

Roma1 N.1094

Maggio 1998 (Versione corretta, Gennaio 1999)

Errori e incertezze di misura
- rassegna critica e proposte per l'insegnamento -

Giulio D'Agostini

Dipartimento di Fisica dell'Università di Roma "La Sapienza"

E_mail: dagostini@roma1.infn.it

URL: <http://www-zeus.roma1.infn.it/~agostini/>

Indice

1	Introduzione	5
2	Inevitabilità delle incertezze	7
3	Raccomandazioni ISO/BIPM	8
4	Cause delle incertezze di misura	10
5	Valutazioni usuali delle incertezze	13
6	Critica della “teoria degli errori massimi”	14
6.1	$\Delta y = \sum_i \left \frac{\partial y}{\partial x_i} \right \Delta x_i$	14
6.2	Regola della mezza divisione	18
6.3	$\Delta t = 0.2 s?$	21
6.4	Imperativo categorico di riportare le “barre di errore”	21
6.5	Rette di massima e minima pendenza	21
7	Critica degli “errori statistici”	24
8	Riassumendo	25
9	Da dove ricominciare?	26
10	Probabilità soggettiva	28
11	Formalismo sulle variabili casuali: “$f(x)$”, $E(X)$ e σ	29
12	Valutazione dell’incertezza di misura: schema generale	30
13	Imparare dagli esperimenti: il problema dell’induzione	32
14	Dalla probabilità degli effetti alla probabilità delle cause	33
14.1	Verosimiglianza	33
14.2	Probabilità iniziale e probabilità finale	34
14.3	Paura dei “pregiudizi”? Inevitabilità di principio e frequente irrelevanza pratica delle prior	34
14.4	Scorciatoia al ragionamento bayesiano: il cane e il cacciatore	35
15	Imparare dall’esperienza	36
16	Risultati delle misure dirette in assenza di errori sistematici	37
16.1	Condizioni di ripetitività	37
16.2	Singola osservazione con σ_r nota	39
16.3	n osservazioni indipendenti con σ_r nota	39
16.4	Caso di σ_r ignota	41
16.4.1	Misure ripetute della stessa grandezza fisica	41
16.4.2	Singole misure di grandezze fisiche variabili (grafici)	42

16.4.3	Bisogna sempre ripetere le misure? Rarità delle situazioni in cui σ_r sia completamente ignota	43
17	Propagazione delle incertezze	43
18	Come tener conto degli errori sistematici	46
18.1	Condizioni di riproducibilità	46
18.2	Correzione dei risultati per tener conto di errori sistematici noti - calibrazioni	47
18.3	Incetnze dovute all'inesatta conoscenza dell'entità di un possibile errore sistematico	47
18.4	Imperfetta conoscenza delle costanti di calibrazioni e dei parametri di influenza	47
18.4.1	Errore di zero (offset)	47
18.4.2	Errore di scala	49
18.4.3	Importanza delle misure per differenza	50
18.5	Casi di errore di più difficile schematizzazione	51
18.6	Incetzza su un fattore di influenza	52
18.7	Propagazione senza derivate	53
18.8	Calibrazione, intercalibrazione e "randomizzazione"	53
19	Coefficiente di correlazione	54
19.1	Valutazione pratica di ρ dovuto ad errori di calibrazione	54
19.2	Propagazione di varianze e covarianze	55
20	Raccomandazioni BIPM/ISO	57
21	Valutazione delle incetnze di tipo B	59
22	Esempi numerici	61
23	Fit di andamenti lineari	64
23.1	Grafici e parametri della retta	65
23.2	Inferenza bayesiana e metodo dei minimi quadrati	67
23.3	Analisi grafica	67
23.4	Effetto degli errori sistematici	70
23.5	Esempio numerico	72
23.6	Altri tipi di fit	75
23.7	"Riabilitazione del computer"	75
23.8	Rette di taratura?	76
24	Note sulla didattica	77
25	Conclusioni	79
26	Nota bibliografica	80
A	Capacità di interpolazione fra le tacche e incetzza di lettura	82
A.1	Errore ed incetzza di misura (commento all'esperienza di interpolazione fra le tacche)	82

B	Tempo di reazione e misure di cronometraggio	84
C	Distribuzioni triangolari	85
D	Teorema del limite centrale	87
E	Deviazione standard della media aritmetica	87
	E.1 Stime di segmenti alla lavagna	87
	E.2 Media di valori distribuiti uniformemente	88
F	Soluzione dell'esercizio di interpolazione fra le tacche	88

1 Introduzione

Scrivendo questa nota ho cercato di venire incontro a tutti coloro che trovano motivi di insoddisfazione nella trattazione dei cosiddetti “errori di misura”: *studenti* poco convinti di quanto viene loro insegnato, in quanto spesso privo di consistenza logica e di corrispondenza con l’esperienza di laboratorio; *laureandi e neolaureati* che, inseriti nella ricerca, sia pura che applicata, sperimentano l’inadeguatezza dei metodi appresi per far fronte alle analisi complesse che si presentano e sono confusi dalle tante “ricette monouso”, spesso contraddittorie, che trovano in libri, note tecniche e articoli; *ricercatori* che, delusi dall’incongruenza fra teoria e pratica, affermano francamente di non usare la statistica o, addirittura, di non essere interessati al “calcolo degli errori”; *docenti universitari* che si rendono conto, specialmente se interagiscono con ambienti di ricerca internazionali, che le cose stanno cambiando ed in effetti provano a riaggiornare i corsi, pur ostacolati dalla mancanza di testi validi e da colleghi con i quali devono interagire per uniformità di programma; *insegnanti* delle scuole medie superiori in dubbio su come comportarsi nei corsi di laboratorio e indecisi se affrontare o meno il discorso delle incertezze di misura, con quali metodi e a quale livello.

Questo lavoro è basato sull’esperienza acquisita nella ricerca (nel campo della Fisica Subnucleare) e nell’insegnamento, arricchita da interazioni, dirette o indirette, con metrologi e probabilisti, da discussioni con colleghi di diverse nazionalità e dal feedback ricevuto in seminari e corsi di perfezionamento impartiti in Italia e all’estero. Negli ultimi anni ho anche beneficiato di contatti con gli insegnanti dell’AIF (Associazione per l’Insegnamento della Fisica) di Roma e con gli studenti del Corso in Perfezionamento in Didattica della Fisica.

La trattazione non pretende di avere nessun carattere di completezza. Ha però il vantaggio di offrire una vista d’insieme che permette al lettore la possibilità di confrontare i metodi che vengono criticati con quanto viene proposto. E’ pertanto opportuno fare alcuni chiarimenti preliminari sul taglio del lavoro e sulle conoscenze che si presuppongono da parte del lettore.

- L’impostazione della parte di rassegna critica è deliberatamente provocatoria, con la speranza di suscitare le reazioni di chi non condivide questo punto di vista e di innescare un dibattito costruttivo su eventuali temi controversi dal quale beneficiare tutti, studenti in primo luogo.
- Una trattazione delle incertezze di misura che fornisca soltanto formule pratiche, anche se ritenute ragionevoli dai più e supportate dalle raccomandazioni delle massime organizzazioni di metrologia, costituisce una costruzione vacillante se non si presenta in modo coerente, con delle basi sulle quali eventualmente convenire e dalle quali derivare le conseguenze logiche.
- Le basi di partenza di questa impostazione sono semplicemente:
 1. la constatazione dell’inevitabile stato di *conoscenza incertezza* cui l’induzione dà luogo (ricollegabile alla famosa “critica di Hume” di tale processo);

2. l'accettazione della validità del concetto di *probabilità* per classificare la plausibilità delle affermazioni quando si è in stato di incertezza (riconducibile alla visione originale di probabilità la quale è stata ripresa e consolidata dopo i primi decenni di questo secolo ed è attualmente nota come "*soggettiva*" o "*bayesiana*").

Una volta accettati questi presupposti, tutto il resto viene derivato come conseguenza logica. Si tratta soltanto di formalismi, approssimazioni ed eventuali scorciatoie per semplificare la trattazione dei normali problemi di routine.

- Si presuppone che il lettore sia già stato esposto alla problematica dell'incertezza di misura e che abbia una conoscenza elementare del calcolo delle probabilità. In particolare, si assume che egli abbia le nozioni di base per trattare le variabili casuali (" $f(x)$ ", " $E(X)$ ", " σ ", distribuzione uniforme e normale). Alcuni di questi concetti verranno reinterpretati alla luce della probabilità soggettiva, ma i risultati formali del calcolo delle probabilità sono gli stessi dell'approccio convenzionale.
- Anche se il concetto di funzione densità di probabilità è importante per capire il ragionamento che si svilupperanno, la parte applicativa farà uso solo di deviazioni standard, in quanto l'ipotesi di normalità (distribuzione gaussiana) è spesso soddisfacentemente soddisfatta nei casi di routine. Anche il teorema di Bayes, sul quale sarà basato il processo di aggiornamento della conoscenza, sarà scavalcato con argomentazioni intuitive.
- Venendo alle applicazioni della cosiddetta inferenza bayesiana all'incertezza di misura, saranno mostrate procedure alternative a quelle criticate nella parte iniziale. Molti riconosceranno metodi che già conoscono e usano, spesso visti però come una tecnica *ad hoc* per risolvere un problema particolare.
- Infine verranno dati dei suggerimenti per la didattica.
- Il lavoro si conclude con un'appendice in cui vengono riportati alcuni argomenti accessori e illustrate alcune semplici esperienze.

Ovviamente sarò felicissimo di avere scambi di idee con tutti coloro che avranno critiche, commenti o proposte sul testo.

Giulio D'Agostini

Roma, Maggio 1998

Questa versione corretta ha beneficiato dei commenti di Maria Grazia Ianniello e dell'attentissima lettura di Marco Schioppa.

Roma, Gennaio 1999

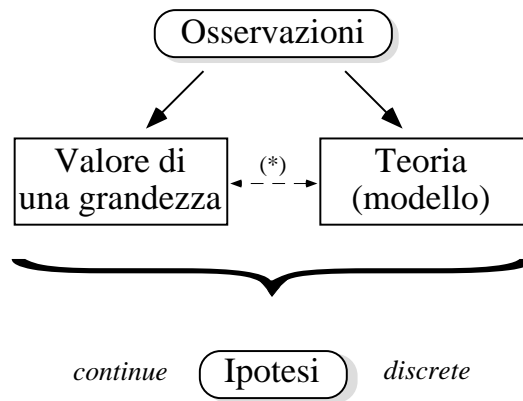


Figura 1: Dalle osservazioni alle ipotesi. La relazione fra valore della grandezza e teoria sta ad indicare che in genere le grandezze hanno significato soltanto all'interno di una teoria o un modello.

2 Inevitabilità delle incertezze

Cerchiamo di capire quali sono le ragioni di fondo che, nella ricerca scientifica, conducono ad uno stato di incertezza. La figura 1 schematizza l'attività del fisico o di qualsiasi altro ricercatore. Dai dati sperimentali si cerca di determinare il valore di una certa grandezza o di stabilire quale teoria descrive meglio i fenomeni osservati. In realtà entrambi i processi possono essere visti come due aspetti dello stesso problema: come passare dalle osservazioni alle ipotesi. Infatti i due problemi possono essere riformulati nei seguenti modi:

- A** quali valori sono (più) compatibili con la definizione della grandezza oggetto della misura, avendo letto certi numeri sugli strumenti (e subordinatamente a tutte le conoscenze sugli strumenti e sulla grandezza in questione)?
- B** quale teoria è (più) compatibile con i fenomeni osservati (e subordinatamente alla credibilità della teoria basata su argomenti formali, estetici e di semplicità¹)?

La sola differenza fra i due processi di apprendimento è che, mentre nel secondo caso si ha a che fare generalmente con un piccolo numero di ipotesi, nel primo caso il numero di ipotesi è virtualmente infinito (le grandezze assumono i valori numerici con continuità, almeno in linea di principio).

Il motivo per cui non si arriva mai alle condizioni ideali di certezza, ovvero tali che soltanto una delle tante (o infinite) ipotesi sia da ritenersi *vera* e tutte le altre *false*, può essere compreso analizzando lo schema che segue.

- A:** Per quanto riguarda la determinazione del valore di una grandezza si dice comunemente che l'incertezza sia dovuta ad inevitabili “*errori di misura*” (fluttuazioni della risposta, calibrazione degli strumenti, condizioni ambientali non perfettamente note, etc.);

¹Si vedrà come queste postille di subordinazione delle conclusioni scientifiche a conoscenze e “pregiudizi” a priori giocano un ruolo fondamentale nei processi di misura e nell'accettazione di teorie da parte della comunità scientifica.

B: Quando si tratta di una teoria possiamo distinguere due casi:

- (B₁) La legge è probabilistica, ovvero *“le osservazioni non sono una mera conseguenza logica della teoria”*. Un classico esempio è quello della genetica. Un esempio più semplice è quello del lancio di una moneta. E’ noto dal calcolo delle probabilità che, lanciando una moneta regolare, le due sequenze di testa (T) e croce (C)

TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT (1)

TTCCCCCTTTTCTTCCTTTCTCTTC (2)

hanno la stessa probabilità. Quindi sarà impossibile arrivare a conclusioni certe sulla regolarità di una moneta ignota pur avendo osservato una sequenza di lunghezza arbitraria²;

- (B₂) La legge è deterministica. Non è difficile convincersi come questa denominazione sia valida solo in linea di principio. Infatti, in tutti i casi, *“le osservazioni dipendono anche da molti altri fattori esterni alla teoria”*, siano essi condizioni iniziali e ambientali, errori sperimentali, e così via. Ne segue quindi che le inevitabili incertezze su questi fattori rendono la relazione teoria→osservazione di tipo probabilistico anche in questo caso.

3 Raccomandazioni ISO/BIPM

Il fatto che a qualcuno la frase “le incertezze sono dovute ad errori di misura” possa suonare come una tautologia è un indizio della forte disomogeneità di linguaggio e di metodologia riscontrabile nel campo degli errori e delle incertezze di misura. Questo è in effetti il caso.

Indicativo dello stato di confusione su questo argomento è il recente sforzo - tuttora in corso - delle massime organizzazioni di metrologia per suggerire dei criteri generali di comportamento. In particolare, nel 1993 l’Organizzazione Internazionale per la Standardizzazione (ISO) ha pubblicato una *“Guida all’espressione dell’incertezza di misura”*. Essa è basata sulle raccomandazioni dell’Ufficio Internazionale di Pesi e Misure (BIPM) e fornisce una descrizione delle procedure suggerite, con esempi pratici. La *Guida* accenna anche la base teorica sulla quale essa si fonda, ma in forma molto breve, dato il carattere succinto di collezione di norme, tipico delle pubblicazioni ISO.

La *Guida* riassume comunque, in un centinaio di pagine, 15 anni di lavoro delle più autorevoli organizzazioni mondiali di metrologia:

²Ma dopo l’osservazione della prima delle sequenze è forte il sospetto che si tratti di una moneta con due teste, qualora ci siano delle buone ragioni per far sorgere un simile dubbio (ad esempio non si è modo di verificare direttamente la regolarità della moneta). Il concetto di probabilità servirà a quantificare il grado di tale sospetto.

BIPM	Bureau International des Poids et Mesures
IEC	International Electrotechnical Commission
IFCC	International Federation of Clinical Chemistry
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
IUPAP	International Union of Pure and Applied Physics
OIML	International Organization of Legal Metrology

Oltre alle organizzazioni internazionali che sponsorizzano l'iniziativa, aderiscono molti istituti nazionali ad esse affiliati, come ad esempio: la tedesca DIN (Deutsches Institut für Normung); l'americano NIST (National Institute of Standards and Technology); l'italiano UNI (Ente Italiano per l'Unificazione).

Se andiamo a consultare la *Guida ISO* riguardo i termini *incertezza* ed *errore* troviamo:

- Incetezza: “un parametro che caratterizza la dispersione dei valori che possono essere ragionevolmente attribuiti al misurando”;
- errore: “differenza fra il risultato di una misura e un valore vero del misurando”.

Si noti che:

1. la definizione ISO di incertezza chiarisce il concetto; per quanto riguarda la definizione operativa si fa uso della “*incertezza standard*”, che indica la deviazione standard (σ) dei possibili valori che può assumere il misurando (ciascuno di essi pesato con il suo “grado di fiducia”, nel senso che sarà chiarito nel paragrafo 10);
2. per quanto riguarda l'errore, è evidente come esso sia generalmente ignoto, in quanto relativo “al” valore vero della grandezza di interesse (“al” è fra virgolette a causa delle implicazioni del prossimo punto);
3. l'uso dell'articolo indeterminativo davanti a “valore vero” è intenzionale e ha delle profonde ragioni “filosofiche” che saranno brevemente illustrate nel seguito.

La definizione ISO di valore vero è infatti:

- valore vero: “un valore compatibile con la definizione di una data grandezza particolare”³

Questa definizione potrà sembrare vaga, ma, a pensarci bene, è quella più pragmatica (si veda anche il punto 1 della lista delle cause delle incertezze di misura del prossimo paragrafo).

³ “Grandezza particolare” (altezza di una certa torre, massa dell'elettrone, accelerazione di gravità a Roma) è in contrapposizione con “grandezza generale” (lunghezza, massa, accelerazione).

4 Cause delle incertezze di misura

Vale la pena di riportare la lista delle possibili cause di incertezza nelle misure, così come compare nella *Guida*, con l'aggiunta di alcuni esempi esplicativi. Essa rappresenta una specie di “decalogo” da scorrere ogni volta che si cercano le possibili sorgenti di errore.

1. *Incompleta definizione del misurando.*

Ad esempio la “percentuale di potassio nell'acqua del Mar Adriatico” non definisce completamente il misurando e il risultato può dipendere da dove è prelevato il campione. Lo stesso vale per l'“accelerazione di gravità al livello del mare”, in quanto essa dipende anche dalla latitudine e, potendo eseguire la misura con precisione infinita, dal punto esatto (quale è il livello del mare?). Si capisce quindi come ogni indeterminazione sulla definizione si riflette su una infinità di valori che soddisfano la definizione (vedi definizione del valore vero⁴). Nel paragrafo 22 sarà trattato numericamente il caso di “densità dell'aria”.

2. *Imperfetta realizzazione della definizione del misurando.*

“Vita media di decadimento dell'isotopo X ” e “sezione efficace di un neutrone bombardato da un fascio di elettroni” sono definizioni univoche, almeno in linea di principio. In pratica non è facile ottenere un campione assolutamente puro di sostanza o effettuare delle misure nelle condizioni ideali della definizione.

Per fare altri esempi, si pensi a: “accelerazione di un corpo lungo un piano inclinato privo di attrito” e “periodo di un pendolo semplice di lunghezza l ”. Qui si fa chiaramente riferimento ad astrazioni di cui gli apparati sperimentali sono imperfette realizzazioni.

3. *Campione non rappresentativo, ovvero il campione misurato non rappresenta il misurando definito*

Caso classico sono i sondaggi per stimare - misurare - la “percentuale della popolazione in possesso di un certo carattere” (in senso lato). Non avendo a disposizione le risorse economiche e il tempo per eseguire un'indagine adeguata, o in mancanza di un modello teorico per la scelta del campione si rischia di effettuare un sondaggio su coloro che sono caratterizzati da un'altra proprietà comune dalla quale può dipendere il carattere oggetto della ricerca (ad esempio si può rischiare di intervistare solo amici o concittadini, o solo coloro che vedono la televisione ad una certa ora o che trascorrono il pomeriggio a casa). Altro esempio è quello dell'analisi chimico-fisica di un quadro che comporta la distruzione di un campione di tela. Le informazioni che si ricavano da un lembo periferico - più facilmente ottenibile per l'analisi - possono differire da quelle ottenibili da parti artisticamente più interessanti del quadro (convincereste il Louvre a cedervi un occhio della Gioconda?).

⁴Si capisce quindi come la definizione di valore vero come “quello che si otterrebbe dopo una serie infinita di misure con strumentazione ideale” non è migliore di quella ISO, anzi, questa dà l'illusione che questo valore sia, almeno idealmente, unico, mentre la definizione ISO tiene conto che le misure vengono eseguite in condizioni reali e con tutte le cause di incertezza che saranno elencate in questo paragrafo.

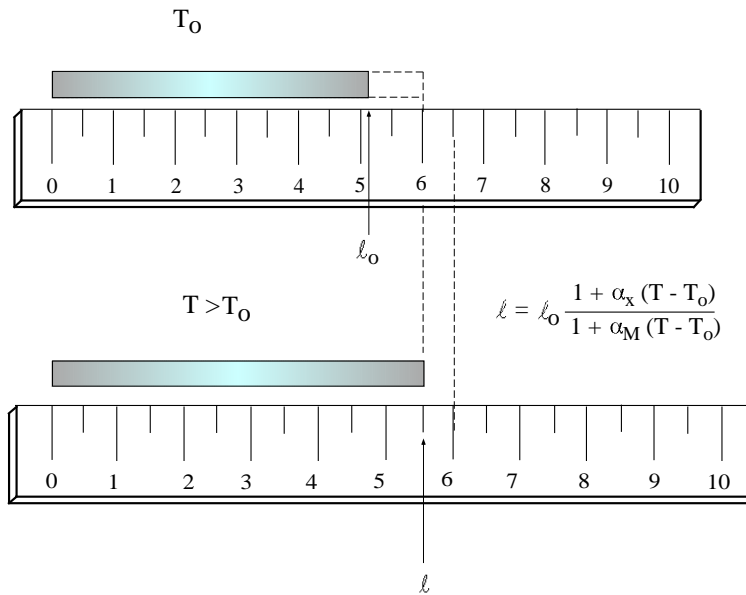


Figura 2: Esempio di errore sistematico dovuto all'utilizzo di uno strumento a temperatura diversa da quella nominale.

4. *Imperfetta conoscenza delle condizioni ambientali di influenza o inadeguata conoscenza degli effetti di tali condizioni*

Ad esempio una misura di precisione può essere falsata dalla non esatta conoscenza della temperatura ambientale. La figura 2 mostra, ad esempio, l'errore introdotto nella misura, se questa viene eseguita ad una temperatura T diversa da quella di riferimento T_0 . Conoscendo i coefficienti di dilatazione termica è possibile correggere il valore ottenuto (come mostrato in figura), ma ogni eventuale incertezza sul valore di T (e, in misura minore, su α) si "propaga" nell'incertezza sulla lunghezza l .

5. *Errore di lettura di uno strumento.*

La lettura delle scale analogiche dipende dall'acuità visiva e dall'abilità di stima dello sperimentatore. E' da notare inoltre che la qualità delle interpolazioni della lettura dipendono molto dalle condizioni di lavoro (illuminazione e facilità di lettura) e dal fatto che la grandezza fisica sia statica o rapidamente variabile con il tempo. Non è inoltre da trascurare la dipendenza dall'importanza che lo sperimentatore dà a priori alla qualità della determinazione di tale grandezza fisica. Infatti è inutile sforzarsi a leggere i decimi di millimetro per dare al falegname le dimensioni di un tavolo.

Da queste considerazioni ne segue che l'incertezza da associare all'errore di misura non è univocamente determinata dal tipo di strumento.

6. *Risoluzione finita o soglia di discriminazione dello strumento.*

Ad esempio, se la lettura avviene con uno strumento digitale si è limitati alla cifra meno significativa del display anche se la qualità del *segnale di*

misura è tale da essere significativamente sensibile a variazioni di valori ben minori dell'entità dell'ultima cifra del display.

Come esempio numerico prendiamo un dispositivo elettronico che produce una variazione nella tensione di uscita di 17 mV se la pressione cambia di 1 mbar. Assumiamo che a 1018 mbar il *segnale di misura* sia pari a 3.714 V e che sia convertito direttamente in mbar e mostrato su un display digitale a 4 cifre e che indica il valore direttamente in mbar. Una variazione di ± 0.2 mbar non produce nessuna variazione dell'indicazione ("1018") pur causando una variazione di ± 3.4 mV sul segnale di misura.

7. *Valori inesatti dei campioni e dei materiali di riferimento.*

Si pensi ad esempio ad una massa campione con graffi, polvere e ossidazioni, oppure ad una soluzione campione di pH che si è contaminata dal momento della sua preparazione.

I campioni servono a calibrare (o ricalibrare) gli strumenti. Ogni incertezza sul valore del campione si riflette sulla *costante di calibrazione* e quindi su tutte le misure che saranno eseguite con tale strumento. Le incertezze su queste misure saranno quindi correlate.

8. *Valore inesatto di costanti e altri parametri che intervengono nell'analisi dei dati.*

Spesso le misure indirette dipendono da costanti e parametri misurati dallo stesso sperimentatore, da suoi colleghi o semplicemente riportate su articoli o libri. Ogni incertezza su queste grandezze si *propaga* su quelle misurate.

9. *Approssimazioni e assunzioni che intervengono nel metodo e nella procedura di misura.*

Ad esempio, nel modello teorico elementare che descrive l'oscillazione del pendolo sono usualmente trascurati gli effetti che derivano dal fatto che l'angolo di oscillazione è diverso da zero. Se si conoscono i termini correttivi si può ottenere, in linea di principio, il valore $T(\alpha = 0)$ dal valore misurato $T(\alpha = \alpha_0)$. Se invece si è *coscienti* di effetti che non si è in grado di calcolare, oppure se si può effettuare soltanto una stima grossolana degli stessi, allora si dovrà introdurre un ulteriore contributo all'incertezza.

10. *Variazioni in osservazioni ripetute del misurando sotto condizioni di misura apparentemente identiche.*

Queste variazioni sono legate ai così detti *errori casuali*.

Si noti che:

- non è facile effettuare una suddivisione netta fra le diverse sorgenti di incertezza e, in particolare, tutte quelle dei punti 1-9 hanno una qualche influenza sul punto 10 in quanto difficilmente esse produrranno esattamente un identico effetto durante una serie di misure;
- nella maggior parte delle misure lo sperimentatore fa parte integrante del processo di misura. Il risultato dipende molto da abilità ed esperienza di chi effettua le misure, sia per quanto riguarda la manualità

nell'operare gli strumenti nella lettura di strumenti analogici, che per la capacità di vagliare i vari contributi all'errore di misura. Quindi è difficile separare l'incertezza "intrinseca" dovuta allo strumento da quella di altri contributi.

5 Valutazioni usuali delle incertezze

Vediamo ora quali sono le tecniche di valutazione delle incertezze ("errori") usate dalla maggior parte dei laureati in materie scientifiche e tuttora insegnate nelle scuole secondarie superiore e anche all'Università (si noti come alcuni metodi variano leggermente a seconda dell'Università di provenienza; nel seguito si fa riferimento a quello che, grosso modo, rappresenta lo standard romano).

Cominciamo esaminando attentamente la seguente lista di nozioni tipiche. Esse saranno analizzate in dettaglio nei prossimi paragrafi. Questa lista era stata originariamente elaborata per l'introduzione ad una serie di seminari tenuti alla fine del 1996 ad insegnanti dell'AIF e serviva a valutare il background comune.

1. "Propagazione degli errori massimi":

$$\Delta y \approx \left| \frac{\partial y}{\partial x_1} \right| \Delta x_1 + \left| \frac{\partial y}{\partial x_2} \right| \Delta x_2 + \dots \quad (3)$$

ad esempio

$$y = x_1 \pm x_2 \Rightarrow \Delta y = \Delta x_1 + \Delta x_2$$

2. "Propagazione degli errori statistici":

$$\sigma^2(y) = \left(\frac{\partial y}{\partial x_1} \right)^2 \sigma^2(x_1) + \left(\frac{\partial y}{\partial x_2} \right)^2 \sigma^2(x_2) + \dots \quad (4)$$

ad esempio

$$y = x_1 \pm x_2 \Rightarrow \sigma(y) = \sqrt{\sigma^2(x_1) + \sigma^2(x_2)}$$

3. Regola della "mezza divisione":

$$\Delta x = \frac{1}{2} \text{ divisione} \quad (5)$$

4. I punti sperimentali vanno riportati sui **grafici sempre con le "barre di errore"**.

5. **Rette di massima e di minima pendenza.**

6. Avendo eseguito un numero n “abbastanza grande” di misure, il **risultato** va riportato come

$$\mu \stackrel{?}{=} \bar{x} \pm \sigma \quad (6)$$

oppure

$$\mu \stackrel{?}{=} \bar{x} \pm \frac{\sigma}{\sqrt{n}} \quad (7)$$

Inoltre, molto importante:

che significato si attribuisce a queste due espressioni?

Il risultato del sondaggio (confermato sostanzialmente in altri seminari e corsi di perfezionamento) è stato che:

- i partecipanti conoscevano molto bene questi concetti, con eccezione del punto 2;
- per quanto riguarda il punto 3, alcuni mostravano addirittura una preferenza ad una stima dell’incertenza più conservativa ($\Delta x = 1$ divisione);
- sulla scelta fra $\pm\sigma$ e $\pm\sigma/\sqrt{n}$, c’era una netta preferenza ad utilizzare come incertezza il valore della deviazione standard delle singole misure anziché dividerlo per \sqrt{n} (“altrimenti diventa troppo piccolo”). Comunque, a parte il numero da mettere nell’espressione, c’era l’unanime consenso che l’espressione stesse a significare “una certa probabilità che μ sia compreso nell’intervallo”.

6 Critica della “teoria degli errori massimi”

Passiamo ora in rassegna i concetti e le procedure che abbiamo illustrato, cercando di capire su cosa sono fondate e cosa implicano.

$$6.1 \quad \Delta y = \sum_i \left| \frac{\partial y}{\partial x_i} \right| \Delta x_i$$

Questa espressione starebbe a significare che

se siamo “praticamente certi” che il valore vero x_{v_i} è compreso nell’intervallo dato da $x_i \pm \Delta x_i$, ne segue che siamo “praticamente certi” che il valore vero di y_v è compreso nell’intervallo dato da $y \pm \Delta y$.

E’ opinione comune che, affinché la formula sia valida, debba valere $\Delta x_i \ll x_i$ (giustificazione usuale). Se accettiamo per buona tale espressione di “propagazione lineare degli errori massimi” e i presupposti sui quali essa si basa andiamo incontro ad incongruenze, come mostrano gli esempi che seguono.

