statica}}$

Possiamo fare anche il conto passando al continuo. Definiamo la molla come un mezzo elastico di sezione S e lunghezza ℓ_0 . La sua densità è quindi $\rho = m_m/S\ell_0$. Introduciamo il modulo di Young della molla:

$$E = \frac{k\ell_0}{S} = \frac{k\ell_0}{m_m/\rho\ell_0} = \frac{k\rho\ell_0^2}{m_m}$$

misurato in $(N m^{-2})$, ovvero in Pascal. La legge di Hooke scritta usando il modulo di Young diventa

$$\frac{F}{S} = E \frac{\Delta\ell}{\ell}$$

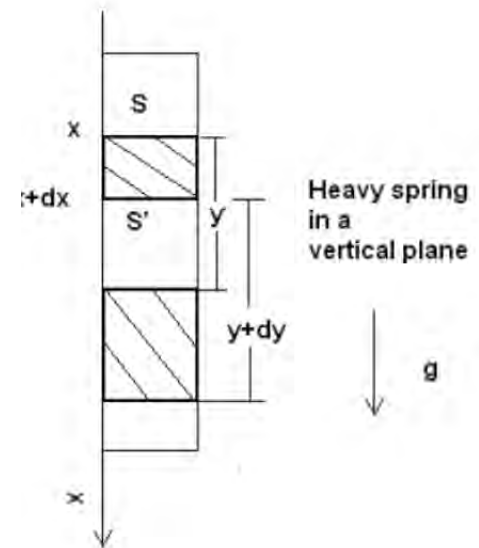
Consideriamo un elemento dx della molla alla distanza x dal vincolo. A causa della massa delle molla sottostante tale elemento si sposterà di y e la lunghezza dell'elemento sarà pari a dy .

Ad x la costante elastica della parte da 0 ad x della molla sarà:

$$k(x) = \frac{ES}{x}$$

Mentre la massa della molla sottostante è:

$$m(x) = m_m \frac{\ell_0 - x}{\ell_0}$$



Quanto vale $m_{\text{eff. statica}}$

In base a ciò l'allungamento y sarà:

$$y = \frac{m(x)g}{k(x)} = m_m g \frac{\ell_0 - x}{\ell_0} \frac{1}{ES} x$$

O, allo stesso modo, l'allungamento dell'elemento dx alla distanza x sarà:

$$dy = m_m g \frac{\ell_0 - x}{\ell_0} \frac{1}{ES} dx$$

Dall'integrazione otteniamo l'allungamento totale della molla:

$$\Delta \ell = \int dy = \int_0^{\ell_0} m_m g \frac{\ell_0 - x}{\ell_0} \frac{1}{ES} dx = \frac{m_m g}{ES} \frac{1}{\ell_0} \int_0^{\ell_0} (\ell_0 - x) dx = \frac{m_m g}{ES} \frac{1}{\ell_0} \left(\ell_0^2 - \frac{\ell_0^2}{2} \right) = \frac{m_m}{2} g \frac{\ell_0}{ES}$$

Ricordando che $E = \frac{k\ell_0}{s}$ ottengo

$$\Delta \ell = \frac{m_m g}{2} \frac{g}{k} \quad \text{ovvero nuovamente} \quad m_{\text{eff. statica}} = \frac{m_m}{2}$$

Misura della costante elastica – Metodo dinamico

Partiamo dalla legge di Hooke:

$$mg = k\Delta\ell = k(\ell_{eq} - \ell_0)$$

Spostando la molla dalla posizione di equilibrio raggiunta applico un'ulteriore forza. Se lascio poi andare la massa questa si metterà ad oscillare rispetto alla posizione di equilibrio ℓ_{eq} con equazione:

$$m\ddot{x} + kx = mg$$

La soluzione di questa equazione differenziale sarà data da una soluzione dell'equazione omogenea:

$$m\ddot{x} + kx = 0 \quad \text{o} \quad \ddot{x} + \frac{k}{m}x = 0$$

Sommata ad una soluzione particolare dell'equazione non omogenea. Le condizioni iniziali sono:

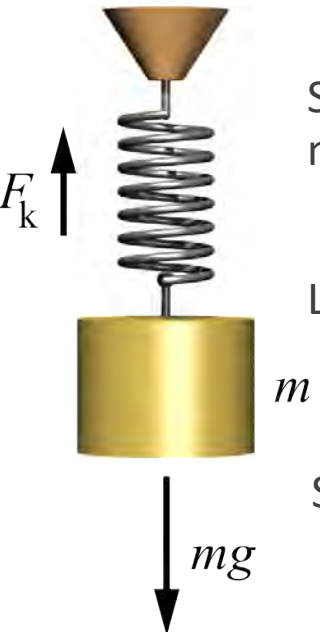
$$\begin{aligned}\dot{x}(t=0) &= 0 \\ x(t=0) &= x_0\end{aligned}$$

La soluzione dell'equazione differenziale omogenea è del tipo $x(t) = A \sin(\omega t + \varphi)$ mentre una soluzione particolare dell'equazione non omogenea è la soluzione costante:

$$x = \frac{mg}{k} = \ell_{eq} - \ell_0$$

La soluzione completa è quindi:

$$x(t) = A \sin(\omega t + \varphi) + \ell_{eq} - \ell_0$$



Misura della costante elastica – Metodo dinamico

La soluzione completa è quindi:

$$x(t) = A \sin(\omega t + \varphi) + \ell_{eq} - \ell_0$$

Con:

$$\begin{aligned}\dot{x}(t = 0) &= 0 \\ x(t = 0) &= x_0\end{aligned}$$

Abbiamo:

$$\begin{aligned}\dot{x} &= A\omega \cos(\omega t + \varphi) = 0 \\ \ddot{x} &= -A\omega^2 \sin(\omega t + \varphi)\end{aligned}$$

Da cui:

$$-A\omega^2 \sin(\omega t + \varphi) + \frac{k}{m} A \sin(\omega t + \varphi) + \frac{k}{m} [\ell_{eq} - \ell_0] = g$$

Dobbiamo perciò avere:

$$\omega = \sqrt{\frac{k}{m}}$$

Ovvero la pulsazione non cambia a causa della forza di gravità.

Misura della costante elastica – Metodo dinamico

La condizione iniziale $\dot{x}(t = 0) = 0$ fornisce:

$$\cos(\varphi) = 0 \quad \text{o} \quad \varphi = \frac{\pi}{2}$$

Mentre applicando $x(t = 0) = x_0$ otteniamo:

$$x_0 = A + \ell_{eq} - \ell_0$$

Da cui

$$A = x_0 - (\ell_{eq} - \ell_0)$$

Per una soluzione completa data da:

$$x(t) = [x_0 - (\ell_{eq} - \ell_0)] \sin\left(\sqrt{\frac{k}{m}}t + \frac{\pi}{2}\right) + (\ell_{eq} - \ell_0)$$

Il fatto che $\omega = \sqrt{\frac{k}{m}}$ ci permette di calcolare la costante elastica della molla dalla misura del periodo di oscillazione:

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{m}{k}}$$

Ma in questo caso la massa è tutta la massa appesa alla molla, inclusa la massa del supporto.

Quanto vale $m_{\text{eff. dinamica}}$

Nella misura della costante elastica con metodo dinamico non è possibile trascurare la massa della molla in relazione alla stima del periodo di oscillazione. Occorrerà valutare il contributo di m_m anche in questo caso come $m_{\text{eff. dinamica}}$.

Calcoliamo la variazione dell'energia cinetica dK della molla a causa della massa dm di un elemento di lunghezza dx che si muove con velocità istantanea v durante il moto armonico della molla:

$$dK = \frac{1}{2} dm v^2$$

Poiché vale $dm = m_m \frac{dx}{\ell}$, mentre la velocità alla distanza x sarà legata alla velocità all'estremo ℓ della molla V da:

$$v = \frac{x}{\ell} V$$

Possiamo scrivere

$$dK = \frac{1}{2} \left(\frac{x}{\ell} V \right)^2 \frac{m_m}{\ell} dx$$

Quanto vale m . $m_{\text{eff.}}$ dinamica

L'energia cinetica totale della molla in oscillazione si ottiene integrando da 0 alla lunghezza della molla ℓ :

$$K = \frac{1}{2} \left(\frac{V}{\ell} \right)^2 \frac{m_m}{\ell} \int_0^{\ell} x^2 dx$$

Che vale

$$K = \frac{1}{2} \left(\frac{V}{\ell} \right)^2 \frac{m_m \ell^3}{3} = \frac{1}{2} V^2 \frac{m_m}{3}$$

Sommando anche la massa appesa alla molla (che si muove anch'essa con velocità V) abbiamo un'energia cinetica totale

$$K = \frac{1}{2} V^2 \left(m + \frac{m_m}{3} \right)$$

Abbiamo quindi che la massa effettiva della molla da sommare alla massa appesa è nel caso dinamico pari a

$$m_{\text{eff. dinamica}} = \frac{m_m}{3}$$

Modello sbagliato?

Ho comunque un modo per evitare di applicare un modello (come quello della massa dinamica efficace). Immaginiamo di misurare i periodi di oscillazione per la molla con il solo piattello e per la molla con il piattello e la massa m_i ; abbiamo:

$$T_i^2 = 4\pi^2 \frac{\frac{m_m}{3} + m_p + m_i}{k} \quad \text{e} \quad T_{p,i}^2 = 4\pi^2 \frac{\frac{m_m}{3} + m_p}{k}$$

La differenza tra i due periodi e' indipendente dal modello usato per stimare la massa efficace della molla:

$$T_i^2 - T_{p,i}^2 = 4\pi^2 \frac{m_i}{k}$$

E da questo il valore della costante elastica k

$$k = 4\pi^2 \frac{m_i}{T_i^2 - T_{p,i}^2}$$

I periodi con il solo piattello vanno rimisurati per ogni misura con masse m_i diverse, in modo da avere un set di misure indipendenti.

Metodo Dinamico con attriti

Nel caso in cui non si possano trascurare gli attriti, dobbiamo inserire un termine proporzionale a dx/dt che dipenderà dalla viscosità del mezzo (e da eventuali attriti meccanici nella molla stessa). Consideriamo per semplicità l'equazione omogenea (abbiamo visto la soluzione della non omogenea):

$$m\ddot{x} + C\dot{x} + kx = 0$$

Ovvero:

$$\ddot{x} + \frac{C}{m}\dot{x} + \frac{k}{m}x = 0$$

Con condizioni iniziali:

$$x(t = 0) = x_0 \quad e \quad \dot{x}(t = 0) = 0$$

E con, per comodità:

$$2\gamma = \frac{C}{m} \quad e \quad \omega_0 = \sqrt{\frac{k}{m}}$$

Ovvero:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0$$

Che è un'equazione differenziale lineare di secondo ordine a coefficienti costanti.

Metodo Dinamico con attriti

La teoria delle equazioni differenziali ci dice che esistono **due soluzioni linearmente indipendenti** e che la soluzione generale è una combinazione lineare di queste.

Per trovare le due soluzioni, pensiamo ad x come una variabile complessa, $x(t) \rightarrow z(t)$. Poiché i coefficienti sono reali, se $z(t)$ è soluzione di $\ddot{z} + 2\gamma\dot{z} + \omega_0^2 z = 0$, allora $x(t) = \mathbb{R}[z(t)]$ è soluzione dell'equazione in x .

Dalla teoria, la forma della soluzione è:

$$z(t) = Ae^{st}$$

Posso ottenere la costante complessa s , osservando che vale:

$$\frac{d}{dt} Ae^{st} = Ase^{st} \quad e \quad \frac{d^2}{dt^2} Ae^{st} = As^2 e^{st}$$

Usando ciò riscrivo l'equazione differenziale come una semplice equazione

$$s^2 + 2\gamma s + \omega_0^2 = 0 \quad \text{con soluzioni} \quad s_{\pm} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2}$$

E la soluzione dell'equazione in $z(t)$ risulta essere:

$$z(t) = A^+ e^{s_+ t} + A^- e^{s_- t} = e^{-\gamma t} \left[A^+ e^{\sqrt{\gamma^2 - \omega_0^2} t} + A^- e^{-\sqrt{\gamma^2 - \omega_0^2} t} \right]$$

Metodo Dinamico con attriti

Il caso che ci interessa è quello in cui il coefficiente di attrito è minore della pulsazione:

$$\gamma < \omega_0$$

In questo il caso il discriminante è negativo

$$s_{\pm} = -\gamma \pm i\sqrt{\omega_0^2 - \gamma^2}$$

Ed il moto è descritto dalla legge:

$$x(t) = Ae^{-\gamma t} \cos\left(\sqrt{\omega_0^2 - \gamma^2}t + \phi\right)$$

E' nuovamente un moto oscillatorio smorzato, con A , e ϕ che dipendono dalle condizioni iniziali. Notare che la pulsazione misurata è influenzata dall'attrito:

$$\omega_0 \rightarrow \omega_0 - \frac{c}{2m}$$

Misura di τ

Misurare $\tau = 5T = t_{stop} - t_{start}$ per ogni massa m_i .

- Misurare $\tau = 5T = t_{stop} - t_{start}$ per almeno 10 volte.
- Fare un istogramma con le 10 misure
- Per la trattazione statistica che abbiamo visto i risultati seguono una distribuzione Gaussiana.
 - La miglior stima del valore vero di τ è il valor medio $\bar{\tau} = \frac{1}{10} \sum_{i=1}^{10} \tau_i$
 - La miglior stima della varianza della singola misura (del campione di misure) τ è data da $\sigma_{\tau}^2 = \frac{1}{10-1} \sum_{i=1}^{10} (\tau_i - \bar{\tau})^2$
 - La miglior stima della varianza della media è data da $\sigma_{\bar{\tau}}^2 = \frac{\sigma_{\tau}^2}{10}$
 - Abbiamo quindi $\bar{\tau} \pm \sigma_{\bar{\tau}}$.
 - Tenuto conto che $\bar{T} = \bar{\tau}/5$ e che $\sigma_{\bar{T}} = \sigma_{\bar{\tau}}/5$ abbiamo anche misurato $\bar{T} \pm \sigma_{\bar{T}}$
- Il valore $k = (2\pi)^2 \frac{m}{T^2}$

Misura dello smorzamento

Dobbiamo valutare se possiamo o meno trascurare lo smorzamento nella stima della costante elastica della molla con il metodo dinamico. Anche in questo caso il valore di γ può essere stimato in base all'andamento delle ampiezze massime in funzione del tempo; se prendiamo l'ampiezza massima A_0 al tempo t_0 e le ampiezze massime A_i ai tempi successivi t_i (ogni 10 o 20 o 30 secondi in base al fenomeno) avremo, poiché ai massimi $\cos \alpha = 1$

$$A_i = A_0 e^{-\gamma(t_i - t_0)}$$

Da cui possiamo fare un grafico del valore dei γ_i

$$\gamma_i = \frac{1}{t_i - t_0} \ln \frac{A_0}{A_i}$$

in funzione del tempo t , dopo che abbiamo calcolato l'errore associato (legge di propagazione degli errori massimi);

$$\frac{\Delta \gamma_i}{\gamma_i} = \frac{2\Delta t}{t_i - t_0} + \frac{2\Delta A}{\frac{A_0}{A_i} \ln \frac{A_0}{A_i}}$$

Verifichiamo quindi che l'andamento sia costante; Verifichiamo la compatibilità dei risultati. Infine possiamo stimare il valore di γ (ed il suo errore) da applicare.

Misura della costante elastica – Metodo dinamico

Usando due molle con costanti elastiche diverse:

- Misurare con 5-6 valori di massa diversi (senza superare i limiti di carico delle molle)
- Fare un grafico di \bar{T}_i in funzione di m_i
- Fare un grafico di k_i in funzione di m_i
- Errori dati da (con $\sigma_m = \Delta m / \sqrt{3}$):

$$\frac{\sigma_i^2}{k_i^2} = 4 \frac{\sigma_{\bar{T}_i}^2}{\bar{T}_i^2} + \frac{\sigma_{m_i}^2}{m_i^2}$$

- Misura di k da media pesata usando σ_i^2 come pesi
- In questo caso la massa della molla non può essere trascurata. Possiamo ovviare in due modi
 - Con un modello che descriva l'effetto della massa della molla
 - Usando la prima misura per calcolare k dalla differenza:

$$T_i^2 - T_{0,i}^2 = (2\pi)^2 \frac{m_i}{k_i} \quad \text{da cui} \quad k_i = (2\pi)^2 \frac{m_i}{T_i^2 - T_{0,i}^2}$$

In questo caso (trascurando le masse)

$$\frac{\sigma_i^2}{k_i^2} = 4 \frac{\bar{T}_i^2 \sigma_{\bar{T}_i}^2 + \bar{T}_0^2 \sigma_{\bar{T}_0}^2}{(\bar{T}_i^2 - \bar{T}_{0,i}^2)^2}$$

Misura della costante elastica – commenti

- Nel caso in cui usiamo

$$k_i = (2\pi)^2 \frac{m_i}{T_i^2 - T_{0,i}^2}$$

Le misure non sono tra loro indipendenti (a meno di non misurare ogni volta il periodo a piatto scarico...quindi; **MISURARE + VOLTE il periodo a piatto scarico**). Assumere la misura con la deviazione standard più piccola, o rimisurare ogni volta T_0 , ovvero $T_{0,i}$

- Confrontare la precisione ottenuta mediante i 3 metodi
 - Statico
 - Dinamico con $m + \frac{m_m}{3}$
 - Dinamico con differenza $k_i = (2\pi)^2 \frac{m_i}{T_i^2 - T_{0,i}^2}$
- Commentare la compatibilità delle misure
- Discutere il risultato

Probabilità e statistica

Introduzione

Se noi non fossimo ignoranti non ci sarebbe probabilità , ci potrebbero essere solo certezze. Ma la nostra ignoranza non può essere assoluta, altrimenti non ci sarebbe più probabilità . Così i problemi di probabilità possono essere classificati a seconda della maggiore o minore profondità della nostra ignoranza.

(H. Poincaré)

Meccanica classica: note le proprietà di un corpo (massa, forma, etc.), le condizioni iniziali di moto (posizione, velocità etc.) e le condizioni esterne (campi di forze, condizioni al contorno, etc.) è possibile determinarne in modo esatto il suo comportamento negli istanti successivi.

Ma è sufficiente analizzare il semplice esperimento del lancio di una moneta o di un dado per capire che non è facile prevedere il risultato, ossia l'occorrenza testa/croce o di uno dei 6 numeri nella faccia superiore. In modo simile non è possibile trattare sistemi complessi (come un gas in un cilindro) a partire dai moti delle singole molecole.

Semplicemente con un numero di parametri troppo elevati, modifichiamo il 'metodo scientifico', ovvero rinunciamo a predire l'esatto esito dell'esperimento e ci limitiamo ad esprimere affermazioni sulla plausibilità di ciascuno dei risultati possibili. In altri termini, sebbene si seguiti ad assumere che le leggi "di base" siano di tipo deterministico, il nostro *stato di incertezza* riguardo i dettagli del moto ci impedisce di arrivare a conclusioni *certe*.

Introduzione II

Ci sono poi fenomeni in cui è la stessa meccanica classica a non essere adeguata alla descrizione del processo elementare. Questo succede sulle scale atomiche e inferiori. In questo caso sono le stesse leggi fondamentali che assumono natura aleatoria; **si perde il carattere deterministico** “almeno di principio” della meccanica classica e si deve utilizzare il linguaggio della *meccanica quantistica*.

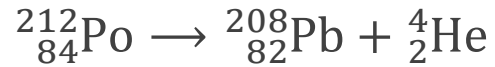
È da notare comunque che, sebbene le due situazioni siano sostanzialmente **differenti dal punto di vista fisico**, esse sono **simili dal punto di vista conoscitivo**: in entrambe siamo in stato di incertezza rispetto ai possibili esiti, anche se nel caso quantistico c'è la convinzione che, pur partendo da un ben preciso stato di preparazione del sistema iniziale, l'evoluzione sia intrinsecamente probabilistica.

Le diversità delle problematiche si riflette sia sul modo di intendere le leggi fisiche che sui metodi usati per valutare la probabilità dei possibili esiti.

La teoria della probabilità è una branca delle discipline matematiche che studia i problemi di definizione e misurazione dell'incertezza.

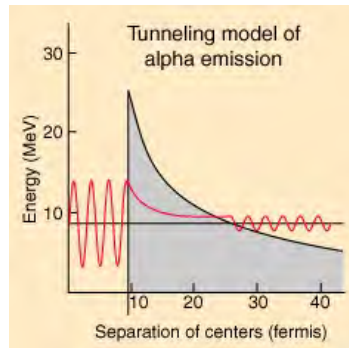
Esempio: decadimento α

Interessa nuclei pesanti, da cui viene emesso un nucleo di elio, ad esempio:

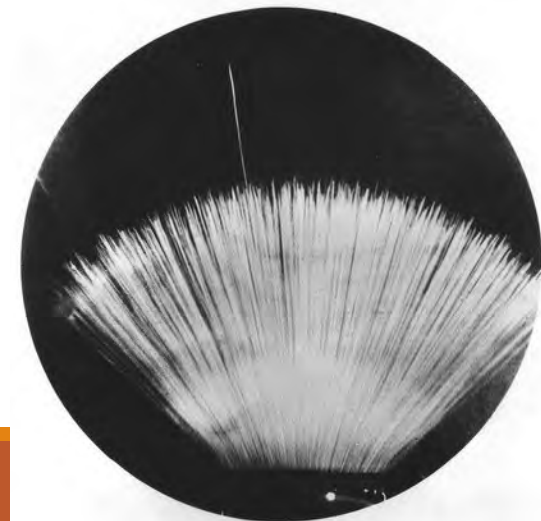
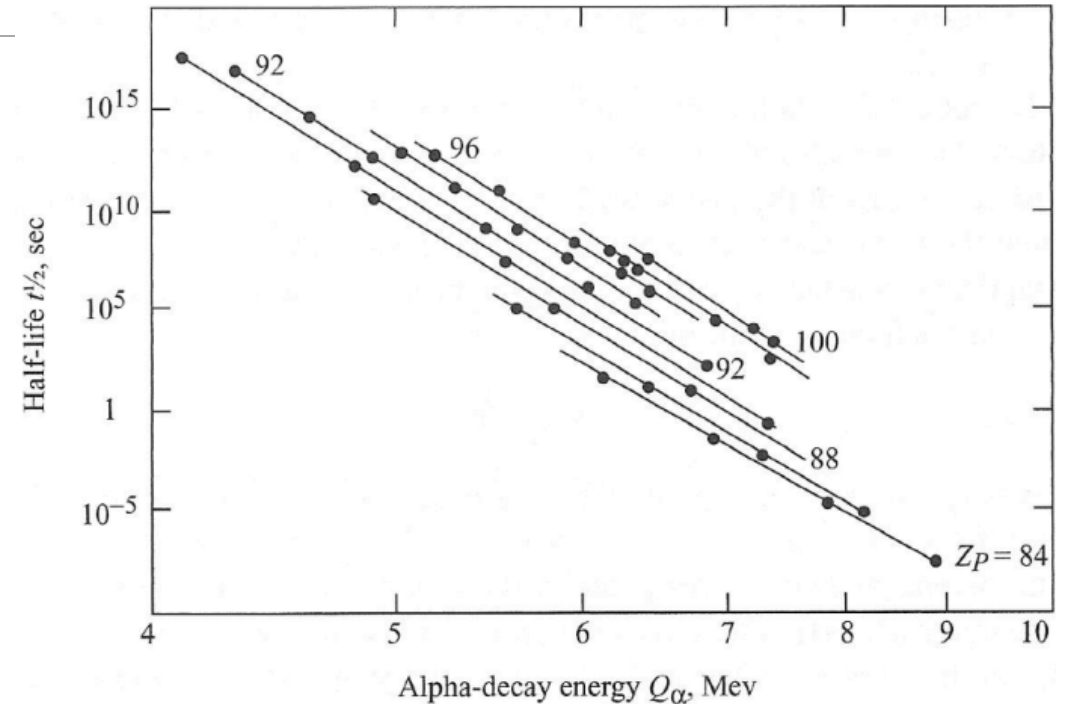


I nuclei α all'interno del nucleo piú grande non dovrebbero uscire poiché non hanno energia sufficiente.

Il decadimento avviene per effetto tunnel (uno dei primi successi della meccanica quantistica)



Come descritto nel 1929 da Gamow, Condon and Gurney.



Nascita della definizione di probabilità

L'individuazione di un preciso riferimento cronologico per la nascita della **definizione di probabilità** è una questione ancora aperta.

Una delle tappe fondamentali sembra essere il 1654. Antoine Gombaud, cavaliere di Méré, un nobile francese appassionato di giochi d'azzardo, richiama l'attenzione di **Pascal** sull'apparente contraddizione emersa in un popolare gioco d'azzardo, consistente nel **lancio ripetuto di una coppia di dadi**.

Il problema che Gombaud sottopone a Pascal è calcolare **la probabilità di ottenere almeno un "doppio 6" nell'arco di 24 lanci** (una regola operativa diffusa tra i giocatori dell'epoca suggeriva che scommettere sull'evento "il doppio 6 si verifica almeno una volta su 24 lanci" sarebbe stato vantaggioso, ma i calcoli di Gombaud portavano a conclusioni opposte).

Tali problemi portarono **Pascal e Fermat** ad uno **scambio epistolare**, nel quale furono formulati per la prima volta alcuni principi fondamentali della teoria della probabilità. Infatti, nessuna formulazione teorica generale era stata sviluppata prima di questo carteggio.

Christian Huygens, venuto a conoscenza di questa corrispondenza, pubblica nel 1657 il primo trattato di teoria della probabilità, dal titolo ***De ratiociniis in ludo aleae***.

Tra gli studiosi ai quali si devono fondamentali contributi alla teoria della probabilità in questo periodo i contributi più importanti sono quelli di Jacob Bernoulli (1654-1705) e Abraham De Moivre (1667-1754).

Perché studiare la probabilità

L'inizio della teoria delle probabilità, chiamata all'epoca la "dottrina della sorte", avviene quindi nel XVII secolo, come risposta a due classi di problemi, legate rispettivamente ai giochi d'azzardo e alle assicurazioni. Grandi matematici e statistici come Fermat (1601-1665), Laplace (1749-1827) e Bernoulli (1654-1705) discutono spesso nelle memorie di esperimenti eseguiti con dadi, ed uno dei loro obiettivi era proprio quello di fornire strategie vincenti.

Nel secondo caso si rendeva necessaria per le assicurazioni la stima della probabilità di morte di un individuo di una certa età, ovvero la probabilità che egli potesse sopravvivere un determinato numero di anni dalla stipula del contratto.

Questi due diversi contesti hanno dato luogo a due diversi metodi per valutare la probabilità, o meglio, a due "definizioni" di probabilità talvolta in contrapposizione fra di loro: la "matematica" e la "sperimentale" (o "empirica"); la "classica" e la "frequentista"; quella "a priori" e quella "a posteriori".

La definizione ISO

Nel 1993 l'Organizzazione Internazionale per la Standardizzazione (ISO) ha pubblicato una serie di guide ([“Guides in metrology”](#)) per stabilire regole generali per valutare e esprimere l'incertezza della misura applicabili su un vasto spettro di misure. Nella guida introduttiva ([Evaluation of measurement data](#)) Vi si legge:

... In contrast to this frequency-based point of view of probability, an equally valid viewpoint is that **probability is a measure of the degree of belief that an event will occur.**

For example, suppose one has a chance of winning a small sum of money D and one is a rational bettor. One's degree of belief in event A occurring is $p = 0.5$ if one is indifferent to this two betting choices:

1. receiving D if event A occurs but nothing if it does not occur;
2. receiving D if event A does not occur but nothing if it does occur.

Recommendation INC-1 (1980) upon which this Guide rests implicitly adopts such a viewpoint of probability since it views expressions such as equation (E.6) as the appropriate way to calculate the combined standard uncertainty of a result of a measurement.

Calcolo delle probabilità

Oggetto della teoria matematica sviluppata nel Calcolo delle Probabilità è un generico *esperimento casuale*, la cui singola esecuzione è chiamata *prova* dell'esperimento.

Il *risultato* (o *esito*) della prova si indica con ω . L'insieme di tutti i possibili esiti costituisce lo *spazio campione* Ω associato all'esperimento casuale.

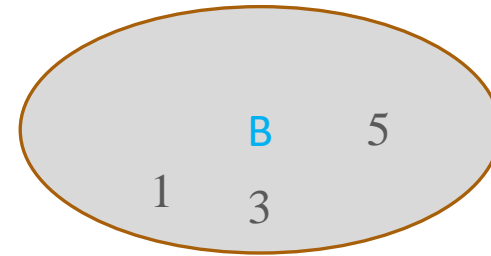
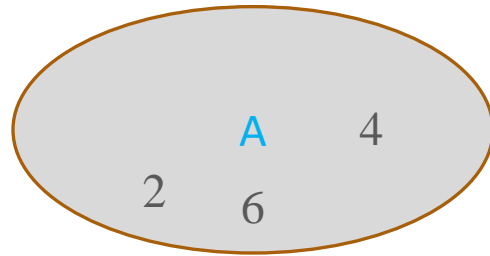
Un *Evento* A relativo al medesimo esperimento è un certo insieme di risultati ω , ovvero un sottoinsieme dello spazio campione Ω .

Se un risultato $\omega \in A$, si dice che esso realizza l'evento A . Se l'insieme $A \subset \Omega$ è costituito da un solo elemento ω , allora quest'ultimo prende il nome di *evento elementare*; altrimenti A è un *evento composto*.

Eventi

Eventi disgiunti: A e B sono disgiunti se **NON** possono verificarsi simultaneamente

- Esempio: Lancio un dado; $A = \{\text{pari}\}$, $B = \{\text{dispari}\}$



Evento certo: si verifica con certezza

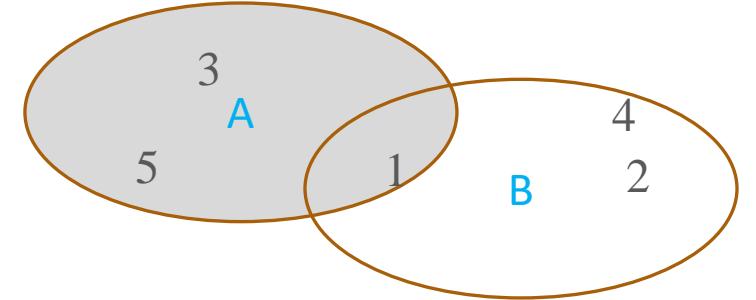
- Esempio: Lancio un dado ed esce un numero tra 1 e 6

Eventi

Unione ("somma") di eventi: $C = A \cup B$ è l'evento che si verifica se si verifica A oppure B

Esempio: $A = \{\text{dado dispari}\}$, $B = \{\text{dado} \leq 4\}$

$A \cup B = \{\text{dado} \leq 5\}$



Intersezione ("prodotto") di eventi: $C = A \cap B$ è l'evento che si verifica se si verificano sia A che B

Esempio: $A = \{\text{dado dispari}\}$, $B = \{\text{dado} \leq 2\} \Rightarrow A \cap B = \{1\}$

Differenza di eventi: $C = A - B$ o A/B è l'evento costituito dai casi di A che non appartengono a B

Esempio: $A = \{\text{dado dispari}\}$, $B = \{\text{dado} \leq 2\} \Rightarrow A - B = \{3, 5\}$

Evento negato: $B = \bar{A}$ è l'evento che si verifica se e solo se non si verifica A.

Esempio: $A = \{\text{dado dispari}\}$, $B = \bar{A} = \{\text{dado pari}\}$.

Eventi

Eventi disgiunti o incompatibili: (Insiemi disgiunti sono insiemi che non hanno alcun elemento in comune) la probabilità dell'evento $C = A \cap B$ è pari a zero $P(A \cap B) = 0$

Esempio: $A = \{\text{dado dispari}\}$, $B = \{\text{dado pari}\}$

$A \cap B = \{\text{dado pari e dispari allo stesso tempo}\}$

Eventi indipendenti: $P(A \cap B) = P(A) \cdot P(B)$ la probabilità che si verificano sia l'evento A che l'evento B è data dal prodotto delle due probabilità

Esempio, probabilità che esca un doppio 6 nel lancio di due dadi: $A = \{6 \text{ in dado } 1\}$,

$$B = \{6 \text{ in dado } 2\} \implies P(A \cap B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Definizione di probabilità

CLASSICA O 'A PRIORISTICA' (originariamente formulata da Laplace):

Dato uno spazio **finito** di eventi Ω , la **probabilità di un evento E** , indicata con **$P(E)$** , è uguale al **rapporto tra il numero dei casi favorevoli e quello dei casi possibili**, posto che gli eventi siano tutti equiprobabili.

$$P(E) = \frac{\text{casi favorevoli}}{\text{casi possibili}}$$

Si chiama anche **probabilità a priori** di un evento. Questa definizione è sufficiente per i casi più elementari.

A tale definizione può essere mossa la seguente critica: non appena la situazione si complica leggermente diventa necessario tenere conto che alcuni casi possono essere più o meno "favorevoli". Quindi si aggiunge alla definizione la clausola "**purché siano ugualmente possibili**".

Definizione di probabilità

FREQUENZISTICA, formulata da Venn (1834-1923), emerge da un ragionamento a posteriori fondato sull'osservazione dei risultati di un esperimento:

$$P(E) = \lim_{N \rightarrow \infty} \frac{n}{N}$$

dove n = numero di volte in cui si verifica l'evento E , risultato di un certo esperimento; N = numero totale di volte in cui si ripete l'esperimento; n/N è la frequenza relativa statistica dell'evento E .

Questo limite significa che **se si compiono più serie di prove**, con N sempre più grande, **il rapporto n/N tende a stabilizzarsi** intorno ad un certo valore, **con oscillazioni sempre più piccole man mano che N cresce**. Si propone di verificarlo sperimentalmente, ad esempio per l'evento "testa" nel lancio di una moneta.

Questa definizione è buona per le applicazioni, ma non per una rigorosa costruzione matematica.

Storicamente...

Il primo tentativo di valutare la probabilità fuori dall'ambito dei giochi d'azzardo fu motivato dal **calcolo delle pensioni**.

L'**ammontare del vitalizio** da corrispondere all'assicurato **dipende** infatti, oltre che dal capitale o da altri fattori economici, **dalla probabilità di morte dell'assicurato in funzione della sua età**. Una valutazione realistica (oggettiva) della probabilità è resa necessaria onde evitare perdite economiche nei casi di eccessive sottostime o sovrastime. **Considerando la sopravvivenza di ciascuna persona da un anno all'altro, ci sono due modalità elementari, ma esse sono per fortuna non equiprobabili e quindi la "definizione" di Laplace è inapplicabile.**

Il problema fu risolto compilando delle tabelle di mortalità per le varie età e stimando la probabilità dalla frequenza (ovvero da quante volte quel tipo di evento si è verificato nel passato).

"Essendo costretti dalla consuetudine a trasferire il passato al futuro in tutte le nostre inferenze, quando il passato si è manifestato del tutto regolare e uniforme ci aspettiamo un evento con la massima sicurezza e non lasciamo posto a qualche altra supposizione contraria. . . . Sebbene diamo la preferenza a quello che è stato trovato più usuale e crediamo che questo effetto si verificherà, non dobbiamo trascurare gli altri effetti, ma dobbiamo assegnare a ciascuno di essi un particolare peso e autorità in proporzione a come lo abbiamo trovato più o meno frequente" [Hume]

Dal punto di vista storico le due “definizioni” di probabilità appena incontrate, legate fra loro dalla “legge empirica del caso” hanno indotto molti a:

- confondere il concetto di probabilità con il suo metodo di valutazione;
- ritenere che il **valore di probabilità sia oggettivo**, cioè che sia insito nella natura delle cose e non dipenda da chi lo valuta;
- credere che si possa parlare di probabilità solo in questi due casi, molto particolari e riduttivi, rispetto alla complessità del mondo reale. **Sarebbero quindi esclusi da argomentazioni probabilistiche tutti quegli eventi per i quali è impossibile eseguire l’inventario dei casi possibili e di quelli favorevoli o per i quali non è possibile ripetere “infinite volte” l’esperimento nelle identiche condizioni.**

Non è difficile convincersi che **la probabilità dipende dallo stato di conoscenza** (o stato di informazione) **di chi la valuta**:

- la probabilità $1/6$ di ciascuno dei risultati del lancio di un dado assume la regolarità del dado. Ciò nonostante un dado reale avrà necessariamente delle asimmetrie (ad esempio dovute al diverso numero di puntini incisi sulle facce) e qualcuno può credere a valori diversi da $1/6$ se ha una conoscenza meno vaga di quel dado;
- la situazione in cui il dado è stato già lanciato, ma riparato dalla vista di chi deve stimare la probabilità, non cambia la valutazione. La probabilità (del dado regolare) è sempre $1/6$ anche se dal punto di vista fisico l'esito è già determinato.
- Se invece si riesce a sbirciare e a leggere il valore di una faccia laterale la situazione cambia immediatamente: la probabilità della faccia vista e del suo complemento a 7 si annullano e quelle delle altre facce aumentano ad $1/4$.

Definizione soggettivista di Probabilità

Sviluppata indipendentemente da Ramsey (1903-1930) e De Finetti (1906-1985), l'impostazione soggettiva afferma *che la probabilità di un evento ω è una misura p del grado di fiducia che un individuo attribuisce al verificarsi di ω sulla base delle sue opinioni ed informazioni sull'evento e il principio di coerenza.*

De Finetti preferiva utilizzare una definizione più operativa: *la probabilità di un evento ω , secondo l'opinione di un individuo I , è il prezzo p che I giudica equo pagare per riscuotere un importo unitario nel caso in cui ω si verifichi.*

Quindi una misura di un'opinione personale (adottata da ISO):

la probabilità di un evento è il grado di fiducia che si ha nel verificarsi di esso

Per esempio: la probabilità che in un processo giudiziario l'imputato sia giudicato colpevole è una misura della nostra conoscenza dei fatti e della nostra abilità deduttiva. Tale definizione si formalizza adottando lo schema tipico delle scommesse **regolate da condizioni di equità**: la probabilità dell'evento è misurata dal prezzo che un individuo ritiene equo pagare per ricevere "1" se l'evento si realizza, e "0" se non si verifica.

Definizione "assiomatica" di probabilità:

Formulata da Kolmogorov (1903-1987), è considerata la definizione più consistente sul piano matematico, in quanto si fonda su alcuni principi generali di teoria della misura.

Dato un evento $A \subseteq \Omega$ la misura di probabilità P è una funzione che associa ad A un numero $P(A)$ (la probabilità di A) non negativo $P: \Omega \rightarrow \mathbb{R}^+$ che gode delle seguenti proprietà (assiomi)

1. $P(A) \geq 0$
2. $P(\{\text{evento certo}\}) = 1$
3. Se A e B sono eventi disgiunti (incompatibili), cioè $P(A \cap B) = 0$, allora $P(A \cup B) = P(A) + P(B)$
4. Completa additività: Data una famiglia composta da un'infinità numerabile di eventi $\{A_i\} = \{A_i; i = 1, 2, \dots\}$ a due a due incompatibili, tali cioè che $P(A_i \cap A_k) = 0 \forall i \neq k$

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

Definizione "assiomatica" di probabilità:

Osservazione: Dato un insieme di eventi, ci sono più modi di assegnare le probabilità

Esempio: Moneta con testa T o croce C.

$$\left. \begin{array}{l} P(T \cup C) = (\{\text{evento certo}\}) = 1 \\ P(T \cup C) = P(T) + P(C) \end{array} \right\} \Rightarrow P(C) = 1 - P(T)$$

Posso assegnare $P(T) = 0.5$ oppure $P(T) = p$, $P(C) = 1 - p$

(al variare di p ho tutte le possibili monete truccate)

Il calcolo delle probabilità insegna come maneggiare le probabilità supponendo di conoscerle. La loro stima è affare della statistica

Proprietà della misura di probabilità

- Probabilità dell'evento negazione:

$$P(\bar{A}) = 1 - P(A)$$

- Estremi della misura di probabilità:

$$0 \leq P(A) \leq 1$$

- Teorema delle probabilità totali: Dati due eventi $A, B \in \Omega$ compatibili, caratterizzati cioè da intersezione non vuota $P(A \cap B) \neq 0$

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

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

- Probabilità in una relazione di inclusione: Se l'evento A è incluso nell'evento B ($A \subseteq B$) allora

$$P(A) \leq P(B)$$

Esempio: Calcolare la probabilità che un dado onesto dia un numero pari o minore di 3

$$A = \{pari\} = \{2, 4, 6\} \implies P(A) = \frac{1}{2}$$

$$B = \{dado < 3\} = \{1, 2\} \implies P(B) = \frac{1}{3}$$

$$A \cap B = \{2\} \implies P(A \cap B) = \frac{1}{6}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{4}{6}$$

Probabilità condizionata

Se la valutazione di probabilità dipende dallo stato di informazione, non ha senso parlare di una “probabilità assoluta”. Si può soltanto parlare di **probabilità condizionata ad una certa informazione**. Indichiamola genericamente con $P(E|I)$, “probabilità di E dato lo stato di informazione I ”, o “probabilità di E data I ”.

Quando si parla di $P(E)$ senza aggiungere altro si fa riferimento a circostanze convenzionali, oppure - implicitamente - allo stato di informazione di chi la valuta.

- quando si dice che la probabilità della faccia di un dado sia $1/6$ si sta assumendo che dado e lancio siano perfettamente regolari;
- quando si dice che la probabilità di **testa** nel lancio di una moneta sia $1/2$ si assume che la moneta sia regolare, che il lancio sia “fatto a caso” (non è irragionevole pensare ad un prestigiatore in grado di controllare il lancio) e che inoltre la moneta non cada verticale e non vada in un tombino (ne’ che sia catturata al volo).

