

Lezione n.22

MISURE, STRUMENTI DI MISURA, INCERTEZZE Sperimentali

MISURA DI UNA GRANDEZZA FISICA

- Misurare una grandezza fisica significa, in generale, individuare il **numero** che esprime il **rapporto** tra la grandezza da misurare e *una della stessa natura, assunta come unità di misura*:

$$x = \frac{\text{grandezza}}{\text{unità'}}$$

- Scritto così sembra semplice, in realtà solo raramente individuare il valore di questo rapporto è una operazione semplice, più spesso è una operazione indiretta.
- E' fondamentale corredare il valore numerico trovato x con l' unità $[x]$ altrimenti x non ha alcun significato:

Misura di X = $x [x]$

TIPI DI PROCEDURA DI MISURA

- ✖ **Misura diretta:** con uno strumento che misura *quella* grandezza.
 - + Es.: Lunghezza con metro, corrente elettrica con amperometro,...
- ✖ **Misura indiretta:** non si misura la grandezza di interesse ma una o più *altre* grandezze, *con strumenti che le misurano direttamente*, e si risale alla grandezza desiderata mediante una *legge nota*:

$$y = f(x_1, x_2, x_3, \dots)$$

- + Es.: Accelerazione $a=f(s,t)$ da misure dirette di s con metro e t con cronometro

OPERAZIONI/STRUMENTI DI MISURA (DIRETTA)

- ✖ **Confronto diretto:** si *accosta/giustappone* un campione della grandezza alla grandezza da misurare, eventualmente all' interno di uno strumento più complesso.
 - + Es.: Lunghezza con metro, massa con bilancia a piatti
- ✖ **Lettura diretta:** si fa *interagire* lo strumento con la grandezza, in modo che nello strumento avvenga un *processo fisico*, che produce sullo strumento la lettura di una grandezza *diversa da quella da misurare* (tipicamente lunghezza, angolo, digit) ma correlata a quella da misurare mediante una relazione nota, dal costruttore direttamente “*implementata*” nello strumento, nel senso che lo strumento è fornito di una scala **tarata** direttamente in unità della grandezza da misurare
 - + Es.: bilancia a indice (dinamometro), amperometro

La maggior parte degli strumenti per misure elettriche e ottiche sono strumenti **tarati**.

STRUMENTI TARATI

- ✖ Uno strumento **tarato** è costituito da:
 - + un dispositivo sensibile alla grandezza da misurare, **rivelatore**
 - + un dispositivo indicatore, *indice* (ago, tacca luminosa) che mostra, generalmente all'occhio dello sperimentatore, il risultato della misura su una scala graduata (*lettura analogica*, indicazione continua) o su un numeratore elettronico (*display a lettura digitale*, indicazione discontinua).
 - + Tra il rivelatore e l'indice a volte si interpone un dispositivo **trasduttore** che trasforma la risposta del rivelatore in variazioni di una grandezza più facilmente osservabile dallo sperimentatore, quale lunghezza, angolo.

INCERTEZZE Sperimentali

- Se il risultato di una misura (diretta), comunque effettuata, è un numero, corredato dall'opportuna unità di misura ...

~~Misura di $X = x / x /$~~

- ... tale numero non è però un *numero esatto* in senso matematico.
- Ciò perché l'operazione del misurare è soggetta a una ineliminabile **incertezza sperimentale Δx , quindi:**

Misura di $X = x \pm \Delta x [x]$

- Parte essenziale della esecuzione di una misura è non soltanto determinare il più accuratamente x ma anche determinare il più accuratamente possibile l' **incertezza Δx** .

INCERTEZZA E ERRORE - 1

Concettualmente:

- ✖ Diciamo x_v , il valore “**vero**” della grandezza che stiamo per misurare.
- ✖ Diciamo x il valore misurato (il risultato della misura).
- ✖ In generale x sarà diverso da x_v , (ancorchè -sperabilmente- vicino a x_v , per almeno **due** motivi...
 1. Limitazioni intrinseche dello strumento
 2. Fluttuazioni casuali nell’ interazione tra grandezza e strumento
- ✖ ...e talvolta anche per un **terzo** motivo:
 3. Una errata taratura dello strumento o una alterata risposta fisica dello strumento
- ✖ La differenza $e = x_v - x$ è l’ errore di fatto commesso in quella misura
- ✖ Naturalmente, tutto questo è solo concettuale, perché x_v non lo conosciamo (altrimenti non ci sarebbe motivo di eseguire la misura) e dunque non potremo conoscere e .
- ✖ Il punto però è che si deve quantificare la bontà della misura. Allora di deve **stimare** e .
- ✖ La corretta determinazione dell’ **incertezza Δx** costituisce la migliore stima possibile dell’ **errore** inconoscibile e e ne determina dei limiti, il cui significato è in parte deterministicamente probabilistico.
- ✖ Dunque non determiniamo l’ **errore sperimentale**, ma lo stimiamo determinando certi suoi limiti, e questi limiti sono chiamati l’ **incertezza sperimentale**.

continua ➔

INCERTEZZA E ERRORE - 2

La differenza è illustrata graficamente in questo schizzo:

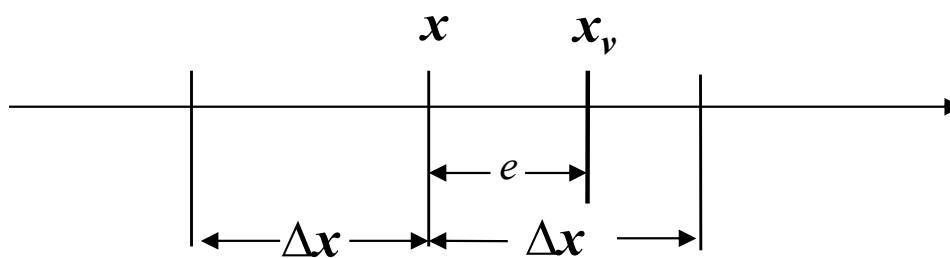


Figura 1: Differenza fra incertezza e errore: $x_v \in (x - \Delta x ; x + \Delta x)$

SORGENTI DI ERRORE → CONTRIBUTI ALL' INCERTEZZA

- ✖ Errore assoluto massimo a priori (o sensibilità di misura) → ε_x
 - Intrinseca, ineliminabile limitazione dello strumento: la min quantità della grandezza che lo strumento o il dispositivo in uso è in grado di apprezzare.
- ✖ Errori accidentali (o casuali, o statistici) → σ_x
 - Imprevedibili variazioni nella interazione tra strumento e grandezza, in più o in meno con uguale probabilità, per natura non determinabili esattamente, ma stimabili con metodi *statistici*
- ✖ Errore sistematico → η_x
 - Erronea taratura dello strumento o alterazione nella sua risposta
 - Erronea predisposizione dell' apparato sperimentale, ...

→ Incertezza sperimentale:
$$\Delta x = \sqrt{\varepsilon_x^2 + \sigma_x^2 + \eta_x^2}$$

DETERMINAZIONE DEI CONTRIBUTI ALL' INCERTEZZA - 1

✖ Dunque per ogni misura (diretta) di una grandezza X bisogna determinare l' incertezza Δx . Ciò implica in linea di principio 3 operazioni, che di seguito approfondiremo:

1. Valutazione di ε_x

- → caratteristiche degli strumenti
 - →→ caratteristiche degli strumenti elettrici

2. Stima di σ_x

- → ripetizione misure
 - →→ analisi con metodi statistici dei dati raccolti

3. Indagine su η_x

continua ➔

DETERMINAZIONE DEI CONTRIBUTI ALL' INCERTEZZA - 2

❖ AVVERTENZE GENERALI:

1. La determinazione di questi contributi non è una scienza esatta e su **alcuni dei passi coinvolti** non è raro trovare opinioni diverse di autori (di testi) o sperimentatori diversi. Pertanto quello che si dà sono criteri o linee guida, ma non sempre è possibile né corretto dare una prescizione obbligatoria assoluta.
2. Una certa dose di buon senso e di adattabilità a specifici casi concreti è richiesta come parte della “formazione” dello sperimentatore.
3. Come conseguenza di (1) è essenziale che quando si riporta il risultato di una misura si descriva anche dettagliatamente tutta la procedura di analisi che ha portato ai valori numerici finali in modo che si sappia esattamente come si è deciso di valutare ogni contributo.
4. E’ ugualmente essenziale riportare i dati originari non analizzati, in modo che eventualmente un altro sperimentatore possa condurre una nuova analisi degli stessi con criteri diversi e giungere a un suo risultato finale.
5. Se le analisi sono tutte corrette i risultati finali potranno essere comunque differenti ma, di solito, lo saranno non “troppo”.
6. ***Ogni volta che sia possibile metterò io stesso in evidenza casi in cui esistono scuole di pensiero differenti su uno stesso passo.***

VALUTAZIONE DI ε_x : CARATTERISTICHE DEGLI STRUMENTI - 1

- ✖ L' **errore assoluto massimo a priori** ε_x (detto anche **sensibilità di misura**) è definito come la più piccola quantità di grandezza capace di fare variare in maniera riproducibile la risposta dello strumento.
- ✖ Questa apparentemente semplice definizione si traduce in una varietà di casistiche pratiche. Per affrontarle, consideriamo alcune caratteristiche che giocheranno un ruolo nella individuazione del valore di ε_x da adottare.
- ✖ *Inoltre, in apparati complessi costituiti da più strumenti che concorrono cooperativamente alla misura, ε_x sarà da valutare con riferimento alla procedura nel suo complesso piuttosto che a un dato strumento.*
- ✖ **Campo di misura:** è dato dalla differenza tra il valore massimo, detto **portata o fondo scala (f.s.)**, e il valore minimo, detto **soglia**, della grandezza da misurare che lo strumento può apprezzare. Esistono strumenti che hanno più di una portata. Le diverse portate sono prefissate dal costruttore, e possono essere impostate a scelta dello sperimentatore, mediante manopole, interruttori etc.
- ✖ **Sensibilità di lettura** I_x : il min intervallo della scala predisposta dal costruttore, o una unità della più piccola cifra mostrata se lo strumento ha un display digitale.
- ✖ **Sensibilità dello strumento** α_x : il rapporto tra spostamento dell' indice (o variazione del numero digitale mostrato) e variazione della grandezza che lo causa. Per esempio, se la scala è di angoli: $\alpha_x = \delta\theta/\delta x$.
 - + Se il processo fisico utilizzato per la misura non coinvolge una legge lineare, α_x può variare lungo la scala. Per esempio se $\theta(x)$ è una funzione quadratica, $\theta(x)=kx^2$, allora $\alpha_x=2kx$, quindi cresce lungo la scala, cioè bastano variazioni via via più piccole di x per produrre uguali spostamenti dell' indice: la scala si "dirada" verso i valori di x più grandi.

continua ➔

VALUTAZIONE DI ε_x : CARATTERISTICHE DEGLI STRUMENTI - 2

- ✖ La caratteristica più facilmente accessibile all'utilizzatore di uno strumento, sia a confronto diretto che a lettura diretta o tarato, è ovviamente la sua **sensibilità di lettura l_x** .
- ✖ Sarebbe ideale dunque che gli strumenti fossero costruiti in modo che $l_x = \varepsilon_x$.
- ✖ Spesso però questo non avviene, talvolta per buoni motivi.

VALUTAZIONE DI ε_x : STRUMENTI A CONFRONTO DIRETTO - 1

- ✖ In strumenti molto semplici, in genere in quelli a confronto diretto, ε_x , che è legata alla risposta strutturale dello strumento, è generalmente molto più piccola di qualunque scala leggibile che si riesca a realizzare.
 - + Es.: un regolo, strutturalmente, ha sensibilità di misura pari alla dimensione di un atomo del suo materiale, cioè se potessimo vedere i singoli atomi, giustapponendo il regolo a un oggetto da misurare, il confronto avrebbe un livello di dettaglio (**sensibilità**) pari alle dimensioni di un atomo del regolo. Cioè potrei fare affermazioni del tipo “questo oggetto è lungo $3\ 549\ 327\ 512 \pm 1$ atomi del mio regolo”, Ma nessuna scala leggibile potrebbe avere tacche spaziate così finemente, a meno di non corredare il regolo di un ingombrante microscopio elettronico. Nei migliori “regoli” da laboratorio (i **calibri**) il **nonio aiuta** nella direzione giusta, portando la sensibilità di lettura da 1 mm dei regoli comuni a valori più piccoli come 10^{-4} o 10^{-5} m.
- ✖ In strumenti del genere (calibri, palmer, goniometri con nonio) dunque la limitazione **pratica** non è quella intrinseca allo strumento ma al fatto che deve essere letto da un operatore umano.
- ✖ Nelle scale praticamente realizzate, si ritiene in generale che quando l’ “indice” (l’ indice può anche non essere parte dello strumento ma p.es. uno spigolo dell’ oggetto) sta tra due tacche lo sperimentatore, utilizzando la massima cura possibile, cioè ponendosi di fronte al tratto interessato per evitare errori di parallasse, sia in grado di dire almeno se l’ indice sta nella prima o nella seconda parte dell’ intervallo tra due tacche, e quindi che sia – in definitiva - in grado di apprezzare $l_x/2$.

continua ➔

VALUTAZIONE DI ε_x : STRUMENTI A CONFRONTO DIRETTO - 2

- ✖ In tali casi si pone quindi, in definitiva, $\varepsilon_x = l_x/2$. Perchè? Perchè anche se strutturalmente lo strumento potrebbe fare di meglio, l' operatore che lo legge commette un errore che a priori stimiamo abbia **questo** valore massimo..
- ✖ Nell' eventualità che la scala dello strumento sia molto "larga", come p. es. un regolo con una scala in cm o in mezzi cm (ne abbiamo e ne useremo in Lab2) si ritiene che l' operatore sia in grado di giudicare la posizione dell' "indice" entro frazioni anche più piccole dell' intervallo tra due tacche, quali 1/4 o anche 1/10. In definitiva, la valutazione di quale frazione dell'intervallo tra due tacche riesce a apprezzare è lasciata al singolo sperimentatore.

