

Appunti di Laboratorio di Meccanica

Mattia Miotto, Lorenzo Monacelli

19 maggio 2012

Indice

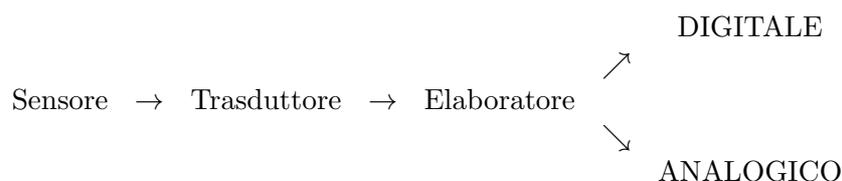
| | | |
|----------|---|-----------|
| 1 | Strumenti di misura | 2 |
| 1.1 | Sensibilità di uno strumento | 2 |
| 1.2 | Misure di lunghezza | 3 |
| 1.2.1 | Il nonio | 4 |
| 1.3 | Indeterminazione delle misure | 5 |
| 1.3.1 | Indeterminazione delle misure indirette | 6 |
| 2 | Calcolo delle probabilità | 8 |
| 2.1 | Nozioni fondamentali | 8 |
| 2.2 | Un pó di formule! | 9 |
| 2.2.1 | Gli assiomi fondamentali | 10 |
| 3 | Funzioni di distribuzione di probabilità | 11 |
| 3.1 | La funzione di distribuzione binomiale o di Bernoulli | 12 |
| 3.2 | La funzione di Poisson | 13 |
| 4 | Funzioni di densità di probabilità | 16 |
| 4.1 | La funzione di Gauss! | 17 |
| 4.2 | Valore aspettato | 19 |
| 4.3 | Scarto | 19 |
| 4.4 | Varianza | 20 |
| 5 | Funzioni a due variabili | 21 |
| 6 | Minimi Quadrati | 23 |

Capitolo 1

Strumenti di misura

Uno strumento di misura è un apparato che permette di descrivere la realtà in modo quantitativo.

È formato di solito da un sensore¹, un trasduttore², un elaboratore³ e un'uscita che permette di leggere il segnale.



Ogni strumento di misurazione ha delle caratteristiche, come la *Prontezza*⁴ o la *Portata*⁵.

Precisione e Sensibilità sono due caratteristiche molto importanti dello strumento, che hanno un significato preciso, non sempre uno strumento con un'alta precisione e alta sensibilità è lo strumento adatto per eseguire una misura

1.1 Sensibilità di uno strumento

Se G è una grandezza da misurare e \mathcal{R} è la risposta dello strumento, \mathcal{R} è una funzione di G :

$$\mathcal{R} = f(G, \lambda_i)$$

La risposta dipende infatti anche da un certo numero di parametri λ_i che variano in modo incontrollabile, quindi anche se G fosse sempre costante, otterremo sempre diverse risposte.

¹Un oggetto che subisce una modificazione da parte della misura che si vuole ottenere

²Legge le modifiche apportate al sensore e trasforma l'informazione in un segnale elettrico

³Trasforma il segnale elettrico in uno digitale o analogico

⁴Tempo di risposta dello strumento

⁵Massima grandezza misurabile dallo strumento

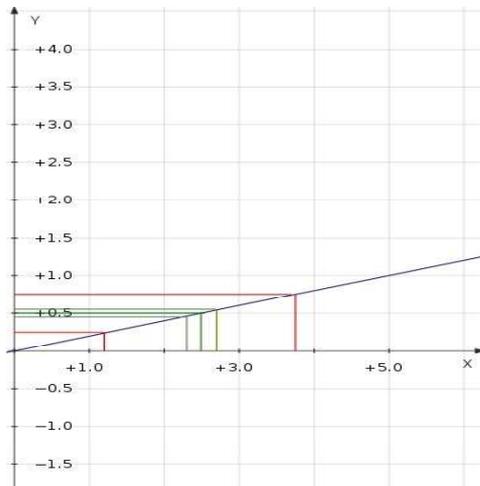
Definiamo la **Sensibilità** di uno strumento la variazione della risposta rispetto alla variazione della grandezza:

$$\text{Sensibilità} = \frac{df(G, \lambda_i)}{dG}$$

Vediamo di studiare cosa succede con uno strumento che ha molta sensibilità e uno che ne ha poca.

Ipotizziamo per semplicità che la sensibilità sia una costante positiva:

Figura 1.1: Bassa sensibilità, le linee verdi indicano l'oscillazione della misura (asse x) le linee rosse indicano l'oscillazione della misura dovuta ai fattori λ_i (asse y)



Come si può osservare dalla Figura 1.1 uno strumento con sensibilità bassa ha una misura che varia poco, e quindi maggiore precisione di uno strumento con sensibilità alta.

1.2 Misure di lunghezza

Una misura di lunghezza può essere ottenuta per comparazione con il metro, definito come la decimiliardesima parte di quarto di meridiano terrestre. Quando effettuiamo una misura essa cadrà sempre tra un millimetro e il

successivo, per questo quando andiamo a indicare una misura effettuata con un metro la indicheremo così:

$$G = (7.45 \pm 0.05) \text{ cm}$$

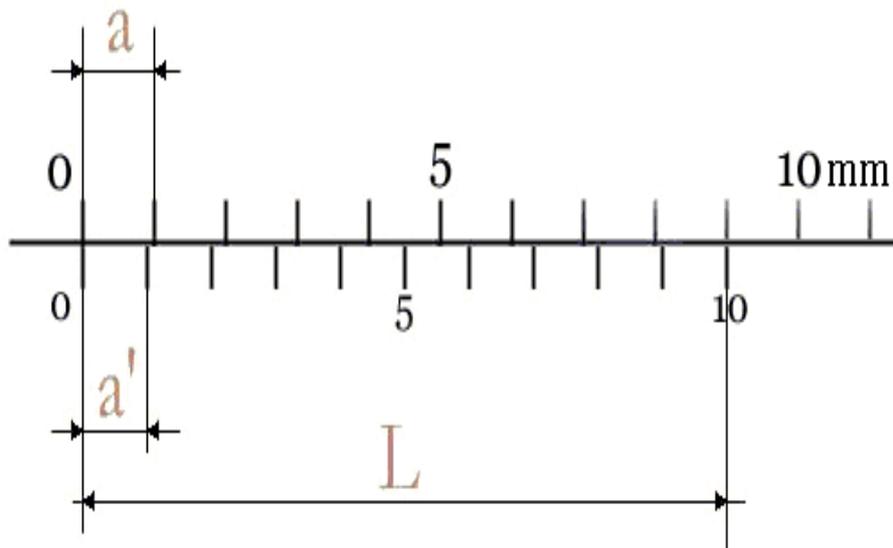
Abbiamo misurato una misura il cui valore è compreso tra 3 e 4 millimetri. Il $\Delta G = \pm 0.05$ indica la precisione dello strumento di misura.