Probabilità condizionata

Dati due eventi A ed M con $P(M) \neq 0$, la **probabilità di A condizionata da M** (probabilità che si verifichi A se si è verificato M) è:

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

Esempio: Trovare la probabilità che un dado dia $\{<4\}$ sapendo che il risultato del lancio è stato pari:

$$P(\{< 4\}|\{pari\}) = \frac{P(\{1, 2, 3\} \cap \{pari\})}{P(\{2, 4, 6\})} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

Teorema della probabilità totale: Se M_1, M_2, \dots, M_n sono eventi disgiunti (cioè $P(M_i \cap M_j) = 0 \forall i \neq j$) e l'unione $M_1 \cup M_2 \cup \dots \cup M_n = \{\text{evento certo}\}$ allora:

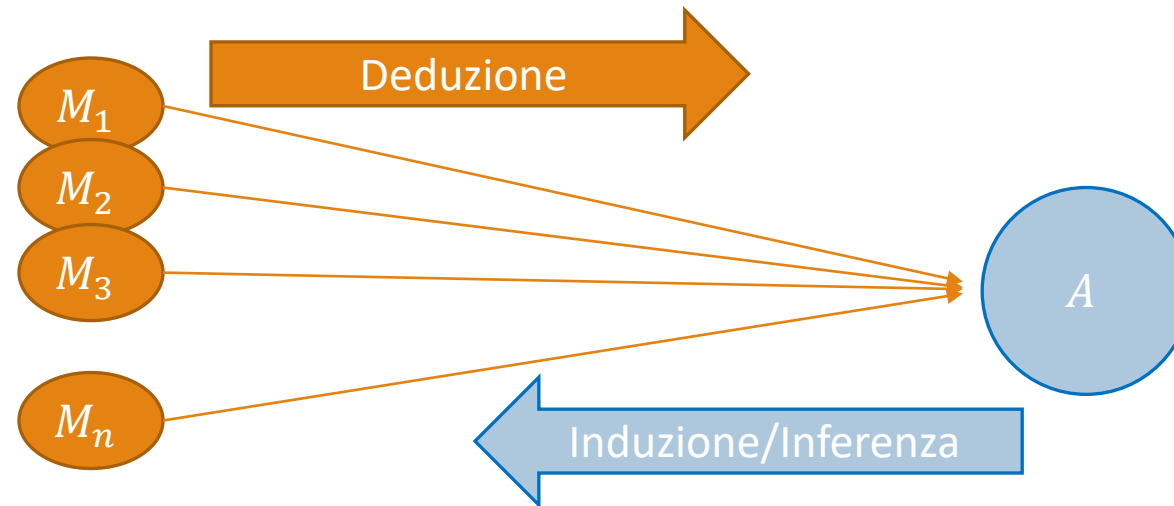
$$P(A) = \sum_{i=1}^n P(A|M_i)P(M_i)$$

Probabilità delle cause

Un modo interessante di rileggere la probabilità condizionata è di pensare **il condizionante causa dell'evento (visto come effetto)**. Ovvero: **abbiamo osservato A ; qual è la probabilità che la causa sia M_i ?** Questo vale, ad esempio, se si considerano gli eventi condizionati:

- “auto rubata” | “auto nuova e di valore”;
- “l'ago della bilancia si posiziona su 1000.00 g” | “chilogrammo campione su bilancia di laboratorio”;
- “si registrano tot conteggi in un contatore di radioattività” | “la radioattività ambientale vale”;

se si riesce a trovare una regola per **invertire la probabilità** e valutare $P(M|A)$ a partire da $P(A|M)$ questa potrà essere utilizzata per imparare dall'esperienza



Teorema di Bayes o probabilità delle cause

Formulato da Thomas Bayes (1702-1761), pubblicato postumo due anni dopo la sua morte in «Essays towards solving a problem in the doctrine of chances» e da Laplace nel 1774, è alla base di innumerevoli applicazioni pratiche e di un particolare approccio statistico, detto appunto bayesiano

Consideriamo un evento A e una classe completa di ipotesi M_i , ovvero che siano esaustive e mutuamente esclusive:

$$\begin{cases} M_i \cap M_j = \emptyset \quad \forall (i \neq j) \\ \bigcup_i M_i = \Omega \end{cases}$$

Applicando la formula per la probabilità condizionata

$$P(A \cap M_i) = P(M_i)P(A|M_i) = P(A)P(M_i|A) \quad \text{ricordo che} \quad P(A \cap M_i) = P(M_i \cap A)$$

Da cui (se $P(A) \neq 0$):

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

Probabilità di una causa se si è verificato un effetto

Teorema di Bayes

Ricordando che l'insieme di alternative M_1, M_2, \dots, M_n che partizionano in modo completo lo spazio degli eventi Ω (ovvero $M_i \cap M_j = \emptyset \forall i \neq j$ e $\bigcup_{i=1}^n M_i = \Omega$) permette di esprimere $P(A) = \sum_{j=1}^n P(A|M_j)P(M_j)$ ricaviamo la seguente espressione per la probabilità condizionata:

$$P(M_i|A) = \frac{P(A|M_i)P(M_i)}{P(A)} = \frac{P(A|M_i)P(M_i)}{\sum_{j=1}^n P(A|M_j)P(M_j)}$$

Dove:

- $P(M)$ è la probabilità a priori o [probabilità marginale](#) di M . « $P(M)$ a priori» significa che non tiene conto di nessuna informazione riguardo A .
- $P(M|A)$ è la [probabilità condizionata](#) di M , noto A . Viene anche chiamata probabilità a posteriori, visto che è derivata o dipende dallo specifico valore di A .
- $P(A|M)$ è la probabilità condizionata di A , noto (avvenuto) M .
- $P(A)$ è la probabilità a priori di A , e funge da [costante di normalizzazione](#).

Intuitivamente, il teorema descrive il modo in cui le opinioni nell'osservare M siano arricchite dall'aver osservato l'evento A .

Viene impiegato per calcolare la probabilità di una causa che ha scatenato l'evento verificato.

Teorema di Bayes

Esempio: Carta di controllo. M = causa speciale, A = punto fuori dai limiti di controllo.

- Se c'è una causa speciale la carta segnala un fuori controllo con probabilità $P(A|M) = 0.5$
- In assenza di cause speciali, la probabilità di falso allarme è $P(A|\bar{M}) = 0.0027$
- La probabilità di una causa speciale è $P(M) = 0.05$.

Sapendo che è stato segnalato un fuori controllo, quale è la probabilità che ci sia effettivamente una causa speciale? *Devo calcolare $P(M|A)$*

$$P(A) = P(A|M)P(M) + P(A|\bar{M})P(\bar{M}) = 0.5 \cdot 0.05 + 0.0027 \cdot 0.95 = 0.025 + 0.002565 = 0.0276$$

Da cui...

$$P(M|A) = \frac{P(A|M)P(M)}{P(A)} = \frac{0.025}{0.0276} = 0.907$$

Indipendenza

Ricordo che due eventi A e B si dicono indipendenti se:

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

Osservazione: Se A e B sono indipendenti:

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

Interpretazione: L'informazione che B si è verificato non ha nessuna influenza sulla probabilità che si verifichi A.

Osservazione: Disgiunzione e indipendenza sono due concetti diversi!

Esempio: Dado onesto; A = {pari}, B = {dispari}

- A e B sono disgiunti
- $P(A \cap B) = P(\{\text{evento impossibile}\}) = 0 \neq P(A) \cdot P(B)$

Osservazione: In molti casi l'indipendenza viene postulata in base a ragioni fisiche. In tal caso, mi permette di calcolare $P(A \cup B)$ (probabilità dell'evento congiunto) in base alla sola conoscenza di $P(A)$ e $P(B)$.

Se non c'è indipendenza, la conoscenza di $P(A)$ e $P(B)$ è insufficiente a determinare $P(A \cap B)$

Variabili Casuali

Dato lo spazio degli eventi elementari ω (o insieme di tutte le possibili modalità relative ad un determinato fenomeno, o spazio campione Ω) una “variabile casuale” (random variable) o aleatoria è un’applicazione/funzione X che associa gli elementi ω di Ω a numeri reali $X(\omega) = x$.

Variabile Casuale X : un esperimento casuale il cui esito è un numero reale x

Esempi:

1. Lancio di un dado. Possibili esiti: $x = 1, 2, 3, 4, 5, 6$.
2. Fermo per la strada una persona a caso e ne misuro la statura.
3. L'errore di misura compiuto da un sensore in una determinata misurazione.

Le variabili casuali possono essere discrete, se l’insieme dei possibili valori è finito o numerabile, o continue, se, **almeno idealmente**, possono assumere qualsiasi valore reale appartenente ad un intervallo continuo, che può anche essere $(-\infty, +\infty)$.

Nota: Nei tre esempi, ripetendo l'esperimento non ho nessuna garanzia di ottenere lo stesso risultato (casualità).

Funzione di ripartizione

Data una **variabile casuale** X a valori reali, **la funzione che fa corrispondere ad un valore x , la probabilità cumulative $P(X \leq x)$ che la variabile casuale assuma un valore $\leq x$ viene detta funzione di ripartizione**, o distribuzione cumulativa. La funzione di ripartizione è indicata con F_X :

$$F_X: \mathbb{R} \rightarrow [0,1] \quad F_X(x) := P(X \leq x)$$

La funzione di ripartizione è definita sia per le variabili casuali discrete che per le variabili casuali continue.

[**Osservazione:** Se due variabili casuali hanno la medesima funzione di ripartizione esse si dicono somiglianti.]

Proprietà della funzione di ripartizione:

1. $0 \leq F_X(x) \leq 1$
2. Nel caso in cui il dominio sia \mathbb{R} $\lim_{x \rightarrow -\infty} F_X(x) = 0$ e $\lim_{x \rightarrow +\infty} F_X(x) = 1$ con il secondo limite che è il risultato della condizione di normalizzazione (ovvero della probabilità di osservare l'evento certo).
3. $F_X(x)$ è **monotona crescente in senso lato**, ovvero $\forall x_1, x_2 \in \mathbb{R} \mid x_1 < x_2$ vale $F_X(x_1) \leq F_X(x_2)$

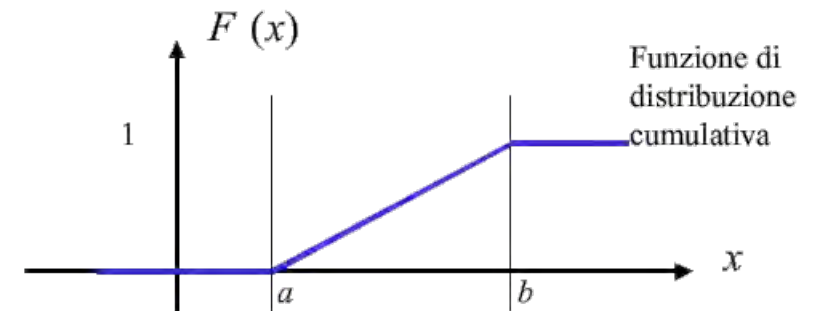
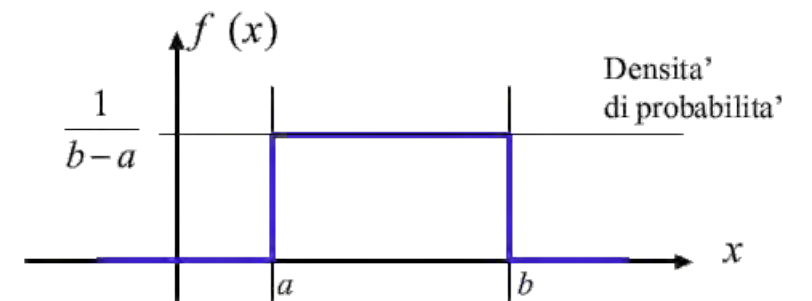
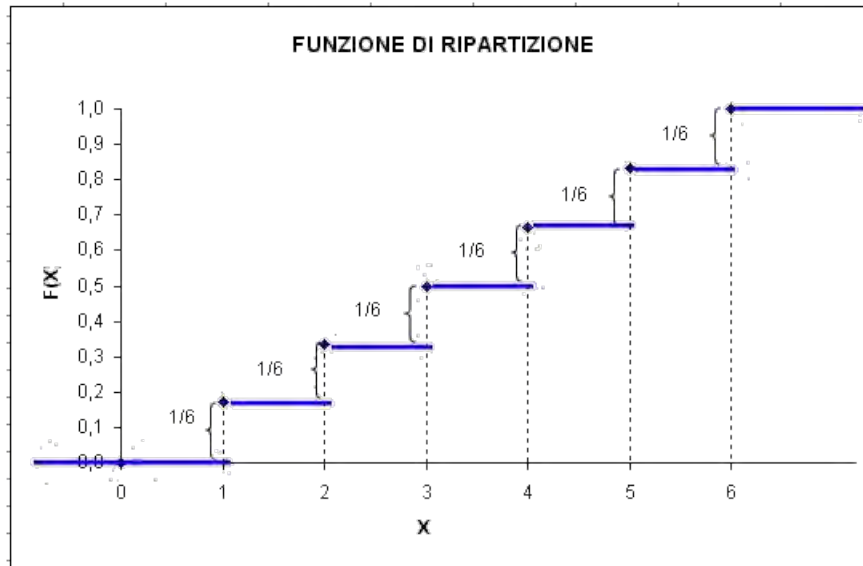
Funzione di ripartizione

4. $P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$ dunque la funzione di ripartizione consente di stabilire la probabilità che la variabile casuale semplice X assuma valori compresi in intervalli di tipo $(x_1, x_2]$ dove $x_1, x_2 \in \mathbb{R}$, con $x_1 < x_2$. A partire da questo posso calcolare anche le altre probabilità, ad esempio: $P(x_1 \leq X < x_2) = F_X(x_2) - F_X(x_1) + P(X = x_1) - P(X = x_2)$
5. Nel caso di una variabile casuale discreta, la funzione di ripartizione è continua a destra: $\lim_{x \rightarrow x_0^+} F_X(x) = F_X(x_0)$. Abbiamo anche: $\lim_{x \rightarrow x_0^-} F_X(x) \neq F_X(x_0)$ cioè la funzione di ripartizione presenta dei punti di discontinuità di 1° specie (o salti).

ESEMPIO: nel lancio di un dado non truccato gli unici risultati possibili sono le facce da 1 a 6, ciascuna con probabilità 1/6. La funzione di massa di probabilità di questa variabile casuale X e la funzione di ripartizione sono riportate nella seguente tabella:

X	$f(x) = P(X)$	$F_X(x)$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6

Esempi:



Densità di probabilità

Data la variabile casuale continua a valori reali $X: \Omega \rightarrow \mathbb{R}$ che associa valori dallo spazio campionario all'intervallo (a, b) $-\infty \leq a < b \leq +\infty$, la **funzione di densità di probabilità** (PDF) o **funzione di distribuzione di probabilità** è la funzione $f_X: \mathbb{R} \rightarrow \mathbb{R}^+$ che ad ogni x associa il limite per dx che tende a 0, del rapporto tra la probabilità che la variabile casuale assuma valori nell'intervallo $(x, x + dx]$ e l'ampiezza dx .

In simboli:

$$f_X: \mathbb{R} \rightarrow \mathbb{R}^+: x \rightarrow \lim_{dx \rightarrow 0} \left[\frac{P(x < X \leq x + dx)}{dx} \right]$$

La funzione di densità in x , allora, rappresenta quanto vale la probabilità "intorno ad x " in rapporto all'ampiezza di tale "intorno". Il termine funzione di densità, serve proprio ad evocare quanto è densa la probabilità.

Ogni evento deve essere ricondotto all'unione, negazione o intersezione di intervalli del tipo $(-\infty, x]$.
Abbiamo:

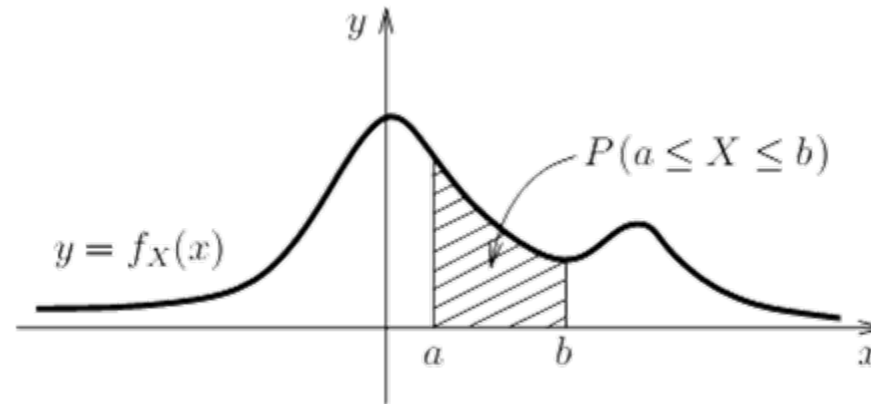
$$P(X \leq x) = \int_{-\infty}^x f_X(x') dx'$$

Oppure

$$P(a \leq X \leq b) = \int_a^b f_X(x) dx$$

Densità di probabilità

La probabilità che una variabile aleatoria continua X assume valori in un intervallo reale $[a, b]$ è data dall'area sottesa al grafico della funzione di densità.



Osservazioni:

La funzione di densità, **non è una probabilità**, è però una funzione legata alla probabilità, perché se voglio calcolare la probabilità che la variabile casuale continua X appartenga ad un intervallo, basta che faccia l'integrale della funzione di densità o:

$$dP = f(x)dx$$

La dimensione della f_X è l'inverso della dimensione di x ovvero $[x]^{-1}$.

Densità di probabilità

Proprietà della funzione di densità:

1. Una funzione di densità non può mai assumere valori negativi, ossia $f_X(x) \geq 0$ ciò assicura che la probabilità che X cada in un qualsiasi intervallo sia non-negativa.
2. L'area totale sottesa alla funzione è uguale a 1, ossia:

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1$$

(in quanto quest'integrale rappresenta la probabilità dell'evento certo). Il fatto che l'integrale della funzione di densità nell'intervallo $(-\infty, +\infty)$ valga 1 viene denominato: **condizione di normalizzazione**. Per avere questo, ossia per avere la convergenza dell'integrale, deve valere $\lim_{x \rightarrow \pm\infty} f_X(x) = 0$.

3. La probabilità che la variabile casuale continua X assuma un particolare valore dell'intervallo è uguale a zero $P(X = x)$. Ciò è dovuto al fatto che **ad un singolo valore corrisponde ad un intervallo di ampiezza nulla**, quindi la corrispondente area è anch'essa nulla. Questo per esempio implica che non ha influenza l'inclusione o meno degli estremi dell'intervallo nel calcolo della probabilità, ossia:

$$P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = P(a < X \leq b)$$

Osservazione: La funzione di densità può essere costante o non costante.

Esempio di densità di probabilità

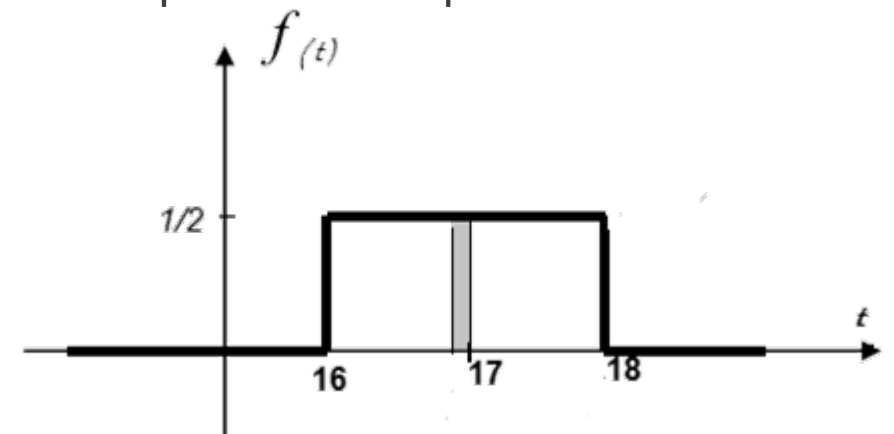
Ci viene preannunciata una **visita** che potrebbe arrivare in un **istante** non meglio precisato, fra le **16:00** e le **18:00**. Per un impegno preso in precedenza **ci dobbiamo assentare dalle 16:45 alle 17:00**. **Qual è la probabilità che la visita arrivi mentre non ci siamo?**

L'istante dell'arrivo della visita è una variabile casuale X . Tutti gli istanti dalle 16:00 alle 18:00 sono equiprobabili. Fuori da questo intervallo la probabilità è zero. Dunque è intuitivo considerare X come una **variabile casuale continua**, la cui densità f_x ha un **valore costante sull'intervallo [16,18]** ed ha il valore zero fuori di questo intervallo. Quanto deve valere la costante? Deve essere tale da soddisfare la relazione:

$$\int_{16}^{18} c dt = 1$$

ovvero l'area del rettangolo con base [16,18] e l'altezza c sia 1. Dunque $2c = 1$ e perciò abbiamo $c = 1/2$. La funzione di densità della variabile X , sarà:

$$f_X(t) = \begin{cases} 0 & \text{per } t/h < 16 \\ \frac{1}{2} & \text{per } 16 \leq t/h \leq 18 \\ 0 & \text{per } t/h > 18 \end{cases}$$



Valore di aspettazione

Le funzioni **distribuzioni di probabilità** o le funzioni **densità di probabilità** sono di fondamentale importanza nei fenomeni casuali in quanto descrivono le caratteristiche globali del fenomeno stesso e la loro conoscenza permette di stabilire a priori la probabilità che una misura cada in un determinato intervallo.

Le **caratteristiche principali** delle distribuzioni di probabilità sono riassunte tramite parametri o **indici di posizione** (che dicono attorno a quali valori è centrata la distribuzione) e di **dispersione** (che sono legati alla “**larghezza**” della distribuzione)

Fra gli indici di posizione, particolare importanza ha il “**valore di aspettazione**” (o valore atteso, o valor medio, o media, o speranza matematica, ...) **della variabile casuale**.

Valore di aspettazione

Data una variabile casuale X , con $x \in [a, b]$ con funzione **densità di probabilità** $f_X(x)$, e una generica funzione $g(x)$, si definisce **valore di aspettazione** della funzione $g(x)$ per la funzione di distribuzione $f_X(x)$ la quantità

$$E[g(x)] = \int_a^b g(x) f_X(x) dx$$

Il valore di aspettazione, che è un numero, indipendente da x , si ottiene quindi integrando i valori della $g(x)$ pesati con la densità di probabilità di $f(x)$.

Il significato del valore di aspettazione come media pesata con la probabilità è ancora più evidente nel caso in cui la variabile casuale sia discreta.

$$E[g(x)] = \sum_{i=1}^n g(x_i) p_i \quad \text{con} \quad \sum_{i=1}^n p_i = 1$$

Valore di aspettazione

Dalla definizione di valore di aspettazione è evidente che:

$$E[a] = a$$

$$E[ag(x)] = aE[g(x)]$$

$$E[a_1g_1(x) + a_2g_2(x)] = a_1E[g_1(x)] + a_2E[g_2(x)]$$

Dove a , a_1 e a_2 sono costanti.

Il **valore di aspettazione della variabile casuale x** , spesso indicato con μ_x , è quindi:

$$\mu_x = E[x] = \int_{\Omega} x f_X(x) dx \quad \text{o} \quad \mu_x = E[x] = \sum_{i=1}^n x_i p_i$$

L'espressione del valore di aspettazione è molto simile alla definizione di centro di massa

Altri indici di posizione

Oltre al valore di aspettazione μ_x vengono introdotti altri valori 'di posizione':

- il 'valore modale' o **moda**: valore di x per cui $f_X(x)$ è massima;
- **i quantili** che specificano la probabilità che la variabile casuale assuma valori in determinati intervalli. I **quantili di ordine $\alpha(x)$** , dove $0 \leq \alpha \leq 1$, è il valore di x per cui

$$\alpha = F(x) = \int_{-\infty}^x f_X(x) dx$$

- il 'valore mediano' o **mediana**: valore di x per il quale $F(x) = \int_{-\infty}^x f_X(x) dx = 0.5$ (quantile $x_{0.5}$);
- i quartili (quantile $x_{0.25}$ e $x_{0.75}$), ovvero i valori di x per cui $F(x) = \int_{-\infty}^x f_X(x) dx = 0.25$ o $F(x) = \int_{-\infty}^x f_X(x) dx = 0.75$

Nel caso di funzioni di distribuzioni simmetriche **valore di aspettazione, valore modale e valore mediano coincidono.**

Varianza e deviazione standard

Tra i possibili indici di dispersione, il più usato è la “varianza” della variabile casuale x , in genere indicata con σ_x^2 o $V(x)$

$$\sigma_x^2 = E[(x - \mu_x)^2] = \int_{\Omega} (x - \mu_x)^2 f(x) dx$$

Viene inoltre definita la **deviazione standard** σ_x come la radice quadrata della varianza $\sigma_x = \sqrt{\sigma_x^2}$. La deviazione standard ha le stesse dimensioni della variabile casuale x ed è un indice della larghezza della dispersione dei valori attorno al valore di aspettazione μ_x .

La varianza gode di alcune proprietà notevoli:

- Può essere scritta come: $\sigma_x^2 = E[x^2] - E[x]^2$
- Soddisfa alla disuguaglianza di Čebyšëv: $P(|x - \mu_x| > \lambda\sigma_x) \leq \frac{1}{\lambda^2}$

Ad esempio, la probabilità che x non appartenga all'intervallo $(\mu_x - 2\sigma_x, \mu_x + 2\sigma_x) \leq 1/4 = 0.25$ (o viceversa la probabilità che x appartenga all'intervallo $(\mu_x - 2\sigma_x, \mu_x + 2\sigma_x) > 0.75$) indipendentemente dalla funzione di distribuzione di x .

Momenti di funzioni di distribuzione

Valore di aspettazione e varianza di una variabile casuale sono due “**momenti**” della funzione densità di probabilità $f(x)$.

I momenti rispetto all'origine e al valore di aspettazione di x sono rispettivamente i valori di aspettazione (quindi quantità numeriche) delle potenze di x e di $(x - \mu_x)$. **Il momento algebrico di ordine k (o k -esimo momento algebrico) e il momento centrale di ordine k sono definiti come:**

$$\mu_k^* = E[x^k] \quad e \quad \mu_k = E[(x - \mu_x)^k]$$

Il momento di ordine 0 vale 1 nel primo caso ($x^0 = 1$, e quindi ho la sola condizione di normalizzazione) e nel secondo caso ($(x - \mu_x)^0 = 1$).

Si vede in modo altrettanto semplice che $\mu_1^* = \mu_x$, $\mu_1 = 0$, $\mu_2 = \sigma_x^2$. Il momento centrale μ_3 è invece legato alla asimmetria della funzione densità di probabilità, ed è uguale a zero per funzioni simmetriche.

I momenti caratterizzano completamente le funzioni di distribuzione, e si può dimostrare sotto ipotesi non particolarmente restrittiva, **che due funzione densità di probabilità con gli stessi momenti coincidono.**

Distribuzioni di probabilità e funzioni di distribuzione

In alcuni casi, facendo delle ipotesi sul fenomeno statistico considerato, introducendo cioè un modello probabilistico, è possibile ricavare le distribuzioni di probabilità o le funzioni di distribuzione per le variabili casuali che intervengono nel fenomeno.

Le probabilità ottenute possono essere poi confrontate con risultati sperimentali da ripetizioni della stessa misura, fornendo informazioni sulla validità del modello probabilistico.

Alcune di queste funzione densità di probabilità sono particolarmente importanti e diffuse.

Vediamo quelle più legate agli argomenti trattati durante il corso.

Distribuzione uniforme discreta

Una **variabile casuale X discreta ha distribuzione uniforme se tutti i valori che può assumere sono equiprobabili**. Se x può assumere i valori $x_i, i = 1, \dots, n$, le probabilità relative ai diversi valori dovranno essere uguali: $p_i = p$ per ogni i . Usando la condizione di normalizzazione, si ottiene **$p = 1/n$** .

Il valore di aspettazione di x

$$\mu_x = E[x] = \sum_{i=1}^n p x_i = \frac{1}{n} \sum_{i=1}^n x_i$$

è la media aritmetica dei possibili valori, e la varianza:

$$\sigma_x^2 = E[(x - \mu_x)^2] = \sum_{i=1}^n p (x_i - \mu_x)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \mu_x^2$$

è la media dei quadrati degli scarti dal valore di aspettazione.

Esempi

Come già detto, nel lancio di una moneta, i possibili valori di X sono solo due: 1 o 2 (T o C) e a priori ci si aspetta che siano equiprobabili, e che quindi x abbia una distribuzione uniforme con $p_i = 0.5$. Allora

- Valore di aspettazione è $\mu_x = 0.5 \times \sum_{i=1}^2 x_i = 0.5 \times 1 + 0.5 \times 2 = 1.5$
- Deviazione standard vale $\sigma_x^2 = 0.5 \times \sum_{i=1}^2 x_i^2 - 1.5^2 = 0.5 + 2 - 0.25 = 2.25$ o $\sigma_x = 1.5$.

Nel lancio di un dado, i possibili valori di X sono sei: 1, 2, 3, \dots , 6 e di nuovo a priori ci si aspetta che siano equiprobabili, e che quindi X abbia una distribuzione uniforme con $p_i = 1/6$. In questo caso abbiamo:

- Valore di aspettazione è $\mu_x = 3.5$
- Deviazione standard vale $\sigma_x^2 \cong 2.9$ o $\sigma_x \cong 1.7$.

Distribuzione uniforme continua

Una variabile casuale continua definita nell'intervallo (a, b) ha **funzione di distribuzione uniforme** se in (a, b) è $f(x) = \text{cost.}$, in modo che la **probabilità che x assuma valori in intervalli diversi ma di stessa ampiezza contenuti in (a, b) sia la stessa**. Per la condizione di normalizzazione, deve essere $h = 1/(b - a)$, e quindi

$$f(x) = \frac{1}{b - a}$$

il valore di aspettazione è il punto medio dell'intervallo

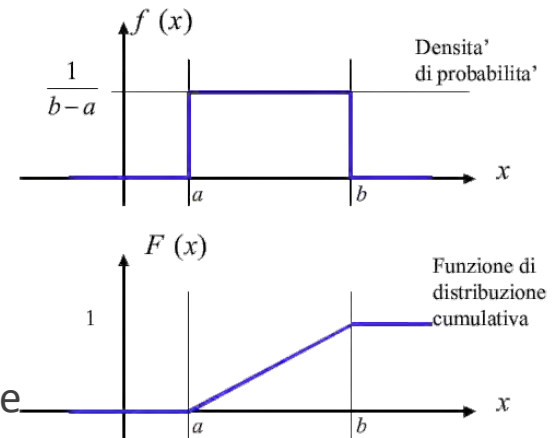
$$\mu_x = \int_a^b x f(x) dx = \frac{1}{b - a} \int_a^b x dx = \frac{1}{b - a} \frac{b^2 - a^2}{2} = \frac{a + b}{2}$$

La varianza può essere calcolata allo stesso modo con conti leggermente più lunghi e si ottiene

$$\sigma_x^2 = E[x^2] - E[x]^2 = \int_a^b x^2 f(x) dx - \mu_x^2 = \frac{1}{b - a} \int_a^b x^2 dx - \left(\frac{a + b}{2}\right)^2 = \frac{(b - a)^2}{12}$$

La deviazione standard vale quindi $\sigma_x = \frac{b - a}{2\sqrt{3}}$. La probabilità che l'intervallo $(\mu_x - \sigma_x, \mu_x + \sigma_x)$ contenga il valore vero è

$$P(\mu_x - \sigma_x < x < \mu_x + \sigma_x) = 2 \frac{1}{b - a} \frac{b - a}{2\sqrt{3}} = \frac{1}{\sqrt{3}} = 0.577$$



* Il tempo di attesa di un evento periodico (come l'arrivo di un autobus a una certa fermata in condizioni di traffico costante) idealmente misurato con uno strumento di sensibilità infinita ha una distribuzione uniforme.

Risoluzione di lettura e distribuzione uniforme

La **distribuzione uniforme** viene utilizzata quando mancano informazioni sulla distribuzione della variabile casuale all'interno di un intervallo **e può essere utilizzata anche nel caso di una misura di una grandezza fisica con uno strumento digitale.**

Se ad esempio si misura una massa M , ottenendo un valore $m = 12.01 \text{ g}$ con un'incertezza dovuta alla risoluzione di lettura di 0.01 g ed errore di risoluzione di $\Delta m = 0.005 \text{ g}$, si dirà che $M = 12.010 \pm 0.005 \text{ g}$, intendendo che l'intervallo $(12.005 \text{ g}, 12.015 \text{ g})$ include il valore "vero" di M , senza però avere indicazioni su eventuali valori più probabili all'interno dell'intervallo.

Volendo associare un'incertezza statistica alla misura, cosa in certi casi particolarmente utile, si può ipotizzare che M abbia una distribuzione uniforme (benché abbia un valore ben definito) nell'intervallo $(12.005 \text{ g}, 12.015 \text{ g})$ e quindi valore di aspettazione $\mu_M = 12.010 \text{ g}$ e deviazione standard

$$\sigma_M = \frac{\Delta m}{\sqrt{3}} = 0.003 \text{ g}$$

Si potrà quindi dire che $M = 12.010 \pm 0.003 \text{ g}$ sapendo che il contenuto di probabilità di questa affermazione è del 58%.

Distribuzione binomiale (o di Bernoulli)

Consideriamo **un fenomeno statistico che possa verificarsi soltanto secondo due modalità mutuamente esclusive**: A (favorevole, successo) e \bar{A} (sfavorevole, insuccesso).

Supponiamo che la probabilità che il fenomeno si verifichi secondo la modalità A sia p , e quindi che la probabilità che il fenomeno si verifichi secondo la modalità \bar{A} sia $q = 1 - p$. Assumendo che gli eventi siano tutti indipendenti, la probabilità che su n eventi ci siano k successi (o che, se il fenomeno viene misurato n volte, k volte si verifichi secondo la modalità A) è data da

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

Dove $\binom{n}{k} = C_{n,k}$ è il numero delle combinazioni semplici di n elementi presi k alla volta, con elementi non ripetuti.

Distribuzione binomiale

Questa distribuzione può essere ottenuta tenendo conto che:

- La probabilità che i **primi k** eventi siano **favorevoli** e i rimanenti $n - k$ siano **sfavorevoli**, essendo gli eventi indipendenti, è il prodotto di k volte la probabilità che il singolo evento sia favorevole e di $n - k$ volte la probabilità che il singolo evento sia sfavorevole, cioè $P_k = p^k q^{n-k}$.
- Per qualsiasi altra sequenza di k eventi favorevoli e $n - k$ eventi sfavorevoli (ad esempio il primo favorevole, il secondo sfavorevole, poi $k - 1$ favorevoli e infine $n - k - 1$ sfavorevoli, ecc.), la probabilità è la stessa: $P_k = p^k q^{n-k}$.
- Per ottenere la probabilità cercata bisogna quindi moltiplicare P_k per il numero di **tutte le possibili combinazioni** in cui si hanno k eventi favorevoli e $n - k$ eventi sfavorevoli, cioè per

$$\binom{n}{k} = \frac{n!}{k! (n - k)!} = \frac{n(n - 1)(n - 2) \cdots (n - k + 1)}{k!} = \frac{n(n - 1)(n - 2) \cdots (n - k + 1)}{k(k - 1)(k - 2) \cdots 1}$$

$$\binom{n}{0} = \frac{n!}{0! n!} = 1$$

$$\binom{n}{n} = \frac{n!}{n! 0!} = 1$$

Cenni di calcolo combinatorio

Immaginiamo di avere un evento A che abbia una probabilità p di essere favorevole ed una probabilità $q = 1 - p$ di essere sfavorevole. Possiamo chiederci quale sia la **probabilità di avere un certo numero r di eventi favorevoli in n tentativi**

- Nel caso in cui ci chiediamo la probabilità di n successi (insuccessi) la risposta è semplice, perché posso ottenere tale risultato p^n (q^n) in un solo modo. Con una successione di successi (insuccessi).
- Nel caso in cui mi chiedo la probabilità di avere r eventi che si presentano in modo favorevole (e di conseguenza $n - r$) in modo contrario la risposta si complica perché tale risultato può essere ottenuto in più modi. Posso avere, ad esempio i primi r eventi tutti positivi, ed i secondi $n - r$ tutti negativi, o viceversa o varie combinazioni. In ogni caso la probabilità di una singola combinazione r favorevoli ed $n - r$ sfavorevoli sarà $p^r q^{n-r}$.
- Il problema si riduce quindi a trovare il numero di combinazioni che producono questo risultato. Tale numero viene generalmente designato come **combinazione semplice di r elementi in classe n** ed indicato con $C_{n,r}$ o $\binom{n}{r}$, n su r , con questa seconda notazione denominata come «**coefficiente binomiale**». Il calcolo di questo coefficiente è oggetto della parte dell'algebra chiamata calcolo combinatorio.

Cenni di calcolo combinatorio

Supponiamo avere di n oggetti che possiamo sistemare in r spazi, con $r \leq n$, con la possibilità di mettere un solo oggetto in ognuno di questi spazi.

- Nel primo spazio possiamo sistemare 1 tra gli oggetti e quindi abbiamo n possibili combinazioni.
- Nel secondo spazio, una volta sistemato il primo oggetto, rimaniamo con $n - 1$ possibili combinazioni, perciò vi sono $n(n - 1)$ possibilità per riempire i primi due spazi
- Ci sono $(n - 2)$ modi per riempire il terzo spazio e così via fino agli $(n - r + 1)$ modi che ho per riempire l'ultimo spazio degli r iniziali
- Indicando quindi con $D_{n,r}$ il numero di modi con cui sistemare gli n oggetti negli r spazi si hanno le cosiddette **disposizioni semplici di n elementi r a r** . Questa sarà pari a $D_{n,r} = \prod_{k=0}^{r-1} n - k$. Se poi moltiplico ambe i membri per il fattoriale $(n - r)!$, ovvero $(n - r)! D_{n,r} = (n - r)! \prod_{k=0}^{r-1} n - k = n!$ Ottengo:

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

- Ricordiamo la definizione di fattoriale:

$$n! = \prod_{k=1}^n k = 1 \cdot 2 \cdot 3 \cdots n$$

Vale $0! = 1$ perché posso sistemare solo in un modo 0 elementi (inoltre, per ricorrenza poiché $1! = 1$, e visto che $n! = n \cdot (n - 1)!$, ho che $1 = 1! = 1 \cdot 0! \Rightarrow 0! = 1$)

Cenni di calcolo combinatorio

Quindi r oggetti in r spazi possono apparire in parecchie disposizioni. Infatti posso prendere gli elementi A, B e C in tre spazi come (ABC), (ACB), (BAC), (BCA), (CAB), (CBA). Usando la definizione precedente nel il caso particolare $n = r$ ho infatti: $D_{r,r} = r!$

In effetti le **permutazioni semplici** (senza ripetizioni) di r oggetti, indicate con P_r non sono altro che disposizioni semplici con $n = r$; **le permutazioni di un insieme di r oggetti sono le disposizioni semplici di tali oggetti di lunghezza r .**

Una permutazione di un insieme di oggetti è una presentazione ordinata, cioè una sequenza, dei suoi elementi nella quale ogni oggetto viene presentato una ed una sola volta.

Con questo apparato possiamo tornare alle nostre combinazioni semplici;

Si chiama **combinazione semplice** una presentazione di elementi di un insieme nella quale **non ha importanza l'ordine dei componenti**, ovvero, non si può ripetere lo stesso elemento più volte solo variando l'oggetto al suo interno.

Cenni di calcolo combinatorio

All'interno di ognuna delle combinazioni di $C_{n,r}$ semplici ho quindi $r!$ disposizioni di $D_{n,r}$ possibili ovvero:

$$D_{n,r} = r! C_{n,r}$$

Ed a partire da questo possiamo rispondere al numero di combinazioni ovvero

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

Ed infine rispondere alla nostra domanda iniziale: qual è la **probabilità $P(n, r)$** di avere un certo numero r di **eventi favorevoli in n tentativi**.

Da quanto visto abbiamo:

$$P(n, r) = \binom{n}{r} p^r q^{n-r} = \frac{n!}{r! (n-r)!} p^r q^{n-r}$$

Breve cenno sui fattoriali 1

Fattoriale di un numero negativo:

$$n! = \prod_{i=0}^n (n - i)$$

$$(c)^n n! = \prod_{i=0}^n c(n - i)$$

$$(-1)^n n! = \prod_{i=0}^n -(n - i) = (-n)!$$

Fattoriale di un numero reale positivo (Eulero 1730, introduce la funzione Pi Π):

$$x! = \Pi(x) = \int_0^1 (-\ln q)^x dq = \int_0^\infty t^x e^{-t} dt \quad (i. e. \text{ponendo } t = -\ln q)$$

Con la proprietà

$$\Pi(x + 1) = (x + 1)\Pi(x)$$

Breve cenno sui fattoriali 2

Sempre Eulero estende nel 1768 la definizione a tutti i numeri reali introducendo la funzione Gamma Γ , con uno spostamento di una unità della variabile:

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

La funzione Γ è legata alla funzione Π dalla relazione seguente:

$$\Pi(x) = \Gamma(x + 1) = x!$$

Vale:

$$z\Gamma(z) = \Gamma(z + 1)$$

Da questa, visto che $\Gamma(1) = 1$ ottengo:

$$\Gamma(n + 1) = n!$$

In questo modo sono definite le combinazioni semplici:

$$\binom{z}{k} = \frac{\Gamma(z + 1)}{k! \Gamma(z - k + 1)}$$

Importanti relazioni tra le funzioni Gamma sono le seguenti (Formola di riflessione di Eulero):

$$\Gamma(z)\Gamma(1 - z) = \frac{\pi}{\sin \pi z} \quad \Gamma(z)\Gamma(-z) = -\frac{\pi}{\sin \pi z}$$

Inciso: la notazione ! per fattoriale fu introdotta solo nel 1808 da Kramp; **Inciso 2:** successivamente è stato esteso anche ai numeri complessi

Inciso3: la notazione Gamma Γ fu introdotta dal Laplace

Breve cenno sui fattoriali 3

Importanti relazioni tra le funzioni Gamma sono le seguenti:

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z} \quad \Gamma(z)\Gamma(-z) = -\frac{\pi}{\sin \pi z}$$

La prima relazione permette di calcolare $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. Infatti vale:

$$\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) = \frac{\pi}{\sin \frac{\pi}{2}} = \pi$$

La seconda permette di calcolare $\Gamma\left(-\frac{1}{2}\right) = -\sqrt{\pi}$. Infatti vale:

$$\Gamma\left(\frac{1}{2}\right)\Gamma\left(-\frac{1}{2}\right) = -\frac{\pi}{\sin \frac{\pi}{2}} = -\pi$$

Ed infine:

$$\Gamma\left(n + \frac{1}{2}\right) = \left(n - \frac{1}{2}\right)\Gamma\left(n - \frac{1}{2}\right) = \left(n - \frac{1}{2}\right)\left(n - \frac{3}{2}\right)\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

Che posso riscrivere come:

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n-1)(2n-3)\cdots 1}{2^n}\Gamma\left(\frac{1}{2}\right) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}$$

Distribuzione binomiale (o di Bernoulli)

L'insieme delle probabilità $P(k; n, p)$ costituisce la distribuzione binomiale (o di Bernoulli) per la variabile casuale discreta k ; questa distribuzione di probabilità dipende da **due soli parametri**: n , il numero totale di eventi, e p , la probabilità che un evento sia favorevole.