In conclusione:

Detta **n** la frazione di intervallo che lo sperimentatore si sente in tutta coscienza in grado di apprezzare, la **sensibilità di misura** ε_x in uno strumento a confronto diretto, o anche a lettura diretta ma semplice, sarà data da

$$\varepsilon_x = l_x/n$$

Se lo strumento è **digitale**, in ogni caso ε_x non può essere altro che l_x , la più piccola cifra mostrata, cioè $\varepsilon_x = l_x$.

NOTA: secondo alcuni (v. p.es. Foti-Gianino, p. 29) negli strumenti digitali bisogna adottare $\varepsilon_x = l_x/2$ nel caso in cui lo strumento mostra un numero stabile o oscillante tra 2 cifre, e $\varepsilon_x = 0.7l_x$ nel caso particolare che la cifra mostrata oscilli di ± 1 unità.

VALUTAZIONE DI ε_x : STRUMENTI ELETTRICI

- ✖ Negli strumenti elettrici (tarati!) si verifica spesso la condizione opposta: la sensibilità dello strumento è peggiore di una scala graduata leggibile (in tal caso in angoli) che si potrebbe realizzare o delle cifre che un display potrebbe mostrare. Capita cioè che lo strumento risponde in maniera riproducibile a una variazione di tensione, corrente elettrica, etc, solo se tale variazione è grande abbastanza da corrispondere a un angolo relativamente ampio, certamente maggiore del min che un umano potrebbe leggere senza difficoltà, pertanto tali strumenti hanno spesso scale con tacche piuttosto distanziate.
- ✖ In fase di costruzione il costruttore valuta con dei test la reale **sensibilità di misura ε_x** . Sembra dunque piuttosto semplice tracciare le tacche sulla scala semplicemente spaziate proprio di ε_x , o se più piace di $2\varepsilon_x$.
- ✖ Ma non sempre questo avviene. I motivi sono due:
 1. Spesso ε_x è un valore non semplice, p.es. 0.375 V. Sarebbe scomoda e non rapida una lettura con tacche spaziate di 0.375. Ben più pratica e rapida è la lettura se le tacche sono spaziate di -diciamo- 0.1 o 0.2 o 0.5.
 2. Molti strumenti hanno un' unica scala per la lettura ma più portate; in tal caso la distanza tra due tacche potrebbe corrispondere al valore di ε_x al massimo in una sola di tali portate.
- ✖ Allora per gli strumenti elettrici è invalsa l' abitudine di riportare sullo strumento un parametro atto a dedurre ε_x in qualunque portata, se ve ne è più di una: la

classe di precisione: C_p

VALUTAZIONE DI ε_x : CLASSE DI PRECISIONE

- La classe di precisione esprime ε_x come percentuale del fondo scala:

$$C_p = \frac{\varepsilon_x}{(f.s.)} \cdot 100$$

- Sembra un giro vizioso ma in realtà in questo modo basta che il costruttore dichiari un solo numero (C_p) perché lo sperimentatore possa dedurre ε_x in qualunque fondo scala adoperato:

$$\varepsilon_x = \frac{C_p}{100} \cdot (f.s.)$$

- C_p viene normalmente riportato nel quadrante di lettura o più raramente sul manuale dello strumento.

Esempio pratico: Un voltmetro ha fondo scala 3 / 10 / 50 Volt e $C_p=1.5$. Viene adoperato nel f.s.=3 Volt. Qual' è la sua sensibilità di misura ?

$$\varepsilon_x = \frac{1.5}{100} \cdot (3) = 0.045 V$$

Ma se lo stesso venisse adoperato nel f.s.=10V, avrebbe $\varepsilon_x=0.15V$

VALUTAZIONE DI ε_x IN STRUMENTI ELETTRICI

ATTENZIONE!

- Molti sperimentatori ritengono che l' **errore assoluto massimo a priori** da utilizzare come ingrediente di Δx in una misura elettrica sia quello derivante dalla **classe di precisione** come illustrato nella slide precedente.
- Ma un' altra scuola di pensiero ritiene che vi siano due contributi a ε_x . Uno relativo alla risposta intrinseca dello strumento, espresso da C_p come illustrato prima, e un secondo derivante dal possibile errore che può commettere lo sperimentatore nella lettura, valutato come nella lettura degli strumenti a confronto diretto come l_x/n o una unità dell' ultima cifra se digitale. In tal caso ε_x si deve valutare così:

$$\varepsilon_x = \sqrt{\left(\frac{C_p}{100} \cdot (f.s.)\right)^2 + \left(\frac{l_x}{n}\right)^2}$$

- Gli appartenenti alla prima scuola di pensiero ritengono inutile il secondo contributo dato che normalmente non vi è alcuna difficoltà nella lettura della scala e in ogni caso questo contributo sarebbe trascurabile rispetto al primo.

STIMA DI σ_x : VARIAZIONI CASUALI

- ✖ L' **errore casuale** σ_x (detto anche **errore accidentale** o **statistico**) è definito come l' imprevedibile variazione nella interazione tra strumento (o apparato) e grandezza da misurare.
- ✖ Per sua natura quindi il suo contributo all' incertezza Δx non è determinabile se si effettua la misura una sola volta, dato che in tal caso non si potrà osservare alcuna **variazione**.
- ✖ Per stimare σ_x è dunque indispensabile ripetere una misura più volte, nelle stesse condizioni. La precisione “nelle stesse condizioni” è importante in quanto altrimenti una variazioni del risultato potrebbe essere imputabile alla variazione delle condizioni piuttosto che alla casuale variazione nell' interazione tra strumento e grandezza.
- ✖ Ma anche ripetendo una misura più volte, le variazioni in questione saranno rilevabili solo se sono più grandi della **sensibilità di misura** ε_x altrimenti le variazioni concettualmente ci sono ma lo strumento a disposizione non ha sensibilità sufficiente per rilevarle.
- ✖ Dunque, ammesso che si siano controllate al meglio le condizioni della misura per ridurre al minimo le variazioni di tali condizioni e che lo strumento non sia guasto, uno strumento è – paradossalmente- migliore se i risultati della sua misura, eseguita più volte, fluttuano: significa che ha una **sensibilità di misura** ε_x abbastanza piccola da apprezzare le naturali variazioni casuali nella propria interazione con la grandezza!
- ✖ Dunque per stimare il contributo **casuale** σ_x all' incertezza Δx si deve in linea di principio sempre provare a ripetere una data misura più volte, nelle stesse condizioni.

STIMA DI σ_x : CASO IN CUI E' TRASCURABILE

- Se la ripetizione fornisce sempre lo stesso risultato, si conclude che σ_x non è stimabile in quanto è $<\varepsilon_x$ e si pone $\sigma_x=0$.
- In tal caso come valore misurato si riporta quello letto e come incertezza:

$$\Delta x = \sqrt{\varepsilon_x^2 + \eta_x^2}$$

- [NOTA: su η_x dobbiamo ancora investigare]
- Questa scelta porta comunque a una stima ragionevole di Δx in quanto in tal caso σ_x^2 sarebbe trascurabile rispetto a ε_x^2 .

Con la pratica, per alcuni tipi di strumenti e/o di misure, diviene spesso evidente che tale ripetizione porterebbe sempre agli stessi valori in quanto lo strumento a disposizione non ha sensibilità sufficiente per rilevare le variazioni casuali e non sarà necessario ripetere effettivamente la misura.

Questo è spesso il caso per misure dirette di grandezze elettriche.

STIMA DI σ_x : ANALISI STATISTICA - 1

- Se la ripetizione NON fornisce sempre lo stesso risultato, si conclude che σ_x è stimabile in quanto è $>\varepsilon_x$.
- In tal caso la migliore determinazione del risultato della misura è la media aritmetica dei valori letti:

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

- Ovvero, raggruppando i valori uguali tra loro e detto q_i il numero di volte che ciascun valore distinto si presenta e $f_i = q_i/n$ la frequenza con cui si presenta:

$$x_m = \frac{\sum_{i=1}^r q_i x_i}{\sum_{i=1}^r q_i}$$

$$x_m = \sum_{i=1}^r f_i x_i$$

- Mentre come σ_x si prende la **deviazione standard**:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - x_m)^2}{n - 1}}$$

continua ➔

STIMA DI σ_x : ANALISI STATISTICA - 2

- ✖ Le basi teoriche per le scelte precedenti sul valore da riportare x_m e su σ_x sono le seguenti:
 1. La statistica mostra che se le variazioni nei risultati sono davvero dovute a fluttuazioni casuali, allora al crescere del numero di misure (nelle stesse condizioni sperimentali) cioè per $n \rightarrow \infty$, il valor **medio** tende al valore “vero” inconoscibile della grandezza:

$$x_m \rightarrow x_v$$

- 2. Inoltre, la statistica mostra anche che per $n \rightarrow \infty$ i valori misurati tendono a presentare una **distribuzione gaussiana** centrata proprio sul valor medio:

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

- ✖ Naturalmente nella pratica non si possono fare ∞ misure. In generale è bene farne il più possibile. In ogni caso per ogni dato numero di misure effettivamente fatte, la **deviazione standard** σ_x risulta essere la migliore stima praticamente accessibile del parametro σ della distribuzione gaussiana.
- ✖ In tale distribuzione, σ rappresenta la distanza ai due lati del valore centrale a cui si hanno i punti di flesso della curva e questi, matematicamente, delimitano la regione attorno al valore centrale che contiene il 68.3% dei valori.
- ✖ In conclusione la deviazione standard adottata come σ_x rappresenta una stima e non una determinazione esatta dell' **errore casuale**. In tale senso è anche chiamata un *indice di precisione*.
 - + Tale indice indica l' intervallo entro cui, assumendo il processo casuale, è presumibile che starebbero ~68.3 % delle misure, se si continuasse a misurare sempre nelle stesse condizioni.
 - + Ovviamente tale parametro è tanto meglio stimato quante più misure si sono già fatte.
 - + Autoconsistentemente, si dovrebbe constatare che la frazione di misure fatte che effettivamente ricadono nell' intervallo $[x_m - \sigma_x, x_m + \sigma_x]$ è tanto più vicino al 68.3% quante più misure si sono fatte.

continua ➔

STIMA DI σ_x : ANALISI STATISTICA - 3

- ✖ Ricordiamo che è possibile definire un altro indice di precisione statistico, lo scarto quadratico o deviazione standard **della media**:

$$\sigma_m = \frac{\sigma_x}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n (x_i - x_m)^2}{n(n-1)}}$$

- ✖ Tale definizione sarà giustificata più avanti. Essa differisce nel significato da σ_x in maniera piuttosto sottile:
 - + σ_x è un indice della dispersione dei singoli valori attorno alla media: il 68.3% circa dei valori giace nell' intervallo $[x_m - \sigma_x, x_m + \sigma_x]$.
 - + σ_m è un indice di **dispersione delle medie**, cioè dice che una nuova media di una nuova serie di misure ha una probabilità del 68.3% di ricadere nell' intervallo $[x_m - \sigma_m, x_m + \sigma_m]$, ossia di differire dalla precedente meno di σ_m .
- ✖ Secondo alcuni è σ_m e non σ_x l' indice da utilizzare per determinare l' incertezza Δx .

INDAGINE SU η_x : CONSIDERAZIONI GENERALI

- ✖ L' **errore sistematico** η_x è definito come una erronea taratura dello strumento, o alterazione nella sua risposta, o erronea predisposizione dell' apparato sperimentale

Per sua natura, affligge le misure con una differenza, o un fattore, “*sistematici*”, cioè sempre dello stesso valore e segno, rispetto al valore vero.

- ✖ Per sua natura si può sospettare la sua presenza solo se:
 1. c' è un valore **atteso** per la misura e il valore **misurato**, corredato da una incertezza costruita con i soli due contributi già trattati,
$$\Delta x = \sqrt{\varepsilon_x^2 + \sigma_x^2}$$
, ne differisce significativamente;
 2. oppure si dispone di più strumenti diversi per misurare la stessa grandezza e i loro risultati non concordano, tenuto conto dei **Δx** di ciascuno calcolati con i soli due contributi già trattati.

