

Appunti di Statistica

Mirolò Manuele / Alessio Brusini

a.a. 2025/26

Indice

0.1	Definizioni di probabilità	2
0.2	Definizione assiomatica di probabilità	2
0.3	Proprietà della misura di probabilità	2
0.3.1	Probabilità condizionata	3
0.3.2	Probabilità delle cause	3
0.4	Teorema di Bayes	3
0.5	Indipendenza	3
0.6	Variabili casuali e funzione di ripartizione	3
0.6.1	Funzione di ripartizione	4
0.7	Densità di probabilità	4
0.7.1	Proprietà della funzione densità di probabilità	5
0.8	Valore di aspettazione	5
0.9	Varianza	6
0.10	Momenti delle funzione di distribuzione	6
0.11	Distribuzione uniforme discreta	6
0.12	Distribuzione uniforme continua	6
0.13	Cenni di calcolo combinatorio e fattoriale	7
0.14	Distribuzione binomiale	8
0.15	Distribuzione di Poisson	8
0.16	Distribuzione di Maxwell-Boltzmann	9
1	Funzione di distribuzione di Gauss	12
1.1	Esempio di distribuzione di Gauss	13
1.2	Dimostrazione Euriatica	14
2	Funzione di distribuzione per variabili casuali	14
3	Funzione di distribuzione congiunta	17
3.1	Stima dei parametri	17
3.2	Fit di una retta	18
3.3	Caso non lineare	19
4	Confronto istogramma/gaussiana	19
5	Lezione 06/10/2025	20
5.1	Istogramma	20
5.2	In caso di limitazione dello strumento	20
5.3	Propagazione degli errori	21

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)

Il prof consiglia per l'analisi dei Dati, root del Cern. <https://root.cern.ch>

La probabilità è la misura del grado della credenza che un certo evento avvenga.

L'oggetto della teoria è un certo *esperimento causale*, la cui esecuzione è detta *prova*, il risultato è indicato con ω , l'insieme di tutti gli ω possibili è detto *spazio campione* Ω . Un *Evento* A è un sottoinsieme di Ω , se $\omega \in A$ allora esso realizza l'Evento. Se esso è l'unico elemento costituente di A allora esso è l'*evento elementare*, altrimenti A è un *evento composto*. Altre nomenclature sono legate alla teoria insiemistica (per esempio due eventi A e B sono disgiunti se l'insieme A non interseca l'insieme B).

0.1 Definizioni di probabilità

Si possono dare due definizioni di probabilità:

CLASSICA : dato uno spazio **finito** Ω , la probabilità è il rapporto:

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

Questa funzione è utile quando i casi risultano essere **ugualmente possibili**.

FREQUENTISTICA : emerge dal ragionamento fondato sui risultati di un esperimento:

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

dove n è il numero di volte che si verifica un evento, mentre N è le volte che viene ripetuto l'esperimento.

0.2 Definizione assiomatica di probabilità

Dato un evento $A \subseteq \Omega$ la misura di probabilità P è una funzione $P : \Omega \rightarrow \mathbf{R}$ le cui caratteristiche sono:

- $P(A) \geq 0$
- $P(\text{evento certo}) = 1$
- Se A e B sono eventi **disgiunti** (incompatibili), cioè $P(A \cap B) = 0$, allora $P(A \cup B) = P(A) + P(B)$, che nella forma generalizzata risulta essere:

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

0.3 Proprietà della misura di probabilità

- Probabilità negazione: $P(\bar{A}) = 1 - P(A)$
- Estremi della misura di probabilità: $0 \leq P(A) \leq 1$
- Teorema delle probabilità totali (con $A \cap B \neq \emptyset$):

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

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

- Probabilità in relazione d'inclusione (con $A \subseteq B$): $P(A) \leq P(B)$

0.3.1 Probabilità condizionata

La probabilità che si verifichi A se si verifica B :

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

Il **Teorema delle probabilità totale** mi enuncia che se ho n eventi e si ha che

$$P(M_i \cap M_j) = 0 \quad \forall i \neq j \quad \& \quad \bigcup_{i=1}^n M_i = \Omega$$

dove Ω è lo spazio campionario; allora:

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

Mentre per l'indipendenza si ha $P(A \cap B) = P(A)P(B)$

0.3.2 Probabilità delle cause

Ci chiediamo quale sia la probabilità che M_j sia la causa di un certo evento A .

0.4 Teorema di Bayes

Dato un certo set di ipotesi (M_i) e un certo evento A il teorema mi dice come la probabilità a "posteriori" $P(M_i|A)$ cambia dopo che l'evento A si è verificato.

Consideriamo un evento A e una classe completa di ipotesi (M_j):

$$\begin{cases} M_i \cap M_j = 0 \\ \bigcup_j M_j = \Omega \end{cases}$$

Utilizzando il teorema della probabilità condizionata e il fatto che $P(A \cap M_i) = P(M_i \cap A)$ posso ricavare la probabilità che una causa sia dovuta a un determinato evento

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

Dove:

$P(M)$: probabilità a priori/marginali (non tiene conto di A)

$P(M | A)$: probabilità condizionata/a posteriori, nota A

$P(A | M)$: probabilità condizionata di A , nota M

$P(A)$: è la probabilità di A a priori, serve da costante di normalizzazione

0.5 Indipendenza

Si dicono due eventi indipendenti: $P(A \cap B) = P(A) \cdot P(B)$ e quindi $P(A|B) = P(A)$

Per esempio: nella misura del pendolo fermare il pendolo ad ogni misura rende le misure indipendenti fra di loro (una misura di un periodo sbagliato non influisce sulla misura successiva).

0.6 Variabili casuali e funzione di ripartizione

Dato lo spazio di eventi elementare ω una *variabile casuale* è una funzione X tale che $X(\omega) = x$ dove $x \in \mathbb{R}$, dunque una funzione che mi prende un campione casuale dal mio spazio campionario e gli associa un numero (es: valore, altezza, ...).

Essa può essere discreta (numerabile) oppure continua, se almeno idealmente posso considerare Ω continuo.

