

Enrico Pierobon

# Laboratorio di fisica 1



Introduzione all'analisi dati per il

laboratorio di fisica

Università degli studi di Trento

## Premessa

Questa trattazione si propone di essere un supporto allo studente in quanto racchiude le nozioni fondamentali svolte a lezione. La semplicità e la chiarezza vogliono essere al centro di questo documento a scapito della precisione e del formalismo matematico così da permettere al lettore di chiarire eventuali dubbi e di prepararsi al meglio per l'esame. Se questo documento ti piacerà ne sarò felice, se invece ti farà schifo ti auguro di migliorarlo visto che i sorgenti sono liberi. Buona fortuna.

# Indice

<b>I</b>	<b>Statistica</b>	<b>1</b>
<b>1</b>	<b>Analisi dati laboratorio</b>	<b>3</b>
1.1	Misura di una grandezza fisica . . . . .	3
1.1.1	Strumenti di misura . . . . .	3
1.2	Incertezza . . . . .	4
1.2.1	Classificazione incertezza di misura . . . . .	4
1.2.2	Rappresentazione dell'incertezza . . . . .	5
1.2.3	Tipi di intervallo di incertezza . . . . .	5
1.3	Come trattare gli errori . . . . .	6
1.3.1	Errore di risoluzione . . . . .	6
1.3.2	Errore sistematico o di tipo $B$ . . . . .	7
1.3.3	Errore casuale o di tipo $A$ . . . . .	7
1.4	Istogrammi . . . . .	9
1.4.1	Normalizzazione in altezza . . . . .	9
1.4.2	Normalizzazione in area . . . . .	10
1.4.3	Scelta del binning . . . . .	10
1.4.4	Parametri statistici . . . . .	10
1.4.5	Istogrammi cumulativi . . . . .	11
1.4.6	Proprietà asintotiche degli istogrammi . . . . .	11
1.5	Stima dei parametri della distribuzione limite . . . . .	12
1.5.1	Esempio: distribuzione uniforme . . . . .	13
1.5.2	Esempio: distribuzione di Gauss o normale . . . . .	14
<b>2</b>	<b>Incertezza</b>	<b>15</b>
2.1	Propagazione incertezza . . . . .	15
2.1.1	Relazione lineare . . . . .	15
2.2	Applicazioni della propagazione dell'incertezza . . . . .	18
2.2.1	Media campionaria . . . . .	18
2.2.2	Misura di un multiplo di una grandezza $x$ . . . . .	20
2.2.3	Composizione di contributi di incertezza di varia natura	20
2.3	Compatibilità tra misure . . . . .	21
2.3.1	Intervallo d'incertezza massima . . . . .	21
2.3.2	Intervallo d'incertezza tipo . . . . .	22

2.3.3	Problema dei falsi allarmi . . . . .	22
2.4	Informazione da incertezze differenti . . . . .	23
2.4.1	Media pesata . . . . .	23
2.5	Propagazione via metodo grafico . . . . .	24
2.5.1	Caso unidimensionale . . . . .	24
2.5.2	Esempi . . . . .	25
2.6	Estensione a funzioni di più grandezze . . . . .	26
2.6.1	Esempi . . . . .	26
2.6.2	Propagazione dell'incertezza su misure dirette non statisticamente indipendenti . . . . .	27
<b>3</b>	<b>Regressione lineare e FIT</b>	<b>29</b>
3.1	Metodo di minima e massima pendenza . . . . .	29
3.2	Metodo dei minimi quadrati . . . . .	30
3.2.1	Caso di proporzionalità diretta . . . . .	31
3.2.2	Riepilogo procedura fit minimi quadrati . . . . .	32
3.2.3	Valutazione della legge . . . . .	33
3.3	Test del Chi2 . . . . .	33
3.3.1	Calcolo del Chi2 dal grafico . . . . .	34
3.3.2	Test del Chi2: casi possibili . . . . .	34
3.3.3	Determinazione a posteriori delle incertezze . . . . .	36
3.4	Esempi di fit . . . . .	36
3.4.1	Legge costante . . . . .	36
<b>II</b>	<b>Probabilità</b>	<b>39</b>
<b>4</b>	<b>Introduzione ai fenomeni aleatori</b>	<b>41</b>
4.1	Fenomeni aleatori . . . . .	41
4.2	Teoria della probabilità . . . . .	42
4.3	Definizione assiomatica della probabilità . . . . .	43
4.3.1	Interpretazione della probabilità classica o a priori . . . . .	44
4.3.2	Interpretazione della probabilità frequentista o a posteriori . . . . .	44
4.3.3	Interpretazione soggettiva o Bayesiana della probabilità . . . . .	45
4.3.4	Probabilità condizionata . . . . .	46
4.3.5	Probabilità condizionata al prodotto di eventi . . . . .	49
4.3.6	Teorema della probabilità totale . . . . .	49
4.4	Teorema di Bayes e interpretazione soggettiva . . . . .	50
4.4.1	Esempio: individui sani/malati . . . . .	52
<b>5</b>	<b>Calcolo combinatorio e prove ripetute</b>	<b>55</b>
5.1	Calcolo combinatorio . . . . .	55
5.1.1	Introduzione . . . . .	55

5.1.2	Permutazioni . . . . .	56
5.1.3	Disposizioni . . . . .	56
5.1.4	Combinazioni . . . . .	57
5.2	Applicazione alle prove ripetute . . . . .	58
5.2.1	Distribuzione di probabilità binomiale o di Bernoulli .	58
5.2.2	Distribuzione di probabilità multinomiale . . . . .	59
<b>6</b>	<b>Variabili aleatorie</b>	<b>61</b>
6.1	Distribuzione aleatoria discreta . . . . .	61
6.1.1	Distribuzione binomiale . . . . .	62
6.1.2	Passeggiata a caso . . . . .	65
6.1.3	Distribuzione di Poisson . . . . .	67
6.2	Distribuzione aleatoria continua . . . . .	70
6.2.1	Densità di probabilità . . . . .	71
6.3	Parametri di caratterizzazione . . . . .	71
6.3.1	Momenti della distribuzione . . . . .	72
	<b>Indice analitico</b>	<b>77</b>



Parte I

Statistica





# Capitolo 1

## Introduzione all'analisi dati di laboratorio

### 1.1 Misura di una grandezza fisica

Per misura di una grandezza fisica si intende la procedura operativa per determinare il valore numerico che corrisponde al rapporto fra quantità della grandezza e unità di misura. È inoltre fondamentale stimare l'incertezza associata alla misura. Esistono due grandi distinzioni delle misure:

- **Misure dirette:** misure effettuate confrontando direttamente la grandezza con l'unità di misura
- **Misure indirette:** misure in relazione algebrica con le misure dirette.

#### 1.1.1 Strumenti di misura

Per strumento di misura si intende quell'oggetto che permette il confronto della grandezza con l'unità di misura. A questo oggetto si associano delle caratteristiche fondamentali tipiche di ogni strumento:

- **Risoluzione:** è definita come la più piccola variazione  $\Delta X$  che lo strumento è in grado di rilevare sulla grandezza da misurare.
- **Range:** è definito come l'intervallo massimo sul quale lo strumento è in grado di lavorare.
- **Sensibilità:** è definita come il rapporto tra

$$\frac{\Delta U}{\Delta X}$$

- **Tempo di risposta:** è definito come il tempo per cui lo strumento tende ad allinearsi con la nuova misurazione qualora si presentasse una

discontinuità. Generalmente segue un andamento esponenziale del tipo:

$$U(t) = U(0) + \Delta U(1 - e^{-\frac{t}{\tau}})$$

Dopo un tempo  $3\tau$ , dove  $\tau$  è definito come il tempo di risposta, generalmente si è attorno al 5% di differenza con il valore asintotico.

## 1.2 Incertezza

Come detto precedentemente una misura deve sempre riportare la propria incertezza<sup>1</sup>. Le cause che normalmente generano incertezza sono:

- **Cause intrinseche:**
  - Le fluttuazioni casuali sono proprietà del fenomeno che studio
  - Sto trascurando informazioni importanti sul mio modello
- **Processo di misura:**
  - *Incertezza introdotta dallo strumento di misura:* solo il fatto che lo strumento abbia una risoluzione finita comporta incertezza.
  - *Interazione operatore-strumento.*
  - *Interazione con l'ambiente dello strumento:* possibili cause ambientali come temperatura, umidità, pressione atmosferica.
  - *Scelta della metodologia di misura:* a parità di strumento-operatore-ambiente la strategia di misura gioca un ruolo molto importante.

### 1.2.1 Classificazione incertezza di misura

Le cause di incertezza si possono classificare in tre grandi sottogruppi:

- **Incertezza dovuta alla risoluzione della misura:** come affermato in precedenza, nessun strumento ha risoluzione infinita il che produce un errore non trascurabile.
- **Errori tipo A o errori casuali:** fluttuazioni imprevedibili del risultato con ripetizione.
- **Errori tipo B o errori sistematici:** è un errore che affligge ogni misurazione in maniera continua e costante, per identificare errori di questo tipo è necessario ripetere l'esperimento in condizioni differenti.

---

<sup>1</sup>Per una trattazione più completa si veda capitolo 2

### 1.2.2 Rappresentazione dell'incertezza

Riuscire a predire l'incertezza a cui è affetto l'esperimento è un aspetto fondamentale di ogni relazione scientifica. Esistono vari metodi per considerare il contributo degli errori ai dati tra cui:

- **Incetenza assoluta:** prevede di fissare un valore centrale o di riferimento  $X_0$  e assegnare un valore di incertezza in positivo e in negativo rispettivamente  $\delta X_+$  e  $\delta X_-$  così da ottenere un intervallo di incertezza del tipo:

$$X = X_0 + \delta X_+ - \delta X_- \quad (1.1)$$

oppure

$$[X_0 - \delta X_-; X_0 + \delta X_+] \quad (1.2)$$

È possibile che gli intervalli di incertezza coincidano  $\delta X := \delta X_+ = \delta X_-$  formando quindi un intervallo simmetrico centrato in  $X_0$

$$X = X_0 \pm \delta X \quad (1.3)$$

- **Incetenza relativa:** è definita come il rapporto tra l'incetenza  $\delta X$  e il valore di riferimento  $X_0$  così da divenire:

$$X = X_0 \left( 1 \pm \frac{\delta X}{X_0} \right) \quad (1.4)$$

### 1.2.3 Tipi di intervallo di incertezza

L'intervallo di incertezza può rappresentare concetti molto diversi:

- **Intervallo di incetenza massima:** è l'intervallo che include tutti i dati prelevati

$$X \in [X_0 - \delta X; X_0 + \delta X] \quad \forall X \quad (1.5)$$

- **Incetenza tipo o standard:** si propone di misurare la fluttuazione tipica di ogni dato al fine di stimare la fluttuazione totale attorno al valore centrale  $X_0$ .

$$\sigma X := \sqrt{\langle (X_i - X_0)^2 \rangle} \quad (1.6)$$

La definizione di  $\sigma$  misura infatti in media lo scarto quadratico tra ogni valore e il valore centrale.

- **Intervalli di confidenza:** sono intervalli assunti rifacendosi alla probabilità che ha un dato di cadere in un certo intervallo;

- scelgo la probabilità  $P_o$  che mi interessa
- determino la larghezza dell'intervallo tenendo conto che:

$$P(X \in [\quad]) = P_o \quad (1.7)$$

## 1.3 Come trattare gli errori

### 1.3.1 Errore di risoluzione

Definiamo la risoluzione dello strumento di misura come  $\Delta X_s$  e come risoluzione della procedura di misura  $\Delta X_m$ . Esistono procedure di misura in grado di ridurre notevolmente  $\Delta X_m$  fino ad arrivare alla condizione  $\Delta X_m \ll \Delta X_s$ . Ad esempio se dovessimo misurare lo spessore di un foglio risulterebbe molto più produttivo misurare lo spessore di  $n$  fogli e poi dividere la lunghezza  $l$  per  $n$  ottenendo così una riduzione drastica dell'errore di procedura di misura  $\Delta X_m$ . Osserviamo inoltre che se la misura dovesse essere affetta da solo errore di risoluzione allora la ripetizione della misura in condizione identiche restituirà sempre lo stesso valore. Posso quindi escludere errori di tipo  $A$  ma non errori di tipo  $B$ .

Si può interpretare l'errore di risoluzione, come visto in precedenza, in molteplici maniere:

- **Incertezza massima di risoluzione:** ovvero assegno un intervallo massimale che compre la totale distribuzione dei dati<sup>2</sup>. Otterrei quindi un intervallo di incertezza dove se  $\Delta X_m$  è l'errore di risoluzione, l'intervallo  $X$  è dato da:

$$X = X_0 \pm \frac{\Delta X_m}{2} \quad (1.8)$$

- **Incertezza associata ad una probabilità:** ovvero, ipotizzando che il valore  $X$  si distribuisce all'intervallo di incertezza massima<sup>3</sup> si fissa una probabilità e si risale agli estremi dell'intervallo.
- **Incertezza tipo (standard) di risoluzione:** dobbiamo dare significato a  $\sigma$  definito come:

$$\sigma_{ris} := \sqrt{\langle (X_i - X_0)^2 \rangle} \quad (1.9)$$

1.  $X_0 :=$  Valore di lettura dello strumento.
2.  $X_i :=$  Possibili valori veri di  $X$  non percepibili a causa della risoluzione dello strumento assumendo equiprobabilità nell'intervallo di incertezza massimo.
3.  $\langle .. \rangle :=$  Media integrale dovuta all'assunzione di intervallo equiprobabilistico.

---

<sup>2</sup>Come nel caso 1.2.3 di incertezza massima.

<sup>3</sup>Come nel caso 1.2.3 di intervallo di confidenza

$$\sigma_{ris}^2 = \frac{1}{\Delta X_m} \int_{X_0 - \frac{\Delta X_m}{2}}^{X_0 + \frac{\Delta X_m}{2}} (X - X_0)^2 dx \quad (1.10)$$

$$\stackrel{(\dagger)}{=} \frac{1}{\Delta X_m} \int_{-\frac{\Delta X_m}{2}}^{\frac{\Delta X_m}{2}} t^2 dt \quad (1.11)$$

$$= \frac{1}{3\Delta X_m} \left[ \left( \frac{\Delta X_m}{2} \right)^3 - \left( -\frac{\Delta X_m}{2} \right)^3 \right] \quad (1.12)$$

$$= \frac{\Delta X_m^2}{12} \quad (1.13)$$

Dove su  $(\dagger)$  è stata applicata la sostituzione  $t \stackrel{s}{=} X - X_0$

$$\sigma_{ris} = \frac{\Delta X_m}{\sqrt{12}} \simeq 0.3\Delta X_m \quad (1.14)$$

Da cui si ottiene che l'errore di risoluzione è:

$$X \simeq X_0 \pm 0.3\Delta X_m \quad (1.15)$$

che è anche l'intervallo a cui corrisponde il 58% di confidenza

### 1.3.2 Errore sistematico o di tipo B

L'errore tipo è un errore che si ripete per ogni misurazione causato generalmente dall'inadeguatezza del modello dell'esperimento o dalle condizioni ambientali trascurate o fuori controllo. Vi sono due principali vie da seguire per porre rimedio ad un errore di tipo B: una più teorica che agisce sul modello dell'esperimento l'altra più sperimentale:

- **Rimedio teorico:** raffino il modello dell'esperimento oppure considero il fattore scatenante calcolandone la perturbazione. Questa procedura va a riscuotersi sul calcolo delle incertezze tramite congetture e ipotetici modelli di distribuzione di probabilità degli errori.
- **Rimedio sperimentale:** ripeto l'esperimento in condizioni differenti ovvero tento di rendere misurabile la fluttuazione. Quest'altra procedura considera l'intervallo di incertezza stimabile alla stessa maniera di errori tipo A.

### 1.3.3 Errore casuale o di tipo A

Stimatore per eccellenza e senz'altro più intuitivo è il  $\sigma$  definito come:

$$\sigma := \sqrt{\langle (x_i - x_0)^2 \rangle} \quad (1.16)$$

1.  $x_0 :=$  media aritmetica dei dati o campioni  $x_i$
2.  $\langle .. \rangle :=$  interpretabile anch'essa come media aritmetica

Otteniamo quindi

$$\sigma^* := \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - m^*)^2} =: \sqrt{D^*} \quad (1.17)$$

dove  $\sigma^*$  è detta **deviazione standard campionaria** e  $D^*$  è detta **varianza campionaria**

### Interpretazione della varianza campionaria

$$D^* = \frac{1}{N} \sum_{i=1}^N (x_i - m^*)^2 = m^*[(x_i - m^*[x])^2] \quad (1.18)$$

$$= \frac{1}{N} \sum_{i=1}^N (x_i^2 - 2x_i m^*[x] + m^{*2}) \quad (1.19)$$

$$= \frac{1}{N} \sum_{i=1}^N x_i^2 - \frac{2}{N} \left( \sum_{i=1}^N x_i \right) m^*[x] + \frac{1}{N} \sum_{i=1}^N m^{*2}[x] \quad (1.20)$$

$$\stackrel{(\dagger)}{=} m^*[x^2] - 2m^*[x]m^*[x] + m^{*2}[x] \quad (1.21)$$

$$= m^*[x^2] - m^{*2}[x] \quad (1.22)$$

Dove su  $(\dagger)$  si è sfruttato il fatto che:

$$\frac{1}{N} \sum_{i=1}^N x_i = m^*[x]$$

La varianza campionaria non è altro che la differenza tra la media del quadrato dei campioni e la media quadrati dei campioni.

Se dovessi considerare esponenti  $k$  di  $m^*[(x_i - x_0)^k]$  funzionerebbe comunque con  $k \neq 2$ ?

- $k = 1$ : si avrebbe il conseguente annullamento dello scarto e quindi non porterebbe alcuna informazione.
- $k = 2n \quad n \in \mathbb{Z}^{++}$ : si avrebbe in risalto le fluttuazioni isolate grandi, poiché l'errore tipo  $A$  presenta fluttuazioni continue ma non improvvise non è un buon stimatore.
- $k = 2n + 1 \quad n \in \mathbb{Z}^{++}$ : si avrebbero dei membri negativi che andrebbero a compensare le fluttuazioni positive e quindi a falsare la corretta stima della deviazione.

- $|(x_i - x_0)|$ : il valore assoluto risolverebbe il problema dei membri negativi ma porterebbe la  $\sigma^*$  a legarsi più di frequente a distribuzioni di probabilità molto comuni come la gaussiana.

## 1.4 Istogrammi

L'istogramma è un grafico che si propone di misurare la forma della distribuzione, informa sulla frequenza con cui si sono verificati differenti valori della grandezza. Definiamo dei parametri fondamentali per l'istogramma:

- *il binning*: è la divisione sull'asse delle  $x$  in intervalli di uguale dimensione. Ogni intervallo avrà valore centrale  $X_j$  e una lunghezza  $\Delta X_j$ .
- $\mathcal{N}$ : è il numeri di bin.
- $N$ : è il numero totale dei dati

Sull'asse delle  $x$  si pongono i diversi bin, mentre sull'asse delle  $y$  si pone il conteggio dei compresi nel rispettivo bin.

