Sulla metodologia della misura e sul concetto di probabilità: una introduzione critica di alcuni aspetti epistemologici

Paolo Agnoli, Dicembre 2002

INDICE

PREMESSA

Questo scritto è stato elaborato come tesina integrativa alla prova di esame del corso di Bioetica tenuto all'Università Roma 2 dalla prof.ssa Gigliotti e dal dott. Aportone nell'anno accademico 2002-2003.
Nel quadro della tesi di fondo del volume di riferimento del primo modulo del corso (Vineis, 1999), ovvero quella dell'insostenibilità di una contrapposizione tra "questioni di fatto" e "questioni di valore", questo scritto vuole introdurre criticamente alcuni attuali sviluppi nel campo della metodologia della misura. L'impossibilità di fatto di separare interamente un problema conoscitivo da un problema etico può rendere a mio avviso utile una simile introduzione anche per un successivo obiettivo (lasciato eventualmente ad una ulteriore tesina, e peraltro ampiamente discusso nel libro di Vineis) più *direttamente* pertinente alla bioetica, quello di analizzare e discutere le azioni reciproche esistenti tra i modelli epistemologici ed i modelli etici.

INTRODUZIONE

Weights and Measures may be ranked among the necessaries of life to every individual of human society. They enter into the economical arrangements and daily concerns of every family. They are necessary to every occupation of human industry; to the distribution and security of every species of property; to every transaction of trade and commerce; to the labors of the husbandman; to the ingenuity of the artificer; to the studies of the philosopher; to the researches of the antiquarian; to the navigation of the mariner; and the marches of the soldier; to all the exchanges of peace, and all the operations of war. The knowledge of them, as in established use, is among the first elements of education, and is often learned by those who learn nothing else, not even to read and write. This knowledge is riveted in the memory by the habitual application of it to the employments of men throughout life.

John Quincy Adams, Report to the Congress, 1821

Science and technology have advanced in more than direct ratio to the ability of men to contrive methods by which phenomena which otherwise could be known only through the senses of touch, hearing, taste, and smell have been brought within the range of visual recognition and measurement and thus become subjects to that logical symbolization without which rational thought and analysis are impossible.

William N. Ivins, Jr.

La metodologia della misura come la intendiamo attualmente è stata sviluppata in un periodo relativamente recente, ma metodi primitivi di misurazione sono stati da tempi molto antichi alla base della capacità dell'uomo di interagire e soprattutto di *incidere sul mondo circostante*. Si può dire che da sempre la misura è stato uno dei procedimenti della vita quotidiana necessari alla sopravvivenza e alle relazioni sociali. L'utilizzo delle provviste, il raccolto, la caccia, il baratto, la definizione del suolo da coltivare etc…, sono tutte attività in cui è facile ritrovare la necessità di misurare.

Personalmente credo che se si vuole ricercare una spiegazione unificante al progresso tecnologico e scientifico che le diverse collettività umane hanno realizzato nel corso della storia, tale spiegazione va proprio individuata nel perseguimento, da parte dell'uomo *inteso in senso sociale*, dell'espansione della propria capacità di incidere su quanto lo circonda, più che in una disinteressata curiosità individuale. La costruzione di imbarcazioni, ad esempio, dagli antichi tipi di canoe alle navette spaziali, che cos'è se non la ricerca della soluzione al problema di estendere la mobilità dell'uomo ad un dominio diverso dal suo, decisivo *dal punto di vista economico* per l'appropriazione di beni altrimenti irraggiungibili?

Non è affatto raro, anche ai nostri giorni, imbattersi in quello che secondo me è soltanto un pregiudizio ben radicato: *la scienza è una dimensione pura, la tecnica è buona o cattiva* secondo l'uso che se ne fa. Non credo che ciò sia vero. La tecnica non è l'applicazione della scienza: la scienza non inaugura se stessa per contemplare il mondo, la scienza inaugura se stessa per trasformare il mondo e quindi l'intenzione tecnica è già nello sguardo scientifico. La scienza si compone di *fare*(intervenire) almeno tanto quanto di *sapere*(rappresentare)(vedi anche Galimberti, 1999).

Gli uomini hanno dovuto affrontare problemi pratici sin dagli albori, cominciando con il più importante di tutti: rimanere vivi. *Gli stessi organismi viventi, nella loro grande diversità, sono differenti soluzioni al problema di sopravvivere in un ambiente che cambia*.

L'uomo può vivere solo grazie alla sua azione che da subito diventa ideativa e progettuale: la tecnica è l'essenza dell'uomo.