È possibile però aumentare ulteriormente la precisione della misura fatta usando alcuni strumenti come il *nonio* o il *calibro*.

1.2.1 Il nonio

Prendiamo la scala principale che divide un centimetro in 10 millimetri. Il nonio decimale è uno strumento che divide un centimetro in $n + 1$ parti uguali (vedi Figura 1.2).

Figura 1.2: Il Nonio ($L = 10\text{cm}$, $a = 1\text{mm}$ $a_1 = \frac{1}{11}\text{cm}$)



$$L = n \cdot a = n \cdot a_1$$

$$a_1 = \left(1 - \frac{1}{n}\right) a$$

Per effettuare una misura usando il nonio basta affiancarlo vicino al metro posizionandolo *dopo* l'oggetto da misurare, e contando dopo quante tacche troviamo una tacca del nonio coincidere con una del metro (nella formula s tacche).

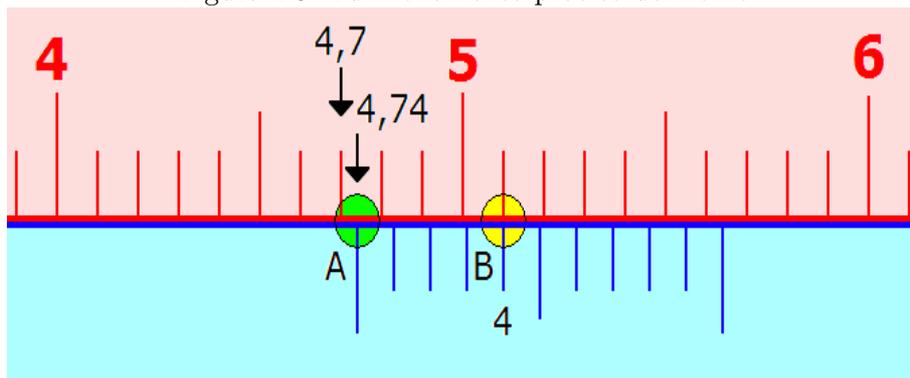
$$L = n \cdot a + s \cdot a - s \cdot a_1 = n \cdot a + s \cdot a - \left(1 - \frac{1}{n}\right) \cdot a_1$$

Quindi:

$$L = na + \frac{s}{n}a = a \left(n + \frac{s}{n} \right)$$

La Figura 1.3 mostra il funzionamento.

Figura 1.3: Funzionamento pratico del nonio



Oltre il nonio possono essere usati altri strumenti per migliorare ulteriormente la precisione della misura effettuata, come il calibro.

1.3 Indeterminazione delle misure

Se una misura diretta viene effettuata con uno strumento poco sensibile ci accorgeremo che la misura sarà la stessa ogni volta che la eseguiamo, man mano che aumenta la precisione dello strumento ci si accorge che le misure tendono a variare e a oscillare continuamente.

Abbiamo tre casi differenti. Il caso in cui la sensibilità dello strumento è talmente bassa da non percepire le oscillazioni della grandezza da misurare, ci troveremo quindi a misurare sempre la stessa grandezza.

$$G \pm \Delta G$$

Dove ΔG è proprio la precisione dello strumento, G la grandezza misurata.

Vice versa, se il nostro strumento è abbastanza accurato da percepire le fluttuazioni della grandezza dovremo comportarci in maniera diversa.

Una volta effettuate N misure da essere soddisfatti dobbiamo scegliere G e ΔG nel modo seguente:

$$G = \frac{1}{n} \sum_{i=1}^N G_i \quad \Delta G = \sqrt{\frac{\sum_{i=1}^N (G_i - G)^2}{N \cdot (N - 1)}}$$

Dove G risulta essere il valor medio di tutte le misure eseguite. Questo risultato fu ottenuto per la prima volta da Gauss.

Infine, terzo e ultimo caso, quando la precisione dello strumento non è molto elevata, o abbiamo poche misure. In questo caso non possiamo affidarci come nel caso precedente a raffinate teorie statistiche, ma solo al buon senso, e quindi indicare come errore assoluto la differenza tra la massima e la minima misura ottenuta:

$$G = \frac{1}{n} \sum_{i=1}^N G_i \quad \Delta G = \frac{G_{\max} - G_{\min}}{2}$$

Chiaramente il caso più interessante è quello in cui le misure siano molte e sufficientemente precise da consentirci di calcolare la varianza della misura con la formula di Gauss.

Si nota subito che in questo caso il valore ΔG tende a zero per N molto grande. Si potrebbe quindi pensare di ottenere una misura infinitamente precisa prendendo un infinito numero di misure, infatti analiticamente vale la seguente espressione:

$$\lim_{N \rightarrow \infty} \sqrt{\frac{\sum_{i=1}^N (G_i - \bar{G})^2}{N \cdot (N - 1)}} = 0$$

In realtà questa è una astrazione, perché la varianza non può comunque mai scendere sotto l'incertezza dello strumento. Questo significa che c'è sempre un limite sotto il quale ΔG non può scendere.

1.3.1 Indeterminazione delle misure indirette

Abbiamo visto come calcolare nei vari casi l'indeterminazione ΔG delle misure dirette, ma come ottenere lo stesso risultato per misure indirette⁶.

Ovvero immaginiamo di dover calcolare il volume di un cilindro, e abbiamo misurato il raggio e la sua altezza.

$$\begin{cases} r \pm \Delta r \\ h \pm \Delta h \end{cases}$$

Sapendo che il volume del cilindro vale $V = \pi r^2 h$, come possiamo stimare ΔV ?

stimiamo il valore:

$$V + \Delta V = \pi[r^2 h + \Delta r^2 h + 2r \Delta r h + r^2 \Delta h + \Delta r^2 \Delta h + 2r \Delta r \Delta h]$$

Poiché $\Delta r \ll r$ e $\Delta h \ll h$ eliminiamo tutti i valori in cui abbiamo due Δ qualcosa moltiplicati, perché trascurabili.

$$V + \Delta V = \pi r^2 h + 2\pi r h \Delta r + \pi r^2 \Delta h$$

⁶Le misure indirette possono essere ottenute tramite l'impiego di funzioni geometriche o fisiche, come il calcolo della velocità o di un volume

Il primo termine è proprio il volume, quindi semplifichiamo

$$\Delta V = 2\pi r h \Delta r + \pi r^2 \Delta h$$

Si può notare come questa formula sia molto simile a quella del differenziale totale della funzione volume. Sia infatti definita la funzione $V(r, h) = \pi r^2 h$ il differenziale totale dv è uguale a:

$$dv = \frac{\partial V}{\partial r} dr + \frac{\partial V}{\partial h} dh$$

In particolare passiamo dagli infinitesimi ai valori finiti otteniamo:

$$\Delta v = \frac{\partial V}{\partial r} \Delta r + \frac{\partial V}{\partial h} \Delta h$$

È facile verificare che le i valori che appaiono nella formula di prima sono proprio le derivate parziali fatte rispetto al raggio e all'altezza del volume.