La condizione di normalizzazione è

$$N = \sum_{j=1}^{\mathcal{N}} n_j^* \quad (1.23)$$

Gli istogrammi non possono essere confrontati tra di loro a meno che non abbiano lo stesso binning e lo stesso numero di dati, per ovviare a questo problema si introduce la normalizzazione in area e in altezza.

### 1.4.1 Normalizzazione in altezza

La normalizzazione in altezza comporta la sostituzione dei valori sull'asse delle  $y$  con  $p_j^*$  detta frequenza campionaria definita come:

$$p_j^* := \frac{n_j^*}{N} \quad (1.24)$$

Con questa normalizzazione è possibile confrontare istogrammi con numero di dati differente ma pur sempre con lo stesso binning.  $p_j^*$  infatti è detta **frequenza campionaria** perché simile alla probabilità che il valore cada all'interno del bin  $j$ -esimo. Anche in questo caso, vale una condizione di normalizzazione ovvero:

$$\sum_{j=1}^{\mathcal{N}} p_j^* = \frac{\sum_{j=1}^{\mathcal{N}} n_j^*}{N} = 1 \quad (1.25)$$

### 1.4.2 Normalizzazione in area

La normalizzazione in area prende in considerazione anche il valore medio del bin  $\Delta X_j$  e definisce la **frequenza o densità campionaria**  $f_j^*$

$$f_j^* := \frac{p_j^*}{\Delta X_j} = \frac{n_j^*}{N \Delta X_j} \quad (1.26)$$

Istogrammi normalizzati in area si possono confrontare liberamente con altri istogrammi a binning e numero di dati differenti a patto che anch'essi siano normalizzati in area. La condizione di normalizzazione si rifà alla (1.25) ovvero:

$$\sum_{j=1}^{\mathcal{N}} f_j^* \Delta X_j^* = 1 = \sum_{j=1}^{\mathcal{N}} p_j^* \quad (1.27)$$

### 1.4.3 Scelta del binning

La scelta del binning è fondamentale per una corretta interpretazione dell'istogramma, un binning troppo stretto non porterebbe informazione sulla distribuzione ma sul singolo dato mentre un binning troppo largo, sarebbe inutile poiché non renderebbe possibile l'analisi delle caratteristiche relative alla distribuzione. Bisogna optare per una scelta di compromesso ovvero scegliere il binning che più esalta la forma della distribuzione; è bene scegliere un binning  $\Delta X_j \sim \sqrt{N}$  vedremo più avanti il motivo

### 1.4.4 Parametri statistici

Anche qui ci ritroviamo a dare significato ad  $m^*$  e  $D^*$ .

$$m^*[x] := \frac{1}{N} \sum_{i=0}^N x_i \stackrel{(\dagger)}{\simeq} \frac{1}{N} \sum_{j=1}^{\mathcal{N}} x_j n_j^* = \sum_{j=1}^{\mathcal{N}} x_j p_j^* = \sum_{j=1}^{\mathcal{N}} x_j f_j^* \Delta X_j \quad (1.28)$$

Notiamo che in  $(\dagger)$  ci si è serviti dell'approssimazione:

$$x_i \simeq x_j n_j^*$$

$$D^*[x] := \frac{1}{N} \sum_{i=1}^N (x_i - m^*[x])^2 \stackrel{(\dagger)}{\simeq} \frac{1}{N} \sum_{j=1}^{\mathcal{N}} (x_j - m^*[x])^2 n_j^* \quad (1.29)$$

$$= \sum_{j=1}^{\mathcal{N}} (x_j - m^*[x])^2 p_j^* = \sum_{j=1}^{\mathcal{N}} (x_j - m^*[x])^2 f_j^* \Delta X_j \quad (1.30)$$

in  $(\ddagger)$  ci si è serviti invece di:

$$(x_i - m^*[x])^2 \simeq (x_j - m^*[x])^2 n_j^*$$



### 1.4.5 Istogrammi cumulativi

Gli istogrammi cumulativi sono categorie di istogrammi che riportano sull'asse  $y$  il conteggio cumulativo ovvero il conteggio come gli istogrammi normali a cui va sommato il conteggio di tutti i bin precedenti. Piuttosto utilizzati tra gli istogrammi cumulativi sono gli istogrammi cumulativi in frequenza campionaria che riportano sull'asse delle  $y$  la frequenza cumulativa di ogni bin. Sono particolarmente indicati per l'utilizzo di quantili ovvero rette tracciate da una percentuale di  $p_j^*$  *cumulativo* che vanno ad intercettare il rispettivo bin  $j$ -esimo così da identificare ad occhio che probabilità c'è di trovarsi al di sopra o sotto di un certo valore di probabilità. Il percentile o il quantile rispettivamente al 50% o 0.5 merita il nome particolare di *mediana*.

### 1.4.6 Proprietà asintotiche degli istogrammi

Per limite asintotico degli istogrammi si intende  $N \rightarrow \infty$  ovvero un numero di campioni infinito ottenendo così l'**istogramma limite**. L'istogramma limite ha la proprietà di assumere la forma corretta della distribuzione perciò, trattandosi di una condizione irrealizzabile in laboratorio, si utilizzano gli *stimatori* per avvicinarsi il più possibile alle proprietà dell'istogramma limite. Per esempio, stimatore di  $f_j$  ovvero la densità limite riferita al bin  $j$ -esimo è  $f_j^*$ . La relazione è la seguente:

$$f_j^* \xrightarrow{N \rightarrow \infty} f_j \quad (1.31)$$

Quindi  $f_j^*$  è uno stimatore di  $f_j$ . Il concetto di istogramma limite non è altro che un'astrazione più grande essendo legata al binnig; si dice **distribuzione limite** se soddisfa:

$$N \rightarrow \infty \quad (1.32)$$

$$\Delta X_j \rightarrow 0 \quad (1.33)$$

Ovvero oltre alla proprietà precedente per l'istogramma limite, si richiede un binning infinitesimo. In questo caso, la densità campionaria diventa:

$$f_j^* \xrightarrow{N \rightarrow \infty; \Delta X_j \rightarrow 0} f_j(x) \quad (1.34)$$

dove  $f_j(x)$  è detta distribuzione limite di  $x$  o densità di probabilità di  $x$  con unità di misura  $\frac{1}{[x]}$ . La condizione di normalizzazione passando al limite della distribuzione diventa:

$$\sum_{j=1}^N f_j^* \Delta X_j^* = 1 \xrightarrow{N \rightarrow \infty; \Delta X_j \rightarrow 0} \int_{\text{dominio di } x} f(x) dx = 1 \quad (1.35)$$

## 1.5 Stima dei parametri della distribuzione limite

Come accennato in precedenza, il raggiungimento sperimentale della distribuzione limite è fuori discussione, per tanto si utilizzano gli **stimatori**, ovvero particolari funzioni che tendono ad assomigliare ai parametri veri della distribuzione. Tra gli stimatori introdotti prima, abbiamo:

- media campionaria <sup>4</sup>

$$m^*[x] \simeq \sum_{j=1}^{\mathcal{N}} x_j f_j^* \Delta X_j \xrightarrow{N \rightarrow \infty; \Delta X_j \rightarrow 0} m[x] := \int_{\text{dom. di } x} x f(x) dx \quad (1.36)$$

- varianza campionaria <sup>5</sup>

$$D^*[x] \simeq \sum_{j=1}^{\mathcal{N}} (x_j - m^*[x])^2 f_j^* \Delta X_j \xrightarrow{N \rightarrow \infty; \Delta X_j \rightarrow 0} D[x] := \quad (1.37)$$

$$:= \int_{\text{d. di } x} f(x)(x - m)^2 dx \quad (1.38)$$

Fino ad adesso abbiamo utilizzato  $D^*$  come stimatore di varianza campionaria; esiste però uno stimatore migliore ovvero  $\tilde{D}$  definito come:

$$\tilde{D} := \frac{N}{N-1} D^* = \frac{1}{N-1} \sum_{i=1}^N (x_i - m^*)^2 \quad (1.39)$$

Risulta uno stimatore migliore per vari motivi:

- Con un solo dato  $\tilde{D}$  non è definito ed impedisce la misura.
- $\tilde{D} > D^*$  dato che  $\frac{N}{N-1} < 1$  e poiché  $(x_i - m^*)^2$  in media<sup>6</sup> è minore di  $(x_i - m)^2$ .
- $D^* \xrightarrow{N \rightarrow \infty} \tilde{D}$  quindi le proprietà asintotiche sono conservate.

Come nel caso di  $\tilde{D}$  lo stimatore  $\sigma$  diventerà

$$\tilde{\sigma} := \sqrt{\tilde{D}} = \quad (1.40)$$

$$= \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - m^*)^2} \simeq \sqrt{\frac{1}{\mathcal{N}-1} \sum_{j=1}^{\mathcal{N}} (x_j - m^*[x])^2 f_j^* \Delta X_j} \quad (1.41)$$

<sup>4</sup>Si veda l'equazione (1.28) per la dimostrazione.

<sup>5</sup>Si veda l'equazione (1.29) per la dimostrazione.

<sup>6</sup>Questo fatto è dovuto ad  $m^*$  poiché è più centrato sui i campioni rispetto  $m$ .

### 1.5.1 Esempio: distribuzione uniforme

Nel caso di equiprobabilità delle fluttuazioni:

$$f(x) = \begin{cases} \frac{1}{\Delta X} & \text{per } x \in [x_{\min}, x_{\max}] \\ 0 & \text{altrove} \end{cases} \quad (1.42)$$

Dove  $\Delta X := x_{\max} - x_{\min}$  e quindi per soddisfare la condizione di normalizzazione:

$$1 = \int_{x_{\min}}^{x_{\max}} \frac{1}{\Delta X} dx \Leftrightarrow \Delta X = x_{\max} - x_{\min} \quad (1.43)$$

Nel caso in cui la distribuzione si uniforme  $m[x]$  diventa semplicemente:

$$m[x] = \int_{x_{\min}}^{x_{\max}} x \frac{1}{\Delta x} dx = \frac{1}{2\Delta x} [x_{\max}^2 - x_{\min}^2] \quad (1.44)$$

$$= \frac{1}{2\Delta x} [x_{\max} - x_{\min}] \Delta x \quad (1.45)$$

$$= \frac{1}{2} (x_{\max} + x_{\min}) \quad (1.46)$$

Mentre la varianza campionaria  $D[x]$ :

$$D[x] = \int_{x_{\min}}^{x_{\max}} (x - m)^2 \frac{1}{\Delta x} dx \stackrel{(\dagger)}{=} \frac{\Delta x^2}{12} \quad (1.47)$$

Dove su  $(\dagger)$  è stato utilizzato il risultato della (1.14)

Di conseguenza per  $\sigma[x]$  si ottiene:

$$\sigma[x] = \frac{\Delta x}{\sqrt{12}} \quad (1.48)$$

Osservazioni:

- La massima fluttuazione possibile di  $x$  rispetto a  $m[x]$  è  $\simeq 1.73 \sigma[x]$  infatti

$$\sigma[x] = \frac{\Delta x}{\sqrt{12}} = \frac{2}{\sqrt{12}} \frac{\Delta x}{2} \Rightarrow \frac{\Delta x}{2} = \frac{\sqrt{12}}{2} \sigma = \sqrt{3} \sigma \quad (1.49)$$

- La probabilità di ottenere un valore di  $x \in [m - \sigma, m + \sigma]$  è  $\simeq 58\%$  poiché

$$p(x \in [m - \sigma, m + \sigma]) = \int_{m-\sigma}^{m+\sigma} \frac{1}{\Delta x} dx = \frac{1}{2\Delta x} 2\sigma = \frac{2}{\Delta x} \frac{\Delta x}{\sqrt{12}} = \frac{1}{\sqrt{3}} \quad (1.50)$$

### 1.5.2 Esempio: distribuzione di Gauss o normale

La distribuzione di **Gauss** è la distribuzione più comune in natura, moltissimi fenomeni di origine aleatoria si possono ricondurre a questa distribuzione definita come:

$$f: \mathbb{R} \rightarrow \mathbb{R} \quad f(x) := \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{(x-m)^2}{2\sigma^2} \right] \quad (1.51)$$

Dipende da 2 parametri ovvero:

- $m$  detto valore medio o speranza matematica
- $\sigma$  detta deviazione standard

e anche per la gaussiana vale la condizione di normalizzazione:

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \quad (1.52)$$

Analizziamo ora alcune probabilità specifiche:

- $\text{prob} \{x \in [x_\alpha, x_\beta]\} = \text{area sottesa da } x_\alpha \text{ ad } x_\beta$
- $\text{prob} \{x \in [x - \sigma, x + \sigma]\} = \int_{m-\sigma}^{m+\sigma} f(x)dx \simeq 1 - \frac{1}{3} \simeq 68\%$
- $\text{prob} \{x \in [x - 2\sigma, x + 2\sigma]\} = \int_{m-2\sigma}^{m+2\sigma} f(x)dx \simeq 1 - \frac{1}{20} \simeq 95\%$
- $\text{prob} \{x \in [x - 3\sigma, x + 3\sigma]\} = \int_{m-3\sigma}^{m+3\sigma} f(x)dx \simeq 1 - \frac{1}{400} \simeq 99.74\%$

Quindi, con pochi dati posso misurare solo la campana della distribuzione mentre per misurare le code necessito di moltissimi dati.

## Capitolo 2

# Incertezza

In questo capitolo analizzeremo come propagare l'incertezza in misure dirette e indirette <sup>1</sup> e come discutere la compatibilità tra misure e predizione teoriche.

### 2.1 Propagazione incertezza

Supponendo che  $x$ ,  $y$ ,  $z$  siano misure dirette e che l'applicazione  $Q$  sia in relazione con le misure  $Q = Q(x, y, z)$  spiegheremo come calcolare l'incertezza su  $Q$  a seconda del tipo di legge.

#### 2.1.1 Relazione lineare

Supponiamo il caso più semplice ovvero una relazione lineare:

$$Q = a + bx + cy + dz \dots \quad (2.1)$$

con  $a, b, c, d..$  coefficienti costanti. Prima di stimare l'incertezza su  $Q$  necessito della stima valore medio o  $m[x]$ .

- Stima del valore medio di  $Q$  tramite:

---

<sup>1</sup>Per definizione di misure dirette-indirette si veda 1.1

– Media campionaria:

$$m^*[Q] := \frac{1}{N} \sum_{i=1}^N (a + bx_i + cy_i + dz_i \dots) \quad (2.2)$$

$$= \frac{1}{N} \sum_{i=1}^N a + \frac{1}{N} \sum_{i=1}^N bx_i + \frac{1}{N} \sum_{i=1}^N cy_i + \frac{1}{N} \sum_{i=1}^N dz_i \dots \quad (2.3)$$

$$= a + b \left( \frac{1}{N} \sum_{i=1}^N x_i \right) + c \left( \frac{1}{N} \sum_{i=1}^N y_i \right) + d \left( \frac{1}{N} \sum_{i=1}^N z_i \right) \dots \quad (2.4)$$

$$m^*[Q] = a + b m^*[x] + c m^*[y] + d m^*[z] \dots \quad (2.5)$$

$$= a + bx_0 + cy_0 + dz_0 \dots \quad (2.6)$$

– Speranza matematica:

$$m[Q] = \int_{\text{dom di } Q} Q f(Q) dQ \quad (2.7)$$

$$= \int_{\text{dom di } Q} (a + bx + cy + dz \dots) f(Q) dQ \quad (2.8)$$

$$\stackrel{(\dagger)}{=} a + bm[x] + cm[y] + dm[z] \dots \quad (2.9)$$

Dove su  $(\dagger)$  si utilizza il passaggio al limite asintotico di  $N \rightarrow \infty$  di (2.6). Il risultato è complicato da dimostrare ma molto intuitivo.

• Stima dell'incertezza di  $Q$  tramite:

– Deviazione standard campionaria  $\tilde{\sigma}$ :

$$\tilde{\sigma} := \sqrt{\tilde{D}[Q]} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (Q_i - Q_0)^2} \quad Q_0 := m^*[Q] \quad (2.10)$$

– Deviazione standard della popolazione  $\sigma[Q]$ :

$$\sigma[Q] := \sqrt{D[Q]} = \sqrt{\int_{\text{dominio}} (Q - m[Q])^2 f(Q) dQ} \quad (2.11)$$

Nota: per semplicità ora consideriamo il caso di  $\tilde{\sigma}$  in presenza di errori casuali per poi passare al limite asintotico  $N \rightarrow \infty$ . Definiamo anche la *covarianza campionaria*  $\tilde{\sigma}_{xy}$  che ci sarà utile in seguito. Introduciamo anche il concetto di grandezze *statisticamente indipendenti*.

### Grandezze statisticamente dipendenti

Per meglio comprendere la covarianza è necessario introdurre il concetto di **grandezze statisticamente indipendenti** ovvero:  $x$  e  $y$  si dicono statisticamente indipendenti se non sussiste alcuna relazioni fra le fluttuazioni di (2.12) e (2.13)

$$(x_i - m[x]) \quad (2.12)$$

$$(y_i - m[y]) \quad (2.13)$$

In altre parole se il valore di una fluttuazione non mi da alcuna informazione sul valore dell'altra fluttuazione. Attenzione che  $m[x]$  e  $m[y]$  possono essere in relazione, l'importante è che non ci sia legame tra le due **fluttuazioni**.

Definiamo la **covarianza campionaria**  $\tilde{\sigma}_{xy}$ :

$$\tilde{\sigma}_{xy} := \frac{1}{N-1} \sum_{i=1}^N (x_i - x_0)(y_i - y_0) \quad (2.14)$$

$$\sigma_{xy}^* := \frac{1}{N} \sum_{i=1}^N (x_i - x_0)(y_i - y_0) \quad (2.15)$$

da cui

$$\tilde{D}[Q] = \frac{1}{N-1} \sum_{i=1}^N (a + bx_i + cy_i + \dots - a - bx_0 - cy_0 - \dots)^2 \quad (2.16)$$

$$= \frac{1}{N-1} \sum_{i=1}^N (bx_i - bx_0)^2 + \frac{1}{N-1} \sum_{i=1}^N (cy_i - cy_0)^2 + \dots \quad (2.17)$$

+ *altri termini quadratici...*

$$+ \frac{2}{N-1} \sum_{i=1}^N (bx_i - bx_0)(cy_i - cy_0) + \dots \quad (2.18)$$

+ *altri termini doppi prodotti...*

$$= b^2 \tilde{D}[x] + c^2 \tilde{D}[y] + \dots \text{ altri termini di varianza} \quad (2.19)$$

$$+ 2bc \tilde{\sigma}_{xy} + 2bd \tilde{\sigma}_{xz} + \dots \text{ altri termini di covarianza} \quad (2.20)$$

$$\tilde{D}[Q] \xrightarrow{N \rightarrow \infty} D[Q] = b^2 D[x] + c^2 D[y] + \dots + 2bc \sigma_{xy} + 2bd \sigma_{xz} + \dots \quad (2.21)$$

Analizzando la definizione di covarianza abbinata alla definizione di statisticamente indipendenti otteniamo che:

$$(N-1)\tilde{\sigma}_{xy} = \sum_{i=1}^N (x_i - x_0)(y_i - y_0) \stackrel{(\dagger)}{\simeq} 0 \quad (2.22)$$

Una diretta conseguenza della indipendenza statistica è che la sommatoria avrà in media termini negativi e positivi che si bilanceranno a vicenda da cui segue  $(\dagger)$ . Questa compensazione farà crescere il valore della sommatoria meno rapidamente che  $\propto N$  dimostreremo infatti che è  $\propto \sqrt{N}$  cioè:

$$\tilde{\sigma}_{xy} \xrightarrow{N \rightarrow \infty} 0 = \sigma_{xy} \quad (2.23)$$

$$\forall N \quad |\tilde{\sigma}_{xy}| \ll \tilde{D}[x] + \tilde{D}[y] \quad (2.24)$$

Quindi  $\tilde{\sigma}_{xy}$  è trascurabile rispetto alle varianze. Per concludere, **nel caso di grandezze statisticamente indipendenti vale:**

$$D[Q] = b^2 D[x] + c^2 D[y] + \dots \quad (2.25)$$

$$\tilde{D}[Q] = b^2 \tilde{D}[x] + c^2 \tilde{D}[y] + \dots \quad (2.26)$$

Osservazioni:

- Dimostreremo che per grandezze statisticamente vale la seguente uguaglianza:

$$\begin{aligned} \left\langle \sum (x_i - x_0)(y_i - y_0) \right\rangle &= \sum \langle (x_i - x_0)(y_i - y_0) \rangle \\ &= \sum \langle x_i - x_0 \rangle \langle y_i - y_0 \rangle \end{aligned} \quad (2.27)$$

- La stima della propagazione dell'incertezza è valida solo per le incertezze tipo *A* mentre è errata per le incertezze che mostrano regolarità (come errori tipo *B*).
- L'espressione  $D[Q]$  è più generale e può essere applicata a qualunque modello teorico di incertezza.
- Ricordarsi che per le grandezze non statisticamente indipendenti va presa in considerazione l'effetto della covarianza.