- La distribuzione binomiale è normalizzata (**ricordiamo infatti il binomio di Newton**)

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

$$\sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1$$

- Il valore di aspettazione di k è $E[k] = np$, infatti

$$p \frac{\partial}{\partial p} (p + q)^n = p \frac{\partial}{\partial p} \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \Rightarrow np(p + q)^{n-1} = \sum_{k=0}^n k \binom{n}{k} p p^{k-1} q^{n-k}$$

$$np = \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k}$$

Varianza della binomiale

- La varianza è $\sigma_k^2 = npq$ infatti:

$$p^2 \frac{\partial^2}{\partial p^2} (p + q)^n = p^2 \frac{\partial^2}{\partial p^2} \sum_{k=0}^n \binom{n}{k} p^k q^{n-k}$$
$$p^2 n(n-1)(p+q)^{n-2} = \sum_{k=0}^n k(k-1) \binom{n}{k} p^2 p^{k-2} q^{n-k}$$
$$p^2 n(n-1) + \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k} = p^2 n(n-1) + np = \sum_{k=0}^n k^2 \binom{n}{k} p^k q^{n-k} = E[k^2]$$

$$\sigma_k^2 = E[k^2] - E[k]^2 = p^2 n^2 - p^2 n + np - p^2 n^2 = np - p^2 n = np(1-p)$$

da cui

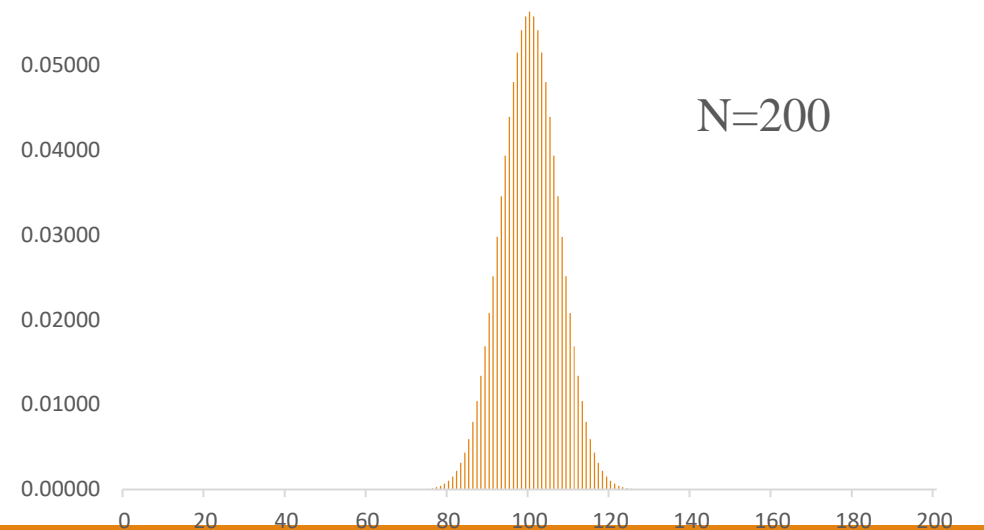
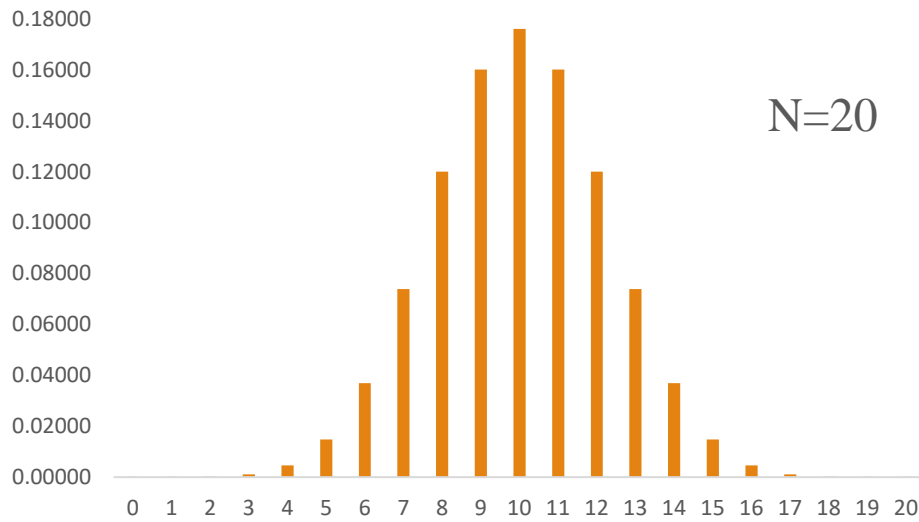
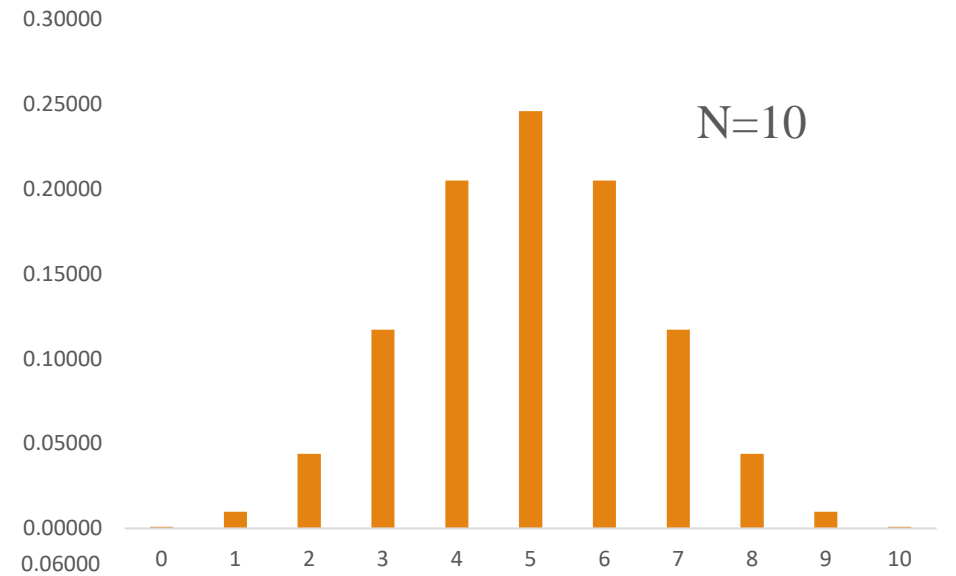
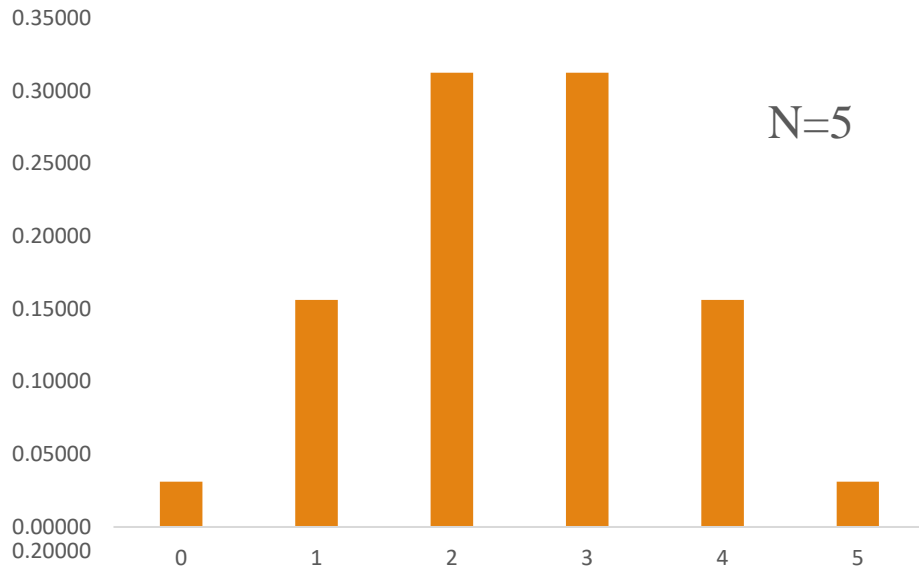
$$\sigma_k^2 = E[k^2] - E[k]^2 = npq$$

Applicazioni

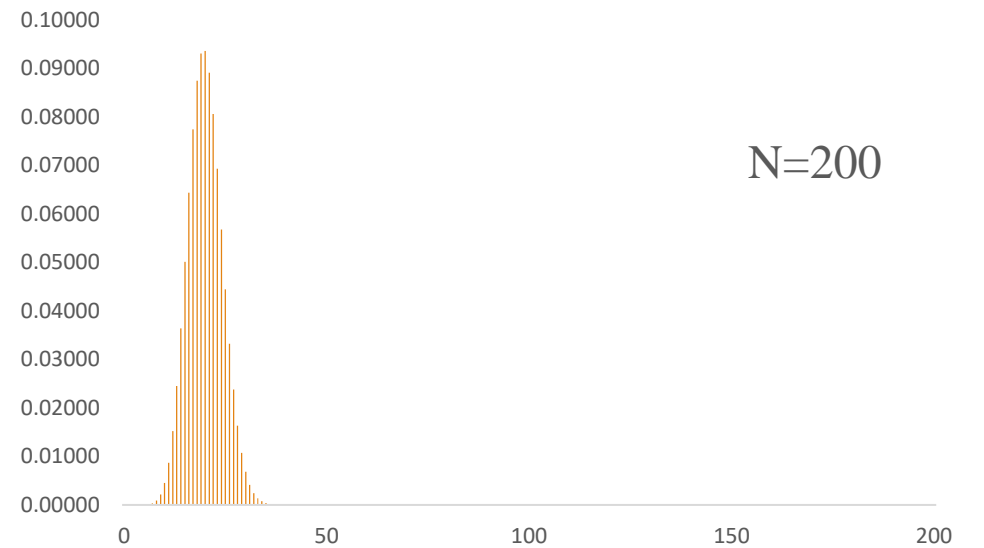
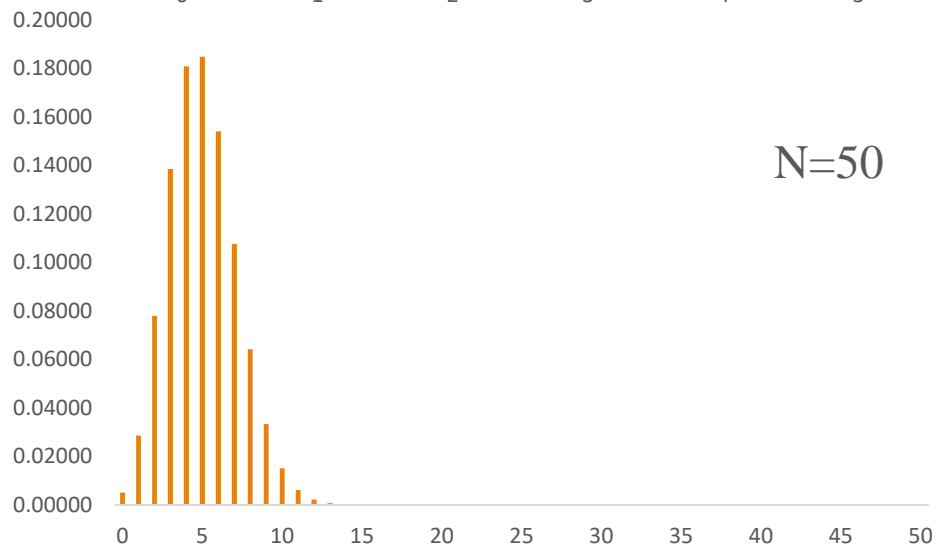
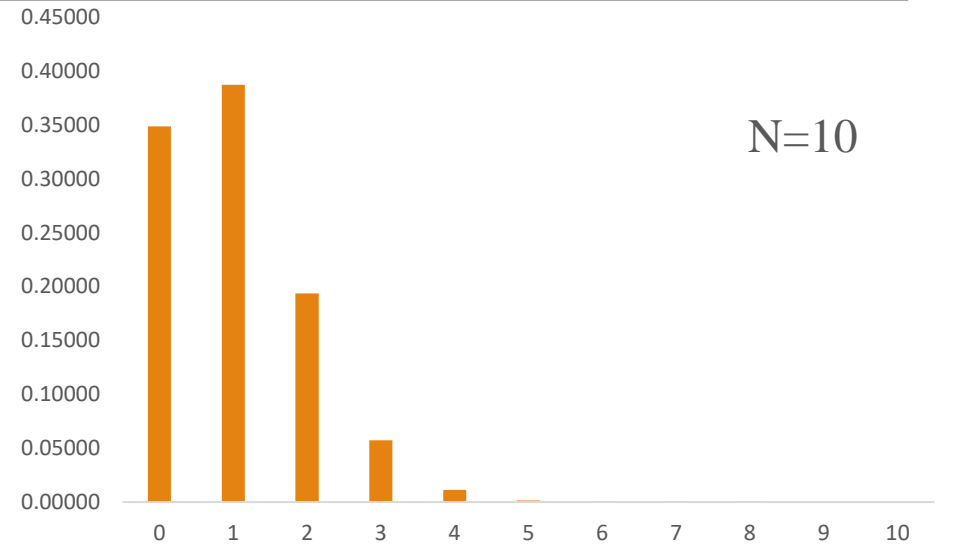
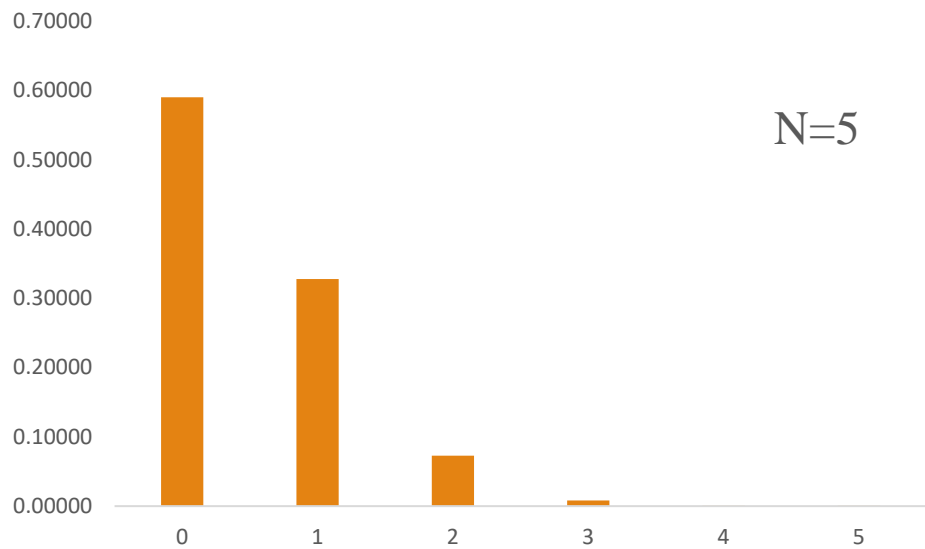
La distribuzione binomiale ha moltissime applicazioni, dal lancio del dado (ad es probabilità di ottenere in N lanci k volte il numero 4) al numero di eventi previsto nei diversi intervalli di un istogramma trattata nell'esempio.

La sua importanza sta anche nel fatto che, al limite per $n \rightarrow \infty$ permette di ottenere la funzione di distribuzione di Gauss se p rimane costante e la distribuzione di Poisson se np rimane costante, entrambe usatissime in fisica.

Esempi ($p = 0.5$)



Esempi ($p = 0.1$)



Distribuzione binomiale e istogrammi

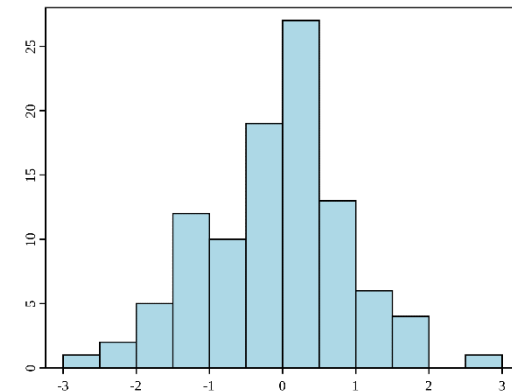
Consideriamo un **particolare intervallo i di un istogramma** e consideriamo **favorevole** un evento **se il valore cade nell'intervallo** scelto. Su un totale di n eventi indipendenti, la probabilità di averne n_i in quell'intervallo è data dalla distribuzione binomiale. Il numero di eventi attesi nell'intervallo considerato è $E[n_i] = np_i$, dove p_i è la probabilità che il singolo evento cada nell'intervallo considerato.

La varianza di n_i è $\sigma_{n_i}^2 = np_i(1 - p_i) \approx n_i(1 - n_i/n)$ e quindi la deviazione standard è $\sigma_{n_i} = \sqrt{n_i(1 - n_i/n)}$

La probabilità p_i di un evento favorevole può non essere nota a priori (può non essere nota la distribuzione di probabilità corrispondente all'istogramma). Una volta costruito l'istogramma, però, si può prendere come valore stimato di p_i il valore $\hat{p}_i = n_i^m/n$ dove n è il numero totale di eventi e n_i^m è il numero misurato di eventi nell'intervallo considerato in n prove.

La varianza del numero di eventi nell'intervallo considerato diventa $\sigma_{n_i}^2 = n_i^m(1 - n_i^m/n)$. Nel caso $p \rightarrow 0$, che corrisponde a un numero elevato di intervalli, si ha $\sigma_{n_i} = \sqrt{n_i^m}$.

L'incertezza assoluta aumenta con n_i , e quindi con n , ma quella relativa diminuisce.



Distribuzione di Poisson

La distribuzione di Poisson (o degli eventi rari) è una funzione di distribuzione fondamentale nella fisica nucleare, ed è un caso limite della distribuzione binomiale.

Si voglia determinare la probabilità di osservare k eventi in un dato intervallo di tempo t , nelle seguenti ipotesi:

- a) La presenza/assenza dell'evento al tempo t non dipende dalla storia degli eventi prima del tempo t
- b) La probabilità di un singolo evento aumenta proporzionalmente alla lunghezza del tempo considerato, secondo una costante μ con dimensioni $[t]^{-1}$ intrinseca del fenomeno considerato
- c) La probabilità di due eventi contemporanei è 0

Allora la distribuzione degli eventi segue la distribuzione di Poisson:

$$P_k(t) = \frac{(\mu t)^k}{k!} e^{-\mu t} \quad \text{o} \quad P_k(m) = \frac{m^k}{k!} e^{-m}$$

$P_k(m)$ indica la probabilità di avere k eventi quando in media se ne hanno m

Distribuzione di Poisson

La distribuzione di Poisson:

$$P_k(t) = \frac{(\mu t)^k}{k!} e^{-\mu t} \quad \text{o} \quad P_k(m) = \frac{m^k}{k!} e^{-m}$$

E' normalizzata:

$$\sum_{k=0}^{\infty} P_k(m) = \sum_{k=0}^{\infty} \frac{m^k}{k!} e^{-m} = \frac{1}{e^m} \sum_{k=0}^{\infty} \frac{m^k}{k!} = 1, \text{ come vedo sviluppando } e^m = \sum_{l=0}^{\infty} \frac{m^l}{l!}$$

Per m sufficientemente grandi (> 5) la distribuzione di Poisson tende ad una distribuzione gaussiana.

Ha il seguente valore di aspettazione $E[k] = \hat{k} = \mu t$ e varianza $E[(k - \hat{k})^2] = \sigma^2 = \mu t$ o

$$E[k] = \hat{k} = m \text{ e } E[(k - \hat{k})^2] = \sigma^2 = m$$

$$\sum_{k=0}^{\infty} k P_k(m) = \sum_{k=0}^{\infty} k \frac{m^k}{k!} e^{-m} = m \sum_{k=1}^{\infty} \frac{m^{k-1}}{(k-1)!} e^{-m} = m$$

Distribuzione di Poisson

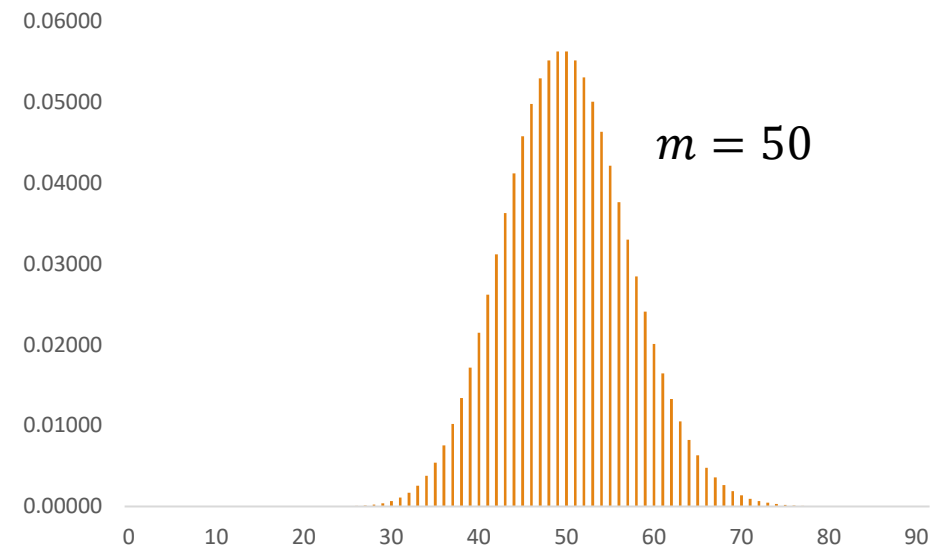
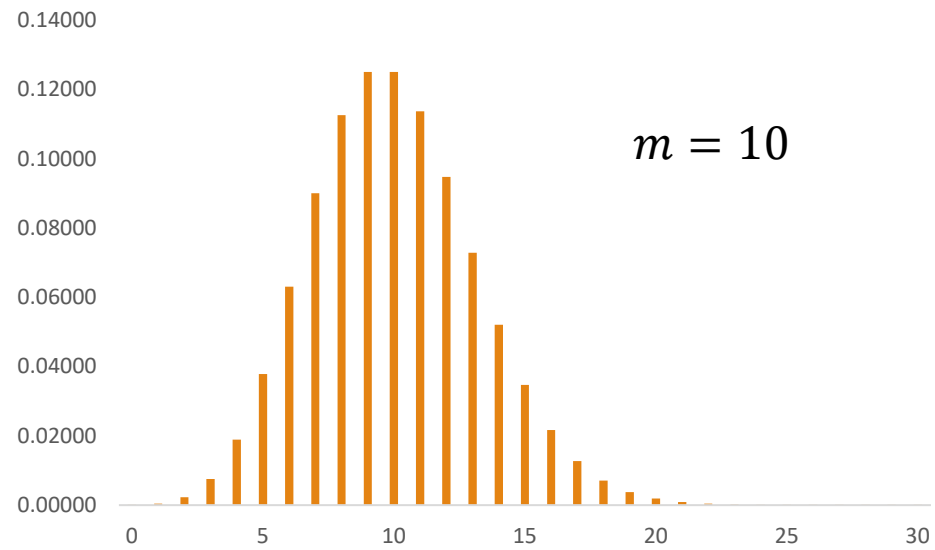
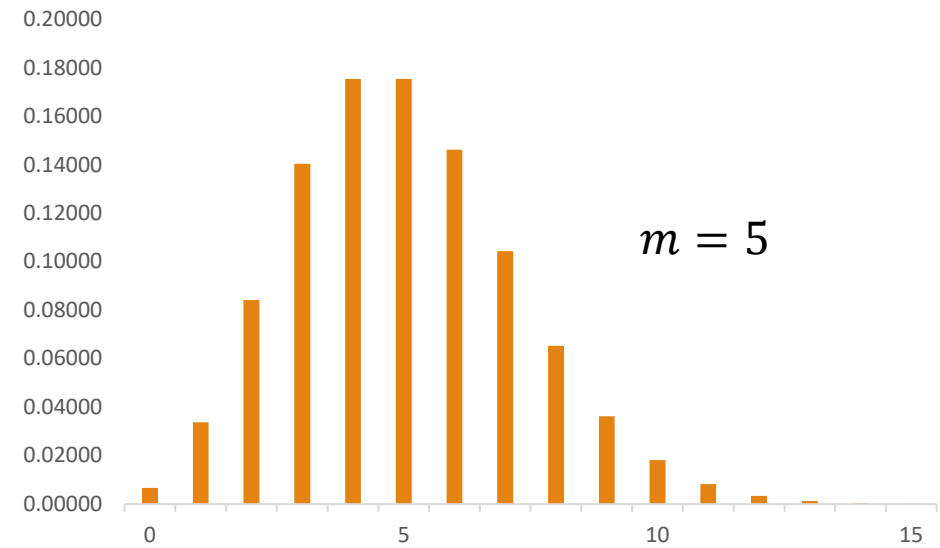
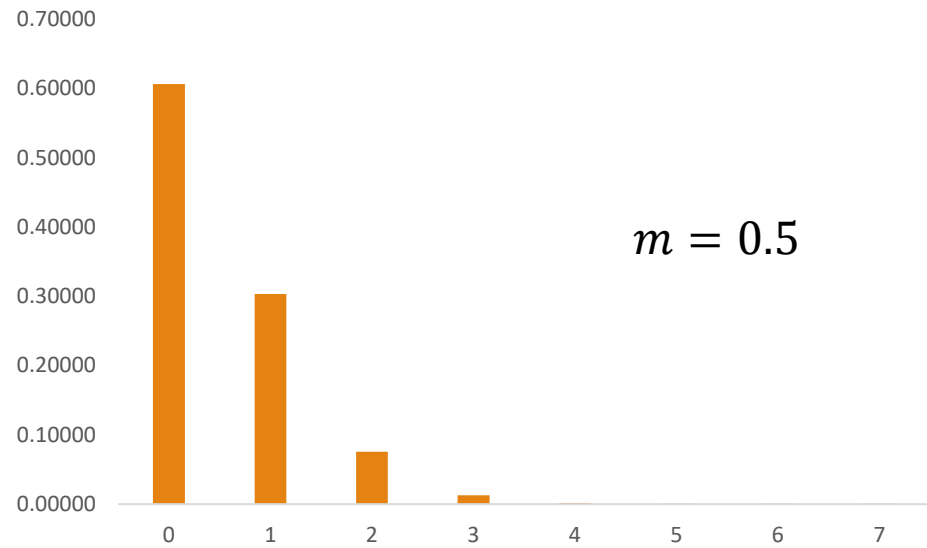
$$E[(k - \hat{k})^2] = \sigma^2 = m:$$

$$\begin{aligned} E[k^2] &= \sum_{k=0}^{\infty} k^2 P_k(m) = \sum_{k=0}^{\infty} k^2 \frac{m^k}{k!} e^{-m} = me^{-m} \sum_{k=1}^{\infty} k \frac{m^{k-1}}{(k-1)!} = \\ &= me^{-m} \left(\sum_{k=1}^{\infty} (k-1) \frac{m^{k-1}}{(k-1)!} + \sum_{k=1}^{\infty} \frac{m^{k-1}}{(k-1)!} \right) = \\ &= me^{-m} \left(m \sum_{k=2}^{\infty} \frac{m^{k-2}}{(k-2)!} + \sum_{k=1}^{\infty} \frac{m^{k-1}}{(k-1)!} \right) = me^{-m} \left(m \sum_{j=0}^{\infty} \frac{m^j}{j!} + \sum_{i=0}^{\infty} \frac{m^i}{i!} \right) = \\ &= me^{-m} (me^m + e^m) = m^2 + m \end{aligned}$$

Da cui:

$$\sigma^2 = E[(k - \hat{k})^2] = E[k^2] - E[k]^2 = m^2 + m - m^2 = m$$

Esempi



Distribuzione di Poisson: P_0

Calcoliamo la probabilità di avere 0 eventi nell'intervallo di tempo $(0, t + dt]$. La probabilità di avere un evento in dt è:

$$dP_1 = \mu dt \quad \text{quindi} \quad dP_0 = 1 - \mu dt$$

Per cui, data $P_0(t)$, posso scrivere la probabilità di avere 0 eventi in $(0, t + dt]$ come il prodotto della probabilità di avere 0 eventi in $(0, t]$ e di non avere eventi in $(t, t + dt]$:

$$P_0(t + dt) = P_0(t)(1 - \mu dt)$$

Posso quindi scrivere:

$$\frac{dP_0}{dt} = \frac{P_0(t + dt) - P_0(t)}{dt} = \frac{P_0(t)(1 - \mu dt) - P_0(t)}{dt} = -\mu P_0(t)$$

Ovvero:

$$\frac{dP_0}{P_0} = -\mu dt$$

Da cui ottengo (con la condizione al contorno $P_0(0) = 1$):

$$P_0(t) = e^{-\mu t}$$

Distribuzione di Poisson: P_1

La probabilità di avere 1 evento nell'intervallo di tempo $(0, t + dt]$ si calcola con una procedura simile. Ricordando che $dP_1 = \mu dt$, la probabilità di avere 1 evento in $(0, t + dt]$ è data dalla somma della probabilità di avere 1 evento in $(0, t]$ e nessun evento in $(t, t + dt]$ e dalla probabilità di avere 0 eventi in $(0, t]$ ed 1 evento in $(t, t + dt]$:

$$P_1(t + dt) = P_1(t)(1 - \mu dt) + P_0(t)\mu dt$$

Da cui:

$$\frac{P_1(t + dt) - P_1(t)}{dt} = \frac{P_1(t)(1 - \mu dt) + P_0(t)\mu dt - P_1(t)}{dt} = -\mu[P_1(t) - P_0(t)]$$

Ovvero:

$$\frac{dP_1}{dt} = -\mu P_1(t) + \mu P_0(t)$$

Che ha come soluzione:

$$P_1(t) = \mu t e^{-\mu t}$$

Nel caso generico k ho

$$\frac{dP_k}{dt} = -\mu[P_k(t) - P_{k-1}(t)] \quad \text{con soluzione} \quad P_k(t) = \frac{(\mu t)^k}{k!} e^{-\mu t}$$

Origine della distribuzione di Poisson

Partiamo dalla distribuzione di Bernoulli:

$$P(k; n, p) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k! (n-k)!} p^k q^{n-k}$$

E valutiamo il caso in cui il prodotto $np = \text{cost} = m$, sia costante al tendere di $n \rightarrow \infty$. Scriviamo la probabilità come $p = m/n$:

$$P(k; n, p) = \frac{n!}{k! (n-k)!} \left(\frac{m}{n}\right)^k \left(1 - \frac{m}{n}\right)^{n-k}$$

Che possiamo riscrivere come:

$$\begin{aligned} P(k; n, p) &= \frac{m^k}{k!} \frac{n!}{n^k (n-k)!} \left(1 - \frac{m}{n}\right)^n \left(1 - \frac{m}{n}\right)^{-k} \\ &= \frac{m^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \left(1 - \frac{m}{n}\right)^n \left(1 - \frac{m}{n}\right)^{-k} = \frac{m^k}{k!} e^{-m} \end{aligned}$$

Prendendo il limite $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} = 1; \lim_{n \rightarrow \infty} \left(1 - \frac{m}{n}\right)^{-k} = 1; \text{ mentre } \lim_{n \rightarrow \infty} \left(1 - \frac{m}{n}\right)^n = e^{-m} \text{ da } \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$$

Distribuzione di Maxwell-Boltzmann

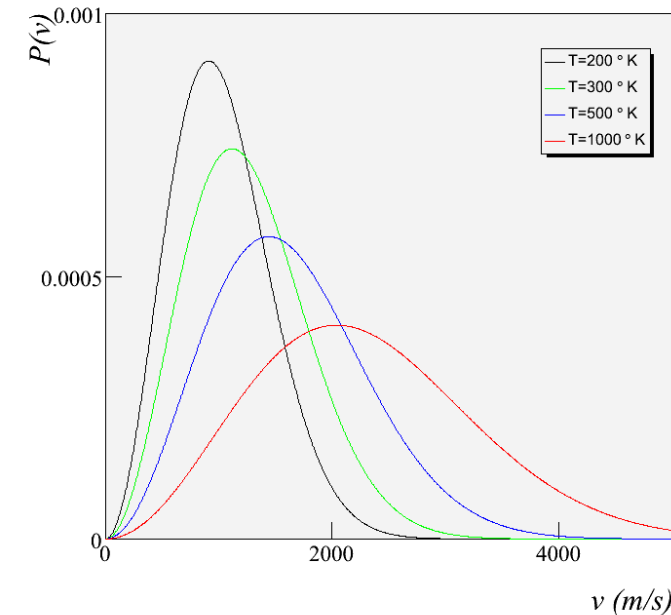
- Prendiamo un gas di peso molare $M = N_A m$ all'equilibrio termodinamico con temperatura T . Ricordiamo inoltre che il valore della costante dei gas $R = N_A k \approx 8.31 \text{ J K}^{-1} \text{ mol}^{-1}$.
- Abbiamo una relazione che lega la velocità quadratica media delle molecole di un gas con la sua temperatura:

$$v_{qm} = \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3RT}{M}}$$

- In realtà le singole molecole hanno una velocità che si discosta da quella media.
- Maxwell nel 1859 trovò la legge di distribuzione di queste velocità: [$f(v)$ -densità di probabilità]:

$$f(v) = 4\pi \left(\frac{M}{2\pi RT} \right)^{3/2} v^2 \exp \left[-\frac{Mv^2}{2RT} \right] = 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} v^2 \exp \left[-\frac{mv^2}{2kT} \right]$$

- All'aumentare della temperatura il picco (valore più probabile della velocità) si sposta verso velocità più elevate.
- Le curve sono normalizzate in modo tale che l'integrale della probabilità sia sempre uguale a 1.
- **La distribuzione fu derivata da Maxwell con un ragionamento ingegnoso...**



Funzione di distribuzione

- Ricaviamo la funzione di distribuzione di Maxwell. Introduciamo dapprima alcune definizioni:

$$\begin{aligned}\rho(\vec{r}) &= (dm/dV) && - \text{densita' del gas} \\ \eta(\vec{r}) &= (\rho(\vec{r})/m) && - \text{densita' di particelle} \\ dn &= \eta(\vec{r}) dV && - \# \text{ di particelle in } dV\end{aligned}$$

- Possiamo ricavare la funzione di distribuzione delle particelle nello spazio stimando:

$$f(\vec{r}) = \frac{\eta(\vec{r})}{n_T}$$

- La probabilità di trovare una particella nell'elemento di volume dV vale $f(\vec{r})dV$
- Vale condizione di normalizzazione che l'integrale della probabilità esteso al volume unitario sia uguale a uno:

$$\int_{V=1} f(\vec{r})dV = 1$$

- Le molecole si muovono in uno spazio delle fasi a 6 dimensioni: 3 dimensioni spaziali e 3 dimensioni per la velocità. In questo contesto non siamo interessati alla distribuzione delle molecole nello spazio, in quanto per ipotesi esse sono distribuite in maniera omogenea, ma ci interessa la loro distribuzione nello spazio delle velocità

-
- Introduciamo quindi una densità di particelle nello spazio delle velocità $\eta(v)$, **in modo analogo a allo spazio tridimensionale**. Di conseguenza il numero di particelle dn che hanno velocità compresa tra v e $v + dv$ **nello spazio tridimensionale delle velocità**, vale:

$$dn(\vec{v}) = \eta(\vec{v})d^3v$$

- Introduciamo ora la funzione di distribuzione delle velocità $f(\vec{v})$ definita nel modo seguente e con la normalizzazione indicata:

$$f(\vec{v}) = \frac{\eta(\vec{v})}{n_T} \qquad \int_{\Omega_v} f(\vec{v})d\vec{v} = 1$$

- Se il **gas non è sottoposto a sollecitazioni esterne**, la distribuzione delle velocità deve essere isotropa, vale a dire che la funzione di distribuzione delle velocità $f(\vec{v})$ deve essere **funzione solo del modulo della velocità** ma non della sua direzione.
- Dato che l'energia cinetica di una molecola vale $mv^2/2$, cioè è anch'essa una funzione del modulo della velocità, allora ne consegue che la funzione di distribuzione delle velocità si può esprimere come una **funzione dell'energia cinetica K** delle molecole.
- Il numero di particelle dn comprese nell'elemento $d\vec{v} = d^3v$ dello spazio delle velocità vale:

$$dn(v) = n_T p(K)dv_x dv_y dv_z$$

che in coordinate polari si può scrivere:

$$dn(v) = n_T p(K)v^2 dv d\Omega = n_T p(K)v^2 dv \sin \theta d\theta d\varphi$$

- Dato che la $p(K)$ non dipende dalle variabili angolari θ e φ , queste due variabili si possono integrare:

$$dn(v) = n_T p(K) v^2 dv \sin \theta d\theta d\varphi = 4\pi n_T p(K) v^2 dv$$

- Consideriamo un singolo urto tra due molecole aventi inizialmente energie K_1 e K_2 e successivamente K_1' e K_2' . La probabilità che ciò accada è proporzionale alla densità $p(K_1)$ e alla densità $p(K_2)$

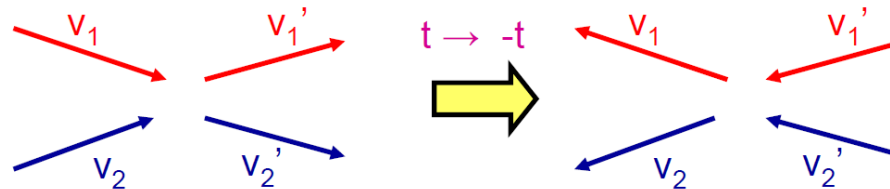
$$P = C p(K_1)p(K_2)$$

dove C è una costante di proporzionalità che dipende dal tipo di interazione che consideriamo. Nel caso di un urto puramente meccanico, essa dipende ad esempio dalle dimensioni delle molecole (sezione d'urto e/o parametro d'impatto).

- In maniera analoga si possono considerare due molecole aventi energie K_1' e K_2' che abbiano dopo l'urto energie K_1 e K_2 . La probabilità che ciò accada è:

$$P' = C' p(K_1')p(K_2')$$

- L'urto elementare tra le due molecole è invariante per inversione temporale.



Ne consegue che le proprietà “dinamiche” dei due urti devono essere le stesse, ovvero le due costanti C e C' devono essere identiche. Se così non fosse i due processi sarebbero distinguibili tra di loro.

-
- Il gas è in equilibrio termodinamico. Ovvero non può aumentare il numero di molecole di energia K'_1 e K'_2 a scapito di quelle di energie K_1 e K_2 e viceversa. Ne consegue quindi che le due probabilità P e P' devono essere le stesse. Poiché anche C e C' sono anche uguali, si ha:

$$p(K_1)p(K_2) = p(K'_1)p(K'_2)$$

- Nell'urto elastico l'energia cinetica delle due molecole si conserva:

$$K_1 + K_2 = K'_1 + K'_2 = \text{Const.}$$

- Se la particella acquista energia x nell'urto, ovvero: $K'_1 = K_1 + x$, la conservazione impone $K'_2 = K_2 - x$, per cui:

$$p(K_1)p(K_2) = p(K_1 + x)p(K_2 - x)$$

- L'unica funzione che soddisfi questa condizione è:

$$p(K) = A \exp[-\beta K] = A \exp\left[-\beta \frac{mv^2}{2}\right]$$

Con $\beta > 0$.

- Quindi:

$$dn(v) = n_T 4\pi p(K) v^2 dv = n_T 4\pi A \exp\left[-\beta \frac{mv^2}{2}\right] v^2 dv$$

Parametri della distribuzione

- Le costanti A e β vanno determinate imponendo delle condizioni al contorno:
 - A) normalizzazione (dove n è il numero di particelle per unità di volume)

$$\int_V dn(v) = \int_0^{\infty} n_T 4\pi A \exp\left[-\beta \frac{mv^2}{2}\right] v^2 dv = n_T$$

- Energia totale del gas (per unità di volume):

$$\int_V K dn(v) = \int_0^{\infty} n_T 4\pi A \frac{mv^2}{2} \exp\left[-\beta \frac{mv^2}{2}\right] v^2 dv = \frac{3}{2} n_T kT$$

Maxwell (...segue...)

- Il risultato delle due condizioni al contorno dà per la densità di probabilità $p(v)$ l'espressione seguente:

$$p(v)dv = 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} v^2 \exp\left[-\frac{mv^2}{2kT}\right] dv$$

questa è la **funzione di distribuzione delle velocità di Maxwell-Boltzmann**.

- La funzione presenta un massimo per il valore: $v_M = \sqrt{2kT/m}$ che indica la velocità più probabile; la possiamo anche riscrivere come:

$$kT = \frac{1}{2} m v_M^2$$

- Utilizzando la velocità più probabile, la funzione di distribuzione di Maxwell si può riscrivere nel modo seguente:

$$p(v)dv = \frac{4}{\sqrt{\pi}} \frac{v^2}{v_M^3} \exp\left[-\frac{v^2}{v_M^2}\right] dv$$

Maxwell (...segue...)