> Siamo fratelli e sorelle delle nostre macchine. Menti e utensili si sono affinati tra loro sin dal momento in cui la pietra di un cavernicolo abituato a rovistare tra gli scarti si ruppe in maniera netta, consegnando così il primo bordo tagliente nelle mani di un cacciatore. La scaglia di ossidiana e il chip di silicio sono colpiti dalla luce dello stesso fuoco da bivacco che si è tramandato di mano in mano sin dai primordi della mente umana.(Dyson,2000, 12)

Certo, la tecnologia a volte crea problemi più grossi di quelli che risolve. Un bellissimo libro dello storico Edward Tenner(Tenner, 1996) è una lista di soluzioni con effetti secondari disastrosi e progetti andati completamente storti. Ma anche Tenner sarebbe d'accordo con il fisico e divulgatore scientifico Alan Cromer quando questi afferma:

> E' la produzione di utensili, piuttosto che il bipedismo, a distinguere gli esseri umani dagli animali.[…]Un'attività così complessa come la manifattura di utensili, che deve essere tramandata di generazione in generazione, richiede come condizione preliminare lo sviluppo di una caratteristica distintiva dell'uomo ancora più fondamentale: la cultura.(Cromer, 1993, 67)

Ma cosa si deve intendere qui, in prima istanza, per *cultura*? La comune premessa a tutti i differenti sforzi creativi dell'uomo é stata sempre quella di *osservare il mondo circostante e conservare e gestire le informazioni considerate pertinenti allo scopo prefissato.*

La metodologia della misura, oltre ad un'ovvia importanza pratica, presenta quindi implicazioni concettualmente fondamentali nei processi conoscitivi. Il processo attraverso cui si sono formati i concetti sulle misure è una componente molto importante dell'evoluzione delle rappresentazioni umane del mondo, della formazione dei sistemi di classificazione e dei concetti astratti.

Forse non è inutile ricordare che in greco *Logos* significa anche proporzione, *legge dei rapporti*. Ed è la ricerca di queste leggi che apre alla misura, *ovvero alla sapienza*. La proporzione, la misura del visibile apre per i Greci, se così posso dire, alla misura dello spirito.

Si può quindi davvero affermare che:

La nozione di misura, il modo di intenderla, perfino la sua concreta grandezza, sono tutte categorie fondamentali del pensiero umano.(Kula, 1987, 268)

Questo lavoro presenta una breve introduzione alla metodologia della misura attualmente in uso tra gli scienziati e i tecnici sperimentali e si sofferma su alcune considerazioni critiche che vogliono servire a delineare gli sviluppi oggi in atto, *sviluppi che stanno proponendo un vero e proprio cambiamento di "paradigma" teoretico*. E credo sia del tutto inutile sottolineare il potere euristico di qualsiasi controversia scientifica.
Cercherò in particolare di chiarire come la teoria standard di trattazione degli errori stia evolvendo sulla base di una critica al concetto convenzionale stesso di probabilità e alla adozione della cosiddetta *statistica bayesiana*. Questa statistica, come vedremo, è basata sull'idea intuitiva che la probabilità quantifica il grado di fiducia attribuito all'occorrenza di un evento. Le "credenze" fanno parte naturale del "fare" scienza e l'ammettere che esistono non necessariamente pregiudica la percepita oggettività delle scienze stesse.

In other words, one needs only to look closely at how frontier science makes progress, instead of seeking refuge in an idealized concept of objectivity[…]My preferred motto on this matter is *"no one should be allowed to speak about objectivity unless he has had 10-20 years working experience in frontier science, economics, or any other applied field"*[…]the statistician D. Berry […] has amused himself by counting how many times Hawking uses 'belief', 'to believe', or synonyms, in his *'A brief history of time'*. The book could have been entitled *'A brief history of beliefs'*, pointed out Berry in his talk.(D'Agostini, July 1999, 123)