0.6.1 Funzione di ripartizione

Data una variabile casuale X , la funzione F che fa corrispondere a un valore x la probabilità cumulativa $P(X \geq x)$ viene detta funzione di ripartizione

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

La funzione di ripartizione è definita sia per le variabili casuali discrete che per le variabili casuali continue. Proprietà delle funzioni di ripartizioni:

1. $0 \leq F_X(x) \leq 1$
2. nel caso che il dominio sia \mathbf{R} , allora $\lim_{x \rightarrow -\infty} F_X(x) = 0$ e $\lim_{x \rightarrow \infty} F_X(x) = 1$
3. È una funzione monotona crescente;
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 $x_1, x_2 \in \mathbf{R}$, 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:

$$\lim_{x \rightarrow x_0^+} F_X(x) = F_X(x_0).$$

mentre a sinistra abbiamo una discontinuità di 1° specie (o salto):

$$\lim_{x \rightarrow x_0^-} F_X(x) \neq F_X(x_0)$$

Questo è dovuto al fatto che la probabilità di un evento elementare è diversa da zero, quindi la funzione di ripartizione ha un salto in corrispondenza di ogni valore che può assumere la variabile casuale discreta (pensa lancio dei dati dove cambia la probabilità)

In un intervallo posso avere una **Probabilità uniforme**, ossia eventi in cui la funzione di probabilità è costante (ex nel caso dei dadi la funz. di ripartizione è a gradini)

0.7 Densità di probabilità

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

In simboli:

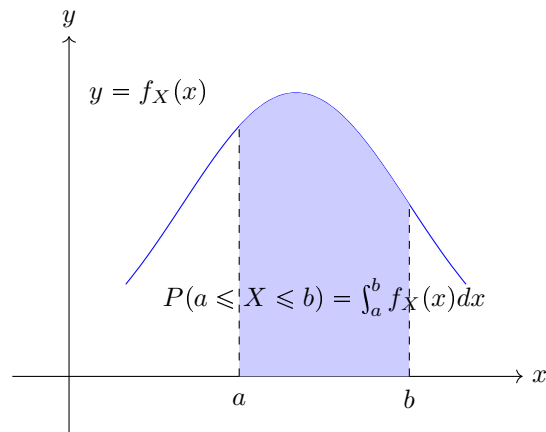
$$f_X : \mathbf{R} \rightarrow \mathbf{R}^+ : x \mapsto \lim_{dx \rightarrow 0} \left[\frac{P(x < X \leq x + dx)}{dx} \right] = \lim_{dx \rightarrow 0} \frac{F_X(x + dx) - F_X(x)}{dx}$$

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

Attenzione a non confondere la funzione di densità con una probabilità. Essa mi determina la densità di probabilità che la variabile casuale sia nell'intervallo dx .

Definiamo la **Funzione di distribuzione cumulativa**:

$$F(x) = \int_{x_{min}}^x f_X(t) dt$$



0.7.1 Proprietà della funzione densità di probabilità

- non può essere mai negativa (altrimenti avremmo una probabilità negativa per qualche X)
- condizione di normalizzazione:

$$\int_{-\infty}^{\infty} f(x) = 1$$

dunque si deve avere che: $\lim_{x \rightarrow \pm \infty} f_X(x) = 0$

- la probabilità che una variabile casuale assuma un particolare valore dell'intervallo è $P(X = x) = 0$, infatti ad un singolo valore corrisponde un intervallo con ampiezza nulla; questo mi porta a poter escludere gli estremi

0.8 Valore di aspettazione

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".

Data una variabile casuale $X \in [a, b]$, la funzione di densità a lei associata $f_X(x)$; andando a considerare una qualsiasi funzione $g(x)$, posso definire il valore di aspettazione di $g(x)$ come:

$$E[g(x)] = \int_a^b g(x)f_X(x)dx \quad o \quad E[g(x)] = \sum_{i=1}^n g(x_i)p_i$$

vado dunque a pesare i valori di $g(x)$ pesati con la densità di probabilità.

Il valore di aspettazioni ha proprietà simili a quelle di una applicazione lineare.

Per la variabile casuale continua il **valore di aspettazione** è dato da

$$\mu_x = E[x] = \int_{\Omega} x f_X(x) dx$$

μ sarà detto *valore medio*. Ragionamento analogo vale per le variabili casuali discrete.

Ovviamente si ha $E[\mu_x] = \mu_x$, poichè μ_x è una costante e la sua dimensione è quella variabile casuale. Si tratta di un "indice di posizione" all'interno del dominio.

Altri indici di posizione sono:

Moda : valore di x per cui f_X è max;

Quantili : probabilità che la variabile casuale cada in determinati intervalli:

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

0.9 Varianza

Tra i vari indici di dispersione la più usata è la varianza, che viene definita come:

$$\sigma^2 = E(g(x) - E(g(x)))^2$$

svolgendo i calcoli:

$$\sigma^2 = \int (x - \mu_x)^2 f_X(x) dx = \int (x^2 - 2x\mu_x + \mu_x^2) f_X(x) dx = \int x^2 f(x) dx - 2\mu_x \int x f(x) dx + \mu_x^2 \int f(x) dx$$

dove μ_x è una costante; inoltre l'integrale del secondo termine non è altro che la definizione di μ_x , abbiamo quindi che:

$$\sigma^2 = \int x^2 f(x) dx - \mu_x^2 = E[x^2] - E[x]^2$$

quindi $E[(x - \mu_x)^2] = E[x^2] - E[x]^2$

Questa è la prima importante proprietà della varianza, la seconda invece riguarda la disuguaglianza di Cebyshev

$$P(|x - \mu_x| > \lambda \sigma_x) \leq \frac{1}{\lambda^2}$$

Mentre definiamo la *deviazione standard* come $\sigma = \sqrt{\sigma^2}$

Example 1. *Varianza della media:*

$$y = y(x) = \frac{1}{n} \sum_{i=1}^n x_i$$
$$V(x) = \sum_{i=1}^n \left(\frac{\partial y}{\partial x_i} \right)^2 \sigma^2 = \frac{\sigma^2}{n}$$

0.10 Momenti delle funzione di distribuzione

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

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

Di conseguenza μ non è altro che il momento di ordine 1 rispetto all'origine, mentre σ^2 è il momento di ordine 2 rispetto alla media.

I momenti caratterizzano completamente le mie funzioni di distribuzioni, tanto che possiamo dire, sotto ipotesi non particolarmente restrittive, che due funzioni con gli stessi momenti coincidono.

0.11 Distribuzione uniforme discreta

Se tutti i valori che può assumere una certa variabile X sono equiprobabili, utilizzando la condizione di normalizzazione si ottiene $p = 1/n$, quindi

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

ovvero la media aritmetica dei possibili valori (come nel dado). La varianza è $\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2$

0.12 Distribuzione uniforme continua

Una variabile casuale continua su (a,b) ha funzione di distribuzione uniforme, se $f_X(x) = \text{cost}$. Per la condizione di normalizzazione si deve avere $f_X(x) = \frac{1}{b-a}$.