1. Se $x_1 = 0.0 \pm 0.5$ e $x_2 = 0.5 \pm 0.5$ quanto vale $\Delta(x_2 - x_1)$? (La seconda condizione non è più valida.)

2. Misuriamo due spessorini, uno di 1 mm e l'altro di 2 mm (valori "esatti"), con un righello aventi divisioni di 1 mm. Otteniamo $x_1 = 1.0 \pm 0.5$ mm e $x_2 = 2.0 \pm 0.5$ mm, da cui $x_2 - x_1 = 1 \pm 1$ mm. Come si recita in questi casi, le due misure sono "uguali entro gli errori". Ciò nonostante, una qualsiasi ispezione visuale suggerisce che uno spessore è circa il doppio dell'altro. Nessuno potrà giurare che il rapporto fra i due sia esattamente 2: potrebbe essere 1.8, 1.9, 2.0, 2.1, 2.2, o forse 1.7 o 2.3, ma sicuramente sono esclusi i valori prossimi a 1. Si ottiene quindi un risultato formale in netta contraddizione con quanto si crede: una conclusione paradossale!
3. Consideriamo un termometro a mercurio, avente divisioni di 0.1°C e di cui sappiamo che potrebbe essere scalibrato al più di 0.6°C . Consideriamo le seguenti letture, lasciando sospese le incertezze e le successive elaborazioni:

$$T_1 = 22.00 \dots \pm \dots ^\circ\text{C} \quad (8)$$

$$T_2 = 23.00 \dots \pm \dots ^\circ\text{C} \quad (9)$$

$$T_2 - T_1 = \dots \pm \dots ^\circ\text{C} \quad (10)$$

La risposta usuale a questo quesito è che ΔT_1 e ΔT_2 sono pari a 0.6°C , mentre

$$T_2 - T_1 = 1.0 \pm 1.2 ^\circ\text{C}.$$

(Qualcuno, sospettando un tranello, azzarda un $\Delta(T_2 - T_1) = 0.6^\circ\text{C}$.) Non è difficile convincersi che, mentre incertezze di 0.6°C su ciascuna misura sono ragionevoli, se intese come "errori massimi", quella sulla differenza non è affatto sensata. La calibrazione assoluta non può avere alcun effetto sulla differenza fra valori di temperatura così prossimi. Alla luce delle considerazioni del punto precedente, possiamo affermare che la stima più ragionevole dell'incertezza su $T_2 - T_1$ sia inferiore a 0.1°C (per arrivare ad valore numerico bisognerà premettere altre considerazioni e saperne di più sul termometro, sulle condizioni di misura e su chi ha eseguito le letture).

4. Torniamo ora all'espressione "praticamente sicuri":

- cosa significa?
- cosa si paga se non è vero (se dovesse risultare che il valore vero è al di fuori dell'intervallo indicato, o almeno "molto al di fuori", visto che non si trattava di certezza assoluta)?
- è quello che serve veramente?

Analizziamo quest'ultimo punto. Prendiamo, come esempio, la somma di tante grandezze di uguale valore e incertezza (tanto per semplificare i conti):

$$\Delta x_i = \underline{\text{costante}} = \Delta x$$

$$x_i = \underline{\text{costante}} = x$$

La somma degli n valori e la sua incertezza, calcolata usando la (3), sono

$$y_n = \sum_{i=1}^n x_i = nx$$

$$\Delta y_n = \sum_{i=1}^n \Delta x_i = n \Delta x$$

Confrontiamo questo risultato con quanto si ottiene mediante un piccolo programma di simulazione⁵, assumendo che il valore vero delle x_i potrebbe essere in qualsiasi punto entro l'intervallo $x_i \pm \Delta x_i$. La figura 3 mostra i risultati di 10000 simulazioni, per $n=1, 2, 3, 5, 10, 20$ e 50. Per comodità l'asse delle ascisse è preso fra $y_n - \Delta y_n$ e $y_n + \Delta y_n$ dati dalla formula precedente. Come si vede dalla figura, è senz'altro corretto affermare di essere “praticamente certi” che il risultato sia in quell'intervallo, ma, al crescere di n , la prudenza è tale che il risultato si è “impoverito” rispetto alle sue potenzialità originarie.

Si potrebbe obiettare che in pratica si fanno solo poche misure. Questo può essere vero in una semplice esperienza di laboratorio, ma nel mondo reale la propagazione delle incertezze è in principio illimitata: ognuno utilizza informazioni precedentemente ricavate da lui o da altri, e le conclusioni verranno utilizzate da altri ancora, etc. (nessuno fa una misura per incorniciare il risultato a casa, senza nessuna influenza per altri⁶ ...).

Riassumendo, possiamo affermare che l'uso della cosiddetta “teoria” degli *errori massimi* conduce a

- una tendenza a sovrastimare le incertezze;
- all'impossibilità di trattare propriamente gli effetti delle correlazioni.

Ora, qualcuno potrebbe pensare che l'effetto delle correlazioni possa essere una finezza e che la sovrastima delle incertezze sia da ritenere addirittura essere un pregio. Se gli esempi precedenti, che hanno mostrato come facilmente si arriva a sovrastime di un ordine di grandezza non dovesse bastare, citiamo la *Guida ISO* in proposito:

“The method [*quello raccomandato dalla Guida*] stands, therefore, in contrast to certain older methods that have the following two ideas in common:

- The first idea is that the uncertainty reported should be 'safe' or 'conservative' (...) In fact, because the evaluation of the uncertainty of a measurement result is problematic, it was often made deliberately large.

⁵Non ci sarebbe alcun bisogno di simulare il processo al computer, dato che la soluzione può essere ottenuta analiticamente mediante il calcolo delle probabilità, ma l'esperienza mi insegna che le simulazioni possono essere più convincenti per alcune persone.

⁶Ci si potrebbe chiedere: come mai questo processo non porta ad un collasso? Semplicemente perché nei laboratori non si seguono queste regole e, invece di nascondere la testa nella sabbia degli errori massimi, si cerca di ricalibrare in continuazione strumenti e procedure. Questo è quanto dovrebbe imparare subito anche lo studente

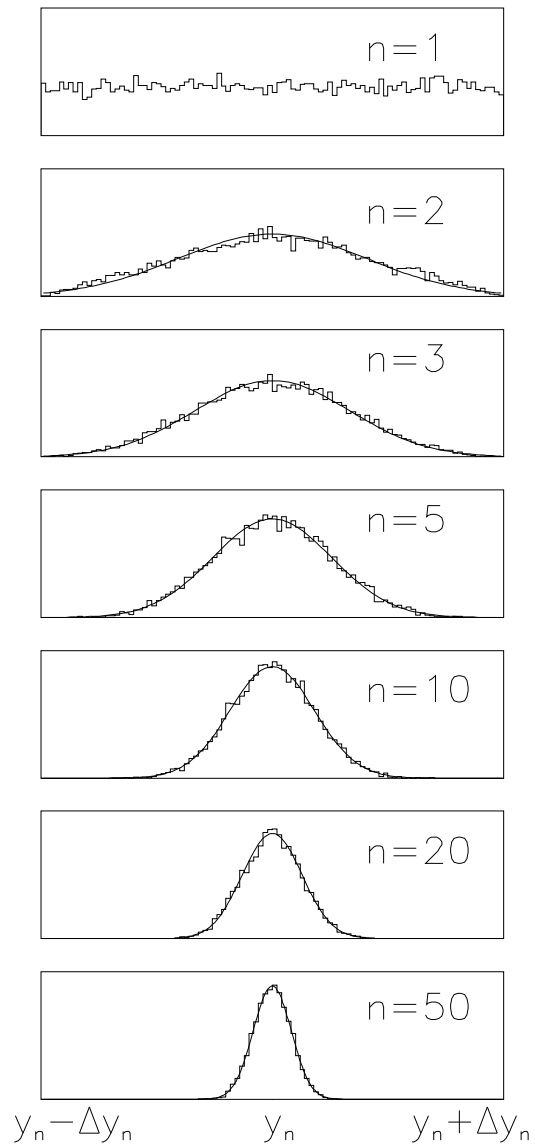


Figura 3: Simulazione della distribuzione del valore vero ottenuta sommando n risultati aventi gli stessi limiti di errore. Per confronto viene anche riportata la distribuzione normale avente come media il centro dell'intervallo e deviazione standard $\sqrt{n}/\sqrt{12}$ (vedi appendice sul teorema del limite centrale).

• ...

(...) if the 'maximum error bound' (the largest conceivable deviation from the putative best estimate) is used (...) the resulting uncertainty (...) will be unusable by anyone wishing to incorporate it into subsequent calculations (...)”.

Comunque, il motivo principale per cui vanno evitate le sovrastime delle incertezze è che in questo caso è più facile arrivare a risultati in accordo (artificiosamente) con valori noti o con quelli di altri esperimenti. Questo impedisce di identificare i possibili effetti sistematici che possono distorcere il risultato (si ricordi che spesso dietro gli errori sistematici c'è quasi sempre della Fisica: dispersioni termiche, rumore elettromagnetico, approssimazioni rozze, etc.), o di scoprire addirittura una nuova fenomenologia (ma questo non capita nelle esperienze di laboratorio didattico ...). Aumentare artificialmente le incertezze equivale a rifiutarsi di imparare. Farlo per “paura di sbagliare” è puerile⁷.

6.2 Regola della mezza divisione

Questa è una delle regole più radicate nella mente di chi ha seguito corsi di teoria della misura, una sorta di dogma al quale credere, scarsamente supportato (se preso alla lettera) da giustificazioni teoriche o pratiche. In realtà è abbastanza semplice convincersi che:

- $\Delta x = \frac{1}{2}$ divisione non corrisponde all'errore di lettura⁸: provare per credere! Ad esempio, una semplicissima esperienza consiste nel fare delle misure con un calibro e confrontare il valore stimato interpolando fra le tacche (distanziate un millimetro) con quello letto sul nonio (vedi tabella 4 in Appendice). Il risultato che si ottiene è ben lontano da un errore di mezza divisione. Si notano scarti tipici al più dell'ordine di un decimo di divisione e la deviazione standard tipica degli scarti interpolazione-nonio è inferiore al decimo di divisione, con un massimo di frequenza intorno a ≈ 0.7 decimi, un valore niente affatto casuale alla luce di quanto vedremo fra breve.

In effetti, questo è in linea con la tradizione classica (tuttora in voga in giro per il mondo) che raccomanda di sforzarsi di leggere fra le divisioni. Mostriamo, come curiosità, una figura tratta da un articolo su *Nature* (del 14 marzo 1996, Vol. 380, pag. 101) sull'astronomo Nevil Maskelyne.

⁷Questo aspetto psicologico non riguarda soltanto gli studenti. Non è raro vedere anche nella ricerca avanzata risultati in sorprendente accordo fra di loro o con predizioni teoriche nonostante le loro enormi barre di incertezza, o fisici sperimentali preoccupati se i loro valori differiscono di un paio di deviazioni standard da una “solida predizione” o da un risultato precedente.

⁸A volte lo si sente chiamare anche **errore di sensibilità**, o addirittura semplicemente “sensibilità” (in una nota per studenti si legge testualmente: “l'indeterminazione su tali grandezze può essere presa pari alla sensibilità del termometro impiegato, ovvero mezza tacca”). In questo caso “sensibilità” starebbe per “risoluzione” (vedi norma DIN 1319, 2, 9, che incontreremo fra poco). E' raccomandabile utilizzare il termine “sensibilità” per indicare ... la sensibilità, ovvero, detto alla buona, “il rapporto fra la variazione della risposta e la variazione dello stimolo”.

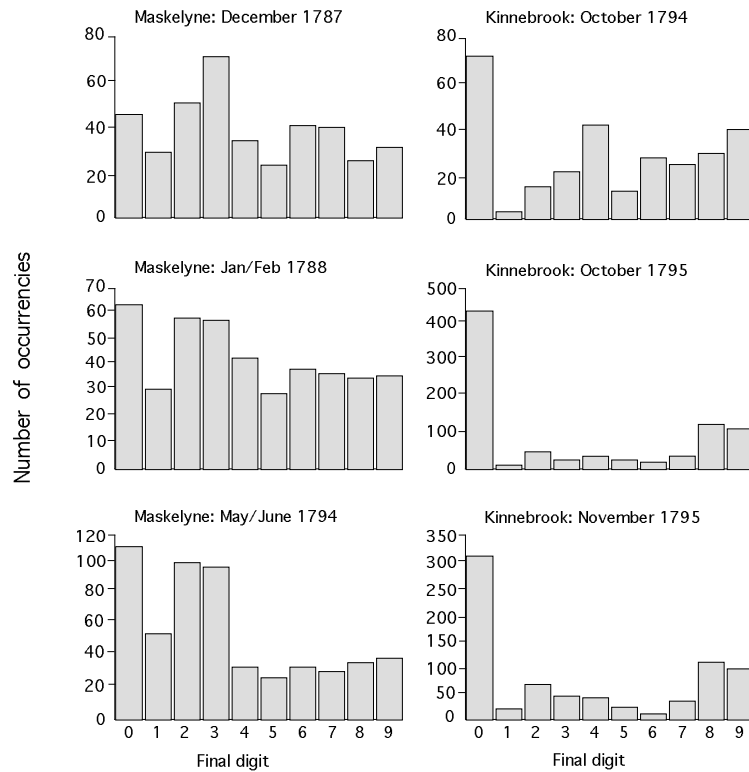


Figura 4: Istogramma dell'ultima cifra significativa nei dati di Maskelyne e in quelli del suo assistente.

Questi licenziò il suo assistente, accusandolo di non essere accurato nelle letture (qualcuno insinua che questo non sia stato il motivo principale, ma per noi è irrilevante). La figura 4 riporta la distribuzione dell'ultima cifra (ovvero quella stimata interpolando fra le tacche) delle misure di Maskelyne e del suo assistente. Si noti come anche quest'ultimo abbia una certa tendenza ad arrotondare un po' troppo, o a prediligere certe cifre, ma niente a che vedere con la superficialità di Kinnebrook.

E' interessante vedere cosa raccomandano le varie norme degli istituti di metrologia a proposito degli strumenti a lettura analogica.

- “Line scales mainly have a scale numbering with regular spacing and are mostly intended for a continuous indication of measured values”. (DIN 1319, part 2, 6.1.1) [Per “continuo” si intende che la quantizzazione della lettura alla mezza divisione è arbitraria.]
- “Unduly small scale spacing (less than approx. 0.7 mm) should be avoided, since such scales are tiring to read and in particular the estimating of tenths is impossible so that the observation is rendered less certain.” (DIN 1319, part 2, 6.3)
- “In some areas of metrology the term “resolution” is used. This is understood to mean the small change in the value of the measurand which is necessary to produce a perceptible (often specified) small change in the

response (in the case of measuring instruments with scale indication, for example, 1/5 of the scale interval)". (DIN 1319, part 2, 9)

- “In un formato per osservatore umano l’incertezza di lettura dipende dalle caratteristiche costruttive della scala e dell’indice, dalle modalità d’osservazione, dal rumore eventuale e dall’abilità dell’osservatore. Per esempio se si ammette che un osservatore di normale abilità, leggendo lo strumento nella posizione appropriata, possa stimare 1/5 di divisione, si indicherà come incertezza di lettura ± 0.1 divisioni.”⁹ (UNI 4546, 5.5)

Ne segue che, quando le condizioni di misura lo permettono, bisogna sforzarsi a leggere fra le tacche¹⁰.

- $\Delta x = \frac{1}{2}$ divisione non corrisponde all’errore di calibrazione (si sente ripetere spesso “il costruttore ha disegnato le tacche in modo tale che . . .”). Per convirsene, è sufficiente leggere le norme (ISO, DIN, UNI) a cui i costruttori di strumenti si devono attenere. Si scopre allora come il possibile errore di calibrazione possa essere, in taluni casi, ben inferiore al decimo di divisione (nel caso dei righelli, per esempio), mentre in altri si arriva addirittura ad alcune divisioni (il caso di alcuni termometri). Ad esempio, la norma su “*Attrezzi da disegno - Modalità di controllo e precisione per squadre, righe e multidecimetri*” (UNI 5131) riporta:

- *Sulla lunghezza l della parte millimetrata è ammessa una tolleranza di*

$$\pm \frac{0.20 l}{1000}.$$

Quindi per i normali righelli abbiamo tolleranze di circa ± 0.1 mm per letture fatte a fondo scala (ed, in ogni caso, le letture prossime sono necessariamente correlate, in quanto un eventuale difetto dello strumento si ripercuote in entrambe).

Un ultimo commento sulla regola della mezza divisione: essa implicherebbe che

- questo sia il solo errore in gioco, mentre, come visto precedentemente, sono molte le cause di incertezza e molto spesso lo sperimentatore fa parte integrante del processo di misura;
- i vari errori siano non correlati (per quanto riguarda il successivo uso nelle propagazioni).

Da questa regola discende inoltre l’imperativo di riportare i punti sperimentali sui grafici sempre con le loro “barre di errore”, che sarà commentato fra breve.

⁹Essere praticamente sicuri che il valore sia entro il 1/5 di divisione, vuol dire che, se ci si sforza al interpolare al meglio, ci si aspetta una deviazione standard dell’errore di lettura di circa $0.2/\sqrt{12}$ divisioni, compatibile al valore di ≈ 0.7 che si osserva sperimentalmente.

¹⁰**Perché non cambiare strumento?** Domanda legittimissima. Il problema è che questo non è sempre possibile. Quindi è importante, all’occorrenza, imparare a sfruttare tutta la potenzialità degli strumenti a disposizione. Queste dovrebbero essere le **regole del gioco** sulle quale sviluppare un corso di teoria e pratica di valutazione delle incertezze di misure.

6.3 $\Delta t = 0.2\text{ s}$?

Accenniamo rapidamente ad un'altra regola non giustificata. Si dice spesso che l'errore dovuto nelle misure di cronometraggio manuale sia di due decimi di secondo, dovuto ai riflessi umani.

Anche se è vero che il tempo medio di riflesso è di circa 0.2 s (intorno a 180 ms per studenti mediamente svegli) questo non ha niente a che vedere con l'errore sulla misura di cronometraggio. Questo dipende invece dall'entità delle fluttuazioni rispetto al ritardo medio: se, per assurdo, il dito di uno studente rispondesse anche 1 secondo dopo lo stimolo, ma senza fluttuazioni, l'errore sarebbe nullo. Quindi, ancora una volta la misura dipende dalla persona (ci aspettiamo che Max Biaggi si comporterà meglio di qualsiasi vincitore di premio Nobel ...) e dalle condizioni di misura. In Appendice (paragrafo B) vengono illustrate delle semplici esperienze per permettere a ciascuno di valutare le proprie capacità. Effettuandole si potranno osservare errori tipici di alcuni centesimi di secondo¹¹.

6.4 Imperativo categorico di riportare le “barre di errore”

Nessuno mette in dubbio l'importanza di riportare sui grafici i valori misurati con le relative barre di incertezza. Il solo problema è che queste barre dovrebbero essere veramente associate ad una incertezza, in modo consistente con la sua definizione. Purtroppo questo non è vero se, come succede spesso:

- si utilizza la regola della mezza divisione come punto di partenza;
- le incertezze su grandezze misurate indirettamente vengono valutate con la propagazione lineare delle incertezze.

Anche se si facesse uso di altri criteri e procedure meno criticabili, sia per l'incertezza dovuta allo strumento che per la legge di propagazione, partire da tali incertezze implica trascurare altri fattori che intervengono nell'incertezza e che possono essere più importanti di quella di lettura e di eventuale calibrazione dello strumento.

Quello che si fa generalmente nel mondo della ricerca è riportare sul grafico semplicemente i punti osservati e valutare l'incertezza dalla dispersione dei dati (“*residui*”), lungo un andamento noto (o ipotizzato) dei dati sperimentali. La figura 5 mostra due grafici relativi al comportamento di una bilancia di altissima precisione, pubblicati da T. Quinn¹², direttore del BIPM (tanto per prendere un ricercatore al di sopra di ogni sospetto ...).

6.5 Rette di massima e minima pendenza

Anche la procedura di stimare i parametri di un andamento lineare dalle cosiddette rette di massima e minima pendenza è un derivato della “teoria degli

¹¹Si noti inoltre che, quando si misura una grandezza fisica (X) in funzione del tempo (t), non ha molto senso parlare di errori su t e su X , in quanto ogni differenza dell'istante di lettura dal tempo nominale si rifletterà in un errore sulla grandezza fisica. Quindi, ai fini del risultato finale, è più che ragionevole attribuire tutto l'errore a X e considerare t esente da errore (si veda anche il paragrafo 16.4.2).

¹²T.J. Quinn, “*The beam balance as an instrument for very precise weighing*”, Meas. Sci. Technol., **3**(1992), 141.

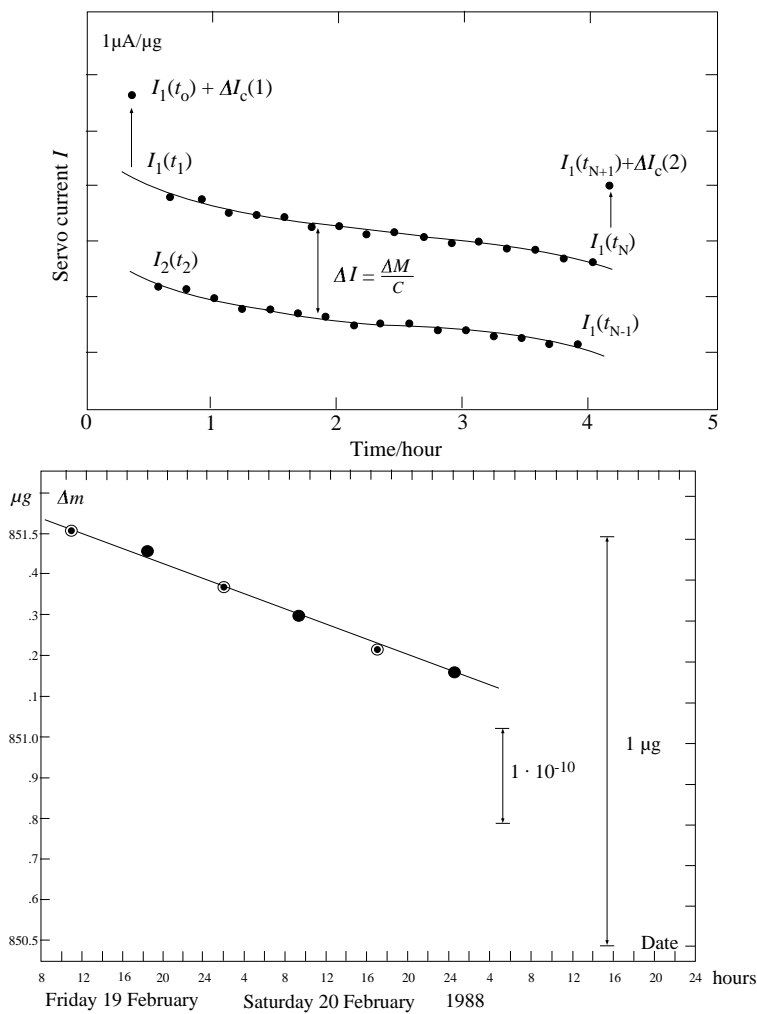


Figura 5: Grafici relativi ad una bilancia di altissima precisione (deviazione standard di 4 parti su 10^{12}) pubblicati dal direttore del BIPM: si noti l'assenza delle barre di incertezza.

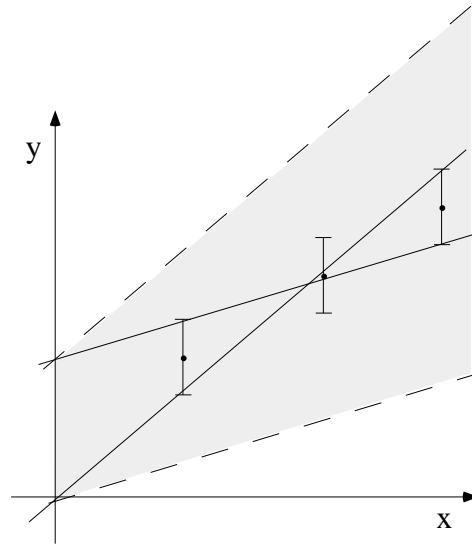


Figura 6: Regione di plausibilità della legge fisica che descrive i dati sperimentali, come risulta dal risultato dato secondo la procedura delle rette di massima e minima pendenza.

errori massimi”, ove l’intervallo di valori fra le barre è considerato certo. Da questo punto di vista c’è almeno una certa coerenza logica in quello che, alla luce dell’esperienza, si rivela come un paradosso che infastidisce gli studenti:

- peggiore è l’accordo fra i punti sperimentali (con le loro barre di incertezza), migliore (e spesso in modo imbarazzante, alla luce di valori veri noti) è la precisione sui parametri.

A questo è da aggiungere che:

- il risultato simultaneo dei due parametri è spesso “non informativo”, nel senso che si presta ad interpretazioni molto vaghe e non accettabili alla luce dei dati sperimentali. Infatti, nel presentare un risultato con

$$\begin{cases} m = m_0 \pm \Delta m \\ c = c_0 \pm \Delta c \end{cases}$$

(m e c stanno rispettivamente per coefficiente angolare e intercetta) si perdono informazioni: chi non ha accesso al grafico può immaginare un andamento vero che può essere compreso, senza alcuna preferenza, fra

$$y = (m_0 - \Delta m)x + (c_0 - \Delta c)$$

e

$$y = (m_0 + \Delta m)x + (c_0 + \Delta c).$$

Questo corrisponde, in alcune esperienze didattiche ad aver misurato più o meno niente (vedi ad esempio figura 6).

- Il risultato dipende dalla scelta degli assi e non è invariante per traslazione. Ad esempio, la regione di plausibilità si riduce se si sceglie l’origine dell’asse delle ascisse al centro dei punti. In particolare, con un po’ di fortuna (o di sfortuna?) si può ottenere $\Delta c = 0$! Questo è dovuto al fatto che con questa procedura non c’è alcun modo di tener conto delle correlazioni fra i parametri, correlazioni inevitabili in quanto essi sono ottenuti dalle stesse informazioni di partenza.

7 Critica degli “errori statistici”

L’altra regola di propagazione di incertezze generalmente nota (ma, al dire il vero, non troppo fra gli insegnanti di scuola media) è quella cosiddetta degli “errori statistici”, che riportiamo per comodità:

$$\sigma^2(y) = \left(\frac{\partial y}{\partial x_1}\right)^2 \sigma^2(x_1) + \left(\frac{\partial y}{\partial x_2}\right)^2 \sigma^2(x_2) + \dots \quad (11)$$

Essa è decisamente meglio di quella precedente, se non altro in quanto si sostituiscono probabilità a incertezze. Ma all’atto pratico anche questa formula presenta i suoi problemi.

- Innanzitutto è da premettere il dato di fatto che molti studenti studiano questa formula in modo astratto, senza nessuna applicazione durante l’intero corso di laurea, e quindi si crea un atteggiamento di diffidenza nei suoi confronti. E difatti, alla prima occasione in cui si tenta di applicarla, nascono i problemi.
- Seconda premessa è che la (11) non è completa, essendo valida soltanto nel caso in cui le x_i sono indipendenti, condizione che è violata qualora le grandezze sono misurate con lo stesso strumento, un caso tutt’altro che astratto.
- Comunque, il primo problema legato a tale formula è quello di interpretazione. Per qualcuno potrà sembrare un cavillo filosofico, ma in realtà è un punto cruciale. Cosa significa $y \pm \sigma(y)$? La stragrande maggioranza delle persone interpellate sono concordi nell’affermare che (assunto un modello gaussiano) essa voglia indicare

$$P[y - \sigma(y) \leq y_v \leq y + \sigma(y)] = 68 \% : \quad (12)$$

“c’è il 68 % di probabilità che il valor vero di y si trova nell’intervallo $y \pm \sigma(y)$ ”.

Quando poi si chiede *cosa sia la probabilità* si ottengono risposte tipiche (“casi favorevoli su casi possibili” e “limite della frequenza”) che non contemplan affermazioni probabilistiche sui valori veri, così come sono espresse dalla (12).

- Un problema pratico tipico è quello di “cosa mettere nelle $\sigma(x_i)$ ” della (11). Siccome questa formula deriva dal calcolo delle probabilità, applicato alle variabili casuali, le x_i e la y che entrano nella formula devono

avere il significato di variabile casuale e le $\sigma(x_i)$ quello di deviazione standard. Quindi, se non si associano variabili casuali ai valori veri, l'uso della (11) è arbitrario.

- Nel caso di n misure ripetute, si impara che le $\sigma(x_i)$ vanno calcolate come “ σ/\sqrt{n} ” (nonostante si incontra ancora qualcuno diffidente del fattore $1/\sqrt{n}$ e che preferisce ometterlo per “non avere errori troppo piccoli”). Purtroppo, non sempre è possibile effettuare molte misure che mostrino una variabilità da manuale dei valori letti. Come comportarsi, ad esempio, se:
 - si effettua una sola misura ($n = 1$)?
 - si legge un grandissimo numero di volte (“ $n \rightarrow \infty$ ”) lo stesso valore (ad esempio 3.512 V su uno strumento digitale)?
- Come comportarsi se sono presenti anche “errori sistematici”?
- Come valutare e gestire le correlazioni fra diverse misure introdotte, ad esempio, da errori sistematici comuni?

La conseguenza di questi problemi tecnici (usualmente quello di principio sull'interpretazione della probabilità non viene nemmeno preso in considerazione) è che in genere gli studenti imparano delle formule che poi non utilizzeranno e seguitano a lavorare con gli errori massimi¹³.

8 Riassumendo

Volendo fare il punto prima di proseguire, diciamo che la situazione appare sconcertante: la teoria degli “errori massimi” è visibilmente incongruente dal punto di vista teorico e insoddisfacente dal punto di vista pratico; non si sa bene come comportarsi con quella degli “errori statistici”.