INDAGINE SU η_x : CASISTICA

- ✖ In generale si presenteranno (almeno) 4 possibili casi:
 1. Se ne scopre la presenza e se ne determina il valore.
 - In tal caso non ha senso inserirlo nell' incertezza Δx : se so il suo valore, correggo direttamente i dati (sottraendogli, aggiungendogli o moltiplicandoli/dividendoli per, il fattore)
 2. Ne sospetto la presenza ma non ho alcun modo di valutarlo quantitativamente.
 - Anche in questo caso non viene, naturalmente, inserito nell' incertezza Δx , ma allora sono indispensabili commenti che illustrino il “sospetto”.
 3. Determino oltre ogni ragionevole dubbio che non c' è.
 - Ovviamente, ai fini del calcolo di Δx , si pone $\eta_x=0$.
 4. Non sono sicuro che c' è, oppure temo che ci sia ma non ne posso determinarne un valore univoco, ma posso “stimarlo” ragionevolmente.
 - In questi casi e solo in questi si procede:
 - a una sua **stima presumibile** "ragionevole"
 - o alla sua **valutazione quantitativa** ➔

INDAGINE SU η_x : VALUTAZIONE

- + In alcuni apparati sperimentali, è possibile stimare gli errori sistematici indirettamente, non perché i valori trovati appaiano in sé sospetti, ma perché p.es. dovrebbero presentare delle caratteristiche generali (di simmetria, di distribuzione), che non si ritrovano nei dati.
 - La deviazione dalla simmetria o distribuzione attesa può essere usata come *indice* dell' errore sistematico presente nell' apparato.
 - La procedura per determinare tale *indice* dipende dal singolo caso.

Esempio:

Misura della distanza focale di una lente convergente con il metodo di Bessel.

- Una volta determinato η_x :
 - secondo alcuni autori esso va computato insieme a ε_x e σ_x per determinare Δx come già visto:

$$\Delta x = \sqrt{\varepsilon_x^2 + \sigma_x^2 + \eta_x^2}$$

- secondo altri va indicato separatamente, cioè Δx presenterà due addendi:

$$\Delta x = \sqrt{\varepsilon_x^2 + \sigma_x^2} + \eta_x$$

INDAGINE SU η_x : RANDOMIZZAZIONE - 1

- ✖ Se si hanno a disposizione più strumenti o apparati diversi per misurare la stessa grandezza, si può stimare l' **errore sistematico attribuibile alla misura** mediante la cosiddetta **randomizzazione degli errori sistematici:**
 - ✖ Si esegue la misura con ciascuno dei vari strumenti.
 - ✖ Se con ciascuno strumento separatamente si è stabilito che gli errori casuali sono trascurabili, allora le eventuali differenze di valori saranno dovute solo a **differenze sistematiche** tra gli strumenti.
 - ✖ Non conosciamo l' errore sistematico di ciascuno strumento individualmente, ma è ragionevole presumere **essi siano distribuiti casualmente** tra gli strumenti
 - ✖ Allora si potranno trattare **statisticamente** gli m valori ottenuti con m strumenti diversi come un insieme di m misure della stessa grandezza ottenute con **un singolo strumento virtuale**.
 - ✖ Si calcola allora la deviazione standard σ_x di tali valori. Essa costituisce la migliore stima disponibile dell' errore sistematico dello **strumento virtuale**, quindi dell' η_x da attribuire alla misura. Essa non è l' errore sistematico di alcun particolare strumento ma rappresenta correttamente il contributo sistematico alla determinazione la più accurata possibile dell' incertezza Δx in quanto ci fa passare da nessuna notizia sull' eventuale errore sistematico di ogni strumento individuale a una valutazione dell' errore sistematico attribuibile alla misura **effettuata con il complesso degli strumenti disponibili**.
 - ✖ Nel caso che con qualcuno degli strumenti risultino apprezzabili gli errori casuali σ_i , con quelli si ripetono le misure più volte, mentre con gli eventuali strumenti a essi insensibili si rileva la misura una sola volta.

continua ➔

INDAGINE SU η_x : RANDOMIZZAZIONE - 2

- ✖ Alla fine si mettono *tutti i dati raccolti* in un unico insieme e si calcolano:
 1. Il loro valor medio complessivo.
 - ✖ Per calcolare tale valore medio si vuole dare maggiore peso alle misure effettuate con (eventuali) strumenti di sensibilità di misura ϵ_i migliore. Si utilizzerà quindi la **media pesata** in cui come peso per ciascuna misura si usa l'*inverso del quadrato dell' ϵ_i* dello strumento con cui è effettuata.
 2. La deviazione standard σ'_x di tutti i dati globalmente.
- ✖ Così facendo:
 - ✖ gli ϵ_i sono già stati utilizzati per costruire i pesi: $p_i = \frac{1}{\epsilon_i^2}$
 - ✖ gli eventuali σ_i di uno o più strumenti sono di fatto inglobati nell' unico σ'_x che si calcola come deviazione standard dell' intero insieme di dati: σ'_x contiene sia la stima di tali σ_i relativi a singoli strumenti sia la stima delle *differenze tra strumenti* dovute agli **errori sistematici degli stessi**.
 - ✖ Quindi $\Delta x = \sqrt{\cancel{\epsilon_x}^2 + \sigma_x^2 + \eta_x^2}$
- ✖ In conclusione, il risultato di una misura effettuata con randomizzazione sarà dunque:

$$\text{Misura di } X = x_{mp} \pm \sigma'_x [x]$$

in cui x_{mp} è la media pesata definita così:

$$x_{mp} = \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i}$$

Lezione n.23

PROCEDURE DI ELABORAZIONE E ANALISI DI DATI Sperimentali

ESPRESSIONE FINALE DEL RISULTATO DI UNA MISURA DIRETTA

- Le precedenti considerazioni sulla determinazione di ciascuno dei 3 possibili contributi all' **incertezza sperimentale Δx** porta alle seguenti "ricette" finali per esprimere il risultato di una misura diretta, che comprendono sia i criteri per determinare il valore **x** da riportare come **risultato** della misura sia quelli per stimare l' **incertezza Δx** con cui corredarlo.
- Tutte tali ricette si conformano al formato generale

$$\text{Misura di } X = x \pm \Delta x [x]$$

in cui sia **x** che **Δx** sono valutati il più accuratamente possibile, e ne costituiscono casi particolari.

ASSEGNAZIONE RISULTATO A UNA MISURA DIRETTA - 1

- Se abbiamo un solo strumento o apparato, la ripetizione (effettiva o presunta) delle misure ha dato sempre la stessa lettura [dunque $\sigma_x < \varepsilon_x$] e non si deve considerare un errore sistematico η_x , detto x_l il valore letto:

$$\text{Misura di } X = x_l \pm \varepsilon_x [x]$$

- Se abbiamo un solo strumento o apparato, la ripetizione (effettiva o presunta) delle misure ha dato sempre la stessa lettura [dunque $\sigma_x < \varepsilon_x$] e si è stimato l' errore sistematico η_x , detto x_l il valore letto, secondo la scuola di pensiero cui vogliamo aderire scriviamo o:

$$\text{Misura di } X = x_l \pm \sqrt{\varepsilon_x^2 + \eta_x^2} [x]$$

- ...oppure:

$$\text{Misura di } X = x_l \pm \varepsilon_x \pm \eta_x [x]$$

- In particolari casi si può determinare il segno dell' errore sistematico anche se il valore può solo essere stimato (altrimenti correggeremmo direttamente i dati): allora davanti a η_x non ci sarà \pm ma soltanto $+$ oppure soltanto $-$. In questo caso soltanto l' ultima espressione è applicabile.

ASSEGNAZIONE RISULTATO A UNA MISURA DIRETTA – 2.1

- Se abbiamo un solo strumento o apparato, la ripetizione (effettiva) delle misure ha dato letture diverse [dunque $\sigma_x > \varepsilon_x$] e non si deve considerare un errore sistematico, detta x_m la media aritmetica dei valori letti, secondo la scuola di pensiero cui vogliamo aderire scriviamo o:

$$\text{Misura di } X = x_m \pm \sqrt{\varepsilon_x^2 + \sigma_x^2} [x]$$

- ...oppure:

$$\text{Misura di } X = x_m \pm \sqrt{\varepsilon_x^2 + \sigma_m^2} [x]$$

ASSEGNAZIONE RISULTATO A UNA MISURA DIRETTA – 2.2

- ✖ Se abbiamo un solo strumento o apparato, la ripetizione (effettiva) delle misure ha dato letture diverse [dunque $\sigma_x > \varepsilon_x$] e si è stimato l' errore sistematico η_x , detta x_m la media aritmetica dei valori letti, secondo le scuola di pensiero cui vogliamo aderire scriveremo una delle seguenti 4 espressioni, i cui differenti significati sono già stati illustrati:

+ 1.

$$\text{Misura di } X = x_m \pm \sqrt{\varepsilon_x^2 + \sigma_x^2 + \eta_x^2} [x]$$

+ 2.

$$\text{Misura di } X = x_m \pm \sqrt{\varepsilon_x^2 + \sigma_m^2 + \eta_x^2} [x]$$

+ 3.

$$\text{Misura di } X = x_m \pm \sqrt{\varepsilon_x^2 + \sigma_x^2} \pm \eta_x [x]$$

+ 4.

$$\text{Misura di } X = x_m \pm \sqrt{\varepsilon_x^2 + \sigma_m^2} \pm \eta_x [x]$$

- ✖ In particolari casi si può determinare il segno dell' errore sistematico anche se il valore può solo essere stimato (altrimenti correggeremmo direttamente i dati): allora davanti a η_x non ci sarà \pm ma soltanto $+$ oppure soltanto $-$. In questo caso soltanto le ultime due espressioni sono utilizzabili.

ASSEGNAZIONE RISULTATO A UNA MISURA DIRETTA - 3

- Se abbiamo più strumenti o apparati, e dunque è stato possibile effettuare la randomizzazione, ricordiamo che il risultato finale della misura sarà espresso da:

$$\text{Misura di } X = x_{mp} \pm \sigma_x [x]$$

in cui x_{mp} è la media pesata definita così:

$$x_{mp} = \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i}$$

e i pesi valgono:

$$p_i = \frac{1}{\varepsilon_i^2}$$

MISURE INDIRETTE

- Sia una grandezza fisica W per la quale non è disponibile uno strumento o apparato che la misuri direttamente, ma della quale è nota una relazione con altre grandezze, diciamo X, Y, Z, \dots che invece sono misurabili direttamente:

$$W = f(X, Y, Z, \dots)$$

- Allora è certamente possibile pervenire a una misura indiretta di W misurando direttamente, ciascuna con un appropriato strumento o apparato, le grandezze X, Y, Z, \dots
- Dobbiamo individuare le procedure per esprimere sia il *risultato* della misura, w , sia l' *incertezza sperimentale* Δw con cui corredarlo:

$$\text{Misura di } W = w \pm \Delta w [w]$$

- Il risultato non presenta problemi: detti x, y, z, \dots i risultati delle misure delle grandezze X, Y, Z, \dots (letti, medie o medie pesate che ciascuno di essi sia), sarà semplicemente:

$$w = f(x, y, z, \dots)$$

INCERTEZZA IN MISURE INDIRETTE

- ✖ Ogni regola ragionevole per determinare Δw deve tenere conto sia delle incertezze con cui sono misurate (direttamente) le grandezze componenti ($\Delta x, \Delta y, \Delta z, \dots$,) sia di come W dipende da ciascuna di esse, nel senso che dovrebbe avere più peso, p.es. l' incertezza su una grandezza dalla quale W dipende quadraticamente piuttosto che una da cui dipende linearmente, etc.
- ✖ Se consideriamo le incertezze sperimentali, dal punto di vista matematico, come *variazioni* (del valore misurato rispetto al valore vero) allora l' operazione matematica che descrive il modo in cui una *variazione* di W dipende da *variazioni* della grandezze componenti è il **differenziale**:

$$dW = \frac{\partial W}{\partial x} dx + \frac{\partial W}{\partial y} dy + \frac{\partial W}{\partial z} dz + \dots$$

- ✖ Questa espressione, immaginando di calcolare le (funzioni) derivate per i valori di X, Y, Z, \dots risultati delle misure, e rimpiazzare i differenziali (dx, dy, dz, \dots), che sono entità matematiche, con le incertezze ($\Delta x, \Delta y, \Delta z, \dots$,), che sono quantità determinabili sperimentalmente con le procedure illustrate in precedenza, suggerisce il criterio per propagare le singole incertezze nell' incertezza cercata Δw .
- ✖ Sulla implementazione pratica di questo criterio esistono però, di nuovo, diverse possibilità alternative e diverse scuole di pensiero.

DETERMINAZIONE DI ΔW – ALTERNATIVE

- ✖ Le scuole di pensiero si dividono sia sulla formula precisa per combinare le incertezze dirette (Δx , Δy , Δz , ...), sia sulle regole per combinare i vari possibili contributi (ε_i , σ_i , η_i , ...) a ciascuna incertezza diretta Δi .
- ✖ Passerò in rassegna le molteplici possibili procedure alternative.

DETERMINAZIONE DI Δw - 1

- Un primo possibile approccio è il seguente:

1. Misuro direttamente le grandezze componenti in **maniera totalmente indipendente l' una dall' altra**, come cioè se ciascuna fosse l' unico obiettivo della misura. Applicando le procedure viste precedentemente perverrà alle seguenti determinazioni finali:

$$\text{Misura di } X = x \pm \Delta x [x]$$

$$\text{Misura di } Y = y \pm \Delta y [y]$$

$$\text{Misura di } Z = z \pm \Delta z [z]$$

etc....

2. Per determinare Δw mi metto nelle condizioni più prudenti, in cui tutte le grandezze componenti contribuiscono con tutta l' ampiezza del proprio intervallo di incertezza. La formula da usare è in tal caso la seguente:

$$\Delta w = \left| \frac{\partial W}{\partial x} \right| \Delta x + \left| \frac{\partial W}{\partial y} \right| \Delta y + \left| \frac{\partial W}{\partial z} \right| \Delta z + \dots$$

- Il motivo per cui si usano i valori assoluti delle derivate è che qualche derivata potrebbe essere negativa ma si vuole costruire un semi-intervallo di incertezza Δw tutto positivo: è il \pm davanti al Δw poi che estende tale intervallo ai due lati del valore misurato finale di w .
- Sottolineo che ciascun Δi è la combinazione di uno o più dei contributi ($\varepsilon_i, \sigma_i, \eta_i, \dots$) ma tale combinazione è effettuata internamente in ciascun Δi , separatamente e indipendentemente. Solo le Δi finali sono poi combinate insieme in questa formula.