Il valore medio è dunque:

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

Per la varianza invece si ottiene che:

$$\sigma_x^2 = E[x^2] - E[x]^2 = \frac{(b-a)^2}{12}$$

Nota: Per cambiare dalla trattazione degli errori massimi a quella statistica si ha che la mia misura eseguita avrà come risultato: $x \pm \Delta x$, quindi $\sigma_x^2 = \frac{(b-a)^2}{12}$ e $\sigma_x = \frac{b-a}{2\sqrt{3}}$. In questo caso $b-a = 2\Delta x$, quindi $\Delta x = \sigma_x\sqrt{3}$.

La probabilità che x sia nell'intervallo è: $P(\mu_x - \sigma_x < x \leq \mu_x + \sigma_x) = \frac{1}{\sqrt{3}}$

Infatti:

$$P(\mu_x - \sigma_x < x \leq \mu_x + \sigma_x) = \int_{\mu_x - \sigma_x}^{\mu_x + \sigma_x} f(x) dx = \frac{1}{b-a} \int_{\mu_x - \sigma_x}^{\mu_x + \sigma_x} dx = \frac{1}{b-a} 2\sigma_x = \frac{1}{\sqrt{3}}$$

0.13 Cenni di calcolo combinatorio e fattoriale

Si chiama combinazione semplice quella in cui non importa l'ordine degli elementi e il numero di combinazioni $C_{n,r}$ è dato da:

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

Con la definizione di fattoriale:

$$n! = \prod_{k=1}^n k = 1 \cdot 2 \cdot 3 \cdots n = \prod_{i=0}^{n-1} (n-i)$$

Fattoriale di un numero negativo:

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

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

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

Fattoriale di un numero positivo:

$$x! = \Pi(x) = \int_0^\infty t^x e^{-t} dt$$

Che presenta la proprietà:

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

Sempre Eulero introdusse la funzione Gamma Γ , con uno spostamento di unità nelle variabili rispetto a Π :

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

Difatti questa funzione è legata a Π da $\Pi(x) = \Gamma(x+1) = x$ e vale che $z\Gamma(z) = \Gamma(z+1)$

Da questa funzione sono definite le combinazioni semplici del tipo:

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

Importanti relazioni tra le funzioni Gamma sono:

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

Da queste ricavo che $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ e $\Gamma(-\frac{1}{2}) = -\sqrt{\pi}$, l'ultima importante relazione è invece:

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

0.14 Distribuzione binomiale

Nel caso che un fenomeno possa verificarsi soltanto con due modalità mutualmente escluse (successo p o insuccesso q), la probabilità che un determinato evento si verifichi è:

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

quindi la probabilità di avere:

- n successi è p^n
- 0 è q^n
- 1 è pq^{n-1}
- $n-1$ è $p^{n-1}q$

Le **proprietà** della distribuzione binomiale sono le seguenti:

- la distribuzione binomiale è normalizzata, infatti la sommatoria non è altro che il binomio di Newton: $\sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$;
- il valore di aspettazione di k è: $E[k] = np$
- la varianza è $\sigma_x^2 = E[x^2] - E[x]^2 = npq$

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

Supponiamo di avere n oggetti da poter essere messi in r spazi, avrò n possibilità per riempire il primo spazio, $n-1$ per il secondo e così via, quindi il numero di modi possibili sarà $D_{n,r} = \prod_{k=0}^r n-k$, moltiplicando per $(n-r)!$ si ottiene $n!$

La sua importanza sta anche nel fatto che, a $\lim_{n \rightarrow \infty}$ permette di ottenere la funzione di **distribuzione di Gauss** se $p = cost$ e la **distribuzione di Poisson** se $np = cost$, entrambe usatissime in fisica.

Consideriamo un intervallo i di un istogramma e consideriamo favorevole un evento che cade nell'intervallo, su un totale di eventi n . Il numero di eventi attesi nell'intervallo considerato è $E[n_i] = np_i$

La varianza di n_i è $\sigma_{n_i}^2 = np_i(1-p_i) \simeq n_i(1-n_i/n)$.

Posso prendere come valore stimato di p_i il valore $\hat{p}_i = n_i^m/n$, dunque per una probabilità $p \rightarrow 0$; essa corrisponde a $\sigma_{n_i} = \sqrt{n_i^m}$.

0.15 Distribuzione di Poisson

È un caso limite della fisica nucleare, e serve per indicare la probabilità di osservare k eventi in un determinato intervallo t :

- la presenza/assenza dell'evento non dipende da quelli prima di t
- la probabilità di un evento è proporzionale a t
- la probabilità di avere due eventi contemporaneamente è nulla

Si tratta del caso in cui i successi p della binomiale siano piccoli e il numero di eventi $n \rightarrow \infty$.

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

La probabilità di avere 0 eventi in $(0, t+dt)$:

$$P_0(t+dt) = P_0(t)(1-\mu dt) \rightarrow P_0(t) = e^{-\mu t}$$

dove ho impostato che $P_0(0) = 1$ e sottraggo μdt che è la probabilità di avere un evento in dt ($e^{-\mu dt} \simeq 1 - \mu dt$ tolgo il termine quadratico risultante)

La probabilità di avere un evento è invece:

$$P_1(t + dt) = P_1(1 - \mu dt) + P_0(\mu dt) \rightarrow \frac{dP_1}{dt} = -\mu(p_1(t) - p_0(t))$$

Continuando in maniera ricorsiva trovo che la probabilità di avere "m" eventi in un determinato intervallo è:

$$\frac{dP_m}{dt} = -\mu(p_m(t) - p_{m-1}(t))$$

che ha come soluzione:

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

dove $m = \mu t = E[k]$ mi rappresenta il valore di aspettazione di un certo k

La distribuzione di Poisson naturalmente è normalizzata (ricordando lo sviluppo di Taylor dell'esponenziale), infatti:

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

e il valore di aspettazione non è altro che:

$$E[k] = \sum_{k=0}^{\infty} k P_k = m$$

e la varianza è:

$$\sigma^2 = E[k^2] - E[k]^2 = m$$

Si può notare che la distribuzione di Poisson non è altro che il limite per $n \rightarrow \infty$ della distribuzione di Bernulli.