In linea generale, se $G(a, b)$ è una misura indiretta il valore dell'indeterminazione ΔG può essere ottenuto a partire dall'indeterminazione delle misure a, b attraverso l'impiego del *differenziale totale* della funzione.

$$\begin{aligned} & a \pm \Delta a \\ & b \pm \Delta b \\ & G(a, b) \pm \Delta G \\ \Delta G &= \left| \frac{\partial G}{\partial a} \right| \Delta a + \left| \frac{\partial G}{\partial b} \right| \Delta b \end{aligned} \quad (1.1)$$

I moduli servono a descrivere che il contributo di ogni variazione non può che essere positivo all'aumento del valore ΔG .

Indeterminazione delle misure indirette gaussiane

Il caso visto prima è evidente se noi stiamo usando misure costanti. Ma come comportarci nel caso in cui le misure del raggio e dell'altezza del cilindro sono fluttuanti?

Per fortuna anche in questo caso esiste un potente mezzo statistico che consente di calcolare velocemente con una formula quanto vale l'indeterminazione totale della misura. Il valore principale G varrà come la funzione calcolata nei punti medi:

$$G = G(\bar{a}, \bar{b})$$

e l'indeterminazione sarà qualcosa di molto simile alla misura dell'invarianza.

$$\begin{cases} \sigma_a = \sqrt{\frac{\sum (a_i - \bar{a})^2}{N(N-1)}} \\ \sigma_b = \sqrt{\frac{\sum (b_i - \bar{b})^2}{N(N-1)}} \end{cases}$$

$$\Delta G = \sqrt{\left(\frac{\partial G}{\partial a} \bar{a} \right)^2 \cdot \sigma_a^2 + \left(\frac{\partial G}{\partial b} \bar{b} \right)^2 \cdot \sigma_b^2} \quad (1.2)$$

Capitolo 2

Calcolo delle probabilità

2.1 Nozioni fondamentali

Per iniziare, definiamo evento, tutto ciò che accade e classifichiamo gli eventi in tre tipologie:

- Eventi **certi**
- Eventi **impossibili**
- Eventi **aleatori** o **casuali**

Per ovvi motivi, il nostro interesse si concentrerà sugli eventi aleatori, in quanto sono proprio questi, che stanno alla base dello studio della probabilità! Si definisce **universo**, o spazio dell'evento, l'insieme che ha per elementi tutti i possibili modi in cui l'evento può manifestarsi; Se consideriamo l'evento, lancio di un dado a sei facce, avremo che il nostro universo, è l'insieme $U = 1, 2, 3, 4, 5, 6$, e un evento sarà un qualsiasi sottoinsieme di U . In particolare un sottoinsieme formato da un solo elemento si chiama *evento elementare*.

Ricordando le proprietà degli insiemi:

1. La somma di due insiemi, è l'insieme che ha per elementi tutti gli elementi del primo insieme più tutti gli elementi del secondo insieme, con gli elementi comuni presi una volta sola.

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

2. Il prodotto di due insiemi è l'insieme di tutti gli elementi che appartengono sia al primo che al secondo insieme.

$$(A * B) = (A \cap B)$$

3. La differenza fra due insiemi é l'insieme formato da tutti gli elementi del primo insieme meno gli elementi in comune con il secondo insieme.

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

Consideriamo, per esempio, l'universo dato da un mazzo di carte da poker, rappresentato nella tabella. Il mazzo di carte puó essere considerato come combinazione dei 4 semi, $\{\clubsuit, \diamond, \heartsuit, \spadesuit\}$, con le 13 carte, $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K\}$.

| | | | | | | | | | | | | | |
|--------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|----------------|----------------|----------------|
| \clubsuit | $\clubsuit 1$ | $\clubsuit 2$ | $\clubsuit 3$ | $\clubsuit 4$ | $\clubsuit 5$ | $\clubsuit 6$ | $\clubsuit 7$ | $\clubsuit 8$ | $\clubsuit 9$ | $\clubsuit 10$ | $\clubsuit J$ | $\clubsuit Q$ | $\clubsuit K$ |
| \diamond | $\diamond 1$ | $\diamond 2$ | $\diamond 3$ | $\diamond 4$ | $\diamond 5$ | $\diamond 6$ | $\diamond 7$ | $\diamond 8$ | $\diamond 9$ | $\diamond 10$ | $\diamond J$ | $\diamond Q$ | $\diamond K$ |
| \heartsuit | $\heartsuit 1$ | $\heartsuit 2$ | $\heartsuit 3$ | $\heartsuit 4$ | $\heartsuit 5$ | $\heartsuit 6$ | $\heartsuit 7$ | $\heartsuit 8$ | $\heartsuit 9$ | $\heartsuit 10$ | $\heartsuit J$ | $\heartsuit Q$ | $\heartsuit K$ |
| \spadesuit | $\spadesuit 1$ | $\spadesuit 2$ | $\spadesuit 3$ | $\spadesuit 4$ | $\spadesuit 5$ | $\spadesuit 6$ | $\spadesuit 7$ | $\spadesuit 8$ | $\spadesuit 9$ | $\spadesuit 10$ | $\spadesuit J$ | $\spadesuit Q$ | $\spadesuit K$ |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | J | Q | k |

Tabella 2.1: Esempio di universo, le 52 carte del poker!