- Usando la funzione di distribuzione di Maxwell-Boltzmann, posso ricavare la velocità media (in modulo) delle molecole:

$$\mu_v = E[v] = \frac{4}{\sqrt{\pi}} \frac{1}{v_M^3} \int_0^{\infty} v^3 \exp\left[-\frac{v^2}{v_M^2}\right] dv = \frac{2}{\sqrt{\pi}} v_M = \sqrt{\frac{8kT}{m}}$$

- Mentre la velocità quadratica media introdotta in precedenza vale:

$$v_{qm} = \sqrt{\langle v^2 \rangle} = \sqrt{\int_0^{\infty} v^2 p(v) dv} = \sqrt{\frac{3kT}{m}}$$

Distribuzione di Maxwell-Boltzmann per v

- Si può osservare che nella funzione di distribuzione di probabilità si possono identificare tre termini:

$$p(v)dv = \left\{ 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} \right\} \left\{ \exp \left[-\frac{mv^2}{2kT} \right] \right\} \{v^2 dv\}$$

- La costante di normalizzazione assicura che l'integrale della $p(v)$ esteso a tutte le velocità possibili sia uguale a 1.
- La statistica di Boltzmann tiene conto della probabilità di occupazione di uno stato avente una data energia in funzione della temperatura del sistema.
- Infine lo spazio delle fasi tiene conto della molteplicità del livello energetico, vale a dire in quanti "modi" è possibile ottenere una configurazione che abbia la stessa energia. Nel nostro caso questo corrisponde al volume di una cortecchia di sfera di raggio interno v e spessore dv . A tutti i punti di questo volume corrisponde la medesima energia $K = mv^2/2$.

Distribuzione di Maxwell-Boltzmann per K

- Abbiamo visto la funzione di distribuzione relativa alle velocità; usando le conoscenze sulla trasformazione della funzione di distribuzione per grandezze derivate possiamo ricavare la distribuzione di Maxwell in funzione dell'energia cinetica $K = mv^2/2$; da:

$$p(v)dv = 4\pi \left(\frac{m}{2\pi RT}\right)^{3/2} \exp\left[-\frac{mv^2}{2RT}\right] v^2 dv$$

Usando la funzione inversa

$$v = \sqrt{\frac{2K}{m}} \quad \text{e} \quad \frac{1}{2} dv^2 = v dv = \frac{1}{m} dK$$

Ottengo:

$$p(K) = 4\pi \left(\frac{1}{2\pi kT}\right)^{3/2} \exp\left[-\frac{K}{kT}\right] \sqrt{K} dK = 2 \left(\frac{1}{kT}\right)^{3/2} \sqrt{\frac{K}{\pi}} \exp\left[-\frac{K}{kT}\right]$$

Funzione di distribuzione di Maxwell-Boltzmann

- La funzione di Maxwell-Boltzmann dipende da un solo parametro a :

$$f(x; a) = \sqrt{\frac{2}{\pi}} \frac{x^2}{a^3} e^{-\frac{x^2}{2a^2}}$$

- Il valore di aspettazione di x vale:

$$E[x] = \mu_x = 2a \sqrt{\frac{2}{\pi}}$$

- Mentre la varianza vale:

$$E[(x - \mu_x)^2] = \sigma^2 = \frac{a^2(2\pi - 8)}{\pi}$$

Funzione di distribuzione di Gauss

Nel 1818 fu chiesto a Gauss di compiere la rilevazione geodetica del Regno di Hannover; tale studio cartografico lo portò a sviluppare la distribuzione gaussiana degli errori, chiamata anche variabile casuale normale, usata per descrivere la misura degli errori casuali.

La funzione di distribuzione di Gauss (o distribuzione normale) è una delle più importanti in statistica. Dipende da 2 parametri, indicati con μ e σ^2 , ed ha la forma

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

La funzione $f(x)$ tende a zero per $x \rightarrow \pm\infty$, è simmetrica rispetto a μ ed ha due punti di flesso per $x = \mu \pm \sigma$.

- La condizione di normalizzazione è rispettata: $\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$
- Il valore di aspettazione vale: $\mu_x = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu$
- La varianza vale: $\sigma_x^2 = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x - \mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2$

Normalizzazione della Gaussiana

$$I_x = \int_{-\infty}^{+\infty} e^{-x^2} dx$$

Se vado nel piano:

$$I^2 = I_x I_y = \int_{-\infty}^{+\infty} e^{-x^2} dx \int_{-\infty}^{+\infty} e^{-y^2} dy = \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy$$

Passo dalle coordinate cartesiane a quelle polari: $x = r \cos \theta$ $y = r \sin \theta$ e $dx dy = |J_\varphi| d\theta dr = r d\theta dr$

Ricordando che $J_\varphi = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$

$$= \int_0^{+\infty} r e^{-r^2} dr \int_0^{2\pi} d\theta = \frac{1}{2} (2\pi) \int_0^{+\infty} e^{-r^2} dr^2 = \pi \Rightarrow I = \sqrt{I^2} = \sqrt{\pi}$$

Significato probabilistico

I due parametri μ e σ^2 hanno quindi proprio il significato di valore di aspettazione e varianza di x . Poiché la distribuzione di Gauss è molto usata, i valori dei suoi integrali su intervalli finiti (e la sua funzione cumulativa o di ripartizione) si trovano tabulati. In particolare si ha che:

$$P(\mu - \sigma < x < \mu + \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mu-\sigma}^{\mu+\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 0.683$$

$$P(\mu - 2\sigma < x < \mu + 2\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mu-2\sigma}^{\mu+2\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 0.954$$

$$P(\mu - 3\sigma < x < \mu + 3\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mu-3\sigma}^{\mu+3\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 0.997$$

$n\sigma$	Prob.
1	0.683
2	0.954
3	0.997
4	$1 - 6.3 \times 10^{-5}$
5	$1 - 5.7 \times 10^{-7}$

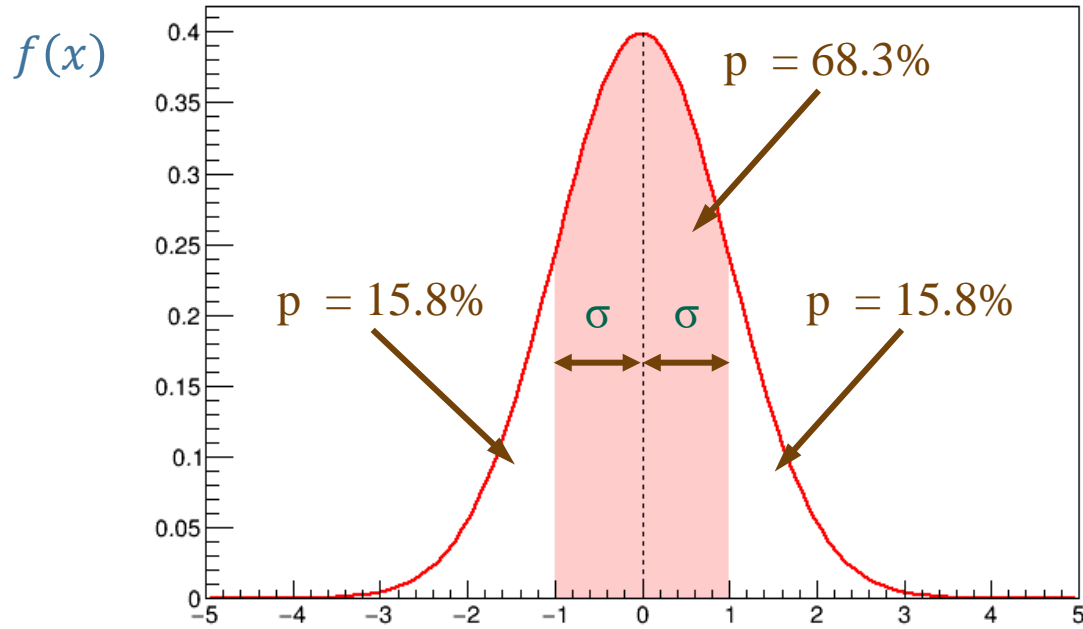
cioè la variabile casuale x ha una probabilità del 68.3% di assumere un valore nell'intervallo $(\mu - \sigma < x < \mu + \sigma)$, ecc.

Nel caso in cui è $\mu_x = 0$ e $\sigma_x^2 = 1$, la funzione di distribuzione di Gauss è detta funzione di distribuzione normale standard:

$$N(0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Tavole della funzione di ripartizione della variabile Normale Standardizzata:

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



z	Seconda cifra decimale di z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997

Funzione di distribuzione di Gauss, distribuzione binomiale ed errori accidentali

Si può dimostrare che, se la variabile casuale k ha una distribuzione binomiale

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

la variabile casuale derivata z definita come:

$$z = \frac{k - E[k]}{\sigma_k} = \frac{k - np}{\sqrt{np(1-p)}}$$

Con p costante tende a seguire una funzione di distribuzione normale standard quando $n \rightarrow \infty$.

Questo processo di limite è alla base della dimostrazione che la distribuzione degli errori accidentali nelle misure ripetute è data dalla funzione di distribuzione di Gauss. Nelle ipotesi che

- tutte le misure siano affette da errori accidentali che consistono in una differenza (casuale) tra valore misurato x e valore vero a (assumendo esista) della grandezza fisica. Se u è il generico errore in una misura, è $x = a + u$;
- gli errori accidentali siano l'effetto di un numero n molto elevato di fenomeni, ciascuno dei quali introduce un errore elementare molto piccolo ε costante;
- questi errori elementari sono tutti uguali in valore assoluto e hanno tutti la stessa probabilità ($p = 0.5$) di essere positivi o negativi;
- gli errori elementari sono tutti indipendenti tra loro,

Funzione di distribuzione di Gauss, distribuzione binomiale ed errori accidentali

Sia ε l'errore elementare che possiamo avere in una misura soggetta ad errori accidentali ed n il numero totale di errori elementari di cui è soggetta una singola misura. La probabilità che il contributo all'errore totale di un singolo ε sia positivo è del 50% (e la stessa probabilità che il contributo sia negativo).

La probabilità che l'errore accidentale u abbia un valore $u_k = \varepsilon k - (n - k)\varepsilon$, cioè sia dovuto al contributo positivo di k errori elementari e a quello negativo dei rimanenti $n - k$, è data da:

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

Essendo $p = 0.5$ e quindi $E[k] = np = 0.5n$, si ottiene $E[u_k] = \varepsilon\{E[k] - (n - E[k])\} = 0$, come intuitivamente ci si aspettava. Al limite per $n \rightarrow \infty$ con p costante, u_k diventa una variabile continua u con funzione di distribuzione gaussiana:

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

Se, per una certa misura σ_u^2 fosse nota, potremmo calcolare la probabilità che l'errore accidentale in una operazione di misura cada in uno specifico intervallo. In generale non è questo il caso, e σ_u^2 deve essere stimata a partire dai risultati di misura.

Dimostrazione euristica

Abbiamo visto che nel caso di k deviazioni elementari positive su n totali il valore dell'errore è:

$$u_k = (2k - n)\varepsilon$$

Con probabilità:

$$P(k) = \frac{n!}{k!(n-k)!} \left(\frac{1}{2}\right)^n$$

Se invece $k+1$ cause introducono deviazioni positive

$$u_{k+1} = (2k + 2 - n)\varepsilon \quad \text{e} \quad P(k+1) = \frac{n!}{(k+1)!(n-k-1)!} \left(\frac{1}{2}\right)^n$$

Se prendiamo l'intervallo tra queste due grandezze, il valore centrale è la semisomma mentre la larghezza dell'intervallo la è differenza:

$$u = \frac{1}{2}(u_{k+1} + u_k) = (2k - n + 1)\varepsilon$$

$$\Delta u = (u_{k+1} - u_k) = 2\varepsilon$$

Dimostrazione euristica

Assumiamo che la probabilità sia data dal valore medio di $P(k)$ e $P(k + 1)$, vero per $\Delta u \rightarrow 0$.

La densità di probabilità $y(u)$ dell'errore per l'intervallo considerato è:

$$y(u) = \frac{P}{\Delta u} \quad \text{con} \quad P = \frac{1}{2} [P(k + 1) + P(k)]$$

Posso quindi calcolare anche:

$$\Delta y = \frac{\Delta P}{\Delta u} \quad \text{con} \quad \Delta P = [P(k + 1) - P(k)]$$

Dividendo la seconda equazione per la prima posso cioè scrivere:

$$\frac{\Delta y}{y} = \frac{\Delta P}{P}$$

Dimostrazione euristica

Vale:

$$\Delta P = \left(\frac{1}{2}\right)^n \left[\frac{n!}{(k+1)!(n-k-1)!} - \frac{n!}{k!(n-k)!} \right] = \left(\frac{1}{2}\right)^n \frac{n!}{k!(n-k-1)!} \left[\frac{1}{k+1} - \frac{1}{n-k} \right] =$$
$$\left(\frac{1}{2}\right)^n \frac{n!}{k!(n-k-1)!} \frac{n-2k-1}{(k+1)(n-k)}$$

$$P = \frac{1}{2} \left(\frac{1}{2}\right)^n \left[\frac{n!}{(k+1)!(n-k-1)!} + \frac{n!}{k!(n-k)!} \right] = \frac{1}{2} \left(\frac{1}{2}\right)^n \frac{n!}{k!(n-k-1)!} \frac{n+1}{(k+1)(n-k)}$$

Da cui:

$$\frac{\Delta y}{y} = \frac{\Delta P}{P} = 2 \frac{n-2k-1}{n+1} = 2 \frac{n-2k-1}{n+1} \frac{\varepsilon^2}{\varepsilon^2}$$

Ricordiamo che:

$$u = (2k - n + 1)\varepsilon$$
$$\Delta u = 2\varepsilon$$

Dimostrazione euristica

Quindi:

$$\frac{\Delta y}{y} = \frac{\Delta P}{P} = 2 \frac{n - 2k - 1}{n + 1} = 2 \frac{n - 2k - 1}{n + 1} \frac{\varepsilon^2}{\varepsilon^2} = - \frac{u \Delta u}{(n + 1) \varepsilon^2} = - \frac{u \Delta u}{\sigma^2}$$

Avendo posto:

$$\sigma^2 = (n + 1) \varepsilon^2$$

E facendo l'ipotesi che σ^2 sia una quantità finita anche al limite $n \rightarrow \infty$, ovvero che ε diventi infinitesimo a questo limite. Al limite $n \rightarrow \infty$, ovvero abbiamo quindi:

$$\frac{dy}{y} = - \frac{u du}{\sigma^2}$$

La soluzione dell'equazione differenziale è:

$$\ln y = - \frac{u^2}{2\sigma^2} + C \quad \text{ovvero} \quad y(u) = A e^{-\frac{u^2}{2\sigma^2}} \quad \text{con} \quad \int_{-\infty}^{+\infty} y(u) du = 1$$

La condizione di normalizzazione permette di definire il valore di A e scrivere:

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

Dimostrazione formale

La dimostrazione che la **distribuzione di Bernoulli** tende alla funzione di Gauss quando il numero di tentativi $n \rightarrow \infty$ (noto come il **Central Limite theorem di de Moivre e Laplace**) è piuttosto lunga e laboriosa.

- Fu inizialmente dimostrato per il caso $p = 1/2$ da Abraham de Moivre in *The Doctrine of Chances* nel 1738
- Pierre Simon Laplace ha riscoperto la dimostrazione e l'ha estesa al caso $p \neq 1/2$ in *Théorie Analytique des Probabilités* nel 1812. Anche Gauss contribuì alla dimostrazione più generale.
- Laplace ha anche scoperto **una generalizzazione** che ha chiamato *Teorema del limite centrale*, molto più generale rispetto al caso della distribuzione binomiale, ma la sua dimostrazione non era rigorosa.
- Nel 1901-1902 il matematico russo **Aleksandr Liapunof** ha dato la prima prova rigorosa del teorema del limite centrale.
- Nel 1922 **George Polya** ha per primo usato il termine 'limite centrale'

Teorema di de Moivre-La Place

Siano $a, b \in \mathbb{R}$ con $a < b$, e consideriamo una variabile aleatoria binomiale S_n , che possa assumere tutti i valori tra 0 e n con $n \in \mathbb{Z}$. La probabilità che S_n assuma il valore k è quindi

$$P(S_n = k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

Con p costante e $q = 1 - p$. Al limite di $n \rightarrow \infty$, la probabilità che la variabile aleatoria

$$\frac{S_n - np}{\sqrt{np(1-p)}}$$

cada nell'intervallo $(a, b]$ è data da:

$$\lim_{n \rightarrow \infty} P_n \left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

O, che è lo stesso, introdotta la variabile $y \in \mathbb{R}$, e definito $k(y) = np + y\sqrt{np(1-p)}$

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{k(y)} \binom{n}{j} p^j (1-p)^{k(y)-j} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{np + y\sqrt{np(1-p)}} e^{-\frac{(x-np)^2}{2np(1-p)}} dx$$

Cenni sulla dimostrazione

Per dimostrare il teorema c'è la necessità di usare diversi lemmi:

- i. Normalizzazione della distribuzione normale $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1$
- ii. $\frac{1}{\sqrt{2\pi}(1+y)} (e^{-y^2/2} - e^{-(y+1)^2/2}) \leq \frac{1}{\sqrt{2\pi}} \int_y^{+\infty} e^{-\frac{x^2}{2}} dx \leq \frac{1}{\sqrt{2\pi}y} e^{-\frac{y^2}{2}}$
- iii. $\lim_{y \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_y^{+\infty} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}y} e^{-\frac{y^2}{2}}$
- iv. La convergenza del teorema di Moivre-Laplace è uniforme sia in a che in b
- v. Per ogni $n > 0$ si può scrivere $n! = \sqrt{2\pi n} n^{n+1/2} e^{-n} (1 + \varepsilon_n)$, esiste sempre un valore reale A tale che $|\varepsilon_n| < \frac{A}{n}$ (formula di Sterling)

Cenni sulla dimostrazione

vi. de Moivre-Laplace Binomial Point Mass Limit:

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(k-np)^2}{2np(1-p)}} (1 + \delta_n(k))$$

dove per $a > 0$ vale $\lim_{n \rightarrow \infty} \max_{|k-np| < a\sqrt{n}} |\delta_n(k)| = 0$

vii. Se $[a, b]$ è un intervallo di \mathbb{R} e sia f una funzione definita su \mathbb{R} che sia 0 al di fuori di $[a, b]$ e continuo in $[a, b]$, allora per ogni t :

$$\lim_{h \rightarrow 0, h \geq 0} h \sum_{k=-\infty}^{\infty} f(t + kh) = \int_a^b f(x) dx$$

Gauss come limite della binomiale

Sia definito un intervallo $K_n = [a\sqrt{np(1-p)}, b\sqrt{np(1-p)}]$, prendiamo una variabile aleatoria S_n con distribuzione binomiale allora la probabilità che $S_n - np \in K_n$ è data da:

$$\begin{aligned} P(S_n - np \in K_n) &= \sum_{k=0}^n 1_{K_n}(k - np) \binom{n}{k} p^k q^{n-k} \\ &= \frac{1}{\sqrt{2\pi np(1-p)}} \sum_{k=0}^n 1_{K_n}(k - np) e^{-\frac{(k-np)^2}{2np(1-p)}} (1 + \delta_n(k)) \end{aligned}$$

All'aumentare di n la sommatoria può essere sostituita con le somme su tutto \mathbb{Z} ed otteniamo

$$\frac{1}{\sqrt{2\pi}} \frac{1 + \delta_n}{\sqrt{np(1-p)}} \sum_{k \in \mathbb{Z}} 1_{[a,b]} \left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{\frac{np}{1-p}} \right) e^{-\frac{1}{2} \left(\frac{k}{\sqrt{np(1-p)}} - \sqrt{np-p} \right)^2}$$

Dei 2 termini $1 + \delta_n$, mentre il termine con 1 converge ad un valore finito, il termine con δ_n converge a 0 e può essere trascurato.

Con $h = 1/\sqrt{np(1-p)}$ e $f(x) = (1/\sqrt{2\pi})e^{-x^2/2}$, per il lemma vii dimostro il teorema per un intervallo finito $[a, b]$

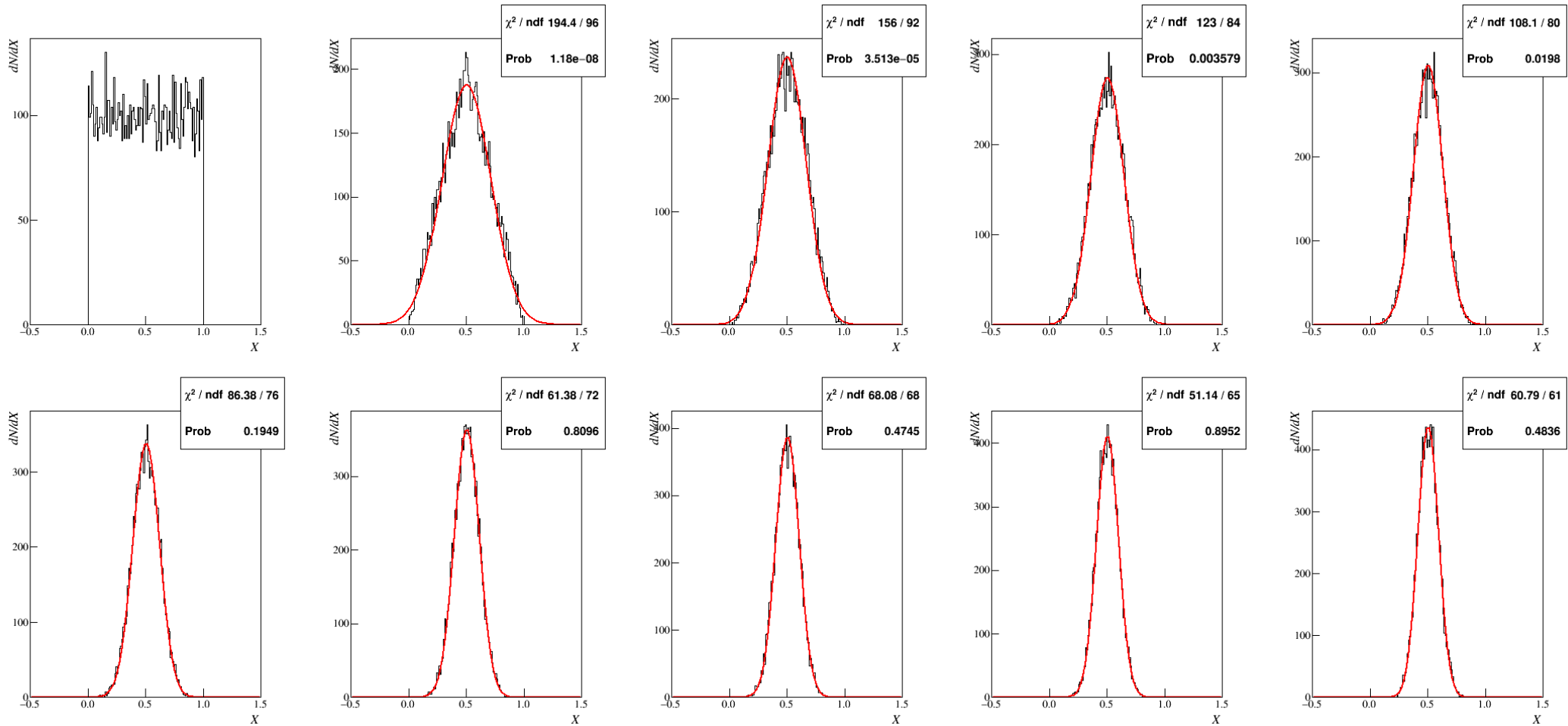
Gauss come limite della binomiale

- Il prossimo passo consiste nel dimostrare che il limite vale anche quando $a \rightarrow -\infty$ e b è finito e reale
- Si dimostra che il teorema vale anche per a finito e reale e $b \rightarrow +\infty$
- Che dimostra la validità del teorema su $[-\infty, +\infty]$

- Dimostrazione completa disponibile in rete (per i *mathematically oriented*), ad esempio, sul sito:

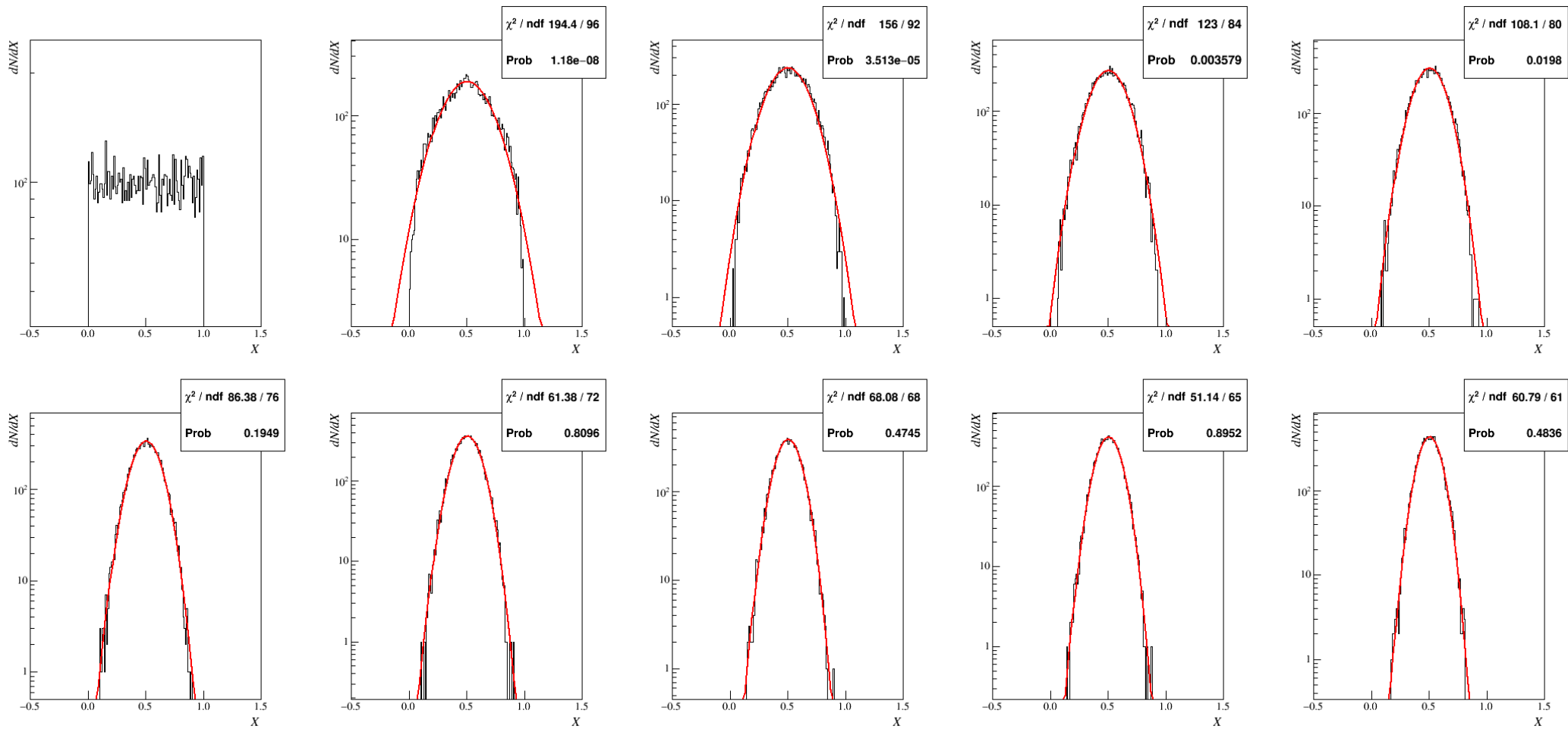
<http://www.math.unl.edu/~sdunbar1/ProbabilityTheory/Lessons/BernoulliTrials/DeMoivreLaplaceCLT/demoivrelaplaceclt.pdf>

Limite centrale



La media di N variabili casuali R_n converge ad una distribuzione Gaussiana, indipendentemente dalle distribuzioni di origine

Limite centrale



Funzioni di distribuzione per più variabili casuali

Generalizzando quanto visto per una variabile casuale, nel caso in cui il fenomeno statistico sia caratterizzato da n variabili casuali, si introduce la “densità di probabilità composta” o “congiunta” $f(x_1, x_2, \dots, x_n)$ definita in modo che la probabilità che le variabili casuali assumano valori in determinati intervalli sia data da:

$$P(x_1 < X_1 < x_1 + dx_1, \dots, x_n < X_n < x_n + dx_n) = f(x_1, x_2, \dots, x_n) dx_1 \cdots dx_n$$

Come nel caso di una variabile casuale, la funzione $f(x_1, x_2, \dots, x_n)$ deve essere limitata e definita positiva in tutto il campo di definizione delle variabili casuali $\Omega_{\vec{x}} = \{(a_1, b_1), \dots, (a_n, b_n)\}$ e deve essere normalizzata, ciò deve essere

$$\int_{\Omega_{\vec{x}}} f(x_1, x_2, \dots, x_n) dx_1 \cdots dx_n = 1$$

La funzione di distribuzione per la generica variabile casuale x_i , che da la densità di probabilità per x_i indipendentemente dal valore (o per ogni valore) delle altre variabili casuali è la funzione di distribuzione “marginale”

$$f_{X_i}(x_i) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, x_2, \dots, x_n) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n$$

dove l'integrazione è su tutte le variabili $x_j \neq x_i$

Funzioni di distribuzione per più variabili casuali

Generalizzando le espressioni introdotte per il caso di una variabile casuale, valori di aspettazione e varianze sono dati da:

$$E[x_i] = \mu_i = \int_{\Omega_{\vec{x}}} x_i f(x_1, x_2, \dots, x_n) dx_1 \cdots dx_n$$

$$E[(x_i - \mu_i)^2] = \sigma_i^2 = \int_{\Omega_{\vec{x}}} (x_i - \mu_i)^2 f(x_1, x_2, \dots, x_n) dx_1 \cdots dx_n$$

Nel caso di più variabili casuali si deve tener conto anche della correlazione tra le variabili. Ciò viene fatto introducendo la “covarianza” di due variabili casuali:

$$E[(x_i - \mu_i)(x_j - \mu_j)] = \text{cov}(x_i, x_j) = \int_{\Omega_{\vec{x}}} (x_i - \mu_i)(x_j - \mu_j) f(x_1, x_2, \dots, x_n) dx_1 \cdots dx_n$$

Nel caso di n variabili abbiamo n valori di aspettazione e n varianze, e, vista la simmetria $\text{cov}(x_j, x_i) = \text{cov}(x_i, x_j)$, abbiamo $(n^2 - n)/2$ o $n(n - 1)/2$ covarianze

Funzioni di distribuzione per più variabili casuali

La covarianza può anche essere scritta come

$$\text{cov}(x_i, x_j) = E[x_i x_j] - \mu_i \mu_j$$

(basta sviluppare il doppio prodotto nell'integrale).

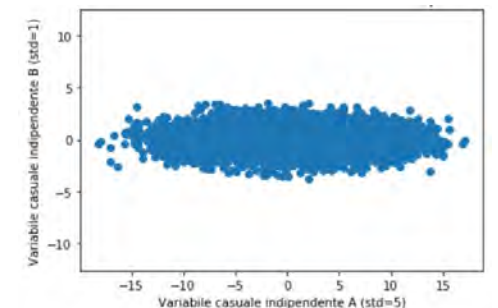
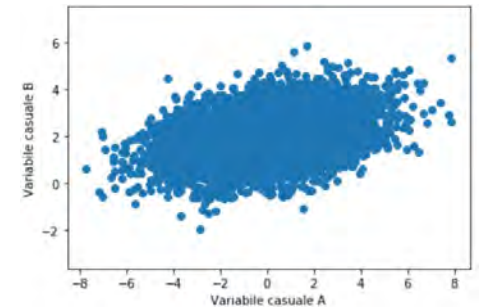
In genere, per esprimere la correlazione statistica tra due variabili viene utilizzato il “**coefficiente di correlazione**”:

$$\rho(x_i, x_j) = \frac{\text{cov}(x_i, x_j)}{\sigma_i \sigma_j}$$

rapporto tra la covarianza e il prodotto delle deviazioni standard. Deve essere:

$$-1 \leq \rho(x_i, x_j) \leq 1$$

Nel caso $\rho(x_i, x_j) = 1(-1)$ si ha correlazione (anti correlazione) completa tra le due variabili, mentre **per variabili indipendenti vale $\rho_{ij} = 0$** .



Matrice delle covarianze

Possiamo scrivere in forma matriciale le covarianze associate al nostro set di variabili indicando con

$$V_{ij} = \text{cov}(x_i, x_j)$$

gli elementi di tale matrice.

Per come è costruita la matrice è simmetrica, infatti $\text{cov}(x_i, x_j) = \text{cov}(x_j, x_i)$, ed ha come elementi diagonali le varianze delle diverse variabili $\text{cov}(x_i, x_i) = \sigma_i^2$:

$$V = \begin{pmatrix} \sigma_1^2 & \cdots & \text{cov}(x_1, x_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(x_n, x_1) & \cdots & \sigma_n^2 \end{pmatrix}$$

Funzioni di distribuzione per più variabili casuali

Casi limite:

- Correlazione +1. La correlazione di una variabile aleatoria X con se stessa:

$$\rho(x_i, x_i) = \frac{\text{cov}(x_i, x_i)}{\sigma_i^2} = \frac{E[x_i x_i] - \mu_i \mu_i}{\sigma_i^2} = \frac{\sigma_i^2}{\sigma_i^2} = 1$$

- Correlazione -1. La correlazione di una variabile aleatoria X con una variabile aleatoria $Y = -X$

$$\rho(x_i, -x_i) = \frac{\text{cov}(x_i, -x_i)}{\sigma_i^2} = \frac{-E[x_i x_i] + \mu_i \mu_i}{\sigma_i^2} = \frac{-\sigma_i^2}{\sigma_i^2} = -1$$

- Correlazione 0. La correlazione di due variabili aleatorie indipendenti è 0:

$$\rho(x_i, x_j) = \frac{\text{cov}(x_i, x_j)}{\sigma_i \sigma_j} = \frac{E[x_i x_j] - \mu_i \mu_j}{\sigma_i \sigma_j} = \frac{E[x_i]E[x_j] - \mu_i \mu_j}{\sigma_i \sigma_j} = \frac{\mu_i \mu_j - \mu_i \mu_j}{\sigma_i \sigma_j} = 0$$

Funzioni di distribuzione per più variabili casuali

Una **semplificazione notevole** nella trattazione di fenomeni statistici che coinvolgono più variabili si ha nel caso in cui le **variabili casuali siano statisticamente indipendenti**. Infatti, ricordando che ($P(A \cap B) = P(A) \cdot P(B)$ se A e B sono eventi indipendenti) è:

$$P(a_1 < x_1 < a_1 + dx_1, \dots, a_n < x_n < a_n + dx_n) \\ = P(a_1 < x_1 < a_1 + dx_1) \cdot P(a_2 < x_2 < a_2 + dx_2) \cdots P(a_n < x_n < a_n + dx_n) = f_1(a_1)dx_1 \cdots f_n(a_n)dx_n$$

dove $f_i(x_i)$ è la funzione di distribuzione della variabile x_i , normalizzata. La funzione di distribuzione delle n variabili casuali è quindi fattorizzata:

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2) \cdots f_n(x_n)$$

Ovvero, valori di aspettazione e varianze sono date dalle espressioni introdotte nel caso di una variabile casuale:

$$E[x_i] = \mu_i = \int_{a_i}^{b_i} x_i f_i(x_i) dx_i \quad \text{e} \quad E[(x_i - \mu_i)^2] = \sigma_i^2 = \int_{a_i}^{b_i} (x_i - \mu_i)^2 f_i(x_i) dx_i$$

Inoltre

$$E[(x_i - \mu_i)(x_j - \mu_j)] = \text{cov}(x_i, x_j) = \int_{a_i}^{b_i} x_i f_i(x_i) dx_i \int_{a_j}^{b_j} x_j f_j(x_j) dx_j - \mu_i \mu_j = 0$$

Variabili indipendenti con funzione di distribuzione di Gauss

Particolarmente interessante per le applicazioni nel corso di Laboratorio 1.

Supponiamo che le variabili x_i siano indipendenti e abbiano tutte funzione di distribuzione Gaussiana:

$$f_i(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}}$$

In questo caso la funzione di distribuzione congiunta è:

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_i(x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}} = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} e^{-\sum_{i=1}^n \frac{(x_i-\mu_i)^2}{2\sigma_i^2}}$$

Se poi le varianze e i valori medi sono rispettivamente uguali a σ e μ per tutte le variabili, si ottiene:

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}}$$

Funzioni di variabili casuali

Se $y = y(x)$ è una funzione della variabile casuale x , anche y deve essere una variabile casuale e, come tale, avrà una sua funzione di distribuzione. Indicando con $f(x)$ e $g(y)$ le funzioni di distribuzione di x e y , nel caso di corrispondenza biunivoca deve essere

$$f(x_a)dx = P(x_a < x < x_a + dx) = P(y_a < y < y_a + dy) = g(y_a)dy$$

dove $y_a = f(x_a)$. La determinazione di $g(y)$ si riduce quindi ad una trasformazione di variabile. Se esiste la funzione inversa $x = x(y)$, allora

$$g(y)dy = f(x(y))dx(y) = f(x(y)) \left| \frac{dx(y)}{dy} \right| dy$$

ovvero

$$g(y) = f(x(y)) \left| \frac{dx(y)}{dy} \right|$$

Il modulo della derivata fa sì che la funzione densità di probabilità $g(y)$ sia definita positiva.

Funzioni di variabili casuali

Nel caso in cui la funzione $y = y(x)$ non sia **biunivoca**, cioè se esistono m valori x_j tali che $y = y(x_j)$, la funzione di distribuzione di y è data da

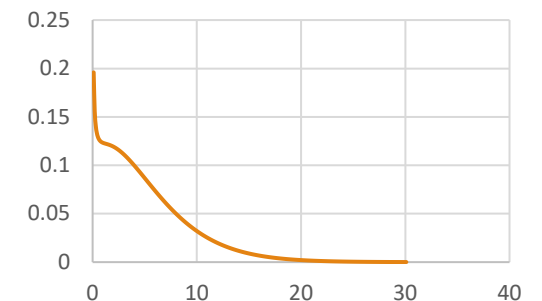
$$g(y) = \sum_{j=1}^m f(x_j(y)) \left| \frac{dx_j(y)}{dy} \right|$$

Ad esempio, nel caso $y = ax + b$, e x è distribuito secondo $f(x)$, sarà $x(y) = \frac{(y-b)}{a}$; quindi

$$g(y) = f(x(y)) \left| \frac{dx(y)}{dy} \right| = \frac{f((y-b)/a)}{|a|}$$

Mentre, nel caso $y = ax^2$ (con $a > 0$), ho due valori possibili per x ; $x = +\sqrt{y/a}$ e $x = -\sqrt{y/a}$, per cui

$$g(y) = f\left(+\sqrt{\frac{y}{a}}\right) \frac{1}{2\sqrt{ya}} + f\left(-\sqrt{\frac{y}{a}}\right) \frac{1}{2\sqrt{ya}}$$



Funzioni di variabili casuali

Come ulteriore esempio si può considerare il caso di una **variabile casuale x con distribuzione $N(\mu, \sigma)$** . La variabile

$$y = \frac{(x - \mu)}{\sigma}$$

ha distribuzione normale standard $N(0,1)$.

Infatti è

$$x = \sigma y + \mu$$

e

$$\frac{dx}{dy} = \sigma$$

e quindi

$$g(y) = f(x(y)) \left| \frac{dx(y)}{dy} \right| = \sigma \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\sigma y + \mu - \mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

Funzioni di variabili casuali

Non sempre la funzione di distribuzione $g(y)$ può essere calcolata facilmente e/o ha un'espressione semplice, che permetta di calcolare con un'integrazione non numerica il valore di aspettazione e la varianza di y .

Il valore d'aspettazione di y è dato da:

$$\mu_y = E[y] = \int_{-\infty}^{+\infty} yg(y)dy$$

Ma se sono note la $f(x)$ e la relazione $y(x)$ posso scrivere (teorema della media)

$$\mu_y = E[y] = \int_{-\infty}^{+\infty} yg(y)dy = E[y(x)] = \int_{-\infty}^{+\infty} y(x)f(x)dx$$

$$\sigma_y^2 = E \left[(y(x) - \mu_y)^2 \right] = \int_{-\infty}^{+\infty} (y(x) - \mu_y)^2 f(x)dx$$

Funzioni di variabili casuali

Il calcolo dell'integrale non è pratico in molti casi. Inoltre, alle volte **non conosco la funzione di distribuzione x ma conosco (calcolo) media e varianza μ_x e σ_x^2** ; posso calcolare media μ_y e varianza σ_y^2 di $y = y(x)$?