In particolare, questa rassegna critica dovrebbe aver fatto sorgere al lettore dei seri dubbi su

- validità del concetto stesso di errore massimo;
- dogma della 1/2 divisione;
- propagazioni lineari degli errori massimi;
- imperativo categorico delle barre d'errore;
- retta di massima e minima pendenza;
- validità delle affermazioni probabilistiche sui valori veri;

¹³Qualcuno prova a trasformare “errori massimi” in “errori statistici”, considerando $\Delta x = 3\sigma(x)$ e, nella direzione opposta, $\sigma(x) = \Delta x/\sqrt{6}$ (assumendo una distribuzione uniforme del valore vero di x entro $2\Delta x$). La seconda trasformazione è ragionevolissima se veramente si crede che x possa assumere qualsiasi valore entro $\pm\Delta x$, sebbene questo credere sia in contrasto con le interpretazioni usuali di probabilità. La trasformazione inversa ($\sigma \rightarrow \Delta$), con l'uso successiva delle propagazioni lineari è invece assurdo in quanto in contrasto con le proprie credenze (gli errori massimi assumono, tacitamente, indifferenza entro $\pm\Delta$).

- uso pratico della propagazione degli errori statistici.
- terminologia varia.

Anticipiamo, tanto per fare una lista completa,¹⁴ altri punti che affronteremo:

- concetto di probabilità;
- “dogma dell’immacolata osservazione”¹⁵;
- classificazione dell’incertezza in “statistica” e “sistemica”.

Si tratta quindi di rivedere completamente il processo che permette di imparare dai dati sperimentali. Parafrasando Kant[?] si può dire (sostituendo le parole in corsivo con quelle fra parentesi):

“Tutti i *metafisici* (fisici) sono quindi solennemente e legittimamente sospesi dalle loro funzioni fino a tanto che abbiano soddisfacentemente risposto alla domanda come *sono possibili le conoscenze sintetiche a priori* (è possibile imparare dalle osservazioni)?”

Chiaramente, questa citazione è da prendere in modo scherzoso (per quanto riguarda l’invito a sospendere le attività . . .), ma, riflettendoci bene, si nota come essa sia più pertinente di quanto si possa pensare inizialmente. Infatti la *critica di Hume* al problema dell’*induzione*, che aveva “interrotto il sonno dogmatico” del grande filosofo tedesco, è sopravvissuta anche all’analisi dello stesso Kant¹⁶. Ritourneremo su questo argomento fra qualche paragrafo.

9 Da dove ricominciare?

Per ricostruire una teoria delle incertezza di misura che non soffra di tutte i problemi mostrati, partiamo da alcune considerazioni.

1. Un po’ in analogia del “cogito” cartesiano, a questo punto, l’unica affermazione sulla quale è difficile non essere d’accordo è quanto detto nell’introduzione:

¹⁴Un altro punto molto critico, ma su cui non entreremo, è quello legato ai cosiddetti “test di ipotesi”.

¹⁵Ho sentito questa espressione dal neurologo Sergio Della Sala durante la conferenza annuale del CICAP (Comitato Italiano per il Controllo delle Affermazioni sul Paranormale), Padova, Novembre 1997. L’originale dovrebbe essere di Umberto Eco, forse in “Kant e l’ornitorinco” . . .

Il significato che attribuisco a questa espressione in questo contesto sarà chiaro a partire dal paragrafo 13: la mera osservazione empirica (“un numero su un display”) non accresce la Conoscenza, se questa informazione viene avulsa dal contesto di “credenze” che contornano misurando, strumento di misura e processo di misura.

¹⁶E’ molto espressivo il commento di Bruno de Finetti sull’insuccesso di Kant a far fronte alla critica di Hume: “Ma le reazioni contro ogni chiarificazione intelligente sono sempre pronte e pieno di sacro zelo, in difesa della sacra ottusità: ecco il povero Kant affannarsi a tamponare la falla aperta da Hume ed a rabberciare la sconnessa fabbricazione tradizionale, dove il ragionamento induttivo si vuole a forza ricollegato e inserito, al pari di quello deduttivo, nelle strutture anguste della logica del certo”.

il processo di induzione dalle osservazioni ai valori di grandezze fisiche conduce ad affermazioni che, inevitabilmente, sono affette da un certo grado di incertezza.

2. Il concetto *naturale* sviluppato dalla mente umana per quantificare la plausibilità delle affermazioni in situazioni di incertezza è quello di probabilità.

Si tratta quindi di costruire una teoria probabilistica (probabilistica e non, genericamente, “statistica”) dell’incertezza di misura.

Questi due punti di partenza sembrano assolutamente ragionevoli, ma il secondo appare in contraddizione con la critica sull’interpretazione probabilistica del risultato, avanzata nel paragrafo precedente. In realtà questo non è un vero problema, ma soltanto un prodotto di una visione distorta (cioè diversa da quella naturale) del concetto di probabilità. In effetti la maggior parte dei fisici stessi, pur “credendo” che la probabilità sia “il rapporto fra casi favorevoli e casi contrari” o “limite della frequenza”¹⁷, si stupiscono quando vengono a sapere che l’affermazione

$$P\left(\bar{x} - \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}}\right) = 68 \% \quad (13)$$

è illegittima¹⁸ (indichiamo con μ il valore vero e con x il valore osservato). Infatti, secondo la statistica convenzionale, non hanno senso affermazioni probabilistiche sul valor vero. Esso sarebbe un valore “costante, ma ignoto”. In tale approccio si può affermare soltanto che

$$P\left(\mu - \frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + \frac{\sigma}{\sqrt{n}}\right) = 68 \% . \quad (14)$$

Ma questa è un’affermazione probabilistica su \bar{x} , dati μ e σ . Non sono ammesse invece affermazioni probabilistiche su μ , sebbene sia a queste che lo sperimentatore faccia riferimento quando esegue un esperimento per “diminuire lo stato di incertezza su μ ”.

¹⁷Assumiamo che il lettore sia al corrente delle “definizioni” standard, quelle che si studiano comunemente

¹⁸Nessun testo serio di probabilità convenzionale riporta la (13). Si parla invece di “intervallo di fiducia”, che ha però tutt’altro significato, anche se diversi testi e molti insegnanti ne suggeriscono un’interpretazione probabilistica. Senza entrare nei dettagli, ad uso di chi è familiare con questi concetti, chiariamo brevemente come la ragione di fondo di questa contraddizione sia da ricercarsi nel rifiuto di accettare l’interpretazione di probabilità come grado di fiducia. Il concetto frequentistico di intervallo di fiducia è quindi una sorta di forzatura inventata per caratterizzare l’incertezza in un modo consistente con la visione frequentistica di probabilità (vedi nel seguito). Purtroppo - è un dato di fatto - tentare di classificare lo stato di incertezza evitando il concetto di probabilità conduce a fraintendimenti. Emblematico di questi ben noti problemi è quanto risultava da una tavola rotonda fra statistici americani alla quale ho assistito in occasione del loro congresso annuale del 1996: “*i nostri studenti non capiscono gli intervalli di fiducia*”.

10 Probabilità soggettiva

Quindi il primo concetto da rivedere è quello di probabilità. Nei limiti di spazio di questa trattazione, diciamo soltanto che:

- sostanzialmente si assume che il concetto di probabilità sia primitivo, ovvero vicino a quello del senso comune. Per dirlo in un modo scherzoso, il concetto di probabilità è quello che si ha “prima di andare a scuola” e che si seguita ad usare inconsciamente dopo, “nonostante quello che si è appreso”¹⁹;
- detto altrimenti, la probabilità è una *misura del grado di fiducia* (o di credenza, in inglese *degree of belief*) che una qualsiasi affermazione risulti essere vera;
- il valore di probabilità va da 0 a 1 per affermazioni che vanno dall’impossibile al certo;
- siccome quanto più si crede che una certa affermazione sia vera, tanto più si è disposti a scommettere su di essa, la *scommessa “coerente”*²⁰ può essere utilizzata per definire operativamente il valore di probabilità;
- è facile convincersi che le “definizioni” standard non possono invece definire il concetto di probabilità in quanto esse assumono il concetto di probabilità. Prese alla lettera, esse sono infatti definizioni circolari:

– (*definizione combinatoria*)

$$p = \frac{\text{numero dei casi favorevoli}}{\text{numero dei casi possibili (se egualmente probabili)}};$$

– (*definizione frequentista*)

$$p \approx \frac{\text{numero delle prove favorevoli}}{\text{numero totale delle prove}},$$

quando sono state effettuate un grande numero di prove nelle stesse condizioni (equiprobabilità);

¹⁹Questa osservazione deriva dalla constatazione che, come detto, la maggior parte dei fisici interpellati sia convinta in buona fede della legittimità della (13), pur sostenendo che la probabilità sia il “limite della frequenza”.

²⁰Senza entrare nel dettaglio, chiariamo brevemente cosa si intende per “coerente” (o “*re-versible*”): una volta fissate le quote di scommessa pro e contro l’evento (proporzionali alla probabilità dell’evento e del suo opposto), deve essere indifferente allo scommettitore il verso della scommessa: se c’è una netta propensione pro, vuol dire che bisogna alzare la quota in favore dell’evento; nel caso opposto bisogna alzare l’altra quota. Il rapporto delle quote, in condizione di indifferenza sul verso da scegliere, è una valutazione del rapporto delle probabilità. Quindi il valore della probabilità è dato dalla quota di scommessa sull’evento divisa per il totale delle quote.

Si può dimostrare che la coerenza fornisca (come teoremi!) le regole sintattiche della probabilità analoghe a quelle espresse dai ben noti assiomi. Inoltre da essa si deriva anche la relazione che lega probabilità condizionata alla probabilità congiunta evento-condizionante e a quella del condizionante (mentre nell’approccio assiomatico questa formula è una definizione, con il risultato di produrre conseguenze paradossali). Un’altro aspetto importante della coerenza è che essa fa sì che le valutazioni soggettive siano tutt’altro che “arbitrarie”.

- le due “pseudo-definizioni” sono prontamente recuperate come regole di valutazione della probabilità, qualora colui che effettua la valutazione ritenga che le clausole siano soddisfatte.

Il concetto di probabilità che abbiamo brevemente illustrato è quello della cosiddetta *probabilità soggettiva*, ad indicare che il valore dipende dallo stato di informazione del *soggetto* che effettua la valutazione. Anche se questo approccio alla probabilità può sembrare inizialmente sospetto (l’aggettivo “soggettivo” suona decisamente male per tutti coloro che, in perfetta buona fede, desiderano il progresso della scienza, nel modo più “oggettivo” possibile), esso è quello meglio fondato e il più produttivo, come anche riconosciuto nella già citata *Guida ISO*:

“(. . .) In contrast to this frequency-based point of view of probability an equally valid viewpoint is that probability is a measure of the *degree of belief* that an event will occur.
(. . .) Recommendation INC-1 (. . .) implicitly adopts such a viewpoint of probability (. . .)”.

Il punto di forza di questa interpretazione della probabilità è, oltre al recupero del concetto intuitivo, la possibilità di fare affermazioni probabilistiche su qualsiasi evento, indipendentemente dal fatto di avere un problema perfettamente simmetrico (“casi possibili e casi favorevoli”) o di poter ripetere l’esperimento un grande numero di volte (“limite della frequenza”).

11 Formalismo sulle variabili casuali: “ $f(x)$ ”, $E(X)$ e σ

Prima di procedere alle applicazioni, ricordiamo brevemente la terminologia sulle variabili casuali (nel linguaggio della probabilità soggettiva).

- Una *variabile casuale*, o *numero aleatorio*, è qualsiasi numero rispetto al quale si è in stato di incertezza. Facciamo due esempi nel contesto delle misure:
 1. Pongo un chilogrammo campione su una bilancia di laboratorio con indicazione (digitale) dei centesimi. Che valore leggerò (in grammi)? 1000.00, 999.95, 1000.03 . . . ?
 2. Leggo su una bilancia di laboratorio 2.315 g. Quanto vale il valore vero della massa del corpo? 2.311, 2.312, . . . 2.315, . . . 2.319, . . . ?

Nel primo caso la variabile è la lettura x (subordinatamente ad un certo valore vero); nel secondo caso la variabile è il valore vero μ (subordinatamente ad un certo valore letto).

- Ai possibili valori della grandezza viene associata una funzione $f(x)$ che quantifica il grado di fiducia ad essi assegnato. Quindi scrivere che $f(x_1) > f(x_2)$ sta ad indicare che si crede più a x_1 che a x_2 . A seconda che la variabile x sia *discreta* o *continua*, $f(x)$ ha l’accezione di *funzione di probabilità* o di *funzione densità di probabilità*.

- Tutte le proprietà di $f(x)$ apprese nei corsi convenzionali rimangono valide nell'approccio soggettivista. In particolare si ricorda che il *valore atteso*, indicato con $E(X)$ e calcolato come *media* dei possibili valori di x pesati con $f(x)$, dà il *baricentro* della distribuzione; la deviazione standard, indicata con σ e calcolata come *radice quadrata del momento di inerzia* della distribuzione (leggi *varianza*: “media dei quadrati degli scarti”), fornisce la dispersione di valori che è possibile attendersi dalla variabile.
- Tutte le distribuzioni di variabile casuale sono subordinate ad un certo stato di informazione. Utilizzando i due esempi precedenti possiamo perciò scrivere

$$\begin{aligned} f(x) &\longrightarrow f(x | \mu = 1000.00) \\ f(\mu) &\longrightarrow f(\mu | x = 2.315), \end{aligned}$$

ove “|” si legge “dato”, “subordinatamente a”, etc.

- L'intero stato di incertezza sui valori della grandezza di interesse è espresso da $f(\mu)$. Da questa funzione è possibile calcolare la probabilità che la grandezza abbia un valore compreso in un certo intervallo. Per semplicità, spesso si riassume lo stato di informazione di $f(\cdot)$ in due soli numeri: $E(\cdot)$ e σ . E' interessante l'interpretazione conoscitiva di queste due grandezze. Esse possono essere viste come la *previsione* e l'*incertezza di previsione* del numero aleatorio.
- Fra le distribuzioni di probabilità, quella più importante per la trattazione delle incertezze di misura è indubbiamente la ben nota *gaussiana*, che assumiamo nota. Altre distribuzioni interessanti per le applicazioni sono la *distribuzione uniforme* e la *distribuzione triangolare*. Anche la distribuzione uniforme è generalmente ben conosciuta. Ricordiamo soltanto che essa ha una deviazione standard pari alla sua larghezza divisa $\sqrt{12}$. Utilizzando la semilarghezza Δ , si ha:

$$\sigma(\text{distr. uniforme}) = \frac{2 \Delta}{\sqrt{12}} = \frac{\Delta}{\sqrt{3}}.$$

Essendo la distribuzione triangolare meno nota, essa è descritta in Appendice C). Nel caso “isoscele” (triangolare simmetrica) di semiampiezza Δ la deviazione standard vale

$$\sigma(\text{triangolare}) = \frac{\Delta}{\sqrt{6}}.$$

12 Valutazione dell'incertezza di misura: schema generale

Avendo presentato tutti gli ingredienti necessari, vediamo come sviluppare una teoria delle incertezze di misure. Nei prossimi paragrafi affronteremo il problema a diversi gradi, cercando di limitare il formalismo e le complicazioni matematiche.

1. Il primo passo consiste nella valutazione dello stato di incertezza su una grandezza misurata *direttamente*, ossia il cui valore sia letto su uno strumento, senza compiere altri calcoli. Si terrà conto del fatto che a volte possono essere eseguite più misure indipendenti sullo stesso misurando nelle stesse condizioni.

Prima di passare alla parte puramente applicativa sono necessarie ancora delle premesse di carattere generale sul processo di apprendimento. Dobbiamo infatti:

- capire cosa vuol dire, dal punto di vista generale, riaggiornare lo stato di incertezza sul valore del misurando;
 - introdurre brevemente il problema delle inversioni di probabilità, ovvero di come passare dalla distribuzione di probabilità dell'osservabile alla distribuzione di probabilità della grandezza fisica;
 - accennare al meccanismo di aggiornamento bayesiano;
 - risolvere in modo intuitivo, anziché formale, l'apprendimento bayesiano nel caso di semplici problemi di routine.
2. Il secondo passo consiste nel propagare l'incertezza su grandezze misurate direttamente a grandezze misurate indirettamente. Anche in questo caso ci accontenteremo di medie e deviazioni standard.
 3. Il terzo passo riguarda la trattazione delle incertezze dovute ad errori sistematici di valore ignoto. Questo non è altro che un aspetto particolare della propagazione delle incertezze (sulle costanti di calibrazioni e altri fattori di influenza). Bisognerà però tenere conto dei seguenti fatti:
 - innanzitutto occorre convincersi che, anche se le incertezze sulle costanti di calibrazioni sono spesso descritte da un modello non gaussiano e valutate in modo abbastanza approssimativo, si può arrivare a distribuzioni finali gaussiane (teorema del limite centrale, vedi appendice D);
 - in secondo luogo, le correlazioni indotte dagli errori sistematici su diverse grandezze misurate con lo stesso strumento rendono complicato il calcolo della propagazione delle incertezze. Occorre quindi trovare delle scorciatoie per risolvere il problema, almeno nei casi semplici, e indicare una via generale (magari approssimativa).
 4. Sarà quindi introdotta la classificazione BIPM/ISO delle incertezze di misura in tipo *A* e tipo *B*, facendo notare la coerenza con quanto sviluppato in questo scritto.
 5. Come ultimo passo, considereremo la determinazione dei parametri della retta che meglio si adatta ai punti sperimentali. Nonostante la tendenza attuale di delegare tali compiti a programmi automatici su computer, raccomanderemo l'uso dell'analisi grafica, più istruttiva delle scatole nere informatiche e che possono portare a risultati numerici confrontabili a quelli ottenibili con "fit mediante i minimi quadrati".

13 Imparare dagli esperimenti: il problema dell'induzione

Ogni misura è eseguita lo scopo di accrescere la conoscenza di chi la esegue e di chi ha interesse a quella specifica conoscenza. Questi possono essere una certa comunità scientifica, un medico che ha prescritto una certa analisi o un commerciante che deve acquistare un prodotto. E' anche chiaro che la necessità stessa di eseguire misure indica che ci si trovava in uno stato di incertezza su qualcosa di interesse. Questo "qualcosa" può essere una costante fisica o una teoria sull'origine dell'universo; lo stato di salute di un paziente; la composizione chimica di un nuovo prodotto. In tutti i casi la misura ha lo scopo di *modificare* un certo stato di conoscenza.

Si sarebbe tentati di dire addirittura "acquisire", anziché "modificare", lo stato di conoscenza, come ad indicare che la conoscenza possa essere creata dal nulla nell'atto della misura. Non è difficile convincersi, invece, che nella maggior parte dei casi si tratta invece soltanto di un aggiornamento alla luce di fatti nuovi e di un certo raziocinio. Prendiamo ad esempio la misura della temperatura di una stanza, effettuata con un termometro digitale - tanto per escludere contributi soggettivi alla lettura dello strumento - e supponiamo di ottenere 21.7°C . Anche se si potrà dubitare del decimo di grado, indubbiamente la misura è servita a restringere l'intervallo di temperature ritenute plausibili prima della misura - quelle compatibili con la sensazione di "ambiente confortevole". In base alla conoscenza del termometro usato, o dei termometri in generale, ci saranno valori di temperatura in un certo intervallo intorno a 21.7°C ai quali crediamo di più e valori al di fuori ai quali crediamo di meno.

E' però altresì chiaro che se il termometro avesse indicato, a parità di sensazione fisiologica, 17.3°C si sarebbe tentati a ritenere che esso non funzioni bene. Non si avrebbero invece dubbi sul suo malfunzionamento se avesse indicato 2.5°C !

I tre casi corrispondono a tre diversi gradi di aggiornamento della conoscenza. Nell'ultimo caso, in particolare, l'aggiornamento²¹ è nullo.

Il processo di apprendimento dalle osservazioni sperimentali è chiamato dai filosofi *induzione*. Probabilmente a molti lettori sarà anche noto che in filosofia esiste l'irrisolto "problema dell'induzione" dovuto alla critica di Hume a tale processo. Questa può essere sintetizzata affermando che l'induzione non è "giustificata", nel senso che è impossibile dimostrare, con la stessa forza di un teorema matematica, che da certe osservazioni possano seguire *necessariamente* determinate conclusioni scientifiche. L'approccio probabilistico che abbiamo appena intrapreso sembra essere l'unica via d'uscita a tale critica²².

²¹Ma anche in questo caso si è imparato qualcosa, cioè che il termometro non funziona . . .

²²Molto spesso si pensa che l'unico metodo scientifico valido sia quello della falsificazione. Non ci sono dubbi che, se una teoria non è in grado di descrivere i risultati di un esperimento, essa vada scartata o modificata. Ma poiché non è possibile dimostrare la certezza di una teoria, diventa impossibile decidere fra tutte le (infinite) ipotesi non falsificate. Il metodo probabilistico permette di fornire una scala di credibilità a tutte le ipotesi considerate (o rapporti di credibilità fra ogni coppia di ipotesi). Un caso in cui il metodo di falsificazione è completamente inadeguato è quello relativo agli incertezze di misura. Infatti, prendendo alla lettera tale metodo, si sarebbe autorizzati soltanto a verificare se il valore osservato sullo

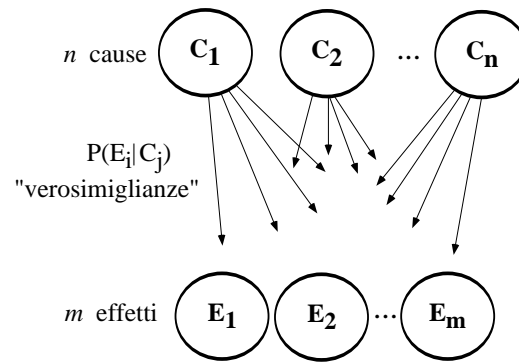


Figura 7: Relazioni cause-effetti viste in termini di condizionanti e eventi condizionati

14 Dalla probabilità degli effetti alla probabilità delle cause

Per formalizzare il discorso appena fatto, occorre associare variabili casuali sia ai possibili valori delle grandezze fisiche che ai valori osservabili sullo strumento. Fatto ciò, si tratterà di imparare come inferire la distribuzione di probabilità del valore vero, ossia come valutare, per ogni possibile valore della grandezza, un corrispondente grado di fiducia.

14.1 Verosimiglianza

Cominciamo con la distribuzione di probabilità dei valori osservabili, indicati con x . Come detto,

$$f(x|\mu)$$

sta per la funzione densità di probabilità (x è una variabile continua, dal punto di vista pratico) di osservare²³ un certo valore x , dato un determinato valore vero μ . Tutti i possibili valori di μ possono essere visti come le infinite *cause* responsabili del valore x osservato (il loro *effetto*). La figura 7 dovrebbe aiutare ad illustrare il problema.

La funzione $f(x|\mu)$ ci dà la verosimiglianza che μ possa causare x e per questo è chiamata semplicemente *verosimiglianza*. Essa va stimata dalla conoscenza del comportamento dello strumento e, più in generale, dell'insieme di tutte le procedure di misura. Molto spesso si utilizza per la verosimiglianza un modello gaussiano, giustificato un po' dall'esperienza e soprattutto dalle aspettative teoriche, basate sul teorema del limite centrale. Consideriamo,

strumento è compatibile o no con un valore vero, niente di più. Si capisce come, con queste premesse, non si possa fare molta strada.

²³Attenzione a non confondere la probabilità di osservare un certo valore x , subordinatamente ad un certo valore di μ , con la probabilità del valore che è stato effettivamente osservato. Essendo questo un numero certo (a meno di non essere ubriachi), ad esso non si applica il concetto di probabilità. Così pure, si faccia attenzione a non chiamare $f(x|\mu)$ "probabilità che x venga da μ " (il nome corretto - trascurando il fatto inessenziale che si tratta di una densità di probabilità e non di una probabilità - è "probabilità di x , dato un certo valore μ ", che è chiaramente ben altra cosa!).

quindi nel seguito, per semplicità, una verosimiglianza del tipo

$$f(x|\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (15)$$

con σ che non dipende dal valore di μ (stiamo assumendo che lo strumento risponda nello stesso modo a tutti i possibili valori di μ).

14.2 Probabilità iniziale e probabilità finale

Una volta fissata la funzione di verosimiglianza e un valore osservato x , si tratta di costruire la $f(\mu|x)$. Per arrivare in modo euristico alla formula generale, consideriamo soltanto due possibili valori di μ . Se, in base alle nostre conoscenze, riteniamo i due valori ugualmente probabili, ci sembrerà naturale protendere per il valore per il quale la verosimiglianza di osservare x è maggiore. Ad esempio, se $\mu_1 = -1$, $\mu_2 = 10$ e $x = 2$, si è tentati a credere che l'osservazione sia dovuta più verosimilmente alla causa μ_1 che alla causa μ_2 . Se però la grandezza di interesse è definita positiva, la causa μ_1 crolla da causa più probabile a causa impossibile. Ci sono poi casi intermedi in cui, per motivi legati all'esperienza precedente, si tende a credere *a priori* più ad una causa che all'altra. Ne segue che il grado di fiducia risultante di un certo valore di μ sarà proporzionale sia alla verosimiglianza che esso produca il valore osservato che al grado di fiducia che si attribuiva a μ prima dell'osservazione²⁴:

$$f(\mu|x) \propto f(x|\mu) \cdot f_o(\mu). \quad (16)$$

Questo è uno dei modi di scrivere il teorema di Bayes, che ha un ruolo centrale nelle inferenze probabilistiche. L'inessenziale fattore di proporzionalità è ricavato dalla condizione di normalizzazione (l'integrale su tutti i possibili valori di μ deve dare 1). $f_o(\mu)$ è chiamata distribuzione *iniziale*, o *a priori* (o più sinteticamente, in inglese, "*prior*") mentre $f(\mu|x)$ è la distribuzione *finale*, o *a posteriori*, ove il "prima" e il "dopo" è rispetto alla nuova osservazione x e non è da intendersi in modo strettamente temporale. La funzione $f_o(\mu)$ riassume lo stato di incertezza su μ alla luce di tutte le conoscenze a disposizione, a parte il verificarsi del dato sperimentale x . Quindi distribuzione iniziale e finale dovrebbero essere scritte, più precisamente come:

$$\begin{aligned} f_o(\mu) &\rightarrow f(\mu|I_o) \\ f(\mu|x) &\rightarrow f(\mu|"I_o\&x"). \end{aligned}$$

14.3 Paura dei "pregiudizi"? Inevitabilità di principio e frequente irrilevanza pratica delle prior

Molti possono rimanere perplessi al pensiero che le conclusioni scientifiche possano dipendere dal "pregiudizio" sulla grandezza fisica ("pregiudizio" ha

²⁴Si noti l'uso dello stesso simbolo $f(\cdot)$ per indicare funzioni di diverse variabili, anche se sarebbe formalmente più corretta una scrittura della (16) del tipo:

$$\varphi(\mu|x) \propto \mathcal{L}(x|\mu) \cdot \pi(\mu),$$

con $\varphi(\cdot)$, $\mathcal{L}(\cdot)$ e $\pi(\cdot)$ che ricordano dal nome, rispettivamente, la finale, la verosimiglianza (in inglese likelihood) e la prior.

correntemente un significato prevalentemente negativo, ma in realtà significa semplicemente un giudizio a priori, basato su una esperienza precedentemente acquisita). Non potendo addentrarci con la dovuta profondità in questo aspetto interessante del problema, aggiungiamo alcune note esplicative. Premettiamo una citazione di Poincaré²⁵, non tanto per tentare di convincere mediante il principio di autorità, quanto perché il concetto è espresso con molta chiarezza.

“Un effetto potrebbe essere prodotto dalla causa a o dalla causa b . L’effetto è appena stato osservato. Ci domandiamo la probabilità che sia dovuto alla causa a . Questa è una probabilità di causa a *posteriori*. Ma non la potrei calcolare, se una convenzione più o o meno giustificata non mi dicesse in anticipo qual’è la probabilità a *priori* che la causa a entri in gioco”.

Detto altrimenti, il contributo delle probabilità a priori è cruciale nei problemi di inferenza. Questo non deve però spaventare, in quanto:

- è assolutamente ragionevole trarre le conclusioni non in modo meccanico, ma alla luce della ragione;
- nelle misure di routine l’intervallo di “accettanza a priori” dei possibili valori è talmente ampio, rispetto alla larghezza della verosimiglianza, che in pratica è come se tutti i possibili valori di μ fossero ritenuti a priori ugualmente possibili. La prior viene allora ad essere assorbita nella costante di normalizzazione:

$$f(x | \mu) \cdot f_{\circ}(\mu) \xrightarrow{\text{prior molto vaga}} f(x | \mu); \quad (17)$$

- quando invece questo non è vero (ad esempio se si usa uno strumento con il quale non si ha confidenza, oppure se si deve valutare il risultato di una misura eseguita da persona inesperta) è assolutamente legittimo credere più ai propri pregiudizi che al dato empirico. E’ infatti molto più facile che uno studente sbagli la misura che scopra una nuova legge fisica. Poincaré fa un bell’esempio di questo tipo di valutazioni nel suo libro “*Scienza e ipotesi*” dal quale sono state tratte anche le altre citazioni. Racconta infatti delle molte soluzioni al problema della quadratura del cerchio sottoposte all’Accademia di Francia da sconosciuti e che venivano cestinate senza che neanche si perdesse tempo a vagliarle. Questo comportamento - commenta Poincaré - deriva dall’aver soppesato la probabilità a priori che “ci sia un pazzo più in Francia” con quella che uno sconosciuto potesse risolvere un problema sul quale avevano fallito eminenti matematici.

14.4 Scorciatoia al ragionamento bayesiano: il cane e il cacciatore

Terminiamo questa succinta introduzione all’inferenza bayesiana mostrando un ragionamento intuitivo che aiuta a capire meglio cosa si intenda per in-

²⁵H. Poincaré, “*Scienza e Ipotesi*”, molto interessante il capitolo XI sul calcolo delle probabilità.

versione di probabilità e quindi ad accettare con serenità il contributo delle probabilità a priori nelle inferenze.