## 2.2 Applicazioni della propagazione dell'incertezza

### 2.2.1 Media campionaria

Nel caso di misure ripetute in condizioni identiche e indipendenti calcolo lo scarto tipo di  $m^*[x]$  osservando che

$$m^* = \frac{1}{N} \sum_{i=1}^N x_i \quad (2.28)$$



combinazione lineare di

$$\frac{1}{N} \quad (2.29)$$

e

$$\sum_{i=1}^N x_i \quad (2.30)$$

diventa da (2.25) quindi:

$$\sigma[m^*] = \sqrt{\frac{1}{N^2}D[x_1] + \frac{1}{N^2}D[x_2] + \cdots + \frac{1}{N^2}D[x_n]} \quad (2.31)$$

$$\stackrel{(\dagger)}{=} \sqrt{\frac{N D[x]}{N^2}} = \frac{1}{\sqrt{N}} \sqrt{D[x]} = \frac{1}{\sqrt{N}} \sigma[x] \quad (2.32)$$

Dove su  $(\dagger)$  si è sfruttato il fatto che gli  $x_i$  sono misure ripetute nelle stesse condizioni  $\Rightarrow D[x_i] = D[x]$ .

Il fattore  $\frac{1}{\sqrt{N}}$  è detto **fattore di riduzione** dell'incertezza di  $m^*$  rispetto all'incertezza tipo del singolo campione. Si ricava quindi che

$$m^*[x] \pm \sqrt{\frac{1}{N} \frac{1}{N-1} \sum_{i=1}^N (x_i - m^*[x])^2} = m^*[x] \pm \frac{\tilde{\sigma}[x]}{\sqrt{N}} \quad (2.33)$$

che è l'intervallo d'incertezza tipo nel caso di errori tipo  $A$ .

Osserviamo inoltre che:

- Il contributo dell'incertezza casuale  $\propto \frac{1}{\sqrt{N}}$
- La riduzione di  $\sigma[m^*] \propto \frac{1}{\sqrt{N}}$  si applica solo agli errori di tipo  $A$  infatti per tutti gli errori che si verificano con regolarità

$$x_i = \delta x_i + \delta x + x_0 \quad (2.34)$$

dove:

- $\delta x_i :=$  *fluttuazione tipo A*
- $\delta x :=$  *fluttuazione uguale per tutti i valori*
- $\delta x_0 :=$  *valore vero*

otteniamo che

$$m^*[x] = \delta x + x_0 \quad (2.35)$$

per cui da

$$\delta x \gg \sigma_{casuali}[x] \quad (2.36)$$

e quindi si conclude con

$$\sigma[m^*] = \sigma[x] \quad (2.37)$$

cioè ho la stessa deviazione standard del singolo dato.

### 2.2.2 Misura di un multiplo di una grandezza $x$

Partiamo dalla relazione <sup>2</sup>  $Q = Mx$  dove

- $M \in \mathbb{Z}$
- $Q := \text{valore misurato}$
- $x := \text{valore calcolato}$
- $x_0 := \frac{Q_0}{M}$

Abbiamo dai risultati precedenti (2.25) e (2.26) che:

$$D[Q] = M^2 D[x] \quad (2.38)$$

$$\tilde{D}[Q] = M^2 \tilde{D}[x] \quad (2.39)$$

rispettivamente per un intervallo di incertezza tipo su  $x$

$$x_0 \pm \frac{\sigma[Q]}{M} \quad (2.40)$$

e per intervallo di incertezza casuale

$$x_0 \pm \frac{\tilde{\sigma}[Q]}{M} \quad (2.41)$$

Notare che in questo caso la riduzione di incertezza va come  $\propto \frac{1}{M}$  e quindi conviene misurare  $Q = Mx$  piuttosto che fare  $M$  misure separate di  $x$  per poi calcolarne la media.

### 2.2.3 Composizione di contributi di incertezza di varia natura

È molto probabile trovarsi nelle condizioni di dover considerare incertezza generata da diversi fattori; è quindi probabile che si tratti di eventi statisticamente indipendenti. Per calcolare la fluttuazione totale  $\delta x$  è sufficiente sommare i contributi di incertezza come della funzione totale  $\delta x$

$$\delta x = \delta x_A + \delta x_B + \delta x_{ris} \quad (2.42)$$

dove

- $\delta x_A := \text{errore casuale che dipende dal campione}$
- $\delta x_B := \text{errore sistematico}$

---

<sup>2</sup>Ad esempio la durata di  $M$  periodi di pendolo consecutivi

- $\delta x_{ris} := \text{errore di risoluzione}$

Grazie alle (2.25) ed (2.26) otteniamo che:

$$D_{totale}[x] = D_A[x] + D_B[x] + D_{ris}[x] \Rightarrow \sigma_{tot}[x] = \sqrt{\sigma_A^2[x] + \sigma_B^2[x] + \sigma_{ris}^2[x]} \quad (2.43)$$

Ora sappiamo che: **le incertezze dovute ad effetti indipendenti tra loro si compongono quadraticamente.**

Osserviamo che:

- $\sigma_{tot} \geq \sigma_{ris}$  la deviazione standard totale non può scendere sotto la deviazione standard di risoluzione
- se  $\sigma_A[m^*] \simeq \sigma_{ris}[x] \Rightarrow \sigma_{tot} \simeq \sqrt{2\sigma_{ris}^2[x]} \simeq \sqrt{2}\sigma_{ris}[x]$  aumentando il numero di misure per esempio di 4 ottengo:

$$\sigma_A[m^*] \rightarrow \frac{1}{2}\sigma_{ris}[x] \Rightarrow \sigma_{tot} = \sqrt{\frac{1}{4}\sigma_{ris}^2 + \sigma_{ris}^2} \simeq 1.1\sigma_{ris} \quad (2.44)$$

Il vantaggio nel ripetere le misure è molto piccolo rispetto al costo. Conviene quindi progettare l'esperimento in maniera tale da

$$\sigma_A[m^*] \gtrsim \sigma_{ris} \Rightarrow \frac{\sigma_A[x]}{\sqrt{N}} \gtrsim \frac{\Delta X_m}{\sqrt{12}} \Rightarrow N \lesssim \frac{\sigma^2[x]}{\Delta X_m^2} 12 \quad (2.45)$$

Così ottengo più incertezza di tipo  $A$  con costi accettabili.

Se lo scopo dell'esperimento è analizzare la forma della distribuzione nelle sue caratteristiche rare più misure indipendenti prendo meglio è.

## 2.3 Compatibilità tra misure

Per decidere se misure o predizioni teoriche sono compatibili tra di loro supponendole:  $x_a$  e relativa incertezza  $\delta x_a$ ,  $x_b$  relativa incertezza  $\delta x_b$  prima di tutto devono verificare

$$m[x_a] = m[x_b] \quad (2.46)$$

Se le speranze matematiche sono compatibili allora si può considerare la differenza  $R := x_a - x_b$  che deve essere:

$$R \simeq 0 \quad (2.47)$$

### 2.3.1 Intervallo d'incertezza massima

Se gli intervalli d'incertezza massima sono sovrapposti allora non posso escludere che le misure siano compatibili tra di loro quindi l'ipotesi del sono compatibili vince. Considero inoltre

$$|R| \leq \delta x_a + \delta x_b \quad (2.48)$$

se è verificata allora tengo buona l'ipotesi di compatibilità.

### 2.3.2 Intervallo d'incertezza tipo

Considero  $x_a \pm \sigma_a$  e  $x_b \pm \sigma_b$  indipendenti. Se utilizzo il sistema utilizzato 2.3.1 mi imbattere in falsi allarmi:

### 2.3.3 Problema dei falsi allarmi

Il problema di falso allarme si verifica perché la probabilità che

$$m[R] \simeq 0 \quad (2.49)$$

e nonostante ciò

$$R \notin [-\sigma_R, +\sigma_R] \quad (2.50)$$

è troppo alta. Con questo criterio chiamo troppo frequentemente un falso allarme di incompatibilità<sup>3</sup> anche se in realtà  $x_a$  è compatibile con  $x_b$ . La soluzione per i falsi allarmi è quella di allargare l'intervallo tramite un fattore di copertura  $k$  in maniera tale da:

$$R \leq k\sigma_R = k\sqrt{\sigma_A^2 + \sigma_B^2} \quad (2.51)$$

- $k$  alto è una condizione prudente perché abbassa la probabilità di falso allarme.
- $k$  basso è più potente perché permette di riconoscere incompatibilità più piccole.

In ogni caso si raccomanda  $k \gtrsim 2$ . La scelta comunque è responsabilità dello scienziato, è soggettiva e dipende da una valutazione costi/benefici. Per non inquinare la probabilità di falso allarme **bisogna decidere a priori il fattore di copertura  $k$** , se si dovesse sbirciare, la decisione della soglia potrebbe essere influenzata dalla tentazione di accondiscendere a qualche aspettativa.

- Nella comunicazione scientifica bisogna:
  - dare la conclusione su incompatibilità e relativo falso allarme del fattore di copertura scelto a priori
  - scrivere a che fattore di copertura corrisponde il valore trovato per  $r$ .
- Prestare attenzione che la relazione fra probabilità di falso allarme e  $k$  è conosciuta solo fino ad un certo valore di  $k$ , per valori superiori bisogna affidarsi ad estrapolazioni. Per esempio nel caso della distribuzione normale:

---

<sup>3</sup>Se seguissi una distribuzione gaussiana avrei il 32% di falso allarme

- $k \simeq 3$  probabilità di falso allarme  $\simeq \frac{1}{370}$
- $k \simeq 4$  probabilità di falso allarme  $\simeq \frac{1}{1.6 \times 10^4}$
- $k \simeq 5$  probabilità di falso allarme  $\simeq \frac{1}{1.7 \times 10^6}$
- $k \simeq 6$  probabilità di falso allarme  $\simeq \frac{1}{5 \times 10^8}$

## 2.4 Comporre l'informazione di risultati con incertezze differenti

Qualora ci si trovasse nel misurare la stessa grandezza in condizioni differenti per stabilire o meno la dipendenza dalle nuove condizioni.

### 2.4.1 Media pesata

Assumiamo:

- $N :=$  misure indipendenti di  $x_i$ .
- $m[x_i] = m$ : stessa speranza matematica.
- $\sigma[x_i] = \sigma_i$ : le deviazioni standard tra loro diverse.

La media campionaria è inservibile nel caso di  $\sigma$  diversi tra loro si introduce quindi la **media pesata** definita come:

$$m_w := \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i} \quad (2.52)$$

dove  $w_i$  sono i pesi<sup>4</sup>. Al denominatore vale la condizione di normalizzazione affinché:

$$m[m_w] = \frac{1}{\sum w_i} \sum w_i m[x_i] = m \quad (2.53)$$

Propagando l'incertezza come da (2.25) di  $m_w$  di  $\sigma_i$  si ottiene la sua varianza:

$$\sigma^2[m_w] = \frac{1}{(\sum w_i)^2} \sum_{i=1}^N w_i^2 \sigma^2[x_i] \quad (2.54)$$

Riepilogando le condizioni necessarie all'utilizzo della media pesata:

- $\sigma[w_m]$  è corretto se le  $\sigma_i$  sono corrette.
- La media pesata ha senso solo se le  $x_i$  sono compatibili tra loro ovvero hanno tutte la stessa speranza matematica.

---

<sup>4</sup>La media campionaria è interpretabile come una media pesata di peso unitario

- $x_i$  devono essere indipendenti.

La media pesata da liberà nella scelta dei pesi, posso usare criteri soggettivi<sup>5</sup>. La scelta corretta dei pesi determina la bontà del risultato in particolare se la distribuzione è di gauss la scelta ottimale è:

$$w_i = \frac{1}{\sigma_i^2} \quad (2.55)$$

Infatti minimizza l'incertezza statistica su  $m_w$  facendo diventare:

$$m_w = \frac{\sum_{i=0}^N \frac{x_i}{\sigma_i^2}}{\sum_{i=0}^N \frac{1}{\sigma_i^2}} \quad (2.56)$$

$$\sigma[m_w] = \frac{1}{\sqrt{\sum_{i=0}^N \frac{1}{\sigma_i^2}}} \quad (2.57)$$

Otteniamo quindi:

$$\sigma[m_w] = \frac{1}{\sum_{i=0}^N \frac{1}{\sigma_i^2}} \left( \sum_{i=0}^N \frac{\sigma_i^2}{\sigma_i^4} \right)^{\frac{1}{2}} = \frac{1}{\sum_{i=0}^N \frac{1}{\sigma_i^2}} \left( \sum_{i=0}^N \frac{1}{\sigma_i^2} \right)^{\frac{1}{2}} \quad (2.58)$$

### Confronto tra media campionaria e media pesata

La differenza fondamentale è nella stima di  $\sigma$ . Nella media campionaria  $\sigma[m^*]$  è calcolato dalle fluttuazioni<sup>6</sup> degli  $x_i$ . Nella media pesata invece è calcolato dalle stime di  $\sigma_A$  dei dati e dagli  $w_i$  inoltre non considera le fluttuazioni fra i dati e richiede più informazioni; è quindi più potente ma meno robusta<sup>7</sup>. Se le stime degli  $\sigma[x_i]$  non sono molto diverse fra loro le loro differenze potrebbero essere dovute soltanto a incertezze casuali degli stimatori, conviene quindi considerarle uguali fra loro ed utilizzare la media campionaria perché più robusta.

## 2.5 Propagazione dell'incertezza: interpretazione grafica

### 2.5.1 Caso unidimensionale

Per semplicità consideriamo il caso unidimensionale lineare del tipo

$$Q := a + bx \quad (2.59)$$

<sup>5</sup>Credo più ad una misura piuttosto che ad un'altra.

<sup>6</sup>Si veda l'equazione (1.29).

<sup>7</sup>Risente la dipendenza dalla veridicità delle informazioni.

È facile verificare che:

$$\sigma[Q] = |b|\sigma[x] \quad (2.60)$$

Il valore assoluto della pendenza della legge lineare converte  $\sigma[x] \rightarrow \sigma[Q]$   
Generalizzando ora una legge non lineare:

$$Q := f(x) \quad (2.61)$$

Sappiamo che  $\delta x$  dipende dal punto  $x_o$  in cui è calcolato e che  $\delta x \propto$  pendenza locale derivata di  $f(x)$ . Non è sbagliato pensare ad una approssimazione lineare di  $f(x)$  in un intorno di  $x_0$ :

$$f(x) \stackrel{(\dagger)}{\simeq} f(x_0) + \left( \frac{df}{dx} \right)_{x=x_0} (x - x_0) \quad (2.62)$$

$$\simeq a + b(x - x_0) \quad (2.63)$$

Dove su  $(\dagger)$  è stato utilizzato il primo sviluppo in serie di Taylor.

Da qui posso applicare le formule di propagazione già viste<sup>8</sup>.

$$\sigma[Q] \simeq \left| \frac{dQ}{dx} \right|_{x=x_0} \sigma[x] \quad (2.64)$$

e

$$m[Q] \simeq Q_0 = f(x_0) \quad (2.65)$$

Queste formule però non sono valide in due casi:

- Se sono in un punto di massimo, minimo o flesso

$$\frac{df}{dx} = 0 \quad (2.66)$$

Devo quindi continuare a sviluppare in serie di Taylor.