0.16 Distribuzione di Maxwell-Boltzmann

Si considera un gas con peso molare $M = N_A m$ in equilibrio termodinamico a temperatura T . La costante dei gas è $R = N_A k \approx 8.31 \text{ J K}^{-1} \text{ mol}^{-1}$.

La velocità quadratica media delle molecole del gas è legata alla temperatura dalla relazione:

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

Tuttavia, le singole molecole hanno velocità che si discostano da questo valore medio.

Maxwell nel 1859 derivò la distribuzione di probabilità delle velocità $f(v)$:

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

All'aumentare della temperatura, il picco della distribuzione (velocità più probabile) si sposta verso valori più elevati.

Le curve sono normalizzate in modo che l'integrale della probabilità sia sempre uguale a 1.

La distribuzione fu ottenuta da Maxwell attraverso un ragionamento particolarmente ingegnoso:

- introduciamo prima alcune definizioni:

$$\rho(\vec{r}) = \frac{dm}{dV} \quad \text{densità del gas,}$$

$$\eta(\vec{r}) = \frac{\rho(\vec{r})}{m} \quad \text{densità di particelle,}$$

$$dn = \eta(\vec{r}) dV \quad \text{numero di particelle in } dV.$$

- da cui ricaviamo che la distribuzione delle particelle nello spazio è:

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

- introduciamo una densità di particelle nello spazio delle velocità, in modo analogo a quello tridimensionale:

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

- introduciamo ora la funzione di distribuzione normalizzata della velocità $f(\vec{v})$:

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

- se la distribuzione della velocità è isotropa, allora la funzione di distribuzione dipende solo dal modulo e non solo dal verso;
- Il numero di particelle dn comprese nell'elemento $d\vec{v} = dv_x dv_y dv_z$ dello spazio delle velocità vale:

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

che in coordinate e angoli polari si può scrivere:

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

- dato che ho dei valori che non dipendono dagli angoli posso trovare che:

$$dn(v) + c = 4\pi n_T p(K) v^2 dv$$

Nota: Ricordiamo la definizione di angolo solido:

Vogliamo sapere quante molecole hanno il vettore velocità compreso fra \vec{v} e $\vec{v} + d\vec{v}$. Passiamo alle coordinate polari sferiche (r, θ, ψ) , l'area A identificata dal vettore \vec{v} sarà l'intersezione della superficie sferica di due cerchi la cui latitudine differisce per $d\theta$ e due la cui longitudine differisce per $d\psi$. Di conseguenza si ha:

$$dA = (r d\psi)(r \sin(\theta) d\theta)$$

L'angolo solido sarà definito da:

$$d\omega = \frac{dA}{r^2} = \sin(\theta) d\theta d\psi$$

Il numero di particelle con velocità $v + dv$ sarà: $dn(v) \frac{d\omega}{4\pi}$ essendo 4π l'angolo solido massimo. Questa espressione ci dice che non c'è una direzione preferenziale per le velocità.

Se consideriamo inoltre che le molecole possono scontrarsi fra di loro prima di colpire l'area A , dobbiamo limitarci a considerare un cilindro di lunghezza $vd\tau$ dove τ è talmente piccolo che non avvengono urti fra le molecole, esso vale $dV = vd\tau \cos(\theta) dA$.

Il numero di molecole che urterà l'area dA sarà $dn(v) \frac{d\omega}{4\pi} \frac{dV}{v}$.

- La probabilità che l'urto accada è

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

ovvero è proporzionale alla probabilità di avere le due particelle con K_1 e K_2 e dopo K'_1 e K'_2 .

Posso ora considerare due particelle che hanno energia cinetica iniziale

$$P = C p(K'_1) p(K'_2)$$

e, e che dopo urto abbiamo K_1 e K_2 .

Il processo può essere visto da ambo le parti, $C = C'$

- il gas è essendo in equilibrio termodinamico deve essere che l'energie cinetiche si equivalgono e si conservano, dunque:

$$P = p(K_1) p(K_2) = p(K'_1) p(K'_2) = p(K_1 + x) p(K_2 - x)$$

dove x rappresenta un incremento (\pm) dell'energia

- dunque l'unica funzione che soddisfi questi requisiti è:

$$p(K) = A e^{-\beta \frac{mv^2}{2}}$$

dove β (> 0) è una costante che mi serve per rendere il mio esponente adimensionale. Quindi:

$$dn(v) + c = 4\pi n_T A e^{-\beta \frac{mv^2}{2}} v^2 dv$$

Nota:

le costanti A e β sono da determinare, per la prima integriamo da 0 a ∞ (tutte le velocità), arriviamo a

$$A = \left(\frac{\frac{1}{2}m\beta}{\pi} \right)^{\frac{1}{2}}$$

quindi per β utilizziamo la definizione della media quadratica di tutte le velocità

$$\langle v^2 \rangle = \frac{1}{N} \int_0^\infty v^2 dn(v).$$

Sostituendo l'espressione di $dn(v)$, si ha

$$\langle v^2 \rangle = 4\pi A^3 \int_0^\infty v^4 \exp[-\beta(\frac{1}{2}mv^2)] dv$$

Il cui valore è: $\frac{3}{m\beta}$

Dalle supposizioni abbiamo che l'unica energia è quella cinetica, dal teorema di equipartizione dell'energia, essendoci 3 gradi di libertà:

$$\frac{1}{2}m\langle v^2 \rangle = \frac{3}{2}k\theta \rightarrow \frac{1}{2}m \frac{3}{m\beta} = \frac{3}{2}k\theta,$$

cioè

$$\beta = \frac{1}{k\theta}.$$

- imponendo le condizioni di contorno, ovvero $\int_V dn_v = n_T$ e $\int_V K dn_v = \frac{3}{2}n_T K T$ ottengo che:

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

- la funzione presenta un massimo in $v_M = \sqrt{\frac{2kT}{m}} \rightarrow kT = \frac{1}{2}mv_M^2$, la velocità più probabile!!
- 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:

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

Costante di normalizzazione	Probabilità di occupazione del livello energetico	Spazio delle fasi
-----------------------------	---	-------------------

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

Essa mi garantisce:

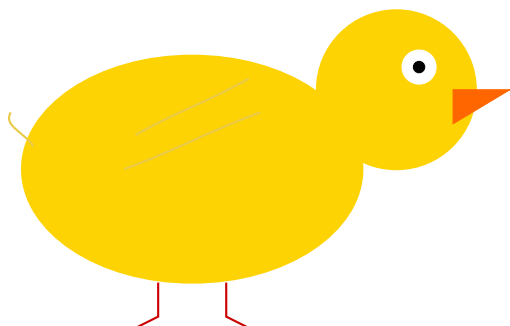
- La costante di normalizzazione assicura che l'integrale della $p(v)$ esteso a tutte le velocità possibili sia uguale a 1.
- La statistica di Boltzmann tiene conto della probabilità di occupazione di uno stato avente una data energia in funzione della temperatura del sistema.
- Infine lo spazio delle 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 $K = \frac{mv^2}{2}$.