Consideriamo un cacciatore che si aggira in un bosco con il suo cane sempre in continuo movimento intorno a lui. Supponiamo che la probabilità che il cane si trovi entro un raggio di 100 m dal cacciatore sia del 50%. Osserviamo il cane in una certa posizione: cosa possiamo dire sulla p[osizione dove si trova il cacciatore? Nessuno esiterà a dire che, al 50%, si troverà entro 100 m dal cane. Chiaramente il cacciatore sta per μ e il cane per l'osservazione x . Ma non è difficile convincersi che per arrivare in modo intuitivo a questo risultato, si sta tacitamente assumendo che il cacciatore possa essere, a priori, in ogni punto del bosco. Le cose cambiano se il cane sta costeggiando un fiume, se corre in una certa direzione con la preda in bocca o se è dentro un terreno recintato (ad esempio a oltre 100 metri dal filo spinato) in cui lui può entrare e il cacciatore no. Detto più chiaramente, si sta assumendo una *distribuzione iniziale uniforme* (del cacciatore nella foresta) e una verosimiglianza simmetrica. Ogni variazione da questo modello porta a conclusioni differenti. Nel seguito considereremo semplici misure di routine in cui il modello cacciatore→cane e cane→cacciatore funziona secondo l'inversione intuitiva che abbiamo descritto. Ci sono però dei problemi (specialmente in fisica di frontiera) in cui questo è tutt'altro che vero.

15 Imparare dall'esperienza

A questo punto cerchiamo di ricapitolare, senza formule, cosa abbiamo imparato:

la qualità della conoscenza di una grandezza fisica, dopo aver effettuato delle osservazioni sperimentali, dipende dalla verosimiglianza che un valore della grandezza possa aver prodotto le osservazioni e da quanto si sapeva a priori sulla grandezza fisica (prima delle nuove osservazioni).

Analizziamo i due contributi:

- la verosimiglianza descrive lo stato di conoscenza su
 - strumentazione;
 - condizione ambientali e fattori di influenza;
 - contributo dello sperimentatore
 - etc. etc. (vedi “decalogo” ISO al paragrafo 4)
- la prior sui possibili valori della grandezza fisica implica una buona conoscenza della fenomenologia sulla quale si sta investigando.

Quindi l'insegnamento di fondo di questo approccio si riconduce a quello che tutti i fisici sanno già:

per ottenere risultati scientifici di qualità è necessario avere familiarità con tutti gli aspetti sperimentali della misura e una approfondita conoscenza della fisica.

E' soltanto il bilanciamento fra questi due contributi che permette di accettare un risultato, confrontarlo con altri, ripetere le misure, calibrare la strumentazione, etc., e, in conclusione, produrre risultati utili per la comunità scientifica.

Si noti quanto sia cruciale il ruolo normativo della scommessa coerente. Essa infatti estirpa dal campo di lavoro tutte le contaminazioni dogmatiche (regola della mezza divisione, etc), responsabilizza lo sperimentatore sulle proprie affermazioni e, in ultima analisi, fa appello alla sua onestà scientifica. Se qualcuno afferma, ad esempio, che, al 68 %, un tavolo è lungo 120.00 ± 0.08 mm, deve sentirsi sicuro di questa affermazione quanto lo è della possibilità di estrarre una pallina bianca da una scatola che contiene 68 palline bianche e 32 nere. Se lo è di più (come spesso capita) o di meno vuol dire che è disonesto, nel senso che è cosciente che la sua affermazione trarrà in inganno gli eventuali utilizzatori del suo risultato.

E' interessante citare il seguente avvertimento della *Guida* ISO, in linea con quanto detto:

“Although this *Guide* provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty, and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value”.

Infine, lo slogan bayesiano “imparare dall’esperienza” ci fornisce anche delle indicazioni sulla didattica del laboratorio di Fisica. Ne parleremo brevemente nel paragrafo 24.

16 Risultati delle misure dirette in assenza di errori sistematici

Veniamo ora a come riportare i risultati delle misure dirette, ovvero misure effettuate con uno strumento opportunamente calibrato che fornisce direttamente il valore della grandezza, senza la necessità di effettuare altri conti. Questo è il livello più basso del processo di misura. Eventualmente da queste misure se ne possono ottenere altre, per via indiretta, mediante successive elaborazioni dei risultati di questo stadio.

E' noto - lo ripetiamo per fissare le idee e aggiungere delle precisazioni - che ripetendo più volte delle misure della stessa grandezza e nelle *stesse condizioni* si ottengono valori diversi a causa degli inevitabili errori casuali. Le cause di questi errori sono quelle elencate nel paragrafo 4.

16.1 Condizioni di ripetitività

Prima di procedere, è importante definire un po' meglio cosa si intende per “stesse condizioni”. Esse sono più propriamente denominate *condizioni di ripetitività* e comprendono:

- stesso procedimento di misura;
- stesso osservatore;
- stessi strumenti, utilizzati nelle stesse condizioni;
- stesso luogo;
- ripetizione delle misure in un breve periodo di tempo.

La bontà dell'accordo fra risultati di misure successive effettuate in queste condizioni è indice della *ripetitività* dei risultati.

In caso di totale assenza di errori sistematici, le letture seguono una distribuzione pressoché gaussiana intorno al valore vero. A questa distribuzione viene associata la funzione di verosimiglianza di cui si è parlato precedentemente.

Per quanto riguarda la valutazione a priori, possiamo dire che se l'esperimento è ben pianificato (scelta di strumenti, procedure, etc.), l'entità degli errori di misura è molto minore dell'incertezza con la quale è possibile conoscere il valore del misurando prima dell'operazione di misura (ad esempio, dovendo misurare l'esatta lunghezza di una barra di circa un metro, è difficile trovare delle persone per le quali l'intervallo di "accettanza del risultato" sia inferiore di una decina di centimetri, "infinitamente maggiore" di quello dovuto ai possibili errori di un normale strumento di misura²⁶). Quindi, in queste circostanze, possiamo utilizzare l'inversione di probabilità intuitiva del cane e del cacciatore, bypassando l'uso esplicito del teorema di Bayes²⁷.

Come primo passo, assumiamo di conoscere la deviazione standard che descrive l'entità delle fluttuazioni. Chiamiamola σ_r , ove r sta a ricordare "ripetitività", o anche "*random*", (*casuale*). Si noti che essa contiene l'eventuale contributo delle fluttuazioni di stima di lettura degli strumenti analogici. Altra assunzione che facciamo è che σ_r non dipenda dal valore di μ nell'intervallo di interesse (a volte succede invece che una procedura di misura abbia una precisione dipendente dal valore del misurando).

Consideriamo prima il caso semplice di una singola misura e quindi quello di molte misure ripetute.

²⁶Le prior non devono mai sparire completamente dalla mente, ma devono servire a vigilare attentamente il flusso dei dati e intervenire al minimo sospetto che qualcosa non vada!

²⁷Una prior uniforme e una verosimiglianza gaussiana producono, in virtù della (17), il seguente risultato:

$$\begin{aligned} f(\mu | x) &\propto f(x | \mu) \propto \exp \left[-\frac{(x - \mu)^2}{2 \sigma_r^2} \right] \\ &\propto \exp \left[-\frac{(\mu - x)^2}{2 \sigma_r^2} \right], \end{aligned}$$

in cui nell'ultimo passaggio sono stati invertiti μ e x , al fine di ricordare che la variabile della nuova funzione è μ e non più x (questo diventa il parametro che dà il centro della distribuzione). Ne segue che il valore vero è distribuito intorno al valore osservato secondo una gaussiana avente la stessa deviazione standard della verosimiglianza:

$$f(\mu | x) = \frac{1}{\sqrt{2 \pi} \sigma_r} \exp \left[-\frac{(\mu - x)^2}{2 \sigma_r^2} \right].$$

16.2 Singola osservazione con σ_r nota

Si osserva x . Siccome si crede che i possibili valori dell'osservabile abbiano una distribuzione di probabilità gaussiana con

$$\sigma(x) = \sigma_r$$

intorno a μ , dall'inversione di probabilità si ha che i possibili valori di μ hanno una analoga distribuzione di probabilità intorno a x con

$$\sigma(\mu) = \sigma(x) = \sigma_r.$$

Ne segue che

$$\begin{aligned} \mu &= x \pm \sigma_r && \text{al } \approx 68\% \text{ di probabilità} \\ &= x \pm 2\sigma_r && \text{al } \approx 95\% \text{ di probabilità} \\ &= \dots \end{aligned}$$

Nel seguito considereremo come standard l'incertezza data ad una deviazione standard, essa è chiamata *incertezza standard*[13].

Concludiamo questo paragrafo invitando a riflettere al significato probabilistico di espressioni del tipo $\mu = x \pm \sigma_r$ sul quale è già stato detto precedentemente e a diffidare da interpretazioni del tipo "se io ripetessi la misura un grande numero di volte, allora nel 68,3% dei casi otterrei risultati compresi nell'intervallo $x \pm \sigma_r$ ": oltre a forzature interpretative si commette un errore numerico di $\sqrt{2}$: come mai?

16.3 n osservazioni indipendenti con σ_r nota

Cominciamo con due osservazioni, ad esempio x_1 e x_2 . Innanzitutto è chiaro che, alla luce di ciascuna delle due osservazioni si avrebbe:

$$\begin{aligned} \mu|_{x_1} &= x_1 \pm \sigma_r \\ \mu|_{x_2} &= x_2 \pm \sigma_r, \end{aligned}$$

ove $\mu|_{x_i}$ sta ad indicare " μ subordinato all'osservazione di x_i " (si è preferito scrivere così, anziché μ_1 e μ_2 , per ricordare che stiamo parlando della stessa grandezza fisica). Cosa possiamo dire ora su μ alla luce di x_1 e x_2 ?

$$\mu|_{x_1 \cap x_2} = ?$$

Dal punto di vista generale si possono percorrere tre strade diverse.

1. Si può considerare il risultato $f(\mu | x_1)$ come prior antecedente l'osservazione di x_2 e inserirla nella formula dell'inferenza bayesiana. In modo simmetrico si può rovesciare l'ordine di osservazione: prima x_2 e poi x_1 . Se la procedura è corretta le conclusioni finali non devono dipendere dall'ordine.
2. In alternativa, si può considerare la verosimiglianza di osservare, dato un valore di μ , la coppia $\{x_1, x_2\}$ e applicare il teorema di Bayes all'inversione:

$$f(x_1, x_2 | \mu) \rightarrow f(\mu | x_1, x_2).$$

3. C'è infine un terzo modo di procedere, euristico, in quanto assume che la soluzione di prova sia quella buona e fornisce soltanto il valore dell'incertezza. Per semplicità illustriamo questo (per gli altri c'è poco da spiegare, si tratta solo di fare i conti ...).

Assumiamo, ragionevolmente, che il valore centrale intorno al quale è distribuito $\mu | x_1 \cap x_2$ sia la media aritmetica:

$$\bar{x} = \frac{x_1 + x_2}{2}.$$

Subordinatamente ad un certo valore di μ anche la media aritmetica \bar{x} è una variabile casuale, in quanto è funzione di variabili casuali. Anch'essa, per simmetria, ha una distribuzione di probabilità intorno a μ . In questo caso il processo di inferenza è del tipo²⁸

$$f(\bar{x} | \mu) \rightarrow f(\mu | \bar{x}).$$

La differenza rispetto al caso precedente è nella diversa deviazione standard di \bar{x} intorno a μ . Anche se per arrivare all'esatto fattore di riduzione si rimanda ad un testo di calcolo delle probabilità (vedi anche Appendice), il risultato può essere giusticato dicendo che, *essendo x_1 e x_2 indipendenti, le fluttuazioni della media tendono a compensarsi*. Quantitativamente si ottiene che

$$\sigma(\bar{x}) = \frac{\sigma(x)}{\sqrt{2}} = \frac{\sigma_r}{\sqrt{2}}.$$

Quindi, utilizzando ancora una volta il ragionamento di inversione di probabilità intuitivo, possiamo dire che

$$\sigma(\mu) = \sigma(\bar{x}) = \frac{\sigma_r}{\sqrt{2}}.$$

In effetti si può dimostrare che tutte e tre le strade conducono al medesimo risultato. Nel caso generale di n osservazioni indipendenti, si ha:

$$\sigma(\mu) = \frac{\sigma_r}{\sqrt{n}},$$

da cui:

$$\mu = \bar{x} \pm \frac{\sigma_r}{\sqrt{n}}. \quad (18)$$

Quindi, la soluzione giusta al quesito 6 del paragrafo 5 è quella con " $1/\sqrt{n}$ ". Ciò nonostante, è ragionevole preoccuparsi di un'incertezza che tende a zero all'aumentare di n . Questo induce molti a diffidare di questa formula e ignorare il \sqrt{n} . Ciò è dovuto alla mancanza di una visione globale del problema. La formula (18) è assolutamente corretta, ma non tiene conto di altre possibili cause di incertezza. Queste vanno identificate e il loro contributo va combinato insieme a σ/\sqrt{n} in maniera appropriata (vedi paragrafo 18).

²⁸Il fatto che il solo valore della media aritmetica sia in grado di produrre una inferenza statistica della stessa qualità dei singoli valori osservati è legato al concetto statistico di "sufficienza".

16.4 Caso di σ_r ignota

Il caso precedente è indubbiamente quello più semplice. Quando invece σ_r è ignota, essa va stimata dagli stessi dati, o da altre informazioni (esperimenti analoghi, esperienza di colleghi, etc.). Nel caso di assoluta ignoranza (ma è difficile trovarsi in tale stato!) è assolutamente impossibile effettuare stime dell'incertezza e quindi presentare risultati in modo scientificamente corretto. Un caso di assoluta ignoranza potrebbe essere quello, assurdo, di tentare di stimare l'incertezza partendo da un numero locato in una memoria di un computer, scritto con un grandissimo numero di cifre in quanto risultato di una operazione matematica, senza sapere a quale esperimento si riferisca, con quale strumentazione è stato ottenuto e così via.

Consideriamo ora due casi tipici che capitano in laboratorio.

16.4.1 Misure ripetute della stessa grandezza fisica

Questo caso si riconduce a quello di σ nota quando il numero di misure è abbastanza grande:

- il valore vero è in prossimità della media aritmetica, con una incertezza che decresce come $1/\sqrt{n}$;
- la deviazione standard dell'errore statistico ("valore osservato meno valore vero") può essere stimato dalla deviazione standard degli scarti dei singoli valori osservati rispetto alla media²⁹:

$$\sigma_r \approx \sigma_n(\Delta),$$

con $\Delta_i = x_i - \bar{x}$.

Quando invece n è piccolo (al di sotto della decina, tipicamente), nascono altre complicazioni, in quanto:

- innanzitutto $\sigma(\Delta)$ non è una buona stima di σ_r (si pensi al caso limite di $n = 1$);
- in secondo luogo, la verosimiglianza dipende dal parametro ignoto σ_r sul quale c'è stato di incertezza: la banale inversione di probabilità tipo cane-cacciatore non è più ovvia.

Come si può intuire, il problema diventa complicato. La soluzione usuale della statistica convenzionale consiste in

1. aumentare la deviazione standard stimata per tenere conto che la media è legata ("vincolata") ai valori stessi e che quindi la deviazione standard tende ad essere sottostimata rispetto a quanto si otterrebbe disponendo

²⁹Si ricorda che, ai fini del calcolo pratico essa è valutata come:

$$\sigma_n^2(\Delta) = \sigma_n^2(x) = \overline{x^2} - \bar{x}^2,$$

ove σ_n^2 sta ad indicare, secondo la convenzione delle calcolatrici tascabili, che la varianza è calcolata come media dei quadrati degli scarti.

di un campione più numeroso (si pensi al caso limite $n = 1$); si preferisce allora usare σ_{n-1}

$$\sigma_n \longrightarrow \sigma_{n-1} = \sqrt{\frac{n}{n-1}} \sigma_n;$$

2. cambiare il tipo di distribuzione finale: dalla gaussiana alla cosiddetta t di Student.

In realtà, anche se questi metodi vanno in qualche modo “nella direzione” giusta, bisogna fare attenzione a non prenderli troppo alla lettera. Ad esempio, se si osservano due valori che differiscono di 0.3 mm e si applica ciecamente questo metodo ne risulta un intervallo di incertezza di quasi 10 cm qualora si richiedesse un “livello di confidenza” del 99.9%. Qualsiasi meccanico troverebbe ridicola questa conclusione. Molto spesso, quando n è veramente dell’ordine dell’unità, può essere più sensato quello che si sapeva su σ_r prima della misura di quanto si possa ricavare dai dati stessi (su questo punto ritorneremo fra breve). Quando invece il problema è veramente critico è essenziale ripetere più volte le misure. Quando infine n è già dell’“ordine di 10” (ma anche 5-6 può andare abbastanza bene), l’inversione gaussiana diventa abbastanza ragionevole.

16.4.2 Singole misure di grandezze fisiche variabili (grafici)

Quando si devono fare molte misure di una grandezza in funzione di un’altra per poi riportare su un grafico l’andamento, può diventare veramente lungo e noioso ripetere molte volte le misure per ciascun punto delle ascisse. In taluni casi non è nemmeno pensabile (si immagini ad esempio la misura di una grandezza rapidamente variabile con il tempo). Questo è il caso tipico in cui non ha senso riportare delle barre di incertezza sul grafico. Al contrario, è lo stesso grafico che permette di stimare le incertezze (con alcune ipotesi e previa attenta ispezione visuale). Se, ad esempio, le incertezze sulle ascisse sono trascurabili³⁰, e quelle sulle ordinate sono (=“si pensa che siano”) costanti (o modellizzabili in qualche modo semplice) è possibile valutare σ_r delle ordinate (o “ σ_{y_r} ”, per chi preferisce ricordare nel simbolo che si tratta di ordinate) dagli scarti fra il singolo valore osservato e l’“andamento medio” dei punti (leggi: la curva “motivata fisicamente” che approssima meglio i dati sperimentali). Questa procedura si chiama *metodo dei residui* (vedi figura 5 e paragrafo 23).

Da quanto detto in questo paragrafo si capisce la critica, avanzata nel paragrafo 6.4, all’obbligo di riportare immediatamente i punti sperimentali con le barre di incertezza. Nella maggior parte dei casi di laboratorio questa regola rigida è semplicemente assurda, in quanto sono i grafici stessi che devono servire a stimare le incertezze.

³⁰In realtà questo è una condizione non necessaria, legata ad un modo semplicistico di vedere le cose: anche se ci sono incertezze sulle ascisse, queste possono essere riflesse su quelle delle ordinate e la soluzione pratica non cambia.

16.4.3 Bisogna sempre ripetere le misure? Rarità delle situazioni in cui σ_r sia completamente ignota

I casi precedentemente trattati potrebbero indurre a credere che, ogni volta che si deve determinare il valore di una certa grandezza fisica, si debba necessariamente ripetere la misura tante volte, o avere effettuato nel passato una lunga serie di misure di quel misurando in quelle condizioni, o almeno aver effettuato misure in condizioni prossime a quelle di interesse. In realtà, questo non è assolutamente necessario, come ben sanno anche meccanici, falegnami e muratori. Ognuno di loro, nel suo ambito e con i propri strumenti, è perfettamente cosciente di quanto credere al valore misurato. In effetti è veramente raro il caso in cui un professionista faccia delle misure senza avere nessuna idea dell'errore (o dell'errore percentuale) che può commettere e, quindi, della corrispondente incertezza sul valore della grandezza di interesse. L'esperienza acquisita su misure analoghe permette di valutare σ_r anche nei casi in cui si sia eseguita una sola misura. Questo valore può essere quindi utilizzato nella stima delle incertezze, come è stato sempre fatto, più o meno coscientemente, dagli sperimentatori e come raccomandato nella *Guida ISO* (vedi paragrafo 21).

Terminiamo con un'ultima raccomandazione, dedicata a tutti coloro che affermano, a ragione, che nella loro attività di ricerca “è impossibile il calcolo degli errori”. Sono perfettamente d'accordo con loro se si riferiscono al modello standard stereotipato di deviazione standard, propagazione, etc., ovvero a schematizzare il processo inferenziale nel paradigma prior-verosimiglianza-posterior. In questi casi non banali, l'*extrema ratio* è una valutazione puramente soggettiva, alla luce della regola normativa della coerenza, eseguita rispondendo a domande del tipo: “quanto credo - io - nel numero che ho ottenuto?”; “quanto è ampio l'intervallo di valori tale da essere in stato di indifferenza rispetto all'eventualità che il valore vero vi sia compreso o no?” (Questo definisce un intervallo al 50% di probabilità.). Ci si rende conto allora che in realtà, molto spesso, si sa molto di più di quello che si pensava. Questo è il motivo per cui anche quelli che “non calcolano gli errori” poi, in pratica, quando sono messi alle strette con queste domande, limitano le cifre con cui forniscono il risultato, o si esprimono sull'eventuale accordo con valori teorici o con altri risultati, sebbene i numeri sono diversi fra loro in senso matematico. Questo ultimo discorso verrà formalizzato meglio nel paragrafo 21, quando si discuterà la valutazione delle cosiddette incertezze di tipo B.

17 Propagazione delle incertezze

Il problema della propagazione delle incertezze è molto più semplice dei precedenti, almeno per quanto riguarda le questioni di principio o le stime soggettive. Infatti “c'è poco da pensare”: si fa semplicemente uso delle tecniche del calcolo delle probabilità per propagare l'incertezza su *variabili di partenza* in quella sulle variabili derivate. La propagazione entra quindi in gioco quando si effettuano, genericamente parlando, *misure indirette*³¹.

³¹Vedremo come entrano in gioco anche più per valutare effetti di errori sistematici di misure dirette.

		b		
		4	5	6
a	9	13	14	15
	10	14	15	16
	11	15	16	17

Tabella 1: Combinazione di 3 valori di a con 3 valori di b che danno luogo a 5 possibili valori della somma $c = a + b$. Se si assume l'equiprobabilità di a e di b si arriva ad una distribuzione di probabilità del tipo triangolare (discreta).

Prima di andare avanti c'è da fare una osservazione sulla notazione. Per alleggerire le formule indicheremo con x , y , etc. direttamente i valori veri delle grandezze (precedentemente indicati con μ_x , μ_y , etc.) e non più quelli osservati. Con μ_x e μ_y saranno invece indicati i valori attesi dei valori veri (in pratica le medie aritmetiche ottenute dalle misure dirette).

Per capire bene il problema, partiamo da variabili discrete. Per semplicità prendiamo due grandezze, a e b , che possono assumere soltanto tre valori, con distribuzione uniforme. Ad esempio: $a_1 = 9$, $a_2 = 10$, $a_3 = 11$; $b_1 = 4$, $b_2 = 5$, $b_3 = 6$. Essendo tutti i valori equiprobabili abbiamo: $f(a_i) = f(b_i) = 1/3$. Se adesso siamo interessati alla variabile $c = a + b$, l'incertezza sul valore di a e di b si propaga sul valore di c .

Il caso discreto con tre soli valori possibili permette di seguire il “flusso di incertezza”, come mostrato in tabella 17. La variabile c può essere un numero compreso fra 13 e 16, ma a differenza di a e di b , i valori non sono tutti equiprobabili. Infatti, mentre i valori estremi si possono verificare per una particolare coppia di a e di b , ci sono più coppie che possono produrre gli altri valori. In particolare, il valore $c = 15$ è quello più probabile semplicemente perché esso può essere ottenuto da possibili coppie.

Un caso analogo, leggermente più complicato, è mostrato in figura 8. Si tratta delle distribuzioni di probabilità della somma degli esiti di 1, 2 e 3 dadi. Si noti il graduale l'addensamento della probabilità nei valori centrali, dovuta ad un semplice effetto combinatorio. Per questo motivo la deviazione standard non cresce linearmente con l'ampiezza massima della distribuzione. Ad esempio, combinando due distribuzioni uniformi fra 0 e 1 (è il limite di un dado con infinite facce), non si ottiene $2/\sqrt{12}$, bensì $1/\sqrt{6}$, $\sqrt{2}$ volte più piccola (si riconosce la deviazione standard di una distribuzione triangolare!). Quelle che invece crescono linearmente sono le varianze ($2 \times 1/12 = 1/6$).

Si capisce inoltre come, per simmetria, la distribuzione delle differenze intorno al valore centrale debba essere uguale a quella delle somme.

Quindi, per due variabili indipendenti si ottiene la seguente regola di propagazione:

$$\sigma^2(x \pm y) = \sigma^2(x) + \sigma^2(y). \quad (19)$$

Consideriamo successivamente una trasformazione di scala: $y = cx$. Anche la scala delle possibili fluttuazioni si trasforma nello stesso modo e, poiché il segno di c è ininfluente, si ottiene:

$$\sigma(cx) = |c| \sigma(x). \quad (20)$$

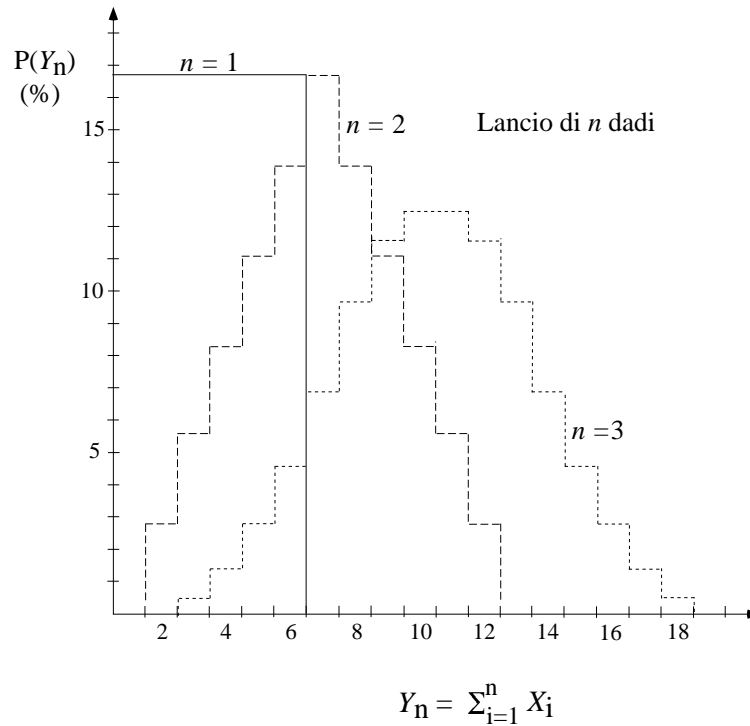


Figura 8: Distribuzione della somma dei risultati ottenuti dal lancio di n dadi. La concentrazione della probabilità al centro della distribuzione è dovuta all'elevato numero di combinazioni risultanti in valori della somma intermedi e giustifica qualitativamente il teorema del limite centrale.

Combinando i risultati espressi dalle formule (19) e (20) si ottiene la regola generale della varianza di una combinazione lineare di variabili casuali.

$$\sigma^2 \left(\sum_i c_i x_i \right) = \sum_i c_i^2 \sigma^2(x_i). \quad (21)$$

E' importante notare che questa regola dipende soltanto dalla definizione di varianza e non dal tipo di distribuzione di probabilità delle variabili casuali.

Dalla (21) si ottiene la regola generale per una funzione qualsiasi, mediante linearizzazione intorno ai valori attesi. Infatti se indichiamo con $Z = Z(X, Y)$ la generica funzione delle due variabili casuali X e Y , abbiamo

$$\begin{aligned} z = z(x, y) &\approx z(\mu_x, \mu_y) + \left. \frac{\partial z}{\partial x} \right|_{\mu_x, \mu_y} (x - \mu_x) + \left. \frac{\partial z}{\partial y} \right|_{\mu_x, \mu_y} (y - \mu_y) + \dots \\ &\approx k + \left. \frac{\partial z}{\partial x} \right|_{\mu_x, \mu_y} x + \left. \frac{\partial z}{\partial y} \right|_{\mu_x, \mu_y} y + \dots \\ \sigma^2(z) &= \left(\frac{\partial z}{\partial x} \right)^2 \sigma^2(x) + \left(\frac{\partial z}{\partial y} \right)^2 \sigma^2(y), \end{aligned} \quad (22)$$

dove k contiene tutti i termini che non dipendono dalle variabili casuali e che quindi sono ininfluenti ai fini del calcolo della varianza. L'ultimo passaggio

è stato ottenuto facendo uso della (21). E' generalmente sottointeso che le derivate vadano calcolate nel punto di migliore stima di x e di y . Il caso generale va da sé.

Si ricordi che la (22) è basata su una linearizzazione. La funzione deve essere abbastanza lineare un certo numero di deviazioni standard intorno alle migliori stime delle variabili di partenza. Questo è generalmente vero se le σ sono molto minori delle stime. Se la funzione è lineare non c'è nessun vincolo sul valore di σ . Ad esempio, se $z = x + y$, con $x = 0.1 \pm 0.7$ e $y = 0.0 \pm 1.0$, si ha $z = 0.1 \pm 1.2$.

Per quanto riguarda l'uso della formula di propagazione, si raccomanda di fare una lista dei contributi all'incertezza totale dovuti a ciascun termine da cui la grandezza finale dipende. Questo permette di capire quale contributo sia maggiormente responsabile e sul quale bisogna intervenire al momento di pianificare un nuovo esperimento. Quindi la formula (22) può essere riscritta nel seguente modo, didatticamente più valido:

$$\sigma(z) = \left| \frac{\partial z}{\partial x} \right| \sigma(x) \oplus \left| \frac{\partial z}{\partial y} \right| \sigma(y), \quad (23)$$

ove con “ \oplus ” si è indicata l'operazione di *somma in quadratura*. E' inoltre importante dare alle derivate il significato di coefficiente di sensibilità[13].

18 Come tener conto degli errori sistematici

Oltre agli errori casuali che abbiamo descritto precedentemente ogni misura può essere affetta da errori sistematici, dei quali è impossibile accorgersi operando in condizioni di ripetitività. Essi infatti si mantengono pressoché costanti durante il periodo di tempo durante il quale sono ripetute le misure e quindi non contribuiscono alla dispersione dei dati rispetto alla media.

18.1 Condizioni di riproducibilità

Anche le cause degli errori sistematici sono quelle della lista del paragrafo 4.

Per rendersi conto di eventuali errori sistematici bisogna misurare lo stesso misurando (ovvero la stessa grandezza fisica definita operativamente nello stesso modo) in *condizioni di riproducibilità*.³² Per fare questo bisogna variare:

- principio di misura;
- metodo di misura;
- osservatore;
- strumenti di misura;
- condizioni di utilizzazione degli strumenti;

³²La *Guida ISO* definisce *reproducibility* (of results of measurements) “closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement” (i risultati si intendono già corretti per eventuali errori sistematici noti).