Prendiamo ora i due insiemi, $A = \{K\}$ e $B = \{\clubsuit\}$, osserviamo che:

- $(A + B) = 16$ elementi
- $(A * B) = 1$ elemento
- $(A + \bar{B}) = 39$ elementi (dove \bar{B} indica il complementare di B, in questo caso $\bar{B} = \{\diamond, \spadesuit, \heartsuit\}$)

2.2 Un pó di formule!

Vediamo ora come possiamo rappresentare matematicamente il concetto di probabilitá. Definiamo:

1. La probabilitá p come il numero di eventi favorevoli, n , diviso il numero di eventi equiprobabili N , ottenendo come risultato un numero nell'intervallo $[0, 1]$, in particolare p vale 0 se l'evento é impossibile, 1 se l'evento é certo.

$$p = \frac{n}{N}$$

2. La frequenza f come il rapporto fra il numero di successi, m , ed il numero totale di esperienze fatte, M ; il vantaggio di ricorrere alla frequenza risiede nel non dover considerare solo gli eventi equiprobabili.

$$f = \frac{m}{M}$$

Bisogna specificare che le due formule, per quanto si presentino molto simili, assumono significati intrinseci, profondamente differenti! Infatti mentre la prima si riferisce a tutto l'universo degli eventi, la seconda é limitata al numero di esperienze fatte e solo rendendo questo numero infinitamente grande, otterremo lo stesso risultato. Si ha infatti che:

$$p = \lim_{M \rightarrow \infty} \frac{m}{M}$$

2.2.1 Gli assiomi fondamentali

Tutta la probabilità si basa su tre assiomi fondamentali:

1. Ad ogni evento posso assegnare un numero p , che rappresenta la probabilità di quell'evento.
2. Quando l'evento é formato da tutto l'universo U allora $P(U) = 1$
3. Siano A e B due sottoinsiemi incompatibili di U allora $P(A + B) = P(A) + P(B)$

Tramite questi tre assiomi si puó dimostrare che:

- La probabilità dell'insieme vuoto é nulla, $P(V) = 0$, infatti si ha che:
 $U = U + \{\} \rightarrow P(U) = P(U + V) = P(U) + P(V) = 1 \rightarrow P(V) = 0$
- Inoltre, $P(\bar{A}) = 1 - P(A)$, infatti: $U = A + \bar{A}$ e quindi,

$$P(U) = P(A) + P(\bar{A}) \rightarrow 1 = P(A) + P(\bar{A})$$

- Infine se consideriamo due eventi compatibili, abbiamo che $P(A+B) = P(A) + P(B) - P(A * B)$, infatti se $A * B = C$:

$$A+B = (A-C)+C+(B-C) \rightarrow P(A+B) = P(A-C)+P(C)+P(B-C)$$

$$\text{Ma, } A = (A - C) + C \rightarrow P(A) = P(A - C) + P(C) \rightarrow P(A - C) = P(A) - P(C) \text{ e } B = (B - C) + C \rightarrow P(B) = P(B - C) + P(C) \rightarrow P(B - C) = P(B) - P(C)$$

Da cui, con le opportune sostituzioni si ricava:

$$P(A+B) = P(A) - P(C) + P(C) + P(B) - P(C) \rightarrow P(A+B) = P(A) + P(B) - P(C)$$

Capitolo 3

Funzioni di distribuzione di probabilità

Introduciamo il concetto di distribuzione di probabilità, ricorrendo ad un esempio, consideriamo l'evento, somma del lancio di due dadi. Se indichiamo con n , questa somma avremo $n \in [2, 12]$, in particolare: Infatti, ogni evento

| Somma lanci | Risultati possibili | Probabilità evento |
|-------------|---------------------|--------------------|
| $n = 2$ | 1,1 | $\frac{1}{36}$ |
| $n = 3$ | 1,2 2,1 | $\frac{2}{36}$ |
| $n = 4$ | 1,3 2,2 3,1 | $\frac{3}{36}$ |
| $n = 5$ | 1,4 2,3 3,2 4,1 | $\frac{4}{36}$ |
| ... | ... | ... |
| $n = 12$ | 6,6 | $\frac{1}{36}$ |

Tabella 3.1: Risultati dei due lanci

elementare ha probabilità $1/6$ di verificarsi, essendo i lanci indipendenti le probabilità si moltiplicano, e poi si sommano nel caso l'evento si possa manifestare in più modi. La funzione di distribuzione avrà dominio $I = \{x \in N | 2 < x < 12\}$, e massimo in corrispondenza del valore più alto di probabilità, come mostra il grafico 3.1.

E' evidente che se si considera la somma di tutte le probabilità, si otterrà il valore 1, infatti è certo che lanciando due volte un dado, la somma dovrà per forza essere compresa in $[2, 12]$.

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

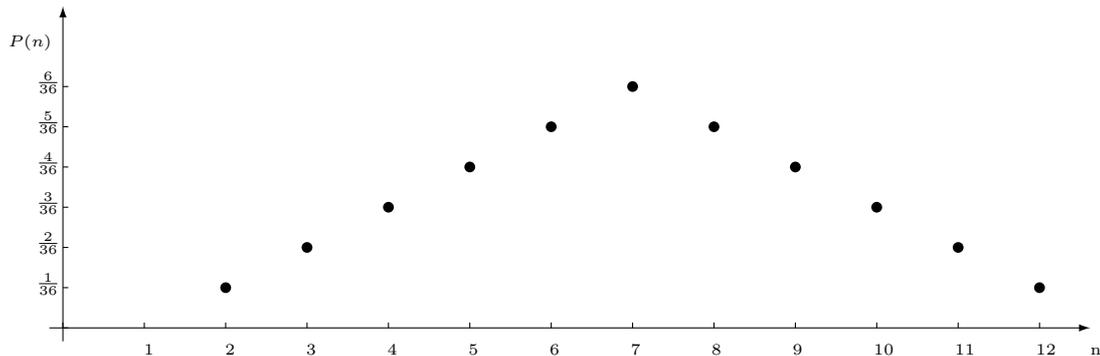


Figura 3.1: Grafico di distribuzione di probabilità

3.1 La funzione di distribuzione binomiale o di Bernoulli

Consideriamo un evento E , tale evento può accadere o può non accadere, associamo al verificarsi di E una probabilità p , e supponiamo che le condizioni del verificarsi o del non verificarsi di E , si ripetano N volte. E' lecito a questo punto chiederci, su N volte quale è la probabilità che E si verifichi n volte! Come prima cosa notiamo che se $P(E) = p$ allora $P(\bar{E}) = 1 - p$ e che se E si verifica n volte allora \bar{E} si verifica $(N - n)$ volte! Allora abbiamo:

$$\underbrace{E \times \dots \times E}_{n \text{ volte}} \times \underbrace{\bar{E} \times \dots \times \bar{E}}_{N-n \text{ volte}} \implies P(n) = p^n \cdot (1 - p)^{N-n}$$

Bene armati della formula che abbiamo ricavato, proviamo a verificarla in un caso semplice.