1 Funzione di distribuzione di Gauss

Data x variabile casuale:

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

Quarck!



Essa rispetta le proprietà delle funzioni di distribuzione:

- è normalizzata

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

- il valore di aspettazione è μ

$$E[x] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu_x$$

- la varianza vale σ^2

$$Var(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x - \mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma_x^2$$

Per centrarla in 0 scrivo $y = x - \mu$ e $dy = dx$, ottengo

$$f(y; 0, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y)^2}{2\sigma^2}}$$

per centrarla in 0 e con $\sigma^2 = 1$ pongo $y = \frac{x-\mu}{\sigma}$ e $dy = \frac{dx}{\sigma}$ si ha:

$$f(y; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y)^2}{2}}$$

questa è detta *distribuzione normale standard*.

E' interessante notare che in questo caso la funzione di ripartizione ha l'interessante proprietà:

$$F(-x) = 1 - F(x)$$

Se vogliamo trovare la probabilità che x stia in un certo intervallo $[a, b]$ avremo:

$$P(a \leq x \leq b) = P(b \leq x) - P(a \leq x) = F(b) - F(a)$$

Per la variabile standardizzata y :

$$P\left(\frac{x-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) - P\left(\frac{x-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) = F\left(\frac{b-\mu}{\sigma}\right) - F\left(\frac{a-\mu}{\sigma}\right)$$

Scegliendo come estremi $a = \mu - n\sigma$ e $b = \mu + n\sigma$:

$$P(\mu - n\sigma \leq x \leq \mu + n\sigma) = F(n) - F(-n) = 2F(n) - 1$$

Si può concludere che risulta normale trovare valori fuori dall'intervallo $\pm\sigma$ poichè, per esempio, per $n = 3$ la probabilità è di 0.980. Graficamente si ha che:

- Nel grafico a campana il valore che massimizza f è μ , il valore a mezza altezza:

$$f(x) = \frac{1}{2} \frac{1}{\sqrt{2\pi}\sigma}$$

la cui larghezza vale: $2\sqrt{2\ln(2)}\sigma$

1.1 Esempio di distribuzione di Gauss

Data una distribuzione binomiale abbiamo che la variabile casuale ha il seguente valore:

Nelle ipotesi che:

- l'errore delle misure è $u = a - x$ (x =valore misurato, a = valore vero)
- l'errore è dovuto da n errori, introducendo un errore finale ε

Il contributo di ε sarà sia positivo che negativo diviso a 50/50 (legge dei grandi numeri), infatti $u_k = \varepsilon k - (n - k)\varepsilon$, dove k mi dà un contributo positivo, mentre $(n-k)$ negativo. Di conseguenza il suo valore di aspettazione $E[u_k] = 0$, per il limite di $n \rightarrow \infty$ ho una distribuzione gaussiana centrata in 0.

1.2 Dimostrazione Euriatica

Vogliamo dimostrare che la distribuzione degli errori accidentali è gaussiana. Abbiamo visto che nel caso di k deviazioni elementari positive su i totali il valore dell'errore e la probabilità sono:

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

se consideriamo che $k+1$ valori producono deviazione positive posso fare una media tra le due trovando che:

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

Per $\Delta U \rightarrow 0$, posso assumere che il valore sia dato dal valore medio di $P(k)$ $P(k+1)$; la **densità di probabilità** dell'errore per l'intervallo considerato:

$$y(u) = \frac{P}{\Delta U} \quad P = \frac{1}{2}(P(k) + P(k+1))$$

mentre:

$$\Delta y = \frac{\Delta P}{\Delta U} \quad P = (P(k) - P(k+1))$$

Da cui posto alcuni conti posso trovare che:

$$\frac{\Delta y}{y} = \frac{\Delta P}{P} = -\frac{u \Delta u}{\sigma^2}$$

con $\sigma^2 = (n+1)\varepsilon^2$, che diventa una quantità finita al limite, ho che:

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

Breve riassunto:

- defi. variabile casuale discreta o continua
- def. funzione di ripartizione $F_x(x)$
- def. di funzione di distribuzione e di densità di variabili casuali
- Valori di aspettazione come indici (posizione/dispersione)
- esempi di funzioni di distribuzioni

2 Funzione di distribuzione per variabili casuali

Se la funzione di densità di probabilità dipende da più variabili si ha, per il valore di aspettazione

$$E(x_1, \dots, x_n) = \int \vec{x} f(\vec{x}) d\vec{x}$$

dove $\vec{x} = (x_1, \dots, x_n)$.

Per un singolo valore x_i si ha

$$E(x_i) = \int x_i f(\vec{x}) d\vec{x}$$

Si definisce allora l'elemento di posizione ij della *matrice della covarianza*:

$$V_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)] = \int (x_i - \mu_i)(x_j - \mu_j) f(\vec{x}) d\vec{x}$$

si ha allora che:

- $\sigma_i^2 = V_{ii} = \sigma^2$
- la *covarianza* è un qualsiasi elemento della matrice $i \neq j$

E' utile definire il *coefficiente di correlazione*:

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

Si ha così una nuova definizione di variabili indipendenti:

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

$$E(x_i, x_j) = E(x_i)E(x_j) \rightarrow Cov(x_i, x_j) = 0 \text{ e } \rho(x_i, x_j) = 0$$

è la probabilità che la n variabili causali cadano all'interno di certi intervalli contemporaneamente, essa ha le stesse proprietà della funzione di distribuzione usuale, si può così definire la funzione di distribuzione marginale per una variabile x_i , ottenuta integrando la funzione citata sopra sui vari intervalli dove $x_j \neq x_i$

Posso costruire il valore di aspettazione di due variabili causali (**covarianza**), il cui valore è 0 se le variabili sono indipendenti fra loro

$$Cov(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Se le variabili sono indipendenti la funzione di distribuzione di Gauss diviene:

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

La distribuzione congiunta

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

Se le varianze sono i valori medi:

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

Data $y = y(x)$ variabile casuale dipendente da un'altra variabile casuale x , nel caso di **corrispondenza biunivoca** (ex. $y = ax - b$) si deve avere:

$$f(x_a)dx = g(y_a)dy$$

dove $y_a = f(x_a)$, allora la determinazione di $g(y)$ si riduce ad una trasformazione di variabile. Ossia le funzioni di distribuzioni devono essere identiche, se esiste la funzione inversa $x = x(y)$ si ha:

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

Nel caso invece di funzioni **non biunivoche**, cioè esistono m valori x_i tali che $y = y(x_i)$ (ex. $y = ax^2$), allora "sommo le diverse componenti biunivoche":

$$g(y) = \sum_{i=0}^n f(x_i(y)) \left| \frac{dx_i(y)}{dy} \right|$$

Se è difficile calcolare la $g(y)$, utilizzo il valore di aspettazione

$$\mu_y = E[y] = \int y g(y) dy = E[y(x)] = \int y(x) f(x) dx$$

$$\sigma_y^2 = E[(y(x) - \mu_y)] = \int (y(x) - \mu_y)^2 f(x) dx$$

Calcolo dunque il valore di aspettazione e la varianza

Se l'andamento è lineare $y(x) = ax + b$ si ha:

$$E[y(x)] = aE(x) + b$$

Utilizzando questo è facile verificare che:

$$\sigma_y^2 = E[y^2] - \mu_y^2 = a^2 \sigma_x^2$$

Che è quanto si ottiene dalla legge di propagazione della varianza.

Se $y(x)$ non è lineare utilizzo l'approssimazione lineare intorno a μ_x e tronco al secondo ordine:

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

Dove il secondo termine scompare perchè $E[x] = \mu_x$

Per la varianza

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

Esattamente ciò che si ottiene con la **legge di propagazione della varianza**.

Sapendo che per variabili indipendenti $E[x_1, x_2] = E[x_1]E[x_2]$

Per funzioni a più variabili

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

Ricordiamo inoltre che se $y = \sum_{i=1}^n x_i^{\alpha_i}$ allora:

$$\frac{\sigma_y^2}{y^2} = \sum_{i=1}^n \alpha_i^2 \left(\frac{\sigma_{x_i}}{x_i} \right)^2$$

da cui si può ricavare quali sono gli errori trascurabili.

Vediamo il caso di n misure:

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

La varianza sarà:

$$\sigma_x^2 = \sum_{i=1}^n \left(\frac{\partial \bar{x}}{\partial x_i} \right)^2 \sigma_{x_i}^2$$

Essendo che questa derivata vale sempre $\frac{1}{n^2}$ allora:

$$\sigma_x^2 = \frac{1}{n^2} \sum_{i=1}^n \sigma_{x_i}^2 \rightarrow \sigma_x = \frac{\sigma_i}{\sqrt{n}}$$

Se ho compiuto 2 misure e mi chiedo se esse siano **compatibili**, costruisco una funzione derivata come $y = x_1 - x_2$ (differenza fra le misure), supponiamo $\sigma_1 \neq \sigma_2$ e $E[x_1] = E[x_2] = \mu$

$$E[y] = E[x_1 - x_2] = 0$$

$$\sigma_y^2 = \sum_{i=1}^n \left(\frac{\partial y}{\partial x_i} \right)^2 \sigma_{x_i}^2 = \sigma_1^2 + \sigma_2^2$$

La funzione di distribuzione Gaussiana

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

Si ha compatibilità quando

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

3 Funzione di distribuzione congiunta

Da un set di misure indipendenti posso costruire la "funzione di distribuzione congiunta" (prodotto fra le funzioni gaussiane delle n variabili):

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} e^{-\sum_{i=1}^n \frac{(x_i - \mu_i)^2}{2\sigma_i^2}} \text{ per } \mu \text{ e } \sigma = \text{cost.} \rightarrow \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}}$$

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

3.1 Stima dei parametri

Posso ottenere una stima dei parametri della funzione di distribuzione, voglio che sia a minima varianza (rispetto ai dati sperimentali che ho raccolto) e che il valore sia centrato nel valore vero, posso usare:

- **Metodo della massima verosimiglianza** (o Likelihood)
- **Metodo dei minimi quadrati**, meno preciso rispetto al primo

Dato un set di misure ho che la funzione di distribuzione cumulativa (o congiunta) $f(x_1, \dots, x_n, \lambda_1, \dots, \lambda_n)$, dove λ_i sono i parametri da stimare. Voglio trovare i valori $\lambda_1, \dots, \lambda_n$ che minimizzano il valore di $\mathfrak{L} = f(x_1^m, \dots, x_n^m, \lambda_1, \dots, \lambda_n)$, dove x_i^m sono i valori misurati.

Questo perchè voglio massimizzare la probabilità di trovare i miei valori misurati all'interno di $x_i \pm \Delta x$

In realtà però si punta a massimizzare il logaritmo di \mathfrak{L} , per esempio per delle misure con distribuzione gaussiana

$$\ln(\mathfrak{L}) = \ln \frac{1}{2\pi\sigma^n} - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}$$

Le equazioni di Likelihood sono:

$$\frac{\partial \mathfrak{L}}{\partial \mu} = 0 \quad \frac{\partial^2 \mathfrak{L}}{\partial \sigma^2} < 0$$

Da cui si ottengono:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad e \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

ma facendo qualche calcolo

$$(x - \mu)^2 = (x - \bar{x} + \bar{x} - \mu)^2 = (x - \bar{x})^2 + -2(\bar{x} - \mu)(x - \bar{x}) + (\bar{x} - \mu)^2$$

da cui:

$$E[(x - \mu)^2] = E[(x - \bar{x})^2] + E[2(\bar{x} - \mu)(x - \bar{x})] + E[(\bar{x} - \mu)^2]$$

$$\sigma^2 = E[(x - \bar{x})^2] + \frac{\sigma^2}{n}$$

e raccogliendo

$$\sigma^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

Supponiamo di avere più misure con σ differenti, in questo caso allora la misura x_1 avrà una funzione di distribuzione gaussiana e la distribuzione congiunta sarà una produttoria di queste funzioni. In questo caso la funzione di distribuzione manterrà la sommatoria anche per i σ_i :

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

Utilizzando le equazioni di Likelihood si ottengono:

$$\mu = \frac{\sum_{i=1}^n \frac{x_i^m}{\sigma_i^2}}{\sum_{j=1}^n \frac{1}{\sigma_j^2}}$$