DETERMINAZIONE DI $\Delta W - 2$

- ✖ Un secondo possibile approccio ([io preferisco questo](#)) è il seguente:
 1. Misuro direttamente le grandezze componenti in maniera **totalmente indipendente l' una dall' altra**, come cioè se ciascuna fosse l' unico obbiettivo della misura. Applicando le procedure viste precedentemente perverrò alle seguenti determinazioni finali:

$$\text{Misura di } X = x \pm \Delta x [x]$$

$$\text{Misura di } Z = z \pm \Delta z [z]$$

$$\text{Misura di } Y = y \pm \Delta y [y]$$

etc....

2. Per determinare Δw ricordo che ogni Δi rappresenta l' intervallo entro cui il valore vero inconoscibile della relativa grandezza i_v dovrebbe stare. Ora è improbabile che per tutte le grandezza componenti il valore vero stia dalla stessa parte dell' intervallo e a simile distanza dal valore misurato: piuttosto, **statisticamente, lato e distanza a cui il valore vero sta dovrebbero essere distribuiti casualmente tra le grandezze componenti**. Quindi per la grandezza misurata indirettamente un intervallo ridotto con criteri statistici sarà statisticamente sufficiente a contenerne il valore vero inconoscibile w_v . La formula da usare è in tal caso la seguente:

$$\Delta w = \sqrt{\left(\frac{\partial W}{\partial x} \Delta x\right)^2 + \left(\frac{\partial W}{\partial y} \Delta y\right)^2 + \left(\frac{\partial W}{\partial z} \Delta z\right)^2 + \dots}$$

3. Sottolineo che ciascun Δi è la combinazione di uno o più dei contributi ($\varepsilon_i, \sigma_i, \eta_i, \dots$) ma tale combinazione è effettuata internamente in ciascun Δi , separatamente e indipendentemente. Solo le Δi finali sono poi combinate insieme in questa formula.

DETERMINAZIONE DI ΔW – 3.1

- Un terzo approccio considera che:

- Una combinazione di questo tipo:

$$\Delta w = \sqrt{\left(\frac{\partial W}{\partial x} \Delta x\right)^2 + \left(\frac{\partial W}{\partial y} \Delta y\right)^2 + \left(\frac{\partial W}{\partial z} \Delta z\right)^2 + \dots}$$

sarebbe rigorosamente corretta, senza assunzione alcuna sulla posizione del valore vero all' interno dell' intervallo di incertezza, per i soli errori casuali σ , nel senso che si può dimostrare matematicamente che:

- amesso che le grandezze componenti siano misurate con strumenti sensibili alle variazioni casuali, cioè che $\sigma_i > \varepsilon_i$
 $\forall i$, l' intervallo che conterrebbe $\sim 68.3\%$ delle misure di w è dato da:

$$\sigma_w = \sqrt{\left(\frac{\partial W}{\partial x} \sigma_x\right)^2 + \left(\frac{\partial W}{\partial y} \sigma_y\right)^2 + \left(\frac{\partial W}{\partial z} \sigma_z\right)^2 + \dots}$$

- Del resto una combinazione dell' altro tipo:

$$\Delta w = \left| \frac{\partial W}{\partial x} \right| \Delta x + \left| \frac{\partial W}{\partial y} \right| \Delta y + \left| \frac{\partial W}{\partial z} \right| \Delta z + \dots$$

sarebbe invece appropriata per i soli errori assoluti max a priori ε oppure per i soli errori sistematici η , per i quali andrebbe così riscritta, rispettivamente:

$$\varepsilon_w = \left| \frac{\partial W}{\partial x} \right| \varepsilon_x + \left| \frac{\partial W}{\partial y} \right| \varepsilon_y + \left| \frac{\partial W}{\partial z} \right| \varepsilon_z + \dots$$

$$\eta_w = \frac{\partial W}{\partial x} \eta_x + \frac{\partial W}{\partial y} \eta_y + \frac{\partial W}{\partial z} \eta_z + \dots$$

- Il motivo per cui per η_w NON si usano i valori assoluti delle derivate è che si vuole propagare il segno di ogni dato η_i all' η_w della grandezza misurata indirettamente esplicitamente e correttamente.

DETERMINAZIONE DI ΔW – 3.2

- Continuando il terzo approccio, le considerazioni precedenti portano a proporre che si propaghino separatamente i contributi di ciascun tipo:

$$\varepsilon_w = \left| \frac{\partial W}{\partial x} \right| \varepsilon_x + \left| \frac{\partial W}{\partial y} \right| \varepsilon_y + \left| \frac{\partial W}{\partial z} \right| \varepsilon_z + \dots$$

$$\sigma_w = \sqrt{\left(\frac{\partial W}{\partial x} \sigma_x \right)^2 + \left(\frac{\partial W}{\partial y} \sigma_y \right)^2 + \left(\frac{\partial W}{\partial z} \sigma_z \right)^2 + \dots}$$

$$\eta_w = \frac{\partial W}{\partial x} \eta_x + \frac{\partial W}{\partial y} \eta_y + \frac{\partial W}{\partial z} \eta_z + \dots$$

- e in un secondo passo, si combinino tali contributi per formare l' incertezza sulla grandezza oggetto della misura:

$$\Delta w = \sqrt{\varepsilon_w^2 + \sigma_w^2} + \eta_w$$

- Si noti che in questo approccio è decisamente opportuno che η_w sia esposto a parte, eventualmente con un suo segno specifico, dato che è stato propagato in una maniera che gli permette di presentare un eventuale segno specifico.

DETERMINAZIONE DI ΔW – 4

- Un quarto approccio deriva dal terzo e io lo definisco l' approccio della speranza:

- I fautori di tale approccio sostengono che generalmente per effettuare una misura indiretta ci si sforza di procurarsi strumenti tali che ciascuna delle grandezze componenti sia misurata (direttamente) o con prevalenza di errori assoluti max a priori o di errori casuali, cioè che sia sempre o $\sigma_i << \varepsilon_i \forall i$ oppure $\sigma_i >> \varepsilon_i \forall i$, di modo che si possa usare o l' una o l' altra delle seguenti due ricette per propagare l' incertezza senza il problema di propagare contributi di natura diversa:

$$\varepsilon_w = \left| \frac{\partial W}{\partial x} \right| \varepsilon_x + \left| \frac{\partial W}{\partial y} \right| \varepsilon_y + \left| \frac{\partial W}{\partial z} \right| \varepsilon_z + \dots$$

$$\sigma_w = \sqrt{\left(\frac{\partial W}{\partial x} \sigma_x \right)^2 + \left(\frac{\partial W}{\partial y} \sigma_y \right)^2 + \left(\frac{\partial W}{\partial z} \sigma_z \right)^2 + \dots}$$

- Io ritengo che l' affermazione alla base (“che generalmente per effettuare una misura indiretta ci si sforza ...”) sia falsa: generalmente, a quanto mi consta, non si riesce affatto a fare così, anzi direi che è un caso piuttosto raro, solo una speranza, mentre in generale ogni misura diretta coinvolta avrà la sua dose indipendente di contributi assoluto max a priori e casuale.
- Basti pensare a una comunissima misura di velocità con strumenti che misurino separatamente spazio e tempo. La misura di s con un regolo ha un Δs che sarà dominato da ε_s , mentre la misura di t con un cronometro a start-stop manuale avrà un Δt dominato da σ_t .
- NOTA:** La propagazione del contributo sistematico usualmente non viene trattata da questi autori.

DETERMINAZIONE DI ΔW – CONCLUSIONE

- In conclusione tranne forse *in casi molto particolari che prenderei in considerazione però volta per volta, specie in casi particolari di comportamento di errori sistematici*, io propongo di adoperare sempre questa semplice procedura per **propagare l'incertezza in una misura indiretta**:

$$\Delta w = \sqrt{\left(\frac{\partial W}{\partial x} \Delta x\right)^2 + \left(\frac{\partial W}{\partial y} \Delta y\right)^2 + \left(\frac{\partial W}{\partial z} \Delta z\right)^2 + \dots}$$

in cui ciascun Δi è la combinazione di uno o più dei suoi contributi (ε_i , σ_i , η_i , ...) ma tale combinazione è effettuata internamente in ciascun Δi , separatamente e indipendentemente dagli altri, seguendo per ciascuna grandezza i le procedure già illustrate.

- Ma propagare separatamente i 3 contributi ciascuno con la tecnica per esso più appropriata e poi metterli insieme per formale **l'incertezza della misura indiretta** è altrettanto valido.

ESERCIZIO: σ DELLA MEDIA, σ_m

- Si perviene alla espressione di σ_m trattando *dal punto di vista matematico* la media come una grandezza “misurata” indirettamente a partire dai singoli valori, che giocano quindi il ruolo di “misure dirette”, e applicando la regola appropriata per la propagazione dei soli σ , che era:

$$\sigma_w = \sqrt{\left(\frac{\partial W}{\partial x} \sigma_x\right)^2 + \left(\frac{\partial W}{\partial y} \sigma_y\right)^2 + \left(\frac{\partial W}{\partial z} \sigma_z\right)^2 + \dots}$$

- In questa applicazione dunque la nostra W è la media definita così:

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

in cui le varie x_i rappresentano le grandezze “misurate direttamente” X, Y, Z, \dots e - ricordando che σ_x è un indice di come le varie x_i sono disperse attorno a x_m - ciascuna x_i è immaginata affetta da un errore casuale σ_x . Allora:

$$\sigma_m = \sqrt{\sum_{i=1}^n \left[\left(\frac{\partial x_m}{\partial x_i} \right)^2 \sigma_x^2 \right]} = \sqrt{n \left[\left(\frac{1}{n} \right)^2 \sigma_x^2 \right]} = \frac{\sigma_x}{\sqrt{n}}$$

ESERCIZIO: L^2 VS. $L \times L$

- Si voglia misurare (indirettamente) l' area di un **oggetto** quadrato mediante misura (diretta) dei lati.

1. Se esprimo l' area come $A=l^2$, dalla procedura per **propagare l'incertezza in una misura indiretta** ...

$$\Delta A = \sqrt{\left(\frac{\partial A}{\partial l} \Delta l\right)^2 + \left(\frac{\partial A}{\partial l} \Delta l\right)^2 + \left(\frac{\partial A}{\partial l} \Delta l\right)^2 + \dots} \quad (1)$$

...ottengo: $\Delta A = \sqrt{\left(\frac{\partial A}{\partial l} \Delta l\right)^2} = \frac{\partial A}{\partial l} \Delta l = 2l \Delta l \quad (2)$

2. Ma se la esprimo come $A=l \times l$ dalla stessa procedura ho:

$$\Delta A = \sqrt{\left(\frac{\partial A}{\partial l} \Delta l\right)^2 + \left(\frac{\partial A}{\partial l} \Delta l\right)^2} = \sqrt{2\left(\frac{\partial A}{\partial l} \Delta l\right)^2} = \sqrt{2(l \Delta l)^2} = \sqrt{2}l \Delta l \quad (3)$$

- La contraddizione è **apparente**. Le 2 espressioni di partenza dell' area sono equivalenti *matematicamente* ma non *sperimentalmente*, mentre la (1) è una formula per misure sperimentali. Sperimentalmente:

- La (2) significa che si misura un solo lato con incertezza Δl e si propaga tale incertezza alla misura dell' area.
- La (3) significa che si misurano **indipendentemente** due lati tra loro perpendicolari con due incertezze Δl_1 e Δl_2 , numericamente *di fatto* uguali $\Delta l_1 = \Delta l_2 = \Delta l$ se si è usato lo stesso regolo, e si propagano tali due incertezze. Tale combinazione dà una incertezza ΔA minore in quanto è improbabile che i **due** valori veri di l stiano dalla stessa parte e alla stessa distanza da quelli misurati. In altre parole, dal punto di vista sperimentale la seconda scelta considera in realtà il **quadrato** come un **rettangolo** di cui misurare indipendentemente l_1 e l_2 , ottenendo eventualmente due valori letti identici: $l_{1f} = l_{2f} = l_f$. Quindi l' incertezza ΔA più compiutamente va riscritta così:

$$\Delta A = \sqrt{\left(\frac{\partial A}{\partial l_1} \Delta l_1\right)^2 + \left(\frac{\partial A}{\partial l_2} \Delta l_2\right)^2} = \sqrt{(l_2 \Delta l_1)^2 + (l_1 \Delta l_2)^2}$$

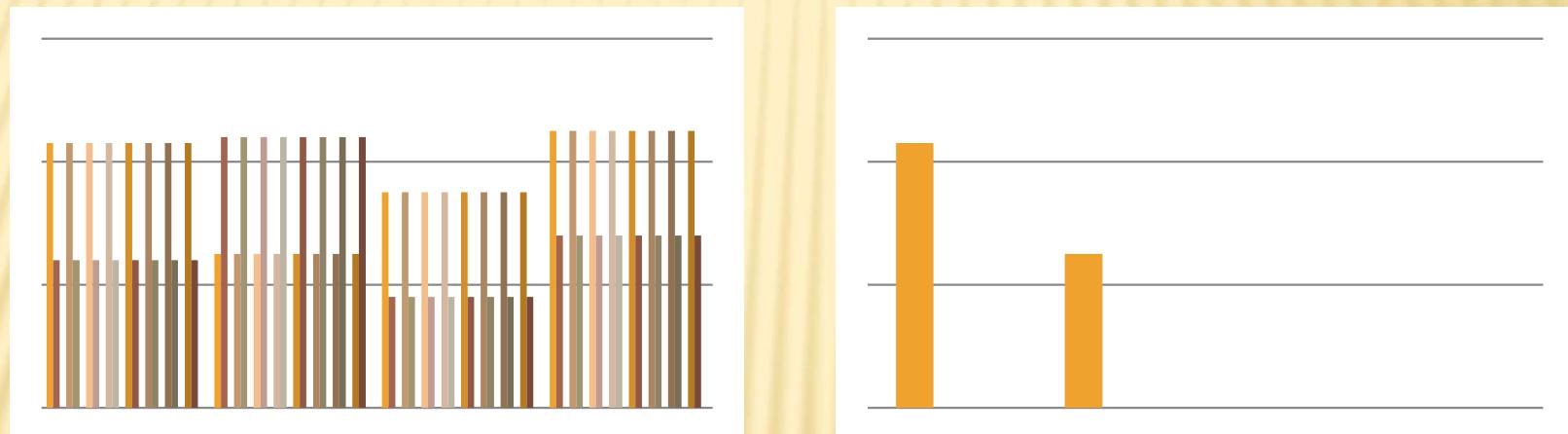
la quale, se $\Delta l_1 = \Delta l_2 = \Delta l$ e nell' ipotesi in cui le due letture siano veramente risultate uguali tra loro, si riduce appunto alla (3).