Voglio chiarire inoltre che questo lavoro non si soffermerà sui problemi posti dalla meccanica quantistica al processo ed ai protocolli della misura. Piuttosto cercherò di sottolineare un'analogia che secondo me deve essere messa in luce fra l'atteggiamento epistemologico nei confronti dell'osservazione scientifica suggerito dal principio di indeterminazione di Heisenberg, per quanto riguarda il mondo atomico, e l'attuale situazione esistente a scale macroscopiche. E' noto che l'immagine del mondo della fisica moderna è cambiata con l'introduzione del principio di indeterminazione. Questo lavoro vuole mettere in luce che il valore *di una qualsiasi operazione di misura* è caratterizzato da una più o meno grande *indeterminazione* rappresentata dagli errori di misura, e che una corretta metodologia può solo consentire di *conoscere statisticamente il valore dell'indeterminazione da associare ad esso*. Le conseguenze epistemologiche che scaturiscono dalla meccanica quantistica hanno quindi, per così dire, *valore generale*.
Le pubblicazioni cui farò riferimento sono sempre citate durante l'esposizione e riportate nella bibliografia, ma vorrei qui esplicitamente segnalare i lavori del fisico prof. Giulio D'Agostini(ed il sito http://zeual1.roma1.infn.it/~agostini/prob+stat.html da lui gestito), che sono stati per me di particolare riferimento per quanto riguarda la discussione sugli sviluppi più recenti

1. CHE COSA SIGNIFICA MISURARE?

Measurement demands some one-one relations between the numbers and magnitudes in question-a relation which may be direct or indirect, important or trivial, according to circumstances.
Bertrand Russell

In un senso molto generale potremmo dare (vedi Campbell 1957, Mandel 1964, Toulmin, 1960) la seguente definizione: *la misura è l'assegnazione di numerali per rappresentare delle proprietà.* In un senso più ristretto qui tratterò principalmente il concetto di misura nelle scienze fisiche,

comprendendo in questa categoria in ogni caso, le applicazioni tecnologiche e i vari campi dell'ingegneria.
Normalmente con fenomeno fisico si suole indicare qualsiasi oggetto, fatto o avvenimento esterno percepito o osservato direttamente, oppure per mezzo di dispositivi particolari. Più precisamente

> un *fenomeno* è una variazione dello stato di cose che ci circonda e che i nostri sensi, o direttamente o per mezzo di strumenti, ci permettono di *osservare*; è quindi una *transizione* da uno stato diciamo *A* a uno *B* in qualche cosa diverso da *A*.(Bernardini, 1974, 4)

Si suole anche supporre che la conoscenza della natura, cioè del mondo esterno, possa essere oggettiva, ovvero indipendente dalla persona che la acquisisce.

> Il poter descrivere con una correlazione di cause-effetti la transizione fra *A* e *B* è il primo passo necessario per comprendere un fenomeno. Naturalmente la descrizione deve implicare solo elementi di giudizio *oggettivi* inerenti al fenomeno, ed essere indipendente dalle caratteristiche fisiologiche dell'osservatore. Inoltre deve potersi esprimere in termini razionali. Il modo più semplice ed immediato, se non l'unico, per poter far questo, è quello di sostituire alle parole dei numeri ossia *misurare* ogni cosa che possa secondo il nostro giudizio partecipare in modo determinante al fenomeno che interessa.(Ibidem)

Che i metodi per conseguire la conoscenza debbono avere questo carattere di oggettività discende da un postulato fondamentale dell'indagine scientifica (postulato di invarianza spazio-temporale) che afferma che i fenomeni naturali sono indipendenti, a parità di condizioni, dal luogo e dal momento in cui vengono osservati(vedi per esempio Severi, 1986). E quindi un'esperienza correttamente eseguita e descritta oggi, deve esser domani sempre riproducibile e dare sempre, nei limiti degli errori di osservazione, lo stesso risultato.

Presupporre questa *oggettività* non è certamente scontato: basti pensare che l'osservatore è parte attiva nel processo conoscitivo, *con tutto il suo complesso di informazioni preesistenti* perché derivanti da esperienze precedenti. Inoltre si può evidentemente supporre che l'osservazione della realtà modifichi la realtà stessa, come anche che l'acquisizione di una mole sufficiente di informazioni richieda del tempo, e non è certo detto che, mentre si svolge questa indagine, tutto rimanga perfettamente costante.

Per ora (l'argomento sarà però ripreso più avanti) prendiamo per buono questo presupposto.