- $\sigma[x]$  è troppo grande e la  $f(x)$  non è linearizzabile con la retta tangente  $(x_0, Q_0)$

### 2.5.2 Esempi

- $Q = x^a$  calcolo  $m[Q]$  che sarà semplicemente:

$$m[Q] \simeq x_0^a \quad (2.67)$$

Invece per  $\sigma[Q]$  ottengo:

$$\sigma[Q] \simeq |a x_0^{a-1}| \sigma[x] \quad (2.68)$$

L'errore relativo in questi casi è l'ideale per controllare a mente cosa ci si aspetta

$$\frac{\sigma[Q]}{Q_0} = |a| \frac{\sigma[x]}{x_0} \quad (2.69)$$

---

<sup>8</sup>Si veda la sezione 2.1.1

- $Q = \ln[x]$  ovviamente definito per  $x > 0$  si ottiene che:

$$\frac{dQ}{dx} = \frac{1}{x} \Rightarrow \sigma[Q] \simeq \frac{\sigma[x]}{x_0} \quad (2.70)$$

## 2.6 Estensione a funzioni di più grandezze

Supponiamo

$$Q := f(x, y) \quad (2.71)$$

Otteniamo che:

$$Q = f(x, y) \simeq f(x_0, y_0) + \left( \frac{\partial Q}{\partial x} \right)_{x_0, y_0} (x - x_0) + \left( \frac{\partial Q}{\partial y} \right)_{x_0, y_0} (y - y_0) \quad (2.72)$$

Che altro non è che lo sviluppo in serie di Taylor al primo ordine dove

$$x_0 := m[x] \quad (2.73)$$

$$y_0 := m[y] \quad (2.74)$$

Quindi vale che:

$$m[Q] \simeq Q_0 = f(x_0, y_0) \quad (2.75)$$

$$\sigma[Q] \simeq \sqrt{\left| \frac{\partial Q}{\partial x} \right|_{x_0, y_0}^2 \sigma^2[x] + \left| \frac{\partial Q}{\partial y} \right|_{x_0, y_0}^2 \sigma^2[y]} \quad (2.76)$$

### 2.6.1 Esempi

- $Q = xy$  abbiamo quindi che:

$$m[Q] \simeq Q_0 = x_0 y_0 \quad (2.77)$$

A patto che  $x$  e  $y$  siano statisticamente indipendenti:

$$\sigma[Q] \simeq \sqrt{|y_0|^2 \sigma^2[x] + |x_0|^2 \sigma^2[y]} \quad (2.78)$$

Con errore relativo:

$$\frac{\sigma[Q]}{|Q_0|} = \sqrt{\left( \frac{\sigma[x]}{|x_0|} \right)^2 + \left( \frac{\sigma[y]}{|y_0|} \right)^2} \quad (2.79)$$

- $Q = x^a y^b$  abbiamo che:

$$m[Q] \simeq Q_0 = x_0^a y_0^b \quad (2.80)$$



E quindi

$$\sigma[Q] \simeq \sqrt{a^2|x_0|^{2(a-1)}|y_0|^{2b}\sigma^2[x] + b^2|x_0|^{2a}|y_0|^{2(b-1)}\sigma^2[y]} \quad (2.81)$$

$$\simeq |Q_0| \sqrt{a^2 \frac{\sigma^2[x]}{|x_0|^2} + b^2 \frac{\sigma^2[y]}{|y_0|^2}} \quad (2.82)$$

Con errore relativo:

$$\frac{\sigma[Q]}{|Q_0|} \simeq \sqrt{\left(a \frac{\sigma[x]}{x_0}\right)^2 + \left(b \frac{\sigma[y]}{y_0}\right)^2} \quad (2.83)$$

### 2.6.2 Propagazione dell'incertezza su misure dirette non statisticamente indipendenti

Si utilizza lo stesso sviluppo per linearizzare:

$$Q = f(x, y) \simeq f(x_0, y_0) + \left(\frac{\partial Q}{\partial x}\right)_{x_0, y_0} (x - x_0) + \left(\frac{\partial Q}{\partial y}\right)_{x_0, y_0} (y - y_0) \quad (2.84)$$

Otengo poi:

$$D[Q] \simeq \left|\frac{\partial Q}{\partial x}\right|_{x_0, y_0}^2 \sigma^2[x] + \left|\frac{\partial Q}{\partial y}\right|_{x_0, y_0}^2 \sigma^2[y] + 2 \left(\frac{\partial Q}{\partial x}\right)_{x_0, y_0} \left(\frac{\partial Q}{\partial y}\right)_{x_0, y_0} \sigma_{xy} \quad (2.85)$$

Dove  $\sigma_{xy}$  è la covarianza<sup>9</sup> che maggioremo con  $D[x]D[y]$ . Sia  $m^*[x]$  operatore di media campionaria, lineare cioè

$$m^*[ax + by] = \frac{1}{N} \sum (ax_i + by_i) = a \frac{1}{N} \sum x_i + b \frac{1}{N} \sum y_i \quad (2.86)$$

$$= am^*[x] + bm^*[y] \quad (2.87)$$

Considero  $A(t)$  definito come

$$A(t) := m^*[\{(x - m^*[x]) + t(y - m^*[y])\}^2] \geq 0 \quad \forall t \in \mathbb{R} \quad (2.88)$$

Considero  $t$  definito come

$$t := -\frac{\sigma_{xy}^*}{D^*[y]} \quad (2.89)$$

Con:

$$\sigma_{xy}^* := m^*[(x - m^*[x])(y - m^*[y])] \quad (2.90)$$

e

$$D^*[y] := m^*[(y - m^*[y])^2] \quad (2.91)$$

---

<sup>9</sup>Per la definizione di covarianza si veda l'equazione (2.15)

$$0 \leq A\left(-\frac{\sigma_{xy}^*}{D^*[y]}\right) = m^* \left[ \left\{ (x - m^*[x]) - \frac{\sigma_{xy}^*}{D^*[y]} (y - m^*[y]) \right\}^2 \right] \quad (2.92)$$

$$= m^* [(x - m^*[x])^2] \quad (2.93)$$

$$+ m^* \left[ -2 \frac{\sigma_{xy}^*}{D^*[y]} (x - m^*[x]) (y - m^*[y]) \right] \quad (2.94)$$

$$+ m^* \left[ \left( \frac{\sigma_{xy}^*}{D^*[y]} \right)^2 (y - m^*[y])^2 \right] \quad (2.95)$$

$$\stackrel{(\dagger)}{=} D^*[x] - 2 \frac{\sigma_{xy}^*}{D^*[y]} \sigma_{xy}^* + \left( \frac{\sigma_{xy}^*}{D^*[y]} \right)^2 D^*[y] \quad (2.96)$$

$$= D^*[x] - 2 \frac{\sigma_{xy}^{*2}}{D^*[y]} + \frac{\sigma_{xy}^{*2}}{D^*[y]} \Rightarrow \sigma_{xy}^{*2} \leq D^*[x] D^*[y] \quad (2.97)$$

Dove su  $(\dagger)$  si è sfruttato il fatto che: il primo membro in (2.93) è  $D^*[x]$ , il secondo membro in (2.94) è la covarianza  $\sigma_{x,y}^*$  e l'ultimo termine in (2.95) è  $D^*[y]$ .

Nel caso non sia facile stimare la covarianza è possibile dare una sovrastima dell'incertezza propagata a  $Q = f(x, y)$

$$D[Q] \simeq \left| \frac{\partial Q}{\partial x} \right|^2 \sigma^2[x] + \left| \frac{\partial Q}{\partial y} \right|^2 \sigma^2[y] + 2 \left( \frac{\partial Q}{\partial x} \right)_{x_0, y_0} \left( \frac{\partial Q}{\partial y} \right)_{x_0, y_0} \sigma_{xy} \quad (2.98)$$

$$\leq \left| \frac{\partial Q}{\partial x} \right|^2 \sigma^2[x] + \left| \frac{\partial Q}{\partial y} \right|^2 \sigma^2[y] + 2 \left( \frac{\partial Q}{\partial x} \right) \left( \frac{\partial Q}{\partial y} \right) \sigma_{xy} \quad (2.99)$$

$$\leq \left| \frac{\partial Q}{\partial x} \right|^2 \sigma^2[x] + \left| \frac{\partial Q}{\partial y} \right|^2 \sigma^2[y] + 2 \left( \frac{\partial Q}{\partial x} \right) \left( \frac{\partial Q}{\partial y} \right) \sigma[x] \sigma[y] \quad (2.100)$$

$$= \left[ \left| \frac{\partial Q}{\partial x} \right| \sigma[x] + \left| \frac{\partial Q}{\partial y} \right| \sigma[y] \right]^2 \quad (2.101)$$

Si ha quindi che:

$$\sigma[Q] \lesssim \left| \frac{\partial Q}{\partial x} \right| \sigma[x] + \left| \frac{\partial Q}{\partial y} \right| \sigma[y] \quad (2.102)$$

ovvero la somma lineare dei sigma pesata con le derivate parziali

## Capitolo 3

# Regressione lineare e FIT

Supponiamo di avere una legge che correli due grandezze fisiche del tipo:

$$y = f(x) \quad (3.1)$$

La procedura che prevede l'analisi dei dati sperimentali al fine di risalire alla legge è detta **FIT** cioè "adattare". Anche in questo caso per semplicità assumiamo una relazione lineare del tipo:

$$y = a + bx \quad (3.2)$$

### 3.1 Metodo di minima e massima pendenza

Se suppongo di avere un grafico raffigurante la legge lineare con tanto di barre d'errore:

- Traccio le **rette di minima e massima** pendenza ovvero le rette passanti per tutte le barre d'errore con pendenza rispettivamente minima e massima.
- Misuro graficamente i parametri di intercetta e pendenza rispettivamente:

$$(a_{min}; b_{min}) \text{ retta di minima pendenza} \quad (3.3)$$

$$(a_{max}; b_{max}) \text{ retta di massima pendenza} \quad (3.4)$$

- Stimo i parametri della legge prendendo la media dei valori:

$$b := \frac{b_{min} + b_{max}}{2} \quad \delta b \sim \frac{|b_{max} - b_{min}|}{2} \quad (3.5)$$

$$a := \frac{a_{min} + a_{max}}{2} \quad \delta a \sim \frac{|a_{max} - a_{min}|}{2} \quad (3.6)$$

Questo metodo ha il vantaggio di essere molto facile da eseguire però ha dei punti a sfavore non trascurabili:

- La decisione della compatibilità delle rette di minima e massima pendenza è soggettiva:
  - nel caso di incertezza tipo, non è assicurato che i valori siano interni alle rette di minima e massima pendenza.
  - solo nel caso di massima incertezza la ricetta non è ambigua.
- L'interpretazione di  $\delta b$  e  $\delta a$  non è chiara: non sono né scarti massimi né errori.
- Mancano criteri oggettivi sulla decisione di compatibilità dei dati con una legge lineare.

### 3.2 Metodo dei minimi quadrati

Il metodo dei minimi quadrati si propone come alternativa al metodo precedente. Questo metodo infatti ha la funzione di trovare i parametri della legge che più si adattano ai dati sperimentali ed è una figata. Consideriamo:

- $(x_i, y_i) := N$  dati sperimentali
- $y(x) := a + bx =$  Legge da stimare

Definiamo la grandezza  $\chi^2$  considerando:

- la discrepanza  $i$ -esima dei dati dalla legge che si traduce in

$$(y_i - y(x_i))^2 = (y_i - a - bx_i)^2 \quad (3.7)$$

- la varianza  $\sigma_{y_i}^2$  del dato  $i$ -esimo

Si ottiene:

$$\frac{(y_i - y(x_i))^2}{\sigma_{y_i}^2} \quad (3.8)$$

Per ora assumiamo che l'unica incertezza sia quella su  $y_i$ , risolveremo dopo il caso generale. Definisco  $\chi^2$  come:

$$\chi^2 := \sum_{i=0}^N \frac{(y_i - y(x_i))^2}{\sigma_{y_i}^2} \quad (3.9)$$

Ora tento di minimizzare il  $\chi^2$  in funzione dei parametri della legge  $y(x)$  nel caso lineare quindi rispetto  $a + bx$ .

$$\begin{cases} \frac{\partial \chi^2}{\partial a} = 0 \\ \frac{\partial \chi^2}{\partial b} = 0 \end{cases} \quad (3.10)$$

### 3.2.1 Caso di proporzionalità diretta

Supponiamo che la legge da adattare sia del tipo:

$$y(x) = bx \quad (3.11)$$

La nuova grandezza  $\chi^2$  sarà quindi

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - bx_i)^2}{\sigma_{yi}^2} \quad (3.12)$$

Che si traduce a una equazione risolvibile per valori di  $b$ . Minimizzando rispetto a  $b$  quindi:

$$0 = \frac{\partial \chi^2}{\partial b} = \sum_{i=1}^N \frac{1}{\sigma_{yi}^2} 2(y_i - bx_i)(-x_i) = 2b \sum_{i=1}^N \frac{x_i^2}{\sigma_{yi}^2} - 2 \sum_{i=1}^N \frac{x_i y_i}{\sigma_{yi}^2} \quad (3.13)$$

Da cui

$$b = \frac{\sum_{i=1}^N \frac{x_i y_i}{\sigma_{yi}^2}}{\sum_{i=1}^N \frac{x_i^2}{\sigma_{yi}^2}} \quad (3.14)$$

L'incertezza su  $b$  ovvero  $\sigma_b$  è:

$$\sigma_b^2 = \sum_{i=1}^N \left| \frac{\partial b}{\partial y_i} \right|^2 \sigma_{yi}^2 = \frac{1}{\left( \sum_{i=1}^N \frac{x_i^2}{\sigma_{yi}^2} \right)^2} \sum_{i=1}^N \left| \frac{x_i}{\sigma_{yi}^2} \right|^2 \quad (3.15)$$

Assumendo che le  $y_i$  siano statisticamente indipendenti

$$\sigma_{yi}^2 = \frac{1}{\left( \sum_{i=1}^N \frac{x_i^2}{\sigma_{yi}^2} \right)^2} \sum_{i=1}^N \left( \frac{x_i^2}{\sigma_{yi}^2} \right) \quad (3.16)$$

Quindi  $\sigma_b$  diventa

$$\sigma_b = \sqrt{\frac{1}{\sum_{i=1}^N \frac{x_i^2}{\sigma_{yi}^2}}} \quad (3.17)$$

Osservazioni:

- $\sigma_b \propto \sigma_y$
- $\sigma_b$  si abbassa la crescere del numero dei dati.

- $\sigma_b \propto (\text{discrepanza}^2 \text{ media degli } x_i \text{ rispetto all'origine})^{-1}$

Per tenere conto delle incertezze sulle  $x_i$  e per renderle confrontabili con le  $y_i$ :

- Propagare l'incertezza da  $x$  a  $y(x)$  se la legge è lineare si operi come su 2.5.1.
- Comporre l'incertezza su  $y_i$  ovvero:

$$\sigma_{yi \text{ tot}}^2 = \sigma_{yi \text{ originale}}^2 + \sigma_{yi \text{ trasferita}}^2 \quad (3.18)$$

Osservazioni:

- Per trasferire l'incertezza serve una stima preliminare di  $b$  o  $\left| \frac{dy}{dx} \right|_{x_i}$  che può venire da una stima tramite rette di massima o minima pendenza oppure da un fit preliminare senza considerare  $\sigma_{xi}$ .
- Basta una stima di  $\sigma_{yi \text{ tot}}$  con una due cifre significative.

### 3.2.2 Riepilogo procedura fit minimi quadrati

- stima preliminare di  $b$  o di  $\left| \frac{dy}{dx} \right|_{x_i}$  tramite:
  - fit preliminare
  - retta di massima e minima pendenza
- trasferire l'incertezza da  $x_i$  a  $y_i$  tramite

$$\sigma_{yi \text{ trasferito}} = |b| \sigma_{xi} \quad \sigma_{yi \text{ trasferito}} \simeq \left| \frac{dy}{dx} \right|_{x_i} \sigma_{xi} \quad (3.19)$$

- controllare se  $\sigma_{yi \text{ trasferito}}$  è trascurabile rispetto  $\sigma_{yi \text{ originale}}$ 
  - Se è trascurabile esegui il fit considerando solo gli errori nativi su  $y_i$
  - Se non è trascurabile rifaccio il fit da capo usando il nuovo  $\sigma_{tot}$

$$\sigma_{yi \text{ tot}}^2 = \sigma_{yi \text{ originale}}^2 + \sigma_{yi \text{ trasferita}}^2 \quad (3.20)$$

- Controllo che il risultato sia compatibile con il valore utilizzato per trasferire l'incertezza.
  - Se è compatibile tengo il risultato del fit.
  - Se non è compatibile trasferisco l'incertezza con il nuovo valore.

### 3.2.3 Valutazione della legge

Minimizzare il  $\chi^2$  permette di calcolare i parametri che meglio si adattano alla legge. Bisogna però verificare la bontà di questo adattamento. Se  $y(x)$  è corretta allora

$$m[y_i] = y(x_i) \quad (3.21)$$

e se la stima della varianza è corretta

$$m[\sigma_{yi}^2] = D[y_i] \quad (3.22)$$

Da cui

$$m[\chi^2] = \sum_{i=1}^N m \left[ \frac{(y_i - y(x_i))^2}{\sigma_{yi}^2} \right] \stackrel{(\dagger)}{=} \sum_{i=1}^N 1 = N \quad (3.23)$$

su  $(\dagger)$  si è utilizzato il fatto che:

$$m[(y_i - y(x_i))^2] = D[y_i] \quad (3.24)$$

In realtà  $m[\chi^2] = \text{gradi di libertà}$  cioè :

$$\nu := N - \text{numero parametri calcolati dai dati} \quad (3.25)$$

Il  $\chi^2$  atteso fluttua di:

$$\sigma[\chi^2] = \sqrt{2\nu} \quad (3.26)$$

sotto le ipotesi:

- $N$  dati indipendenti
- $\sigma_{yi}^2$  siano le varianze degli  $y_i$
- La legge è corretta:  $m[y_i] = y(x_i)$

allora il  $\chi^2$  ha distribuzione limite:

- $m[\chi^2] = \nu$
- $\sigma[\chi^2] = \sqrt{2\nu}$

## 3.3 Test del Chi2

Scopo: verificare come il  $\chi^2$  sperimentale sia compatibile con la distribuzione limite teorica

$$\chi_{teo}^2 \sim \nu \pm \sqrt{2\nu} \quad (3.27)$$

- Discutere la compatibilità calcolando la discrepanza:

$$|R| := |\chi_{oss}^2 - \nu| \leq k\sqrt{2\nu} \quad (3.28)$$

- La relazione fra il fattore di copertura  $k$  e la probabilità di falso allarme è complicata, dipende dal valore di  $\nu$
- Per  $\nu$  bassi può succedere che  $\nu - k\sqrt{2\nu} < 0$  e quindi oltrepassa il bordo del dominio
- Discutere la compatibilità in termini di intervallo corrispondente ad un certa probabilità fissata  $1 - \text{probabilità di falso allarme}$  utilizzando i quantili e i percentili della distribuzione  $\chi^2$ . A parità di scelta di probabilità di falso allarme complessiva, si ha la libertà di assegnare l'intervallo alla coda alta piuttosto che a quella bassa; si può fare come si vuole perfino stabilire un intervallo con un'unica soglia superiore<sup>1</sup>. Non è consigliabile in tutti i casi ripartire la probabilità di falso allarme in parti uguali.

### 3.3.1 Calcolo del Chi2 dal grafico

È bene controllare il risultato del  $\chi^2$  ad occhio quando le barre d'errore sono visibili misurando quanto vale la discrepanza della legge in termini di  $(\text{numero di sigma})^2$ . Nel caso in cui le incertezze non fossero visibili dal grafico è bene plottare i residui dei dati rispetto alla legge. Anche plottare in un istogramma i singoli contributi  $\chi_i^2$  dovuti ai singoli punti può risultare utile. Se viene riscontrato che uno o pochi dati contribuiscono in modo dominante al  $\chi_{oss}^2$  allora:

- Bisogna ricontrollare quelle misure e se non si può
- Considerare di escludere quelle misure a condizione che:
  - venga dichiarato esplicitamente.
  - si tratti di una misura riconducibile a qualche anomalia nell'esperimento e che non sia una caratteristica del fenomeno che sto studiando.

### 3.3.2 Test del Chi2: casi possibili

1.  $\chi_{oss}^2 \in \text{intervallo di compatibilità con la distribuzione teorica}$ . Cioè

$$\nu - k\sqrt{2\nu} \leq \chi_{oss}^2 \leq \nu + k\sqrt{2\nu} \quad (3.29)$$

oppure

$$\chi_{oss}^2 \in [\chi_{inf}^2; \chi_{sup}^2] \quad (3.30)$$

Accetto il risultato del fit:

- la legge è in accordo con i dati

---

<sup>1</sup>Consigliabile nel caso in cui  $\nu \lesssim 8$



- i parametri calcolati dal fit e le loro incertezze sono accettati

2.  $\chi_{oss}^2$  è troppo alto. Cioè

$$\chi_{oss}^2 > \nu + k\sqrt{2\nu} \quad (3.31)$$

oppure

$$\chi_{oss}^2 > \chi_{sup}^2 \quad (3.32)$$

- La legge è sbagliata?
- Sottostimo le incertezze di  $\sigma_{yi}^2$ ?