Il nostro test di verifica, sarà quello di lanciare due volte un dado e vedere con quale probabilità otterrò almeno una volta 1! Nella tabella 3.2,

| Eventi favorevoli | \implies | Probabilità |
|-------------------|---|-----------------|
| n=0 | $\frac{5}{6} * \frac{5}{6}$ | $\frac{25}{36}$ |
| n=1 | $\frac{1}{6} * \frac{1}{6} + \frac{5}{6} * \frac{5}{6}$ | $\frac{10}{36}$ |
| n=2 | $\frac{1}{6} * \frac{1}{6}$ | $\frac{1}{36}$ |

Tabella 3.2: Griglia dei risultati

si sono ricavate le varie probabilità, ricorrendo agli assiomi fondamentali, applicando invece la formula con $p = 1/6$, otteniamo la tabella 3.1

| | | |
|-----|--|-----------------|
| n=0 | $P(0) = \frac{1}{6}^0 (1 - \frac{5}{6})^2$ | $\frac{25}{36}$ |
| n=1 | $P(1) = \frac{1}{6}^1 (1 - \frac{5}{6})^1$ | $\frac{5}{36}$ |
| n=2 | $P(2) = \frac{1}{6}^2 (1 - \frac{5}{6})^0$ | $\frac{1}{36}$ |

I risultati non coincidono se $n=1$; il motivo di questa incongruenza è subito chiaro se si osserva che nel caso $n=1$, ci sono due possibili modi in cui si può verificare l'evento favorevole, infatti il valore 1 può uscire al primo lancio oppure al secondo, cioè l'evento non è univocamente determinato, può verificarsi in più modi, la funzione deve quindi tenere presente questa eventualità, come? Attraverso l'introduzione del coefficiente binomiale! La funzione quindi assume la forma:

$$P_{N,p}(n) = \frac{N!}{(N-n)!n!} \cdot p^n (1-p)^{(N-n)}$$

Tenendo conto che:

$$\sum_{n=0}^N P_{N,p} = 1$$

3.2 La funzione di Poisson

Ora consideriamo un evento che si ripete nel tempo, è lecito chiedersi in un intervallo di tempo, quale è la probabilità che l'evento si verifichi un numero n di volte!

Cominciamo con il dire che:

- Fissato l'intervallo che ci interessa considerare, tutti gli eventi precedenti o successivi a quell'arco di tempo non sono collegati, cioè sono eventi indipendenti
- $P_1(\Delta t) = \lambda \Delta t$ ossia la probabilità che accada una volta è proporzionale all'intervallo scelto
- $P_n(\Delta t) \xrightarrow{n \rightarrow \infty} 0$ mentre la probabilità che accadano un numero tendente all'infinito di eventi in un intervallo piccolo di tempo è infinitesima

Armati di queste ipotesi, cerchiamo di ricavare una formula generale, per induzione.

Partiamo dalla probabilità che non si verifichino eventi nell'intervallo $t + \Delta t$, avremo che:

$$P_0(t+\Delta t) = P_0(t) \cdot P_0(\Delta t) = P_0(t) \cdot (1 - P_1(\Delta t)) \implies P_0(t+\Delta t) - P_0(t) = -\lambda P_0(t) \Delta t$$

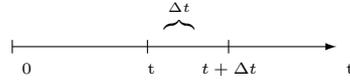


Figura 3.2: Linea del tempo

Ora notiamo che portando Δt al primo membro ottenamo un rapporto incrementale in funzione del tempo:

$$\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t)$$

il risultato è l'equazione differenziale del primo ordine omogenea $P_0'(t) + \lambda P_0(t) = 0$

Ricordando che:

$$y' + ay = 0 \rightarrow \frac{y'}{y} = -a \rightarrow \int \frac{y'}{y} dy = \int -a dt \rightarrow \ln y = -at \rightarrow y = Ae^{-at}$$

Risolvendo la differenziale abbiamo $P_0(t) = e^{-\lambda t}$, poichè il termine generico A , deve valere necessariamente 1.¹ Ora vediamo la probabilità di un evento e per semplicità consideriamo sempre l'intervallo $t + \Delta t$, e ricordiamo che l'evento può verificarsi solo in uno dei due sottointervalli(vedi 3.2), abbiamo,:

$$P_1(t+\Delta t) = P_0(t)*P_1(\Delta t)+P_1(t)*P_0(\Delta t) \rightarrow P_1(t+\Delta t) = e^{-\lambda t}*\lambda\Delta t+P_1(t)(1-\lambda\Delta t)$$

Con le opportune stituzioe fatte sfruttando lipotesi iniziale e quantoo ricavato finora. Svolgendo i pasaggi come nel caso $n = 0$ otteniamo un'altra equazione differenziale, questa volta non omogenea: $P_1'(t) + \lambda P_1(t) = \lambda e^{-\lambda t}$ la cui primitiva è $P_1(t) = \lambda t e^{-\lambda t}$, ora bisognerebbe ripetere le operazioni per $n=3,4,5\dots n$ ma osservare i primi due casi ci porta già a poter generalizzare la formula:

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \text{ ponendo } \lambda t = a \text{ otteniamo } P_n(t) = \frac{a^n}{n!} e^{-a}$$

Tenendo presente che si deve avere:

$$\sum_{n=0}^{\infty} \frac{a^n}{n!} e^{-a} = 1 \text{ ma infatti } e^{-a} \sum_{n=0}^{\infty} \frac{a^n}{n!} = e^{-a} e^a = 1$$

La funzione così trovata è un caso particolare della funzione di Bernoulli in quanto, $P_{N,p}(n) = \frac{N!}{(N-n)!n!} * p^n(1-p)^{(N-n)}$ se:

¹Il valore di A si ottiene se si considera il caso $t = 0$, in quel caso la probabilità che accadano 0 eventi in un tempo anch'esso nullo è pari a 1 ossia alla certezza e quindi $1 = Ae^{-\lambda*0} \rightarrow A = 1$

- $N \rightarrow \infty$
- $p \rightarrow 0$
- $Np = \text{cost}$

Allora si avrà:

$$\frac{N!}{(N-n)!n!} = \frac{1 * 2 * \dots * N}{(1 * 2 * \dots * (N-n))n!} \rightarrow \frac{(N-n) * \dots * N}{n!} \approx \frac{N^n}{n!}$$

E quindi:

$$\frac{N^n}{n!} p^n (1-p)^{N-n} = \frac{N^n}{n!} * \frac{p^n (1-p)^N}{(1-p)^n} \approx \frac{(Np)^n (1-p)^N}{n!}$$

Basta porre $Np = a \rightarrow N = \frac{a}{p}$ e avremo:

$$P_{n,p}(t) = \frac{a^n}{n!} (1-p)^{\frac{a}{p}}$$

Ricordando il limite notevole:

$$\lim_{p \rightarrow 0} (1-p)^{\frac{1}{p}} = e^{-1}$$

Otteniamo nuovamente $P_n(t) = \frac{a^n}{n!} e^{-a}$

² $n! * (1-p)^n \approx n!$

Capitolo 4

Funzioni di densità di probabilità

Finora si sono prese in esame funzioni di distribuzione discrete, ossia definite per valori interi positivi, ora è lecito chiedersi cosa accade se si considerano funzioni definite in tutto R , o comunque in un intervallo $(a, b) \in R$. Per prima cosa, notiamo che $P(x = x_0) = 0$, infatti poiché la variabile può assumere un valore con moltissime cifre decimali (anche infinite), la probabilità che si verifichi quel singolo caso è praticamente nulla, più senso ha chiedersi quanto vale la probabilità in un intervallo, esempio $P(x \in (a, b) | X_0 < x < X_1) = f(x)dx$, in definitiva passare da una variabile discreta ad una continua significa scrivere una funzione di *densità* di probabilità.