Se $y(x)$ è lineare, $y(x) = ax + b$ allora:

$$\mu_y = E[y] = E[ax + b] = aE[x] + b = a\mu_x + b$$

$$\sigma_y^2 = E[(y - \mu_y)^2] = E[y^2] - \mu_y^2 = E[(ax + b)^2] - (a\mu_x + b)^2 = a^2E[x^2] + 2abE[x] + b^2 - (a\mu_x + b)^2$$

Da cui:

$$\sigma_y^2 = a^2E[x^2] + 2ab\mu_x + b^2 - a^2\mu_x^2 - 2ab\mu_x - b^2 = a^2(E[x^2] - \mu_x^2) = a^2\sigma_x^2$$

Viceversa: se $y(x)$ **NON** è lineare posso solo stimare in modo approssimativo la media μ_y e la varianza σ_y^2 linearizzando la $y(x)$ in un intorno di μ_x :

$$y(x) = y(\mu_x) + \left(\frac{dy}{dx}\right)_{x=\mu_x} (x - \mu_x) + \frac{1}{2} \left(\frac{d^2y}{dx^2}\right)_{x=\mu_x} (x - \mu_x)^2 + \dots$$

Funzioni di variabili casuali

$$y(x) = y(x = \mu_x) + \left(\frac{dy}{dx}\right)_{x=\mu_x} (x - \mu_x) + \frac{1}{2} \left(\frac{d^2y}{dx^2}\right)_{x=\mu_x} (x - \mu_x)^2 + \dots$$

Troncando al secondo ordine ottengo per il valore di aspettazione μ_y :

$$\begin{aligned} \mu_y = E[y] &= E \left[y(\mu_x) + \left(\frac{dy}{dx}\right)_{x=\mu_x} (x - \mu_x) + \frac{1}{2} \left(\frac{d^2y}{dx^2}\right)_{x=\mu_x} (x - \mu_x)^2 + \dots \right] = \\ &= E[y(\mu_x)] + \left(\frac{dy}{dx}\right)_{x=\mu_x} E[(x - \mu_x)] + \frac{1}{2} \left(\frac{d^2y}{dx^2}\right)_{x=\mu_x} E[(x - \mu_x)^2] + \dots \\ &\simeq y(\mu_x) + \frac{1}{2} \left(\frac{d^2y}{dx^2}\right)_{x=\mu_x} \sigma_x^2 \end{aligned}$$

Funzioni di variabili casuali

Mentre per la varianza ottengo:

$$\begin{aligned}\sigma_y^2 &= E[y^2] - \mu_y^2 = E \left[\left(y(\mu_x) + \left(\frac{dy}{dx} \right)_{x=\mu_x} (x - \mu_x) + \frac{1}{2} \left(\frac{d^2y}{dx^2} \right)_{x=\mu_x} (x - \mu_x)^2 + \dots \right)^2 \right] - \mu_y^2 = \\ &E \left[(y(\mu_x))^2 \right] + 2y(\mu_x) \left(\frac{dy}{dx} \right)_{x=\mu_x} E[(x - \mu_x)] + \left(\left[\frac{dy}{dx} \right]_{x=\mu_x} \right)^2 E[(x - \mu_x)^2] + \dots - \mu_y^2 = \\ &\approx \left(\left[\frac{dy}{dx} \right]_{x=\mu_x} \right)^2 \sigma_x^2\end{aligned}$$

Ovvero la nota “legge di propagazione della varianza”

Funzioni di variabili casuali

In genere quindi, la conoscenza di valore di aspettazione e varianza di x e μ_x sono sufficienti. Sviluppando in serie y in un intorno di $\mu_x = E[x]$, si ottengono le espressioni (approssimate):

$$\mu_y = E[y] \cong y(x = \mu_x) \cdots \left(+ \frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 \right)$$

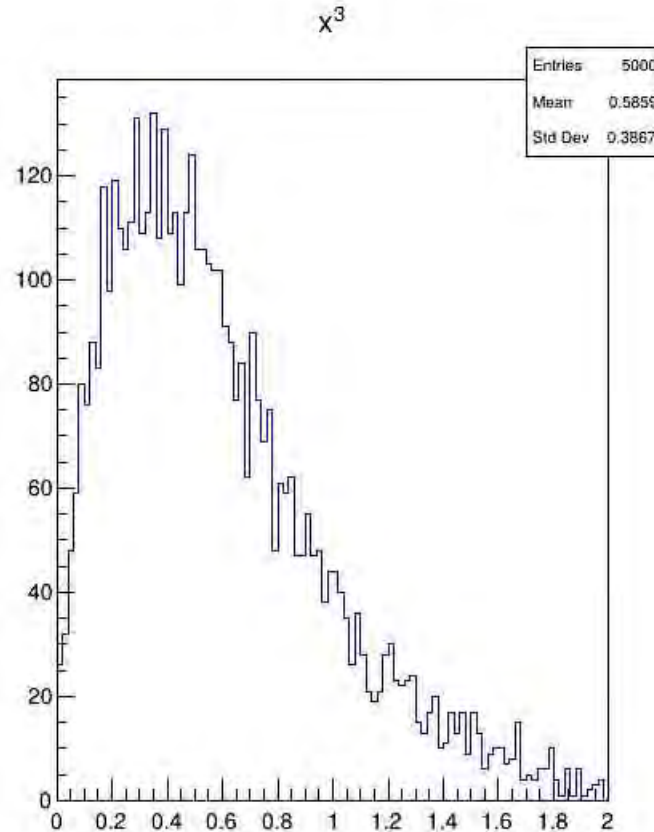
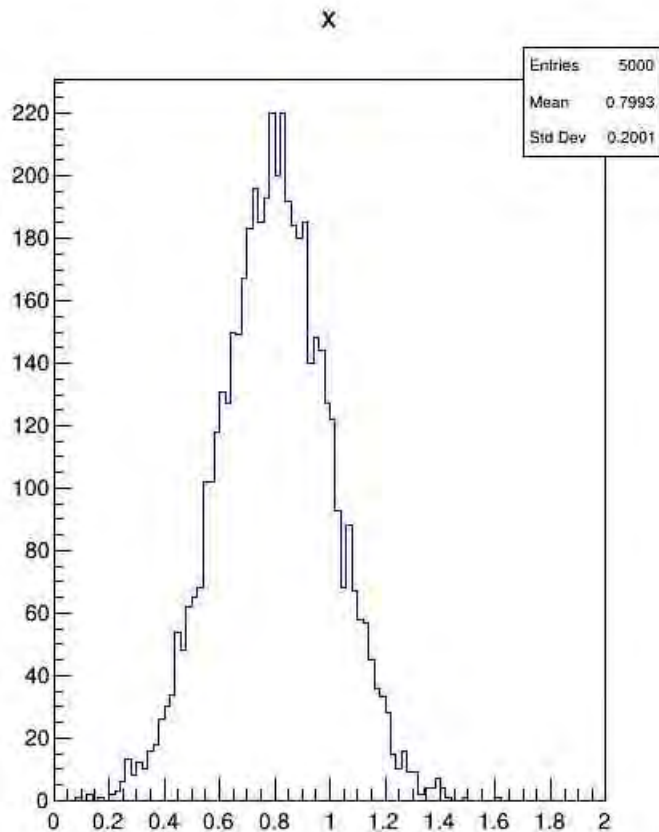
$$\sigma_y^2 = var(y) = \left(\left[\frac{dy}{dx} \right]_{x=\mu_x} \right)^2 \sigma_x^2$$

qualunque sia la funzione di distribuzione di x .

Esempio $Y = X^3$

Misure di x con distribuzione Gaussiana e $\mu_x = 0.8$ e $\sigma_x = 0.2$

$$\mu_y = E[y] \cong y(x = \mu_x) + \frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = \mu_x^3 + 3\mu_x \sigma_x^2 \quad \sigma_y^2 = \text{var}(y) = \left(\left[\frac{dy}{dx} \right]_{x=\mu_x} \right)^2 \sigma_x^2 = 9\mu_x^4 \sigma_x^2 = 0.157$$



$$y(x = \mu_x) = 0.512$$

$$\sigma_y = 0.38$$

$$\frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = 0.096$$

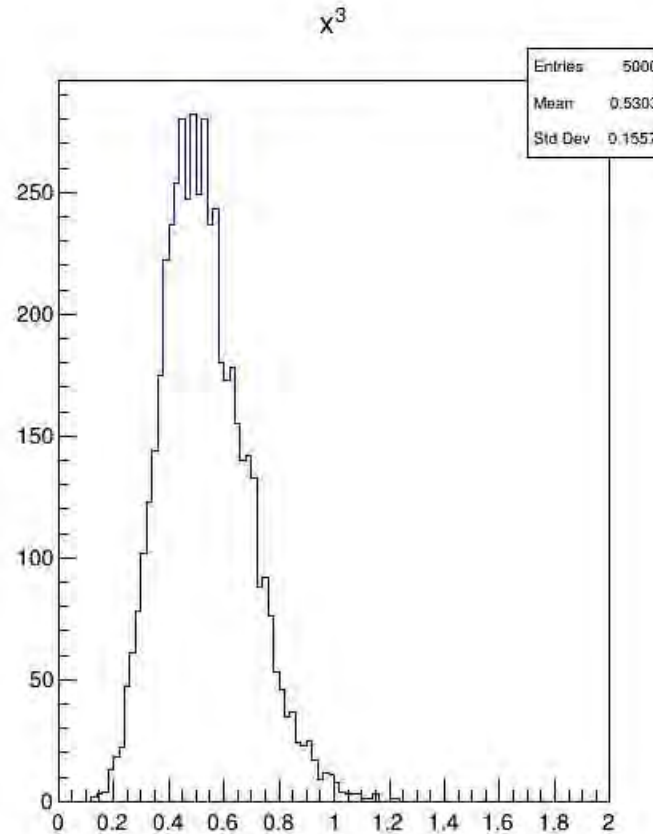
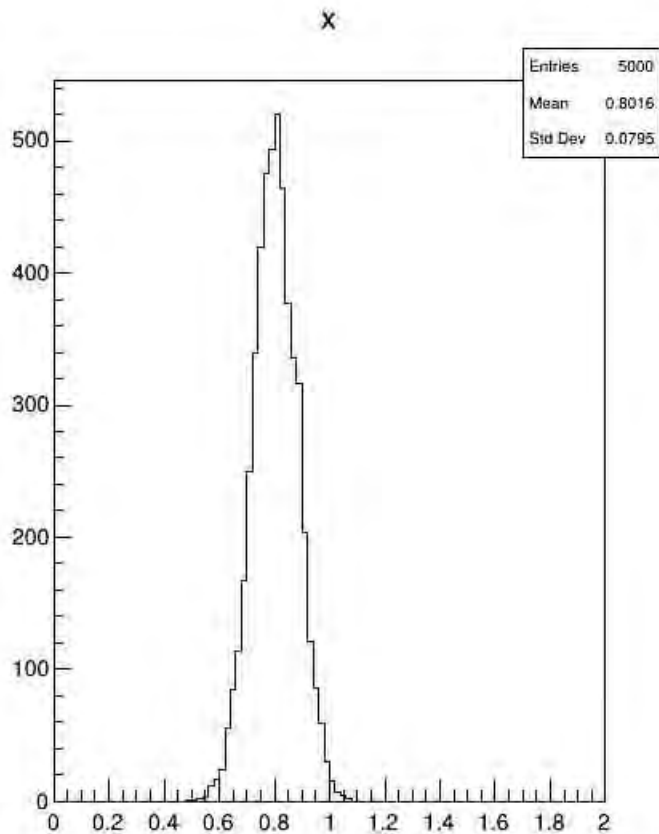
$$y(x = \mu_x) + \frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = 0.608$$

$$\mu_y = 0.608$$

Esempio $Y = X^3$

Misure di x con distribuzione Gaussiana e $\mu_x = 0.8$ e $\sigma_x = 0.08$

$$\mu_y = E[y] \cong y(x = \mu_x) + \frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = \mu_x^3 + 3\mu_x \sigma_x^2 \quad \sigma_y^2 = \text{var}(y) = \left(\left[\frac{dy}{dx} \right]_{x=\mu_x} \right)^2 \sigma_x^2 = 9\mu_x^4 \sigma_x^2 = 0.024$$



$$y(x = \mu_x) = 0.512$$

$$\sigma_y = 0.15$$

$$\frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = 0.015$$

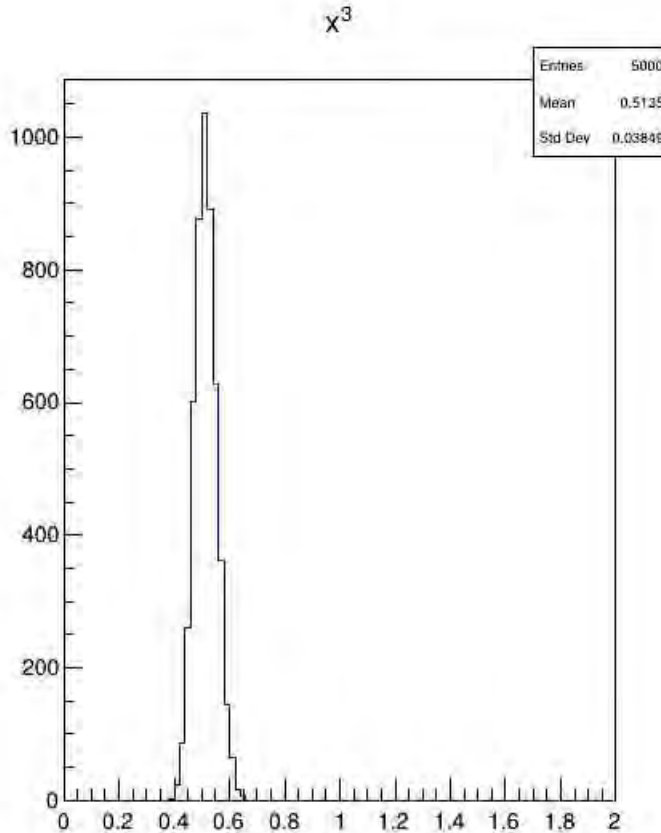
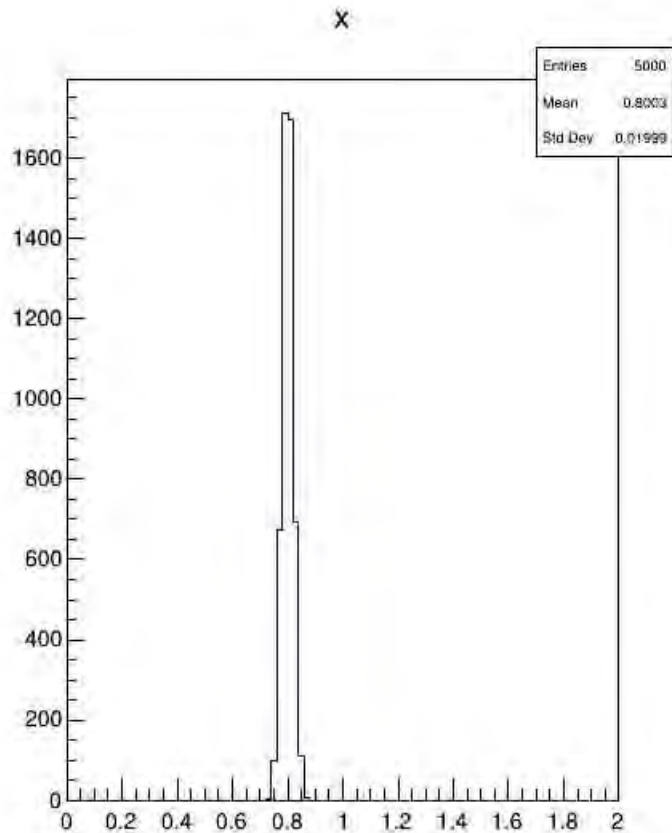
$$y(x = \mu_x) + \frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = 0.530$$

$$\mu_y = 0.530$$

Esempio $Y = X^3$

Misure di x con distribuzione Gaussiana e $\mu_x = 0.80$ e $\sigma_x = 0.02$

$$\mu_y = E[y] \cong y(x = \mu_x) + \frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = \mu_x^3 + 3\mu_x \sigma_x^2 \quad \sigma_y^2 = \text{var}(y) = \left(\left[\frac{dy}{dx} \right]_{x=\mu_x} \right)^2 \sigma_x^2 = 9\mu_x^4 \sigma_x^2 = 0.0015$$



$$y(x = \mu_x) = 0.512$$

$$\sigma_y = 0.04$$

$$\frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = 0.001$$

$$y(x = \mu_x) + \frac{1}{2} \left[\frac{d^2 y}{dx^2} \right]_{x=\mu_x} \sigma_x^2 = 0.513$$

$$\mu_y = 0.513$$

Errori accidentali e misure ripetute

Abbiamo visto che la misura di una grandezza fisica a è affetta da un errore u che è una variabile casuale con funzione di distribuzione gaussiana $N(0, \sigma_u^2)$:

$$h(u) = N(0, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}}$$

Il risultato dell'operazione di misura $x = a + u$ è una funzione di u ed è quindi una variabile casuale con funzione di distribuzione che possiamo calcolare. Poiché $u = x - \mu$, e $du/dx = 1$, la distribuzione di x è

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

Quindi i risultati ottenuti nelle singole operazioni di misura x hanno una funzione di distribuzione di Gauss con valore di aspettazione $\mu_x = \mu$, cioè centrata sul valore vero della grandezza fisica. La varianza è $\sigma_x^2 = \sigma_u^2$. Utilizzando le proprietà della distribuzione di Gauss, se μ e σ_x sono noti, si può calcolare la probabilità che, eseguendo una misura, il risultato x_1 cada in un certo intervallo. In particolare sarà

$$P(\mu - \sigma_x \leq x_1 \leq \mu + \sigma_x) = 0.683$$

Pro memoria

Come già detto,

- se la risoluzione di lettura è bassa, più grande delle fluttuazioni statistiche, le fluttuazioni del valore misurato non sono apprezzabili, **domina l'incertezza "massima" dovuta alla risoluzione di lettura**. La trattazione statistica può comunque effettuata essere usando la distribuzione uniforme, che fornisce una deviazione standard

$$\sigma_x = \Delta x / \sqrt{3}$$

- Viceversa, si può passare dall'incertezza statistica dovuta a errori accidentali all'incertezza massima ponendo $\Delta x = 3\sigma_x$. Infatti, per le proprietà della funzione di Gauss, all'affermazione $x_1 - 3\sigma_x \leq a \leq x_1 + 3\sigma_x$ corrisponde una probabilità di 0.997, molto vicina a 1.

$$\Delta x = 3\sigma_x$$

Funzioni di più variabili casuali

Quando abbiamo a che fare con una funzione composta, $y(x_1, x_2, \dots, x_n)$ non lineare, possiamo seguire lo stesso procedimento di sviluppo in serie usato nel caso della dipendenza da una variabile $y = y(x)$. Sviluppiamo la funzione in un intorno di $(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n})$ fermandoci al secondo ordine :

$$\begin{aligned}
 y(x_1, x_2, \dots, x_n) &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \dots \sum_{k_n=0}^{\infty} \frac{(x_1 - \mu_{x_1})^{k_1} (x_2 - \mu_{x_2})^{k_2} \dots (x_n - \mu_{x_n})^{k_n}}{k_1! k_2! \dots k_n!} \left[\frac{\partial^{k_1 k_2 \dots k_n} y}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \right]_{\vec{x}=\vec{\mu}_x} \\
 &\cong y(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) + (x_1 - \mu_{x_1}) \left[\frac{\partial y}{\partial x_1} \right]_{\vec{x}=\vec{\mu}_x} + (x_2 - \mu_{x_2}) \left[\frac{\partial y}{\partial x_2} \right]_{\vec{x}=\vec{\mu}_x} + \dots + (x_n - \mu_{x_n}) \left[\frac{\partial y}{\partial x_n} \right]_{\vec{x}=\vec{\mu}_x} \\
 &+ \frac{1}{2} (x_1 - \mu_{x_1})^2 \left[\frac{\partial^2 y}{\partial^2 x_1} \right]_{\vec{x}=\vec{\mu}_x} + \frac{1}{2} (x_2 - \mu_{x_2})^2 \left[\frac{\partial^2 y}{\partial^2 x_2} \right]_{\vec{x}=\vec{\mu}_x} + \dots + \frac{1}{2} (x_n - \mu_{x_n})^2 \left[\frac{\partial^2 y}{\partial^2 x_n} \right]_{\vec{x}=\vec{\mu}_x} + \dots \\
 &+ (x_1 - \mu_{x_1})(x_2 - \mu_{x_2}) \left[\frac{\partial^2 y}{\partial x_1 \partial x_2} \right]_{\vec{x}=\vec{\mu}_x} + \dots + (x_i - \mu_{x_i})(x_j - \mu_{x_j}) \left[\frac{\partial^2 y}{\partial x_i \partial x_j} \right]_{\vec{x}=\vec{\mu}_x} + \dots \\
 &= y(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) + \sum_{i=1}^n (x_i - \mu_{x_i}) \left[\frac{\partial y}{\partial x_i} \right]_{\vec{x}=\vec{\mu}_x} + \frac{1}{2} \sum_{i=1}^n (x_i - \mu_{x_i})^2 \left[\frac{\partial^2 y}{\partial^2 x_i} \right]_{\vec{x}=\vec{\mu}_x} + \sum_{i=1}^n \sum_{j=i+1}^n (x_i - \mu_{x_i})(x_j - \mu_{x_j}) \left[\frac{\partial^2 y}{\partial x_i \partial x_j} \right]_{\vec{x}=\vec{\mu}_x}
 \end{aligned}$$

Funzioni di più variabili casuali

Se ora prendo il valore di aspettazione:

$$\begin{aligned} & E[y(x_1, x_2, \dots, x_n)] \\ &= E \left[y(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) + \sum_{i=1}^n (x_i - \mu_{x_i}) \left[\frac{\partial y}{\partial x_i} \right]_{\vec{x}=\vec{\mu}_x} + \frac{1}{2} \sum_{i=1}^n (x_i - \mu_{x_i})^2 \left[\frac{\partial^2 y}{\partial^2 x_i} \right]_{\vec{x}=\vec{\mu}_x} \right. \\ & \quad \left. + \sum_{i=1}^n \sum_{j=i+1}^n (x_i - \mu_{x_i}) (x_j - \mu_{x_j}) \left[\frac{\partial^2 y}{\partial x_i \partial x_j} \right]_{\vec{x}=\vec{\mu}_x} \right] \end{aligned}$$

Ricordando le proprietà del valore di aspettazione:

$$\begin{aligned} & E[y] \\ &= E[y(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n})] + \sum_{i=1}^n E \left[(x_i - \mu_{x_i}) \left[\frac{\partial y}{\partial x_i} \right]_{\vec{x}=\vec{\mu}_x} \right] + \frac{1}{2} \sum_{i=1}^n E \left[(x_i - \mu_{x_i})^2 \left[\frac{\partial^2 y}{\partial^2 x_i} \right]_{\vec{x}=\vec{\mu}_x} \right] \\ & \quad + \sum_{i=1}^n \sum_{j=i+1}^n E \left[(x_i - \mu_{x_i}) (x_j - \mu_{x_j}) \left[\frac{\partial^2 y}{\partial x_i \partial x_j} \right]_{\vec{x}=\vec{\mu}_x} \right] \end{aligned}$$

Funzioni di più variabili casuali

da cui:

$$\mu_y \cong y(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) + \frac{1}{2} \sum_{k=1}^n \left[\frac{\partial^2 y}{\partial^2 x_k} \right]_{\vec{x}=\vec{\mu}_x} \sigma_{x_k}^2 + \sum_{i=1}^n \sum_{j=1, j \neq i}^n \left[\frac{\partial^2 y}{\partial x_i \partial x_j} \right]_{\vec{x}=\vec{\mu}_x} \text{cov}(x_j, x_i)$$

Inoltre si può dimostrare che:

$$E[y^2] - \mu_y^2 \cong \sum_{k=1}^n \left(\left[\frac{\partial y}{\partial x_k} \right]_{\vec{x}=\vec{\mu}_x} \right)^2 \sigma_{x_k}^2 + 2 \sum_{i=1}^n \sum_{j=1, j \neq i}^n \left[\frac{\partial^2 y}{\partial x_i \partial x_j} \right]_{\vec{x}=\vec{\mu}_x} \text{cov}(x_j, x_i) + \dots$$

La seconda espressione è la “legge di propagazione della varianza” nella sua formulazione generale.

Funzioni di più variabili casuali

Nel caso in cui le variabili siano indipendenti, i termini contenenti la covarianza sono nulli, e si ha:

$$\sigma_y^2 \cong \sum_{k=1}^n \left(\left[\frac{\partial y}{\partial x_k} \right]_{\vec{x}=\vec{\mu}_x} \right)^2 \sigma_{x_k}^2$$

Come caso particolare in cui $y = cx_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ si ottiene

$$\frac{\sigma_y^2}{y^2} = \alpha_1^2 \frac{\sigma_{x_1}^2}{x_1^2} + \alpha_2^2 \frac{\sigma_{x_2}^2}{x_2^2} + \dots + \alpha_n^2 \frac{\sigma_{x_n}^2}{x_n^2}$$

Ovvero la legge di propagazione delle varianze relative, il corrispondente della legge di propagazione dell'errore massimo relativo.

Prendiamo

$$z = \sum_{i=1}^n x_i$$

dove le x_i sono variabili casuali indipendenti ed hanno valore di aspettazione $E(x_i) = \mu_i$ e varianza σ_i^2 si ha

$$E[z] = \sum_{i=1}^n \mu_i (= n\mu), \quad \sigma_z^2 = \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n E[(x_i - \mu_i)^2] (= n\sigma^2)$$

queste due relazioni sono valide in generale, qualsiasi siano le funzioni di distribuzione delle variabili x_i . Se poi le x_i hanno distribuzione gaussiana, si può dimostrare che z ha una distribuzione gaussiana, con valore di aspettazione e varianza dati dalle espressioni precedenti.

Media aritmetica di più variabili casuali

Consideriamo ora la variabile casuale media aritmetica di n variabili casuali x_i indipendenti

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

tutte con funzione di distribuzione $f(x; \mu, \sigma^2)$, e quindi anche con stesso valore di aspettazione μ e stessa varianza σ^2 .

Per quanto visto è

$$\mu_{\bar{x}} = E[\bar{x}] = E\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n E[x_i] = \mu$$

e

$$\sigma_{\bar{x}}^2 = \text{var}(\bar{x}) = \frac{1}{n^2} \sum_{i=1}^n \text{var}(x_i) = \frac{1}{n} \sigma^2 \left[\text{sto applicando } \sigma_{\bar{x}}^2 = \sum_{i=1}^n \left(\frac{\partial \bar{x}}{\partial x_i}\right)^2 \sigma^2 \right]$$

La media aritmetica ha quindi una distribuzione centrata attorno al valore di aspettazione delle x_i . La sua varianza è invece n volte minore della varianza delle x_i : **la media ha una dispersione più piccola di $\frac{1}{n}$** attorno al valore centrale delle singole variabili x_i (il cosiddetto “campione”).

Compatibilità delle misure

Se si misura più volte la stessa grandezza fisica, i risultati devono essere compatibili.

- Nel caso della funzione di distribuzione di Gauss che stiamo considerando, i valori possono andare da $-\infty$ a $+\infty$, però la probabilità che assumano valori in intervalli molto lontani da a è molto piccola.
- La probabilità di ottenere un risultato che disti da a più di 3 deviazioni standard è circa 0.3%, per 4 deviazioni standard la probabilità è inferiore a 10^{-4} . La probabilità che due misure di a diano diverse deviazioni standard sarà piccola. Le misure si possono considerare, con alta probabilità, incompatibili.

Esistono test specifici per la compatibilità delle misure (“test di ipotesi”). Per ora ci basta quanto segue. Se $x_1 \pm \sigma_1$ e $x_2 \pm \sigma_2$ sono due misure indipendenti della stessa grandezza fisica, la loro differenza normalizzata alle deviazioni standard, non può essere molto diversa da zero:

$$E[X_1 - X_2] = E[X_1] - E[X_2] = \mu - \mu = 0$$

Considereremo le due misure incompatibili se:

$$|x_1 - x_2| \geq 3\sqrt{\sigma_1^2 + \sigma_2^2}$$

Misure ripetute, secondo approccio

Spesso è conveniente considerare gli n risultati x_i di uno stesso esperimento ripetuto n volte, come n variabili casuali indipendenti.

In base a quanto già visto, le x_i hanno tutte distribuzione gaussiana con stesso valore di aspettazione $\mu_x = a$ (si misura sempre la stessa grandezza fisica) e stessa varianza σ_x^2 (le misure vengono tutte eseguite con stesso metodo e stessi strumenti). Si può quindi scrivere la funzione di distribuzione congiunta

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sigma_x^n} e^{-\frac{\sum_{i=1}^n (x_i - \mu_x)^2}{2\sigma_x^2}}$$

Che permette di calcolare la probabilità di ottenere nelle n misure i valori effettivamente ottenuti.

$$P(x_1^m - \Delta x \leq x_1 \leq x_1^m + \Delta x, \dots, x_n^m - \Delta x \leq x_n \leq x_n^m + \Delta x) = \frac{(\Delta x)^n}{(2\pi)^{n/2} \sigma_x^n} e^{-\frac{\sum_{i=1}^n (x_i - \mu_x)^2}{2\sigma_x^2}}$$

Stima dei parametri

È questo, assieme al “test d’ipotesi”, uno dei capitoli più importanti dell’ “inferenza statistica”, che ha lo scopo di ricavare informazioni sull’intera popolazione partendo da un campione finito di questa. La seguente trattazione del problema della stima dei parametri è trattata in modo molto sintetico, ai soli scopi del corso

La stima dei parametri consiste nel valutare quantitativamente, a partire da un certo numero di misure, i parametri che intervengono in un determinato fenomeno statistico (ad esempio stimare μ e σ della distribuzione di Gauss).

In generale, dato un “campione” di dimensione n (x_1, x_2, \dots, x_n) si tratta di identificare una funzione del campione (la “stima” o “stimatore”) il cui valore di aspettazione sia il valore del parametro che si vuole determinare (lo stimatore è una funzione del campione e quindi è una variabile casuale che su campioni diversi assume valori diversi ed ha una sua funzione di distribuzione).

Lo stimatore può essere identificato usando metodi diversi, a volte equivalenti. È desiderabile, comunque, che abbia alcune proprietà ben definite, tra cui

- usare in modo efficiente tutta l’informazione a disposizione (a varianza minima)
- essere centrato (valore di aspettazione uguale al valore vero del parametro) anche per campioni di dimensione limitata.

Metodo del Maximum Likelihood

Per n sufficientemente grande fornisce stime con funzione di distribuzione di Gauss, a varianza minima, e centrate

Consideriamo un “campione” costituito da n misure $(x_1^m, x_2^m, \dots, x_n^m)$ di una variabile casuale x con funzione di distribuzione $f(x; \lambda_1, \lambda_2, \dots, \lambda_k)$ e funzione di distribuzione congiunta $f(\vec{x}; \lambda_1, \lambda_2, \dots, \lambda_k)$ e supponiamo di voler stimare i k parametri λ_j a partire dal campione.

Il metodo del Maximum Likelihood afferma che la stima migliore dei parametri è costituita da quei valori $\hat{\lambda}_j$ che massimizzano la funzione di Likelihood

$$\mathcal{L}(\lambda_1, \lambda_2, \dots, \lambda_k) = f(x_1^m, x_2^m, \dots, x_n^m; \lambda_1, \lambda_2, \dots, \lambda_k)$$

dove $f(x_1^m, x_2^m, \dots, x_n^m; \lambda_1, \lambda_2, \dots, \lambda_k)$ è la funzione densità di probabilità congiunta dei risultati delle misure e dipende da k parametri. Questa funzione, calcolata sul campione, è la **funzione densità di probabilità degli stimatori**.

Scegliere i parametri in modo tale che \mathcal{L} sia massima significa quindi scegliere i valori dei parametri che massimizzano la probabilità di ottenere i valori misurati della grandezza fisica.

Metodo del Maximum Likelihood

(o: metodo della massima verosimiglianza)

In base a questo metodo, quindi, gli stimatori dei parametri sono quelle funzioni $\hat{\lambda}_j$ del campione che massimizzano il valore della funzione di Likelihood.

Se \mathcal{L} è una funzione continua dei parametri $i = 1, k$, condizione necessaria e sufficiente affinché sia massima è che sia

$$\frac{\partial \ln \mathcal{L}}{\partial \lambda_i} = 0 \text{ per } i = 1, k$$

Queste k equazioni sono dette **equazioni di Likelihood** e gli stimatori dei parametri sono le soluzioni del sistema di k equazioni.

PS: la seconda condizione per cui i valori trovati con le **equazioni di Likelihood** siano massimi è:

$$\left(\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda_i^2} \right)_{\lambda_i = \hat{\lambda}_i} < 0 \text{ per } i = 1, k$$

Esempio 1: stima dei parametri della distribuzione di misure

Supponiamo che le n variabili x_i siano i risultati di n operazioni di misura ripetute nelle stesse condizioni della stessa grandezza fisica di “valore vero” a . Se gli errori accidentali non sono trascurabili, ciascuna variabile casuale x_i ha funzione di distribuzione

$$f(x_i; a, \sigma_x^2) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x_i-a)^2}{2\sigma_x^2}}$$

se le variabili x_i sono indipendenti, la funzione di distribuzione congiunta è

$$f(x_1, x_2, \dots, x_n; a, \sigma_x^2) = \frac{1}{(2\pi)^{n/2} \sigma_x^n} e^{-\frac{\sum_{i=1}^n (x_i-a)^2}{2\sigma_x^2}}$$

La funzione di Likelihood si ottiene sostituendo alle variabili casuali x_1, x_2, \dots, x_n i valori misurati $x_1^m, x_2^m, \dots, x_n^m$

$$\mathcal{L} = f(x_1^m, x_2^m, \dots, x_n^m; a, \sigma_x^2) = \frac{1}{(2\pi\sigma_x^2)^{n/2}} e^{-\frac{\sum_{i=1}^n (x_i^m - a)^2}{2\sigma_x^2}}$$

Il logaritmo del Likelihood vale:

$$\ln \mathcal{L} = -\frac{n}{2} \ln 2\pi\sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n (x_i^m - a)^2$$

Esempio 1: stima dei parametri della distribuzione di misure

Gli stimatori di a e σ_x^2 sono le funzioni del campione tali che

$$\frac{\partial}{\partial \lambda_i} \left(-\frac{n}{2} \ln 2\pi\sigma_x^2 - \frac{1}{2\sigma_x^2} \sum_{i=1}^n (x_i^m - a)^2 \right) = 0$$

ovvero

$$\frac{\partial \ln \mathcal{L}}{\partial a} = \frac{1}{2\sigma_x^2} \sum_{i=1}^n (x_i - a) = 0 \Rightarrow \hat{a} = \frac{1}{n} \sum_{i=1}^n x_i$$

e

$$\frac{\partial \ln \mathcal{L}}{\partial \sigma_x^2} = -\frac{n}{2\sigma_x^2} - \frac{1}{2(\sigma_x^2)^2} \sum_{i=1}^n (x_i - a)^2 = 0 \Rightarrow \hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - a)^2$$

che sono le medie aritmetiche dei valori misurati e degli scarti al quadrato.

Esempio 1: stima dei parametri della distribuzione di misure

Si può verificare che la stima di a è centrata:

$$E[\hat{a}] = \frac{1}{n} \sum_{i=1}^n E[x_i] = \mu$$

Per quanto riguarda invece $\hat{\sigma}_x^2$, nella sua espressione compare a , di cui si conosce solo la stima.. Sostituendo \hat{a} ad a ottengo

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{a})^2 = \frac{1}{n} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2$$

Per calcolare il valore di aspettazione uso la relazione:

$$\sum_{i=1}^n (x_i - a)^2 = \sum_{i=1}^n (x_i - \hat{a})^2 + n(\hat{a} - a)^2$$

Esempio 1: stima dei parametri della distribuzione di misure

Infatti

$$\begin{aligned}\sum_{i=1}^n (x_i - a)^2 &= \sum_{i=1}^n (x_i - \hat{a} + \hat{a} - a)^2 = \sum_{i=1}^n (x_i - \hat{a})^2 + 2 \sum_{i=1}^n (x_i - \hat{a})(\hat{a} - a) + \sum_{i=1}^n (\hat{a} - a)^2 \\ &= \sum_{i=1}^n (x_i - \hat{a})^2 + n(\hat{a} - a)^2 + 2 \sum_{i=1}^n (x_i \hat{a} - \hat{a} \hat{a} - x_i a + \hat{a} a)\end{aligned}$$

dove

$$\begin{aligned}\sum_{i=1}^n (x_i \hat{a} - \hat{a} \hat{a} - x_i a + \hat{a} a) &= \sum_{i=1}^n x_i \hat{a} - n \hat{a} \hat{a} - \sum_{i=1}^n x_i a + n \hat{a} a \\ &= n \hat{a} \left(\frac{1}{n} \sum_{i=1}^n x_i \right) - n \hat{a} \hat{a} - n a \left(\frac{1}{n} \sum_{i=1}^n x_i \right) + n \hat{a} a = 0\end{aligned}$$

Esempio 1: stima dei parametri della distribuzione di misure

Il cui valore di aspettazione è:

$$\begin{aligned} E[\hat{\sigma}_x^2] &= E\left[\frac{1}{n}\sum_{i=1}^n (x_i - \hat{a})^2\right] = \frac{1}{n} E\left[\sum_{i=1}^n (x_i - \hat{a})^2\right] = \\ &\frac{1}{n} E\left[\sum_{i=1}^n (x_i - a)^2 - n(\hat{a} - a)^2\right] = \frac{1}{n} E\left[\sum_{i=1}^n (x_i - a)^2\right] - E[(\hat{a} - a)^2] = \\ &\frac{\sum_{i=1}^n E[(x_i - a)^2]}{n} - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2 \end{aligned}$$

Ovvero:

$$E[\hat{\sigma}_x^2] = \frac{n-1}{n} \sigma^2$$

Esempio 1: stima dei parametri della distribuzione di misure

Quindi

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{a})^2$$

Ha un valore di aspettazione

$$E[\hat{\sigma}_x^2] = \frac{n-1}{n} \sigma_x^2$$

Ed è quindi centrato solo al tendere di $n \rightarrow \infty$, Per ottenere una stima centrata per n qualsiasi, è necessario moltiplicare l'espressione di $\hat{\sigma}_x^2$ per $n/(n-1)$ e definire la varianza del campione come

$$\hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{a})^2$$

Notare che il fattore $1/(n-1)$ fa sì che la stima non sia definita se $n = 1$, come ragionevole.

Esempio 1: stima dei parametri della distribuzione di misure

Posso anche fare il conto esplicito: il valore di aspettazione di $\hat{\sigma}_x^2$ è per definizione

$$E[\hat{\sigma}_x^2] = E \left[\frac{1}{n} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 \right]$$

Sviluppando:

$$E[\hat{\sigma}_x^2] = E \left[\frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{2}{n^2} \sum_{i,j=1}^n x_i x_j + \frac{1}{n^3} \sum_{i,j,k=1}^n x_j x_k \right] =$$
$$\frac{1}{n} \sum_{i=1}^n E[x_i^2] - \frac{2}{n^2} \sum_{i,j=1}^n E[x_i x_j] + \frac{1}{n^2} \sum_{i,j=1}^n E[x_i x_j] = \frac{1}{n} \sum_{i=1}^n E[x_i^2] - \frac{1}{n^2} \sum_{i,j=1}^n E[x_i x_j]$$

Ora, per eventi indipendenti, e per definizione di valori di aspettazione vale:

$$E[x_i x_j] = \begin{cases} E[x_i]E[x_j] = \mu^2 & \text{per } i \neq j \\ E[x_i^2] = \sigma^2 + \mu^2 & \text{per } i = j \end{cases}$$

Esempio 1: stima dei parametri della distribuzione di misure

Sostituendo:

$$E[x_i x_j] = \begin{cases} E[x_i]E[x_j] = \mu^2 & \text{per } i \neq j \\ E[x_i^2] = \sigma^2 + \mu^2 & \text{per } i = j \end{cases}$$

in

$$E[\hat{\sigma}_x^2] = \frac{1}{n} \sum_{i=1}^n E[x_i^2] - \frac{1}{n^2} \sum_{i,j=1}^n E[x_i x_j]$$

ottengo:

$$E[\hat{\sigma}_x^2] = \frac{1}{n} \sum_{i=1}^n (\sigma^2 + \mu^2) - \frac{1}{n^2} \sum_{i,j=1}^n (\delta_{ij} \sigma^2 + \mu^2)$$

$$E[\hat{\sigma}_x^2] = \sigma^2 + \mu^2 - \mu^2 - \frac{1}{n} \sigma^2 = \left(\frac{n-1}{n} \right) \sigma^2$$

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - a)^2 \rightarrow \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{a})^2$$

Esempio 1: stima dei parametri della distribuzione di misure

Date le n misure x_1, x_2, \dots, x_n , la miglior stima del valore di a è data quindi dal valor medio:

$$\hat{a} = \frac{1}{n} \sum_{i=1}^n x_i$$

Le n misure sono distribuite con distribuzione Gaussiana intorno ad \hat{a} con la miglior stima della deviazione standard

$$\hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{a})^2$$

Anche la variabile casuale \hat{a} segue una distribuzione Gaussiana. Possiamo calcolare $\sigma_{\hat{a}}^2$ in modo semplice usando la legge di propagazione della varianza, vista in precedenza, considerando che le variabili x_i sono indipendenti:

$$\sigma_{\hat{a}}^2 = \text{var}(\hat{a}) = \sum_{i=1}^n \left(\frac{\partial \hat{a}}{\partial x_i} \right)^2 \hat{\sigma}_x^2 = \frac{1}{n^2} \sum_{i=1}^n \hat{\sigma}_x^2 = \frac{\hat{\sigma}_x^2}{n} = \frac{n}{n^2} \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{a})^2 = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \hat{a})^2$$

Il risultato della misura della grandezza A scriverà quindi come

$$A = \hat{a} \pm \sigma_{\hat{a}}$$

Con il significato che la probabilità che l'intervallo $(\hat{a} - \sigma_{\hat{a}}, \hat{a} + \sigma_{\hat{a}})$ contenga il valore vero di A è del 68%.