Ovvero una media pesata delle misure per le loro varianze dal termine, in cui la sommatoria del fattore di peso è uguale a 1, le misure più precise hanno peso maggiore.

$$\sigma_\mu = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

Un altro metodo per stimare i parametri è quello dei **minimi quadrati**. Dato un set di misure $\vec{x} = (x_1, \dots, x_n)$ e di conseguenza valori $\vec{y} = (y_1, \dots, y_n)$, il valore vero dei parametri y_i non è noto, ma può essere stimato da $f(\lambda_i, x_i)$. Il principio dei minimi quadrati afferma che la miglior stima dei parametri $(\lambda_1, \lambda_2, \dots, \lambda_n)$ è quella che minimizza la quantità:

$$\chi^2 = \sum_{i=1}^n w_i (y_i - f(x_i))^2$$

dove w_i è un peso, se esso è lo stesso per tutte le misure allora questo termine può essere tolto, se invece esso è un certo σ_i allora esso diventa $\frac{1}{\sigma_i^2}$ ($X^2 = \chi^2$). Infine se le variabili sono correlate avremo che $w_i = \frac{1}{\text{cov}(i,j)}$. Per misure con stessi σ_i e σ_i differenti ottengo:

$$\lambda_i = \sum_{i=1}^n \frac{y_i}{n} \quad \lambda_i = \frac{\sum_{i=1}^n \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

3.2 Fit di una retta

Vogliamo trovare la miglior retta che descrive un set di misure e conseguenti risultati $(x_i, y_i \pm \sigma_i)$, utilizziamo la parametrizzazione: $f_i = mx + q$, i parametri sono m e q. Dovrò risolvere il sistema di equazioni composto da $\frac{\partial X^2}{\partial m} = 0$ e $\frac{\partial X^2}{\partial q} = 0$

Se consideriamo gli errori trascurabili o tutti uguali, possiamo costruire la funzione dei minimi quadrati:

$$X^2 = \sum_{i=1}^N (y_i - f_i)^2 = \sum_{i=1}^N (y_i - \theta_1 - \theta_2 x_i)^2$$

Il minimo si trova ponendo le derivate sia rispetto a θ_1 che θ_2 uguale a 0.

Introducendo i valori medi di x e y :

$$q = \hat{\theta}_1 = \frac{s_y \cdot s_{xx} - s_x \cdot s_{xy}}{N s_{xx} - s_x^2} = \frac{\bar{y} \cdot \overline{x^2} - \bar{x} \cdot \overline{xy}}{\overline{x^2} - \bar{x}^2}$$

$$m = \hat{\theta}_2 = \frac{N s_{xy} - s_x \cdot s_y}{N s_{xx} - s_x^2} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}$$

Gli errori si trovano a partire dalla legge di propagazione della varianza,

$$\sigma_m^2 = \sum_{i=1}^N \left(\frac{\partial m}{\partial y_i} \right)^2 \sigma^2 \sigma_q^2 = \sum_{i=1}^N \left(\frac{\partial q}{\partial y_i} \right)^2 \sigma^2$$

da cui si ottiene (considerando σ tutti uguali):

$$\sigma_q = \sigma^2 \frac{\overline{x^2}}{N(\overline{x^2} - \bar{x}^2)} \quad \sigma_m = \sigma^2 \frac{1}{N(\overline{x^2} - \bar{x}^2)}$$

Mentre nel caso in cui abbiamo σ differenti abbiamo che le operazioni forniscano il valore y_i è:

$$f(y_1, y_2, \dots, y_N) \sim \frac{1}{\prod_{i=1}^N \sigma_i} e^{-\sum_{i=1}^N \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{2\sigma_i^2}}$$

Allora introduciamo la funzione χ^2 che ci permetterà di fare il *test delle ipotesi*, ovvero capire se l'andamento dei nostri dati è lineare o meno. la distribuzione di χ^2 coincide con la funzione di distribuzione cumulativa della variabile casuale X^2 :

$$X^2 \equiv \chi^2 = \sum_{i=1}^N \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{\sigma_i^2}$$

Per aver la miglior stima procediamo con prima, andando a massimizzare la funzione χ^2 , da cui ottengo:

$$q = \hat{\theta}_1 = \frac{s_{xx} \cdot s_y - s_{xy} \cdot s_x}{s \cdot s_{xx} - s_x^2}, \quad m = \hat{\theta}_2 = \frac{s \cdot s_{xy} - s_x \cdot s_y}{s \cdot s_{xx} - s_x^2} \quad m = \hat{\theta}_2 = \frac{s \cdot s_{xy} - s_x \cdot s_y}{s \cdot s_{xx} - s_x^2}$$

dove

$$s_x = \sum_{i=1}^N \frac{x_i}{\sigma_i^2}, \quad s_y = \sum_{i=1}^N \frac{y_i}{\sigma_i^2}, \quad s_{xx} = \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2}, \quad s_{xy} = \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2}, \quad s = \sum_{i=1}^N \frac{1}{\sigma_i^2}$$

Posso scrivere

$$m = \frac{1}{N} \sum \frac{x_i - \bar{x}}{x^2 - \bar{x}^2} y_i \quad q = \frac{1}{N} \sum \frac{y_i - \bar{y}}{x^2 - \bar{x}^2}$$

Per le varianze (utilizzando la legge di propagazione della varianza):

$$\sigma_q^2 = \sigma_{\theta_1}^2 = \frac{s_{xx}}{S \cdot s_{xx} - s x^2} \quad \sigma_m^2 = \sigma_{\theta_2}^2 = \frac{s}{s \cdot s_{xx} - s x^2}$$

3.3 Caso non lineare

Il metodo più semplice è quello di trovare una legge lineare che descrive il mio fenomeno

Example 2. $s = \frac{1}{2}gt^2$ trascuro l'incertezza su s mentre considero l'errore su t , per calcolare g . invertendo ho $t^2 = \frac{2s}{g}$ e prendendo il quadrato mi riconduco ad un'equazione lineare: $t = \sqrt{\frac{2}{g}}\sqrt{s} - t_0$.

$$\sigma_g = \left(\frac{\partial g}{\partial m}\right)^2 \sigma_m^2 \text{ con } m = \sqrt{\frac{2}{g}}, \text{ questo ci fa notare l'importanza del significato assegnato alle variabili}$$

In altri casi posso usare Taylor, oppure sviluppare la derivata in maniera numerica

4 Confronto istogramma/gaussiana

Date 100 misure che cadono in k intervalli di valore centrale x^k e larghezza Δx , ci chiediamo quale sia la funzione di distribuzione che mi definisca la probabilità che un certo dato cada in un certo intervallo (successo), questa è la **funzione di distribuzione binomiale**:

Sapendo che la probabilità per un intervallo k è: $P_k = \binom{n}{l} p^l (1-p)^{n-l}$, dove p è la probabilità di successo e l il numero di successi, so che $E[l] = np$ e $\sigma^2 = np(1-p)$.