- luogo;
- tempo di osservazione.

18.2 Correzione dei risultati per tener conto di errori sistematici noti - calibrazioni

In genere, essendo impossibile per ogni misura effettuare tutte queste variazioni ci si *calibra* su grandezze fisiche il cui valore vero è noto (entro i limiti di accuratezza necessari). Qualora si trovano differenze fra i valori ottenuti e i valori veri si studiano accuratamente tutti i contributi al raggiungimento del risultato, effettuando eventualmente calibrazioni individuali sui singoli pezzi dell'apparato. Alla fine i risultati vengono opportunamente corretti.

18.3 Incertezze dovute all'inesatta conoscenza dell'entità di un possibile errore sistematico

Nel modo semplicistico con cui è stata posta la questione sembra che gli errori sistematici non costituiscano un problema. In realtà non è sempre possibile disporre di campioni di riferimento per ciascuna delle grandezze di interesse. Ne deriva la necessità di utilizzare strumenti commerciali calibrati, anch'essi affetti da un'inevitabile incertezza di calibrazione.

Si noti comunque che, anche potendo far uso di materiale certificato, l'accuratezza non è mai assoluta ed inoltre anche il processo di calibrazione (ottenuto uguagliando due valori entro il limite di risoluzione) è soggetto ad incertezze residue.

In conclusione, bisogna abituarsi a convivere con le incertezze dovute agli errori sistematici e a tenerne conto nel corso delle misure.

18.4 Imperfetta conoscenza delle costanti di calibrazioni e dei parametri di influenza

In questa trattazione semplificata ci occuperemo di

- errori sulle costanti di calibrazioni di zero e di scala dello strumento;
- errori sulle variabili di influenza.

Anche se li abbiamo divisi per convenienza la trattazione generale è assolutamente analoga.

18.4.1 Errore di zero (offset)

Cominciamo con il considerare uno strumento affetto da un possibile errore di zero (ove intendiamo una possibile costante additiva, in inglese *offset*, da aggiungere ai valori letti). Questo caso è illustrato in figura 9.a, dove è mostrata la lettura sullo strumento ("risposta") in funzione del valore della grandezza ("stimolo").

L'effetto di questo tipo di errore è che la curva di risposta potrebbe essere traslata lungo le ordinate, ma di un valore incognito. Infatti, anche se uno strumento è stato calibrato, e quindi crediamo che il valore più plausibile

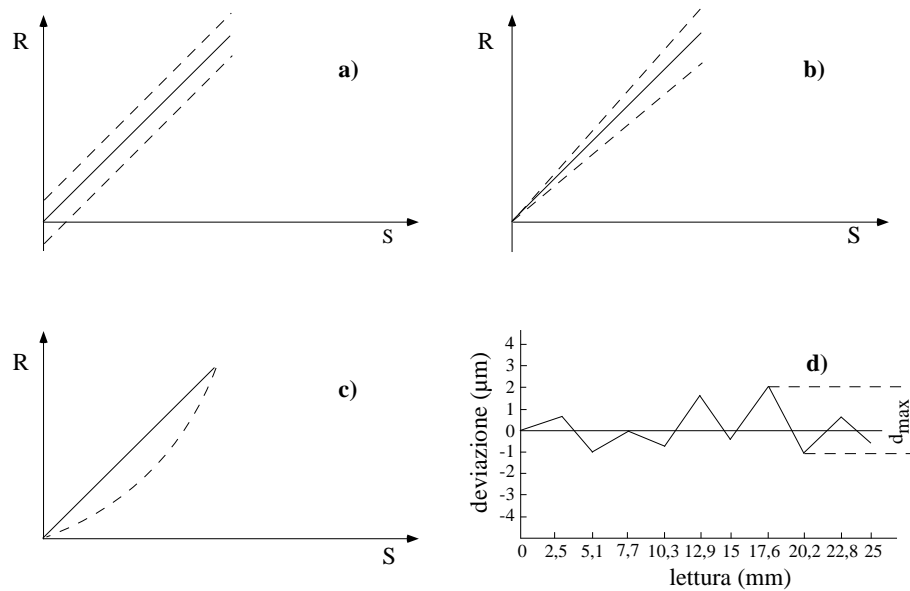


Figura 9: Alcuni tipi di errori sistematici di strumenti: a) errore di zero; b) errore di scala; c) deviazione dalla linearità; d) caso di più difficile modellizzazione (i punti lungo l'ascissa mostrano un esempio di accertamento delle deviazioni). R sta per risposta e S per stimolo.

dell'offset sia 0, la non esattezza della calibratura ci fa ritenere che anche “piccoli” valori intorno a 0 siano possibili.

Assumiamo di potere descrivere i diversi gradi di fiducia dell'offset con una gaussiana³³ centrata in 0 e di deviazione standard σ_z . Chiamiamo inoltre z la variabile casuale ad esso associata (z sta per “zero vero”). Abbiamo detto che, in assenza di errore sistematico, la grandezza ha una incertezza $\sigma(\mu) = \sigma_r/\sqrt{n}$. Facciamo ora questa trasformazione di notazione

$$\mu \longrightarrow \mu_r \quad (24)$$

$$\sigma(\mu) \longrightarrow \sigma(\mu_r) \quad (25)$$

ovvero aggiungiamo il pedice r per indicare che questi sono i valori veri ottenuti tenendo conto dei soli effetti casuali. Per ottenere il valore che tenga conto anche della non perfetta calibratura bisogna sottrarre a μ_r il valor vero dell'offset:

$$\mu = \mu_r - z,$$

ove $z = 0 \pm \sigma_z$. Dalla propagazione delle incertezze otteniamo:

$$\sigma^2(\mu) = \sigma^2(\mu_r) + \sigma_z^2. \quad (26)$$

L'incertezza standard globale è quindi ottenuta combinando in quadratura l'incertezza standard dovuta ai soli effetti casuali con quella dell'offset.

³³In realtà non c'è bisogno che la distribuzione sia normale, in quanto faremo uso soltanto delle proprietà generali della varianza. Anche dal punto di vista pratico, è più frequente il caso di una distribuzione uniforme o triangolare.

Quando si hanno più grandezze misurate con lo stesso strumento si applica la stessa procedura a tutte le grandezze:

$$\sigma^2(\mu_i) = \sigma^2(\mu_{r_i}) + \sigma_z^2.$$

Sorge ora il problema che tutte le incertezze sono correlate: se il valore vero dello zero dovesse valere $Z = z$ tutte le misure sarebbero sbagliate di questo valore. Per studiare l'effetto delle correlazioni, immaginiamo due grandezze, μ_1 e μ_2 e facciamone la differenza e la somma: $D = \mu_1 - \mu_2$, $S = \mu_1 + \mu_2$. Applicando la propagazione delle incertezze che abbiamo visto, semberebbe che

$$\begin{aligned}\sigma^2(D) &\stackrel{?}{=} \sigma^2(\mu_1) + \sigma^2(\mu_2) = \sigma^2(\mu_{r_1}) + \sigma^2(\mu_{r_2}) + 2\sigma_z \\ \sigma^2(S) &\stackrel{?}{=} \sigma^2(\mu_1) + \sigma^2(\mu_2) = \sigma^2(\mu_{r_1}) + \sigma^2(\mu_{r_2}) + 2\sigma_z\end{aligned}$$

Ma c'è qualcosa che non convince; infatti, intuitivamente ci si aspetta che una incertezza dovuta all'offset debba essere ininfluenta sulle differenze: un termometro può anche essere scalibrato di 10°C , ma questo non può influenzare una misura di differenza di temperature. Riscriviamo allora i valori veri a partire da quelli "solo random" più l'effetto dell'offset:

$$\begin{aligned}\mu_1 &= \mu_{r_1} - z \\ \mu_2 &= \mu_{r_2} - z \\ D &= \mu_1 - \mu_2 = \mu_{r_1} - \mu_{r_2} \\ S &= \mu_1 + \mu_2 = \mu_{r_1} + \mu_{r_2} - 2z.\end{aligned}$$

Ne segue quindi

$$\sigma^2(D) = \sigma^2(\mu_{r_1}) + \sigma^2(\mu_{r_2}) \quad (27)$$

$$\sigma^2(S) = \sigma^2(\mu_{r_1}) + \sigma^2(\mu_{r_2}) + 4\sigma_z^2 \quad (28)$$

ben diverse da quelle precedenti e molto più ragionevoli.

18.4.2 Errore di scala

Per il caso di fattore di scala si procede nello stesso modo. In questo caso chiamiamo f il valore vero della costante di scala. Se lo strumento è calibrato al meglio, si crede che

$$f = 1 \pm \sigma_f.$$

Ne segue che

$$\mu = f \mu_r$$

e quindi

$$\sigma^2(\mu) = \sigma^2(\mu_r) + (\mu_r \sigma_f)^2 \quad (29)$$

Anche in questo caso l'incertezza globale è una combinazione quadratica di quella dovuta ai soli effetti casuali e di quella dovuta all'incertezza di calibrazione, la quale è proporzionale al valore del misurando. Come nel caso precedente, più misure effettuate con lo stesso strumento sono correlate. Se

in successive elaborazioni di queste si utilizzassero semplicemente le incertezze globali nella formula di propagazione, senza tener conto degli effetti di correlazione, si otterrebbero risultati errati.

E' istruttivo, nel caso dell'incertezza di scala, considerare il prodotto $P = \mu_1 \cdot \mu_2$, il rapporto $R = \mu_1/\mu_2$ e la differenza $D = \mu_1 - \mu_2$. Lasciando per esercizio i risultati errati che si otterrebbero con il procedimento naïve, ricaviamo i risultati corretti, riscrivendo prima

$$\begin{aligned} P &= \mu_1 \cdot \mu_2 = f^2 \mu_{r1} \mu_{r2} \\ R &= \frac{\mu_1}{\mu_2} = \frac{\mu_{r1}}{\mu_{r2}} \\ D &= \mu_1 - \mu_2 = f (\mu_{r1} - \mu_{r2}). \end{aligned}$$

Si vede subito che l'eventuale errore di scala è ininfluenza nel rapporto; nel prodotto viene amplificato; nelle differenze esso dipende dalla differenza stessa e diventa trascurabile per valori molto vicini. Passando alle incertezze, abbiamo:

$$\sigma^2(P) = \mu_2^2 \sigma_r^2(\mu_1) + \mu_1^2 \sigma_r^2(\mu_2) + (2 \sigma_f \mu_1 \mu_2)^2 \quad (30)$$

$$\sigma^2(R) = \frac{\sigma_r^2(\mu_1)}{\mu_2^2} + \sigma_r^2(\mu_2) \frac{\mu_1^2}{\mu_2^4} \quad (31)$$

$$\sigma^2(D) = \sigma_r^2(\mu_1) + \sigma_r^2(\mu_2) + \sigma_f^2 D^2. \quad (32)$$

18.4.3 Importanza delle misure per differenza

Questi risultati, insieme a quanto imparato nel caso dell'errore di zero, ci insegnano che, nel caso si sia interessati alle differenze, gli eventuali errori di calibrazione dello strumento tendono a semplificarsi e quindi ad essere irrilevanti ai fini del risultato finale. Questo è particolarmente vero quando i valori sono molto vicini.

Il caso delle differenze è importante perché molto spesso sono queste da cui dipendono i risultati finali. In particolare:

- tutte le misure di *termologia* dipendono da differenze di temperatura;
- ogni volta che si è interessati a “*coefficienti angolari*” (derivata di una grandezza in funzione di un'altra) si misurano, sostanzialmente, rapporti fra differenze,

Infine, da questo risultato si impara che, quando è possibile, è preferibile *misurare per differenza*. Questo vale, ad esempio, per:

- misura di massa mediante bilancia: controllare l'indicazione a vuoto; aggiungere artificialmente tare aventi masse diverse per effettuare la misura in diverse zone all'interno della portata, mediando i risultati qualora differiscano fra di loro al fine di semplificare eventuali deviazioni dalla linearità della risposta;
- misura di volumi di liquidi mediante cilindri graduati: fare misure per differenza svuotando parzialmente il recipiente e facendo attenzione a tener conto del menisco del liquido sempre nello stesso modo;

- anche nella misura di lunghezze mediante righelli da disegno è raccomandabile eseguire misure partendo anche da tacche diverse da quelle dello 0, se è richiesta una elevata accuratezza del risultato.

Suggerimenti analoghi valgono anche per misure di differenze di potenziale e simili.

18.5 Casi di errore di più difficile schematizzazione

Per terminare, consideriamo il caso d) di figura 9. Essa indica che le deviazioni della lettura dal valore vero non sono schematizzabili né con un semplice errore di zero né con uno di scala. Il caso riportato è quello di un calibro Palmer, secondo le norme DIN 863. La figura sta ad indicare che il costruttore garantisce che, se lo strumento è usato nel modo corretto e alla temperatura di riferimento, la differenza fra indicazione e valore vero non eccederà i limiti indicati, in tutto l'intervallo di misura dello strumento. La curva spezzata va intesa soltanto in modo illustrativo dell'irregolarità delle deviazioni (a meno che non sia verificata con blocchetti di calibrazione, ma in questo caso le deviazioni andrebbero utilizzate per correggere i dati e non per la valutazione dell'incertezza). Uno strumento che si comporta più o meno in questo modo è il cosiddetto "strumento universale analogico" (tipo tester ICE ancora utilizzato nei laboratori di esperienze didattiche).

Chiaramente, la modellizzazione di questo caso è più problematica, specialmente per quanto riguarda le correlazioni. Le indicazioni di massima che si possono suggerire sono le seguenti:

- per quanto riguarda una singola misura, l'incertezza standard da associare ad una possibile deviazione sistematica è pari alla deviazione standard di una distribuzione uniforme nell'intervallo di tolleranza ($\sigma = \text{larghezza dell'intervallo} / \sqrt{12}$);
- due letture abbastanza ravvicinate sono fortemente correlate e quindi gli effetti sistematici possono essere ignorati se si è interessati alle loro differenze;
- due letture molto distanti possono essere considerate scorrelate;
- se si effettua una serie di misure in funzione di un'altra, al fine di studiarne l'andamento, queste piccole deviazioni possono essere considerate, al fine dell'analisi, come contributi casuali (vedi paragrafo 16.4.2 e primo punto del paragrafo 23.4);
- volendo sfruttare le potenzialità di precisione dello strumento, avviando ai limiti di accuratezza, è necessario eseguire delle calibrazioni assolute su valori di riferimento, o almeno calibrazioni relative con qualche accorgimento³⁴.

³⁴ Ad esempio un voltmetro può essere calibrato in modo relativo (rispetto al valore di fondo scala) utilizzando un partitore di precisione: ogni deviazione dalla linearità sarà imputato al comportamento del voltmetro e la lettura potrà essere corretta. Per fare un buon partitore economico è sufficiente prendere una ventina di resistori all'1% tutti uguali e presi nuovi dalla stessa striscia con cui sono confezionati. Le variazioni relative di resistenza sono ben inferiori all'1% e la loro combinazione riduce ancora di più le incertezze relative.

18.6 Incertezza su un fattore di influenza

Affrontiamo questo argomento con un caso specifico. Supponiamo di dover determinare la lunghezza di una barretta di materiale avente un alto coefficiente di dilatazione termica. Affinché la misura abbia senso, occorre specificare nella definizione del misurando anche il valore della temperatura di riferimento. In genere si utilizza come riferimento $T_o = 20^\circ\text{C}$. Questo non implica che bisogna lavorare esattamente a questa temperatura, essendo ciò impossibile in linea di principio. Il valore di temperatura può essere, al più, nominalmente 20°C , con una incertezza di σ_T ; oppure la misura viene eseguita ad una temperatura $T \pm \sigma_T$ diversa da T_o (ad esempio se non è possibile trasportare il misurando in un piccolo ambiente termostato).

In questi casi è possibile sia

- correggere il valore misurato per effetto della differenza di temperatura;
- associare al risultato un nuovo contributo all'incertezza globale, dovuto all'eventuale incertezza sulla temperatura;

a condizione che sia noto il coefficiente di dilatazione termica

$$\alpha \pm \sigma_\alpha$$

(noto anch'esso con una eventuale incertezza ...).

Indicando con l la lunghezza e $\Delta T = T - T_o$, considerando $\alpha \ll 1$ e assumendo (per semplicità) che lo strumento di misura non subisca la dilatazione termica, si ha (vedi anche figura 2):

$$\begin{aligned} l_T &= l_{T_o}(1 + \alpha \Delta T) \\ l_{T_o} &= \frac{l_T}{(1 + \alpha \Delta T)} \approx l_T(1 - \alpha \Delta T) = \beta l_T \end{aligned} \quad (33)$$

$$\begin{aligned} \sigma^2(l_{T_o}) &= \beta^2 \sigma^2(l_T) + l_T^2 \sigma^2(\beta) \\ &= (1 - \alpha \Delta T)^2 \sigma^2(l_T) + l_T^2 (\Delta T^2 \sigma_\alpha^2 + \alpha^2 \sigma_T^2). \end{aligned} \quad (34)$$

Si noti che:

- anche se $\Delta T = 0$ ($T = T_o$) resta sempre una incertezza dovuta alla non perfetta conoscenza della temperatura;
- se il coefficiente di dilatazione non è ben noto ($\sigma(\alpha)$ “grande”) è raccomandabile cercare di effettuare la misura ad una temperatura molto prossima a quella di riferimento;
- se invece α è “perfettamente noto” il risultato non dipende dalla temperatura alla quale si effettuano le misure.
- se si eseguono più misure a temperature diverse da T_o (anche diverse fra loro), le loro incertezze saranno correlate a causa del comune coefficiente di dilatazione usato per correggere i risultati; nel caso di analisi successive se ne può tenere conto come visto per costanti di calibrazione.

L'esempio mostrato si estende in modo immediato ad altri fattori di influenza, eventualmente anche concomitanti.

18.7 Propagazione senza derivate

Come inciso facciamo notare che è sempre possibile propagare le incertezze linearizzando la dipendenza fra valore della grandezza di interesse e quella di una variabile di influenza mediante sviluppo in serie analitico. Molto spesso può essere più comodo e rapido (o addirittura sarebbe impossibile altrimenti) studiare come varia il valore della grandezza apportando variazioni di $\pm 1 \sigma$ alla variabile di influenza. Se la dipendenza è circa lineare, le variazioni risultanti corrispondono circa alla deviazione standard della grandezza di interesse. Chiaramente, il concetto è lo stesso. Si sta soltanto facendo una derivata con metodo numerico. Va anche da sé che è importante controllare che l'andamento sia circa lineare. In genere un controllo rapido è consistito nel verificare che le deviazioni per $+\sigma$ e $-\sigma$ siano della stessa entità.

18.8 Calibrazione, intercalibrazione e “randomizzazione”

Anche se abbiamo visto come gestire le incertezze che derivano da errori sistematici tipici, è preferibile, nei limiti del possibile, lavorare con strumenti calibrati al meglio, ovvero con possibili errori residui di calibrazione dell'ordine di grandezza (o minori) delle deviazioni standard di ripetibilità σ_r .

Va da sé che strumenti ben calibrati significa strumenti costosi, e non è sempre possibile avere a disposizione tali strumenti per esperienze didattiche (e nemmeno sempre per la ricerca). E' allora opportuno imparare ad operare mediante *intercalibrazioni*. Questa procedura, unita a quelle delle misure per differenza, permette di minimizzare gli errori sistematici strumentali e di semplificare le elaborazioni dei dati. Da quanto detto seguono i seguenti suggerimenti:

1. è bene avere almeno uno strumento di riferimento costoso ad alta “precisione” (in realtà bisognerebbe dire ad alta *accuratezza*): ricalibrando su di esso gli strumenti economici è equivalente ad avere tanti strumenti costosi;
2. quando si devono misurare contemporaneamente diverse grandezze omogenee, è opportuno almeno intercalibrare gli strumenti.

Accenniamo ancora ad un ultimo metodo per ridurre, all'occorrenza, i possibili errori sistematici degli strumenti, qualora non sia possibile effettuare una calibrazione assoluta per diversi valori della grandezza e a; tresì non sia possibile lavorare per differenze, oppure ciò sarebbe possibile ma l'errore dello strumento non è di facile modellizzazione (tipo il caso d) di figura 9).

In questi casi si possono ripetere le misure con diversi strumenti, preventivamente intercalibrati per un valore centrale delle letture. Mediando i risultati si ottiene una compensazione dei diversi fattori casuali di ciascuno strumento (casuale nel solito senso di incerto: anche se la curva di risposta dello strumento rimane fissa, per ciascuno di essi non si sa quale essa sia). Questa tecnica può essere chiamata di *randomizzazione*.

19 Coefficiente di correlazione

A questo punto, avendo introdotto nel paragrafo 18.4 il concetto di correlazione fra valori di grandezze fisiche che hanno incertezze comuni e avendo mostrato la loro importanza nelle inferenze successive, è d'obbligo un brevissimo accenno al *coefficiente di correlazione*. Esso è indicato con $\rho(\mu_1, \mu_2)$, può avere valori compresi fra -1 e 1 ed è atto a misurare il grado di correlazione (lineare³⁵) fra le due grandezze. Il concetto intuitivo (nell'applicazione alle misure) è il seguente

- $\rho(\mu_1, \mu_2) > 0$ sta ad indicare che, se il valore vero μ_1 è maggiore di quello stimato, anche il valore vero μ_2 è maggiore del corrispondente valore vero;
- se $\rho(\mu_1, \mu_2) < 0$, ad una sovrastima di una grandezza corrisponde una sottostima dell'altra;
- se $\rho(\mu_1, \mu_2) = 0$, sovrastime e sottostime dei due valori sono indipendenti.

E' chiaro quindi che, nel caso di errori di zero e di scala, il coefficiente di correlazione fra due grandezze misurate direttamente con lo stesso strumento può essere soltanto ≥ 0 : tutti i valori saranno eventualmente mal determinati nello stesso verso.

Per quanto riguarda l'entità delle possibili sovrastime e sottostime, si può dimostrare che esse sono misurate in termini della deviazione standard:

se μ_1 è sovrastimato di k volte $\sigma(\mu_1)$, allora μ_2 è sovrastimato di k volte $\rho(\mu_1, \mu_2) \sigma(\mu_2)$.

Spesso si fa uso anche di un'altra grandezza per quantificare le correlazioni, sebbene in modo molto meno immediatamente percepibile. Essa è la *covarianza*, indicata con $\text{Cov}(\mu_1, \mu_2)$ e legata al coefficiente di correlazione da

$$\text{Cov}(\mu_1, \mu_2) = \rho(\mu_1, \mu_2) \sigma(\mu_1) \sigma(\mu_2) . \quad (35)$$

Si noti come la covarianza abbia dimensioni che sono il prodotto delle dimensioni delle due grandezze. Per questo è difficile dal suo valore farsi un'idea intuitiva dell'entità delle correlazioni.

19.1 Valutazione pratica di ρ dovuto ad errori di calibrazione

Per ottenere in modo euristico una regoletta pratica per il calcolo del coefficiente di correlazione dovuto ad errori di calibrazione di zero o di scala si può ragionare nel seguente modo:

- per quanto detto precedentemente il coefficiente di correlazione è ≥ 0 ;
- se lo strumento è perfettamente calibrato, non si introduce alcuna correlazione fra le misure e, nel contempo, è nullo il contributo all'incertezza totale;

³⁵Questo è un punto importante, ma sul quale purtroppo non possiamo entrare in dettaglio.

- se l'incertezza dovuta alla calibrazione è molto maggiore di quella dovuta agli effetti casuali, essa domina l'incertezza globale e quindi la correlazione sarà di +1 (“100 %”).

Il coefficiente di correlazione calcolato come

$$\rho = \frac{\text{prodotto delle incertezze "comuni"}}{\text{prodotto delle incertezze totali}}$$

soddisfa questi requisiti (“comuni” sta per “introdotte dall’effetto sistematico comune”). In effetti, si può dimostrare che, per i casi che ci interessano, la formula è corretta.

Otteniamo quindi, per i due tipi di errori considerati:

errore di zero

$$\rho(\mu_1, \mu_2) = \frac{\sigma_z^2}{\sqrt{(\sigma_r^2(\mu_1) + \sigma_z^2) (\sigma_r^2(\mu_2) + \sigma_z^2)}}. \quad (36)$$

errore di scala

$$\rho(\mu_1, \mu_2) = \frac{\mu_1 \mu_2 \sigma_f^2}{\sqrt{(\sigma_r^2(\mu_1) + \mu_1^2 \sigma_f^2) (\sigma_r^2(\mu_2) + \mu_2^2 \sigma_f^2)}}. \quad (37)$$

19.2 Propagazione di varianze e covarianze

Le (36) e (37) permettono di calcolare i coefficienti di correlazione e le covarianze fra grandezze misurate con lo stesso strumento. Ma le correlazioni sono generate anche per altre cause. Infatti, vale in generale:

più grandezze ottenute per misura indiretta a partire da un insieme comune di grandezze di base, ciascuna nota con una sua incertezza, hanno in genere valori correlati.

Le grandezze misurate con lo stesso strumento sono soltanto un caso particolare in cui le stesse informazioni comuni sono le costanti di calibrazione. Altri casi importanti sono:

- più grandezze misurate indirettamente da un insieme di grandezze misurate direttamente;
- i parametri delle curve che descrivono una legge fisica (ad esempio intercetta e coefficiente angolare di un fit lineare).

Per comodità diamo la regola generale di propagazione che tiene conto delle correlazioni fra *variabili di partenza* e fornisce anche le correlazioni fra le *variabili di arrivo*, senza voler dare una dimostrazione o giustificazione. Per semplificare le formule usiamo la seguente notazione:

\mathbf{x}_i sono le n variabili di partenza (corrispondenti alle μ_{X_i} nel formalismo usato finora);

\mathbf{y}_k sono le m variabili di arrivo (corrispondenti a μ_{Y_k});

$\sigma_{\mathbf{x}_{ij}}$: è una notazione compatta che riassume varianze e covarianze:

- se $i = j$ esse sono le n varianze delle x_i :

$$\sigma_{x_{ii}} = \sigma_{x_i}^2 = \sigma^2(X_i).$$

- se $i \neq j$ esse sono le covarianze:

$$\sigma_{x_{ij}} = \text{Cov}(X_i, X_j),$$

- essendo, per ovvi motivi, $\rho(X_i, X_j) = \rho(X_j, X_i)$, ne segue che anche $\text{Cov}(X_i, X_j)$ è uguale a $\text{Cov}(X_j, X_i)$ e quindi $\sigma_{x_{ij}} = \sigma_{x_{ji}}$.

$\sigma_{\mathbf{y}_{kl}}$ sono le analoghe delle $\sigma_{x_{ij}}$ per le Y_k ;

$\mathbf{Y}_k = \mathbf{Y}_k(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ sono le funzioni che legano le Y_k alle X_i ;

le **derivate** si intendono calcolate in corrispondenza dei valori attesi (“migliori stime”) delle X_i .

Con questa notazione la formula compatta di propagazione di varianze e covarianze è:

$$\sigma_{y_{kl}} = \sum_{i,j} \frac{\partial y_k}{\partial x_i} \frac{\partial y_l}{\partial x_j} \sigma_{x_{ij}}, \quad (38)$$

dove la sommatoria si estende alle $n \times n$ combinazioni degli indici i e j . Gli indici k e l variano invece fra 1 e m .

Come esempio, ricaviamo mediante la (38) l'incertezza della somma e della differenza di due grandezze affette da errore di zero comune facendo uso della (29) e della (36). Inoltre possiamo calcolare la covarianza fra S e D (e da questa, banalmente, il coefficiente di correlazione):

$$\begin{aligned} X_1 &= \mu_1 \text{ (vecchia notazione)} \\ X_2 &= \mu_2 \text{ (vecchia notazione)} \\ Y_1 &= S = X_1 + X_2 \\ Y_2 &= D = X_1 - X_2 \end{aligned}$$

$$\begin{aligned} \sigma^2(S) (= \sigma_{y_{1,1}}) &= (+1) \cdot \sigma_{x_{1,1}} + (+1) \cdot \sigma_{x_{1,2}} + (+1) \cdot \sigma_{x_{2,1}} + (+1) \cdot \sigma_{x_{2,2}} \\ &= (\sigma^2(\mu_{r_1}) + \sigma_z^2) + \sigma_z^2 + \sigma_z^2 + (\sigma^2(\mu_{r_2}) + \sigma_z^2) \\ &= \sigma^2(\mu_{r_1}) + \sigma^2(\mu_{r_2}) + 4\sigma_z^2 \end{aligned}$$

$$\begin{aligned} \sigma^2(D) (= \sigma_{y_{2,2}}) &= (+1) \cdot \sigma_{x_{1,1}} + (-1) \cdot \sigma_{x_{1,2}} + (-1) \cdot \sigma_{x_{2,1}} + (+1) \cdot \sigma_{x_{2,2}} \\ &= \sigma^2(\mu_{r_1}) + \sigma^2(\mu_{r_2}) \end{aligned}$$

$$\begin{aligned} \text{Cov}(S, D) (= \sigma_{y_{1,2}}) &= (+1) \cdot (\sigma^2(\mu_{r_1}) + \sigma_z^2) + (-1) \cdot \sigma_z^2 + \\ &\quad (+1) \cdot \sigma_z^2 + (-1) \cdot (\sigma^2(\mu_{r_2}) + \sigma_z^2) \\ &= \sigma^2(\mu_{r_1}) - \sigma^2(\mu_{r_2}). \end{aligned}$$

Abbiamo riottenuto $\sigma(S)$ e $\sigma(D)$ che conoscevamo (vedi (27) e (28)).

Per quanto riguarda la covarianza, si noti come il risultato non sia affatto intuitivo. Questo dovrebbe insegnare che, quando i problemi diventano importanti e complicati, bisogna fare molta attenzione agli effetti di correlazione.

20 Raccomandazioni BIPM/ISO

Confrontiamo ora i risultati ottenuti con quanto raccomandato dal BIPM e dall'ISO.