Esempio 2: media pesata

Supponiamo ora che la grandezza fisica sia stata misurata n volte con tecniche diverse. Le n variabili casuali x_i , risultati delle n misure, avranno quindi distribuzioni con stesso valore di aspettazione a (“valore vero” della grandezza fisica), ma varianze σ_i^2 diverse e note. Le funzioni di distribuzione delle variabili x_i sono

$$f_i(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x_i-a)^2}{2\sigma_i^2}}$$

In questo caso, la funzione di Likelihood è

$$\mathcal{L} = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} e^{-\sum_{i=1}^n \frac{(x_i-a)^2}{2\sigma_i^2}}$$

Dobbiamo massimizzare il Likelihood per ottenere la miglior stima del valore di aspettazione e della varianza di a ; vale

$$\ln \mathcal{L} = \ln \left(\prod_{i=1}^n \sqrt{2\pi}\sigma_i \right) - \sum_{i=1}^n \frac{(x_i - a)^2}{2\sigma_i^2}$$

Esempio 2: media pesata

La miglior stima di a si ottiene da:

$$\frac{\partial \ln \mathcal{L}}{\partial a} = - \frac{\partial \sum_{i=1}^n \frac{(x_i - a)^2}{2\sigma_i^2}}{\partial a} = \sum_{i=1}^n \frac{x_i - a}{\sigma_i^2} = \sum_{i=1}^n \frac{x_i}{\sigma_i^2} - a \sum_{i=1}^n \frac{1}{\sigma_i^2} = 0 \quad \Rightarrow \quad \hat{a} = \frac{\sum_{i=1}^n \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

La varianza di \hat{a} può essere calcolata a partire dalla legge di propagazione della varianza

$$\sigma_{\hat{a}}^2 = \text{var}(\hat{a}) = \sum_{i=1}^n \left(\frac{\partial \hat{a}}{\partial x_i} \right)^2 \hat{\sigma}_{x_i}^2 = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

Queste sono le formule usate quando si vogliono combinare i risultati di diversi esperimenti. Naturalmente i risultati devono essere compatibili.

$$\sigma_{\hat{a}}^2 = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

Misura dell'accelerazione di gravità con pendolo semplice

Istogramma dei valori misurati

Siano state ripetute in condizioni identiche n volte le misure di una certa quantità con risultati dati dagli n valori x_i . Costruendo un istogramma con gli n valori x_i ottenuti, ci si aspetta che l'istogramma assuma una forma che ricordi la forma della funzione di Gauss. Il confronto può essere fatto in modo quantitativo:

- Se si conoscono i valori $\sigma_x = \sigma_a$ e $x = a$ possiamo calcolare la probabilità che un risultato cada nel generico intervallo k -esimo, centrato su x_k^* e di larghezza Δx . Tale probabilità vale:
$$p_k = \int_{x_k^* - \Delta x/2}^{x_k^* + \Delta x/2} f(x_k^*) dx = \Delta x f(x_k^*)$$
- Poiché il numero di eventi atteso segue con buona approssimazione una distribuzione binomiale, il numero di eventi previsto nell'intervallo k -esimo sarà pari a $n_k^c = np_k = n \Delta x f(x_k^*)$, con un'incertezza statistica $\sigma_{n_k} = \sqrt{np_k(1 - p_k)}$.
- Se il numero di misure è sufficientemente grande (maggiore di 10), il contenuto di probabilità associato alla deviazione standard è simile a quello che associamo alla distribuzione di Gauss.
- Posso ora confrontare il numero di eventi calcolato e misurato in ciascun intervallo, considerando la loro differenza in unità di deviazioni standard.

Confronto Istogramma-Gaussiana

Con le 100 misure del periodo del pendolo a 1 m, possiamo fare un confronto tra l'istogramma e la distribuzione Gaussiana.

Abbiamo visto che in caso di distribuzione Gaussiana, Maximum Likelihood o i Minimi quadrati forniscono come miglior stima dei parametri μ e σ della funzione di distribuzione di Gauss:

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

i valori $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$ e $\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2}$. Possiamo confrontare la distribuzione così ottenuta con il nostro istogramma della distribuzione delle misure.

Il numero di eventi contati n_k^m nel k-esimo intervallo di larghezza Δx deve essere confrontato con il numero di eventi atteso nello stesso intervallo a partire dalla funzione di distribuzione di Gauss; dato il valore della funzione nel centro dell'intervallo x_k , $f(x_k) = \frac{1}{\sqrt{2\pi}\hat{\sigma}} e^{-\frac{(x_k - \hat{\mu})^2}{2\hat{\sigma}^2}}$, la probabilità che un evento cada in quell'intervallo è:

$$P_k = P\left(x_k - \frac{\Delta x}{2} < x < x_k + \frac{\Delta x}{2}\right) = \Delta x f(x_k)$$

Il numero di eventi negli intervalli segue una distribuzione binomiale (P_k è la probabilità che l'evento cada nell'intervallo, $1 - P_k$ che cada fuori). Il numero di eventi atteso è $n_k^c = n_{tot} \Delta x f(x_k)$.

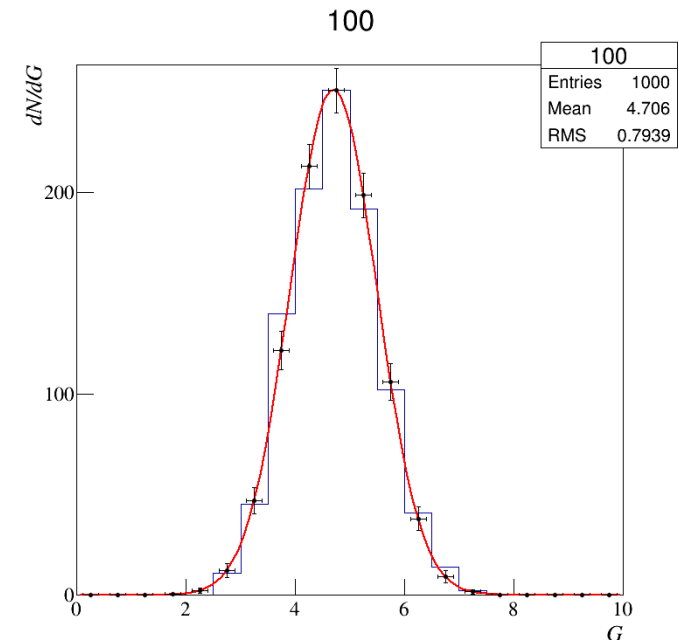
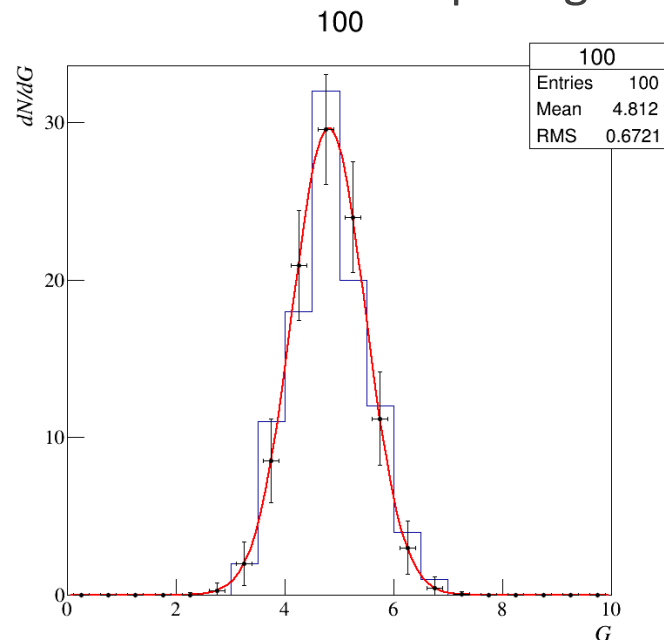
Confronto Istogramma-Gaussiana

La distribuzione del numero di eventi in ogni bin segue la distribuzione binomiale dove P_k è la probabilità che l'evento cada nell'intervallo k-esimo ed $q_k = 1 - P_k$ la probabilità che l'evento non cada nell'intervallo. La deviazione standard della binomiale è

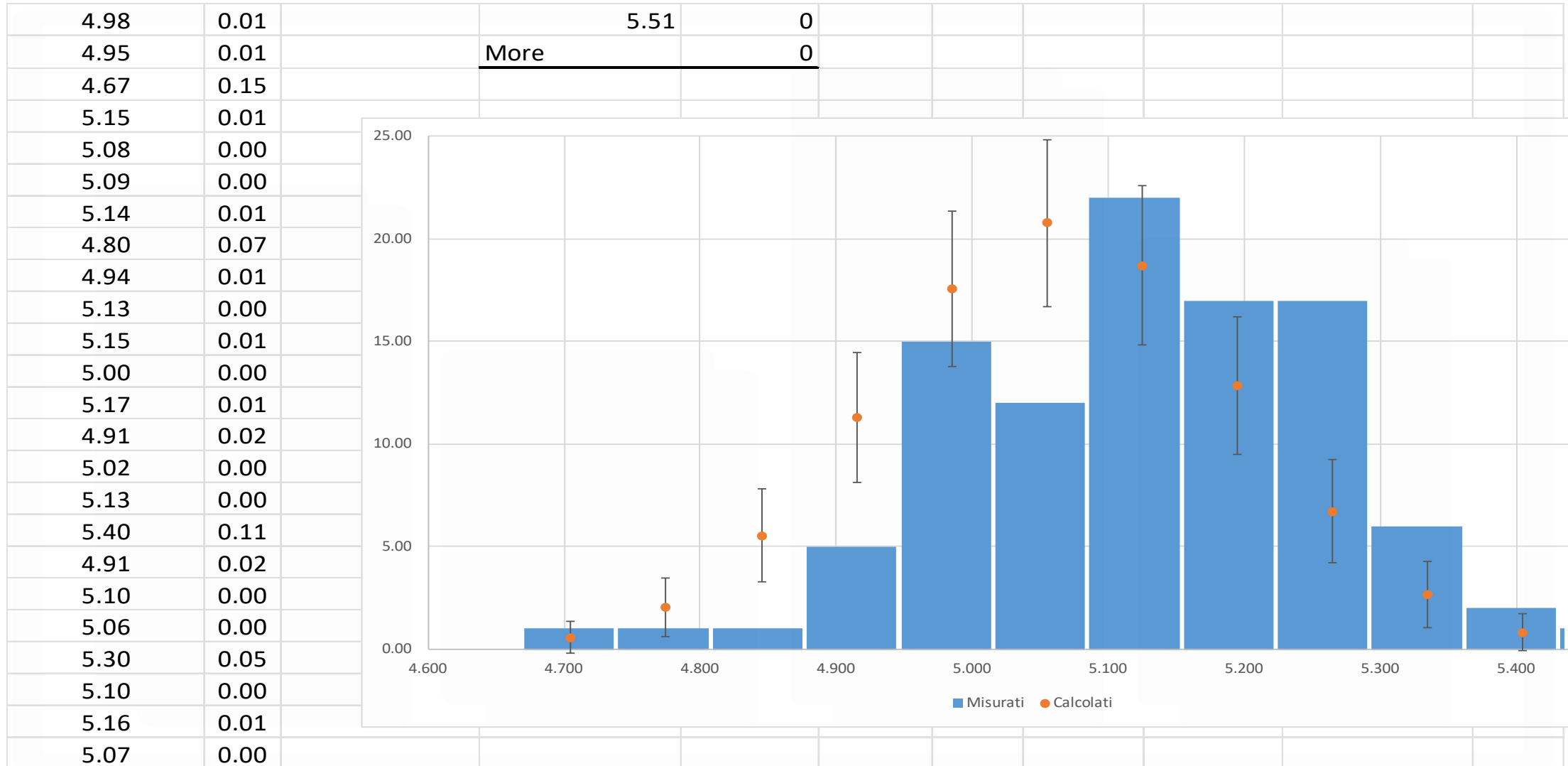
$$\sigma_{n_k} = \sqrt{n_{tot} P_k (1 - P_k)} = \sqrt{n_k^c (1 - n_k^c / n_{tot})} \left(\approx \sqrt{n_k^c} \text{ per } n_{tot} \gg n_k^c \right)$$

Posso quindi confrontare il numero di eventi atteso, ed il numero di eventi effettivamente contato e vedere se per ogni intervallo dell'istogramma vale:

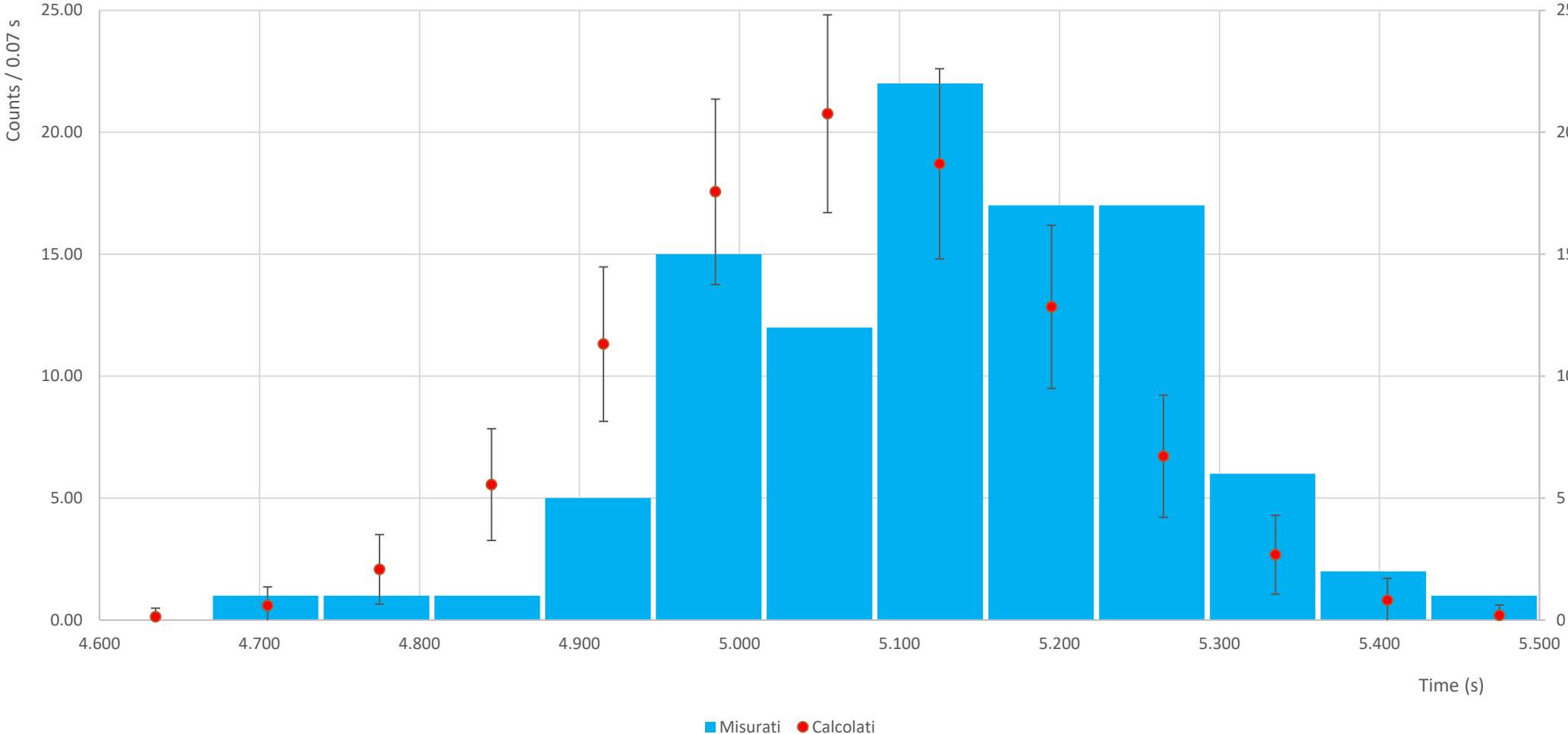
$$|n_k^m - n_k^c| < 3\sigma_{n_k}$$



Example



Example II



Metodo dei Minimi Quadrati

Il metodo dei **Minimi Quadrati (Least-Squares method)** è uno dei metodi tradizionalmente più utilizzati nella stima di parametri, anche se non possiede le proprietà matematiche che sono ascrivibili al metodo della **Massima Verosimiglianza (Maximum Likelihood method)**.

Deve il suo successo al fatto che per un'importante classe di problemi, ovvero quelli in cui **la dipendenza dai parametri è lineare**, fornisce stime «unbias» e con varianza minima, anche nel caso di una dimensione limitata del campione. Parliamo in questo caso del metodo dei **Minimi Quadrati Lineari**.

Principio del metodo dei Minimi Quadrati

Siano dati, per un set di N punti di misura:

$$\vec{x} = \{x_1, x_2, \dots, x_N\}$$

Un set di N risultati (valori) sperimentali indipendenti:

$$\vec{y} = \{y_1, y_2, \dots, y_N\}$$

Di cui non conosciamo i valori veri $\vec{\eta} = \{\eta_1, \eta_2, \dots, \eta_N\}$ ma per cui possiamo assumere un modello teorico che permette di predire questi valori veri per ognuno dei punti di misura attraverso una dipendenza funzionale:

$$f_i = f_i(\theta_1, \theta_2, \dots, \theta_L; x_i) \quad \text{con } L \leq N$$

In questo caso il principio del metodo dei Minimi Quadrati afferma che la miglior stima del set di parametri $\{\theta_1, \theta_2, \dots, \theta_L\}$ è fornita dal valore che rende minima la quantità X^2 :

$$X^2 = \sum_{i=1}^N w_i (y_i - f_i)^2$$

Dove i w_i sono i «pesi» riferiti all'osservabile i -esima.

Indichiamo con $\underline{\hat{\theta}} = \{\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_L\}$ il valore del set di parametri ottenuto dalla minimizzazione di X^2 . Tale set di valori denota quindi la «Stima con il metodo dei minimi quadrati» del set di parametri.

Pesi in casi particolari

Nel caso particolare in cui i valori misurati $\vec{y} = \{y_1, y_2, \dots, y_N\}$ del set di variabili indipendenti siano tutti affetti dalla stessa incertezza (nota ed = σ , o non nota) i pesi all'interno del calcolo del X^2 sono tutti uguali e quindi il loro valore non entra nella minimizzazione:

$$X^2 = \sum_{i=1}^N (y_i - f_i)^2$$

E parliamo in questo caso stima con il metodo dei minimi quadrati non pesati (unweighted Least-Squares estimation).

Nel caso in cui le incertezza sulle y_i (sempre indipendenti) sono tutte diverse tra loro, ma note, il valore dei pesi è normalmente preso pari alla precisione della misura

$$w_i = \frac{1}{\sigma_i^2}$$

E la funzione di X^2 da minimizzare assume la forma:

$$X^2 = \sum_{i=1}^N \left(\frac{y_i - f_i}{\sigma_i} \right)^2$$

Pesi in casi particolari

Infine, nel caso in cui le y_i variabili siano correlate, ovvero nel caso in cui le covarianze $\text{cov}(y_i, y_j)$ siano diverse da zero, e che quindi sia definita una matrice delle covarianze V il valore dei pesi è dato dall'inverso delle covarianze

$$w_i = \frac{1}{\text{cov}(y_i, y_j)} = V_{ij}^{-1}$$

E la funzione di X^2 da minimizzare assume la forma:

$$X^2 = \sum_{i=1}^N \sum_{j=1}^N (y_i - f_i) V_{ij}^{-1} (y_j - f_j)$$

Commenti

- In tutta la discussione abbiamo assunto che gli errori sui punti di misura x_i (cause nei grafici) siano assenti o trascurabili rispetto agli errori sulle y_i (effetti nei grafici).
 - Nel caso il punto di misura definisce una regione che va da x_i a $x_i + \Delta x_i$ l'assunto è che sia possibile comunque fare una precisa valutazione del valore di f_i corrispondente al valore x_i o all'intervallo $x_i + \Delta x_i$.
 - Allo stesso modo il risultato della misura y_i può essere considerato come il valor medio nell'intervallo.
 - In alcuni case (i.e. quando la f_i varia rapidamente in tali intervalli $[x_i - \Delta x_i, x_i + \Delta x_i]$) può essere necessario introdurre l'integrale della f_i nell'intervallo (e corrispondentemente il valore $y_i \Delta x_i$). Il processo di minimizzazione per la stima dei parametri non cambia.
- Il metodo dei minimi quadrati non assume ipotesi sulle funzioni di distribuzione delle osservabili y_i e come tale le sue stime non dipendono dalle distribuzioni. Se però le osservabili seguono delle funzioni di distribuzione Gaussiana la distribuzione di X^2 al minimo X_{min}^2 sarà distribuita come la funzione di distribuzione di **Chi-Quadrato (Chi-Squared) o χ^2** . Ciò permette di valutare la bontà del modello $f_i = f_i(\theta_1, \theta_2, \dots, \theta_L; x_i)$ usato.

Esempi: costante – misure con stesse σ

Immaginiamo di aver fatto un set di misure di una grandezza in identiche condizioni per le cause, ovvero x costante per i diversi y_i e che gli errori sui diversi y_i siano anch'essi costanti e che quindi possiamo usare i minimi quadrati lineari non pesati. Visto che x non cambia, ci aspettiamo un modello costante

$$f_i = \theta_1$$

che descriva un certo set di misure

$$(x, y_1), (x, y_2) \cdots (x, y_N)$$

Se assumiamo che gli errori possano essere trascurati (o che sono tutti uguali) possiamo costruire la funzione X^2 per il nostro set come:

$$X^2 = \sum_{i=1}^N (y_i - f_i)^2 = \sum_{i=1}^N (y_i - \theta_1)^2$$

Il minimo del X^2 per θ_1 si ottiene ponendo uguale a zero la derivata

Esempi: costante – misure con stesse σ

Il minimo del X^2 per il parametro θ_1 si ottiene ponendo uguale a zero la derivata:

$$\frac{\partial X^2}{\partial \theta_1} = \sum_{i=1}^N (2)(y_i - \hat{\theta}_1)(-1) = 0$$

Da cui ottengo:

$$\sum_{i=1}^N y_i - N\hat{\theta}_1 = 0$$

Ovvero:

$$\hat{\theta}_1 = \frac{1}{N} \sum_{i=1}^N y_i$$

Quindi, la miglior stima del valore vero delle N misure y_i è data dalla media aritmetica. Ottengo poi la varianza su $\hat{\theta}_1$ usando la legge di propagazione della varianza.

Esempi: costante – misure con diverse σ

Immaginiamo di aver fatto un set di misure di una grandezza in identiche condizioni per le cause, ovvero x costante per i diversi y_i e che le incertezze su y_i varino σ_i . Visto che x non cambia, ci aspettiamo un modello costante

$$f_i = \theta_1$$

che descriva un certo set di misure

$$(x, y_1), (x, y_2) \cdots (x, y_N)$$

la funzione X^2 per il nostro set viene scritta nel modo seguente:

$$X^2 = \sum_{i=1}^N \left(\frac{y_i - f_i}{\sigma_i} \right)^2 = \sum_{i=1}^N \left(\frac{y_i - \theta_1}{\sigma_i} \right)^2$$

Il minimo del X^2 per θ_1 si ottiene ponendo uguale a zero la derivata

$$\frac{\partial X^2}{\partial \theta_1} = \sum_{i=1}^N (2) \frac{1}{\sigma_i^2} (y_i - \hat{\theta}_1)(-1) = 0$$

Esempi: costante – misure con diverse σ

Da cui ottengo:

$$\sum_{i=1}^N \frac{y_i}{\sigma_i^2} - \hat{\theta}_1 \sum_{i=1}^N \frac{1}{\sigma_i^2} = 0$$

Ovvero:

$$\hat{\theta}_1 = \frac{\sum_{i=1}^N \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2}}$$

Quindi, la miglior stima del valore vero delle N misure y_i è data dalla media pesata con pesi w_i pari a:

$$w_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2}}$$

Ottengo poi la varianza su $\hat{\theta}_1$ usando la legge di propagazione della varianza.

Fit di una retta – «regressione lineare»

Come ultimo esempio dell'uso del metodo dei minimi quadrati (lineari) vediamo come otteniamo i parametri nel caso del fit di una retta in due CASI – 1) errori uguali o trascurabili e 2) incertezze con distribuzioni gaussiane.

Vogliamo cioè trovare la miglior retta

$$y = mx + q$$

che descriva un certo set di misure

$$(x_1, y_1 \pm \sigma_1), (x_2, y_2 \pm \sigma_2) \cdots (x_N, y_N \pm \sigma_N)$$

I nostri parametri sono quindi $\theta_1 = q$ e $\theta_2 = m$ e la parametrizzazione che usiamo è:

$$f_i = \theta_1 + \theta_2 x_i$$

Se assumiamo che gli errori possano essere trascurati, o che sono tutti uguali ($\sigma_i = \sigma$) possiamo costruire la funzione X^2 per il nostro set come:

$$X^2 = \sum_{i=1}^N (y_i - f_i)^2 = \sum_{i=1}^N (y_i - \theta_1 - \theta_2 x_i)^2$$

Il minimo del X^2 per i due parametri θ_1 e θ_2 si ottiene ponendo uguale a zero le rispettive derivate

Fit di una retta – «regressione lineare»

Il minimo del X^2 per i due parametri θ_1 e θ_2 si ottiene ponendo uguale a zero le rispettive derivate:

$$\frac{\partial X^2}{\partial \theta_1} = \sum_{i=1}^N (2)(y_i - \theta_1 - \theta_2 x_i)(-1) = 0$$

$$\frac{\partial X^2}{\partial \theta_2} = \sum_{i=1}^N (2)(y_i - \theta_1 - \theta_2 x_i)(-x_i) = 0$$

Possiamo quindi scrivere le sommatorie separatamente, trascurando il termine costante:

$$\begin{aligned} \sum_{i=1}^N y_i - N\theta_1 - \theta_2 \sum_{i=1}^N x_i &= 0 \\ \sum_{i=1}^N y_i x_i - \theta_1 \sum_{i=1}^N x_i - \theta_2 \sum_{i=1}^N x_i^2 &= 0 \end{aligned}$$

Che costituisce un sistema di equazioni in due incognite θ_1 e θ_1

Fit di una retta – «regressione lineare»

Risolvero quindi il sistema di equazioni in due incognite θ_1 e θ_2 (ad esempio moltiplicando la prima eq per $\sum_{i=1}^N x_i/N$ e sottraendo membri a membro):

$$\begin{cases} N\theta_1 + \theta_2 \sum_{i=1}^N x_i = \sum_{i=1}^N y_i \\ \theta_1 \sum_{i=1}^N x_i + \theta_2 \sum_{i=1}^N x_i^2 = \sum_{i=1}^N y_i x_i \end{cases}$$

Per ottenere la stima dei parametri ovvero $q = \hat{\theta}_1$ e $m = \hat{\theta}_2$. Il risultato è:

$$q = \hat{\theta}_1 = \frac{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i x_i}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i\right)^2}$$

$$m = \hat{\theta}_2 = \frac{N \sum_{i=1}^N y_i x_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i\right)^2}$$

Fit di una retta – «regressione lineare»

Possiamo ulteriormente riscrivere il risultato introducendo in valori medi di x e y ovvero:

$$\bar{x}N = s_x = \sum_{i=1}^N x_i, \quad \bar{y}N = s_y = \sum_{i=1}^N y_i, \quad \overline{x^2}N = s_{xx} = \sum_{i=1}^N x_i^2, \quad \overline{xy}N = s_{xy} = \sum_{i=1}^N y_i x_i$$

Da cui:

$$q = \hat{\theta}_1 = \frac{s_y \cdot s_{xx} - s_x \cdot s_{xy}}{N s_{xx} - s_x \cdot s_x} = \frac{\bar{y} \cdot \overline{x^2} - \bar{x} \cdot \overline{xy}}{\overline{x^2} - \bar{x}^2}$$

$$m = \hat{\theta}_2 = \frac{N s_{xy} - s_x \cdot s_y}{N s_{xx} - s_x \cdot s_x} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}$$

Otengo gli errori sui parametri dalla legge di propagazione della varianza:

$$\sigma_{\hat{\theta}_1}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_1}{\partial y_i} \right)^2 \sigma^2, \quad \sigma_{\hat{\theta}_2}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_2}{\partial y_i} \right)^2 \sigma^2$$

Fit di una retta – «regressione lineare»

Ottengo gli errori sui parametri dalla legge di propagazione della varianza:

$$\sigma_{\hat{\theta}_1}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_1}{\partial y_i} \right)^2 \sigma^2 \qquad \sigma_{\hat{\theta}_2}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_2}{\partial y_i} \right)^2 \sigma^2$$

Con gli errori σ uguali in tutte le misure:

$$\sigma_{\hat{\theta}_1}^2 = \sigma^2 \sum_{j=1}^N \left(\frac{\partial \frac{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i x_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}}{\partial y_j} \right)^2$$

$$\sigma_{\hat{\theta}_2}^2 = \sigma^2 \sum_{j=1}^N \left(\frac{\partial \frac{N \sum_{i=1}^N y_i x_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}}{\partial y_j} \right)^2$$

Fit di una retta – $\sigma_{\hat{\theta}_1}^2$

$$\begin{aligned}
 \sigma_{\hat{\theta}_1}^2 &= \sigma^2 \sum_{j=1}^N \left(\frac{\partial \frac{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i x_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}}{\partial y_j} \right)^2 = \\
 &= \frac{\sigma^2}{\left(N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2 \right)^2} \sum_{j=1}^N \left(\sum_{i=1}^N x_i^2 - x_j \sum_{i=1}^N x_i \right)^2 = \\
 &= \frac{\sigma^2}{\left(N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2 \right)^2} \sum_{j=1}^N \left[\left(\sum_{i=1}^N x_i^2 \right)^2 - 2x_j \sum_{i=1}^N x_i \sum_{i=1}^N x_i^2 + x_j^2 \left(\sum_{i=1}^N x_i \right)^2 \right] = \\
 &= \sigma^2 \frac{N \left(\sum_{i=1}^N x_i^2 \right)^2 - 2 \left(\sum_{i=1}^N x_i \right)^2 \sum_{i=1}^N x_i^2 + \sum_{i=1}^N x_i^2 \left(\sum_{i=1}^N x_i \right)^2}{\left(N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2 \right)^2} = \sigma^2 \frac{\left[N \sum_{i=1}^N x_i^2 - 2 \left(\sum_{i=1}^N x_i \right)^2 + \left(\sum_{i=1}^N x_i \right)^2 \right] \sum_{i=1}^N x_i^2}{\left(N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2 \right)^2} = \\
 &= \sigma^2 \frac{\sum_{i=1}^N x_i^2}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}
 \end{aligned}$$

Fit di una retta – $\sigma_{\hat{\theta}_2}^2$

$$\begin{aligned}
 \sigma_{\hat{\theta}_2}^2 &= \sigma^2 \sum_{j=1}^N \left(\frac{\partial \frac{N \sum_{i=1}^N y_i x_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}}{\partial y_j} \right)^2 = \\
 &= \frac{\sigma^2}{(N s_{xx} - (s_x)^2)^2} \sum_{j=1}^N \left(N x_j - \sum_{i=1}^N x_i \right)^2 = \\
 &= \frac{\sigma^2}{(N s_{xx} - (s_x)^2)^2} \sum_{j=1}^N [N^2 x_j^2 - 2N x_j s_x + (s_x)^2] = \\
 &= \sigma^2 \frac{N^2 s_{xx} - 2N(s_x)^2 + N(s_x)^2}{(N s_{xx} - (s_x)^2)^2} = \sigma^2 \frac{N[N s_{xx} - (s_x)^2]}{(N s_{xx} - (s_x)^2)^2} \\
 \sigma_{\hat{\theta}_2}^2 &= \sigma^2 \frac{N}{(N s_{xx} - (s_x)^2)^2} = \sigma^2 \frac{N}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}
 \end{aligned}$$

Fit di una retta – «regressione lineare»

Abbiamo quindi infine:

$$\sigma_{\hat{\theta}_1}^2 = \sigma^2 \frac{\sum_{i=1}^N x_i^2}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} = \sigma^2 \frac{s_{xx}}{N s_{xx} - (s_x)^2} = \frac{\sigma^2}{N} \frac{\overline{x^2}}{(\overline{x^2} - \bar{x}^2)}$$

$$\sigma_{\hat{\theta}_2}^2 = \sigma^2 \frac{N}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} = \sigma^2 \frac{N}{N s_{xx} - (s_x)^2} = \frac{\sigma^2}{N} \frac{1}{(\overline{x^2} - \bar{x}^2)}$$

Fit di una retta – «regressione lineare»

Vogliamo cioè trovare la miglior retta

$$y = mx + q$$

che descriva un certo set di misure

$$(x_1, y_1 \pm \sigma_1), (x_2, y_2 \pm \sigma_2) \cdots (x_N, y_N \pm \sigma_N)$$

I nostri parametri sono quindi $\theta_1 = q$ e $\theta_2 = m$ e la parametrizzazione che usiamo è:

$$f_i = \theta_1 + \theta_2 x_i$$

Nel caso di errori distribuiti in modo Gaussiano e tra loro diversi (σ_i) La probabilità che le operazioni di misura forniscano il valore osservato y_i e' data da:

$$f(y_1, y_2, \dots, y_N) \sim \frac{1}{\prod_{i=1}^N \sigma_i} e^{-\sum_{i=1}^N \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{2\sigma_i^2}}$$

la funzione X^2 coincide con la distribuzione di χ^2 :

$$X^2 \equiv \chi^2 = \sum_{i=1}^N \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{\sigma_i^2}$$

Il minimo del X^2 per i due parametri θ_1 e θ_2 si ottiene ponendo uguale a zero le rispettive derivate

Fit di una retta – «regressione lineare»

Il minimo del χ^2 per i due parametri θ_1 e θ_2 si ottiene ponendo uguale a zero le rispettive derivate:

$$\frac{\partial \chi^2}{\partial \theta_1} = -2 \sum_{i=1}^N \frac{(y_i - \theta_1 - \theta_2 x_i)}{\sigma_i^2} = 0$$

$$\frac{\partial \chi^2}{\partial \theta_2} = -2 \sum_{i=1}^N \frac{(x_i y_i - x_i \theta_1 - \theta_2 x_i^2)}{\sigma_i^2} = 0$$

Possiamo quindi scrivere le sommatorie separatamente, trascurando il termine costante:

$$\sum_{i=1}^N \frac{y_i}{\sigma_i^2} - \theta_1 \sum_{i=1}^N \frac{1}{\sigma_i^2} - \theta_2 \sum_{i=1}^N \frac{x_i}{\sigma_i^2} = 0$$
$$\sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} - \theta_1 \sum_{i=1}^N \frac{x_i}{\sigma_i^2} - \theta_2 \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} = 0$$

Che costituisce un sistema di equazioni in due incognite θ_1 e θ_1

Fit di una retta – «regressione lineare»

Risolvero quindi il sistema di equazioni in due incognite θ_1 e θ_2 :

$$\begin{cases} \theta_1 \sum_{i=1}^N \frac{1}{\sigma_i^2} + \theta_2 \sum_{i=1}^N \frac{x_i}{\sigma_i^2} = \sum_{i=1}^N \frac{y_i}{\sigma_i^2} \\ \theta_1 \sum_{i=1}^N \frac{x_i}{\sigma_i^2} + \theta_2 \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} = \sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} \end{cases}$$

Otengo (come in precedenza) per $q = \hat{\theta}_1$ e $m = \hat{\theta}_2$:

$$q = \hat{\theta}_1 = \frac{\sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \sum_{i=1}^N \frac{y_i}{\sigma_i^2} - \sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} \sum_{i=1}^N \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} - \left(\sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right)^2}$$

$$m = \hat{\theta}_2 = \frac{\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} - \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \sum_{i=1}^N \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} - \left(\sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right)^2}$$

Fit di una retta – «regressione lineare»

Introduco per semplicità la seguente scrittura

$$s_x = \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \quad s_y = \sum_{i=1}^N \frac{y_i}{\sigma_i^2}$$
$$s_{xx} = \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \quad s_{xy} = \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2} \quad s = \sum_{i=1}^N \frac{1}{\sigma_i^2}$$

Che mi permette di scrivere in forma più sintetica:

$$q = \hat{\theta}_1 = \frac{s_{xx} \cdot s_y - s_{xy} \cdot s_x}{s \cdot s_{xx} - s_x^2}$$
$$m = \hat{\theta}_2 = \frac{s \cdot s_{xy} - s_x \cdot s_y}{s \cdot s_{xx} - s_x^2}$$

Otengo (come prima) gli errori sui parametri dalla legge di propagazione della varianza:

$$\sigma_{\hat{\theta}_1}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_1}{\partial y_i} \right)^2 \sigma_i^2 \quad \sigma_{\hat{\theta}_2}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_2}{\partial y_i} \right)^2 \sigma_i^2$$

Fit di una retta – «regressione lineare»

Ottengo gli errori sui parametri dalla legge di propagazione della varianza:

$$\sigma_{\hat{\theta}_1}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_1}{\partial y_i} \right)^2 \sigma_i^2 \qquad \sigma_{\hat{\theta}_2}^2 = \sum_{i=1}^N \left(\frac{\partial \hat{\theta}_2}{\partial y_i} \right)^2 \sigma_i^2$$

Con gli errori σ diversi nelle varie misure:

$$\sigma_{\hat{\theta}_1}^2 = \sum_{j=1}^N \left(\frac{\partial \left(\frac{\sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \sum_{i=1}^N \frac{y_i}{\sigma_i^2} - \sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} \sum_{i=1}^N \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} - \left(\sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right)^2} \right)}{\partial y_j} \right)^2 \sigma_j^2$$

$$\sigma_{\hat{\theta}_2}^2 = \sum_{j=1}^N \left(\frac{\partial \left(\frac{\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} - \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \sum_{i=1}^N \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} - \left(\sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right)^2} \right)}{\partial y_j} \right)^2 \sigma_j^2$$

Fit di una retta – «regressione lineare»

$$\frac{\partial}{\partial y_j} \frac{\sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \sum_{i=1}^N \frac{y_i}{\sigma_i^2} - \sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} \sum_{i=1}^N \frac{x_i}{\sigma_i^2}}{s \cdot s_{xx} - (s_x)^2} = \frac{1}{s \cdot s_{xx} - (s_x)^2} \left[\frac{1}{\sigma_j^2} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} - \frac{x_j}{\sigma_j^2} \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right]$$

$$\frac{\partial}{\partial y_j} \frac{\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{i=1}^N \frac{y_i x_i}{\sigma_i^2} - \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \sum_{i=1}^N \frac{y_i}{\sigma_i^2}}{s \cdot s_{xx} - (s_x)^2} = \frac{1}{s \cdot s_{xx} - (s_x)^2} \left[\frac{x_j}{\sigma_j^2} \sum_{i=1}^N \frac{1}{\sigma_i^2} - \frac{1}{\sigma_j^2} \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right] = \frac{\frac{x_j}{\sigma_j^2} s - \frac{1}{\sigma_j^2} s_x}{s \cdot s_{xx} - (s_x)^2}$$

$$\frac{\partial \hat{\theta}_1}{\partial y_j} = \frac{1}{s \cdot s_{xx} - (s_x)^2} \left[\sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} - \frac{x_j}{\sigma_j^2} \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right] = \frac{s_{xx} \frac{1}{\sigma_j^2} - \frac{x_j}{\sigma_j^2} s_x}{s \cdot s_{xx} - (s_x)^2}$$

Fit di una retta – «regressione lineare»

Abbiamo cioè:

$$\sigma_{\hat{\theta}_1}^2 = \sum_{j=1}^N \left(\frac{\partial \hat{\theta}_1}{\partial y_j} \right)^2 \sigma_j^2 = \sigma_{\hat{\theta}_1}^2 = \sum_{j=1}^N \left(\frac{S_{xx} \frac{1}{\sigma_j^2} - \frac{x_j}{\sigma_j^2} S_x}{S \cdot S_{xx} - (S_x)^2} \right)^2 \sigma_j^2$$

$$\sigma_{\hat{\theta}_2}^2 = \sum_{j=1}^N \left(\frac{\partial \hat{\theta}_2}{\partial y_j} \right)^2 \sigma_j^2 = \sigma_{\hat{\theta}_2}^2 = \sum_{j=1}^N \left(\frac{\frac{x_j}{\sigma_j^2} S - \frac{1}{\sigma_j^2} S_x}{S \cdot S_{xx} - (S_x)^2} \right)^2 \sigma_j^2$$

Fit di una retta – «regressione lineare»

sviluppiamo:

$$\sigma_{\hat{\theta}_1}^2 = \frac{1}{(s \cdot s_{xx} - (s_x)^2)^2} \sum_{j=1}^N \left[(s_{xx})^2 \frac{1}{\sigma_j^4} - 2s_{xx} \cdot s_x \frac{x_j}{\sigma_j^4} + \frac{x_j^2}{\sigma_j^4} (s_x)^2 \right] \sigma_j^2$$

$$\sigma_{\hat{\theta}_2}^2 = \frac{1}{(s \cdot s_{xx} - (s_x)^2)^2} \sum_{j=1}^N \left[(s)^2 \frac{x_j^2}{\sigma_j^4} - 2s_x \cdot s \frac{x_j}{\sigma_j^4} + \frac{1}{\sigma_j^4} (s_x)^2 \right] \sigma_j^2$$

passaggio successivo, le sommatorie:

$$\sigma_{\hat{\theta}_1}^2 = \frac{(s_{xx})^2 \cdot s - 2s_{xx}(s_x)^2 + s_{xx}(s_x)^2}{(s \cdot s_{xx} - (s_x)^2)^2} = \frac{s_{xx}(s \cdot s_{xx} - (s_x)^2)}{(s \cdot s_{xx} - (s_x)^2)^2}$$

$$\sigma_{\hat{\theta}_2}^2 = \frac{(s)^2 \cdot s_{xx} - 2s(s_x)^2 + s(s_x)^2}{(s \cdot s_{xx} - (s_x)^2)^2} = \frac{s(s \cdot s_{xx} - (s_x)^2)}{(s \cdot s_{xx} - (s_x)^2)^2}$$

Fit di una retta – «regressione lineare»

Da cui si ricava:

$$\sigma_q^2 = \sigma_{\hat{\theta}_1}^2 = \frac{S_{xx}}{S \cdot S_{xx} - S_x^2}$$

$$\sigma_m^2 = \sigma_{\hat{\theta}_2}^2 = \frac{S}{S \cdot S_{xx} - S_x^2}$$

Statistica: che cosa abbiamo imparato

- Distribuzione degli errori casuali: Gaussiana (centrata in 0)
- Distribuzione delle misure affette da errori casuali: Gaussiana (centrata in μ)
 - Stima di μ : $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$
 - Stima di σ del campione: $\hat{\sigma}_\mu^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2$
 - Stima di σ della media: $\hat{\sigma}_{\hat{\mu}}^2 = \frac{1}{n} \hat{\sigma}_\mu^2 = \frac{1}{n} \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2$
- Media pesata per misure con stesso valore di aspettazione ma varianze diverse
- Variabili aleatorie, funzioni di variabili aleatorie (indipendenti)
 - $\mu_y \cong y(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) + \frac{1}{2} \sum_{k=1}^n \left[\frac{\partial^2 y}{\partial^2 x_k} \right]_{\vec{x}=\vec{\mu}_x} \sigma_{x_k}^2$
 - $\sigma_y^2 \cong \sum_{k=1}^n \left(\left[\frac{\partial y}{\partial x_k} \right]_{\vec{x}=\vec{\mu}_x} \right)^2 \sigma_{x_k}^2$
- Significato probabilistico del risultato e compatibilità di misure affette da errori casuali
- Alcune importanti funzioni di distribuzione
- Cenni sull'inferenza statistica

Misura dei moduli di Young e Coulomb dell'acciaio

Deformazioni elastiche

Questa coppia di esperienze è finalizzata allo studio delle **proprietà elastiche dei materiali sottoposti a sforzi di trazione o di taglio**.