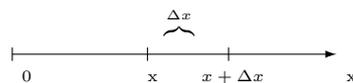


Figura 4.1: Asse degli eventi

Chiamo $F(x) = P(X < x)$ allora si avrà $F(x + dx) - F(x) = F(dx) = f(x)dx$, ma moltiplicando i membri per $\frac{1}{dx}$, ottengo:

$$\frac{F(x + dx) - F(x)}{dx} = f(x) \rightarrow \frac{dF(x)}{dx} = f(x)$$

Quindi si ottiene la funzione di densità cumulativa $F(x) = \int_{-\infty}^x f(x)dx$. Ora sfruttiamo il risultato ottenuto, nel caso semplice $f(x) = \text{cost}$;

In questo caso è molto semplice determinare il valore della probabilità, in quanto sappiamo che l'evento certo vale 1, e quindi:

$$\int_a^b f(x)dx = k(b - a) = 1 \rightarrow k = \frac{1}{(b - a)}$$

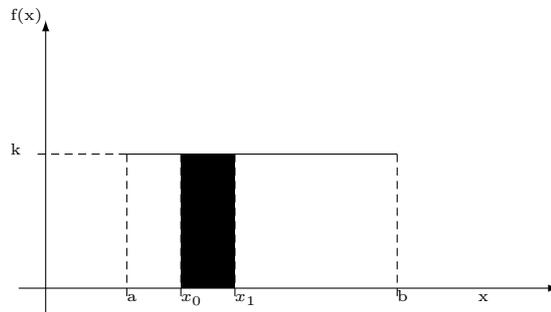


Figura 4.2: Grafico di una distribuzione equiprobabile

Conseguentemente la probabilità in un intervallo (x_0, x_1) è $P = k(x_1 - x_0)$. Mentre per una funzione non costante sarà data dall'integrale della funzione nel dato intervallo.

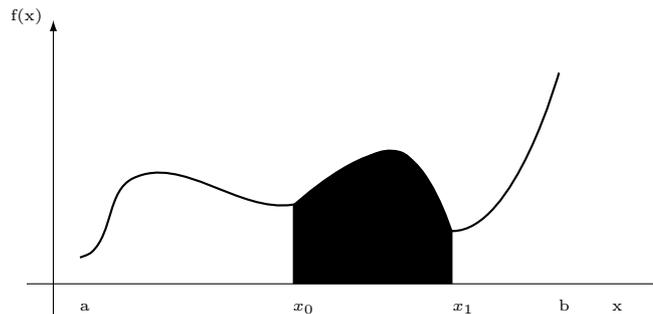


Figura 4.3: La probabilità di f non costante

4.1 La funzione di Gauss!

La funzione di distribuzione elaborata da Gauss:

$$f(x) = Ae^{-h^2(x-m)^2}$$

È definita nell'intervallo $(-\infty, +\infty)$, e dipende da tre parametri, A, h, m. Cerchiamo di definirli effettuando lo studio di funzione, per capirne in primis il significato geometrico. Poichè si deve avere $\int_{-\infty}^{+\infty} f(x)dx = 1$, il parametro A servirà a normalizzare la funzione, cioè a far sì che l'area totale valga 1; Allora avremo, applicando il cambio di variabile:

$$h(x-m) = z \rightarrow x = \frac{z}{h} + m \rightarrow dx = \frac{1}{h} dz \rightarrow \frac{A}{h} \int_{-\infty}^{+\infty} e^{-z^2} dz = \frac{A}{h} \sqrt{\pi} = 1 \rightarrow A = \frac{h}{\sqrt{\pi}}$$

Ora facendo i limiti destro e sinistro abbiamo che:

$$\lim_{x \rightarrow \pm\infty} f(x) = 0$$

Quindi la funzione tende a zero, troviamo i massimi e gli eventuali punto di flesso, facendo le derivate prima e seconda e ponendole uguali a zero...

$$f(x) = Ae^{-h^2(x-m)^2}$$

$$f'(x) = Ae^{-h^2(x-m)^2} \cdot [-2h^2(x-m)] = 0 \rightarrow x = m \text{ è un punto di massimo}$$

$$f''(x) = -2Ah^2(e^{-h^2(x-m)^2} + e^{-h^2(x-m)^2} \cdot -2h^2(x-m)^2) \text{ in } x = m \rightarrow -2Ah^2 < 0$$

Quindi m è un massimo, ora ponendo la derivata seconda uguale a zero otteniamo:

$$-2h^2(x-m)^2 + 1 = 0 \rightarrow (x-m)^2 = \frac{1}{2h^2} \rightarrow x = m \pm \frac{1}{\sqrt{2}h} \text{ sono i due punti di flesso}$$

Ricapitolando i tre parametri identificano:

- m il massimo della curva
- h i valori dei flessi
- A serve a normalizzare l'area

Otterremo quindi il grafico, detto a campana, della funzione di Gauss:

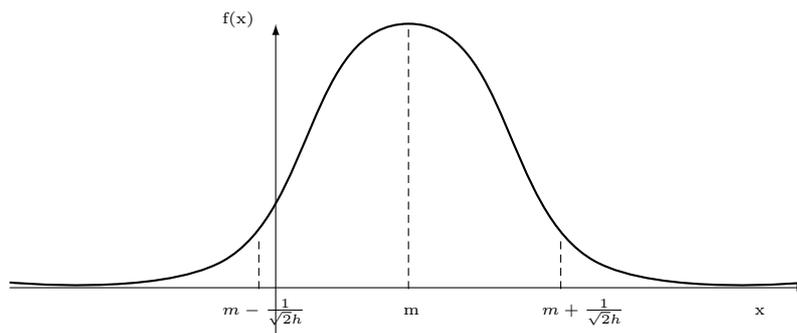


Figura 4.4: Curva di Gauss

Ora per comodità definiamo $\frac{1}{\sqrt{2}h} = \sigma$ e abbiamo la funzione di Gauss nella forma definitiva:

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

Ora, sapendo che l'area sottesa da tutta la curva vale uno, chiediamoci quanto vale l'area e quindi la probabilità in un intervallo generico, e consideriamo l'intervallo centrato nel punto medio e di ampiezza 2σ , cioè quello individuato dalle ascisse dei punti di flesso.