1. *The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:*

A: *those which are evaluated by statistical methods;*

B: *those which are evaluated by other means.*

There is not always a simple correspondence between the classification into categories A or B and the previously used classification into “random” and “systematic” uncertainties. The term “systematic uncertainty” can be misleading and should be avoided.

The detailed report of the uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical result.

La prima raccomandazione afferma, sostanzialmente, che tutte le incertezze possono essere trattate probabilisticamente. La distinzione fra tipo A e tipo B è sottile e potrebbe essere fuorviante per chi pensasse che “metodo statistico” sia sinonimo di “metodo probabilistico”, perché allora sembrerebbe che le incertezze di tipo B non siano probabilistiche. In realtà, qui “statistico” è legato a “dati statistici”, ovvero a osservazioni ottenute ripetendo l'esperimento. Possiamo dire che, a grandi linee e per molte applicazioni, le incertezze di tipo A siano quelle dovute ad errori casuali e quelle di tipo B ad errori sistematici. Ma questo non è sempre vero. Questa è la ragione della strana nomenclatura.

Come esempio di un caso di incertezza di tipo B dovuta ad errori statistici, si pensi al seguente esperimento: eseguo molte misure per tre valori di pH (intorno a 2, 4 e 6) e trovo tre deviazioni standard fra loro compatibili, $\sigma_r = 0.15$; successivamente eseguo una sola misura con lo stesso strumento, nelle stesse condizioni, di un'altra soluzione e ottengo 5.30. L'incertezza dovuta ad errori statistici è chiaramente 0.15, ma essa è di tipo B in quanto proviene da misure su un altro misurando.

2. *The components in category A are characterized by the estimated variances s_i^2 (or the estimated “standard deviations” s_i) and the number of degrees of freedom ν_i . Where appropriate, the covariances should be given.*

Le s_i di cui si parla corrispondono alle stime delle σ_r ottenute mediante misure ripetute sul misurando di interesse (e non note da conoscenze

precedenti). Il numero di gradi di libertà è legato al problema dei piccoli campioni e all'uso della t di Student. Come discusso nel paragrafo 16.4.1, questo è un problema delicato che non andrebbe trattato meccanicamente. In ogni caso, fornire la numerosità del campione utilizzato è sicuramente una buona pratica. Quando non lo si fornisce vuol dire che è "grande". Infine, quando si eseguono simultaneamente coppie (almeno) di misure va anche fornita la stima della varianza (o del coefficiente di correlazione).

3. *The components in category B should be characterized by quantities u_j^2 , which may be considered as approximations to the corresponding variances, the existence of which is assumed. The quantities u_j^2 may be treated like variances and the quantities u_j like standard deviations. Where appropriate, the covariances should be treated in a similar way.*

Chiaramente, questa norma ha senso solo se la probabilità è intesa in senso soggettivo. Vedremo nel seguito come valutare questo contributo. Anche queste incertezze possono essere correlate. Ad esempio se una misura dipende dalla curva di calibrazione di un certo strumento, ricavata da misure precedenti, l'incertezza dipende dall'incertezza sui parametri della curva, i quali sono in genere correlati.

4. *The combined uncertainty should be characterized by the numerical value obtained by applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of "standard deviations".*

Si tratta della propagazione delle incertezze di cui abbiamo parlato. Anche in questo caso è importante il punto di vista secondo il quale anche i valori veri sono variabili casuali.

5. *If, for particular applications, it is necessary to multiply the combined uncertainty by a factor to obtain an overall uncertainty, the multiplying factor used must always be stated.*

L'ultima raccomandazione afferma ancora una volta che l'incertezza dovrebbe essere per convenzione la deviazione standard della distribuzione di probabilità del valore vero. Eventualmente si possono usare altre definizioni di incertezza, ad esempio un intervallo di credibilità del 90, 95 o 99 %, purché venga esplicitamente dichiarato. Ad esempio, sotto ipotesi di distribuzione gaussiana del valore vero intorno al valore stimato questi intervalli possono essere calcolati dalle tabelle della distribuzione normale.

Riassumendo, i presupposti teorici che sono alla base delle raccomandazioni BIPM/ISO sono:

Probabilità soggettiva: permette di parlare di probabilità, e quindi di deviazione standard, dei valori di qualsiasi grandezza fisica; essa è, in particolare, cruciale per:

- incertezze di tipo B;

- propagazione delle incertezze;
- interpretazione probabilistica del risultato.

Incerteza come deviazione standard:

- è “standard”;
- le regole di propagazione si applicano alle varianze e non agli intervalli di credibilità.

Incerteza standard combinata: è ottenuta dalla formula di propagazione delle varianze e fa uso di varianze, covarianze (o coefficienti di correlazione) e derivate prime; si ricordi sempre che si tratta di una formula approssimativa che assume un andamento circa lineare intorno ai valori attesi delle grandezze;

Teorema del limite centrale: fa sì che, sotto opportune condizioni, il valore vero sia distribuito normalmente anche quando sia ottenuto da grandezze la cui incerteza non è descritta da un modello gaussiano.

21 Valutazione delle incertezze di tipo B

Le incertezze di tipo *B* sono indubbiamente quelle più critiche da valutare. Vediamo innanzitutto cosa raccomanda la *Guida*, poi mostriamo degli esempi.

For estimate x_i of an input quantity³⁶ X_i that has not been obtained from repeated observations, the . . . standard uncertainty u_i is evaluated by scientific judgement based on all the available information on the possible variability of X_i . The pool of information may include

- *previous measurement data;*
- *experience with or general knowledge of the behaviour and properties of relevant materials and instruments;*
- *manufacturer’s specifications;*
- *data provided in calibration and other certificates;*
- *uncertainties assigned to reference data taken from handbooks.*

Detto nel linguaggio probabilistico adottato, si cerca di modellizzare, in base alla migliore conoscenza del problema, la distribuzione dei gradi di fiducia (considerando, ad esempio, gli estremi dell’intervallo di valori possibili, se ci sono valori più credibili di altri, e così via) e successivamente se ne ricava la deviazione standard. Anche se la modellizzazione è rozza e la deviazione standard che ne deriva è incerta, è importante notare che:

- le deviazioni standard valutate da piccoli campioni di dati sperimentali non sono meno incerte di quelle ottenute “by scientific judgement”; ad esempio in un modello gaussiano l’incerteza relativa sulla valutazione è approssimativamente uguale a $1/\sqrt{2(n-1)}$. Per $n = 5, 10$ e 20 essa è $\approx 35, 24$ e 16 %;

³⁶Per “grandezza d’ingresso” la *Guida* ISO intende tutte le grandezze che contribuiscono alla valutazione del valore della grandezza di interesse (costanti di calibrazione, parametri di influenza, valori tabulati, risultati di esperimenti precedenti, etc.).

- è preferibile dare le stime più verosimili delle deviazioni standard, invece di sovrastimare per motivi di prudenza: se si danno sempre le migliori stime, le incertezze globali saranno mediamente corrette; se si sovrastima ad ogni passo, saranno mediamente esagerate anche le incertezze globali.
- il modello finale di distribuzione del valore vero è reso circa normale dalla combinazione delle varianze (teorema del limite centrale) indipendentemente dalla forma esatta della distribuzione assunta.

Facciamo degli esempi.

1. Misure di altre grandezze particolari, prossime a quella di interesse ed eseguite nelle stesse condizioni, hanno fornito una deviazione standard di ripetitività σ_r . E' ragionevole assumere per l'incertezza

$$u = \sigma_r .$$

2. Il certificato di calibrazione di un costruttore dichiara che l'incertezza, definita come k deviazioni standard, è " $\pm\Delta$ ":

$$u = \frac{\Delta}{k} .$$

3. Un ricercatore afferma che una certa grandezza vale, al 90%, $\dots \pm \Delta$. Assumendo, ragionevolmente, un modello gaussiano:

$$u = \frac{\Delta}{1.64} .$$

4. Una pubblicazione riporta un risultato come $\bar{x} \pm \Delta$, specificando che la media è stata eseguita con 4 valori (3 gradi di libertà) e che l'incertezza è data al 95%. Si deduce che, verosimilmente, l'intervallo è stato calcolato mediante la t di Student. Ne segue (consultando opportune tabelle):

$$u = \frac{\Delta}{3.18} .$$

5. Un manuale di istruzione dichiara che l'errore massimo che lo strumento fornisce è compreso entro $\pm\Delta$. In mancanza di ulteriori affermazioni, si può assumere una distribuzione uniforme:

$$u = \frac{2\Delta}{\sqrt{12}} = \frac{\Delta}{\sqrt{3}} = 0.58\Delta \approx 0.6\Delta .$$

6. Un parametro è compreso, con la quasi sicurezza, entro $\pm\Delta$, ma si tende a credere più ai valori centrali che a quelli estremi. In questo caso, è più ragionevole ipotizzare una distribuzione triangolare:

$$u = \frac{\Delta}{\sqrt{6}} = 0.41 \Delta \approx 0.4 \Delta .$$

In modo alternativo, l'informazione potrebbe essere anche compatibile con un modello gaussiano con l'intervallo a 2 o 3 σ . Si otterrebbero allora $u = \Delta/2$ o $\Delta/3$, valori a cavallo di quanto ottenuto con la triangolare. Quindi quest'ultima può essere considerata un compromesso per quantificare quello stato di incertezza.

7. Si legge un valore su uno strumento digitale, in cui l'intervallo di scala (la variazione della grandezza associata alla variazione di una unità della cifra meno significativa) vale I . Il valore vero potrebbe essere ovunque nell'intervallo ampio I . Quindi l'incertezza da associare alla quantizzazione della lettura è pari a

$$u = \frac{I}{\sqrt{12}} = 0.29 I \approx 0.3 I.$$

22 Esempi numerici

Facciamo ora degli esercizi numerici su alcuni dei concetti più critici.

1. Riprendiamo il problema del termometro del paragrafo 6. Per dare il risultato occorre valutare la deviazione standard di ripetibilità, che tiene conto anche dell'errore di lettura. Effettivamente bisognerebbe sapere qualcosa di più sulla misura. Come mero esercizio accademico, assumiamo che la misura sia disagiata, addirittura tale che ci sentiamo "tranquilli" entro una divisione. Ne segue $\sigma_r = 0.03^\circ\text{C}$ (punto 5 del paragrafo precedente). Inoltre, possiamo interpretare "scalibrato al più di 0.6°C " come una incertezza descritta da una distribuzione uniforme di semiampiezza 0.6°C e deviazione standard 0.35°C . Abbiamo quindi:

$$\begin{aligned} T_1 &= 22.00 \pm 0.35^\circ\text{C} \\ T_2 &= 23.00 \pm 0.35^\circ\text{C} \\ \rho(T_1, T_2) &= 0.993 \\ T_2 - T_1 &= 1.00 \pm 0.04^\circ\text{C}. \end{aligned}$$

(La risposta naïve per la differenza sarebbe $1.0 \pm 0.5^\circ\text{C}$.)

2. Usando un voltmetro digitale, applicato in due punti di un circuito, si leggono i seguenti valori: 3.512 e 3.508 V. Le misure vengono ripetute più volte e i valori si ripetono esattamente uguali. Le istruzioni del tester affermano che l'errore massimo che il tester può commettere è pari allo 0.2% . Determinare la differenza di potenziale fra i due punti.

Chiaramente, l'incertezza su ciascuno dei punti è 0.007 V ($= 0.002 \times 3.5$ V) ma l'errore di calibrazione non può influenzare la differenza fra valori così vicini, specialmente se è stato verificato che, alternando le letture, effettivamente i valori si ripetono. Nella differenza conta allora soltanto l'incertezza di digitalizzazione:

$$\Delta V = 4.0 \pm 0.4 \text{ mV}.$$

(La risposta naïve sarebbe 4 ± 10 mV. L'incertezza di 0.4 mV è data da $\sqrt{2 \times 1/\sqrt{12}^2}$.)

3. Si effettuano delle misure dello spessore di tre blocchetti con un calibro Palmer fabbricato secondo le norme DIN 863. Esso ha divisioni di $1/100$ di mm e la massima escursione dell'indicazione rispetto al valore vero è di $\pm 2 \mu\text{m}$. Per ogni spessore si eseguono 10 misure, ottenendo i seguenti valori di media e deviazione standard (in mm): $\bar{x}_1 = 0.9832$, $\sigma_1 = 0.0032$,

$\bar{x}_2 = 0.9869$, $\sigma_2 = 0.0028$, $\bar{x}_3 = 1.7341$, $\sigma_3 = 0.0041$. Quanto valgono i tre spessori? Valutare anche la differenza del secondo e del terzo rispetto al primo.

Per ciascuno spessore l'incertezza totale è data dalla combinazione in quadratura dell'incertezza di tipo A ($\sigma_i/\sqrt{10}$) e quella di tipo B ($u = 2\mu\text{m}/\sqrt{3}$). Nelle differenze bisogna tener conto delle eventuali correlazioni. Il tipo di strumento e di misura in oggetto appartengono al caso non "facilmente schematizzabile" discusso nel paragrafo 18.5. Ragionevoli conclusioni sono:

$$\begin{aligned}x_1 &= 0.9832 \pm 0.0015 \text{ mm} \\x_2 &= 0.9869 \pm 0.0015 \text{ mm} \\x_3 &= 1.7341 \pm 0.0017 \text{ mm} \\x_2 - x_1 &= 3.7 \pm 1.3 \mu\text{m} \quad (\text{e non } \pm 2.1 \mu\text{m}) \\x_3 - x_1 &= 0.7509 \pm 0.0023 \text{ mm}\end{aligned}$$

4. Sui dati del problema precedente. Supponiamo che il coefficiente di dilatazione termica del calibro sia $\alpha_c = 1.2 \cdot 10^{-5} \text{ K}^{-1}$ e quello dei blocchetti $\alpha_b = 1.0 \cdot 10^{-4} \text{ K}^{-1}$. Lo sperimentatore purtroppo ha dimenticato di misurare la temperatura, ma, dalla memoria della sensazione fisiologica all'atto della misura, è convinto che essa fosse "molto probabilmente fra 27 e 28 gradi e quasi sicuramente compresa fra 25 e 30 gradi. Come verranno forniti i risultati alla temperatura di riferimento di 20 °C?

I gradi di fiducia sui possibili valori di temperatura possono essere modellizzati con una triangolare o con una gaussiana a 2-3 sigma. Si ha, rispettivamente, nei tre casi: $u_1 = 1.0 \text{ °C}$, $u_2 = 1.3 \text{ °C}$ e $u_3 = 0.8 \text{ °C}$. Prendendo il valore intermedio si ottiene:

$$\Delta T = 7.5 \pm 1.0 \text{ °C}.$$

Applichiamo ora la correzione per la temperatura, ricordando che

$$l = l_0 \frac{1 + \alpha_b \Delta T}{1 + \alpha_c \Delta T}.$$

Invertendo e trascurando i termini di ordine superiore a ΔT (essendo i coefficienti α molto piccoli):

$$\begin{aligned}l_0 &= l \frac{1 + \alpha_c \Delta T}{1 + \alpha_b \Delta T} \\&\approx l (1 + \alpha_c \Delta T - \alpha_b \Delta T) = l [1 - (\alpha_b - \alpha_c) \Delta T]\end{aligned}$$

$$\begin{aligned}\beta &= 1 - (\alpha_b - \alpha_c) \Delta T = 0.99934 \pm 0.00009 \\l_1 &= 0.9826 \pm 0.0015 \text{ mm} \\l_2 &= 0.9862 \pm 0.0015 \text{ mm} \\l_3 &= 1.7330 \pm 0.0017 \text{ mm} \\l_2 - l_1 &= 3.6 \pm 1.3 \mu\text{m} \\l_3 - l_1 &= 0.7504 \pm 0.0023 \text{ mm}\end{aligned}$$

5. “Pesa più un chilogrammo di piombo o un chilogrammo di polistirolo?” suona come una domanda per ingannare i bambini, ma, opportunamente riformulata, mette in imbarazzo i grandi: “se si pongono su una bilancia di precisione un kg di piombo e un kg di polistirolo, in quale caso si leggerà il valore maggiore?” La risposta è ovviamente “il piombo” in quanto il polistirolo risente molto di più della spinta di Archimede dell’aria (si era assunto che la bilancia non operasse sotto vuoto, ovviamente). Facciamo un esempio: un parallelepipedo di polistirolo ha i lati di 36.3 ± 0.2 cm, 37.2 ± 0.2 cm e 36.8 ± 0.2 cm (incertezze scorrelate). Esso è posto su una bilancia perfettamente calibrata, posta in una stanza in cui la densità dell’aria vale esattamente 1.1836 kg m^{-3} (vedremo nel prossimo esercizio quanto questa affermazione sia realistica). Sulla scala digitale si legge 941 g. Determinare massa e densità del polistirolo.

Il volume V vale $(4.97 \pm 0.05) \cdot 10^{-2} \text{ m}^3$ e la massa apparente M_a 0.9410 ± 0.0003 kg, da cui:

$$\begin{aligned} M &= M_a + \rho_{aria} V = 999.8 \pm 0.7 \text{ g} \\ \rho_p &= \frac{M}{V} = \rho_{aria} + \frac{M_a}{V} = \rho_{aria} + \rho_a = 20.12 \pm 0.19 \text{ kg m}^{-3}. \end{aligned}$$

6. Quanto vale la densità dell’aria? Qualcuno si sarà insospettito dal valore “esatto” di 1.1836 kg m^{-3} , specialmente quando in molti libri di testo si trova 1.293 kg m^{-3} . Questo è un caso di incertezza dovuta ad “inesatta definizione del misurando”. Cos’è l’“aria”? Il valore di 1.1836 kg m^{-3} vale per “aria di riferimento”³⁷ (78.10% N₂; 20.94% O₂; 0.92% Ar, ...) con lo 0.06% di CO₂ (invece dell’usuale valore di riferimento di 0.04%) alle seguenti condizioni ambientali: temperatura di 20 °C, pressione atmosferica di 1000 mbar e umidità relativa del 50%. Quanto vale allora l’incertezza sulla densità dell’aria? Essa dipende dall’incertezza sui fattori di influenza.

Facciamo il seguente caso (φ sta per umidità relativa):

$$\begin{aligned} x_{CO_2} &= 0.06 \pm 0.02 \% \\ T &= 20 \pm 1 \text{ }^\circ\text{C} \\ p &= 1000 \pm 20 \text{ mbar} \\ \varphi &= 0.50 \pm 0.10 \text{ (ovvero } 50 \pm 10 \text{ \%)} . \end{aligned}$$

Per fare i conti utilizziamo seguente formula:

$$\rho = (\rho_0 + \varphi A) [1 + 0.0041 (x_{CO_2} - 0.04 \%)] , \quad (39)$$

dove ρ_0 è il valore della densità dell’aria secca, dipendente da temperatura e pressione, e A è un coefficiente che dipende soltanto dalla temperatura (vedi tabella 2). Eseguendo le derivate per via numerica (vedi paragrafo 18.7) troviamo i seguenti contributi all’incertezza totale

$$\begin{aligned} x_{CO_2} = 0.06 \pm 0.02 \% &\rightarrow u_1 = 0.0001 \text{ kg m}^{-3} \\ T = 20 \pm 1 \text{ }^\circ\text{C} &\rightarrow u_2 = 0.0004 \text{ kg m}^{-3} \\ p = 1000 \pm 20 \text{ mbar} &\rightarrow u_3 = 0.0237 \text{ kg m}^{-3} \\ \varphi = 0.50 \pm 0.10 \% &\rightarrow u_4 = 0.0010 \text{ kg m}^{-3} \end{aligned}$$

³⁷I valori che seguono, la formula (39) e la tabella 2 sono presi da F. Kohlrausch, “*Praktische Physik*”, B.G. Teubner Stuttgart 1986. La sezione 72 sulla densità dell’aria è curata da M. Kochsieck dell’istituto tedesco di metrologia di Braunschweig.

T (°C)	p (mbar)			A kg m^{-3}
	980	1000	1020	
19	1.1690	1.1928	1.2167	$-9.87 \cdot 10^{-3}$
20	1.1650	1.1887	1.2125	$-10.47 \cdot 10^{-3}$
21	1.1610	1.1847	1.2084	$-11.09 \cdot 10^{-3}$

Tabella 2: Densità dell'aria secca (in kg m^{-3}) per alcuni valori di temperatura e pressione. A rappresenta il coefficiente di dipendenza dall'umidità relativa (vedi formula (39)).

Si vede quindi come il contributo più importante sia dovuto all'incertezza sulla pressione. Il valore della densità dell'aria con questo stato di conoscenza dei fattori di influenza vale quindi

$$\rho = 1.184 \pm 0.024 \text{ kg m}^{-3}.$$

7. Quanto dovrebbe essere il controllo sulle variabili di influenza per poter affermare che la densità dell'aria sia esattamente 1.1836 kg m^{-3} ?

Affermare che $\rho = 1.1836 \text{ kg m}^{-3}$ sia esatto vuol dire che il solo errore possibile è quello di arrotondamento. Quindi $u(\rho) \lesssim 0.0001/\sqrt{12} = 3 \cdot 10^{-5} \text{ kg m}^{-3}$. Se le quattro incertezze contribuiscono allo stesso modo, ciascuna di esse deve valere al più $u_i \approx 1.5 \cdot 10^{-5} \text{ kg m}^{-3}$. Ne seguono i seguenti requisiti:

$$\begin{aligned} \sigma(x_{CO_2}) &\lesssim 0.003 \% \\ \sigma(T) &\lesssim 0.04 \text{ }^\circ\text{C} \\ \sigma(p) &\lesssim 0.013 \text{ mbar} \\ \sigma(\varphi) &\lesssim 0.15 \%, \end{aligned}$$

condizioni di lavoro tutt'altro che banali! Si capisce allora come, nei casi pratici, non abbia molto senso far riferimento ad una densità dell'aria con più di tre cifre significative.

23 Fit di andamenti lineari

Affrontiamo ora il problema dello studio delle curve che approssimano meglio i punti sperimentali. Per semplicità ci limitiamo ad andamenti lineari, ovvero

$$y = m x + c.$$

Consideriamo, ad esempio, i dati sperimentali, raccolti da un gruppo di studenti, dello studio dell'allungamento e del periodo di una molla appesa a un gancio, in funzione della massa applicata (vedi tabella 3). Consideriamo soltanto la prima serie di misure. Si noti l'assenza, a questo livello, di incertezze associate alle misure. Infatti ogni stima sarebbe arbitraria. I valori sono stati riportati con tutte le cifre che si riuscivano ad apprezzare sugli strumenti. Si noti come gli studenti abbiano deciso, nonostante le raccomandazioni contrarie, che, date le condizioni di lavoro, fosse difficile effettuare letture al di sotto del millimetro. Vedremo nel seguito se, con il senno del poi, avrebbero dovuto sforzarsi un po' e quali sono le conseguenze sul risultato.

n	M (g)	Prima serie		Seconda serie		Terza serie	
		l (mm)	$T \times 10$ (s)	l (mm)	$T \times 10$ (s)	l (mm)	$T \times 10$ (s)
0	63	0	-	0	-	0	-
1	142	0	-	0	-	0	-
2	221	0	-	0	-	0	-
3	300	14	5.01	16	5.09	16	5.19
4	379	32	5.57	33	5.66	33	5.68
5	458	49	6.24	51	6.27	51	6.34
6	536	66	6.78	68	6.82	69	6.94
7	615	85	7.28	86	7.33	87	7.28
8	694	103	7.79	103	7.81	103	7.86
9	773	119	8.13	121	8.31	121	8.24
10	852	137	8.63	139	8.77	139	8.70

Tabella 3: Dati dell'esperienza della molla. n è il numero dei dischetti sospesi, l è l'allungamento ($l_0 = 0$) e T il periodo di oscillazione. Il peso di 10 dischetti è pari a 789 g e la massa della molla stessa (parte oscillante) è pari a circa 63 g.

23.1 Grafici e parametri della retta

I punti sperimentali sono riportati su due grafici (allungamento in funzione della massa e periodo in funzione della radice quadrata della massa, per ovvi motivi). La figura 10 mostra i due grafici. Per quanto riguarda l'elaborazione successiva si può considerare

- un'analisi grafica;
- un'analisi mediante metodo dei minimi quadrati.

Anche se oggi giorno ci sono programmi al computer che “fanno tutto”, essi sono in genere poco raccomandabili dal punto di vista didattico (specialmente se usati nelle prime esperienze). Infatti:

- siccome sono pochi i laboratori che dispongono di un computer per ogni banco di lavoro per l'intera durata dell'esperienza, gli studenti si abituano a separare la fase di raccolta dati da quella di rappresentazione grafica e di elaborazione. Qualsiasi fisico sperimentale raccomanda di fare i grafici immediatamente dopo la presa dati (idealmente addirittura durante), almeno per le prime serie di misure. Questo è il solo modo per controllare che tutto stia procedendo per il verso giusto e che non ci siano sbagli nelle procedure.
- E' più facile sbagliare ad inserire i dati nel computer che a tracciare una retta ad occhio fra i punti sperimentali.
- I programmi sono scatole nere che a volte non usano le assunzioni corrette e quasi mai danno il coefficiente di correlazione fra i parametri, di cruciale importanza per elaborazioni successive (in genere i parametri dei fit lineari, nelle esperienze tipiche di laboratorio, sono correlati all'80-90 %!).

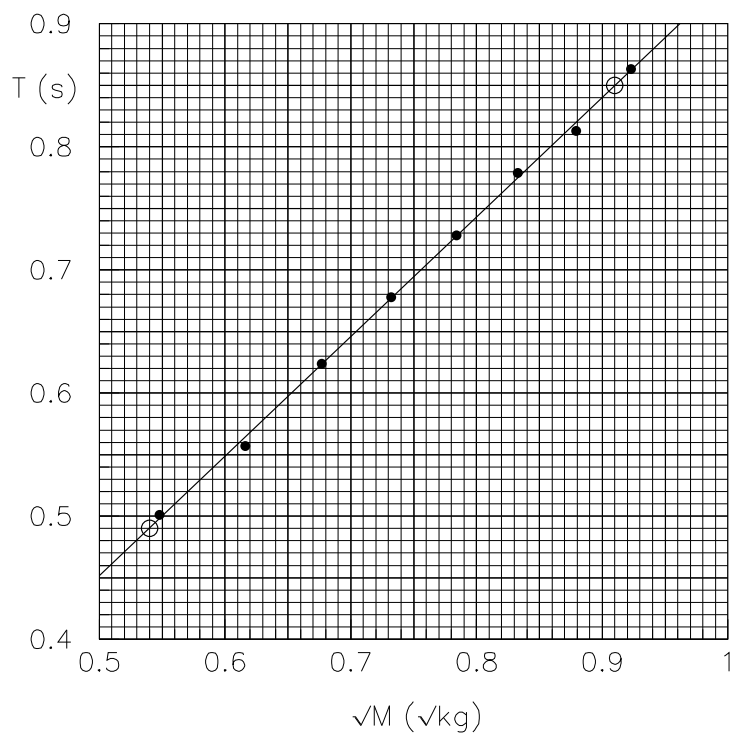
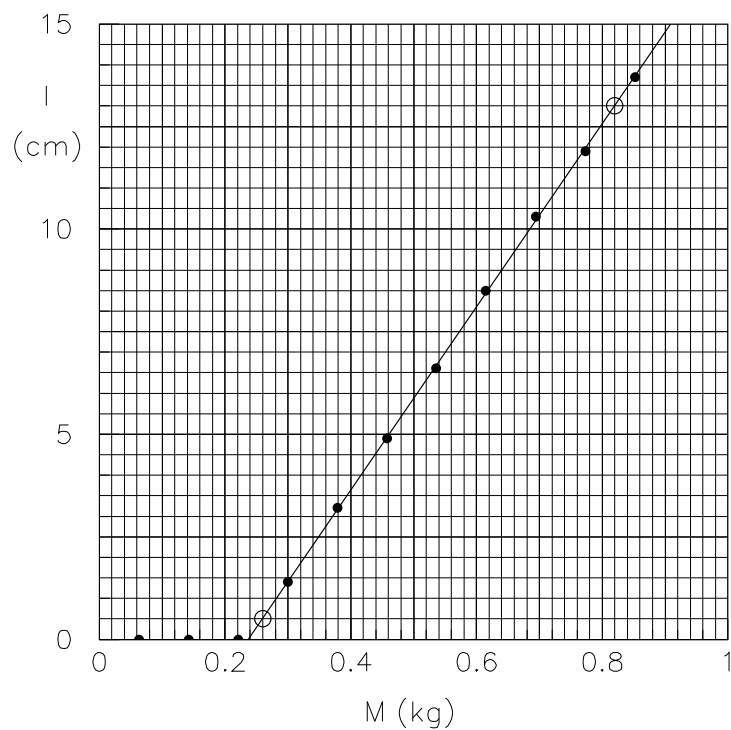


Figura 10: Grafici dell'allungamento in funzione della massa sospesa alla molla e del suo periodo di oscillazione in funzione della radice quadrata della massa. I cerchietti indicano i punti utilizzati per il calcolo dei parametri della retta.

- Gli studenti dovrebbero avere accesso soltanto a programmi ben testati e soltanto dopo che abbiano effettuato a mano tutti i conti in un paio di esercitazioni (una sorta di “patente” che li abiliti a usare i programmi). L’ideale sarebbe che essi possano sviluppare i programmi di analisi di cui necessitano, in quanto i problemi non banali richiedono in genere una soluzione adatta al problema specifico.

23.2 Inferenza bayesiana e metodo dei minimi quadrati

Nel seguito mostreremo come ricavarsi i parametri della retta mediante una procedura grafica che fornisce risultati molto simili a quelli del fit con minimi quadrati (che supponiamo noti, almeno a grandi linee).

Prima di procedere è però opportuno far notare che anche questo metodo può essere fatto discendere dai principi dell’inferenza bayesiana, anche se non staremo a dimostrarlo. Ne segue che, anche in questo caso, le grandezze associate ai parametri della retta possono essere considerate variabili casuali, con tutte le semplificazioni tecniche e interpretative che ne discendono.

23.3 Analisi grafica

Vediamo quali sono i passi necessari per un’analisi grafica che, condotta a termine fino in fondo, produce risultati quantitativi in accordo con quelli ottenibili mediante fit con i minimi quadrati. In molte esperienze, comunque, non è necessario procedere ad un’analisi così accurata come quella proposta (vedi discussione nel paragrafo 24) e ci si può fermare al primo passo.