Con il termine di **corpo rigido** si indica un solido indeformabile, ossia un oggetto che mantiene la propria forma e volume sotto l'azione di qualsiasi sollecitazione. **La nozione di corpo rigido è un'astrazione**.

Tutti i solidi, se sottoposti a sollecitazioni quali trazione, compressione, torsione, scorrimento, si deformano.

Se le sollecitazioni sono intense le deformazioni possono essere permanenti e si può anche arrivare al **cedimento strutturale** (rottura) del materiale.

Deformazioni elastiche

Una deformazione è detta **elastica** se il corpo torna allo **stato originario** quando vengono meno le forze che ne hanno causato la deformazione.

La deformazione è **elastica** se le forze applicate sono inferiori ad un limite che dipende dal materiale, dalla temperatura, dal tipo di deformazione considerata, etc.

Per le deformazioni elastiche si osserva una relazione di **proporzionalità tra sollecitazione e deformazione**.

$$\Delta L \propto F$$

Questo comportamento è noto come **legge di Hooke**.

La **legge di Hooke** è valida per la maggior parte dei minerali, per il vetro, per i materiali ceramici e per i metalli. Per i metalli duttili è vera per carichi modesti.

Allungamenti per trazione

La costante elastica k dipende dal **materiale**, dalla **temperatura** e dalle **caratteristiche geometriche del solido**, ovvero dalla sua lunghezza lungo la direzione di trazione e dalla sezione trasversa rispetto a tale direzione.

$$k(\text{materiale}, T, S, L)$$

Si osserva in particolare che, a parità di lunghezza, la costante elastica di un materiale cresce in **modo proporzionale alla superficie**; raddoppiando cioè la superficie del materiale la costante elastica raddoppia. Per avere lo stesso allungamento devo quindi raddoppiare la forza applicata.

$$k(S, L) \propto S$$

Si osserva anche che, a parità di superficie, la costante elastica decresce in modo inversamente proporzionale alla lunghezza; raddoppiando la lunghezza del filo la costante elastica si dimezza. Per avere lo stesso allungamento devo quindi dimezzare la forza applicata

$$k(S, L) \propto \frac{1}{L}$$

Allungamento per trazione

Possiamo esplicitare queste dipendenze in modo esplicito nello scrivere la costante elastica k

$$F = k(S, L)\Delta L$$

E scrivere, introducendo lo sforzo (stress) $\sigma = F/S$ e il coefficiente di deformazione o allungamento relativo (strain) $\varepsilon = \Delta L/L$

$$\sigma S = k(S, L)L \varepsilon$$



Ovvero

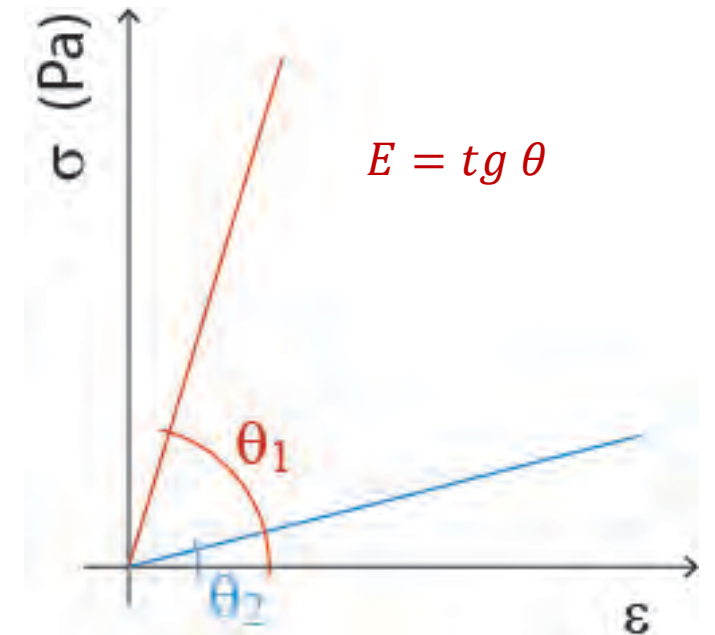
$$\sigma = \frac{k(S, L)L}{S} \varepsilon \quad \Rightarrow \quad \sigma = E_M(T)\varepsilon$$

Dove $E_M(T) = \frac{k(S, L)L}{S}$, misurato in $(N m^{-2})$, ovvero in Pascal, è il **Modulo di Young** del materiale, che non dipende dalle caratteristiche geometriche ma solo dalla temperature.

Rappresentazione grafica legge di Hooke

Il corpo rigido è rappresentato da una retta verticale
(modulo di Young ∞)

Materiale	E (Pa) @ 300 K	E (Pa) @ 800 K	R_s (Pa)
Acciaio	2.1×10^{11}		2.3×10^8
Alluminio	0.7×10^{11}	0.5×10^{11}	3.5×10^8
Ferro	2.0×10^{11}	1.5×10^{11}	
Rame	1.1×10^{11}		
Vetro	0.7×10^{11}		
Calcestruzzo	2.3×10^{10}		
Granito	2.5×10^{10}		



R_s - carico di snervamento

Esempio numerico

Un filo di alluminio con un diametro $d = 2\text{ mm}$ ed una lunghezza $l = 2.0\text{ m}$ si allunga di 1.08 mm quando è sottoposto ad una trazione di 120 N (ovvero ad un carico pari a $120/(\pi \times 10^{-6}) = 4 \times 10^7\text{ Pa}$)

$$\sigma = \frac{F}{S} = \frac{120}{\pi \times 10^{-6}} = 3.82 \times 10^7\text{ Pa} \quad \varepsilon = \frac{\sigma}{E} = \frac{3.82 \times 10^7}{0.7 \times 10^{11}} = 0.54 \times 10^{-3}$$

Ovvero $\Delta L = \varepsilon L = 0.54 \times 10^{-3} \times 2.0 = 1.08 \times 10^{-3}\text{ m}$

Il valore della costante elastica di questo particolare cavo è

$$k = \frac{ES}{L} = \frac{0.7 \times 10^{11} \cdot \pi \times 10^{-6}}{2} \sim 10^5\text{ N/m}$$

Legge di Poisson

Se un corpo è soggetto ad una trazione, oltre ad una elongazione nella direzione di azione della forza, si osserva una riduzione delle dimensioni trasverse (viceversa nel caso di compressione).

Se si considera un cilindro di lunghezza L e raggio r , si osserva, nei limiti di elasticità del materiale che:

$$\frac{\Delta r}{r} = -\nu \frac{\Delta L}{L} \quad \text{dove } \nu \text{ – Coefficiente di Poisson}$$

Per volumi costanti:

$$\Delta[(\pi r^2)L] = 0 \implies 2\pi r \Delta r L = -\pi r^2 \Delta L \implies \frac{\Delta r}{r} = -\frac{1}{2} \frac{\Delta L}{L}$$

Ovvero $\nu = 0.5$. Sperimentalmente $\nu \leq 0.5$, con $\nu = 0.5$ per materiali perfettamente elastici quale il caucciù.

Materiale	ν	Materiale	ν
Vetro	0.25	Rame	0.34
Ferro	0.30	Piombo	0.40
Acciaio	0.27-0.30	Caucciù	0.50

Sforzi da dilatazione termica

Un materiale non vincolato subisce una dilatazione (compressione) termica al crescere (decrescere) della temperatura.

$$\frac{\Delta L}{L} = \alpha \Delta T$$

dove il coefficiente di dilatazione termica α (misurato in K^{-1}) è caratteristico del materiale considerato. Lo sforzo che tale materiale è in grado di esercitare, se soggetto a vincolo, è pari a:

$$\sigma = E\varepsilon = E \frac{\Delta L}{L} = E\alpha\Delta T$$

Per l'acciaio, ciò corrisponde ad uno sforzo pari a $\sigma = 2.1 \times 10^{11} \cdot 1.2 \times 10^{-5} \Delta T$
 $= 2.3 \text{ MPa/K}$

Materiale	α	Materiale	α
Vetro	9×10^{-6}	Rame	17×10^{-6}
Ferro	12×10^{-6}	Alluminio	23×10^{-6}
Acciaio	12×10^{-6}	Piombo	29×10^{-6}

Scorrimento

Se una forza è applicata parallelamente ad una superficie di un corpo, si parla di sollecitazione di taglio.

Il rapporto tra il modulo della forza e l'area della superficie sulla quale la forza agisce, prende il nome di sforzo di taglio.

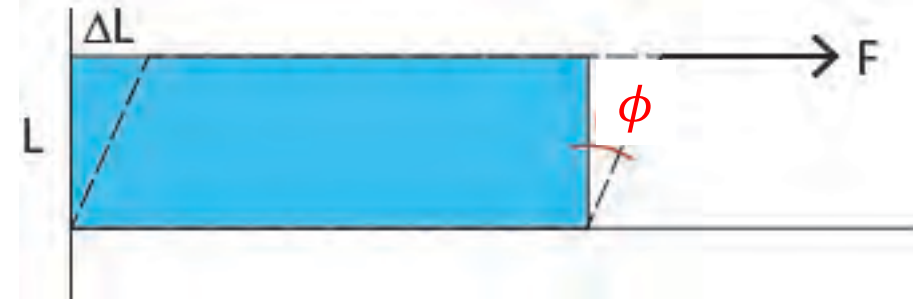
Uno sforzo di taglio tende a provocare una deformazione per scorrimento.

Nei limiti di elasticità del materiale, la legge di Hooke per sforzi di taglio prende la forma di

$$\sigma = G \frac{\Delta L}{L} = G \phi$$

Dove G è il **modulo di rigidità o di taglio**. G è legato al modulo di Young del materiale da

$$G = \frac{E}{2(1 + \nu)}$$



Materiale	G (Pa)	Materiale	G (Pa)
Vetro	$25 \div 32 \times 10^9$	Rame	0.4×10^{11}
Ferro	0.80×10^{11}	Alluminio	0.25×10^{11}
Acciaio	0.83×10^{11}	Piombo	0.34×10^{11}

Esempio numerico

Vogliamo provocare uno scorrimento di $\Delta L = 1 \mu m$, su una barra alta $L = 2 \text{ cm}$ di alluminio, e con una superficie pari a 20 cm^2 e dobbiamo quindi calcolare la forza che dobbiamo esercitare

$$\sigma = G \frac{\Delta L}{L} = 25 \times 10^9 \frac{10^{-6}}{2 \times 10^{-2}} = 12.5 \times 10^5 \text{ Pa}$$

Ovvero $F = \sigma S = 12.5 \times 10^5 \cdot 2 \times 10^{-3} = 2500 \text{ N}$

Torsione

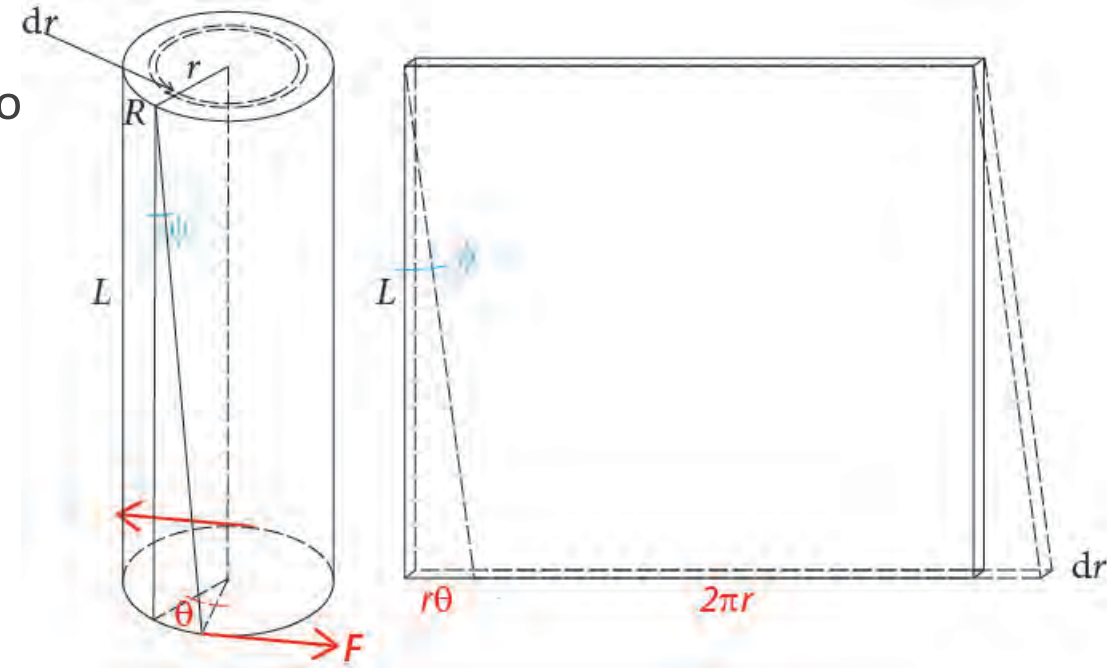
Se si considera un cilindro, e si applica un momento torcente ad una base, mentre l'altra è mantenuta vincolata, si ha una deformazione per torsione. Se il momento è modesto, il cilindro torna nella posizione iniziale una volta che la sollecitazione cessa.

Per effetto della torsione, la base ruota di un angolo θ . Nei limiti di elasticità Hookiana del materiale si osserva proporzionalità tra il momento applicato M e l'angolo di rotazione θ

$$M = k\theta$$

Dove la costante di proporzionalità k è legata al modulo di rigidità del materiale e alle sue caratteristiche Geometriche. **La torsione è apprezzabile se R è piccolo ed L è grande, ossia per oggetti filiformi.**

Per calcolare la relazione tra M e θ nell'ipotesi di filo cilindrico, immaginiamo di sviluppare la corona circolare come un parallelepipedo di altezza L , base $2\pi r$ e spessore dr .



Torsione

Si consideri una corona cilindrica di raggio $r \leq R$ e spessore dr : se si immagina di sviluppare questa corona su di un piano si ottiene un parallelepipedo con dimensioni $2\pi r \times dr \times L$. Per effetto della torsione, una base (quella inferiore nella figura) scorre di una quantità $r\theta$ rispetto all'altra. Lo sforzo applicato è pari a:

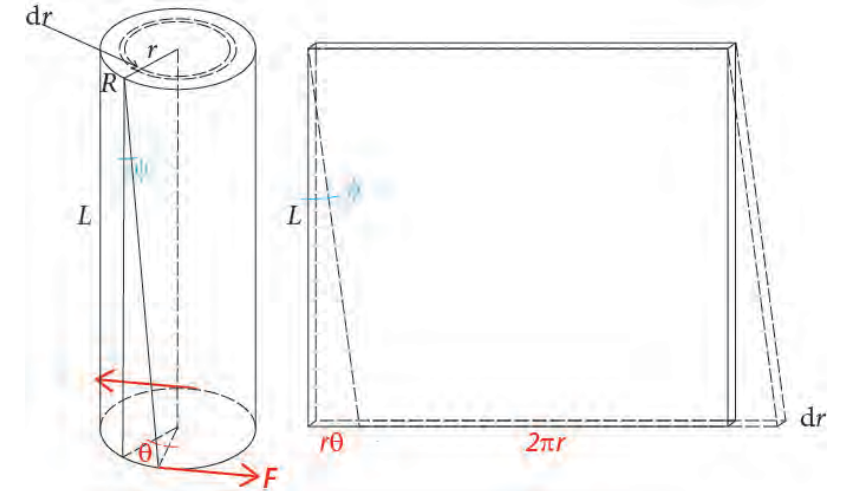
$$\sigma = \frac{dF}{2\pi r dr}$$

Per definizione di **Modulo di Rigidezza di Torsione** G si ha

$$\sigma = G \frac{\Delta L}{L} = G\phi$$

Ma $\Delta L = r\theta$, per cui possiamo scrivere

$$G = \frac{\sigma}{\phi} = \frac{L dF}{2\pi r \Delta L dr} = \frac{L dF}{2\pi \theta r^2 dr} \Rightarrow dF = G \frac{2\pi\theta}{L} r^2 dr$$



Dobbiamo ora integrare su tutto il raggio del filo, per ottenere la forza (o meglio il momento della forza) applicata:

$$dM = rdF = G \frac{2\pi\theta}{L} r^3 dr \Rightarrow M = G \frac{2\pi\theta}{L} \int_0^R r^3 dr = G \frac{2\pi R^4}{4L} \theta = \frac{\pi G}{2} \frac{R^4}{L} \theta$$

Da cui $k = G \frac{\pi R^4}{2L}$. A partire da G viene anche introdotto il **modulo di Coulomb** $\mu = \frac{\pi G}{2}$

$$\Rightarrow M = \mu \frac{R^4}{L} \theta$$

Esempio numerico

Il modulo di Coulomb dell'acciaio è $\mu \sim 10^{11} \text{ Pa}$. Il momento necessario per far ruotare un filo del diametro di $200 \mu\text{m}$ e della lunghezza di 1 m di un angolo $\theta = 1 \text{ mrad}$ (1 mm ad 1 metro) è pari a

$$M = \mu \frac{R^4}{L} \theta = 10^{11} \times \frac{(100 \times 10^{-6})^4}{1} \times 10^{-3} = 10^{-8} \text{ N m}$$

Se all'estremità del filo monto un piattello con braccia da 1 m , la forza da applicare per avere questo momento è pari a:

$$F = 10^{-8} \text{ N}$$

Se immagino di avere 2 masse da 100 g all'estremità di braccia da 1 m appese al filo, e due masse da 10 kg a 10 cm da loro, l'attrazione gravitazionale tra le coppie di masse è:

$$F = G \frac{m_1 m_2}{r^2} = 7 \times 10^{-11} \times \frac{0.1 \times 10}{0.1^2} \sim 10^{-8} \text{ N}$$

Momento torcente:

Applicando il momento M all'estremo libero del filo che ha fatto torcere il filo di θ attorno al proprio asse abbiamo compiuto un lavoro pari a

$$W = \int_0^{\theta} M d\theta' = \frac{1}{2} k \theta^2$$

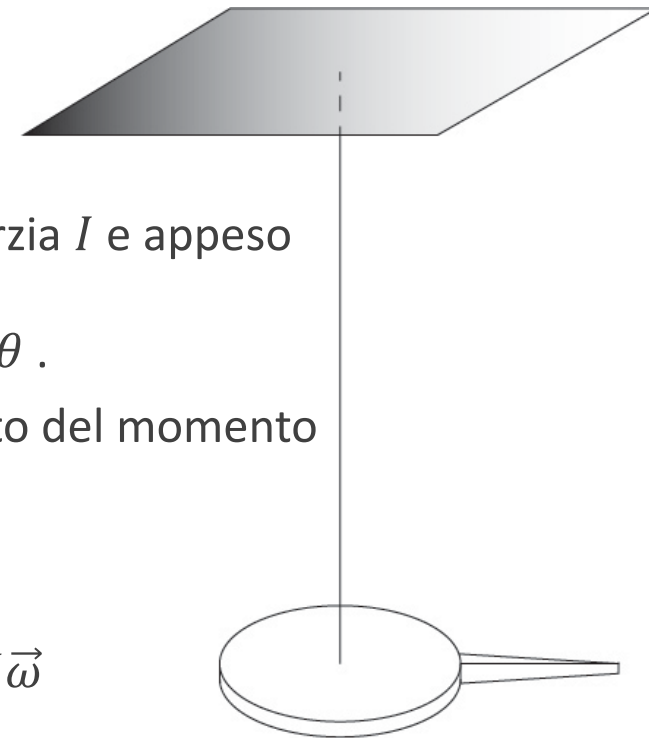
Questo lavoro è immagazzinato nel cilindro come energia potenziale elastica.

Si può realizzare in questo modo il pendolo di torsione:

- Un filo è fissato ad un'estremità al soffitto ed un oggetto di momento di inerzia I è appeso all'estremo libero.
- Per ruotare l'oggetto di un angolo θ occorre applicare un momento $M = k\theta$.
- Quando l'oggetto viene lasciato libero, esso entrerà in oscillazione per effetto del momento meccanico reattivo $-k\theta$ esercitato dal filo.
- Le oscillazioni sono descritte dalla seconda equazione cardinale del moto

$$M = I \frac{d^2\theta}{dt^2} = -k\theta, \quad \text{con} \quad \vec{M} = \frac{d\vec{L}}{dt} \quad \text{ed} \quad \vec{L} = I\vec{\omega}$$

Dove I è il momento d'inerzia rispetto all'asse di rotazione: $I = \int_0^m r_{\perp}^2 dm$



Momento torcente:

Possiamo quindi scrivere

$$\frac{d^2\theta}{dt^2} + \frac{k}{I}\theta = 0$$

Che ha come soluzione (esatta)

$$\theta(t) = \theta_0 \sin\left(\sqrt{\frac{k}{I}}t + \frac{\pi}{2}\right)$$

Il periodo di oscillazione del pendolo è quindi

$$T = 2\pi \sqrt{\frac{I}{k}} = 2\pi \sqrt{\frac{LI}{\mu R^4}}$$

$$T^2 = 4\pi^2 \frac{LI}{\mu R^4}$$



0

Momento d'inerzia

Il momento d'inerzia rispetto al suo asse di un disco di massa M e raggio R e altezza h è

$$I = \int_0^M r^2 dm = 2\pi\rho h \int_0^R r^3 dr = \frac{\pi\rho h}{2} R^4 = \frac{M}{2} R^2$$

Mentre per la corona circolare ottengo

$$I = \int_0^M r^2 dm = 2\pi\rho h \int_{R_i}^{R_e} r^3 dr = \frac{2\pi\rho h}{4} (R_f^4 - R_i^4) = \frac{\rho h \pi (R_f^2 - R_i^2)}{2} (R_f^2 + R_i^2) = \frac{M}{2} (R_f^2 + R_i^2)$$

Ad esempio per 2 corone circolari di 200g, entrambi con un diametro interno di 2 cm ed uno esterno rispettivamente di 4 e 6 cm ho:

$$I_4 = 50 \times 10^{-6} \text{ (kg m}^2\text{)} \text{ e } I_6 = 100 \times 10^{-6} \text{ (kg m}^2\text{)}$$

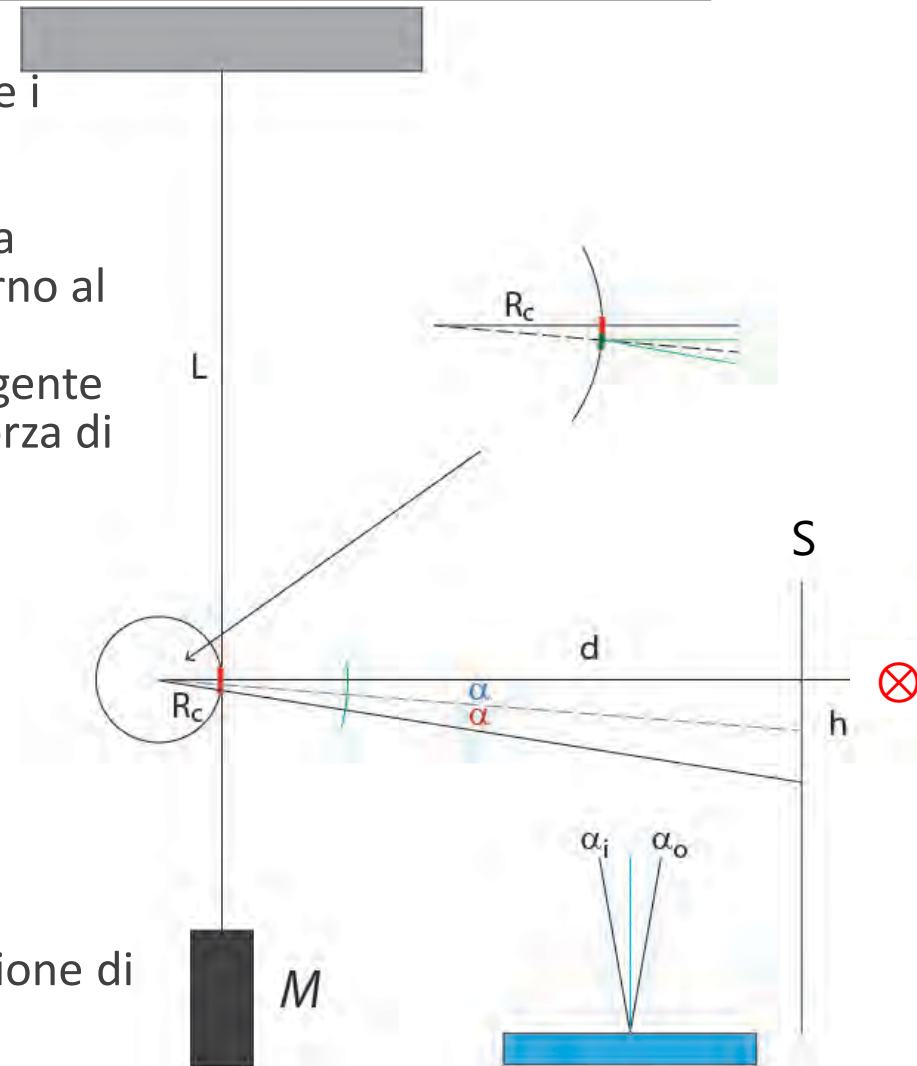
MA NON CONOSCIAMO IL MOMENTO D'INERZIA COMPLETO DEL SISTEMA

Modulo di Young – Setup

Per la misura del modulo di Young dell'acciaio usiamo il **principio della leva ottica** (figura a lato) che permette di amplificare i piccoli effetti dell'allungamento del filo.

Per creare la leva ottica usiamo una sorgente luminosa \otimes posta alla **stessa altezza** dello specchio sul cilindro (che può solo ruotare attorno al suo asse), ad una distanza d dallo specchio. Il filo è vincolato all'estremo superiore, ed avvolto sul cilindro (tangente lungo la verticale) ad una distanza L dal vincolo. Applicando una forza di trazione Mg provochiamo l'allungamento del filo.

- D_f diametro del filo
- R_c raggio del cilindro
- L lunghezza del filo fino al cilindro
- M massa appesa al filo
- S scala graduata
- d distanza specchio/scala graduata
- h spostamento dello spot luminoso successivo all'applicazione di Mg
- α angolo di rotazione del cilindro



Modulo di Young – Setup

Ricordando che $\sigma = E\varepsilon$. Scriviamo

$$\frac{F}{S_f} = E \frac{\Delta L}{L}$$

Ora, per il set-up che abbiamo realizzato abbiamo

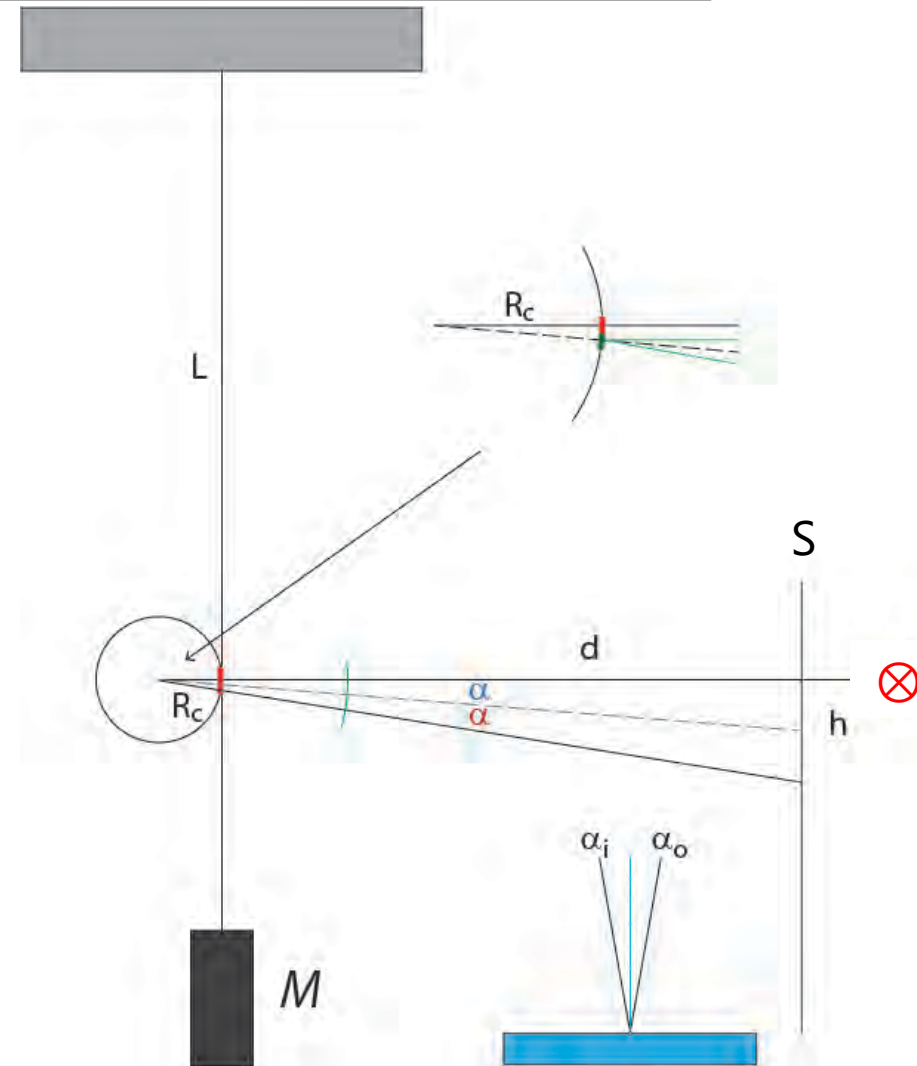
- $\Delta L = R_c \alpha = R_c \frac{1}{2} \tan^{-1} \frac{h}{d}$
- $F = Mg$
- $S_f = \pi(D_f/2)^2$

Che ci permette di ricavare il modulo di Young E in funzione di quantità misurabili

$$E = \frac{8gL}{\pi R_c D_f^2} \cdot \frac{M}{\tan^{-1} \frac{h}{d}}$$

Per piccoli angoli $\tan \alpha \approx \alpha$, possiamo scrivere $\tan^{-1} \frac{h}{d} \approx \frac{h}{d}$

$$E = \frac{8gLd}{\pi R_c D_f^2} \cdot \frac{M}{h}$$



Modulo di Young – Errori

Per valutare E eseguiamo misure con diverse masse M_i che risultano in diverse altezze h_i

$$E_i = \frac{8gLd}{\pi R_c D_f^2} \cdot \frac{M_i}{h_i}$$

Abbiamo solo errori di sensibilità di cui una parte

$$A = \frac{8gLd}{\pi R_c D_f^2}$$

Non varia da misura a misura mentre una parte

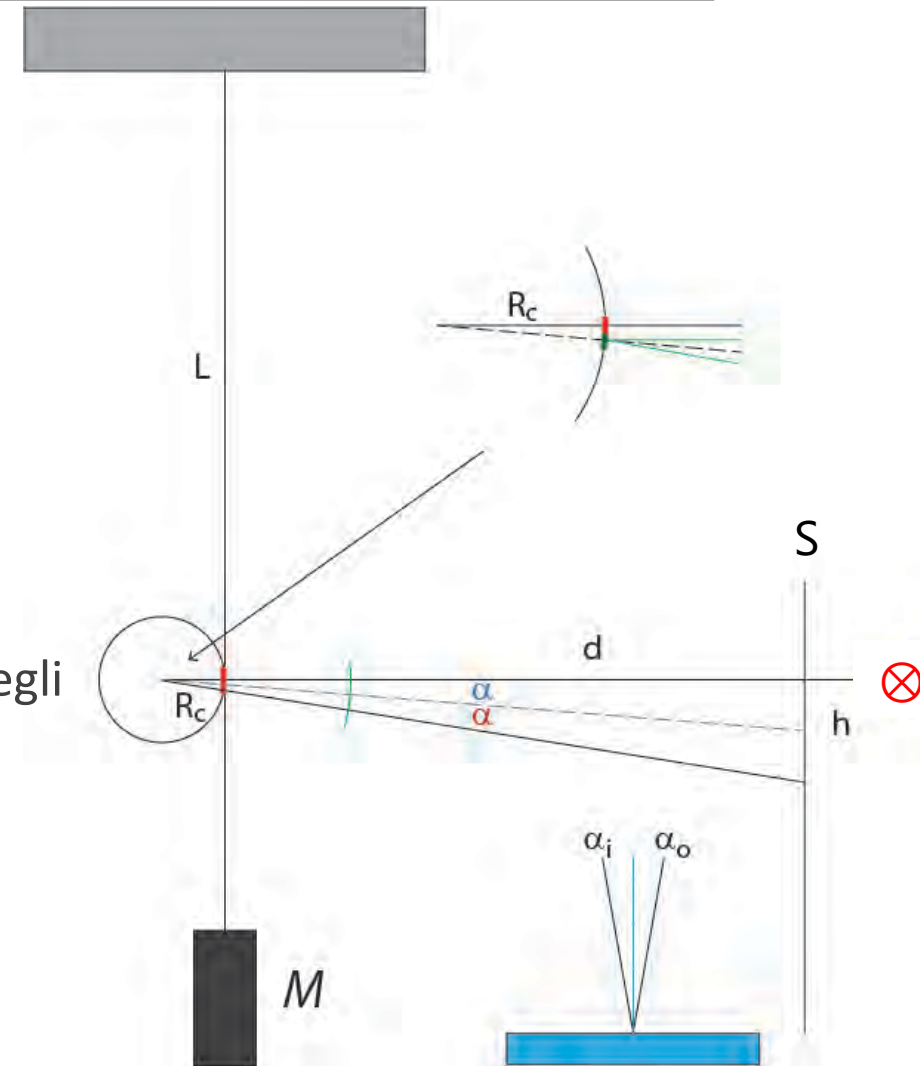
$$B_i = \frac{M_i}{h_i}$$

Cambia da misura a misura. Possiamo usare la propagazione degli errori massimi relativi

$$\frac{\Delta A}{A} = \frac{\Delta g}{g} + \frac{\Delta L}{L} + \frac{\Delta d}{d} + \frac{\Delta R_c}{R_c} + 2 \frac{\Delta D_f}{D_f}$$

e

$$\frac{\Delta B_i}{B_i} = \frac{\Delta M_i}{M_i} + \frac{\Delta h_i}{h_i}$$



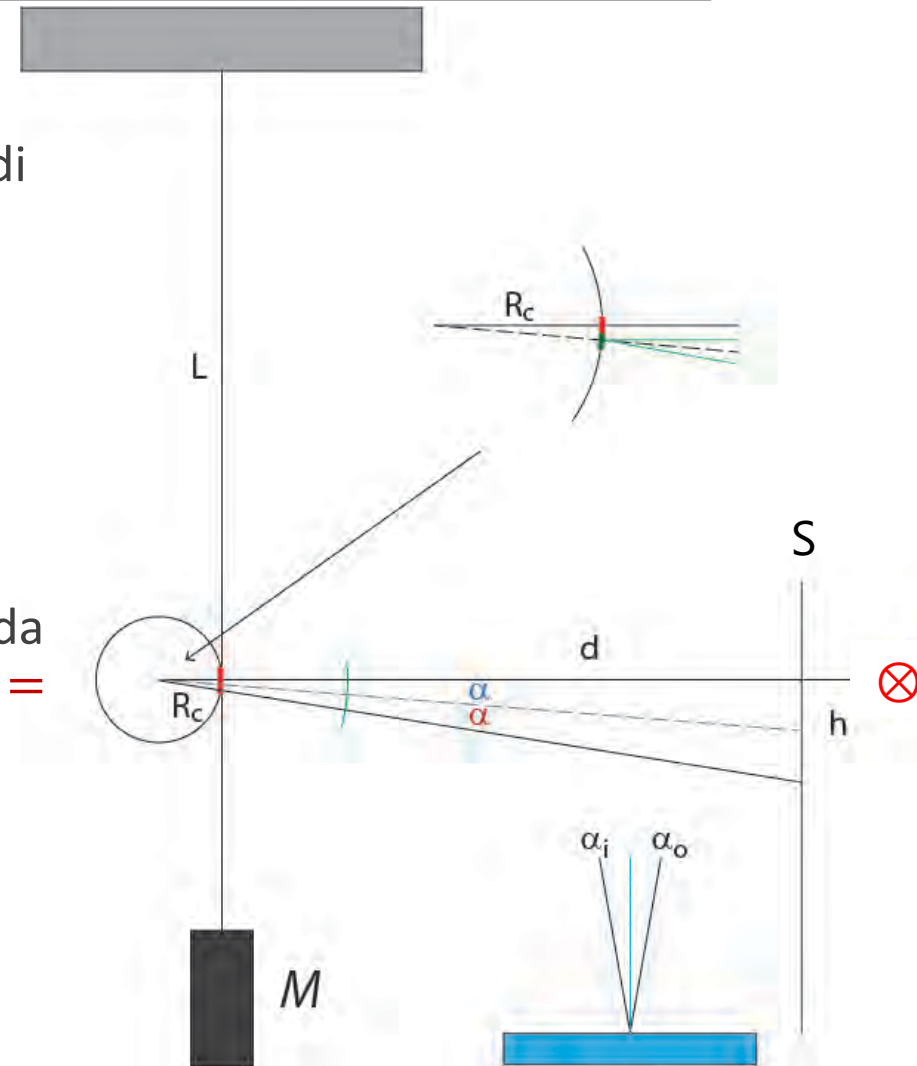
Modulo di Young – Errori

Le misure con masse diverse possono essere usate per una miglior stima di B ovvero (in un modo non completamente giustificato); ricavo la deviazione standard sulle B_i nell'ipotesi di errori di risoluzione $\sigma_{B_i} = \Delta B_i / \sqrt{3}$

$$B = \frac{\sum_{i=1}^n \frac{B_i}{\sigma_{B_i}^2}}{\sum_{i=1}^n \frac{1}{\sigma_{B_i}^2}} \quad e \quad \sigma_B^2 = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_{B_i}^2}}$$

Visto che la misura è dominata dagli errori sulla parte A passo da deviazioni standard ad errori massimi sulla parte B usando $\Delta B = 3 \sigma_B$. Quindi alla fine il modulo di Young è dato da $E = A \cdot B$ e l'errore sul modulo di Young è dato da

$$\Delta E = E \left(\frac{\Delta A}{A} + \frac{\Delta B}{B} \right)$$



Modulo di Young – Grafici

Grafici: ΔL_i vs M_i
 B_i vs M_i

Strumenti:	L, d ed h	Metro lineare	$\Delta x_i = 0.5 \times 10^{-3} \text{ m}$
	R_c	Calibro ventesimale/Palmer	$\Delta R_c = 2.5 \times 10^{-5} \text{ m}$
	D_f	Calibro Palmer	$\Delta D_f = 0.5 \times 10^{-5} \text{ m}$
	M_i	Bilancia	$\Delta M_i = 0.5 \times 10^{-2} \text{ g}$

Limite elastico acciaio: $2 \div 4 \times 10^8 \text{ Pa}$

La posizione iniziale deve essere alla stessa altezza dello specchio e della luce. Ruotare il cilindro per ottenerla

Il filo deve essere teso per non scorrere \Rightarrow pre caricare con 50-100 g il piattello

Verificare che al rimuovere della massa M_i ritroviamo la posizione iniziale

Modulo di Coulomb – Set-up

Per la misura del modulo di Coulomb dell'acciaio usiamo il pendolo di torsione, come illustrato in figura.

Non conosciamo il momento d'inerzia I_0 del sistema del piattello, mentre sappiamo calcolare quello della corona circolare che può essere montata per oscillare sullo stesso asse del piattello (asse del filo).