Non posso accettare i valori calcolati e i parametri, in particolare sono da rigettare le incertezze. Calcolo le incertezze a posteriori ovvero impongo

$$\chi_{oss}^2 = \nu \quad (3.33)$$

3.  $\chi_{oss}^2$  è troppo basso. Cioè

$$\chi_{oss}^2 < \nu - k\sqrt{2\nu} \quad (3.34)$$

oppure

$$\chi_{oss}^2 < \chi_{inf}^2 \quad (3.35)$$

La legge si adatta troppo ai dati ovvero la forma della legge è buona e quindi accetto i parametri ma metto in dubbio le incertezze: potrei aver sovrastimato  $\sigma_{yi}$ . Calcolo a posteriori i  $\sigma_{yi}$  che mettono a posto il  $\chi^2$ .

Sia nel caso 2 che nel caso 3 ricalcolo gli errori  $\sigma_{yi}^2$  a posteriori: se non si vuole rinunciare alla forma della legge bisogna convincere che i nuovi  $\sigma_{iy}^2$  a posteriori sono:

- Corretti, lo dimostro con un nuovo esperimento che li misuri direttamente.
- Plausibili, discutendo quantitativamente che il loro valore è fisicamente ragionevole.

In ogni caso, tutte le volte che accetto una determinazione a posteriori devo rifare il fit e prendere come risultati i nuovi valori dei parametri e relative incertezze. Non ha più alcun senso rifare il test del  $\chi^2$ .

### 3.3.3 Determinazione a posteriori delle incertezze

- Ipotizzando che  $\sigma_{yi}^2$  siano uguali impongo che

$$\frac{1}{\sigma_y^2} \sum_{i=1}^N (y_i - y(x_i))^2 = \nu \quad (3.36)$$

Ottengo che:

$$\sigma_{y \text{ a posteriori}}^2 = \frac{\sum_{i=1}^N (y_i - u(x_i))^2}{\nu} = \frac{\chi_{oss}^2}{\nu} \sigma_{yi \text{ a priori}}^2 \quad (3.37)$$

- Determinare un fattore moltiplicativo  $k$  uguale per tutte le varianze

$$\sigma_{yi \text{ a priori}}^2 \rightarrow k \sigma_{y \text{ a posteriori}}^2 \quad (3.38)$$

Qualora non fosse possibile considerare uguali le varianze sui dati

$$\frac{\chi_{oss}^2}{k} = \sum_{i=1}^N \frac{(y_i - y(x_i))^2}{k \sigma_{yi}^2} = \nu \quad (3.39)$$

Da cui

$$k = \frac{\chi_{oss}^2}{\nu} \quad (3.40)$$

- Altra variante possibile è

$$\sigma_{yi \text{ a priori}}^2 \rightarrow (\sigma_{yi}^2 + \sigma^2)_{a \text{ posteriori}} \quad (3.41)$$

Sostanzialmente aggiungo una varianza uguale a tutti i punti.

$$\sum_{i=1}^N \frac{(y_i - y(x_i))^2}{\sigma_{yi}^2 + \sigma^2} = \nu \quad (3.42)$$

Vado per tentativi nel trovare  $\sigma$

## 3.4 Esempi di fit

### 3.4.1 Legge costante

Supponiamo di adattare i dati sperimentali ad un legge costante del tipo

$$y = a \quad (3.43)$$

Il  $\chi^2$  diventa quindi

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - a)^2}{\sigma_i^2} \quad (3.44)$$

Minimizzo il  $\chi^2$  ovvero

$$0 = \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^N \frac{(y_i - a)}{\sigma_i^2} = -2 \left[ \sum_{i=1}^N \frac{y_i}{\sigma_i^2} - a \sum_{i=1}^N \frac{1}{\sigma_i^2} \right] \quad (3.45)$$

Da cui

$$a = \frac{\sum_{i=1}^N \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2}} \quad (3.46)$$

Mentre le incertezze su  $a$

$$\sigma_a = \sqrt{\frac{\sum_{i=1}^N \frac{1}{\sigma_i^4} \sigma_i^2}{\left( \sum_{i=1}^N \frac{1}{\sigma_i^2} \right)^2}} = \frac{1}{\sqrt{\sum_{i=1}^N \frac{1}{\sigma_i^2}}} \quad (3.47)$$

Questo risultato richiama la media pesata con pesi

$$w_i = \frac{1}{\sigma_i^2} \quad (3.48)$$

Abbiamo ottenuto infatti che la media pesata è il risultato del metodo dei minimi quadrati per questo è fondamentale eseguire sempre il test del  $\chi^2$  sulla media pesata prima di usare i risultati; in particolare se il  $\chi^2$  non è accettabile non posso usare il risultato della media pesata. Nel caso decidessi di tenere buona la legge  $y = a$  ma metto in dubbio le stime di incertezza:

- Non mi fido per nulla delle stime di  $\sigma_i^2$  a priori:
  - Calcolo la media campionaria:

$$m^* = a \quad (3.49)$$

non dipende dalle incertezze a priori

- Calcolo la deviazione

$$\tilde{\sigma} = \frac{1}{N-1} \sum_{i=1}^N (y_i - a)^2 \quad (3.50)$$

notare che al denominatore  $N - 1 = \nu$

Questo procedimento è identico ad ipotizzare varianze uguali e calcolarle a posteriori.

- Non posso considerare uguali le stime  $\sigma_i^2$  a priori:

- Posso tenere la formula della media pesata allargando tutte le incertezze di uno stesso fattore moltiplicativo

$$w_i = \frac{1}{k\sigma_i^2} \quad (3.51)$$

dovrò discutere della plausibilità delle nuove incertezze.

A differenza della procedura precedente questa stima richiede una discussione.

Parte II

Probabilità



## Capitolo 4

# Introduzione ai fenomeni aleatori

### 4.1 Fenomeni aleatori

I fenomeni aleatori o casuali si realizzano in modo non completamente predicibile, la trattazione deterministica è inapplicabile perché:

- Sussistono dei limiti pratici:
  - Difficoltà a ricostruire le condizioni iniziali, questo significa incertezza di predizione sull'evoluzione del fenomeno.
  - Errori di misura<sup>1</sup>.
  - Modello di valutazione troppo complesso<sup>2</sup>.
- La natura del fenomeno non è deterministica;
  - Meccanica quantistica<sup>3</sup>.
- Si è interessati a descrivere quantità medie del fenomeno
  - Meccanica statistica, proprietà macroscopiche di un sistema di molti corpi.

Si hanno a disposizione due approcci complementari:

- Teoria della probabilità: predico i possibili risultati a partire dai modelli teorici.
- Statistica: dalle osservazioni, risalgo ai possibili modelli.<sup>4</sup>

---

<sup>1</sup>Come gli errori di tipo A.

<sup>2</sup>Si pensi al lancio del dado, è improponibile creare un modello adeguato.

<sup>3</sup>Si pensi alla particella come funzione d'onda, non come punto materiale.

<sup>4</sup>Per esempio un FIT.

## 4.2 Teoria della probabilità

Definiamo:

- l'**esito** di una prova come: **risultato di un fenomeno aleatorio**.
- lo **spazio campionario**: generalmente indicato con  $\mathcal{S}$  come l'insieme di tutti gli esiti possibili. Lo spazio campionario può essere costituito da un numero di esiti:
  - finiti come  $N$  misure ripetute
  - infiniti numerabile come il lancio del dado finché non esce il numero 6
  - infinito non numerabile come il tempo necessario prima che una apparecchiatura si guasti
- l'**evento**: sottoinsieme di  $\mathcal{S}$  ovvero un insieme di esiti.

Supponendo  $A$  come evento valgono le seguenti affermazioni:

- Se  $A$  è un **evento certo** allora  $A := \mathcal{S}$
- Se  $A$  è un **evento impossibile** allora  $A := \phi$
- Se  $\bar{A}$  è un **evento complementare** di  $A$  allora  $\bar{A} := \mathcal{S} - A$

Definiamo le principali operazioni tra eventi:

- **Somma di eventi**: supponendo

$$C := A + B \quad (4.1)$$

l'evento somma si realizza quando si realizza  $A$  oppure quando si realizza  $B$ .

$$C = A \cup B \quad (4.2)$$

- **Prodotto di eventi**: supponendo

$$C := AB \quad (4.3)$$

l'evento prodotto si realizza quando si realizza  $A$  e quando si realizza  $B$ .

$$C = A \cap B \quad (4.4)$$

$A$  e  $B$  si dicono **incompatibili** se:

$$A \cup B = \phi \quad (4.5)$$



### 4.3 Definizione assiomatica della probabilità

La probabilità di un evento è un numero reale  $P$  che soddisfa tre assiomi:

1.

$$\forall A \quad P(A) \geq 0 \quad (4.6)$$

2.

$$P(\mathcal{S}) = 1 \quad (4.7)$$

3.

$$\forall A, B \text{ incompatibili} \quad P(A + B) = P(A) + P(B) \quad (4.8)$$

- È una definizione astratta ovvero ci lascia libertà nella assegnazione della probabilità agli eventi
- Dai tre assiomi si ricavano tutte le altre proprietà della probabilità:

—

$$\mathcal{S} = A + \bar{A} \quad (4.9)$$

Da cui

$$P(A + \bar{A}) = P(A) + P(\bar{A}) = 1 \quad (4.10)$$

Quindi è vero che

$$\forall A \quad P(A) \leq 1 \Rightarrow P \in [0, 1] \quad (4.11)$$

Per concludere

$$P(\bar{A}) = 1 - P(A) \quad (4.12)$$

—

$$A + \phi = A \quad (4.13)$$

Da cui

$$P(A) = P(A) + P(\phi) \Rightarrow P(\phi) = 0 \quad (4.14)$$

Ovvero la probabilità dell'evento impossibile è nulla.

—

$$A \subset B \quad (4.15)$$

Implica

$$B = A + (B - A) \Rightarrow P(B) \quad (4.16)$$

$$= P(A) + P(B - A) \geq 0 \Rightarrow P(A) \leq P(B) \quad (4.17)$$

### Probabilità della somma di eventi qualunque

La somma di eventi qualunque è così definita:

$$\forall A, B \subset \mathcal{S} \quad (4.18)$$

Vale che:

$$P(A + B) = P(A) + P(B) - P(AB) \quad (4.19)$$

#### 4.3.1 Interpretazione della probabilità classica o a priori

$$\forall A \subset \mathcal{S} \quad (4.20)$$

Vale che:

$$P(A) := \frac{\text{numero di esiti favorevoli}}{\text{numero di esiti possibili}} = \frac{m_A}{M_S} \quad (4.21)$$

- Viene detta a priori perché non richiede di compiere alcuna osservazione.
- Le ipotesi necessarie per applicare questa interpretazione sono:
  - Gli esiti devono essere **mutualmente esclusivi**<sup>5</sup>.
  - Gli esiti devono essere **equiprobabili**.
- Osserviamo che soddisfa i tre assiomi:

$$1. \quad \frac{m_A}{M_S} \geq 0 \quad (4.22)$$

$$2. \quad \frac{M_S}{M_S} = 1 \quad (4.23)$$

$$3. \quad \frac{m_A + m_B}{M_S} = \frac{m_A}{M_S} + \frac{m_B}{M_S} \quad (4.24)$$

#### 4.3.2 Interpretazione della probabilità frequentista o a posteriori

Si basa sul concetto di frequenza campionaria vista sugli istogrammi<sup>6</sup>.

$$\forall A \subset \mathcal{S} \quad (4.25)$$

Si definisce  $P_A^*$

$$P_A^* := \frac{\text{Numero di volte che si è realizzato } A}{\text{Numero di ripetizione dell'osservazione}} \quad (4.26)$$

---

<sup>5</sup>Si realizzano uno per volta.

<sup>6</sup>Si veda la sezione 1.4.6.

Bisogna specificare però che le ripetizioni devono essere effettuate nelle medesime condizioni e devono essere indipendenti inoltre, si osserva che  $P_A^*$  è una grandezza aleatoria ovvero è soggetta a fluttuazioni nel caso si cambi il set di ripetizioni. È facile intuire che vale la seguente relazione:

$$P(A) := \lim_{N \rightarrow \infty} P_A^* \quad (4.27)$$

Dove  $N$  è il numero di ripetizioni e il limite è inteso nel senso delle probabilità, non del limite puntuale ovvero:

$$\forall \epsilon > 0 \quad P(|P(A) - P_A^*| < \epsilon) \xrightarrow{N \rightarrow \infty} 1 \quad (4.28)$$

- Il limite nel senso delle probabilità è una convergenza molto più debole rispetto alla convergenza puntuale.
- Le fluttuazioni di  $P_A^*$  rispetto a  $P(A)$  rimangono possibili anche con  $N$  grande solo che diventano più improbabili.

Riprendendo la definizione:

- Viene detta a posteriori perché richiede una osservazione.
- Soddisfa i tre assiomi:

1.

$$P_A^* \geq 0 \Rightarrow P(A) \geq 0 \quad (4.29)$$

2.

$$P(\mathcal{S}) = P_{\mathcal{S}}^* = 1 \quad (4.30)$$

3.

$$\forall A, B \text{ incompatibili} \quad P_{A+B}^* = P_A^* + P_B^* \quad (4.31)$$

- Vi è una difficoltà operativa:  $N \rightarrow \infty$  non è una condizione realizzabile quindi bisogna accontentarsi dell'approssimazione:

$$P(A) \simeq P_A^* \quad (4.32)$$

- Il campo di applicabilità è molto più vasto rispetto alla interpretazione a priori.

### 4.3.3 Interpretazione soggettiva o Bayesiana della probabilità

L'interpretazione soggettiva implica la libera possibilità di dissentire in qualsiasi momento; bisogna quindi:

- Compiere assegnazioni di probabilità il più convenienti possibile per venire considerati dalla comunità scientifica.

- Mettere in campo tutte le proprie conoscenze e competenze.

L'interpretazione soggettiva è la più generale possibile:

- Non si può assegnare probabilità ai risultati di osservazioni come nel caso frequentista.
- Si può assegnare probabilità a *teoria, modelli, ipotesi*.
- Bisogna soddisfare i tre assiomi:

1. Banale:

$$\forall A \quad P(A) \geq 0 \quad (4.33)$$

2. Non tanto banale:

$$P(\mathcal{S}) = 1 \quad (4.34)$$

3. Operativamente non banale:

$$\forall A, B \text{ incompatibili} \quad P(A + B) = P(A) + P(B) \quad (4.35)$$

Bisogna suddividere  $\mathcal{S}$  in un set di eventi incompatibili rispettando il secondo assioma.

#### 4.3.4 Probabilità condizionata

Ipotizzando  $B$  un evento non impossibile, si dice **probabilità condizionata dell'evento  $A$  rispetto all'evento  $B$**  la probabilità che si realizzi  $A$  a condizione che si realizzi  $B$ . La probabilità di  $A$  condizionato  $B$  si indica con:

$$P(A|B) := \frac{P(AB)}{P(B)} \quad (4.36)$$

Ovvero:

- Considero  $B$  come nuovo evento certo ovvero come nuovo spazio campionario al posto di  $\mathcal{S}$ .
- Rinormalizzo la probabilità di  $AB$  su questo nuovo spazio campionario. Notare che la definizione trova senso se gli eventi  $A$  e  $B$  sono compatibili altrimenti è banalmente nulla inoltre soddisfa i tre assiomi:

1.

$$\forall A \quad P(A|B) \geq 0 \quad (4.37)$$

2.  $B$  è l'evento certo infatti:

$$P(B|B) = \frac{P(B)}{P(B)} = 1 \quad (4.38)$$

3.  $\forall A_1, A_2$  incompatibili

$$P(A_1 + A_2|B) = \frac{P((A_1 + A_2)B)}{P(B)} = \frac{P(A_1B + A_2B)}{P(B)} \quad (4.39)$$

$$\stackrel{(\dagger)}{=} \frac{P(A_1B)}{P(B)} + \frac{P(A_2B)}{P(B)} = P(A_1|B) + P(A_2|B) \quad (4.40)$$

Dove su  $(\dagger)$  si è sfruttato il fatto che  $A_1$  e  $A_2$  sono incompatibili

Se  $P(A) > 0$  simmetricamente

$$P(B|A) = \frac{P(AB)}{P(A)} \neq P(A|B) \quad (4.41)$$

Infatti il numeratore rimane identico, cambia la normalizzazione rispetto ad  $A$  o  $B$  da cui le seguenti uguaglianze:

$$P(AB) = P(A|B)P(B) = P(B|A)P(A) \quad (4.42)$$

### Teorema di Bayes

Il teorema di Bayes afferma che:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (4.43)$$

### Definizione di eventi indipendenti

Siano  $A$  e  $B$  eventi compatibili non impossibili.  $A$  e  $B$  sono **indipendenti**:

$$\iff P(A|B) = P(A) \quad (4.44)$$

$$\iff P(B|A) = P(B) \quad (4.45)$$

$$\iff P(AB) = P(A)P(B) \quad (4.46)$$

Si ricava da (4.45) è (4.44) che la probabilità di uno non è condizionata dalla probabilità che si realizzi l'altro. Da non confondere con eventi incompatibili:

$$A \cap B = \phi \quad (4.47)$$

### Esempi di probabilità condizionata

- Supponiamo di prendere in esame il lancio di un dado cubico. Definiamo i seguenti eventi:

- $A := \text{esce un numero} < 6$ .  $P(A) = \frac{5}{6}$
- $B := \text{esce un numero pari}$ .  $P(B) = \frac{1}{2}$

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3} \quad (4.48)$$

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{2}{6}}{\frac{5}{6}} = \frac{2}{5} \quad (4.49)$$

- Supponiamo di dover discutere la probabilità di guasto di un sistema "B" composto da sottosistemi  $B_i$  che si attivano uno alla volta in maniera non predicibile ma in sequenza.