Considerando ora una distribuzione di tipo Gaussiano la funzione che la descrive è la classica e la probabilità che un certo valore cada in un intervallo è:

$$P_k = \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} f(x, \hat{\mu}, \sigma^2) dx \simeq f(\bar{x}_k) \Delta x.$$

da cui ottengo il numero di eventi atteso all'interno dell'intervallo n_k e per ogni intervallo ho un certo σ , per ognuno di questi intervalli posso fare un confronto per verificare la compatibilità

$$|n_k^s - n_k^m| \leq 3\sigma_k$$

dove m sta per eventi misurati ed s per eventi attesi, avremo quindi un grafico con un valore atteso di eventi con barra di errore di larghezza σ , dunque in un grafico mi aspetto che:

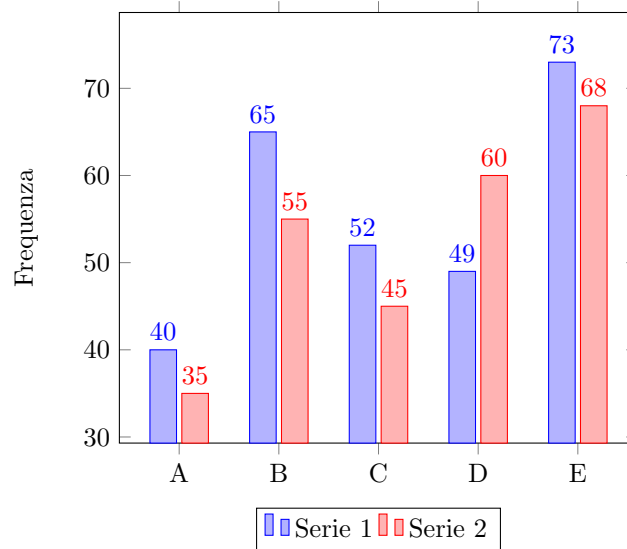
- il 68% deve toccare con le barre di errore il valore di aspettazione
- il 27% si trova a 2σ
- il 5% si trova a 3σ

La deviazione standard della binomiale è circa $\sqrt{n_k^m}$ se $n_{tot} \gg n_k^c$, che in realtà è un'approssimazione che va bene quando $n_k^c > 5$

5 Lezione 06/10/2025

Le nostre misure saranno del tipo x_1, x_2, \dots, x_N , il nostro obiettivo è quello di rappresentare questi dati per capire il fenomeno fisico che stiamo studiando.

5.1 Istogramma



Considero la variabile media x come $x = \frac{1}{N} \sum_{i=1}^N x_i$, gli scarti della media, per definizione devono annullarsi:

$$\frac{1}{N} \sum (\bar{x} - x_i) = 0$$

Per avere maggiori informazioni posso allora definire la varianza della mia variabile casuale x come:

$$\text{var}(x) = \frac{1}{N} \sum (x - \bar{x})^2$$

è utile svolgere il quadrato

$$= \frac{1}{N} (x_i^2 - 2\bar{x}x_i + \bar{x}^2) = \frac{1}{N} \sum x_i^2 - 2\bar{x} \frac{1}{N} \sum x_i + \frac{1}{N} \sum \bar{x}^2 = \bar{x}^2 - 2\bar{x}^2 + \bar{x}^2 = \bar{x}^2 - \bar{x}^2$$

definisco la deviazione standard come $\sigma_x = \sqrt{\text{var}(x)}$, anche se, vista la limitatezza del campione si attua una correzione (di Bessel):

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}$$

5.2 In caso di limitazione dello strumento

Definisco l'errore come la semiampiezza dell'intervallo $\Delta x = \frac{X_{\max} - X_{\min}}{N}$ e il risultato della misura sarà $x \pm \Delta x$.

L'interpretazione statistica in questo caso è che il vero valore della mia misura appartiene a un insieme di misure $\mathbf{X} \in x$ e ad ogni x associo una probabilità $P(x)$, ma non sempre è possibile, per questo uso la *densità di probabilità* definita come:

$$P(x_1 < \mathbf{X} < x_2) = \int_{x_1}^{x_2} P(x) dx$$

ciò vale anche per il caso continuo, e quindi abbiamo la condizione di normalizzazione. Nel caso continuo definisco il valore di aspettazione μ e la varianza σ^2 come $\int_{X_{\min}}^{X_{\max}} (x - \mu)^2 P(x) dx$

Se nel mio intervallo le misure sono *equiprobabili* allora $P(x)$ è distribuzione uniforme:

$$P(x) = \begin{cases} c, & x \in [X_{\min}, X_{\max}] \\ 0, & \text{altrimenti} \end{cases}$$

Se considero la condizione di normalizzazione, ottengo che:

- $c = \frac{1}{X_{\max} - X_{\min}}$
- da cui $\mu = \int_{-\Delta}^{+\Delta} x \frac{1}{2\Delta} dx = 0$
- e $\sigma = \sqrt{\int_{-\Delta}^{+\Delta} (x-0)^2 \frac{1}{2\Delta} dx} = \sqrt{\frac{\Delta^2}{3}}$

5.3 Propagazione degli errori

Supponiamo di avere una funzione $f(x, y)$ con x e y variabili casuali indipendenti, la varianza di f è:

$$\text{var}(f) = \left(\frac{\partial f}{\partial x} \right)^2 \text{var}(x) + \left(\frac{\partial f}{\partial y} \right)^2 \text{var}(y) = \sigma_f^2$$

Example 3. $s = vt + \frac{1}{2}at^2$

$$\sigma_s^2 = \left(\frac{\partial s}{\partial v} \right)^2 \text{var}(v) + \left(\frac{\partial s}{\partial t} \right)^2 \text{var}(t) + \left(\frac{\partial s}{\partial a} \right)^2 \text{var}(a) = t^2 \sigma_v^2 + \frac{1}{4} t^4 \sigma_a^2$$