Bisogna quindi calcolare:

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

Operiamo una sostituzione e poniamo $\frac{x-m}{\sigma} = z$, e quindi:

$$x = \sigma z + m \rightarrow dx = \sigma dz$$

4.2 Valore atteso

Data una funzione di distribuzione, si definisce valore atteso:

- Se la variabile è continua,

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

- Se la funzione è definita per valori discreti,

$$E[x] = \sum_1^N n P_{N,p}(n)$$

4.3 Scarto

Lo scarto invece, è definito come la differenza della variabile con il valore atteso, $\xi = x - E[x]$ (oppure $\xi = n - E[n]$). A questo punto ci si può chiedere quanto vale il valore atteso dello scarto! Abbiamo allora che:

$$E[\xi] = \int_{-\infty}^{+\infty} (x - E[x]) f(x) dx = \int_{-\infty}^{+\infty} x f(x) dx - \int_{-\infty}^{+\infty} E[x] f(x) dx = E[x] \cdot (1 - \int_{-\infty}^{+\infty} f(x) dx) = 0^1$$

Quindi per qualsiasi funzione di distribuzione, si ha che il valore atteso dello scarto è nullo!

¹infatti $\int_{-\infty}^{+\infty} f(x) dx = 1$ in quanto è l'evento certo.

4.4 Varianza

La varianza è il valore aspettato dello scarto al quadrato:

$$Var[x] = E[\xi^2] = E[(x - E[x])^2]$$

Sviluppando:

$$\int_{-\infty}^{+\infty} (x^2 + E[x]^2 - 2xE[x])f(x)dx = \int_{-\infty}^{+\infty} x^2 f(x)dx + E[x]^2 \int_{-\infty}^{+\infty} f(x)dx - 2E[x] \int_{-\infty}^{+\infty} xf(x)dx = E[x^2] + E[x]^2 - 2E[x]^2 = E[x^2] - E[x]^2$$

Inoltre vale la proprietà che:

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n \frac{m}{n} (x_i - \bar{x})^2 = \sum_{i=1}^n f(x_i - \bar{x}) = \sum_{i=1}^n p(x_i - \bar{x}) = Var[x]$$

Basta che il numero di dati sia abbastanza grande! Quindi se $n \rightarrow \infty$

$$Var[n] = \frac{1}{n} \sum_1^n (n - \bar{n}) = S^2$$

Capitolo 5

Funzioni a due variabili

Finora si sono trattate funzioni di distribuzione a una sola variabile, si possono tuttavia considerare funzioni, a più variabili, rappresentanti eventi dipendenti da più fattori casuali.

$$P(x < X < x+dx, y < Y < y+dy, z < Z < z+dz...) = f(x, y, z...)dxdydz...$$

Nel seguito si tratteranno solo funzioni a due variabili, ma tutto quello che verrà detto si può adattare ad un numero n di variabili aleatorie.

Le variabili possono essere:

- dipendenti
- indipendenti
- statisticamente correlate

Per quanto riguarda le variabili dipendenti, è sufficiente studiarne una e poi tramite la relazione che le lega ricavare informazioni sull'altra; quelle statisticamente correlate, verranno trattate in corsi più avanzati; la nostra attenzione si rivolge allo studio delle variabili indipendenti, ossia a quelle, le cui funzioni di distribuzione, possono essere fattorizzate:

$$f(x, y) = h(x) \cdot g(y)$$

Ora la prima cosa da fare è vedere come si comportano il valore atteso e la varianza. Definiamo $z = x + y$, e vediamo il valore atteso $E[z]$, si avrà che:

$$\begin{aligned} E[z] = E[x + y] &= \int_{-\infty}^{+\infty} (x + y) f(x, y) dx dy = \int_{-\infty}^{+\infty} (x + y) h(x) g(y) dx dy = \\ &= \int_{-\infty}^{+\infty} x \cdot h(x) g(y) dx dy + \int_{-\infty}^{+\infty} y \cdot h(x) g(y) dx dy = \\ &= \int_{-\infty}^{+\infty} x \cdot h(x) dx \cdot \int_{-\infty}^{+\infty} y \cdot g(y) dy \rightarrow E[x + y] = E[x] + E[y]^1 \end{aligned}$$

¹Poichè $\int_{-\infty}^{+\infty} x f(x) dx = E[x]$ e $\int_{-\infty}^{+\infty} f(x) dx = 1$

Per la varianza abbiamo che:

$$Var[z] = Var[x + y] = E[z^2] - E[z]^2$$

Il calcolo della varianza si riduce così al calcolo del valore atteso di z^2 :

$$\begin{aligned} E[z^2] &= E[(x + y)^2] = \int_{-\infty}^{+\infty} (x^2 + 2yx + y^2)h(x)g(y)dxdy = \\ &= \int_{-\infty}^{+\infty} x^2 \cdot h(x)g(y)dxdy + \int_{-\infty}^{+\infty} y^2 \cdot h(x)g(y)dxdy + 2 \int_{-\infty}^{+\infty} xy \cdot h(x)g(y)dxdy = \\ &= \int_{-\infty}^{+\infty} x^2 \cdot h(x)dx \cdot \int_{-\infty}^{+\infty} g(y)dy + \int_{-\infty}^{+\infty} y^2 g(y)dy \cdot \int_{-\infty}^{+\infty} h(x)dx + \\ &+ 2 \int_{-\infty}^{+\infty} y \cdot g(y)dy \cdot \int_{-\infty}^{+\infty} x \cdot g(x)dx \rightarrow E[(x + y)^2] = E[x^2] + 2E[x]E[y] + E[y] \end{aligned}$$

Allora mettendo tutto insieme, abbiamo che:

$$\begin{aligned} Var[x + y] &= E[x^2] + E[y^2] + 2E[x]E[y] - E[x + y]^2 = \\ &= Var[x] + Var[y] + E[x]^2 + E[y]^2 + 2E[x]E[y] - E[x]^2 - E[y]^2 - 2E[x]E[y] = Var[x] + Var[y] \end{aligned}$$

Sia il valore atteso che la varianza sono lineari! Questo risultato è molto importante, poichè se x_i è una variabile casuale posso definire:

$$z = \frac{1}{n} \sum_{i=1}^n x_i \rightarrow E[z] = \frac{1}{n} E\left[\sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n E[x_i]$$

E allo stesso modo $Var[z] = \frac{1}{n^2} \sum_{i=1}^n Var[x_i]$. Ora supponiamo che di avere più variabili con lo stesso valore atteso e la stessa varianza, allora:

- $E[z] = \frac{1}{n} \sum_{i=1}^n E[x_i] \rightarrow E[x]$
- $Var[z] = \frac{1}{n^2} \sum_{i=1}^n Var[x_i] \rightarrow \frac{1}{n} Var[x]$

Questo dimostra che è inutile aumentare il numero di misure n , non appena l'incertezza diviene minore della sensibilità dello strumento!