RAPPRESENTAZIONE DATI: ARROTONDAMENTO

- ✖ Misure dirette:
 - + Non riportare più cifre **significative** (le cifre che seguono eventuali 0 iniziali) di quelle effettivamente misurate.
 - **ESEMPIO:** Se si è misurata una lunghezza con una incertezza di **1 cm**, non scrivere $I=3.240\pm0.01\text{ m}$, e neanche $I=3.240\pm0.010\text{ m}$ ma $I=3.24\pm0.01\text{ m}$. Scrivere **3.240** solo se $\Delta I=1\text{ mm}$ cioè se si è in condizioni di scrivere $I=3.240\pm0.001\text{ m}$.
- ✖ Misure indirette:
 - + Dopo avere ricavato il valore di **w** da un calcolo, che può avere un numero arbitrario di cifre decimali, e dopo avere calcolato l' incertezza **Δw propagata**, il cui calcolo anche esso può avere un numero arbitrario di cifre decimali, arrotondare entrambi **armonizzandoli**, in modo che
 1. Δw non abbia più di 1 o 2 cifre **significative**
 2. **w** non abbia più cifre **significative** di quelle che Δw arrotondato indica come effettivamente misurate.
 - **ESEMPIO:** dal calcolo ho $w=3.4529378804$, $\Delta w=0.0416789052$. Arrotondare: $\Delta w=0.04$ o 0.042 e di conseguenza $w=3.45$ o $w=3.453$. Riportare infine:
➤ **$W=3.45\pm0.04 [w]$** o al massimo **$W=3.453\pm0.042 [w]$** .

RAPPRESENTAZIONE DATI: ISTOGRAMMI

- ✗ Negli histogrammi *di frequenza* è importantissima la scelta del numero di ***bins*** (intervalli dell' ascissa):
 - + troppi rendono ciascun ***bin*** poco popolato e l' andamento “zoppicante”;
 - + troppo pochi non fanno *visualizzare* alcun andamento:



- ✗ Due linee guida alternative garantiscono entrambe un corretto raggruppamento dei dati:

$$n_{bins} \approx \sqrt{n_{dati}}$$

$$\text{ampiezza}_{bin} \approx \frac{x_{\max} - x_{\min}}{n_{dati}/3}$$

MISURE IN CORRELAZIONE

- ✖ Una grande frazione di esperimenti di ricerca e di esperienze didattiche ha come obiettivo non la misura diretta o indiretta di **una singola** grandezza fisica, ma lo studio di **una correlazione tra due** (o più) **grandezze fisiche diverse**.
- ✖ Tale studio può avere uno di 3 scopi:
 1. Scoprire una relazione non ancora nota tra le due grandezze
 2. Verificare una legge nota che lega le due grandezze
 3. Misurare (*indirettamente!*) una **terza grandezza** che appare come parametro nella legge nota che lega le due grandezze misurate
- ✖ In generale una tale esperienza utilizza un apparato complesso il quale deve consentire di
 1. Variare a piacere i valori di una delle due grandezze e rilevarli ogni volta.
 2. Causare “automaticamente” una variazione della seconda grandezza quando si varia la prima.
 3. Misurare di volta in volta i valori della seconda grandezza, conseguenti alle variazioni della prima.

MISURE INDIRETTE DA CORRELAZIONE

Quando lo scopo è **determinare una terza grandezza**, generalmente si può fare con due tipi diversi di analisi dei dati.

- ✖ Si consideri come esempio la legge del moto rettilineo uniforme: $s=vt$ (v costante!) in cui si voglia determinare v potendo misurare solo spazi e tempi.
 - ✖ Presumiamo di avere un apparato consistente in un regolo, un cronometro e un dispositivo in cui un corpo si muova a velocità costante, come il viscosimetro a velocità di caduta limite.
 - ✖ Si decidono a piacere tratti di volta in volta diversi sul regolo. La misura di ogni tratto consiste nella lettura di due posizioni, iniziale e finale (decise da noi), e nel calcolo della loro differenza.
 - ✖ Si misuri per ogni tratto deciso il tempo con il cronometro. Questo tempo dipende automaticamente dalla lunghezza del tratto, come richiesto.
1. Si esegue un *best-fit con il metodo dei minimi quadrati*, che trova i parametri della curva scelta (in questo caso la retta) che meglio approssima le posizioni dei punti. Si ricava la grandezza cercata dal parametro in cui è contenuta. In questo caso il coefficiente angolare rappresenta proprio v . E' evidente che questa è una forma di misura *indiretta*.
 2. Si ricava un v_i da ogni coppia di dati (s_i, t_i) [queste NON sono ripetizioni della stessa misura, ma misure diverse con valori che per s_i sono diversi di proposito e per t_i sono diversi di conseguenza!], indi si fa la loro media pesata v_{mp} . Concettualmente questa procedura è equivalente a fare una misura di v con una serie di "apparati" diversi, ciascuno dei quali fornisca una misura *indiretta* della stessa, dunque è una specie di misura *indiretta* ma effettuata mediante *randomizzazione*.
- ✖ In parecchie delle esperienze di Lab2 una grandezza può essere misurata *indirettamente* in questi due modi. Entrambi vanno eseguiti e i due valori ottenuti vanno confrontati, con commento, alla fine.

STUDIO DI UNA CORRELAZIONE

- Dette genericamente **X** e **Y** due grandezze di cui si vuole studiare la correlazione, la tipica **sequenza** di passi di questo studio è la seguente:

NOTA: i dettagli tecnici di ciascun passo saranno richiamati più avanti

1. Raccolta dei dati [coppie (x_i, y_i)] con la metodologia sperimentale detta.
 - Ciascun dato, cioè ciascun elemento di ciascuna coppia è una misura diretta, dunque composta di: *risultato, incertezza e unità di misura* secondo le procedure viste prima.
2. Creazione di un grafico in cui le coppie sono riportate come punti, corredati da barre orizzontali e verticali che rappresentano le incertezze.
 - In qualche caso tali barre non vengono riportate perché troppo corte rispetto alle dimensioni dei simboli e in tal caso va **detto/scritto**.
3. Per esperimento di scoperta:
 1. Formulazione di una o più ipotesi sul tipo di funzione matematica (curva) che potrebbe rappresentare i dati.
 2. Esecuzione di un best-fit con ciascuna delle curve candidate, alla ricerca dei parametri della migliore curva di ogni tipo.
 3. Calcolo del χ^2 per la migliore curva di ciascun tipo trovata dal best fit, al fine di scegliere la funzione che meglio rappresenta i dati.
 4. Se il χ^2 è “troppo alto” per tutte le curve ragionevolmente ipotizzabili si conclude che nessuna correlazione statisticamente significativa è presente tra le due grandezze in quell’ esperimento.
4. Per esperimento di verifica o di misura indiretta di grandezza terza:
 1. Esecuzione di un best-fit con la funzione nota, alla ricerca dei parametri della migliore curva di quel tipo.
 2. Calcolo del χ^2 per quella curva.
 3. Se il χ^2 è “troppo alto” si conclude che la correlazione attesa non è presente da un punto di vista statistico nell’ esperienza eseguita e che dunque la stessa deve essere incorsa in qualche problema, tecnico e/o metodologico.

CORRELAZIONI LINEARI

- ✖ Molte leggi fisiche sono lineari.
 - ✖ Molte di più non lo sono.
 - ✖ Ma molte di queste ultime sono linearizzabili con un cambio di variabili.
- ✖ Es: $V = V_0 e^{-\frac{t}{RC}} \Rightarrow \ln V = \ln V_0 + \left(-\frac{1}{RC}\right) \cdot t \Rightarrow (\ln V) = at + b$
- ✖ Il metodo dei min quadrati per l' individuazione della migliore curva di un dato tipo funzionale è in linea di principio applicabile a qualunque relazione funzionale.
 - ✖ Le formule per i parametri sono piuttosto complesse per funzioni non lineari.
 - ✖ Molti programmi di analisi ormai comprendono un vasto menu di funzioni con cui eseguire best-fits.
 - ✖ Ogni volta che è matematicamente possibile, conviene comunque linearizzare la funzione da usare, in quanto:
 1. L' umano giudica più facilmente, in un grafico, se una serie di punti appare ragionevolmente allineata lungo una retta.
 2. L' esistenza di una correlazione lineare ipotizzata deve essere verificata preventivamente (cioè prima di avventurarsi a trovare i parametri della migliore retta e a calcolare il χ^2) in maniera statisticamente oggettiva mediante il **coefficiente di correlazione lineare r di Bravais-Pearson**.
 - ✖ Sperimentalmente, la linearizzazione comporta:
 1. Trasformare le coppie misurate nelle nuove variabili che dovrebbero rispondere a una legge lineare → le misure, originariamente *dirette*, possono diventare *indirette* e le loro **incertezze** vanno determinate di conseguenza.
 2. Graficare, calcolare r , eseguire il best-fit, calcolare il χ^2 etc, con le nuove coppie trasformate dal cambio di variabili.

STUDIO DI UNA CORRELAZIONE LINEARE

- ✖ Dette genericamente **X** e **Y** due grandezze tra le quali si ipotizza una **correlazione lineare**, da studiare in dettaglio, la tipica **sequenza** di passi di questo studio è la seguente:
- ✖ **NOTA:** i dettagli tecnici di ciascun passo saranno illustrati più avanti
 1. Raccolta dei dati [coppie (x_i, y_i)] con la metodologia sperimentale detta.
 - Ciascun dato, cioè ciascun elemento di ciascuna coppia è una misura diretta, dunque composta di: *risultato, incertezza e unità di misura* secondo le procedure viste prima.
 2. Creazione di un grafico in cui le coppie sono riportate come punti, corredati da barre orizzontali e verticali che rappresentano le incertezze.
 - In qualche caso tali barre non vengono riportate perché troppo corte rispetto alle dimensioni dei simboli e in tal caso va detto/scritto.
 3. Calcolo del **coefficiente di correlazione lineare r di Bravais-Pearson** per verificare la significatività della ipotizzata correlazione lineare
 - Se **r** non supporta l' ipotesi di correlazione lineare essa viene abbandonata e lo studio si ferma qui.
 4. Esecuzione di un best-fit con la funzione **$Y=aX+b$** , alla ricerca dei valori dei parametri **a** e **b** della retta che meglio approssima l' insieme di tutti i punti.
 5. Calcolo del **χ^2** per quella retta.
 1. Se il **χ^2** è “troppo alto” si conclude che la correlazione attesa non è presente da un punto di vista statistico nell’ esperienza eseguita e che dunque la stessa deve essere incorsa in qualche problema, tecnico e/o metodologico.