- Interpreto  $B_i$  come eventi di mutualmente incompatibili
- Associa a  $B_i$  una probabilità:

$$P(B_i) := \text{Frazione media di tempo in cui } B_i \text{ è operativo} \quad (4.50)$$

Dal teorema di Bayes ho che:

$$P(\text{Guasto } B_i) = P(\text{Guasto} | B_i)P(B_i) \quad (4.51)$$

E quindi:

$$P(\text{Guasto } B) = \sum_{i=1}^N P(\text{Guasto } B_i) \quad (4.52)$$

- Supponiamo ora di dover discutere il problema precedente a condizione però che i sottoinsiemi funzionino in parallelo continuamente.

$$P(\text{Guasto } B_i) = P(\text{Guasto} | B_i)P(B_i) \stackrel{(\dagger)}{=} P(\text{Guasto} | B_i) \quad (4.53)$$

Su  $(\dagger)$  si è sfruttato il fatto che  $P(B_i) = 1$

$$P(\text{Guasto } B) = \sum_{i=1}^N P(\text{Guasto } B_i) - \sum_{\text{coppie } i \neq j} P(\text{Guasto } B_i \text{ Guasto } B_j) \quad (4.54)$$

$$+ \sum_{\text{triple } i \neq j \neq k} P(GB_i GB_j GB_k) + \dots \quad (4.55)$$

Se i sottoinsiemi  $B_o$  hanno guasti indipendenti e se  $P(GB_i) \ll 1$  allora si può approssimare:

$$\begin{aligned} P(\text{Guasto } B) &\simeq \\ &\simeq \sum_{i=1}^N P(\text{Guasto } B_i) - \sum_{\text{coppie } i \neq j} P(\text{Guasto } B_i)P(\text{Guasto } B_j) \end{aligned} \quad (4.56)$$

### 4.3.5 Probabilità condizionata al prodotto di eventi

Prendiamo in considerazione i tre eventi realizzabili  $A_1, A_2, A_3$ . Vale l'uguaglianza

$$P(A_3|A_2A_1) = \frac{P(A_1A_2A_3)}{P(A_2A_1)} \quad (4.57)$$

Consideriamo ora il prodotto dei tre eventi:

$$P(A_3A_2A_1) = P(A_3|A_2A_1)P(A_1A_2) = P(A_3|A_2A_1)P(A_2|A_1)P(A_1) \quad (4.58)$$

Per induzione si può estendere al prodotto di  $N$  eventi  $A_1, A_2, \dots, A_N$ :

$$\begin{aligned} P(A_1 \dots A_N) &= \\ &= P(A_N|A_1 \dots A_{N-1})P(A_{N-1}|A_1 \dots A_{N-2}) \dots P(A_2|A_1)P(A_1) \end{aligned} \quad (4.59)$$

### Esempio: restrizione di 5 carte dal mazzo di 52 senza riposizione

Definiamo l'evento  $A$  come:

$$A := \text{Estreggo 5 carte di cuori} \quad (4.60)$$

Possiamo quindi vedere  $A$  come:

$$A = A_1A_2A_3A_4A_5 \quad (4.61)$$

Indicando con  $A_i$  l'estrazione  $i$ -esima di cuori.

$$P(A) = P(A_1)P(A_2|A_1)P(A_3|A_2A_1) \dots P(A_5|A_4A_3A_2A_1) \quad (4.62)$$

Se l'estrazione fosse con ripetizione allora gli eventi  $A_i$  sarebbero indipendenti tra loro e quindi:

$$A = \left(\frac{1}{4}\right)^5 \sim 10^{-3} \quad (4.63)$$

Considerando invece la non riposizione si ottiene:

$$\frac{13}{52} \frac{12}{51} \frac{11}{50} \frac{10}{49} \frac{9}{48} \sim 5 \times 10^{-4} \quad (4.64)$$

### 4.3.6 Teorema della probabilità totale

Sia  $\mathcal{S}$  lo spazio campionario e siano  $A_i$   $n$  eventi tali che:

- $A_i$  incompatibile con  $A_j \quad \forall i \neq j$
- $A_i$  compongano  $\mathcal{S}$  ovvero  $\mathcal{S} = A_1 \cup A_2 \cup \dots \cup A_n$

Allora  $\forall B \subset \mathcal{S}$

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i) \quad (4.65)$$

Dimostrazione:

$$B = \sum_{i=1}^n BA_i \quad (4.66)$$

Con  $BA_i$  incompatibile con  $BA_j \quad \forall i \neq j$

L'impiego principale si ha con il teorema di Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_{i=1}^n P(B|A_i)P(A_i)} \quad (4.67)$$

#### 4.4 Applicazione del teorema di Bayes nell'interpretazione soggettiva della probabilità

Lo scopo è assegnare la probabilità a posteriori di ipotesi/teorie tenendo conto dei risultati di una osservazione sperimentale per decidere quali teorie considerare e quali scartare.

$$\begin{aligned} P(\text{Teoria}_k | \text{Osservazione}) &= \\ &= P(\text{Osservazione} | \text{Teoria}_k) P(\text{Teoria}_k) \frac{1}{P(\text{Osservazione})} \end{aligned} \quad (4.68)$$

$P(\text{Teoria}_k)$ : probabilità a priori della  $\text{Teoria}_k$  come ad esempio:

- $\text{Teoria}_0(a, b)$ : Legge lineare  $y = a + bx$
- $\text{Teoria}_1(A, B)$ : Legge esponenziale  $\log y = A + B \log x$

Sono a priori (i pregiudizi) dello scienziato assegnati sulla base delle conoscenze precedenti all'osservazione: **Soggettivo**.

$P(\text{Osservazione} | \text{Teoria}_k)$ : probabilità di ottenere i risultati dell'esperimento se la  $\text{Teoria}_k$  è vera: ogni  $\text{Teoria}_k$  prevede una distribuzione limite dei possibili risultati delle osservazioni da cui si può calcolare la probabilità della osservazione che si è verificata: **Oggettivo**.

$P(\text{Teoria}_k | \text{Osservazione})$ : probabilità a posteriori della  $\text{Teoria}_k$  dopo aver ottenuto l'osservazione. Se ci si dovesse accontentare di un numero proporzionale alla probabilità si può evitare di calcolare il fattore di normalizzazione. Ad esempio:

- Se fosse sufficiente costruire una classifica delle  $\text{Teorie}_k$



- Selezione bayesiana del modello/teoria fra due teorie alternative:

$$\frac{P(\text{Teoria}_k | Oss)}{P(\text{Teoria}_l | Oss)} = \frac{P(Oss | \text{Teoria}_k)}{P(Oss | \text{Teoria}_l)} \frac{P(\text{Teoria}_k)}{P(\text{Teoria}_l)} \quad (4.69)$$

Dove:

$$\frac{P(\text{Teoria}_k | Oss)}{P(\text{Teoria}_l | Oss)} \quad (4.70)$$

è il rapporto delle probabilità a posteriori:

- $\gg 1$ : fiducia maggiore nella *Teoria<sub>k</sub>*
- $\simeq 1$ : considero le teorie equivalenti
- $\ll 1$ : fiducia maggiore nella *Teoria<sub>l</sub>*

$$\frac{P(Oss | \text{Teoria}_k)}{P(Oss | \text{Teoria}_l)} \quad (4.71)$$

è il Bayes factor ovvero quantifica come l'evidenza sperimentale cambia il rapporto dei priori:

- Gli esperimenti informativi hanno bayes factor:  $\ll 1$  o  $\gg 1$ , gli esperimenti non informativi hanno bayes factor  $\simeq 1$
- Il Bayes factor si usa al posto del test delle ipotesi che sono propri della interpretazione frequentistica
- è facilissimo comporre assieme l'evidenza postata da osservazioni indipendenti: l'osservazione complessiva è l'evento prodotto delle singole osservazioni:

$$Oss_{tot} = \prod P_{oss_i} \quad (4.72)$$

Se sono indipendenti allora

$$P(Oss_{tot} | Teo_k) = \prod P(Oss_i | Teo_k) \quad (4.73)$$

e quindi il Bayes factor totale:

$$\prod P_{\text{Bayes factor singole osservazioni}} \quad (4.74)$$

$$\frac{P(\text{Teoria}_k)}{P(\text{Teoria}_l)} \quad (4.75)$$

Questo termine di normalizzazione è uguale per tutte le *Teorie<sub>k</sub>*. Legge di probabilità totale:

$$P(Oss) = \sum_k P(Oss | Teo_k) P(Teo_k) \quad (4.76)$$

Si sommano i contributi di tutte le possibili *Teorie<sub>k</sub>* alternative pesando con i priori. Può essere molto difficile da calcolare.

#### 4.4.1 Esempio: individui sani/malati

Supponiamo il lo spazio campionario composto da una popolazione di individui che a loro volta possono essere sani o malati. Tramite un test di diagnostica si ottengono due risultati possibili:

- +: c'è problema
- -: non c'è problema

Bisogna conoscere com'è collaudato il test ovvero determinare due parametri:  $P(+|M)$  e  $P(+|Sani)$ . Dai test clinici risulta:

	Sani	M
+	0.5%	99%
-	99.5%	1%
	100%	100%

Dove 0.5% è la probabilità di falso allarme mentre 1% è la probabilità di falso disallarme.

$$P(M|+) = \frac{P(+|M)P(M)}{P(+)} = \frac{P(+|M)P(M)}{P(+|M)P(M) + P(+|S)P(S)} \quad (4.77)$$

Dove  $P(M)$  e  $P(S)$  sono a priori.

- Ipotesi tranquilla:

$$\begin{cases} P(M) = 10^{-3} \\ P(Sani) = 99.9\% \end{cases} \quad (4.78)$$

$$P(M|+) = \frac{0.99 \cdot 10^{-3}}{0.99 \cdot 10^{-3} + 5 \cdot 10^{-3} \cdot 0.999} \simeq \frac{1}{6} \quad (4.79)$$

Il test non è risolutivo.

$$P(S|+) = 1 - P(M|+) = \frac{5}{6} \quad (4.80)$$

- Ipotesi di epidemia:

$$\begin{cases} P(M) = 10\% \\ P(Sani) = 90\% \end{cases} \quad (4.81)$$

$$P(M|+) \simeq 96\% \quad (4.82)$$

$$P(S|M) \simeq 4\% \quad (4.83)$$

È un risultato già più conclusivo.

- Importanza dei priori sul risultato

- $P(M|+)$  e  $P(Sani|+)$  ovvero i posteriori riassumono la fiducia dopo aver eseguito l'osservazione.

I posteriori:

- dipendono fortemente dai priori
- devono diventare i nuovi priori per l'osservazione successiva.

Siano  $Test_1$  e  $Test_2$  due esperimenti di risultati possibili rispettivamente:  $\{+1, -1\}$  e  $\{+2, -2\}$  da cui l'osservazione complessiva è :

$$\begin{Bmatrix} +1+2 & +1-2 \\ -1+2 & -1-2 \end{Bmatrix} \quad (4.84)$$

Applico il teorema di Bayes sull'osservazione complessiva:

$$P(M|+1+2) := \frac{P(+1+2|M)P(M)}{P(+1+2)} \stackrel{(\dagger)}{=} \frac{P(+2|M+1)}{P(+2|+1)} \frac{P(+1|M)P(M)}{P(+1)} \quad (4.85)$$

Dove su  $(\dagger)$  è stato sfruttato il fatto che:

$$P(+1+2|M) = P(+1|M+1)P(+1|M) \quad (4.86)$$

$$P(+1+2) = P(+2|+1)P(+1) \quad (4.87)$$

Dalla (4.85)

$$\frac{P(+1|M)P(M)}{P(+1)} = P(M|+1) \quad (4.88)$$

Ovvero i posteriori della osservazione precedente.

$$P(M|+1+2) = \frac{P(+2|M)}{P(+2)} P(M+1) \quad (4.89)$$

Dove

$$\frac{P(+2|M)}{P(+2)} = \frac{P(+2|M+1)}{P(+2|+1)} \quad (4.90)$$

Il teorema di Bayes dell'osservazione complessiva si ricava sempre dalle definizioni di probabilità condizionata applicate a prodotto di eventi:

$$P(M|+1+2) = \frac{P(+1+2|M)}{P(+1+2)} = \frac{P(+1+2|M)P(M)}{P(+2|+1)P(+1)} \quad (4.91)$$

Dal punti di vista della Bayesian model selection

$$\frac{P(M|+)}{P(\mathcal{S}|+)} = \frac{P(+|M)}{P(+|S)} \frac{P(M)}{P(S)} \quad (4.92)$$

Dove:

$$\frac{P(M|+)}{P(\mathcal{S}|+)} \quad (4.93)$$

- $\ll 1 \Rightarrow$  Ipotesi  $M$  preferita
- $\gg 1 \Rightarrow$  Ipotesi  $S$  preferita

$$\frac{P(+|M)}{P(+|S)} \quad (4.94)$$

bayes factor ( $\simeq 200$  per l'esempio precedente)

$$\frac{P(M)}{P(S)} \quad (4.95)$$

Rapporto dei priori.

Per eseguire più osservazioni indipendenti:

$$\frac{P(M|+1+2)}{P(S|+1+2)} = \frac{P(M+1+2)/P(+1+2)}{P(s+1+2)/P(+1+2)} = \frac{P(+1+2|M)P(M)}{P(+1+2|S)P(S)} \quad (4.96)$$

$$= \frac{P(+1|M)}{P(+1|S)} \frac{P(+2|M)}{P(+2|S)} \frac{P(M)}{P(S)} \quad (4.97)$$

Dove

$$\frac{P(+1|M)}{P(+1|S)} \frac{P(+2|M)}{P(+2|S)} \quad (4.98)$$

È il bayes factor complessivo di esperimenti indipendenti, è semplicemente il prodotto del bayes factors dei singoli esperimenti.

## Capitolo 5

# Calcolo combinatorio e prove ripetute

### 5.1 Calcolo combinatorio

Lo scopo del calcolo combinatorio, nel nostro caso, è di saper calcolare il numero di esiti che appartengono ad uno spazio campionario:

- interpretazione classica della probabilità
- Applicazioni alla prove ripetute: esperimenti condotti nelle medesime condizioni e indipendenti tra loro.

Vi sono applicazioni importanti del calcolo combinatorio in meccanica statistica.

#### 5.1.1 Introduzione

Consideriamo lo spazio campionario  $\mathcal{S} = \bigcup_{i=1}^N S_i$  con  $\{S_i\}$  eventi incompatibili considerando  $E_i$  evento di  $S_i$ . L'evento generale in  $\mathcal{S}$  è

$$E = \{E_1, E_2, E_3, \dots, E_N\} \quad (5.1)$$

Quanti sono gli eventi possibili di  $\mathcal{S}$ ?

$$N_{\mathcal{S}} = N_{S1} \times N_{S2} \times N_{S3} \times N_{S4} \times \dots \times N_{SN} = \prod_{i=1}^N N_{Si} \quad (5.2)$$

#### Esempi

- Consideriamo i possibili numeri di telefono a 6 cifre:

$$N_{Si} = 10 \quad \Rightarrow \quad N_{\mathcal{S}} = 10^6 \quad (5.3)$$

- Consideriamo le possibile triplette di basi azotate<sup>1</sup> ovvero adenina guanina timina citosina

$$N_S = 4^3 = 64 \quad (5.4)$$

- Consideriamo le possibile combinazioni di 20 aminoacidi per fare una proteina piccola ( $\sim 10^2$  aminoacidi)

$$N_s \simeq 20^{100} = 10^{100} 2^{100} = 10^{100} 10^{(\log_{10} 2)100} = 10^{100} 10^{30} \quad (5.5)$$

### 5.1.2 Permutazioni

Per **permutazione** di  $N$  oggetti si intende il numero di modi diversi in cui si può ordinare  $N$  oggetti.

$$P_N := N(N-1)(N-2) \dots 1 = N! \quad (5.6)$$

Il primo termine del fattoriale indica in quanti modi diversi si può prendere  $N$  oggetti, il secondo termine indica i modi per prendere  $N-1$  oggetti e così via.

### 5.1.3 Disposizioni

Per **disposizioni** di  $N$  oggetti *di classe*  $K$  o *da*  $K$  *a*  $K$  si indica il numero di modi in cui è possibile formare un gruppo di  $K$  oggetti a partire da  $N$ , contando come diverse le realizzazioni che differiscono anche solo per l'ordine dei  $K$  oggetti.

$$D_{N,K} := N(N-1)(N-2) \dots (N-K+1) = \frac{N!}{(N-K)!} = \frac{P_N}{P_{N-K}} \quad (5.7)$$

Il valore minimo delle disposizioni si ha per  $K=0$ :

$$D_{N,0} = 1 \quad (5.8)$$

Per  $K=1$

$$D_{N,1} = N \quad (5.9)$$

Il valore massimo si ha per  $K=N$

$$D_{NN} = P_N = N! \quad (5.10)$$

Quando si fa uso delle disposizioni bisogna attenersi a due condizioni:

- considerare gli oggetti distinguibili tra loro
- considerare rilevante l'ordine di comparizione/realizzazione

---

<sup>1</sup>Sono i caratteri con cui è scritto il nostro gene

**Esempio**

Il numero di possibilità della graduatoria di un podio di una gara con  $N$  partecipanti

$$D_{N,3} = N(N-1)(N-2) \quad (5.11)$$

Per esempio se la gara fosse da 8:

$$D_{8,3} = 336 \quad (5.12)$$

**5.1.4 Combinazioni**

Per combinazioni di  $N$  oggetti *di classe*  $K$  o *da*  $K$  a  $K$  si indica il modo in cui si può comporre un gruppo di  $K$  oggetti a partire da  $N$ , considerando equivalenti i modi che differiscono soltanto per l'ordine all'interno del gruppo di  $K$ .

$$C_{N,K} := \frac{D_{N,K}}{P_K} = \frac{N!}{K!(N-K)!} \quad (5.13)$$

$$C_{N,K} := \binom{N}{K} \quad (5.14)$$

La definizione è simmetrica per scambio tra  $K$  e  $N-K$ :

$$C_{N,K} = C_{N,N-K} \quad (5.15)$$

$$C_{N,0} = 1 \quad C_{N,N} = 1 \quad (5.16)$$

$C_{N,K}$  ha massimo per  $K \sim \frac{N}{2}$  Le combinazioni si utilizzano quando:

- Si estraggono oggetti indistinguibili tra loro
- L'applicazione non guarda l'ordine di estrazione

Ovviamente, il coefficiente binomiale è definito sempre per  $K \leq N$

**Esempi**

- Tombola: 90 oggetti, semplifichiamo fissando  $M$  estrazioni. Che probabilità ho di fare cinquina con  $M \geq 5$ ? Non interessa l'ordine delle estrazioni, lo spazio campionario è formato da  $C_{90,M}$  esiti. Quanti esiti sono favorevoli? Fare cinquina implica bloccare 5 numeri del gruppo di  $M$  numeri  $\Rightarrow$  numero di modi in cui i restanti  $M-5$  possono essere estratti:

$$C_{90-5, M-5} = \binom{90-5}{M-5} \quad (5.17)$$

$$P_{Cinquina}(M) = \frac{\binom{90-5}{M-5}}{\binom{90}{M}} \quad (5.18)$$

- Modi diversi per formare i gruppi di laboratorio da 3 supponendo 75 studenti. Il primo gruppo sarà composto così:

$$C_{75,3} = \frac{75 \cdot 74 \cdot 73}{3!} \quad (5.19)$$

Il secondo gruppo

$$C_{72,3} = \frac{72 \cdot 71 \cdot 70}{3!} \quad (5.20)$$

E via così tutti gli altri, si ottiene quindi:

$$\frac{75}{3!^{25}} \quad (5.21)$$

Poiché non interessa l'ordine con cui vengono formati i gruppi si corregge la precedente così da ottenere:

$$\frac{75}{3!^{25} 25!} \sim 6 \times 10^{64} \quad (5.22)$$

## 5.2 Applicazione alle prove ripetute

Applicheremo adesso il calcolo combinatorio alle prove ripetute, in particolare, a prove ripetute indipendenti di una osservazione in condizioni identiche.

### 5.2.1 Distribuzione di probabilità binomiale o di Bernoulli

Il modo più semplice di classificare il risultato di una singola prova è di considerare lo spazio campionario  $S_1$ :

$$S_1 := A + \bar{A} \quad (5.23)$$

Sulla **singola prova**  $P(A)$  è fissata e  $P(\bar{A}) = 1 - P(A)$ . Su **più prove ripetute** abbiamo le seguenti relazioni:

- $\mathcal{S} = \bigcup_i S_i$  con  $S_i$  spazio campionario della prova i-esima
- Il possibile risultato è una n-tupla<sup>2</sup> ad esempio:

$$(A, A, \bar{A}, A, \bar{A}, \dots) \quad (5.24)$$

Abbiamo quindi  $2^N$  possibili esiti.