Si può schematizzare allora, anche se in prima approssimazione, la metodologia della Fisica nel modo seguente:

- è necessario inizialmente individuare o definire il fenomeno che si vuole studiare;
- questo risulta descritto da un certo numero di sue caratteristiche, dette *grandezze fisiche*(p.e. lunghezza, massa, tempo, forza, velocità, densità, temperatura, carica elettrica), ognuna delle quali deve potersi valutare quantitativamente per mezzo di operazioni di confronto con una grandezza ad essa omogenea, assunta come unitaria. Tali operazioni di confronto si chiamano *operazioni di misura* ed i risultati ottenuti si dicono *misure*. Per definire un dato fenomeno, quindi, occorre anche individuare (selezionare)le grandezze necessarie a descriverlo;
- le misure effettuate, opportunamente elaborate, forniscono le informazioni mediante le quali si possono determinare le modalità con cui ogni grandezza, nell'ambito di quel fenomeno, è legata alle altre;
- in questo modo si giunge a determinare le relazioni esistenti fra le grandezze che intervengono in modo essenziale in un fenomeno, stabilendo fra esse dei rapporti, quantitativi, di causa ed effetto. Quando queste relazioni sono legittimamente estendibili a tutta una classe di fenomeni (per es. alla generica caduta di un grave) si enunciano in forma generale e si dicono *leggi*.

Questo modo di rappresentare la metodologia della Fisica è quello di norma ancora oggi offerto nei testi scolastici di scienza. Voglio qui anticipare che è, almeno per quanto riguarda l'ultimo punto, molto discusso il fatto che nella pratica il procedere scientifico si comporti in maniera così lineare.

The road from scientific law to scientific measurement can rarely be traveled in the reverse direction. To discover quantitative regularity one must normally know what regularity one is seeking and one's instruments must be designed accordingly; even then nature may not yield consistent or generalizable results without a struggle.(Kuhn, 1961, 189-190)

Più volte riporterò in questo lavoro una discussione critica sul presupposto di fondo di questa semplicistica schematizzazione. Cionondimeno essa è ora utile nell'introdurre alcune definizioni da cui partire. Secondo questo metodo si tratta di arrivare a codificare un'informazione su una o più caratteristiche prescelte del fenomeno in modo che essa possa essere utilizzata in un secondo tempo, eventualmente da altri.
Una misurazione è basata quindi sulla schematizzazione (modello) del fenomeno per individuarne in modo univoco le proprietà rilevanti e la sua formalizzazione attraverso relazioni (leggi fisiche) tra enti chiamati grandezze fisiche. In altre parole si rappresenta il fenomeno mediante un modello semplice, nel quale le grandezze di interesse compaiono come le sole essenziali alla sua descrizione. Misurare il diametro di una palla significa averla assimilata ad una sfera; il modello considera irrilevanti le variazioni del diametro con la direzione e una sfera è la schematizzazione ad un solo parametro della palla.
Le misure sono quindi la base della conoscenza scientifica, come ci ricorda bene il famoso detto di Lord Kelvin: *If you cannot measure, your knowledge is meager and unsatisfactory.*
Un primo punto che vorrei sottolineare è che *le considerazioni appena fatte non si riferiscono al solo metodo sperimentale* . Spesso si considera l'esperimento come il solo metodo della scienza , ma ciò a ben vedere non è esatto. Anche rimanendo nel campo della Fisica vi sono scienze come l'astrofisica (pensate di ripetere in laboratorio lo scoppio di una supernova !) , la geologia e la meteorologia dove l'osservazione e la classificazione sono i metodi dominanti. E ciò è ovviamente vero per le altre scienze in generale: pensiamo per esempio alla biologia evoluzionistica. Per parlare di scienza in senso moderno, come afferma il biologo e storico della scienza Ernst Mayr

la semplice osservazione, tuttavia non è sufficiente. Fu solo alla fine del diciottesimo secolo che si fece uso per la prima volta di un metodo particolarmente adatto allo studio delle diversità: il metodo comparativo […]. E la differenza tra il metodo sperimentale e il metodo comparativo non è così grande come può apparire[…]. In ambedue i metodi si raccolgono dati e in ambedue l'osservazione gioca un ruolo cruciale. […]Nelle cosiddette scienze osservative, l'osservatore studia gli esperimenti della natura. La differenza principale tra questi due insiemi di osservazioni è che nell'esperimento artificiale si possono scegliere le condizioni e quindi si è in grado di verificare i fattori che determinano il risultato dell'esperimento. In un esperimento della natura , sia esso un terremoto o la produzione di una fauna insulare, il nostro compito principale è di dedurre o di ricostruire le condizioni in cui l'esperimento della natura ha avuto luogo. (Mayr, 1999, 32).

Alcuni addirittura sottolineano i limiti del metodo sperimentale.