DETTAGLI PASSI STUDIO CORRELAZIONI

- ✖ Di seguito richiamo qualche dettaglio sui seguenti passi:
 1. Raccolta dati
 2. Rappresentazione grafica
 3. Coefficiente di correlazione (*per relazioni lineari*)
 4. Best-fit(s) con il metodo dei minimi quadrati
 5. Calcolo del χ^2
 6. Eventuale determinazione dell' incertezza sulla misura *indiretta* di una grandezza terza da un parametro del fit

RACCOLTA DATI

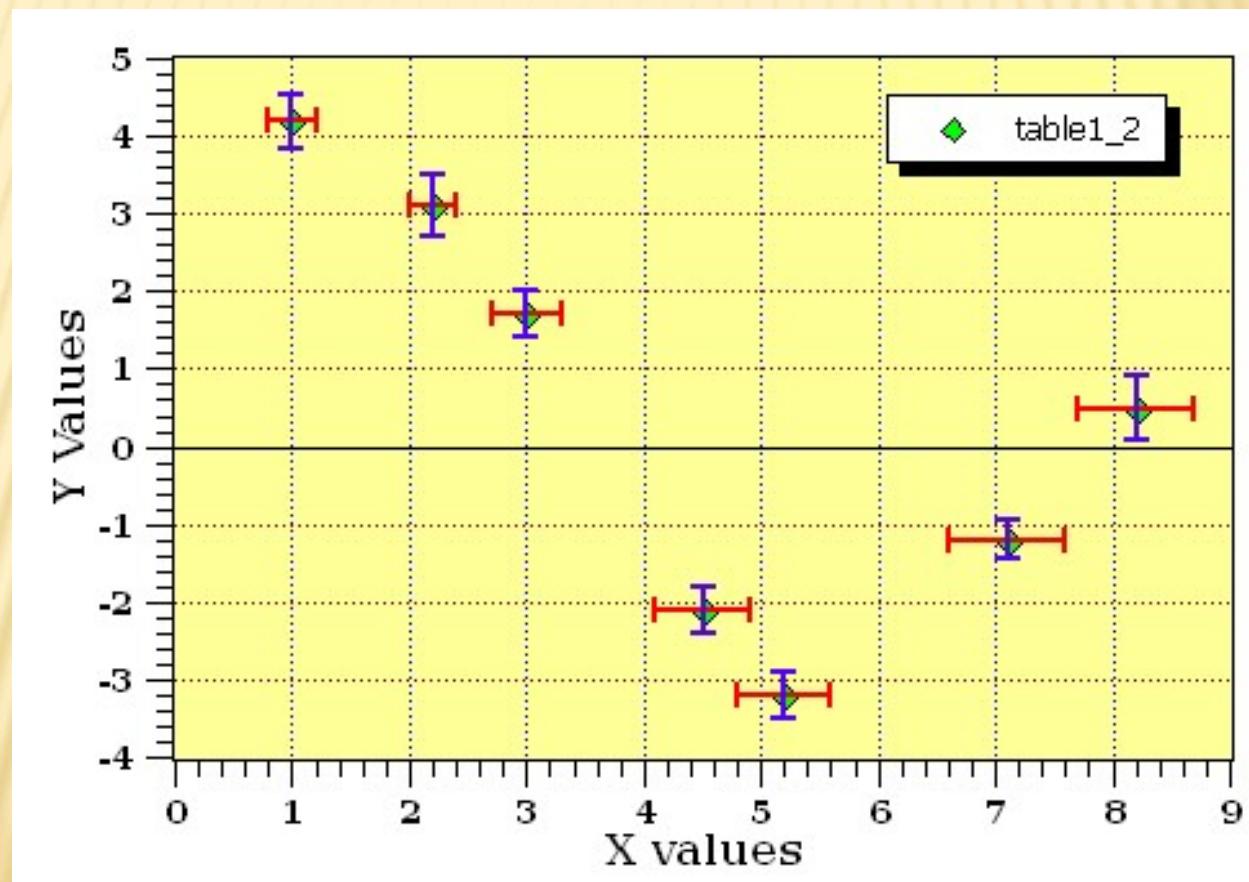
- Tipicamente i dati raccolti per uno studio di correlazione vanno presentati in tabelle, riportando anche le incertezze per ogni singolo dato:

$X [x]$	$\Delta X [x]$	$Y [y]$	$\Delta Y [y]$	
3.45	0.01	740	5	ε_i
4.56	0.04	700	5	$\sqrt{(\varepsilon_i^2 + \sigma_i^2)}$
7.01	0.02	620	5	σ_i
12.22	0.01	533	5	
30.37	0.03	392	5	
63.80	0.04	210	5	
...	

Se le incertezze di una serie risultano tutte uguali si possono omettere dalla tabella e dichiarare nel testo una sola volta

RAPPRESENTAZIONE GRAFICA

- Prima di ulteriori analisi, presentare i dati bruti in grafico, riportando sia i punti che le barre di incertezza:



COEFFICIENTE DI CORRELAZIONE LINEARE - 1

- Se la correlazione ipotizzata o da verificare è lineare, dopo avere comunque presentato i dati in tabella e in grafico, *ma prima di azzardare un best-fit*, è necessario calcolare il **coefficiente di correlazione lineare r di Bravais-Pearson**, per verificare la significatività della ipotizzata correlazione lineare:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

- I simboli di valore medio barrati NON indicano una *media di misure ripetute* nelle stesse condizioni come per stimare l' errore casuale, che ho finora denotato con x_m , ma una media aritmetica dei valori misurati, che sono diversi perché di proposito si sono realizzate condizioni diverse per avere valori diversi della grandezza **X** (e di conseguenza di **Y**).
- Per costruzione, $-1 \leq r \leq 1$. Più r si avvicina a **|1|**, più significativa è la correlazione.
- Va bene, ma dove stabilisco la **soglia di confidenza**, cioè quando la considero **abbastanza** significativa da andare avanti con l' analisi piuttosto che “abbandonarla”?

continua ➔

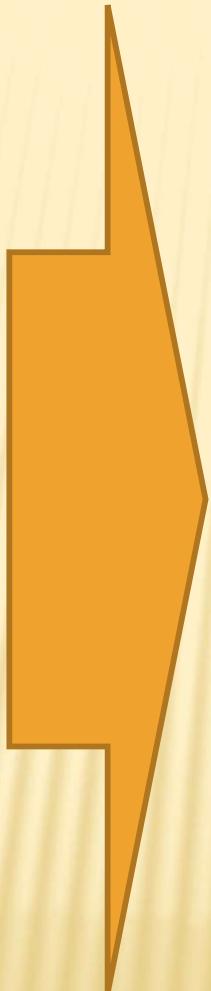
COEFFICIENTE DI CORRELAZIONE LINEARE - 2

- ✖ Devo avere un **criterio di soglia oggettivo**. Questo criterio esiste ma purtroppo non è deterministico (del tipo: “ $r < 0.9$ non è significativa, $r \geq 0.9$ si”) bensì soltanto probabilistico:
- ✖ Statisticamente, *dato il numero di punti sperimentali*, si può dire qual è la **probabilità che un insieme di dati X-Y non correlati da un rapporto causa→effetto, cioè da una legge fisica (lineare) presenti per puro caso un valore di $|r| \geq$ di quello trovato.**
 - Per esempio, se ci sono **5** punti, la probabilità che $|r| \geq 0.8$ per puro caso è del 10%.
- ✖ Dopodiché si definiscono non una, ma due soglie:
 1. Se la suddetta probabilità è **<5%** si dice che nei dati si riscontra una **correlazione lineare significativa**:
 - solo nel 5% dei casi un pari numero di punti in realtà non correlati potrebbe presentare almeno quel valore di **r** .
 2. Se la suddetta probabilità è **<1%** si dice che nei dati si riscontra una **correlazione lineare altamente significativa**:
 - solo nel 1% dei casi un pari numero di punti in realtà non correlati potrebbe presentare almeno quel valore di **r** .
- In genere, si “accetta” la correlazione e si prosegue l’ analisi se in base al valore di **r** la si può dichiarare **significativa**.

continua ➔

COEFFICIENTE DI CORRELAZIONE LINEARE - 3

Abbiamo dunque
bisogno di una
**tabella di
riferimento**
che ci dica la
**probabilità che $|r|$
sia \geq di un valore
dato per un insieme
di n dati scorrelati, al
variare sia del valore
dato, che di n .**



$N \setminus r_m$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
3	100	94	87	81	74	67	59	51	41	29	0
4	100	90	80	70	60	50	40	30	20	10	0
5	100	87	75	62	50	39	28	19	10	3.7	0
6	100	85	70	56	43	31	21	12	5.6	1.4	0
7	100	83	67	51	37	25	15	8.0	3.1	—	0
8	100	81	63	47	33	21	12	5.3	1.7	—	0
9	100	80	61	43	29	17	8.8	3.6	—	—	0
10	100	78	58	40	25	14	6.7	2.4	—	—	0
11	100	77	56	37	22	12	5.1	1.6	—	—	0
12	100	76	53	34	20	9.8	3.9	1.1	—	—	0
13	100	75	51	32	18	8.2	3.0	—	—	—	0
14	100	73	49	30	16	6.9	2.3	—	—	—	0
15	100	72	47	28	14	5.8	1.8	—	—	—	0
16	100	71	46	26	12	4.9	1.4	—	—	—	0
17	100	70	44	24	11	4.1	1.1	—	—	—	0
18	100	69	43	23	10	3.5	—	—	—	—	0
19	100	68	41	21	9.0	2.9	—	—	—	—	0
20	100	67	40	20	8.1	2.5	—	—	—	—	0
25	100	63	34	15	4.8	1.1	—	—	—	—	0
30	100	60	29	11	2.9	—	—	—	—	—	0
35	100	57	25	8.0	1.7	—	—	—	—	—	0
40	100	54	22	6.0	1.1	—	—	—	—	—	0
45	100	51	19	4.5	—	—	—	—	—	—	0
50	100	49	16	3.4	—	—	—	—	—	—	0
60	100	45	13	2.0	—	—	—	—	—	—	0
70	100	41	9.7	1.2	—	—	—	—	—	—	0
80	100	38	7.5	—	—	—	—	—	—	—	0
90	100	35	5.9	—	—	—	—	—	—	—	0
100	100	32	4.6	—	—	—	—	—	—	—	0

BEST-FIT – METODO MINIMI QUADRATI

- Sia $\mathbf{Y} = \mathbf{f}(\mathbf{X})$ la funzione che rappresenta la correlazione ipotizzata. In generale essa conterrà s parametri $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_s$:

$$\mathbf{Y} = \mathbf{f}(\mathbf{m}_k; \mathbf{X})$$

- Si cercano i valori dei parametri che rendono minima la somma dei quadrati degli scarti dei valori misurati y_i dai valori di \mathbf{Y} predetti dalla funzione per \mathbf{X} pari ai valori misurati di x_i , ossia che rendono nulle le seguenti espressioni, al variare di \mathbf{m}_k :

$$\frac{d}{dm_k} \sum_{i=1}^n [y_i - f(m_k; x_i)]^2$$

- L'esplicitazione di questa formula per ogni funzione particolare fornisce le espressioni dei *suoi* migliori parametri \mathbf{m}_k , cioè dei parametri che per *quella forma funzionale* minimizzano le discrepanze tra curva e punti.
- La procedura, di per sé, non dice affatto se questa funzione è appropriata a descrivere i dati, né quale funzione, tra varie alternative, si adatti meglio ai dati. Questo sarà il compito del χ^2 .

BEST-FIT – SCELTA VARIABILI

- ✖ Matematicamente, il metodo dei minimi quadrati ipotizza che la variabile indipendente **X** sia esente da incertezze.
- ✖ Sperimentalmente questo non si verifica mai, dunque quello che nella pratica si deve fare per avvicinarsi alla situazione teorica è:
una scelta consapevole della variabile indipendente.
- ✖ Bisogna cioè, prima di eseguire il best-fit, porre la funzione da utilizzare in forma tale che il ruolo di variabile indipendente **X** sia giocato dalla grandezza la cui *somma* (oppure la *media*, è lo stesso) delle *incertezze relative* è minore:

$$\sum_{i=1}^n \frac{\Delta x_i}{x_i} < \sum_{i=1}^n \frac{\Delta y_i}{y_i} \Rightarrow \frac{\sum_{i=1}^n \frac{\Delta x_i}{x_i}}{n} < \frac{\sum_{i=1}^n \frac{\Delta y_i}{y_i}}{n} \Rightarrow \left\langle \frac{\Delta x_i}{x_i} \right\rangle < \left\langle \frac{\Delta y_i}{y_i} \right\rangle$$

BEST-FIT PESATO O NON?