- Per calcolare la probabilità che si sia verificato l'esito  $A$   $K$  volte su  $N$  tentativi è:

$$P(A)^K P(\bar{A})^{N-K} \quad (5.25)$$

---

<sup>2</sup>Per tupla si intende una collezione di  $N$  oggetti



Qual è quindi la probabilità di ottenere  $K$  volta  $A$  su  $N$  tentativi?  $K$  su  $N$  *successi* è un evento che raccoglie tutti gli esiti in cui  $A$  si verifica  $K$  volte:

$$P(K \text{ successi su } N) = P\left(\sum \text{Esiti corrispondenti}\right) \quad (5.26)$$

$$= \sum P(\text{Esito in cui } A \text{ si è verificato } K \text{ volte}) \quad (5.27)$$

$$= [P(A)^K P(\bar{A}^{N-K})](\text{numero esiti corrispondenti a } K \text{ successi su } N) \quad (5.28)$$

A patto che:

- Gli eventi  $A$  siano indistinguibili tra loro
- La domanda si riferisce a  $K$  *successi su*  $N$  e non fa alcun riferimento all'ordine delle estrazioni

Il numero di esiti corrispondenti è una distribuzione di tipo binomiale quindi:

$$P(K \text{ successi su } N; N, P(A)) = \binom{N}{K} P(A)^K P(\bar{A})^{N-K} \quad (5.29)$$

Il che ci porta ad introdurre la distribuzione di probabilità binomiale o di Bernoulli ovvero la (5.29). La distribuzione di probabilità binomiale è una funzione di una variabile discreta  $K$  e dipende da 2 parametri:

- $N :=$  Numero di prove
- $P(A) :=$  Probabilità di successo nella singola prova

### 5.2.2 Distribuzione di probabilità multinomiale

Nel caso invece si volesse considerare più di 2 eventi alternativi in ogni singola prova ovvero:

$$A_i \cap A_j = \phi \quad \forall i \neq j \quad (5.30)$$

$m$  eventi e

$$\bigcup_{i=1}^m = \mathcal{S} \quad (5.31)$$

qual è la probabilità che si verifichino  $K_1$  volte  $A_1$ ,  $K_2$  volte  $A_2$  e così via su  $N$  tentativi?

- Un esito elementare corrispondente a  $K_1$  volte  $A_1 \dots$ , ha possibilità:

$$P(A_1)^{K_1} P(A_2)^{K_2} \dots P(A_m)^{K_m} \quad (5.32)$$

Dove

$$\sum_{i=1}^m K_i = N \quad (5.33)$$

- Il numero di esiti corrispondenti ad una n-tupla  $K_1, K_2, K_3 \dots K_m$  si contano con le combinazioni:

$$\binom{N}{K_1} \binom{N-K_1}{K_2} \binom{N-K_1-K_2}{K_3} \dots \binom{K_m}{K_m} \quad (5.34)$$

$$= \frac{N!}{(N-K_1)!K_1!} \frac{(N-K_1)}{(N-K_1-K_2)!K_2!} \frac{(N-K_1-K_2)!}{(N-K_1-K_2-K_3)!K_3!} \dots 1 \quad (5.35)$$

$$= \frac{N!}{\prod_i^m K_i!} \quad (5.36)$$

Componendo questi due risultati si ottiene la distribuzione multinomiale definita come:

$$P(K_1, K_2, \dots K_m; N, P(A_1), P(A_2), \dots P(A_m)) := \frac{N!}{\left(\prod_{i=1}^m K_i\right)!} \prod_{i=1}^m P(A_i)^{K_i} \quad (5.37)$$

Questa distribuzione dipende da più variabili  $K_1, K_2 \dots$  ed è funzione dei parametri:

- $N :=$  Numero di prove
- $P(A_i) :=$  Probabilità che  $A_i$  si verifichi

$K$  e  $K_1, K_2, \dots, K_m$  sono esempi di variabile aleatorie

## Capitolo 6

# Variabili aleatorie

Per variabile aleatoria si intende una grandezza che assume valori casuali in maniera non predicibile.

### 6.1 Distribuzione di probabilità di una variabile aleatoria discreta

Una variabile aleatoria discreta è tale se:

- il dominio è un insieme di valori discreti
- in numero di valori può essere finito o infinito

Misure dominate da errori risoluzione sono variabili aleatorie discrete.

La distribuzione di probabilità di una variabile aleatoria discreta si ottiene associando ad ogni possibile valore della variabile aleatoria una probabilità, soddisfacendo la condizione di normalizzazione per qualunque distribuzione:

$$P(\mathcal{S}) = 1 \Leftrightarrow \sum_{k \in \text{Dominio}} P(K) = 1 \quad (6.1)$$

In caso contrario non soddisferebbe il secondo assioma e quindi non sarebbe una probabilità.

Si dice  $F(K)$  funzione cumulativa di probabilità della variabile aleatoria  $K$  se:

$$F(K') := P(K \leq K') = \sum_{K \leq K'} P(K) \quad (6.2)$$

Osserviamo immediatamente che:

- $F(K)$  è una funzione a gradini dove  $\text{gradino}_K = P(K)$  ovvero la probabilità di ottenere il valore  $K$
- $F(K)$  è una funzione strettamente crescente in particolare:

- $F(K') = 0$  se  $K' < \min\{K\}$
- $F(K') = 1$  se  $K' \geq \max\{K\}$

Che risponde a pieno in quanto distribuzione di variabile aleatoria è la binomiale

### 6.1.1 Distribuzione binomiale

La distribuzione binomiale è una distribuzione di variabile aleatoria discreta  $K \in \{0, 1, \dots, N\}$  ed è definita:

$$P(K; N, P(A)) := \frac{N!}{K!(N-K)!} P(A)^K P(\bar{A})^{N-K} \quad (6.3)$$

Associando  $p = P(A)$  e  $q = 1 - p$  si può scrivere in forma compatta come:

$$P(K; N, p) = \frac{N!}{K!(N-K)!} p^K q^{N-K} \quad (6.4)$$

Osserviamo che verifica la condizione di normalizzazione in quando

$$\sum_{K=0}^N \binom{N}{K} p^K q^{N-K} = P(p+q)^N = 1 \quad (6.5)$$

La speranza matematica invece è:

$$mK := \langle K \rangle = Np \quad (6.6)$$

Infatti:

$$\langle K \rangle = \sum_{K=1}^N K \binom{N}{K} p^K q^{N-K} = Np \sum_{K=1}^N \frac{(N-1)!}{(K-1)!(N-K)!} p^{K-1} q^{N-K} \quad (6.7)$$

$$\stackrel{(\dagger)}{=} Np \sum_{S=0}^{N-1} \frac{(N-1)!}{S!(N-1-S)!} p^S q^{N-1-S} \quad (6.8)$$

$$\stackrel{(\ddagger)}{=} Np \quad (6.9)$$

Dove su  $(\dagger)$  ci si è serviti della sostituzione  $S \stackrel{(S)}{=} K-1$  e dell'uguaglianza  $N-K = N-1-(K-1) = N-1-S$  sull'esponente di  $q$ .

Su  $(\ddagger)$  invece si è osservato che il termine della binomiale è la condizione di normalizzazione e quindi unitaria.

Possiamo quindi osservare che se consideriamo  $P_A^* = \frac{K}{N}$  otteniamo che:

$$m[P_A^*] = m\left[\frac{K}{N}\right] = \frac{1}{N} m[K] = \frac{Np}{N} = p \quad (6.10)$$

Cioè la media della frequenza campionaria di  $A$  è la probabilità di realizzazione di  $A$  nella singola prova.

La varianza della variabile aleatoria binomiale  $D[K]$  invece si calcola:

$$D[K] \stackrel{(\dagger)}{=} m[(k - m[k])^2] = m[K^2] - m^2[K] = Npq \quad (6.11)$$

Dove su  $(\dagger)$  ci si è ricondotti al risultato del capitolo 1.3.3

$$m[K^2] = \sum_{K=0}^N K^2 P(K; N, p) = \sum_{K=1}^N K \frac{N!}{(K-1)!(N-K)!} p^K q^{N-K} \quad (6.12)$$

$$\stackrel{(\dagger)}{=} Np \sum_{S=0}^{N-1} (S+1) \frac{(N-1)!}{S!(N-1-S)!} p^S q^{N-1-S} \quad (6.13)$$

$$= Np \left\{ \sum_{S=0}^{N-1} S \binom{N-1}{S} p^S q^{N-1-S} + \sum_{S=0}^{N-1} \binom{N-1}{S} p^S q^{N-1-S} \right\} \quad (6.14)$$

$$\stackrel{(\dagger)}{=} Np \{(N-1)p + 1\} = Np\{Np + 1 - p\} \quad (6.15)$$

Dove su  $(\dagger)$  ci si è serviti della sostituzione  $S \stackrel{(S)}{=} K - 1$

Su  $(\dagger)$  invece si è sfruttato il fatto che il primo termine è  $m[S] = (N-1)p$  mentre il secondo è la condizione di normalizzazione

In conclusione si ottiene che:

$$D[K] = Np\{Np + q\} - (Np)^2 = Npq \quad (6.16)$$

In termini di incertezza relativa:

$$\frac{\sigma[K]}{\langle k \rangle} = \frac{\sqrt{Npq}}{Np} = \frac{\sqrt{q}}{\sqrt{Np}} \quad (6.17)$$

Si osserva che l'incertezza è  $\propto \sqrt{N}$  e quindi la binomiale si stringe al crescere di  $N$  in termini di incertezza relativa

### Applicazioni agli istogrammi

Nel caso degli istogrammi è possibile applicare la distribuzione binomiale a patto che vi siano  $N$  misure indipendenti interpretando il conteggio di ogni bin come una variabile aleatoria binomiale del tipo:

$$P(n_j^*; N, p) \quad (6.18)$$

In particolare lo stimatore di speranza matematica  $m[n_j^*]$  è per definizione:

$$m[n_j^*] = Np_j \quad (6.19)$$

Mentre lo stimatore di scarto tipo è:

$$\sigma[n_j^*] = \sqrt{Np_j(1-p_j)} \stackrel{(\dagger)}{\simeq} \sqrt{Np_j} = \langle n_j^* \rangle \quad (6.20)$$

Dove su  $(\dagger)$  si è ipotizzato di avere sufficienti bin tali che:  $p_j \ll 1$

Si conclude quindi con:

$$n_j^* = Np_j \pm \sqrt{Np_j} = \langle n_j^* \rangle \pm \sqrt{\langle n_j^* \rangle} \quad (6.21)$$

Ovvero l'intervallo d'incertezza tipo che ci si aspetta sul conteggio di ogni bin dell'istogramma<sup>1</sup>

Per istogrammi normalizzati in altezza ricordiamo che:

$$p_j^* = \frac{n_j^*}{N} \quad (6.22)$$

Da cui si ottiene rispettivamente la speranza matematica e lo scarto tipo:

$$m[p_A^*] = P(A) \quad (6.23)$$

$$\sigma[p_A^*] \propto \frac{1}{\sqrt{N}} \xrightarrow{N \rightarrow \infty} 0 \quad (6.24)$$

Risulta notevole che lo scarto si riduca in proporzione alla radice di  $N$ , infatti, questo risultato è una delle tante versioni della legge dei grandi numeri per cui

$$P(A) = \lim_{N \rightarrow \infty} p_A^* \quad (6.25)$$

Da interpretare però nel senso della probabilità ovvero una convergenza molto più debole cioè:

$$\forall \epsilon > 0 \quad P(|p_A^* - P(A)| > \epsilon) \xrightarrow{N \rightarrow \infty} 0 \quad (6.26)$$

Cioè si tratta di convergenza di  $p_A^* \rightarrow P(A)$  in media quadratica

### Applicazione alla verifica del valore della probabilità di un evento

Se si volesse verificare se un dado o una lotteria sono truccati, quante misure  $N$  sarebbero necessarie per determinare  $P(A)$  con un buon livello di accuratezza? Assumendo:

- $p := P(A) :=$  come la probabilità teorica di  $A$
- $p^* :=$  frequenza campionaria di  $A$

---

<sup>1</sup>Ecco svelato il motivo per cui è bene scegliere un  $\Delta X_j \sim \sqrt{n}$  (si veda la sezione 1.4.3)

- $\delta p :=$  tolleranza sul valore di probabilità. Come ripreso più volte, nel caso si operi con incertezza tipo è necessario fissare un fattore di copertura  $k$  in particolare:

$$\delta p := k\sigma[p^*] \quad (6.27)$$

- $p_{vero} :=$  la probabilità vera sconosciuta di  $A$

Diremo che il controllo è andato a buon fine se:

$$P_{vero} \in [p - \delta p; p + \delta p] \quad (6.28)$$

Supponiamo che venga eseguito un test dell'ipotesi:

$$P_{vero} = p = m[p^*] \quad (6.29)$$

Ci si chiede se

$$|p^* - p| \stackrel{?}{\leq} k\sigma[p^*] = \delta p \quad (6.30)$$

Nel caso in cui:

- valga la disequazione si accetta le ipotesi
- non valga la disequazione si rigettano le ipotesi

Bisogna prestare attenzione ad  $p^*$  poiché non è una variabile di tipo gaussiana e quindi decade la probabilità corrispondente ad un certo fattore di copertura bensì è da calcolarsi a partire dalla binomiale:

$$K \frac{\sqrt{pq}}{\sqrt{N}} = \delta p \Rightarrow N = pq \frac{k^2}{(\delta p)^2} \quad (6.31)$$

### 6.1.2 Passeggiata a caso

La passeggiata a casa è una applicazione della distribuzione binomiale di interesse considerevole per i fisici: essa si propone di descrivere la posizione di una molecola, particella o ubriaco che si voglia, dopo un numero  $K$  di passi o urti. Supponiamo che  $d$  sia la distanza che viene percorsa ad ogni passo e che l'evento  $A$  sia associato al passo a destra, l'evento  $\bar{A}$  ovviamente sarà associato al passo a sinistra. Per semplicità di notazione assumiamo che:

$$p := P(A) \quad (6.32)$$

Calcoliamo la posizione  $X_{N,p}$  dopo  $N$  passi con probabilità  $p$

$$X_{N,p} := Kd - (N - K)d = d(2K - N) \quad (6.33)$$

Per quanto riguarda i parametri statistici invece:

- il valore atteso di  $X(K; N, p)$  si ottiene da:

$$m[x] := \sum_{k=0}^N (d(2K - N)) \binom{N}{K} p^K q^{N-K} \quad (6.34)$$

$$= d(2m[K] - N) = d(2Np - N) = Nd(p + p - 1) \quad (6.35)$$

$$= Nd(p - q) \quad (6.36)$$

Osserviamo che per  $p = q = \frac{1}{2}$  il valore atteso è 0 se invece  $p \neq q$  il valore atteso  $m[x] \propto N$ .

- la varianza di  $X(K; N, p)$  si ottiene invece da:

$$D[x] = m[(x - m[x])^2] \quad (6.37)$$

Calcoliamo prima:

$$x - m[x] = d(2K - N) - Nd(p - q) \quad (6.38)$$

$$= d(2K - N - 2Np + N) \quad (6.39)$$

$$= 2d(K - Np) \quad (6.40)$$

$$= 2d(K - m[k]) \quad (6.41)$$

Tornando ora a  $D(x)$  si ottiene che:

$$D[x] = m[(x - m[x])^2] \quad (6.42)$$

$$= m[4d^2(k - m[k])^2] \quad (6.43)$$

$$= 4d^2 D[k] \quad (6.44)$$

$$= 4d^2 Npq \quad (6.45)$$

Da cui

$$\sigma[x] = 2d\sqrt{Npq} \propto \sqrt{N} \quad (6.46)$$

Ovvero con l'aumentare dei tentativi l'incertezza aumenta

Nei processi di diffusione la posizione di un oggetto è sempre più dispersa con il passare del tempo.

$$N \sim \frac{\text{tempo trascorso}}{\text{tempo tipico fra due urti successivi}} \quad (6.47)$$

Ogni scelta casuale nasce da un un urto dell'oggetto

$$d \sim \text{cammino libero medio} = \text{distanza tipica tra urti successivi} \quad (6.48)$$

Nel caso a più dimensioni, se come spesso accade le equazioni del moto sono separabili per le direzioni cartesiane ortogonali ciascuna direzione è interpretabile come una passeggiata a casa unidimensionale.



### 6.1.3 Distribuzione di Poisson

La **distribuzione di Poisson**, definita come

$$P(K; a) := \frac{a^k}{K!} e^{-a} \quad (6.49)$$

Dove  $a > 0$  unico parametro della distribuzione e  $k \in \{0, 1, \dots, \infty\}$ . Si osserva in particolare che la distribuzione di Poisson è il limite asintotico della distribuzione binomiale a condizione che:

$$N \rightarrow \infty \quad (6.50)$$

$$p \rightarrow 0 \quad (6.51)$$

In maniera tale da:

$$Np = \text{costante} =: a \quad (6.52)$$

Consideriamo infatti la binomiale  $P(K; N, p)$  ovvero

$$P(K; N, p) = \binom{N}{K} p^K q^{N-K} \quad (6.53)$$

e prendiamo:

$$p = \frac{a}{N} \quad (6.54)$$

così che:

$$P(K; N, p) = \binom{N}{K} p^K q^{N-K} \quad (6.55)$$

$$= \binom{N}{K} \left(\frac{a}{N}\right)^K \left(1 - \frac{a}{N}\right)^{N-K} \quad (6.56)$$

$$= \frac{N(N-1)\dots(N-K+1)}{N^K} \frac{a^K}{K!} \left(1 - \frac{a}{N}\right)^N \left(1 - \frac{a}{N}\right)^{-K} \quad (6.57)$$

Passando al limite:

$$\lim_{N \rightarrow \infty} P(K; N, p) \stackrel{(\dagger)}{=} \frac{a^K}{K!} e^{-a} \quad (6.58)$$

Dove su  $(\dagger)$  si è sfruttato il fatto che il primo termine di (6.57) tende ad 1, il terzo termine è lo sviluppo dell'esponenziale mentre l'ultimo termine tende ad 1.