Le teorie delle scienze di laboratorio non vengono direttamente confrontate con "il mondo"; esse sono durevoli in quanto sono vere per fenomeni riprodotti - o addirittura creati – da apparati di laboratorio, e sono misurate con gli strumenti che abbiamo costruito. Questo "vero per" non è questione di comparazione diretta fra la teoria e il fenomeno, bensì si basa su ulteriori teorie, in particolare su teorie su come l'apparato lavora e su vari tipi di tecniche per elaborare *i dati* che generiamo[…]Le nostre teorie persistenti e il mondo vanno così d'accordo non tanto perché noi abbiamo capito come il mondo è realmente, quanto piuttosto perché abbiamo conformato le une all'altro.(Hacking, 2001, 35-36)

Ciò che va messo in evidenza comunque è che se, nello studio dei fenomeni naturali , si passa da un metodo classificatorio (ripartizione di un insieme di oggetti o fenomeni in classi) ad un metodo comparativo, il *ragionamento quantitativo* (*come ovviamente nel metodo sperimentale*) diventa predominante. Infatti si introduce una *relazione d'ordine*, cioè un criterio che consente di decidere in modo univoco se, dati due oggetti o fenomeni, il primo possiede una data proprietà in grado minore, uguale o maggiore del secondo. Se la relazione d'ordine gode della proprietà transitiva, si

può stabilire una corrispondenza tra i gradi di quella proprietà ed *un insieme di numeri*, in modo tale che la relazione d'ordine sia preservata (vedi per esempio Fornasini, 2001).
Consideriamo per esempio la scala di Mohs per le durezze dei minerali. La scala è basata sul criterio che un minerale è più duro di un altro se lo scalfisce. La scala elenca dieci minerali in ordine di durezza crescente, associandoli a 10 numeri in ordine crescente:
1.talco 2.gesso 3.calcite 4.fluorite 5.apatite 6.ortoclasio 7.quarzo 8.topazio 9.corindone 10.diamante. Si consideri inoltre che anche le operazioni di conteggio, sia eseguite a mano che con l'ausilio di strumenti, sono misure. Si pensi , per fare un esempio banale, al conteggio del numero di oggetti celesti osservabili in una certa regione del cielo o di animali componenti un particolare gruppo sotto osservazione.
Un secondo punto che vorrei subito introdurre, e su cui si soffermeranno molte altre considerazioni di questo scritto, è che *qualsiasi operazione di misura implica un certo grado di incertezza del suo risultato.*
Saper valutare correttamente l'incertezza di misura è essenziale sia in campo scientifico, per fissare i limiti di validità delle teorie con cui si descrivono i fenomeni naturali, sia in campo tecnologico, per asserire il grado di affidabilità di prodotti e procedure(vedi per esempio Agnoli et al., 1993).
Qui voglio anche introdurre il fatto che nel 1993 la più autorevole organizzazione attiva a livello mondiale nell'opera di standardizzazione, l'*International Organization for Standardization* (ISO) ha pubblicato una guida (*Guide to the Expression of Uncertainty in Measurement*) con uno specifico vocabolario dei termini usati nella metodologia della misura alla cui stesura hanno contribuito esperti nominati da diversi enti nazionali ed internazionali del campo. Quando lo riterrò utile utile farò riferimento a questa guida.
Abbiamo appena visto che per affrontare in modo scientifico lo studio di un determinato fenomeno, occorre all'inizio individuare quali sono le grandezze significative.
Se per esempio il fenomeno da studiare è il moto di una pallina che rotola su una rotaia in discesa, scopriremo ben presto che alcune grandezze, come il dislivello percorso dalla pallina o l'inclinazione della rotaia, influiscono pesantemente sullo svolgimento del fenomeno, mentre altre, come il colore della rotaia e la temperatura dell'aria, non hanno apparentemente alcun effetto sensibile sul moto.
Una volta individuate le grandezze importanti se ne esegue la misura secondo un metodo chiaro e possibilmente ripetibile.
Occorre ora definire in maniera univoca il concetto di grandezza fisica.
Una grandezza fisica può essere definita come una caratteristica misurabile di un sistema, completamente individuata dal risultato di una ben determinata operazione di misura. Il complesso delle operazioni da eseguire per effettuare una misura si chiama *operazione metrica*. Va forse ricordato che il nome metro deriva dal greco *metron*, latino metrum = misura (in senso generale, non specificatamente di lunghezza). La *metrologia* è appunto la scienza della misura.
Quanto è stato finora detto sta a significare che una grandezza si definisce tramite l'insieme di regole atte a misurarla.
La guida ISO definisce una *grandezza* (*misurabile)* come *un attributo di un fenomeno, di un corpo o di una sostanza che può essere distinto qualitativamente e determinato quantitativamente*. Ed inoltre stabilisce che per *misura debba intendersi l'insieme di operazioni che hanno come scopo la determinazione del valore del misurando*, ovvero della grandezza sottoposta a misura.
In particolare si legge.