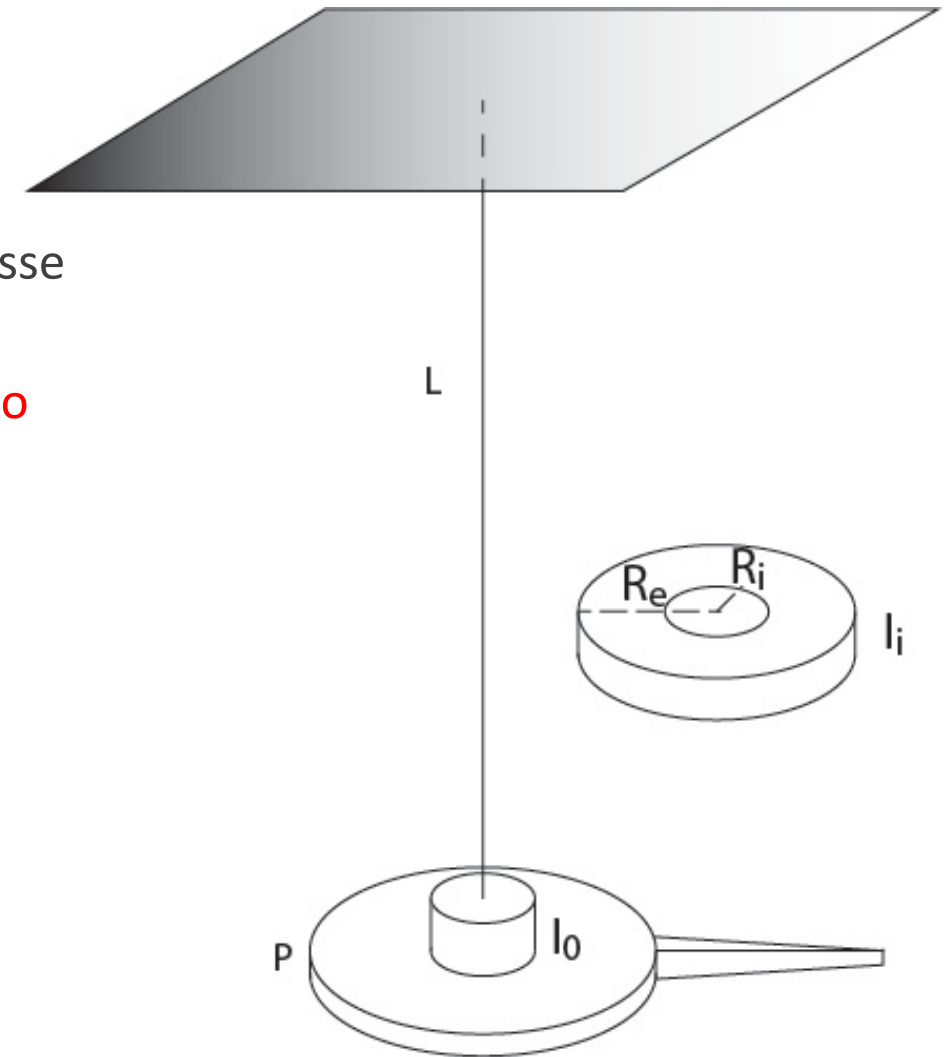
Possiamo però misurare il periodo di oscillazione del piattello scarico e del piattello caricato con la corona circolare, che varranno

$$T_{0,i}^2 = 4\pi^2 \frac{LI_0}{\mu R^4} \quad e \quad T_i^2 = 4\pi^2 \frac{L(I_0 + I_i)}{\mu R^4}$$

Vale quindi

$$T_i^2 - T_{0,i}^2 = 4\pi^2 \frac{LI_i}{\mu R^4}$$

Che permette di calcolare il modulo di Coulomb a partire da quantità misurabili



Modulo di Coulomb – Set-up

Il modulo di Coulomb vale quindi

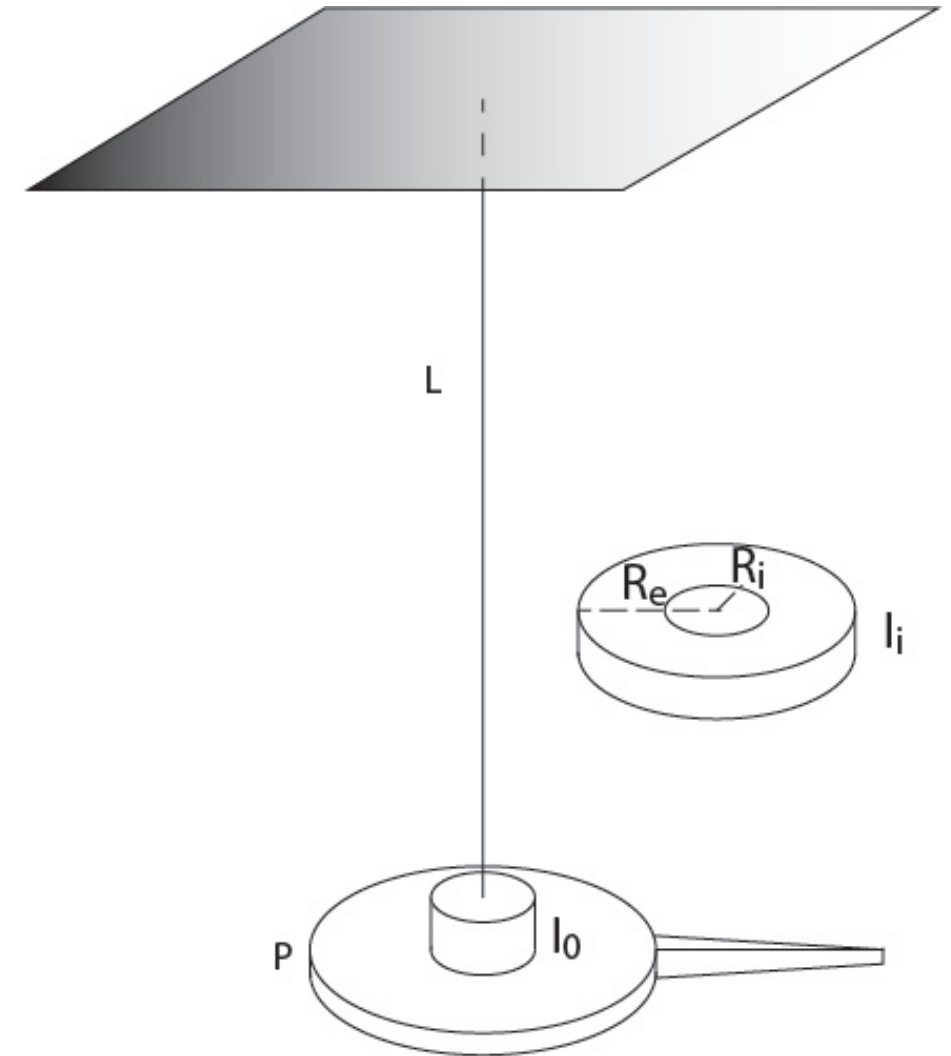
$$\mu = 4\pi^2 \frac{L}{R^4} \cdot \frac{I_i}{T_i^2 - T_{0,i}^2}$$

Scrivendo esplicitamente il momento d'inerzia abbiamo

$$\mu = 2\pi^2 \frac{L}{(D_f/2)^4} \cdot \frac{((D_e/2)^2 + (D_i/2)^2)_i \cdot M_i}{T_i^2 - T_{0,i}^2}$$

Con:

- P piattello di momento d'inerzia incognito I_0
- L lunghezza del filo
- D_f diametro del filo d'acciaio
- $T_{0,i}$ e T_i periodi d'oscillazione
- μ modulo di Coulomb in Pa o $(N m^{-2})$



Misura di T

Come nel caso del pendolo

Misurare $\tau = 4 \div 6T = t_{stop} - t_{start}$ per ogni lunghezza corona circolare e ripetere la misura allo stesso modo a piatto scarico.

- Misurare $\tau = 4 \div 6T = t_{stop} - t_{start}$ per almeno 10 volte per ogni corona circolare.
- Misurare $\tau_0 = 4 \div 6T = t_{stop} - t_{start}$ (piatti scarichi) per almeno 10 volte per ogni corona circolare in modo da avere misure indipendenti.
- Per la trattazione statistica che abbiamo visto i risultati seguono una distribuzione Gaussiana.
 - La miglior stima del valore vero di τ è il valor medio $\bar{\tau} = \sum_{i=1}^{10} \tau_i$
 - La miglior stima della varianza del campione è data da $\sigma_{\tau_i}^2 = \frac{1}{10-1} \sum_{i=1}^{10} (\tau_i - \bar{\tau})^2$
 - La miglior stima della varianza della media è data da $\sigma_{\bar{\tau}}^2 = \frac{\sigma_{\tau_i}^2}{10}$
 - Abbiamo quindi $\bar{\tau} \pm \sigma_{\bar{\tau}}$.
 - Tenuto conto che $\bar{T} = \bar{\tau}/5$ e che $\sigma_{\bar{T}} = \sigma_{\bar{\tau}}/5$ abbiamo anche misurato
$$\bar{T} \pm \sigma_{\bar{T}}$$

Modulo di Coulomb – Errori

Data la misura del modulo di Coulomb

$$\mu = 2\pi^2 \frac{L}{(D_f/2)^4} \cdot \frac{((D_e/2)^2 + (D_i/2)^2)_i \cdot M_i}{T_i^2 - T_0^2}$$

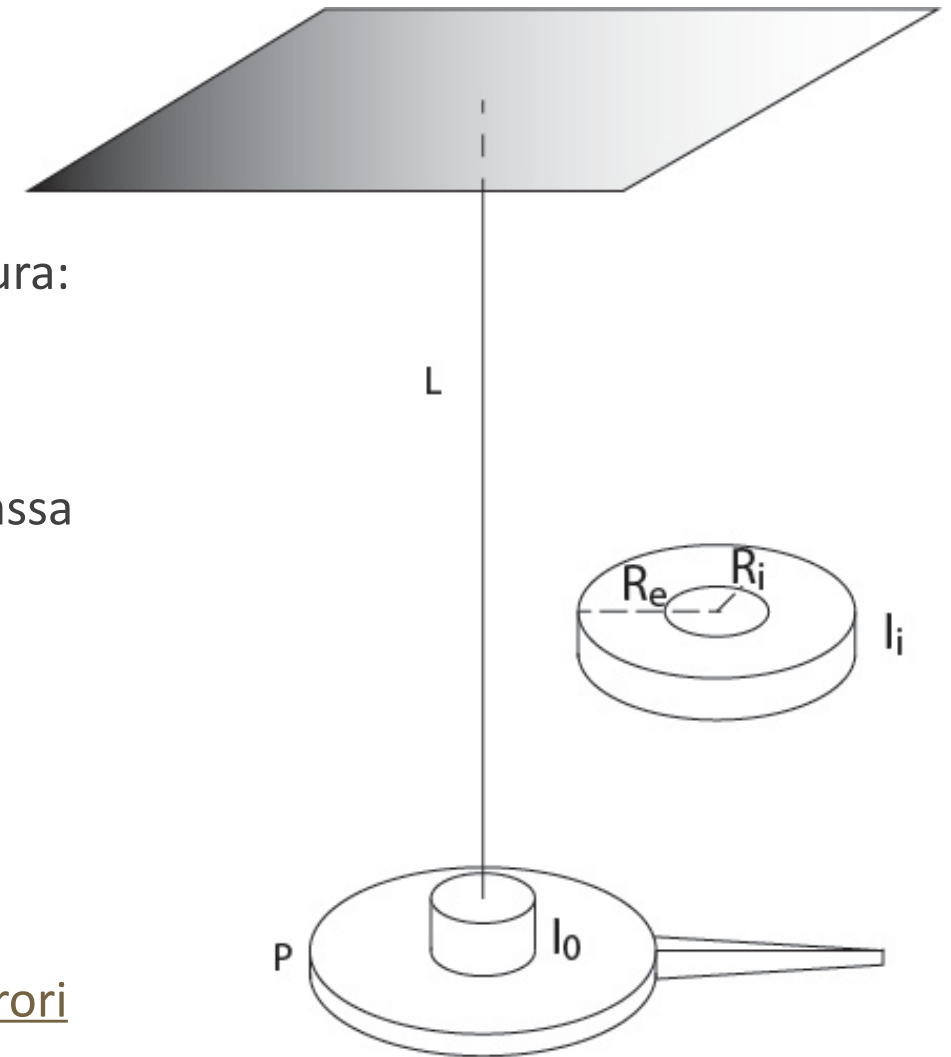
Abbiamo due contributi che non cambiano da misura a misura:

$$A = \frac{L}{(D_f/2)^4}, \text{ con errori } \frac{\Delta A}{A} = \frac{\Delta L}{L} + 4 \frac{\Delta D_f}{D_f}$$

E contributi che cambiano per diverse corone circolari di massa M_i (da notare che se le corone circolari hanno gli stessi diametri, questa parte va aggiunta in A e non in B)

$$B = \frac{((D_e/2)^2 + (D_i/2)^2)_i \cdot M_i}{T_i^2 - T_{0i}^2}, \text{ con errori } \frac{\Delta B}{B} = \frac{\Delta M_i}{M_i} + \frac{\Delta((D_e/2)^2 + (D_i/2)^2)_i}{((D_e/2)^2 + (D_i/2)^2)_i} + \frac{\Delta(T_i^2 - T_{0i}^2)}{T_i^2 - T_{0i}^2}$$

Propagati come errori massimi (nell'ipotesi che siano gli errori dominanti)



Modulo di Coulomb – Errori

Dove

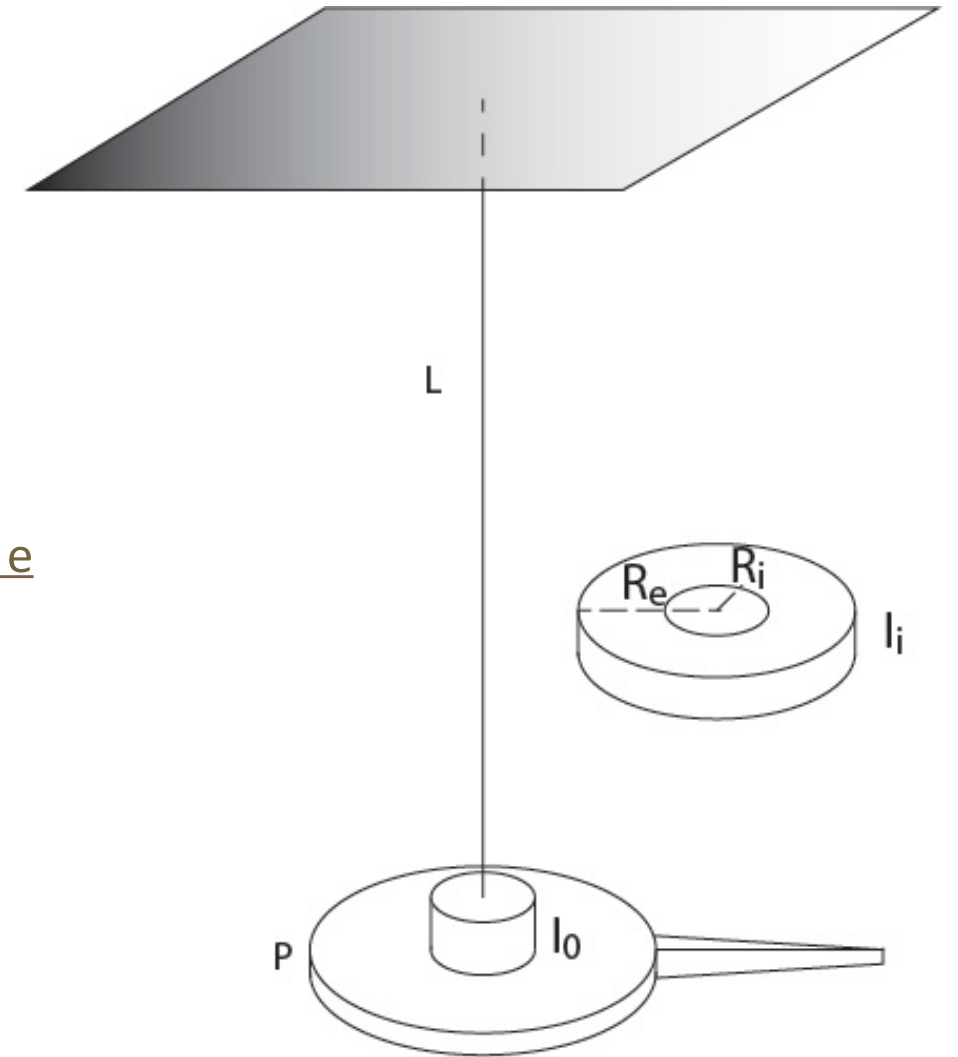
$$\Delta\left(\left(D_e/2\right)^2 + \left(D_i/2\right)^2\right)_i = \frac{1}{2}\left(D_e\Delta D_e + D_i\Delta D_i\right)_i$$

e dove

$$\Delta\left(T_i^2 - T_{0i}^2\right) = 2T_i\Delta T_i + 2T_{0i}\Delta T_{0i}$$

e con l'errore massimo sul periodo ottenuto da $\Delta T_i = 3\sigma_{T_i}$

Nel caso i termini dominanti nell'errore siano statistici trasformare tutti gli errori massimi in deviazioni standard e usare la legge di propagazione della varianza



Modulo di Coulomb – Errori

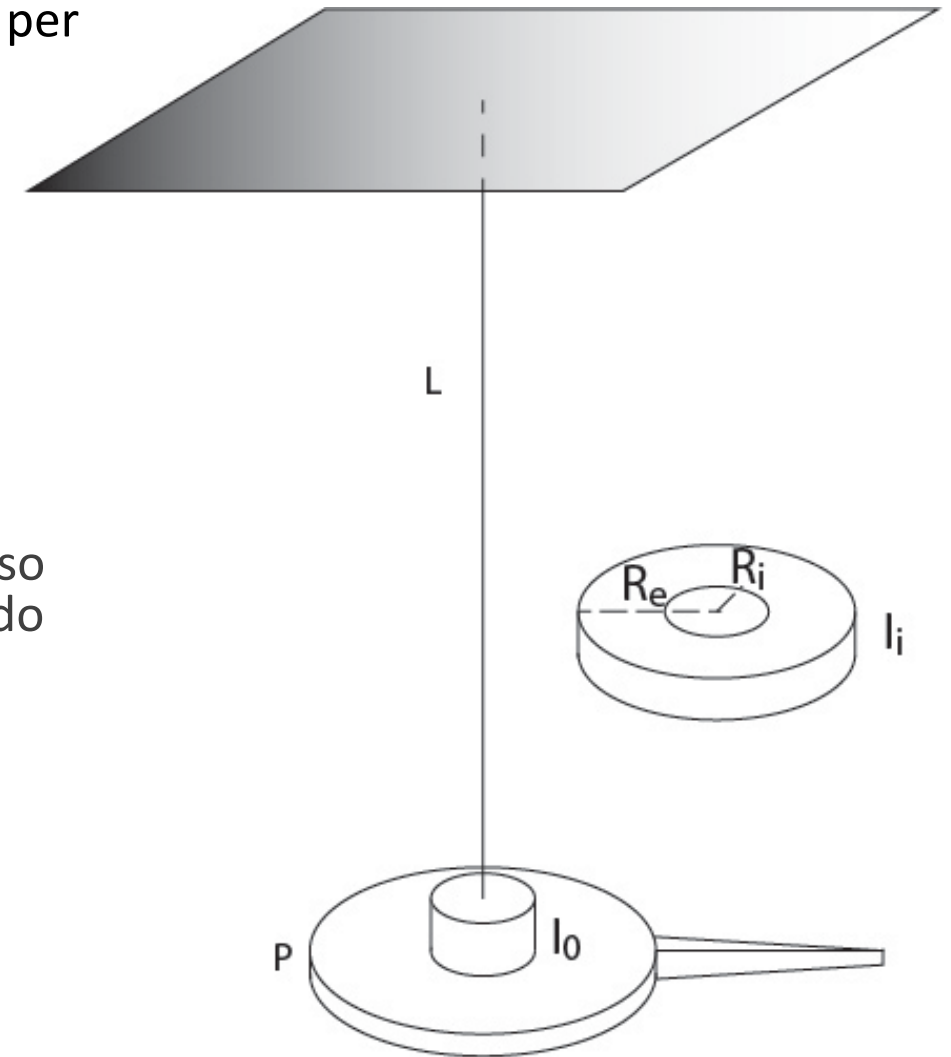
Come per il modulo di Young, posso usare le diverse misure per ottenere la miglior stima di B ovvero; ricavo la deviazione standard sulle B_i nell'ipotesi di errori di risoluzione $\sigma_{B_i} = \Delta B_i / \sqrt{3}$

$$B = \frac{\sum_{i=1}^n \frac{B_i}{\sigma_{B_i}^2}}{\sum_{i=1}^n \frac{1}{\sigma_{B_i}^2}} \quad e \quad \sigma_B^2 = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_{B_i}^2}}$$

Visto che la misura è dominata dagli errori sulla parte A passo da deviazioni standard ad errori massimi sulla parte B usando $\Delta B = 3 \sigma_B$.

Quindi alla fine il modulo di Coulomb è dato da $\mu = A \cdot B$ e l'errore sul modulo di Coulomb è dato da

$$\Delta \mu = \mu \left(\frac{\Delta A}{A} + \frac{\Delta B}{B} \right)$$



Modulo di Coulomb – Grafici

Grafici: $T_i^2 - T_{0i}^2$ vs I_i
 B_i vs I_i

Strumenti:	L	Metro lineare	$\Delta x_i = 0.5 \times 10^{-3} \text{ m}$
	R_e, R_i	Calibro ventesimale	$\Delta R_i = 2.5 \times 10^{-5} \text{ m}$
	D_f	Calibro Palmer	$\Delta D_f = 0.5 \times 10^{-5} \text{ m}$
	M_i	Bilancia	$\Delta M_i = 0.5 \times 10^{-3} \text{ g}$

Materiale	$G \text{ (Pa)}$	$\mu \text{ (Pa)}$	$E \text{ (Pa)}$
Acciaio	0.85×10^{11}	1.3×10^{11}	2.1×10^{11}

Inciso: deformazioni elastiche omogenee

- Deformazione assiale (compressione o trazione):

$$\varepsilon_l = \frac{E}{\sigma}$$

dove E è il modulo di elasticità o modulo di Young.

- Deformazione laterale:

$$\varepsilon_d = -\nu\varepsilon_l$$

La deformazione laterale è proporzionale a quella assiale, e ν è il coefficiente di Poisson.

- Deformazione di volume:

$$\varepsilon_V = \frac{\sigma}{K}$$

dove K è detto modulo di comprimibilità.

- Deformazione di scorrimento o taglio

$$\varepsilon_t = \frac{\tau}{G}$$

dove G è il modulo di rigidità

Sforzo

Per sforzo si intende la forza trasmessa per unità di superficie, nell'intorno di un punto, che si crea a seguito dell'applicazione di sollecitazioni esterne su un sistema, allo scopo di mantenere l'equilibrio; forza non necessariamente perpendicolare alla superficie. Possiamo più facilmente rappresentare lo sforzo come:

- sforzo normale o pressione:

$$\sigma = \vec{F} \cdot \hat{n} = F \cos \theta$$

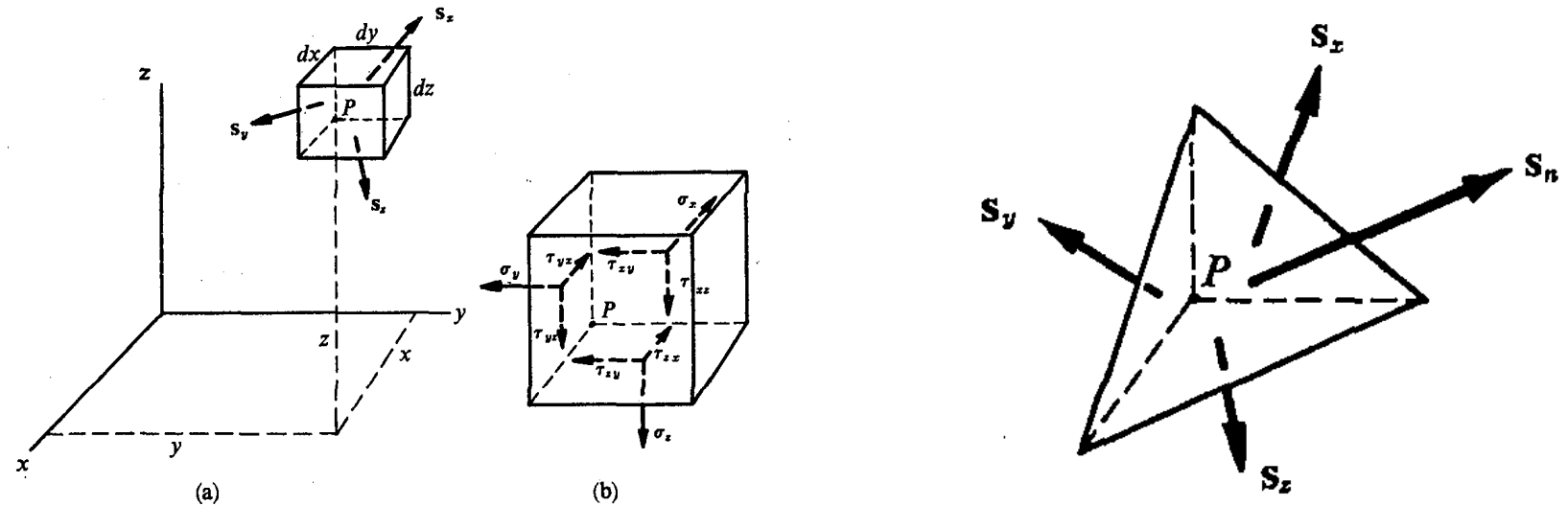
- sforzo tangenziale o di taglio:

$$\tau = \vec{F} \cdot \hat{t} = F \sin \theta$$

dove \hat{n} e \hat{t} rappresentano i versori rispettivamente normale e tangente alla superficie cui è applicata la forza.

Tensore degli sforzi

Nell'analisi degli sforzi è prassi descrivere lo stato di un elemento di volume infinitesimo, che nel caso Cartesiano possiamo identificare con (dx, dy, dz) . Lo stress al punto di coordinate $P = (x, y, z)$ è definito dalle componenti (s_x, s_y, s_z)



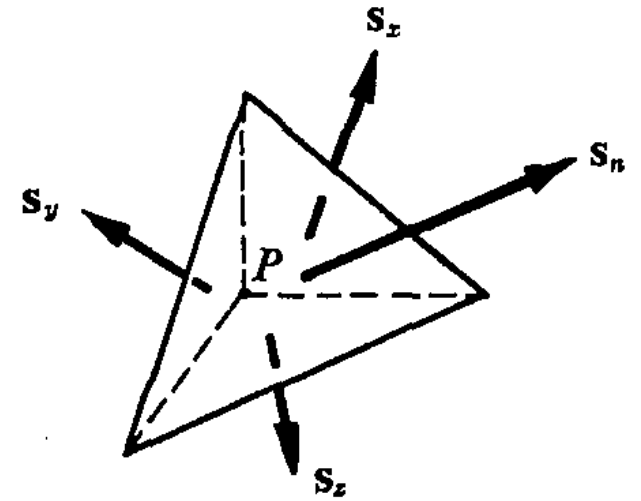
Possiamo quindi scrivere le componenti dello stress sui tre piani; ad esempio

$$s_x = \sigma_x \hat{x} + \tau_{xy} \hat{y} + \tau_{xz} \hat{z}$$

$$s_y = \tau_{yx} \hat{x} + \sigma_y \hat{y} + \tau_{yz} \hat{z}$$

$$s_z = \tau_{zx} \hat{x} + \tau_{zy} \hat{y} + \sigma_z \hat{z}$$

Relazione di Cauchy



Vediamo che relazione intercorre tra queste superfici e una superficie infinitesima generica $d\pi_n$ orientata con vettore uscente \hat{n} .

Consideriamo gli sforzi agenti sulle superfici con ovvi riferimenti degli indici ai vettori: \vec{s}_x , \vec{s}_y , \vec{s}_z , \vec{s}_n ; per la condizione di equilibrio:

$$\vec{s}_x d\pi_x + \vec{s}_y d\pi_y + \vec{s}_z d\pi_z + \vec{s}_n d\pi_n = 0$$

dalla quale otteniamo la relazione di Cauchy:

$$\vec{s}_n = -(\vec{s}_x \cos \alpha + \vec{s}_y \cos \beta + \vec{s}_z \cos \gamma)$$

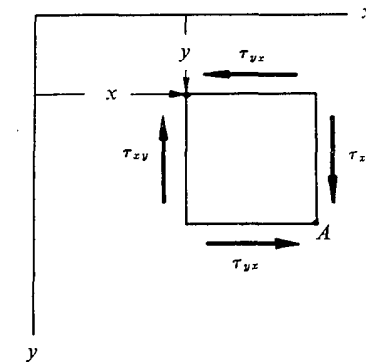
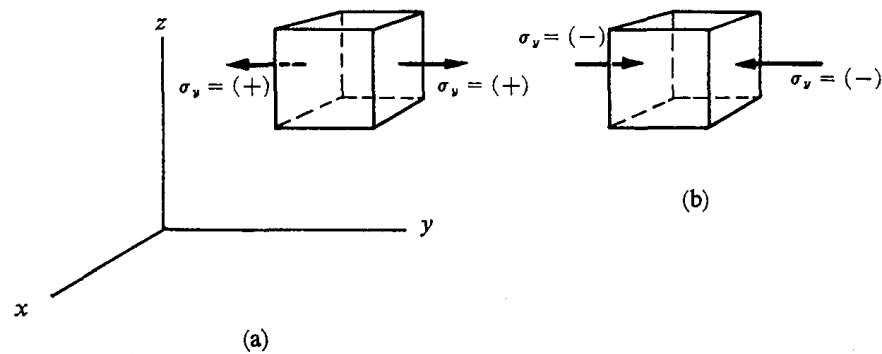
dove si è diviso tutto per dS e sapendo che $\frac{d\pi_x}{d\pi_n} = \cos \alpha, \dots$

Tensore degli sforzi

Otteniamo in questo modo una matrice detta **tensore degli sforzi** sulla generica superficie infinitesima di versore \hat{n} :

$$\Sigma = \begin{pmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{pmatrix}$$

Per il segno usiamo la convenzione di **sforzi positivi in caso di trazioni e come negativi in caso di pressioni**, o equivalentemente stress positivi se uscenti nella direzione positive da facce nella direzione positive e nella direzione negative da facce in direzione negativa . La stessa definizione vale per gli sforzi di taglio.



La condizione di equilibrio del corpo richiede che le forze ed i momenti si annullino tra loro. Per le forze avviene, per i momenti questo richiede:

$$\tau_{xy}(dydz)dx = \tau_{yx}(dxdz)dy \implies \tau_{xy} = \tau_{yx}$$

Deformazioni

- Consideriamo un punto di un sistema continuo isotropo $P(x, y, z)$ non deformato e un altro punto $Q(x + dx, y + dy, z + dz)$ distante da di un tratto sufficientemente piccolo $d\vec{r} = (dx, dy, dz)$. A seguito di una deformazione il punto P si porterà in P' percorrendo $\vec{q} = (q_x, q_y, q_z)$ e Q si porterà in Q' , percorrendo un tratto che differirà da \vec{q} di un tratto sufficientemente piccolo $d\vec{q}$. In pratica il vettore $\overrightarrow{PQ} = (dx, dy, dz)$ si trasformerà in $\overrightarrow{P'Q'} = (dq_x, dq_y, dq_z)$ secondo le relazioni:

$$dq_x = \frac{\partial q_x}{\partial x} dx + \frac{\partial q_x}{\partial y} dy + \frac{\partial q_x}{\partial z} dz$$

$$dq_y = \frac{\partial q_y}{\partial x} dx + \frac{\partial q_y}{\partial y} dy + \frac{\partial q_y}{\partial z} dz$$

$$dq_z = \frac{\partial q_z}{\partial x} dx + \frac{\partial q_z}{\partial y} dy + \frac{\partial q_z}{\partial z} dz$$

Deformazioni

- Possiamo esprimere queste relazioni in forma matriciale con una matrice a nove derivate parziali detta **tensore di deformazione** $\boldsymbol{\varepsilon}$;

$$\begin{pmatrix} dq_x \\ dq_y \\ dq_z \end{pmatrix} = \begin{pmatrix} \frac{\partial q_x}{\partial x} & \frac{\partial q_x}{\partial y} & \frac{\partial q_x}{\partial z} \\ \frac{\partial q_y}{\partial x} & \frac{\partial q_y}{\partial y} & \frac{\partial q_y}{\partial z} \\ \frac{\partial q_z}{\partial x} & \frac{\partial q_z}{\partial y} & \frac{\partial q_z}{\partial z} \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}$$

Ovvero:

$$d\vec{q} = \boldsymbol{\varepsilon} d\vec{r}$$

Deformazioni

- Ogni tensore può essere decomposto in un tensore simmetrico più un tensore antisimmetrico $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_s + \boldsymbol{\varepsilon}_a$

$$\boldsymbol{\varepsilon}_s = \begin{pmatrix} \frac{\partial q_x}{\partial x} & \frac{1}{2} \left[\frac{\partial q_x}{\partial y} + \frac{\partial q_y}{\partial x} \right] & \frac{1}{2} \left[\frac{\partial q_x}{\partial z} + \frac{\partial q_z}{\partial x} \right] \\ \frac{1}{2} \left[\frac{\partial q_y}{\partial x} + \frac{\partial q_x}{\partial y} \right] & \frac{\partial q_y}{\partial y} & \frac{1}{2} \left[\frac{\partial q_y}{\partial z} + \frac{\partial q_z}{\partial y} \right] \\ \frac{1}{2} \left[\frac{\partial q_z}{\partial x} + \frac{\partial q_x}{\partial z} \right] & \frac{1}{2} \left[\frac{\partial q_z}{\partial y} + \frac{\partial q_y}{\partial z} \right] & \frac{\partial q_z}{\partial z} \end{pmatrix}$$

$$\boldsymbol{\varepsilon}_a = \begin{pmatrix} 0 & \frac{1}{2} \left[\frac{\partial q_x}{\partial y} - \frac{\partial q_y}{\partial x} \right] & \frac{1}{2} \left[\frac{\partial q_x}{\partial z} - \frac{\partial q_z}{\partial x} \right] \\ \frac{1}{2} \left[\frac{\partial q_y}{\partial x} - \frac{\partial q_x}{\partial y} \right] & 0 & \frac{1}{2} \left[\frac{\partial q_y}{\partial z} - \frac{\partial q_z}{\partial y} \right] \\ \frac{1}{2} \left[\frac{\partial q_z}{\partial x} - \frac{\partial q_x}{\partial z} \right] & \frac{1}{2} \left[\frac{\partial q_z}{\partial y} - \frac{\partial q_y}{\partial z} \right] & 0 \end{pmatrix}$$

Deformazioni

Il *tensore simmetrico* descrive le seguenti deformazioni:

- **espansioni o allungamenti relativi, rappresentati dagli elementi sulla diagonale principale:**

$$\varepsilon_x = \frac{\partial q_x}{\partial x} \quad \varepsilon_y = \frac{\partial q_y}{\partial y} \quad \varepsilon_z = \frac{\partial q_z}{\partial z}$$

- **scorrimenti o distorsioni, rappresentati dagli elementi fuori dalla diagonale principale:**

$$\varepsilon_{xy} = \varepsilon_{yx} = \frac{1}{2} \left[\frac{\partial q_x}{\partial y} + \frac{\partial q_y}{\partial x} \right]$$

$$\varepsilon_{xz} = \varepsilon_{zx} = \frac{1}{2} \left[\frac{\partial q_x}{\partial z} + \frac{\partial q_z}{\partial x} \right]$$

$$\varepsilon_{yz} = \varepsilon_{zy} = \frac{1}{2} \left[\frac{\partial q_y}{\partial z} + \frac{\partial q_z}{\partial y} \right]$$

- **La traccia del tensore di deformazione, pari alla divergenza dello spostamento, rappresenta un'invariante e viene chiamato coefficiente di dilatazione cubica:**

$$\text{Tr } \varepsilon = \varepsilon_x + \varepsilon_y + \varepsilon_z = \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z} = \nabla \cdot \vec{q}$$

dove ∇ rappresenta l'operatore differenziale vettoriale nabra espresso qui in coordinate cartesiane.

Deformazioni

Il tensore antisimmetrico descrive invece le rotazioni rigide nell'intorno del punto P , che non rappresenta una deformazione. Per esempio in coordinate cartesiane tridimensionali:

$$\varepsilon_{xy} = \frac{1}{2} \left[\frac{\partial q_x}{\partial y} - \frac{\partial q_y}{\partial x} \right]$$

rappresenta la rotazione $d\theta_z$ attorno all'asse z , nell'intorno del punto P .

Relazioni tra sforzi e deformazioni

Quando un materiale viene sottoposto a stress si producono delle deformazioni. Tali deformazioni dipendono dal materiale e sono diverse a seconda del materiale (rame, ferro, legno, cemento, marmo, vetro, gelatina, plastica, argilla, acciaio), storico degli sforzi a cui è stato sottoposto etc. Rimanendo su sforzi elastici e materiali d'interesse del corso introduciamo alcune proprietà comuni:

Un materiale è detto:

- **Isotropico**: se le sue proprietà di deformazione non dipendono dalla direzione di applicazione
- **Omogeneo**: se le sue proprietà di deformazione sono le stesse in qualsiasi suo punto, e, come abbiamo già visto
- **Elastico**: se ritorna alle condizioni iniziali una volta che la sollecitazione esterna viene rimossa.

Legge di Hooke generalizzata

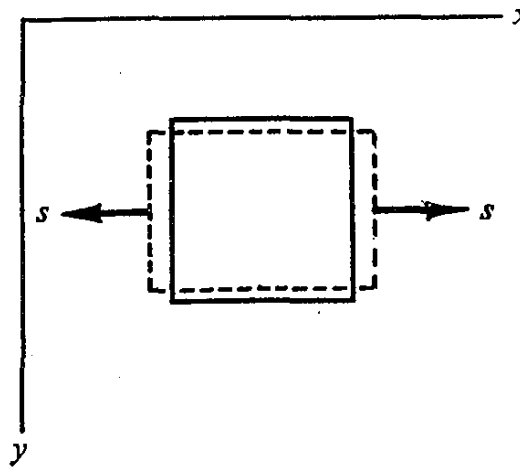
Legge di **Hooke** generalizzata: la relazione lineare tra sforzo e deformazione è chiamata legge di Hooke; nella sua forma più generale ciò equivale a dire che ogni sforzo (ad esempio σ_x) è una funzione lineare delle deformazioni $\varepsilon_x, \varepsilon_y, \dots$ e viceversa:

$$\varepsilon_x = c_1\sigma_x + c_2\sigma_y + c_3\sigma_z + c_4\tau_{yz} + c_5\tau_{xz} + c_6\tau_{yx}$$

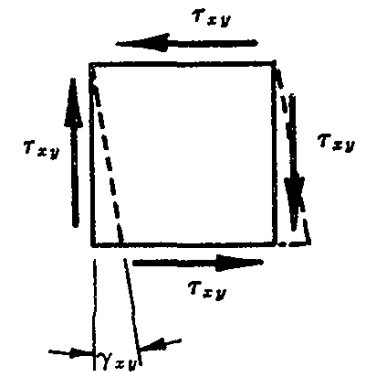
Dove il valore dei coefficienti elastici c_i dipende dai materiali e deve essere determinato sperimentalmente.

La dipendenza è particolarmente semplice nei casi di materiali: isotropici, omogenei ed elastici perché alcuni di tali coefficienti sono nulli ed il risultato ha la forma seguente:

$$\begin{aligned}\varepsilon_x &= \frac{1}{E}\sigma_x - \frac{\nu}{E}\sigma_y - \frac{\nu}{E}\sigma_z & \gamma_{xy} &= \frac{\tau_{xy}}{G} \\ \varepsilon_y &= \frac{\nu}{E}\sigma_x + \frac{1}{E}\sigma_y - \frac{\nu}{E}\sigma_z & \gamma_{yz} &= \frac{\tau_{yz}}{G} \\ \varepsilon_z &= \frac{\nu}{E}\sigma_x - \frac{\nu}{E}\sigma_y + \frac{1}{E}\sigma_z & \gamma_{zx} &= \frac{\tau_{zx}}{G}\end{aligned}$$



(a)



(b)

Legge di Hooke generalizzata

Guardando la figura seguente riusciamo a trovare il legame tra il modulo di rigidità G ed i moduli di Young E e Poisson ν , nel caso in cui vi sia un solo sforzo di taglio $\tau_{xy} = s$.

In tal caso infatti lo sforzo di taglio $\tau_{xy} = s$ è lo stesso che lo sforzo perpendicolare $\sigma_{x'} = +s$ e $\sigma_{y'} = -s$. Viceversa la deformazione è anche la stessa per i due elementi e vale:

$$\varepsilon_{x'} = +\frac{\gamma_{xy}}{2} = +\frac{\tau_{xy}}{2G} = +\frac{s}{2G}$$

$$\varepsilon_{y'} = -\frac{\gamma_{xy}}{2} = -\frac{\tau_{xy}}{2G} = -\frac{s}{2G}$$

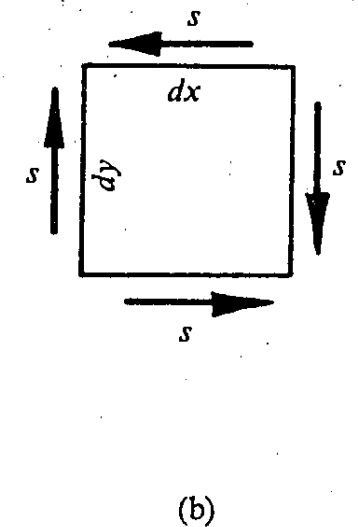
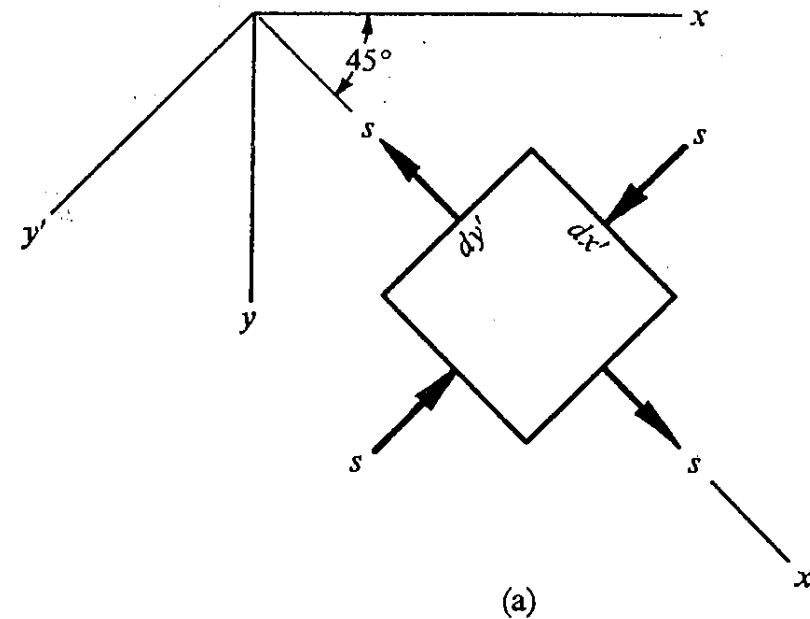
E, dalla legge di Hooke

$$\varepsilon_{x'} = +\frac{\sigma_{x'}}{E} - \nu \frac{\sigma_{y'}}{E} = +\frac{s(1+\nu)}{E}$$

$$\varepsilon_{y'} = +\frac{\sigma_{y'}}{E} - \nu \frac{\sigma_{x'}}{E} = -\frac{s(1+\nu)}{E}$$

Eguagliando le due relazioni ho:

$$\frac{s(1+\nu)}{E} = \frac{s}{2G} \quad \text{o} \quad \mathbf{G} = \frac{\mathbf{E}}{2(1+\nu)}$$



Slides

TROVATE LE SLIDE AGGIORNATE SULLA PAGINA MOODLE2 DEL CORSO

<https://moodle2.units.it/course/view.php?id=7348>

Nelle news del forum