- ✖ Dopo avere deciso quale variabile sarà considerata indipendente, **X**, e di conseguenza quale è quella dipendente, **Y**, bisogna decidere se si deve eseguire il best-fit **pesato** o se è sufficiente quello **non pesato**.
 - Tecnicamente, nulla osta a eseguire sempre il best-fit **pesato**
 - Ma le formule per ricavare i parametri nel caso **pesato** sono più complesse.
- 1. Si può eseguire il best-fit **non pesato** se le incertezze sui vari valori **y_i** misurati sono all'incirca costanti:
- 2. Si deve eseguire il best-fit **pesato** se le incertezze sui vari valori **y_i** misurati sono significativamente diverse tra loro.
 - + L'idea è dare maggior peso ai punti la cui incertezza su **Y** è minore, in modo che la procedura tenda a "avvicinare" di più a essi la curva che troverà.
 - + I pesi da usare per ogni punto sono l'inverso dei quadrati delle incertezze su **Y**:

$$p_i = \frac{1}{\Delta y_i^2}$$

PARAMETRI (PER BEST-FIT LINEARE)

- Funzione con cui fissare: $Y=aX+b$

1. Best-fit pesato:

$$a = \frac{\left(\sum_{i=1}^n p_i \right) \times \left(\sum_{i=1}^n p_i x_i y_i \right) - \left(\sum_{i=1}^n p_i x_i \right) \times \left(\sum_{i=1}^n p_i y_i \right)}{\left(\sum_{i=1}^n p_i \right) \times \left(\sum_{i=1}^n p_i x_i^2 \right) - \left(\sum_{i=1}^n p_i x_i \right)^2}$$

$$b = \frac{\left(\sum_{i=1}^n p_i x_i^2 \right) \times \left(\sum_{i=1}^n p_i y_i \right) - \left(\sum_{i=1}^n p_i x_i \right) \times \left(\sum_{i=1}^n p_i x_i y_i \right)}{\left(\sum_{i=1}^n p_i \right) \times \left(\sum_{i=1}^n p_i x_i^2 \right) - \left(\sum_{i=1}^n p_i x_i \right)^2}$$

$$\Delta a = \sqrt{\frac{\left(\sum_{i=1}^n p_i \right) \times \left(\sum_{i=1}^n p_i (y_i - ax_i - b)^2 \right)}{n - 2}}$$

$$\Delta b = \sqrt{\frac{\left(\sum_{i=1}^n p_i x_i \right)^2 \times \left(\sum_{i=1}^n p_i (y_i - ax_i - b)^2 \right)}{\left(\sum_{i=1}^n p_i \right) \times \left(\sum_{i=1}^n p_i x_i^2 \right) - \left(\sum_{i=1}^n p_i x_i \right)^2}}$$

2. Best-fit non pesato:

$$a = \frac{n \times \left(\sum_{i=1}^n x_i y_i \right) - \left(\sum_{i=1}^n x_i \right) \times \left(\sum_{i=1}^n y_i \right)}{n \times \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right)^2}$$

$$\Delta a = \sqrt{\frac{\left(\sum_{i=1}^n (y_i - ax_i - b)^2 \right)}{n - 2}}$$

$$b = \frac{\left(\sum_{i=1}^n x_i^2 \right) \times \left(\sum_{i=1}^n y_i \right) - \left(\sum_{i=1}^n x_i \right) \times \left(\sum_{i=1}^n x_i y_i \right)}{n \times \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right)^2}$$

$$\Delta b = \sqrt{\frac{\left(\sum_{i=1}^n x_i^2 \right) \times \left(\sum_{i=1}^n (y_i - ax_i - b)^2 \right)}{n \times \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right)^2}}$$

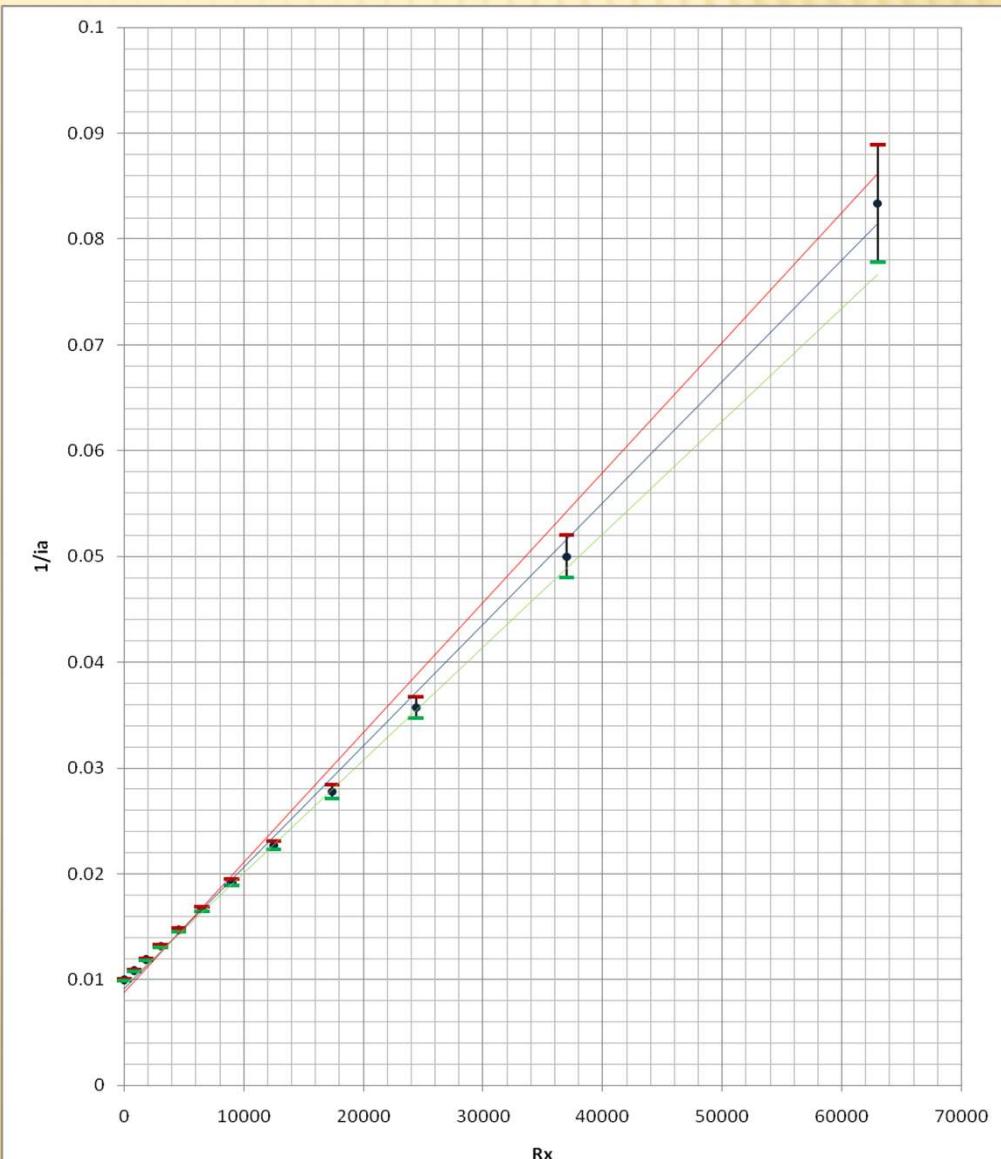
BEST-FIT LINEARE - CASI PARTICOLARI: $b=0$; a, b VINCOLATI

- ✖ Molte leggi fisiche hanno forma funzionale di retta passante per l' origine. La funzione con cui fissare i punti è allora: $\textcolor{red}{Y=aX}$
- ✖ In questo caso si deve comunque effettuare inizialmente il fit con $\textcolor{red}{Y=aX+b}$ e confrontare il valore trovato dal fit, corredato della sua incertezza, $\textcolor{red}{b \pm \Delta b}$, con $\textcolor{red}{0}$.
 1. Se $\textcolor{red}{b-\Delta b < 0 < b+\Delta b}$, b trovato dal fit è statisticamente compatibile con $\textcolor{red}{0}$, la legge prevista è statisticamente verificata dai dati raccolti e si passa allora al fit con la funzione ipotizzata $\textcolor{red}{Y=aX}$ in quanto in tal caso questo fornisce la stima più accurata possibile di $\textcolor{red}{a}$.
 2. In caso contrario vuol dire che ci sono degli errori sistematici nelle misure di $\textcolor{red}{X}$ o di $\textcolor{red}{Y}$ o di **entrambe**.
 - In alcuni casi non sarà possibile decidere se l' errore sistematico affligge $\textcolor{red}{X}$ o $\textcolor{red}{Y}$ o **entrambe**.
 - In altri potrebbe essere possibile risalire alle sorgenti degli errori sistematici, determinarli quantitativamente, correggere i dati, e infine rifare il fit con $\textcolor{red}{Y=aX}$.
- ✖ In maniera analoga, se teoricamente $\textcolor{red}{a}$ o $\textcolor{red}{b}$ sono vincolati ad avere un qualche valore preciso anche se non nullo, cioè se dovrebbero avere valore teorico $\textcolor{red}{a_t}$ o $\textcolor{red}{b_t}$, si deve fare inizialmente comunque il fit con $\textcolor{red}{Y=aX+b}$.
 - Le deviazioni di $\textcolor{red}{a}$ o $\textcolor{red}{b}$ trovati dal fit dai valori previsti teoricamente, $\textcolor{red}{a_t}$ o $\textcolor{red}{b_t}$, possono fare scoprire errori sistematici nelle misure di $\textcolor{red}{X}$ o di $\textcolor{red}{Y}$ o di **entrambe**.
 - Nel caso che sia possibile risalire alle sorgenti degli errori sistematici e determinarli quantitativamente, correggere i dati e rifare il fit vincolato con $\textcolor{red}{Y=aX+b_t}$ oppure $\textcolor{red}{Y=a_tX+b}$.

BEST-FIT LINEARE: ESEMPIO ERRORE SISTEMATICO

- ✗ In questo fit con $Y=aX+b$ risultò $b=0.0092\pm0.0002$, dunque non compatibile con 0 .
- ✗ Tutte le Y (1/corrente) potrebbero essere troppo grandi di $\sim 0.009 \mu A^{-1}$ oppure tutte le X (resistenza) potrebbero essere troppo basse di 10000Ω .
- ✗ Che una resistenza sia starata di 10000Ω sembra improbabile, mentre che un amperometro sia starato di $\sim (0.009)^{-1} \mu A$ è piuttosto comune e potrebbe anche essere difficile da notare sullo strumento.
- Il fit completo ha fatto scoprire un errore sistematico sulla Y !

Analogamente si ragiona per valori di a o b vincolati teoricamente.



PARAMETRI (PER BEST-FIT LINEARE VINCOLATO)

Funzione con cui fissare ➤	$Y=aX$	$Y=aX+b_t$	$Y=a_tX+b$
Best-fit pesato:	$a = \frac{\left(\sum_{i=1}^n p_i x_i y_i \right)}{\left(\sum_{i=1}^n p_i x_i^2 \right)}$ $\Delta a = \sqrt{\frac{\left(\sum_{i=1}^n p_i (y_i - ax_i)^2 \right)}{n - 1}}$	$a = \frac{\left(\sum_{i=1}^n p_i x_i y_i \right) - \left(\sum_{i=1}^n p_i x_i \right) \times b_t}{\left(\sum_{i=1}^n p_i x_i^2 \right)}$ $\Delta a = \sqrt{\frac{\left(\sum_{i=1}^n p_i (y_i - ax_i - b_t)^2 \right)}{n - 1}}$	$b = \frac{\left(\sum_{i=1}^n p_i y_i \right) - \left(\sum_{i=1}^n p_i x_i \right) \times a_t}{\left(\sum_{i=1}^n p_i \right)}$ $\Delta b = \sqrt{\frac{\left(\sum_{i=1}^n p_i (y_i - a_t x_i - b)^2 \right)}{n - 1}}$
Best-fit non pesato:	$a = \frac{\left(\sum_{i=1}^n x_i y_i \right)}{\left(\sum_{i=1}^n x_i^2 \right)}$ $\Delta a = \sqrt{\frac{\left(\sum_{i=1}^n (y_i - ax_i)^2 \right)}{n - 1}}$	$a = \frac{\left(\sum_{i=1}^n x_i y_i \right) - \left(\sum_{i=1}^n x_i \right) \times b_t}{\left(\sum_{i=1}^n x_i^2 \right)}$ $\Delta a = \sqrt{\frac{\left(\sum_{i=1}^n (y_i - ax_i - b_t)^2 \right)}{n - 1}}$	$b = \frac{\left(\sum_{i=1}^n y_i \right) - \left(\sum_{i=1}^n x_i \right) \times a_t}{n}$ $\Delta b = \sqrt{\frac{\left(\sum_{i=1}^n (y_i - a_t x_i - b)^2 \right)}{n(n - 1)}}$

CHI QUADRO – DEFINIZIONE

- ✖ Un best-fit trova la migliore curva di un tipo assegnato ma non dice nulla quantitativamente se effettivamente essa si adatta bene ai dati; inoltre se si eseguono più fits con funzioni assegnate diverse la procedura non “sceglie” la migliore, trova solo i migliori parametri di ciascuna.
- ✖ Un parametro oggettivo (un po' come è oggettivo r , e basato sulla stessa filosofia) è il χ^2 .
- ✖ Dato un fit si definisce questo parametro:

$$\chi^2 = \sum_{i=1}^n \left(\frac{y_i^{fit} - y_i}{\sigma_i} \right)^2$$

in cui y_i^{fit} sono i valori di restituiti dalla funzione fissata in corrispondenza dei valori x_i misurati e y_i sono i valori misurati. σ_i sono i soli contributi casuali alle incertezze Δy_i dei valori y_i misurati.

- ✖ Sulla base dell' ipotesi che gli scarti dei punti misurati dall' andamento matematico trovato dal fit dovrebbero essere di natura accidentale, allora mediamente il rapporto tra scarto dalla curva (numeratore) e errore accidentale della misura (denominatore) dovrebbe valere in valore assoluto ≈ 1 . Dunque teoricamente se la curva con cui si è fissato rappresenta veramente l' andamento dei dati dovrebbe essere $\chi^2 \approx n$, o più precisamente $\chi^2 \approx n-s$ ove è il numero di parametri della curva.
- ✖ Più spesso si usa il “ χ^2 ridotto” definito come $\chi^2/(n-s)$ il quale dunque dovrebbe essere ≈ 1 .