Stima dei parametri

- Si traccia ad occhio la retta, cercando di passare in mezzo a tutti i punti. Questa retta praticamente coincide con quella che si ottiene con i minimi quadrati.
- Si ricavano quindi dal grafico due punti che giacciono sulla retta, che siano ben distanziati e ben leggibili. Da questi si ricavano m e c (l’intercetta si ottiene in genere più agevolmente in modo diretto, come è ben noto).

Per quanto riguarda le cifre con cui rileggere i valori si noti come i punti della retta sono “più stabili” di quelli delle singole misure e quindi possono essere riletti anche con una cifra in più. Si ottengono quindi c e m con il numero di cifre che seguono dalle solite regolette sulle cifre significative.

Stima dell’incertezza sui parametri ripetendo le misure

Il modo più semplice, e che per le prime esperienze è indubbiamente istruttivo, è quello di ripetere più volte la misura e studiare le fluttuazioni dei risultati. Si tenga conto che, non facendo calcoli di “errori massimi” né propagazioni varie, è molto facile ripetere più volte le misure in alcune ore.

Stima dell’incertezza della singola misura dai residui

In realtà non c’è alcun bisogno di ripetere le serie di misure. Se ciascuna

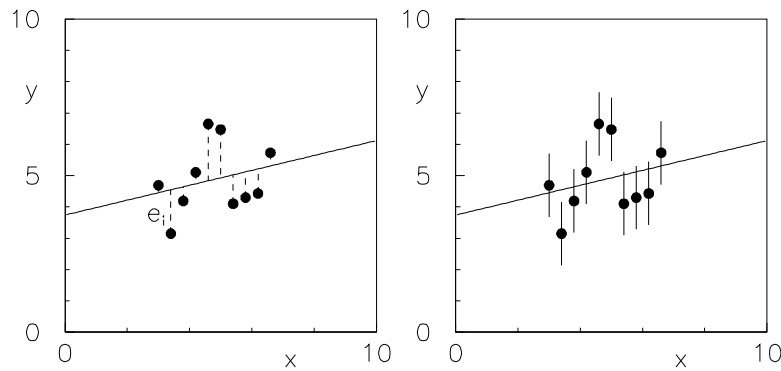


Figura 11: Residui e barre di incertezza ($\pm\sigma_r$).

serie contiene un numero sufficiente di punti (tipicamente, leggermente superiore al numero di parametri che si vogliono valutare) essa racchiude in sé le informazioni necessarie alla valutazione delle incertezze, o almeno a quelle derivanti da errori casuali, mediante il metodo dei residui. Una volta tracciata la retta si può leggere dal grafico, per ogni punto, il residuo e_i , ovvero la differenza fra l'ordinata misurata e il valore della retta in corrispondenza dell'ascissa misurata, come mostrato in figura 11. Si ottiene quindi, dalla media dei quadrati dei residui, la stima della deviazione standard delle ordinate, assumendo che sia la stessa per tutti i punti e attribuendo soltanto alle ordinate le deviazioni dal valore vero:

$$\sigma_r = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n-2}}.$$

Il nome σ_r sta a indicare sia che essa è calcolata dai residui sia che rappresenta l'equivalente della deviazione standard di ripetitività delle misure. Il fattore $n-2$ al posto di n ha la stessa giustificazione dell' $n-1$ nella deviazione standard, tenendo conto che ora ci sono due vincoli fra i dati. Ripetiamo ancora una volta quanto detto a proposito di σ_{n-1} : anche se le ragioni profonde di questa scelta non sono sempre condivisibili, il risultato “va nella direzione giusta”. Anche qui, quando n è dell'ordine della decina, la correzione è influente ai fini pratici.

A questo punto, finalmente si conosce l'errore casuale sulle ordinate in condizioni di ripetitività (nell'ipotesi che quello sulle ascisse sia trascurabile)!

Ovviamente, si può anche fare l'esercizio opposto e attribuire tutto l'errore alle ascisse (senza dover fare tutti i conti, si può propagare σ_r su “ σ_{r_x} ” mediante la derivata: $\sigma_{r_x} = \sigma_r/|m|$). E' interessante notare che, anche se il punto di vista cambia drasticamente, saranno invarianti le conclusioni sulle grandezze fisiche di interesse, legate a coefficiente angolare e intercetta.

Valutazione semplificata di σ_r

Spesso non c'è tempo per rileggere tutti i punti della retta e calcola-

re σ_r dalla somma dei quadrati dei residui. Oppure, semplicemente, ci si accontenta di una sua stima al 20-30 %, (sembra tanto, ma anche un'incertezza del 50 % su σ_r è più che accettabile per molte applicazioni, specialmente se essa è ottenibile in tempi rapidi). In questi casi si possono tracciare, simmetricamente alla retta di migliore stima, due rette parallele tali che contengano i 2/3 circa dei punti. La distanza, lungo l'asse delle ascisse, fra le due rette è approssimativamente uguale a σ_r . Analogamente, si possono considerare la "quasi totalità" dei punti, e considerarlo un intervallo a 2 o a 3 deviazioni standard.

Barre di incertezza

Soltanto a questo punto è lecito riportare le barre di incertezza sui punti del grafico. Ogni barretta verticale è centrata sul punto sperimentale ed ha lunghezza $2\sigma_r$.

Incertezza dei parametri mediante σ_r ricavata dai dati

Nota la deviazione standard da attribuire alle singole fluttuazioni delle y , si possono usare le formule delle incertezza che si ricavano dal metodo dei minimi quadrati, che qui riportiamo per comodità, riscritte in termini delle grandezze che si conoscono e di quanto altro sia facilmente valutabile per via grafica:

$$\sigma(m) = \frac{1}{\sqrt{\text{Var}(x)}} \frac{\sigma_r}{\sqrt{n}} \quad (40)$$

$$\sigma(c) = \sqrt{\frac{\text{Var}(x) + \bar{x}^2}{\text{Var}(x)}} \frac{\sigma_r}{\sqrt{n}} \quad (41)$$

$$\rho(m, c) = -\frac{\bar{x}}{\sqrt{\text{Var}(x) + \bar{x}^2}}. \quad (42)$$

Si noti che $\text{Var}(x)$ non è legata alle incertezze sulle X (convenzionalmente nulle): essa misura invece la dispersione delle x e la sua radice quadrata è legata al cosiddetto "braccio di leva" dei dati sperimentali. E' interessante notare come questo possa essere valutato agevolmente dal grafico se i punti sono circa spazati lungo l'ascissa (caso tipico delle esercitazioni di laboratorio). Approssimando i punti sperimentali ad una distribuzione uniforme si ottiene infatti³⁸

$$\sqrt{\text{Var}(x)} \approx \frac{x_{max} - x_{min}}{\sqrt{12}}.$$

Si faccia inoltre attenzione a non confondere il coefficiente di correlazione fra i parametri (indicato con $\rho(m, c)$) con il coefficiente di correlazione

³⁸Si noti che questa espressione è valida per variabili continue. Per variabili discrete equispaziate fra x_{min} e x_{max} , la formula esatta è

$$\frac{x_{max} - x_{min}}{\sqrt{12}} \sqrt{\frac{n+1}{n-1}},$$

che tende alla deviazione standard del caso continuo quando n è molto grande. Comunque, già per $n = 5$ il fattore correttivo è del 20 % e per $n = 10$ è del 10 %.

fra ascisse e ordinate dei punti sperimentali (indicato con $\rho(x, y)$): la grandezza di interesse ai fini del risultato è $\rho(m, c)$.

Analisi nel baricentro

Osservando le formule (41) e (42) si nota che, se \bar{x} è uguale a zero, il coefficiente di correlazione si annulla e anche l'espressione di $\sigma(c)$ è uguale a σ_r/\sqrt{n} . Questo suggerisce che per semplificare i conti conviene scegliere l'asse delle ascisse in corrispondenza del baricentro dei punti, ovvero effettuare la trasformazione di variabili

$$x' = x - \bar{x};$$

questo è particolarmente comodo se successivamente si deve utilizzare la retta trovata per delle *estrapolazioni* o, in generale, come *curva di taratura*. Chiaramente si otterranno in questo caso valori m' e c' diversi da m e da c ed è facile dimostrare che c' è uguale al baricentro delle \bar{y} .

Se si vuole visualizzare l'incertezza su m' e su c' sarà sufficiente

- tracciare le rette passanti per il baricentro e di pendenze $m' \pm \sigma(m')$;
- disegnare una barra verticale centrata nel baricentro e di semiampiezza $\sigma(c')$.

E' possibile passare poi dai parametri nel sistema del centro di massa a quelli nel sistema originale tenendo conto che $m \pm \sigma(m) = m' \pm \sigma(m')$, $c = c' - \bar{x}m$ e valutando $\sigma(c)$ e $\rho(m, c)$ dalle (41) e (42).

23.4 Effetto degli errori sistematici

La storia in principio non è ancora finita. Cosa abbiamo dimenticato? In questi casi bisogna ripassarsi rapidamente i primi 9 punti del "decalogo" del paragrafo 4 (il punto 10 è l'unico di cui abbiamo tenuto conto, al meglio di quello che si poteva fare).

Consideriamo soltanto l'effetto di errori di calibrazione dello strumento. Vediamo quali sono i possibili contributi all'incertezza. Per comodità ci riferiamo alla figura 9.

Errori sistematici dipendenti dal valore della grandezza

Si tratta del caso d) di figura 9. Di questo si è già tenuto conto, implicitamente, quando è stata valutata σ_r dai residui. Infatti ogni piccola incertezza di taratura nell'intervallo di scala utilizzato influenza σ_r , così pure ogni piccola differenza fra le masse dei diversi dischi.

Errore di zero

- L'errore di zero, sia esso sulle ascisse che sulle ordinate, non ha nessuna influenza sul coefficiente angolare, in quanto esso è valutato come rapporto di differenze:

$$\sigma(m)|_{z_x} = \sigma(m)|_{z_x} = 0.$$

- Un errore di zero sulle ordinate (σ_{z_y}) si riflette direttamente sul valore dell'intercetta, in quanto tutti i punti potrebbero essere traslati coerentemente verso l'alto o verso il basso. Quindi il contributo a $\sigma(c)$ è

$$\sigma(c)|_{z_y} = \sigma_{z_y}.$$

- Un errore di zero sulle ascisse (σ_{z_x}) causerebbe una traslazione orizzontale di tutti i punti. Quindi esso si riflette sull'intercetta attraverso la pendenza della retta. Ne segue che il contributo a $\sigma(c)$ è (trascurando l'effetto di ordine superiore dovuto all'incertezza sulla pendenza stessa):

$$\sigma(c)|_{z_x} = |m|\sigma_{z_x}.$$

Se entrambi i contributi sono presenti, essi vanno considerati in quadratura.

Errore di scala

Per capire il comportamento degli errori di scala, sia sulle ascisse che sulle ordinate, consideriamo due punti sulla retta $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$, in corrispondenza del primo e dell'ultimo punto sperimentale. Da questi punti è possibile ricavarsi m e c :

$$m = \frac{y_2 - y_1}{x_2 - x_1}, \quad (43)$$

$$c = y_1 - \frac{y_2 - y_1}{x_2 - x_1}x_1. \quad (44)$$

Consideriamo ora i fattori di scala f_x e f_y che, come al solito, riteniamo essere

$$\begin{aligned} f_x &= 1 \pm \sigma_{f_x} \\ f_y &= 1 \pm \sigma_{f_y}. \end{aligned}$$

e inseriamoli esplicitamente nelle espressioni di m e di c :

$$\begin{aligned} m &= \frac{f_y(y_2 - y_1)}{f_x(x_2 - x_1)} \\ c &= f_y y_1 - \frac{f_y(y_2 - y_1)}{f_x(x_2 - x_1)} f_x x_1 = f_y \left(y_1 - \frac{y_2 - y_1}{x_2 - x_1} x_1 \right). \end{aligned}$$

Ne concludiamo quindi che

- Il coefficiente angolare risente allo stesso modo dei due fattori di scala:

$$\sigma(m)|_{f_x} = |m|\sigma_{f_x} \quad (45)$$

$$\sigma(m)|_{f_y} = |m|\sigma_{f_y} \quad (46)$$

- L'errore di scala sulle ascisse è ininfluenza sull'intercetta, mentre quello sulle ordinate vi si riflette proporzionalmente:

$$\sigma(c)|_{f_x} = 0 \quad (47)$$

$$\sigma(c)|_{f_y} = |c|\sigma_{f_y} \quad (48)$$

Come al solito, i vari contributi si combinano quadraticamente.

Deviazione dalla linearità

Eventuali deviazioni dalla linearità degli strumenti si riflettono su una deviazione dalla linearità dell'andamento. Non è possibile dare delle regole ad hoc. Qualora si notino tali effetti nei punti sperimentali, bisogna cercare di capire se l'effetto è da attribuire a uno dei due strumenti (o a entrambi), oppure è la legge fisica ad essere inadeguata, o si tratta soltanto di una fluttuazione. Si cerca quindi di sostituire o ricalibrare gli strumenti, di apportare correzioni fenomenologicamente giustificate ai dati sperimentali, oppure si utilizza semplicemente la zona nella quale ci sono dei buoni motivi per presupporre che l'andamento sia quello ipotizzato. Un caso macroscopico è quello del grafico dell'allungamento in funzione della massa di figura 10: la molla non si allunga affatto al di sotto di una massa critica (≈ 200 g) e quindi i primi tre punti non vanno considerati nell'analisi.

La valutazione di effetti sistematici rimane invariata se i parametri della retta sono valutati con i minimi quadrati anziché con l'analisi grafica.

23.5 Esempio numerico

Come esercizio ricaviamoci i parametri dell'andamento dell'allungamento in funzione della massa di figura 10.

1. Si traccia a occhio la retta che meglio passa fra i punti;
2. Si ricavano due punti sulla retta ben leggibili (quelli cerchiati) e da essi i valori di m e di c ³⁹:

$$\begin{aligned} P_1 &= (0.260 \text{ kg}, 0.51 \text{ cm}) \\ P_2 &= (0.820 \text{ kg}, 13.00 \text{ cm}) \\ m &= \frac{(13.00 - 0.51) 10^{-2} \text{ m}}{(0.820 - 0.260) \text{ kg}} = 0.2230 \frac{\text{m}}{\text{kg}} \\ c &= y_1 - m x_1 = -0.0530 \text{ m} = -5.30 \text{ cm}; \end{aligned}$$

3. Utilizzando i valori dei parametri della retta ricavata dal grafico, si calcolano i valori "teorici" dell'allungamento per gli 8 valori di massa. Essi sono, in mm: 13.9; 31.5; 49.1; 66.5; 84.1; 101.8; 119.4; 137.0. Si ottengono quindi i residui dalla differenza fra allungamenti misurati e allungamenti calcolati lungo la retta: 0.1; 0.5; -0.1; -0.5; 0.9; 1.2; -0.4; 0.0. La somma dei loro quadrati è pari a 2.93 mm^2 , da cui si ricava

$$\sigma_r = 0.70 \text{ mm}.$$

Probabilmente gli studenti avrebbero fatto meglio a non accontentarsi della lettura al millimetro. La deviazione standard ottenuta è infatti compatibile con quella dovuta al solo effetto di arrotondamento al millimetro. Siccome tutte le incertezze che seguiranno saranno proporzionali a σ_r , sarebbe stato importante leggere al meglio i valori, a costo di "inventarsi i decimi". L'analisi dei residui sfronda poi, in modo automatico, tutte le cifre superflue.

³⁹Per quanto riguarda il numero di cifre significative, si noti come ne sia stata aggiunta una in più rispetto alle regole usuali. Esse verranno aggiustate in seguito alla luce di $\sigma(m)$ e di $\sigma(c)$. Nel caso in cui l'esperienza non preveda un'analisi completa delle incertezze di misura sarebbe stato sufficiente scrivere $m = 0.223 \text{ m kg}^{-1}$ e $c = -5.3 \text{ cm}$.

4. Conoscendo σ_r , si possono finalmente riportare le barre di incertezza sul grafico. Nel caso in considerazione le barre risulterebbero invisibili sulla figura in quanto confrontabili con le dimensioni dei pallini con i quali sono stati indicati i punti stessi.
5. Per calcolare le incertezze serve conoscere anche il braccio di leva e il baricentro della distribuzione delle ascisse. Dal grafico valutiamo:

$$\begin{aligned}\bar{x} &= 0.576 \text{ kg} \\ \sqrt{\text{Var}(x)} &= \frac{(0.852 - 0.300) \text{ kg}}{\sqrt{12}} = 0.159 \text{ kg}\end{aligned}$$

6. Finalmente abbiamo il risultato:

$$\begin{aligned}m &= 0.2230 \pm 0.0016 \frac{\text{m}}{\text{kg}} \\ c &= -53.0 \pm 0.9 \text{ mm} \quad (= -0.0530 \pm 0.0009 \text{ m}) \\ \rho(m, c) &= -0.96.\end{aligned}$$

Si noti l'altissima correlazione negativa fra i due parametri: detto alla buona, è un po' come se, invece di aver misurato 2 grandezze (nel senso di 2 grandezze indipendenti), ne avessimo misurate 1.04.

7. Confrontiamo il risultato con quanto si ottiene con un fit di minimi quadrati eseguito **al computer**:

$$\begin{aligned}\sigma_r &= 0.66 \text{ mm} \\ m &= 0.2231 \pm 0.0013 \frac{\text{m}}{\text{kg}} \\ c &= -52.8 \pm 0.8 \text{ mm} \quad (= -0.0528 \pm 0.0008 \text{ m}) \\ \rho(m, c) &= -0.95.\end{aligned}$$

L'accordo è più che soddisfacente⁴⁰

8. Per mostrare l'importanza del coefficiente di correlazione ci calcoliamo il valore dell'allungamento previsto per una massa di 1.000 kg⁴¹. Propagando lo stato di incertezza sui parametri della retta sul valore estrapolato, abbiamo:

$$\begin{aligned}\sigma^2(y) &= x^2\sigma^2(m) + \sigma^2(c) + 2x\rho(m, c)\sigma(m)\sigma(c) \\ &= 2.56 + 0.81 - 2.76 \text{ mm}^2 = 0.61 \text{ mm}^2 \\ \sigma(y) &= 0.8 \text{ mm} \\ y &= 170.0 \pm 0.8 \text{ mm}.\end{aligned}$$

⁴⁰Le piccole differenze numeriche sono dovute all'uso della formula approssimata per il braccio di leva. Corretto per il fattore

$$\sqrt{\frac{n+1}{n-1}} = 1.13,$$

esso diventa 0.180 kg, da cui ne segue un risultato praticamente identico a quello ottenuto mediante programma:

$$\begin{aligned}m &= 0.2230 \pm 0.0014 \frac{\text{m}}{\text{kg}} \\ c &= -53.0 \pm 0.8 \text{ mm} \\ \rho(m, c) &= -0.95.\end{aligned}$$

⁴¹Si noti come nella propagazione non si debba tener conto di un'eventuale incertezza sulla massa se essa è simile a quella dei pesetti con i quali sono state effettuate le misure, in quanto questo contributo è già compreso in σ_r .

Tralasciando il coefficiente di correlazione si sarebbe ottenuto $\sigma(y) = 1.8$ mm.

Il computer fornisce invece 170.2 ± 1.5 mm: ???. Mentre la differenza di 0.2 mm sul valore centrale è attribuibile alla lieve differenza fra i valori ottenuti con i due metodi e arrotondamenti interni dei valori, $\sigma(y) = 1.5$ mm è decisamente **sospetta**. Facendo dei controlli ci si accorge che, questo valore è inconsistente con i risultati del fit: **il programma era sbagliato!** Un'ulteriore conferma dell'invito a diffidare delle scatole nere informatiche⁴².

Si noti infine come l'effetto dovuto al termine di correlazione sia ancora più importante per valori all'interno della distribuzione dei punti sperimentali, con un massimo nel baricentro.⁴³ Per una massa di 0.576 kg si prevede un allungamento di 75.4 ± 0.3 mm (che diventerebbe ± 1.3 omettendo le correlazioni).

9. Per quanto riguarda infine le incertezze dovute ad errori sistematici dello strumento, si può dire tranquillamente che essi sono trascurabili:

- le letture degli allungamenti sono state effettuate su un foglio di carta millimetrata solidale con il supporto della molla: deviazioni dell'indicazione dell'ordine di grandezza di 0.7 mm o maggiori sono impensabili;
- i pesetti sono stati misurati tutti insieme, per sottrazione rispetto alla tara, e quindi qualsiasi loro combinazione ha un'incertezza standard ben inferiore ad una parte su mille; anche un'eventuale incertezza di 1 g, proprio a voler esagerare, sulla massa iniziale ($n = 0$), che si riflette su tutti i valori di massa e va perciò considerata come un errore di zero, produce al più un'incertezza sull'intercetta di 0.2 mm; infine, le eventuali differenze⁴⁴ fra un pesetto e l'altro sono già incluse, indirettamente, in σ_r .

Possono essere molto più importanti altri errori che si comportano come errori sistematici costanti in questa serie di misure, ma che possono cambiare effettuandone un'altra. Possono essere dovuti, ad esempio, al modo di operare di chi ha eseguito le misure o ad un suo errore costante di parallasse. Con il metodo

⁴²Questa non è una sceneggiata. Così come è vero che i dati sperimentali sono stati effettivamente presi da studenti e che l'analisi grafica l'ho fatta io prima di verificare i risultati al computer, è anche vero che, in data 25 aprile 1998, al momento di completare questo esercizio, mi sono accorto che il programma dava risultati sbagliati, in quanto dimenticava il termine di correlazione. Si trattava di un programma sviluppato da studenti durante una borsa di collaborazione, i cui risultati erano stati testati (con eccezione delle estrapolazioni, perché sembrava la parte meno critica e sulle quali erano state convenute le formule da usare).

Il valore di $\sigma(y)$ consistente con i risultati del fit dello stesso programma sarebbe dovuto essere 0.6 mm.

⁴³Come abbiamo già fatto notare, il sistema del baricentro è molto conveniente, in quanto ρ si annulla. L'espressione di $\sigma(y)$ è particolarmente semplice e istruttiva:

$$\sigma^2(y) = \sigma^2(c') + x'^2 \sigma^2(m').$$

Si riconosce la combinazione in quadratura dell'incertezza dovuta all'intercetta con quella del coefficiente angolare "proiettata" ad una distanza $|x'|$ dal baricentro.

Siccome $\sigma(y)$ deve essere invariante per traslazioni, antitrasformando da x' a x , otteniamo la formula

$$\sigma^2(y) = \frac{\sigma_r^2}{n} + (x - \bar{x})^2 \sigma^2(m) = \frac{\sigma_r^2}{n} + \frac{(x - \bar{x})^2}{\text{Var}(x)} \frac{\sigma_r^2}{n}.$$

Si vede quindi come la previsione sull'ordinata abbia una precisione che è massima in corrispondenza del baricentro dei punti e si deteriora quando ci si allontana dalla regione in cui sono state effettuate le misure.

⁴⁴Naturalmente, per ottenere la massima accuratezza sulla precisione dei parametri sarebbe stato meglio misurare individualmente ciascuno dei pesetti, al fine di ridurre σ_r , ma a questo livello non ne vale la pena.

grafico proposto, anche una scelta non ottimale della retta migliore produce variazioni nel risultato (tipicamente trascurabili). Questi effetti possono essere studiati soltanto ripetendo le misure (si ricorda che l'ideale sarebbe di poter fare la misura in condizioni di riproducibilità, ovvero tenendo fissa la definizione del misurando, cambiare tutto il resto).

10. Per concludere, tornando al punto 8, cerchiamo di chiarire, anche con esempi numerici, il significato delle diverse deviazioni standard associate alle ordinate che si incontrano effettuando i fit.
- Abbiamo indicato con $\sigma(y)$ la deviazione standard che quantifica l'incertezza della previsione del valore ("vero") di Y , per un dato valore di X . Come abbiamo visto, questa incertezza è minima in corrispondenza del baricentro dei punti sperimentali e aumenta quando ci si allontana da esso.
 - σ_r è dovuta invece alle fluttuazioni attese del singolo valore misurato intorno all'andamento medio dei punti sperimentali. Si sarebbe tentati di dire che essa misura le fluttuazioni dei valori delle y letti avendo fissato X , ma in realtà essa tiene conto anche delle piccole fluttuazioni del valore di Y da un valore di X all'altro.
 - C'è infine un'ultima deviazione standard, che chiamiamo $\sigma_t(y)$ che tiene conto della combinazione dei due effetti. Essa risponde alla domanda: "che valori di y potrei osservare se facessi l'esperimento fissando il valore di x ". L'incertezza sui parametri si riflette sul valore vero di y ed inoltre la singola misura ha un'ulteriore incertezza intorno al valore di y medio. Si può dimostrare che queste due incertezze si combinano, come al solito, quadraticamente:

$$\sigma_t^2(y) = \sigma^2(y) + \sigma_r^2. \quad (49)$$

Facciamo due esempi:

$$\begin{aligned} x = 0.576 \text{ kg} &\Rightarrow \sigma_t(y) = 0.74 \text{ mm} \\ x = 1.000 \text{ kg} &\Rightarrow \sigma_t(y) = 1.1 \text{ mm}. \end{aligned}$$

23.6 Altri tipi di fit

Dato il taglio e le dimensioni di questo scritto, è impossibile affrontare tutti gli aspetti dei fit, o farne una trattazione generale. Si possono incontrare leggi più complicate dell'andamento lineare (casi tipici sono la parabola e l'esponenziale), avere incertezze (veramente) note a priori e generalmente diverse fra di loro, oppure incertezze ignote e diverse fra di loro, ma modellizzabili in qualche modo.

23.7 "Riabilitazione del computer"

Quando il problema diventa più complicato del semplice caso lineare che abbiamo mostrato la trattazione quantitativa rigorosa richiede necessariamente l'uso dei minimi quadrati, possibilmente con l'ausilio di programmi al calcolatore. Anche se alcuni dei casi possono essere trattati a mano, essi necessitano di lunghe elaborazioni e quindi, una volta che si sa bene come comportarsi, soltanto una perdita di tempo. Ad un certo punto diventa inevitabile l'uso del computer, tenendo però conto che:

- devono essere sotto controllo le ipotesi di lavoro del programma ed bisogna essere sicuri che esse corrispondano al caso in questione;
- il programma deve fornire l'incertezza standard dei parametri e il coefficiente di correlazione di ogni coppia di essi; sarebbe opportuno che esso dia anche la deviazione standard σ_r ricavata dai residui, oppure il valore del χ^2 fra curva e punti sperimentali, a seconda che il fit sia stato eseguito con incertezze ignote o note a priori;
- gli eventuali effetti sistematici vanno stimati a parte, ad esempio variando di $\pm 1\sigma$ le costanti di calibrazioni e i fattori di influenza, combinando poi in quadratura (se i vari contributi sono indipendenti) le variazioni propagate sui parametri.

23.8 Rette di taratura?

Molti avranno imparato, oltre alle rette di massima e minima pendenza, le cosiddette *rette di taratura*. Con esse si intendono delle rette “*parallele o convergenti, a seconda della posizione dei punti e dell'entità degli errori, ... che delimitino la massima zona di piano entro la quale sono compresi i punti sperimentali*”⁴⁵. Esse servono a risalire, in modo assolutamente empirico, dalla variabile x alla variabile y , e viceversa, propagando anche le relative incertezze. Nel caso di andamenti non lineari, le rette diventano una specie di striscia a larghezza variabile che segue i punti sperimentali.

Questa procedura di taratura artigianale, che soffre dei problemi del metodo delle rette di massima e minima pendenza, è sostituita, nei metodologie più moderna, dal risultato del fit. Noti i parametri, con le loro incertezze e correlazioni, è possibile passare facilmente da una variabile all'altra. Esempi numerici sono già stati mostrati nel paragrafo 23.5 (punti 8 e 10). In particolare, si faccia attenzione al significato della deviazione standard della variabile di arrivo (punto 10).

⁴⁵M. Severi, “*Introduzione alla sperimentazione fisica*”, Zanichelli, 1985.

24 Note sulla didattica

Probabilmente, chi ha letto attentamente gli ultimi paragrafi sulle applicazioni rimarrà stupito venendo a sapere che, nella maggior parte delle esercitazioni, gli studenti del mio corso non fanno calcoli di incertezze di misura. Questa scelta didattica è basata sugli stessi presupposti a base della teoria di incertezza di misura qui illustrata, riassunta dallo slogan “**imparare dall’esperienza**”, e su alcune considerazioni sulla finalità del corso di esercitazioni di laboratorio.

Ricordiamo che le finalità del laboratorio di Fisica sono molteplici:

1. Prima fra tutte c’è - a mio avviso - l’importanza di verificare di persona delle leggi fisiche, anche se in modo approssimativo e semiquantitativo. Questo avviene, in genere, dopo che tali leggi sono state studiate dal punto di vista teorico, ma è anche proficuo far precedere l’esperienza alla teoria, con grandi vantaggi per l’apprendimento⁴⁶.
2. C’è poi l’aspetto legato allo sviluppo di una certa confidenza e manualità nell’uso degli strumenti di misura e all’apprendimento di tecniche di misura.
3. C’è quindi la parte di interpretazione quantitativa dei risultati e di valutazione delle incertezze, sulla quale abbiamo lungamente parlato.
4. Infine, come sintesi dei punti precedenti, si può arrivare ad una pianificazione qualitativa e quantitativa degli esperimenti.

Alla luce di questi obiettivi e della metodologia di analisi dei dati illustrata, seguono un certo numero di proposte, tutte sperimentate con successo nel mio corso.