²Si rimanda agli appunti sulla varianza

Capitolo 6

Minimi Quadrati

Il metodo dei minimi quadrati è un sistema di interpolazione grafica fondato su solide basi statistiche, che consente di ricavare i parametri che meglio approssimano un rapporto funzionale tra le misure fatte dei campioni statistici.

Tuttavia, perché questo sistema possa essere applicabile occorre soddisfare due condizioni:

1. L'errore relativo sulla coordinata x deve essere nullo:

$$\Delta x \ll x$$

2. La funzione di distribuzione delle y fissato x deve essere una gaussiana:

$$f(y) = \text{Gauss}$$

Per trovare la funzione $f(x)$ che meglio approssima l'esperimento usiamo il criterio di massima verosimiglianza:

Supponendo di conoscere a priori la forma della $f(x)$ vogliamo chiederci quali sono i parametri λ_j che descrivono la funzione $f(x_i/\lambda_j)$ che meglio approssimano i dati sperimentali.

Scriviamo quindi la probabilità di ripetere un'esperimento e per ogni x_i trovare un valore $y_i = f(x_i/\lambda_j)$. Questa probabilità è pari al prodotto di tutte le probabilità di ogni x_i , perché si suppone che le variabili misurate in modo sperimentale siano tra loro indipendenti. Poiché abbiamo supposto all'inizio che le y sono variabili gaussiane questa probabilità sarà:

$$P(i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i - f(x_i/\lambda_j))^2}{2\sigma^2}} dx_i$$

Quindi la probabilità che tutto l'esperimento abbia successo è:

$$P(1) \cdot P(2) \cdot P(3) \cdots P(n)$$

Da cui ricaviamo

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_i - f(x_i/\lambda_j))^2}{2\sigma^2}} dx_i$$

Dove L rappresenta la probabilità che ripetendo l'esperimento troviamo valori di y_i consistenti con quanto predetto dalla funzione $f(x_i/\lambda_j)$ (principio di massima verosimiglianza).

Per trovare i parametri λ_j che offrono una maggiore probabilità, dobbiamo cercare i massimi di L , il che equivale a farne la derivata prima e porla uguale a zero.

Tuttavia derivare una produttoria può essere alquanto complicato, ricorriamo ad un trucchetto. la funzione logaritmo è monotona, per cui se $f(x)$ ha un massimo in x_0 , anche la funzione $\ln(f(x))$ ha un massimo nello stesso punto.

Quindi per semplicità calcoliamo i massimi della funzione $\ln(L)$.

$$\ln(L) = n \ln\left(\frac{1}{\sqrt{2\pi\sigma}}\right) - \sum_{i=1}^n \frac{(y_i - f(x_i/\lambda_j))^2}{2\sigma^2}$$

Notiamo che il primo termine è costante rispetto ai parametri λ_j , perciò si annulla nel processo di derivazione. Il denominatore della sommatoria è una costante che quindi non cambia la posizione del massimo. Per cui rimane da trovare il massimo di:

$$L = - \sum_{i=1}^n (y_i - f(x_i/\lambda_j))^2$$

Il che equivale a trovare il minimo della funzione:

$$L = \sum_{i=1}^n (y_i - f(x_i/\lambda_j))^2$$

Ecco perché questo metodo si chiama principio dei minimi quadrati, in pratica il nostro obiettivo è trovare i parametri λ_j della $f(x_i/\lambda_j)$ che minimizzino i quadrati delle distanze tra il valore misurato y_i e il valore atteso $f(x_i)$.

Passiamo ad un esempio pratico. Supponiamo di voler stimare la forza di una molla in base all'allungamento. La conoscenza della teoria ci dice che $F = kx$, per cui la nostra funzione $f(x) = kx$. Ora proviamo a stimare, dato un certo numero di misure effettuate di x e del corrispondente valore di F il coefficiente k .

$$L = \sum_i (F_i - kx_i)^2$$

Minimiziamo:

$$\frac{\partial L}{\partial k} = \frac{\partial}{\partial k} \sum_i (F_i - kx_i)^2$$

Notate che ho messo il segno della derivata parziale, per il semplice fatto che in genere i parametri usati sono più di uno¹.

$$\frac{\partial L}{\partial k} = \frac{\partial}{\partial k} \sum_i (F_i^2 + k^2 x_i^2 - 2F_i k x_i)$$

$$\frac{\partial L}{\partial k} = \sum_i (2k x_i^2 - 2F_i x_i)$$

Ok, ricordiamoci che stiamo cercando un minimo per cui $\frac{\partial L}{\partial k} = 0$. Possiamo evitare di fare la derivata seconda, in quanto è intuitivo credere che questa funzione non ha massimi (non c'è limite a quanto ci possiamo allontanare dai valori aspettati). I più pignoli procedano a fare la derivata seconda, buona fortuna.

$$\sum_i 2k x_i^2 = \sum_i 2F_i x_i$$

$$k = \frac{\sum_i F_i x_i}{\sum_i x_i^2}$$

Adesso abbiamo la migliore stima di k .

Sarebbe lecito chiedersi qual è l'incertezza da associare al valore di k , o meglio, qual è la varianza di k ?

Per rispondere a questa domanda ricordiamoci la formula di propagazione della deviazione standard:

$$\sigma_k^2 = \sum_i \left(\frac{\partial k}{\partial F_i} \right)^2 \sigma_{F_i}^2$$

Ricordiamoci la derivata di k :

$$\left(\frac{\partial k}{\partial F_i} \right)^2 \sigma_{F_i}^2 = \frac{x_i^2 \sigma_F^2}{(\sum_i x_i^2)^2}$$

Una volta fatta la somma le x_i^2 al numeratore si semplificano con il denominatore e otteniamo

$$\sigma_k^2 = \frac{\sigma_F^2}{\sum_i x_i^2}$$

¹Infatti anche una retta generica ha una funzione del tipo $f(x) = mx + q$. Rimandiamo lo studio del calcolo dei parametri m e q con i minimi quadrati al lettore come esercizio