CHI QUADRO - USO

- ✖ Quanto più $\chi^2/(n-s)$ è **> 1**, tanto più dubbia è l' applicabilità di quella curva ai dati (*non peggiore è il fit, come talvolta si legge!*) .
- ✖ Allora se c' è da scegliere tra curve diverse, si sceglierà quella che restituisce il $\chi^2/(n-s)$ più piccolo.
- ✖ Se per tutte le curve $\chi^2/(n-s) >> 1$ si conclude che nessuna si adatta ai dati.
- ✖ Se si deve valutare una sola forma funzionale, e risulta $\chi^2/(n-s) >> 1$, si conclude che essa non si adatta ai dati.
- ✖ Se risulta $\chi^2/(n-s) << 1$ allora gli errori σ_i inseriti sono stati sovrastimati. Si può “giocare” riducendoli *ad arte* progressivamente fino a che si ottiene $\chi^2/(n-s) \cong 1 \rightarrow$ i valori ridotti rappresentano approssimativamente gli errori accidentali reali.
- ✖ Un caso particolare di questa casistica è che spesso non è possibile inserire nel calcolo del χ^2 i soli errori accidentali in quanto risultarono non rilevabili, cioè $\sigma_i < \varepsilon_i$, almeno per alcuni punti. In tal caso non si può fare altro che usare al loro posto gli ε_i ; se risulta $\chi^2/(n-s) << 1$ di nuovo si riducono i valori di questi ε_i *ad arte*, progressivamente fino a che si ottiene $\chi^2/(n-s) \cong 1 \rightarrow$ i valori così ridotti rappresentano approssimativamente gli errori accidentali σ_i reali, che dunque è **possibile stimare a posteriori!**

MISURA INDIRETTA DA BEST-FIT

- Se lo scopo ultimo del best-fit era di dedurre il valore di una **terza grandezza Z** contenuta in uno dei parametri della curva, l' incertezza ΔZ deve essere dedotta da quella sul parametro di fit mediante la normale propagazione degli errori.
- Le incertezze sulle grandezza effettivamente misurate **X** e **Y** non vengono utilizzate esplicitamente nell' espressione di quella di **Z** , ma si propagano implicitamente, sia attraverso le scelte fatte (della variabile indipendente, se pesato o non) sia tramite l' incertezza sul parametro, che esiste proprio perché la curva non passa esattamente per tutti i punti, e a sua volta questo avviene perché le misure sono affette dalle loro incertezze.
- Detto **m_k** un generico parametro del fit e **Δm_k** la sua incertezza (come **Δa** e **Δb** per la retta), e siano **T, \dots** altre grandezze misurate che possono entrare insieme a **Z** nell' espressione fisica del parametro si ha:

$$z = f(m_k)$$

e:

$$\Delta z = \sqrt{\left(\frac{\partial z}{\partial m_k} \Delta m_k \right)^2 + \left(\frac{\partial z}{\partial t} \Delta t \right)^2 + \dots}$$

Lezione n.24

RICHIAMO: PASSI INDISPENSABILI DELLO STUDIO DI UNA CORRELAZIONE

LISTA PASSI STUDIO CORRELAZIONI

1. Presentazione dati bruti in tabella, **comprese incertezze**.
2. Rappresentazione grafica dei dati bruti, con barre di incertezza **e senza alcun fit né altri parametri**.
 - Serve solo a mostrare che ci può essere una correlazione matematica.
3. Coefficiente di correlazione r (*per relazioni lineari o linearizzate*).
 - Se e solo se r indica una correlazione significativa si prosegue con un fit.
4. Best-fit(s) con il metodo dei minimi quadrati:
 - I. Scelta della variabile indipendente basata sul confronto degli errori relativi delle x e delle y → **eventuale inversione della formula inizialmente adottata, per scambiare x con y** .
 - II. Scelta tra best-fit pesato e non.
 - III. Discussione dell' eventuale termine noto in relazione a errori sistematici: il fit va fatto **inizialmente comunque con tutti i parametri liberi (se lineare, con entrambi i parametri della retta liberi)** e solo dopo avere verificato che il parametro in questione, entro gli errori trovati dal fit, è compatibile con il valore previsto (tipicamente, ma non solamente, 0) **rifare il fit vincolato**, o altrimenti **discutere errori sistematici e se individuati correggere i dati e rifare il fit**.
5. Calcolo del χ^2 .
 - I. Discussione dello stesso e **eventuale riduzione ad arte degli errori assoluti** per arrivare a stimare quelli casuali.
6. *Eventuale* determinazione dell' incertezza sulla misura *indiretta* di una grandezza terza da un parametro del fit.

La mancanza o incompletezza di ciascuno di questi passi comporta una penalizzazione anche importante sul voto attribuito alla relazione

Lezione n.25

PARAGRAFO FINALE RELAZIONI: CONCLUSIONI

CONCLUSIONI

1. No commenti del tipo: «potevo fare di meglio se...»
2. Confronti quantitativi tra valori ottenuti con metodi di misura e/o tecniche di analisi diversi e, se del caso, tra valore/i misurato/i e atteso/i, in grafici simili nell' aspetto a quelli seguenti riportati a titolo di esempio:

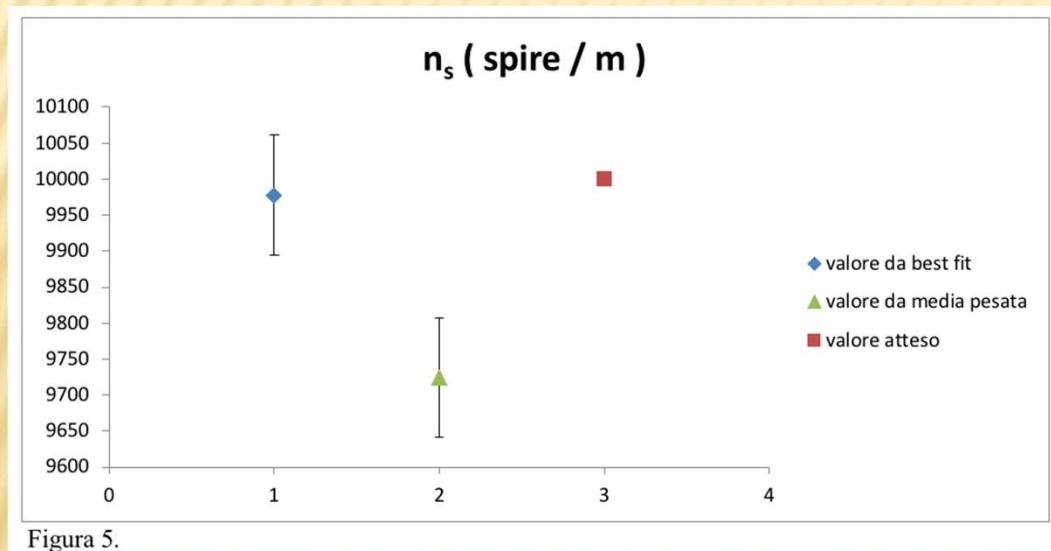
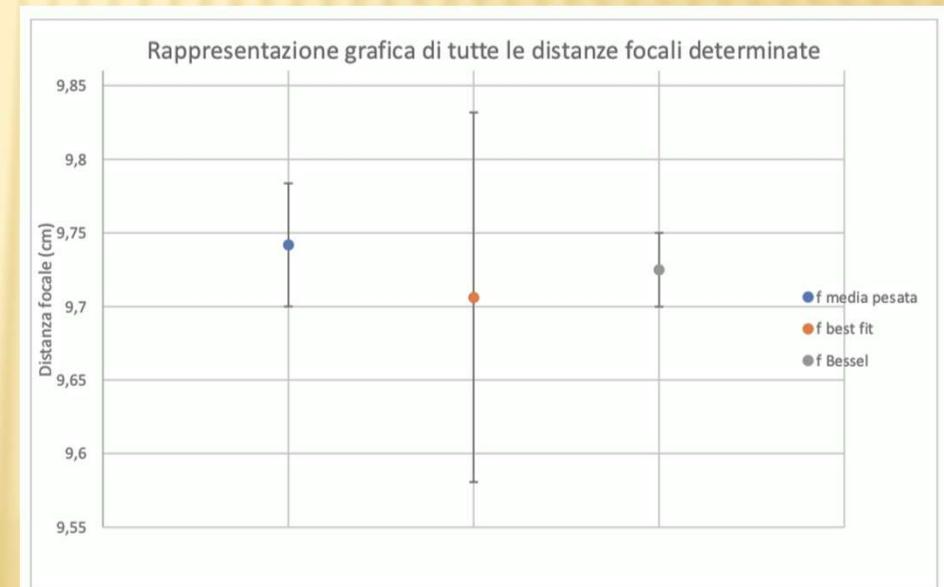
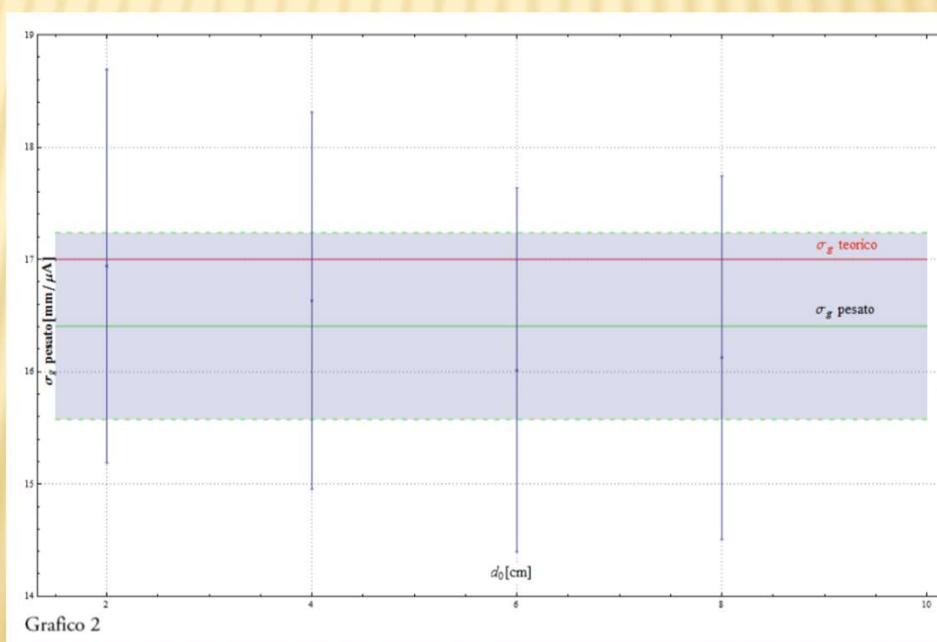
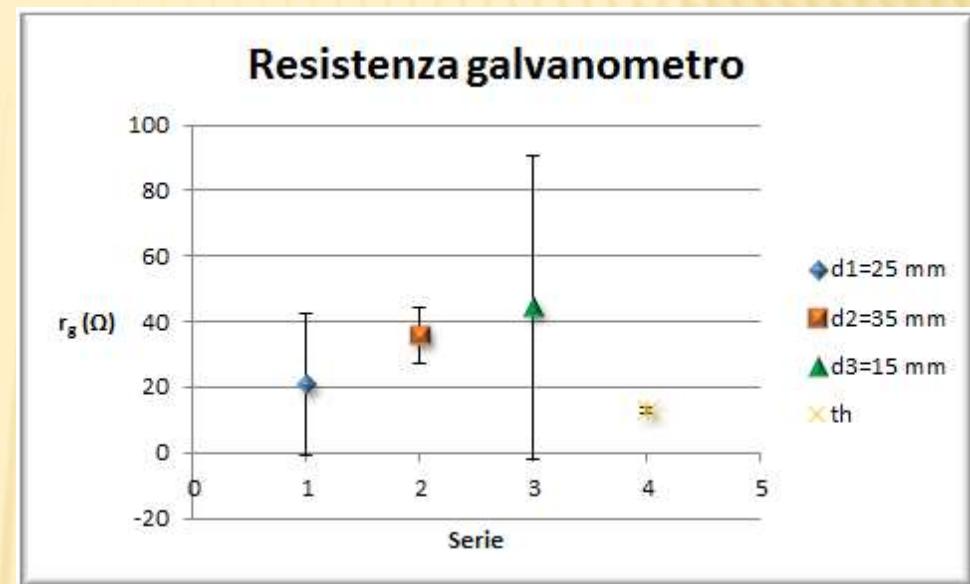
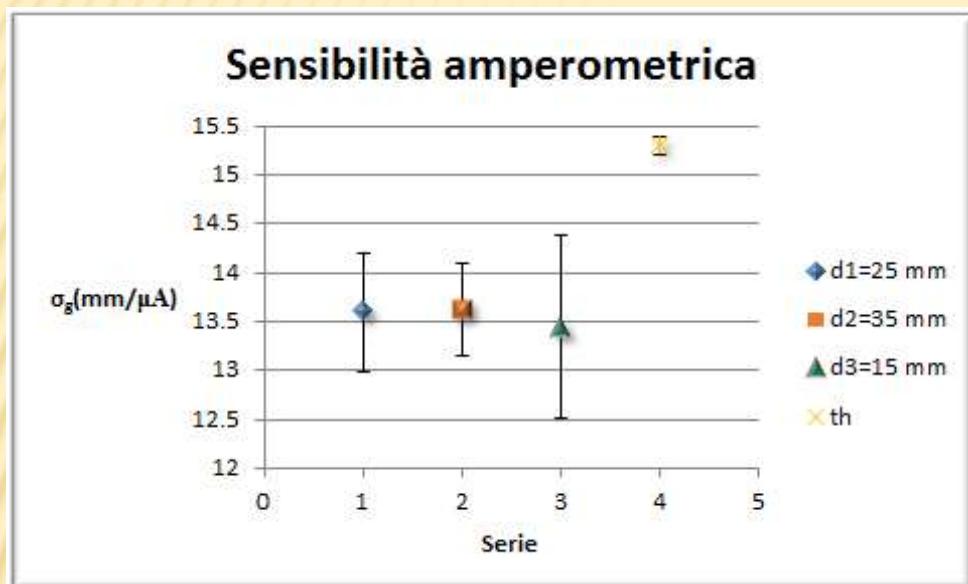


Figura 5.



ALTRI ESEMPI...



Prof. Salvatore Costa: Lezioni di

LABORATORIO DI FISICA 2

FINE