- Gli studenti dovrebbero cominciare a frequentare il laboratorio e a eseguire misure il più presto possibile, affrontando inizialmente semplici esperienze, per acquisire le basi della metodologia scientifica (metodo e ordine nell’acquisizione dei dati, tabelle, grafici, prime elaborazioni quantitative, etc.).
- Finché gli studenti non abbiano sperimentato di prima persona, ripetendo gli esperimenti e/o confrontando fra di loro i risultati, quale sia il significato di errore ed incertezza di misura e non abbiano imparato, in parallelo, il linguaggio per trattare le incertezze (quello della probabilità) non dovrebbero subire una teoria astratta su errori e incertezze. Frasi del tipo “dicesi errore ... ” dovrebbero seguire e non precedere la loro diretta esperienza.
- Ne segue che, nelle esperienze iniziali, è da evitare qualsiasi valutazione di incertezze, al di là di quelle che vengono spontanee agli studenti alla luce delle cifre da loro apprezzabili sugli strumenti, della semplice propagazione del numero di cifre significative e della constatazione

⁴⁶A mio giudizio, sapere “tutto” prima di andare in laboratorio può essere poco stimolante in quanto elimina l’“effetto sorpresa”. Trovo spesso utile informare gli studenti soltanto su come fare la misura ma non su quello che verrà fuori, specialmente se il risultato può essere a prima vista controintuitivo.

che, ripetendo le misure, essi ottengono generalmente risultati diversi. Trovo che sia particolarmente stimolante provocare gli studenti con domande del tipo “quanto ci credi?”, magari rafforzate dalla proposta di scommesse coerenti con poste uguali⁴⁷.

- Siccome procedendo in questo modo gli studenti devono fare pochissimi conti, essi hanno più tempo per ripetere le misure, sviluppando l'intuito per le “verosimiglianze”, o per fare altre esperienze, arricchendo il campo della fenomenologia sperimentabile durante il corso.
- In parallelo gli studenti dovrebbero apprendere le basi del calcolo delle probabilità (e non solo una collezione di formulette) e dell'inferenza statistica. A tale scopo sono molto utili semplici esperienze di statistica o analisi di dati simulati, affinché anche l'apprendimento della teoria della probabilità sia visto come risposta a problemi concreti e non un semplice sviluppo matematico.
- Soltanto essi quando hanno acquisito una certa esperienza di laboratorio e il linguaggio della statistica inferenziale (anche nella forma semplificata secondo la traccia che abbiamo qui illustrato) possono apprezzare appieno una sistemazione più formale dei termini di metrologia, delle cause di errori e della valutazione delle incertezze. Sarà allora molto interessante rianalizzare vecchie esperienze alla luce delle nuove conoscenze, come, ad esempio, confrontare le rette calcolate con i minimi quadrati con quelle tracciate a occhio (è anche istruttivo rileggere il quaderno di laboratorio a distanza di mesi!).
- Non è detto che una volta appreso come calcolare le incertezze di misura essi debbano farlo sempre, o farlo sempre nella maniera più rigorosa: alcune esperienze sono complicate e richiederebbero un tempo enorme; in altre sono più importanti gli aspetti fenomenologici e/o strumentali che l'analisi quantitativa rigorosa dei risultati.

Ad esempio, nelle esperienze sui circuiti elettrici l'interesse primario è quello di familiarizzarsi con la strumentazione e con l'influenza degli elementi circuitali, piuttosto che misurare al percento un'amplificazione o uno sfasamento. Per queste esercitazioni, ogni analisi dell'incertezza che vada oltre un controllo della ragionevolezza delle cifre significative può essere addirittura controproducente.

Quello che è importante è che lo studente si senta sicuro che all'occorrenza, magari con giorni di lavoro, consultando documentazione, informandosi in giro ed eventualmente sviluppando appositi programmi al computer, possa affrontare un'analisi delle incertezze condotta in modo professionale.

⁴⁷Ad esempio, dovendo stimare una temperatura ambiente, se dicono $23 \pm 10^\circ\text{C}$ si scommette pro, se dicono $23.0 \pm 0.1^\circ\text{C}$ si scommette contro, e così via finché non si converge ad risultato di indifferenza che definisce un'intervallo di credibilità al 50 %.

25 Conclusioni

Credo che non ci siano migliori conclusioni riguardo la valutazione delle incertezze di misura che riproporre quanto affermato in proposito dalla *Guida ISO*:

“Although this Guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty, and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value”.

Per quanto riguarda l’insegnamento, aggiungerei che *per gli studenti la valutazione delle incertezze non dovrebbe essere una vessazione*, essendo molteplici le finalità delle esercitazioni di laboratorio.

26 Nota bibliografica

Per una introduzione generale sulla probabilità soggettiva si può consultare la raccolta di articoli pubblicata sotto il titolo di *Probabilità e giochi d'azzardo*[1]. Un buon libro che illustra per esteso i molti aspetti del “ragionare bayesiano” è quello di C. Howson and P. Urbach[2]. Su di esso è basato l'articolo su *Nature* degli stessi autori[3]. Passando a testi veri e propri, si raccomanda quello di Romano Scozzafava[5] e, per chi vuole approfondire lo studio, il classico due volumi di Bruno de Finetti[4]. Un'introduzione meno formale[6, 7], ma più vicina alle applicazioni sulle incertezze di misura, è disponibile sotto forma di dispense presso il Dipartimento di Fisica.

Per quanto riguarda l'introduzione alla metodologia di laboratorio consistente con quanto illustrato qui si rimanda ad altre dispense[8], mentre per chi vuole saperne di più sulla teoria sulle incertezze di misure qui presentata si consiglia la lettura di due note[9, 10] disponibili anche su internet. Per le formule relative ai fit si consigliano altre dispense[11] o il Barlow[12]. Infine, si noti come non ci siano ulteriori referenze riguardo ad alcune delle tecniche presentate (come l'analisi grafica degli andamenti lineari e la trattazione semplificata delle incertezze dovute ad errori sistematici, compreso il caso dei fit) in quanto questa è la prima volta che esse compaiono in forma scritta.

La *Guida* ISO[13] a cui abbiamo fatto più volte riferimento è consultabile presso la biblioteca di Dipartimento.

Infine, chi è interessato al tipo di esercitazioni che svolgono gli studenti del mio corso può procurarsi la raccolta dei “foglietti”[14] che vengono distribuiti agli studenti.

Riferimenti bibliografici

- [1] D. Costantini e P. Monari (curatori), *“Probabilità e giochi d'azzardo”*, Franco Muzzio Editore, 1996. Si vedano, in particolare, i contributi di U. Garibaldi, L. Vianelli (interessante anche i molti rimandi bibliografici) e G. Dall'Aglio.
- [2] C. Howson and P. Urbach, *“Scientific reasoning - the Bayesian approach”*, Open Court, Chicago and La Salle, 1993.
- [3] C. Howson and P. Urbach, *“Bayesian reasoning in science”*, Nature, Vol. 350, 4 April 1991, pag. 371.
- [4] B. de Finetti, *“Theory of probability”*, J. Wiley & Sons, 1974.
- [5] R. Scozzafava, *“La probabilità soggettiva e le sue applicazioni”*, Masson, editoriale Veschi, Roma, 1996. Per una introduzione a livello più elementare si raccomanda anche *“Primi passi in probabilità e statistica”*, Decibel/Zanichelli, 1996.
- [6] G. D'Agostini, *“Probabilità e incertezze di misura - Parte 1: dal concetto di probabilità ai problemi di probabilità inversa”*, Dispense del Dipartimento di Fisica, Dicembre 1997.

- [7] G. D'Agostini, "*Probabilità e incertezze di misura - Parte 2: variabili casuali*", Dispense del Dipartimento di Fisica (disponibile prossimamente).
- [8] G. D'Agostini, "*Laboratorio virtuale - Introduzione alla metodologia di laboratorio*", Dispense del Dipartimento di Fisica, Ottobre 1997.
- [9] G. D'Agostini, "*Probability and measurement uncertainty in Physics - a Bayesian primer*", Nota Interna del Dipartimento di Fisica 1070, Dicembre 1995 (anche DESY-95-242 e hep-ph/9512295). (http://www-zeus.desy.de/zeus_papers/desy_papers.html).
- [10] G. D'Agostini, "*A theory of measurement uncertainty based on conditional probability*", Nota Interna del Dipartimento di Fisica 1079, Novembre 1996 (physics/9611016) (<http://xxx.lanl.gov/ps/physics/9611016>).
- [11] G. D'Agostini, "*Introduzione alla elaborazione statistica dei dati sperimentali*", Dispense del Dipartimento di Fisica, Ottobre 1994 (possibilmente la versione con "contenuto ridotto", essendo la prima parte sostituita dalle dispense [6] e [7]).
- [12] R.J. Barlow, "*Statistics*", John Wiley & Sons, 1989.
- [13] International Organization for Standardization (ISO), "*Guide to the expression of uncertainty in measurement*", Geneva, Switzerland, 1993.
- [14] G. D'Agostini, "*Raccolta dei promemoria per le esercitazioni di laboratorio del corso di 'Fisichetta' per Chimici (M-Z)*", Dispense del Dipartimento di Fisica, Maggio 1998.

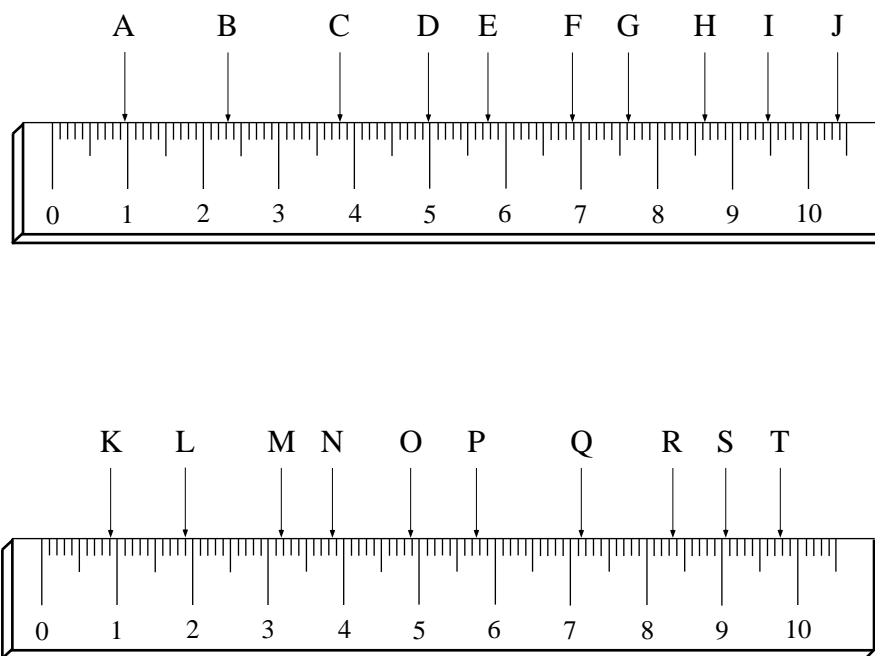


Figura 12: Esercizio di lettura dei decimi di divisione. Le soluzioni sono riportate nel paragrafo F.

A Capacità di interpolazione fra le tacche e incertezza di lettura

Mostriamo ora una semplicissima misura legata alla questione della liceità di leggere il valore indicato da uno strumento stimando un valore fra le tacche di una scala analogica. Uno strumento che si presta bene per tale scopo è il *calibro ventesimale*.

L'esperienza consiste nel posizionare a caso il cursore e nello stimare al meglio il valore misurato coprendo il nonio (eccetto la prima tacca, ovviamente). Il valore *stimato* viene riportato in una tabella. Servendosi poi dell'informazione del nonio si scrive il valore *letto* sulla seconda colonna (è preferibile che la lettura del nonio sia effettuata da un'altra persona). Le differenze fra i due numeri, opportunamente elaborate statisticamente a tempo debito, forniranno una stima quantitativa dell'incertezza di interpolazione fra le tacche. La tabella 4 riporta i dati sperimentali ottenuti da studenti alla loro *prima* esercitazione (in qualche modo i risultati daranno una stima pessimistica della capacità di interpolazione).

La figura 12 permette al lettore di valutare rapidamente la propria capacità di interpolazione senza dover eseguire l'esperienza con un calibro vero.

A.1 Errore ed incertezza di misura (commento all'esperienza di interpolazione fra le tacche)

L'esperienza appena descritta illustra molto bene i concetti di errore e di incertezza di misura, già introdotti a livello intuitivo e sui quali si tornerà in

i	L.T.			P.P.		
	l_S	l_N	Δ	l_S	l_N	Δ
1	18.35	18.45	-0.10	21.90	21.80	+0.10
2	46.00	46.10	-0.10	34.15	34.20	-0.05
3	37.15	37.10	+0.05	72.20	72.25	-0.05
4	28.20	28.25	-0.05	15.25	15.30	-0.05
5	5.60	5.65	-0.05	27.30	27.40	-0.10
6	34.20	34.30	-0.10	56.50	56.55	-0.05
7	48.90	48.75	+0.15	34.40	34.45	-0.05
8	16.75	16.85	-0.10	9.40	9.55	-0.15
9	23.45	23.30	+0.15	47.25	47.40	-0.15
10	72.20	72.30	-0.10	18.90	18.80	+0.10
i	C.P.			S.M.		
	l_S	l_N	Δ	l_S	l_N	Δ
1	40.85	40.85	0.00	30.75	30.70	+0.05
2	62.15	62.10	+0.05	44.95	44.90	+0.05
3	23.20	23.15	+0.05	65.60	65.50	+0.10
4	12.25	12.35	-0.10	33.70	33.80	-0.10
5	59.20	59.30	-0.10	26.90	26.95	-0.05
6	43.70	43.65	+0.05	74.00	74.10	-0.10
7	23.20	23.30	-0.10	7.90	8.00	-0.10
8	37.00	37.10	-0.10	22.75	22.65	+0.10
9	47.35	47.20	+0.15	66.70	66.60	+0.10
10	29.75	29.80	-0.05	43.25	43.35	-0.10
i	C.U.			L.P.		
	l_S	l_N	Δ	l_S	l_N	Δ
1	49.15	49.15	0.00	18.95	18.75	+0.20
2	20.35	20.50	-0.15	15.50	15.55	-0.05
3	22.00	22.00	0.00	27.10	27.20	-0.10
4	29.55	29.80	-0.05	30.50	30.55	-0.05
5	38.15	38.20	-0.05	18.50	19.00	-0.50
6	24.45	24.45	0.00	13.00	13.00	0.00
7	25.00	25.00	0.00	19.20	19.30	-0.10
8	28.85	28.75	+0.10	26.55	26.55	0.00
9	18.50	18.60	-0.10	17.95	17.95	0.00
10	34.25	34.20	+0.05	24.00	24.05	-0.05
i	F.N.			A.M.		
	l_S	l_N	Δ	l_S	l_N	Δ
1	23.75	23.65	+0.10	35.15	35.20	-0.05
2	28.50	28.55	-0.05	41.55	41.45	+0.10
3	33.95	33.95	0.00	39.95	39.90	+0.05
4	13.80	13.70	+0.10	47.35	47.40	-0.05
5	14.85	14.80	+0.05	51.05	51.15	-0.10
6	21.15	21.20	-0.05	63.85	63.85	0.00
7	20.15	20.25	-0.10	35.00	35.05	-0.05
8	7.85	7.75	+0.10	40.95	40.95	0.00
9	19.45	19.45	0.00	87.00	87.10	-0.10
10	18.05	18.00	+0.05	42.35	42.40	-0.05

Tabella 4: Dati sperimentali relativi alla capacità di interpolazione fra le tacche di alcuni studenti (identificati dalle iniziali del nome) alla prima esercitazione di laboratorio. l_S è il valore stimato interpolando a occhio e l_N il valore letto sul nonio. Tutti i valori sono in millimetri. Il valore in grassetto è sospetto: sembra più un errore di scrittura che di valutazione.

termini più formali nel seguito.

La differenza fra i valori stimati e quelli letti al nonio rappresenta l'errore di stima che si commette di volta in volta, in quanto il valore letto al nonio è in buona approssimazione "quello vero". Dalla conoscenza di questi *errori tipici* (da meglio definire) si può risalire all'incertezza della misura derivante dalla sola stima. Consideriamo alcuni casi particolarmente istruttivi.

1. Supponiamo di sapere che l'ipotetica persona A (nessuna di quelle della tabella 4) si sbaglia "in media" di 0.02 cm sia in più che in meno. Se in una misura successiva effettuata in condizioni analoghe lui stimerà 23.54 cm, quanto si crederà a questa affermazione? Ragionevolmente si tenderà a credergli "entro 0.02 cm".
2. Immaginiamo invece che l'altrettanto ipotetica persona B tenda *in media* a sovrastimare le lunghezze di 0.02 cm e che la dispersione degli errori intorno a questo errore sistematico medio sia soltanto di 0.01 cm. Se costui stimerà 15.67 cm chiaramente si tenderà a credere che il valore vero sia "entro 0.01 cm" intorno a 15.65 cm, avendo corretto la sua stima per l'errore medio.
3. Supponiamo che sei persone che si comportino come A stimino 18.33, 18.36, 18.34, 18.31, 18.34, 18.37 cm. Chiaramente, non essendoci nessun motivo per ritenere qualcuna di queste stime più o meno giusta delle altre si tenderà a pensare che "in medio stat virtus". E fin qui va bene. Però se si prova a stimare l'ampiezza dell'intervallo entro cui si può credere che il valore vero sia compreso, l'intuizione potrebbe dare una risposta ("entro 0.03 cm") ben lontana da quella corretta ("entro 0.008 cm"). Per arrivare a queste conclusioni bisognerà prima aver imparato che le stime medie effettuate indipendentemente da 6 persone equivalenti sbagliano mediamente di un fattore $1/\sqrt{6}$ in meno della stima di ciascuna delle persona.
4. Supponiamo di porci nelle condizioni 1., ma di sapere che lo strumento usato non è calibrato perfettamente, ma che può essere scalibrato mediamente dello 0.5 %, in più o in meno. Ciò significa che, anche se la lettura fosse perfetta, si crederà al risultato entro 0.12 cm. Quindi l'incertezza dovuta all'errore di lettura di 0.02 mm diventa trascurabile. E' anche chiaro che non ha nessun senso fare medie fra più letture o sforzarsi di leggere al meglio quando si è coscienti di una incertezza "inevitabile" di questo tipo⁴⁸. Se questa incertezza è intollerabile per le misure che si vogliono fare bisogna ricalibrare lo strumento o procurarsene un altro migliore.

B Tempo di reazione e misure di cronometraggio

Semplici misure complementari a quelle descritte nel paragrafo precedente permettono di valutare l'errore che lo sperimentatore introduce nelle misure

⁴⁸Questa limitazione cade se si effettuano misure su grandezze fisiche aventi valori prossimi fra di loro e siamo interessati soltanto alle loro differenze.

Tempi di riflessi (ms): G.D.									
199	186	189	190	199	193	156	229	196	263
201	176	184	192	176	186	261	196	167	174
181	202	202	181	179	187	209	195	154	173
Tempi di riflessi (ms): D.P.									
188	233	245	195	165	241	201	185	197	206
243	194	194	187	178	187	299	208	243	199
213	199	192	200	196	219	187	176	189	193
Periodo del pendolo (s): G.D.									
1.45	1.46	1.40	1.51	1.44	1.44	1.49	1.47	1.49	1.46
1.45	1.45	1.46	1.46	1.44	1.46	1.50	1.39	1.50	1.48
1.48	1.42	1.46	1.46	1.50	1.50	1.44	1.47	1.47	1.46
Periodo del pendolo (s): D.P.									
1.57	1.44	1.45	1.45	1.41	1.41	1.46	1.47	1.42	1.43
1.40	1.40	1.48	1.46	1.40	1.48	1.46	1.43	1.47	1.43
1.48	1.48	1.45	1.47	1.47	1.46	1.41	1.44	1.41	1.45

Tabella 5: Due situazioni di misure manuali di tempo: tempo di riflesso e cronometraggio applicato a misure del periodo del pendolo. La costanza del periodo durante le due serie di misure è stata monitorata con un cronometro elettronico, il quale ha sempre misurato 1.454 s.

manuali di tempo. Queste permetteranno di valutare le incertezze su misure eseguite in condizioni analoghe.

Il primo esperimento consiste nel misurare il tempo di reazione fra la comparsa di un segnale luminoso sul monitor di un computer e l'istante in cui viene premuto un tasto. Il secondo consiste nel misurare molte volte il periodo di un pendolo che effettua piccole oscillazioni. La costanza del periodo durante le misure viene controllata con un sistema elettronico. I dati sono riportati in tabelle 5 e si riferiscono a due persone.

Chiaramente le situazioni di tempo di reazione è diversa da quella di cronometraggio. Mentre nella prima c'è un ritardo medio rispetto allo stimolo con fluttuazioni che a volte possono diventare anche grandi a causa di piccole distrazioni, nel secondo caso si hanno fluttuazioni positive e negative rispetto al tempo "vero". Inoltre nel secondo caso si combina una fluttuazione sullo *start* e una sullo *stop*.

C Distribuzioni triangolari

Quando la variabile casuale è definita in un certo intervallo, ma ci sono delle ragioni per ritenere che i gradi di fiducia decrescano linearmente dal centro (x_0) verso gli estremi si ha la cosiddetta *distribuzione triangolare* (o di *Simpson*). Anch'essa è molto utile per il calcolo delle incertezza di misura, in quanto in molte circostanze, questo modello può essere più realistico di quello uniforme.

Se la variabile può verificarsi con sicurezza nell'intervallo $x_0 \pm \Delta$, il valore

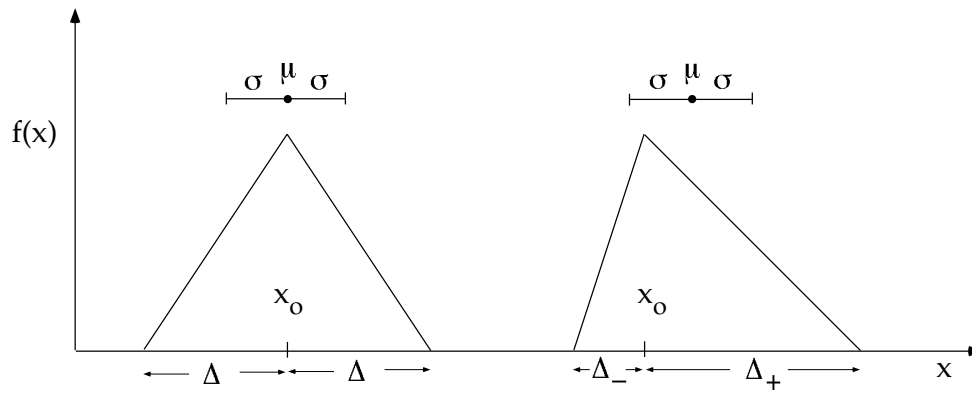


Figura 13: Esempio di distribuzione triangolare simmetrica e asimmetrica.

atteso è x_0 e la deviazione standard vale

$$\sigma = \frac{\Delta}{\sqrt{6}} = 0.41 \Delta \approx 0.4 \Delta. \quad (50)$$

I conti vengono lasciati per esercizio.

C'è un'altra distribuzione triangolare che può avere interessi pratici: il valore al quale si crede di più è x_0 e i gradi di fiducia decrescono linearmente verso gli estremi a e b , ma x_0 non corrisponde con il centro dell'intervallo. Chiamando

$$\Delta_+ = b - x_0 \quad (51)$$

$$\Delta_- = x_0 - a, \quad (52)$$

si ottengono i seguenti risultati per valore atteso e varianza

$$E(X) = x_0 + \frac{\Delta_+ - \Delta_-}{3}. \quad (53)$$

$$\sigma^2 = \frac{\Delta_+^2 + \Delta_-^2 + \Delta_+ \Delta_-}{18}, \quad (54)$$

le cui dimostrazioni vengono lasciate come esercizio.

Quando $\Delta_+ = \Delta_- = \Delta$, si riottengono le formule del caso precedente. E' interessante inoltre notare che, se la differenza fra Δ_+ e Δ_- è piccola si ottiene una deviazione standard circa pari a quella ottenibile con un valore intermedio fra i due:

$$\sigma \approx \frac{\bar{\Delta}}{\sqrt{6}} = \frac{1}{\sqrt{6}} \frac{(\Delta_+ + \Delta_-)}{2},$$

come può essere verificato mediante una espansione in serie della (54).

Un sottocaso particolare della triangolare asimmetrica è quando uno dei due Δ è nullo ed il triangolo diventa rettangolo. Questa distribuzione può modellizzare gradi di fiducia che decrescono linearmente in un certo intervallo. Ad esempio, ci possono essere delle ragioni per ritenere che una grandezza definita non negativa valga molto verosimilmente 0 e che comunque non debba eccedere un certo valore a : si ottiene una previsione di $a/3 \pm a/(3\sqrt{2})$.

D Teorema del limite centrale

Ricordiamo brevemente quanto afferma il teorema del limite centrale: la combinazione lineare di n variabili indipendenti ($Y = \sum_{i=1}^n \alpha_i X_i$) tende ad essere distribuita normalmente, con $\mu_y = \sum_{i=1}^n \alpha_i \mu_{x_i}$ e $\sigma_y^2 = \sum_{i=1}^n \alpha_i^2 \sigma_{x_i}^2$, quando $n \rightarrow \infty$, se: a) le μ_{x_i} sono finite; b) $\alpha_i^2 \sigma_{x_i}^2 \ll \sigma_y^2$ per ciascuna variabile X_i non distribuita normalmente. (Un altro modo di esprimere la seconda condizione è che le σ_{x_i} devono essere dello stesso ordine di grandezza e solo variabili già distribuite normalmente possono fare eccezione; non è invece necessario che siano dello stesso ordine di grandezza anche le μ_{x_i} , in quanto eventuali costanti additive non influenzano la distribuzione).

Si noti che il teorema non dice niente sul numero minimo di componenti necessarie affinché esso sia valido. Dipende dalle distribuzioni delle X_i . Ad esempio, nel caso di variabili aventi la stessa distribuzione uniforme è sufficiente $n \geq 4 - 5$ affinché l'approssimazione sia ragionevole (è molto convincente la figura 3, che può essere reinterpretata, a parte un fattore n come la distribuzione della media). Se le distribuzioni hanno invece un massimo centrale sono sufficienti 2-3 componenti (si noti come 2 triangolari di uguale Δ equivalgono a 4 uniformi). Siccome nelle applicazioni alle incertezze di misura le variabili di partenza sono spesso “quasi gaussiane” la convergenza è generalmente molto rapida⁴⁹.

E Deviazione standard della media aritmetica

Se si hanno n variabili indipendenti X_i , tutte con la stessa distribuzione di probabilità, avente valore atteso μ e deviazione standard $\sigma(X)$, la nuova variabile casuale media aritmetica

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

ha valore atteso μ e deviazione standard

$$\sigma(\bar{X}_n) = \frac{\sigma(X)}{\sqrt{n}}.$$

Questo è un risultato generale che segue dalla *proprietà della varianza* di una combinazione lineare di variabili casuali.

Inoltre, quando n è “abbastanza grande”, la distribuzione della media è normale, in virtù del *teorema del limite centrale*.

Si possono pensare molti semplici esperimenti per “provare” queste leggi della probabilità (in realtà non si “prova” un bel niente, al più si osservano risultati compatibili con esse). Ne proponiamo due.

E.1 Stime di segmenti alla lavagna

Questa prima esperienza è molto semplice ed istruttiva. Si tracciano dei segmenti alla lavagna, orientati disordinatamente e di varie lunghezze (ad esempio

⁴⁹Come controesempio si immagini di avere un milione di variabili indipendenti, aventi distribuzione poissoniana con $\lambda = 10^{-6}$. La somma delle variabili è ancora poissoniana, con $\lambda = 1$, ben lontana da una gaussiana (in questo caso servono 10-20 milioni di contributi).

da 10 a 50 cm circa). Affianco a ciascuno si scrive la sua lunghezza, misurata con un righello. Successivamente si disegna un segmento di lunghezza incognita e si chiede agli studenti di indovinarne la lunghezza e di registrare il valore stimato sul quaderno. Si istogrammano quindi alla lavagna i valori ottenuti se ne calcola la media (o si stima ad occhio il baricentro della distribuzione statistica). Ripetendo l'esperimento più volte si noterà la grande differenza fra gli errori tipici del singolo studente e quello della media della classe. Si può anche notare un miglioramento della capacità di previsione degli studenti nel corso delle misure, come risultato del processo di autocorrezione.

Sarà poi simpatico raccontare che questo era il metodo usato secoli fa dagli artiglieri per stimare la distanza del bersaglio, chiedendo l'opinioni dei soldati e mediando le stime.

Questo esercizio è molto utile per introdurre il discorso sull'inversione della probabilità proponendo domande del tipo: "quanto si può credere al valore stimato da uno studente?"; "quanto si può credere al valore attenuato mediando tutte le stime?"

E.2 Media di valori distribuiti uniformemente

Un'altra esperienza molto semplice consiste nell'utilizzare i numeri casuali della calcolatrice (tasto RAN#, RNDM, o simili). Si possono estrarre n numeri casuali (ad esempio $n = 6$ già può essere sufficiente) e mediarli. Ripetendo tante volte (50-100) questa procedura si può fare alla fine la medie delle medie (per la quale ci si aspetta il valore 0.5) e la deviazione standard delle medie (per la quale ci si aspetta $0.29/\sqrt{n}$). Con una classe numerosa di studenti si possono mettere su un unico istogramma tutte le medie: la curva somiglierà ad una gaussiana⁵⁰.

F Soluzione dell'esercizio di interpolazione fra le tacche

Questi sono i valori riletti, mediante lente d'ingrandimento dotata di scala graduata, su una stampa di qualità standard: A) 0.960; B) 2.320; C) 3.805; D) 4.980; E) 5.770; F) 6.880; G) 7.620; H) 8.630; I) 9.460; J) 10.370; K) 0.915; L) 1.900; M) 3.170; N) 3.850; O) 4.885; P) 5.755; Q) 7.140; R) 8.350; S) 9.050; T) 9.775.

Effetti dovuti a quantizzazioni di stampa potrebbero causare piccoli scarti (± 0.005 cm) da questi valori che però non alterano la sostanza dell'esercizio.

Quanto vale il tuo errore quadratico medio di lettura?

⁵⁰Qualcuno potrebbe parlare di "dimostrazione sperimentale" della proprietà delle medie. Ma a rigore questa non è una dimostrazione. Meglio parlare di "accordo con le previsioni". E' vero però che questo accordo serve a rafforzare la convinzione che le valutazioni a priori siano corrette.