- Non tutte le binomiali tendono alla distribuzione di poisson infatti, considerando le prove ripetute in condizioni identiche in cui si tiene fitto l'evento in questione cioè:

$$p = \text{costante} \Rightarrow m[K] = Np \xrightarrow{N \rightarrow \infty} \infty \quad (6.59)$$

- Molte binomiali possono essere approssimate ad una poisson a condizione che:

$$N \gg 1 \quad (6.60)$$

$$p \ll 1 \quad (6.61)$$

- Verifichiamo la condizione di normalizzazione:

$$\sum_{K=0}^{\infty} \frac{a^K}{K!} e^{-a} = e^{-a} \sum_{K=0}^{\infty} \frac{a^K}{K!} \stackrel{(\dagger)}{=} e^{-a} e^a = 1 \quad (6.62)$$

*Dove su  $(\dagger)$  si è sfruttata la definizione di esponenziale in serie*

- La speranza matematica:

$$m[K] = \sum_{K=0}^{\infty} K \frac{a^K}{K!} e^{-a} = a e^{-a} \sum_{K=1}^{\infty} \frac{a^{K-1}}{(K-1)!} e^{-a} \stackrel{(\dagger)}{=} a \quad (6.63)$$

*Dove su  $(\dagger)$  si è sfruttata la definizione in serie di esponenziale tramite la sostituzione  $s \stackrel{(S)}{=} K-1$*

- Calcoliamo la varianza tramite:

$$D[K] = m[K^2] - m[K]^2 = a \quad (6.64)$$

e quindi, calcoliamo

$$m[K^2] = \sum_{K=0}^{\infty} K^2 \frac{a^K}{K!} e^{-a} = a \sum_{K=0}^{\infty} K \frac{a^{K-1}}{(K-1)!} e^{-a} \quad (6.65)$$

$$\stackrel{(\dagger)}{=} a \sum_{S=0}^{\infty} (S+1) \frac{a^S}{S!} e^{-a} \quad (6.66)$$

$$= a \left( \sum_{S=0}^{\infty} S \frac{a^S}{S!} e^{-a} + \sum_{S=0}^{\infty} \frac{a^S}{S!} e^{-a} \right) \quad (6.67)$$

$$\stackrel{(\dagger)}{=} a(a+1) = a^2 + a \quad (6.68)$$

*Dove su  $(\dagger)$  ci si è serviti della sostituzione  $S \stackrel{(s)}{=} k-1$  mentre su  $(\ddagger)$  si è sfruttato il fatto che il primo termine in parentesi è  $m[s]$  mentre il secondo è la condizione di normalizzazione.*

Si conclude sostituendo all'equazione:

$$D[K] = m[K^2] - m[K]^2 = a^2 + a - a^2 = a \quad (6.69)$$

Osserviamo quindi che anche per la distribuzione di Poisson che  $\sigma[\frac{k}{m[k]}] = \frac{1}{\sqrt{a}} \xrightarrow{a \rightarrow \infty} 0$  ovvero che si stringe all'aumentare delle misure come per la binomiale.

### Processi di Poisson

Per processi di Poisson si intendono delle sequenze di eventi casuali descritti da una coordinata<sup>2</sup> tale da soddisfare le seguenti:

- Gli eventi devono essere indipendenti tra loro.
- Il valore della coordinata di ciascun evento deve essere una variabile aleatoria continua distribuita con densità uniforme entro l'intervallo di coordinata in considerazione.
- All'interno dell'intervallo in considerazione la densità media degli eventi deve essere fissata:

$$m[k] = \lambda \Delta t = \text{costante} \quad (6.70)$$

Dove:

$k :=$  conteggio degli eventi con coordinata appartenete a  $\Delta t$

$\lambda :=$  densità media del processo

$\Delta t :=$  intervallo di coordinata considerato

Sotto queste condizioni il conteggio  $k$  degli eventi con coordinata appartenete ad un  $\Delta t$  è una variabile aleatoria poissoniana e quindi:

$$P(k; \lambda \Delta t) = \frac{(\lambda \Delta t)^k}{k!} \exp[-\lambda \Delta t] \quad (6.71)$$

### Applicazione agli istogrammi

Un'altra applicazione utile della distribuzione di Poisson la si ritrova considerando gli istogrammi. Supponendo il binning fisso e quindi  $p_j$  fisso ciascuna  $n_j^*$  viene considerata come variabile aleatoria binomiale. Quest'ultima, è bene sottolineare, non converge per  $N \rightarrow \infty$  alla distribuzione di Poisson. Per ottenere una convergenza alla distribuzione di Poisson è necessario che il binning sia più fitto possibile così che al crescere di  $N$  si abbia che  $m[n_j^*] \sim \text{costante}$ . Così facendo si ottiene

$$\begin{cases} N \rightarrow \infty \\ P_j \rightarrow 0 \end{cases} \quad (6.72)$$

A condizione che  $N \gg 1$  e che  $n_j^*$  sia *simile* ad variabile aleatorie poissoniane allora:

$$\forall bin_j \quad m[n_j^*] = a_j \quad (6.73)$$

$$\sigma_{n_j^*} = \sqrt{a_j} \quad (6.74)$$

$$P(n_j^*; a_j) \text{ é di poisson} \quad (6.75)$$

---

<sup>2</sup>Si penso ad esempio al tempo o ad una coordinata spaziale...

L'incertezza relativa invece diviene:

$$\frac{\sigma_{n_j}^*}{a_j} = \frac{1}{\sqrt{a_j}} \quad (6.76)$$

## 6.2 Distribuzione di probabilità di una variabile aleatoria continua

Per distribuzione di variabile aleatoria continua si intende l'osservazione di una variabile casuale che ha come dominio un insieme infinito non numerabile di valori. Possiamo considerare come variabili aleatorie continue:

$$x \in \mathbb{R} \quad (6.77)$$

$$x \in [x_1, x_2] \in \mathbb{R} \quad (6.78)$$

$$x \in \mathbb{R}^n \quad (6.79)$$

Per distribuzione cumulativa di una variabile aleatoria continua si intende:

$$F(x') := P(x \leq x') \quad (6.80)$$

Osserviamo le seguenti proprietà per la funzione cumulativa:

1. È adimensionale
2.  $\max\{F(X)\} = 1$
3.  $\min\{F(X)\} = 0$
4.  $F(X)$  è non decrescente cioè se  $x_1 \leq x_2$

$$\{x \leq x_1\} \subset \{x \leq x_2\} \Rightarrow P(x \leq x_1) \leq P(x \leq x_2) \Rightarrow F(x_1) \leq F(x_2) \quad (6.81)$$

5. Se  $x_1 \leq x_2$  allora:

$$P(x_1 \leq x \leq x_2) = P(x \leq x_2) - P(x \leq x_1) = F(x_2) - F(x_1) \quad (6.82)$$

6.  $F(x')$  è continua in  $x'$  se e soltanto se  $P(x = x') = 0$  infatti:

$$P(x = x') = \lim_{x_1 \rightarrow x'} P(X \leq x') - P(x \leq x_1) \quad (6.83)$$

$$= \lim_{x_1 \rightarrow x'} F(x') - F(x_1) = 0 \Leftrightarrow F(x) \text{ è continua} \quad (6.84)$$

Ad esempio se  $F(X)$  è continua su tutto il dominio:

$$P(x = x') = 0 \quad \forall x' \in \text{Dominio} \quad (6.85)$$

Viceversa, una discontinuità di  $F(x)$  in  $x'$  implica:

$$P(x = x') \neq 0 \quad (6.86)$$

### 6.2.1 Densità di probabilità

La distribuzione di densità di probabilità di  $F(x)$  è così definita:

$$f(x) := \frac{d}{dx} F(x) \quad (6.87)$$

Vale anche che:

$$F(x') = \int_{\inf \text{ del dominio}}^{x'} f(x) dx \quad (6.88)$$

Si osserva che la densità di probabilità ha le seguenti caratteristiche:

1. Ha dimensione  $\frac{1}{[x]}$
2. Soddisfa la condizione di normalizzazione infatti:

$$\int_{\text{dom } f} f(x) dx = 1 \quad (6.89)$$

3.  $f(x) \geq 0$  poiché  $F(x)$  è non decrescente<sup>3</sup>
- 4.

$$P(x_1 < x \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx \quad (6.90)$$

## 6.3 Parametri di caratterizzazione per distribuzione di probabilità

Al fine di caratterizzare le distribuzione di probabilità sono state introdotte delle specifiche applicazioni. In particolare le seguenti sono adatte alla misura della posizione

- Speranza matematica<sup>4</sup>
- Mediana è il valore  $x_{med}$  tale che:

$$P(x \leq x_{med}) = 0.5 \Leftrightarrow F(x_{med}) = 0.5 \quad (6.91)$$

In particolare:

$$x_{med} := F^{-1}(0.5) \quad (6.92)$$

<sup>3</sup>Si veda l'osservazione 4 in riferimento alla sezione 6.2.

<sup>4</sup>Si veda il capitolo 2.1.1 nel caso di amnesia improvvisa

- Moda è il valore  $x_{moda}$  tale che:

$$f(x_{moda}) = \max\{f(x)\} \quad (6.93)$$

È bene sottolineare che possono esistere distribuzioni con più di un valore di moda dette appunto multimodali.

Se invece si vuole esaminare la larghezza di una distribuzione è bene utilizzare:

- Varianza e quindi  $\sigma$
- Percentili
- Semi-larghezza a metà altezza
- Larghezza a metà altezza
- Momenti della distribuzione.

### 6.3.1 Momenti della distribuzione

I **momenti della distribuzione** possono essere di due tipi:

- **Momenti iniziali**, generalmente indicati con  $\alpha_s$  sono detti momenti iniziali di ordine  $s$  definiti come:

$$\alpha_s := m[x^s] \quad (6.94)$$

Hanno dimensione  $[x^s]$  e:

- Per variabili aleatorie continue è definito come:

$$\alpha_s := \int_{dom} x^s f(x) dx \quad (6.95)$$

- Per variabili aleatorie discrete è definito come:

$$\alpha_s := \sum_{x_j \in dom} x_j^s P(x_j) \quad (6.96)$$

Bisogna prestare particolare attenzione perché non è assicurata l'esistenza di  $\alpha_s$  di un dato ordine per variabili aleatorie continue perché l'integrale potrebbe non convergere. Se esistessero tutti gli  $\alpha_s$  di ogni ordine, la loro conoscenza è equivalente alla conoscenza della densità di probabilità di  $f(x)$ .

Notare che  $\alpha_0$  è la condizione di normalizzazione e quindi:

$$\alpha_0 := 1 \quad (6.97)$$

$\alpha_1$  rappresenta invece la speranza matematica infatti:

$$\alpha_1 := \int x f(x) dx \quad \sum x_j P(x_j) \quad (6.98)$$

- **Momenti centrati**, generalmente indicati con  $\mu_s$  sono detti momenti centrati di ordine  $s$  definiti come:

$$\mu_s := m[(x - m[s])^s] \quad (6.99)$$

Hanno dimensione  $[x^s]$  e:

- Per variabili aleatorie continue è definito come:

$$\mu_s := \int_{dom} (x - m[x])^s f(x) dx \quad (6.100)$$

- Per variabili aleatorie discrete è definito come:

$$\mu_s := \sum_{x_j \in dom} (x_j - m[x])^s P(x_j) \quad (6.101)$$

Valuteremo ora i principali momenti centrati:

0.  $\mu_0$  corrisponde alla condizione di normalizzazione quindi:

$$\mu_0 = 1 \quad (6.102)$$

1.  $\mu_1$  invece è identicamente nullo infatti:

$$\mu_1 := m[(x - m[x])] = \int (x - m[x]) f(x) dx \quad (6.103)$$

$$= \int x f(x) dx - m[x] \int f(x) dx = m[x] - m[x] = 0 \quad (6.104)$$

2.  $\mu_2$  invece corrisponde alla varianza della distribuzione infatti:

$$\mu_2 := m[(x - m[x])^2] = D[x] = \sigma^2[x] \quad (6.105)$$

3.  $\mu_3$  misura l'asimmetria della distribuzione rispetto alla speranza matematica infatti è nullo nel caso di distribuzioni simmetriche:

$$\begin{aligned} \mu_3 &= \int_{dom} (x - m[x])^3 f(x) dx \\ &= \int_{-\infty}^{m[x]} (x - m[x])^3 f(x) dx + \int_{m[x]}^{\infty} (x - m[x])^3 f(x) dx \quad (6.106) \\ &\stackrel{(\dagger)}{=} \int_{-\infty}^0 y^3 f(x[m] + y) dy + \int_0^{\infty} y^3 f(x[m] + y) dy \end{aligned}$$

Dove su  $\dagger$  si è fatto uso della sostituzione  $y \stackrel{(s)}{=} x - m[x]$

Osserviamo infatti che se in (6.106) gli integrali avessero segno opposto la distribuzione sarebbe simmetrica e quindi  $\mu_3$  nullo.

4.  $\mu_4$  pesa molto di più le grandi fluttuazioni e quindi le code delle distribuzioni

**Coefficiente di asimmetria**

Il **coefficiente di asimmetria**, definito come:

$$\mathcal{S} := \frac{\mu_3}{\sigma^3} \quad (6.107)$$

è adimensionale ed ha lo scopo di misurare gli sbilanciamenti a destra ( $\mathcal{S} > 0$ ) o a sinistra ( $\mathcal{S} < 0$ ). Nel caso di distribuzioni come: *gaussiana*, *densità uniforme* e *binomiale* a patto che  $p = q = 0.5$  Si ha che:

$$\mathcal{S} = 0 \quad (6.108)$$

Il che equivale a dire che sono simmetriche, in particolare, per la binomiale vale che:

$$\mu_3 = [(K - Np)^3] = Npq(p - q) \quad (6.109)$$

Da cui

$$\mathcal{S} = \frac{q - p}{\sqrt{Npq}} \propto \frac{1}{\sqrt{N}} \xrightarrow{N \rightarrow \infty} 0 \quad (6.110)$$

Si osserva quindi che asintoticamente per un numero elevato di misura la binomiale si simmetrizza indipendentemente dal valore di  $p$ .

Anche per la distribuzione di Poisson:

$$\mathcal{S} = \frac{1}{\sqrt{a}} \xrightarrow{a \rightarrow \infty} 0 \quad (6.111)$$

**Coefficiente di appiattimento**

Il **Coefficiente di appiattimento** indicato generalmente con  $\mathcal{A}$  e definito come:

$$\mathcal{A} := \frac{\mu_4}{D^2} - 3 \quad (6.112)$$

Si propone di misurare appunto quando una distribuzione è appiattita ovvero se le code della distribuzione sono più importanti ( $\mathcal{A} > 0$ ) o meno importanti ( $\mathcal{A} < 0$ ) rispetto alla distribuzione gaussiana. Infatti per la distribuzione di Gauss si ha:

$$\frac{\mu_4}{D^2} = 3 \Rightarrow \mathcal{A} = 0 \quad (6.113)$$

In particolare, per la distribuzione binomiale si ha che:

$$\mu_4 = \sum_{K=0}^N (K - Np)^4 P(K; N, p) = Npq(1 + 3Npq - 6pq) \quad (6.114)$$

e quindi

$$\mathcal{A} = \frac{1 - 6pq}{Npq} \propto \frac{1}{N} \xrightarrow{N \rightarrow \infty} 0 \quad (6.115)$$



Mentre per la distribuzione di Poisson

$$\mu_4 = 3a^2 + a \Rightarrow \mathcal{A} = \frac{1}{\sqrt{a}} > 0 \xrightarrow{a \rightarrow \infty} 0 \quad (6.116)$$

Come nel caso precedente di coefficiente di asimmetria, anche in questo caso la binomiale e la poissoniana, per un numero elevato di misure, tendono alle caratteristiche della distribuzione gaussiana infatti si vedrà che per  $N \rightarrow \infty$  sia la binomiale che la distribuzione di poisson diventano la distribuzione gaussiana.

### **Momenti della distribuzione di Poisson**

Applicando ora quanto visto in precedenza i momenti della distribuzione di Poisson sono:

- Coefficiente di asimmetria:

$$\mu_3 = a \Rightarrow \mathcal{S} = \frac{\mu_3}{\sigma_3} = \frac{1}{\sqrt{a}} \xrightarrow{a \rightarrow \infty} 0 \quad (6.117)$$

- Coefficiente di appiattimento:

$$\mu_4 = 3a^2 + a \Rightarrow \mathcal{A} = \frac{\mu_4}{D_3} - 3 = \frac{1}{a} > 0 \quad (6.118)$$

Notiamo che per  $a \rightarrow \infty$  anche la distribuzione di Poisson tende alla gaussiana.



# Indice analitico

- $\chi^2$ , 30
  - Calcolo incertezze a posteriori, 36
  - Casistica del test, 34
  - Test, 33
- Calcolo combinatorio, 55
  - Combinazioni, 57
  - Disposizioni, 56
  - Permutazioni, 56
- Compatibilità tra misure, 21
- Confronto tra media pesata e campionaria, 24
- Covarianza campionaria, 17
- Densità di probabilità, 71
- Distribuzione
  - Coefficiente di asimmetria, 74
  - Coefficiente di Kustosis, 74
  - Coefficienti di appiattimento, 74
  - Momenti centrati, 73
  - Momenti iniziale, 72
  - Parametri di caratterizzazione, 71
- Variabile aleatoria continua, 70
  - Densità di probaibilità, 71
  - Funzione cumulativa, 70
- Variabile aleatoria discreta, 61
  - Binomiale, 58, 62
  - di Bernoulli, 58
  - Di Poisson, 67
  - Funzione cumulativa, 61
- Multinomiale, 59
- Errori
  - Casuale o tipo *A*, 7
  - Classificazione, 4
  - Risoluzione, 6
  - Sistematico o tipo *B*, 7
- Falsi allarmi, 22
  - Distribuzione normale, 22
- Fattore di compertura, 22
- Fenomeni aleatori, 41
- Fit, 29
  - Calcolo incertezze a posteriori, 36
  - Metodo dei minimi quadrati, 30
  - Metodo di massima e minima pendenza, 29
  - Riepilogo procedura, 32
  - valutazione della legge, 33
- Grandezze statisticamente indipendenti, 17
- Incertezza
  - Cause, 4
  - Propagazione, 15
  - Propagazione alla media campionaria, 18
  - Propagazione caso generale, 27
  - Propagazione con funzioni di più variabili, 26

- Propagazione di contributi differenti, 20
- Propagazione in relazione lineare, 15
- Propagazione via metodo grafico, 24
- Rappresentazione, 5
- Sovrastima dell'incertezza, 28
- Tipi di intervallo, 5
- Istogrammi, 9
  - Applicazione della binomiale, 63
  - Binning, 9
  - Cumulativi, 11
  - Densità campionaria, 11
  - Istogramma limite, 11
  - Media campionaria, 10, 12
  - Normalizzazione in altezza, 9
  - Normalizzazione in area, 10
  - Proprietà asintotiche, 11
  - Scelta del bindaggio, 10
  - Stimatori, 12
  - Varianza campionaria, 10, 12
- Legge dei grandi numeri, 64
- Media pesata, 23
- Misure dirette e indirette, 3
- Parametri di caratterizzazione
  - Distribuzione di Poisson, 75
- Passeggiata a caso, 65
- Probabilità, 42
  - Applicazione teorema della probabilità totale, 50
  - Classificazione eventi, 42
  - Condizionata ad eventi, 46
  - Condizionata al prodotto di eventi, 49
  - Definizione assiomatica, 43
  - Esito, 42
  - Eventi indipendenti, 47
  - Evento, 42
  - Incompatibilità tra eventi, 42
  - Interpretazione Bayesiana, 45
  - Interpretazione classica, 44
  - Interpretazione frequentista, 44
  - Prodotto di eventi, 42
  - Somma di eventi, 42
  - Somma di eventi qualunque, 44
  - Spazio campionario, 42
  - Teorema della probabilità totale, 49
  - Teorema di Bayes, 47
- Processi di Poisson, 69
  - Scarto tipo, 7
  - Scarto tipo stimato, 12
  - Speranza matematica, 16
  - Strumenti di misura, 3
    - Caratteristiche, 3
- Variabile aleatorie, 61
- Varianza campionaria, 8