The objective of a measurement is to determine the value of the measurand, that is, the value of the particular quantity to be measured. A measurement therefore begins with an appropriate specification of the measurand, the method of measurement, and the measurement procedure. In general, the result of a measurement is only an approximation or estimate of the value of the measurand and thus is complete only when accompanied by a statement of the uncertainty of that estimate.(ISO,1993, 4)

Cominciamo col considerare un tipo di misura, non sempre possibile, che chiamiamo *misura diretta*. Una misura diretta è una operazione che si effettua confrontando la grandezza da misurare con un‘altra grandezza ad essa *omogenea*, presa come *campione*; cioè misurare una grandezza significa trovare un numero che dica quante volte tale grandezza è più grande o più piccola del campione.
Analizziamo questa definizione: innanzi tutto si deve considerare presupposta, per grandezze soggette ad essere misurate direttamente, la definizione operativa di confronto (uguaglianza o disuguaglianza) e somma. In secondo luogo, va precisato il significato di grandezze fisiche omogenee. Sono omogenee quelle grandezze che possono essere misurate attraverso *lo stesso tipo di operazioni e con le medesime regole*; ciò evidentemente significa che sono definite allo stesso modo.
Infine va scelto il campione con cui confrontare le grandezze ad esso omogenee. La scelta è completamente arbitraria, purché soddisfi ad alcuni requisiti di carattere pratico. In definitiva, misurare una grandezza significa associare ad essa un numero che, riferito a un ben precisato campione, cioè a una ben precisata *unità di misura*, ne fornisce il valore. Ciò vale in generale e non solo per le misure dirette prima definite. Per le misure direttamente ottenibili per confronto, cioè le misure dirette, tale numero rappresenta il rapporto tra il valore della grandezza in esame e quello, assunto come unitario, del campione. Per mettersi in grado di eseguire correttamene una misura, il significato astratto di misura come rapporto va precisato nella seguente maniera: fissato il campione, il confronto con esso fornisce due numeri che delimitano l’intervallo entro cui è compreso il valore della grandezza che stiamo misurando. In maniera intuitiva possiamo dire che il risultato della misura è tanto più preciso quanto più è piccolo tale intervallo. Per comodità si usa introdurre multipli e sottomultipli dell’unità di misura, fino ad avere unità arbitrariamente grandi ed arbitrariamente piccole, purché operativamente definibili. Per operativamente definibili si intende semplicemente, qui e sempre, misurabili.

1.1 Prime riflessioni

The trouble with the world is that the stupid are cocksure and the intelligent are full of doubts.
Bertrand Russell

[...]whether or not a thing is measurable is not something to be decided a priori *by thought alone, but something to be decided only by experiment.*
Richard P. Feynman

L’utilizzo di definizioni operative ha indubbiamente un’enorme vantaggio, innanzi tutto, ma non solo, di ordine pratico. L'uso di una definizione operativa permette di definire esattamente ciò di cui si sta parlando e assicura che, in situazioni diverse, ci si riferisca sempre alla stessa grandezza.

E' successo a più riprese nella storia della Fisica, da Aristotele in poi, che errori concettuali siano stati introdotti per analogie o generalizzazioni ingiustificate o per l'uso ripetuto di modelli impiegati con scarso senso critico. Il superamento di questi errori attraverso un riesame dei concetti fisici basato sulla possibilità' di dare di essi una definizione operativa, ha spesso coinciso coi maggiori progressi della Scienza.
Ne sono di esempio i seguenti casi.
Il calore non è altro che una forma di energia e si può' misurare in vari modi. Ma fino alla metà del secolo scorso il calore era concepito come uno strano fluido, privo di massa, ma indistruttibile, che si trasferiva spontaneamente dai corpi più caldi a quelli più freddi etc. … Nessuno aveva pensato di eseguire una serie di operazioni metriche atte a "osservare" almeno una di queste singolari proprietà. Si deve a Joule il merito di avere dimostrato con una serie di esperienze, che il "calorico", cioè quel tal fluido, si poteva creare e distruggere nei corpi semplicemente ammettendo o sottraendo ad essi delle quantità note e ben definite (cioè misurate) di energia meccanica. Da quelle esperienze ebbe origine il I principio della termodinamica.

L'etere cosmico, mezzo singolare, anch'esso imponderabile ma perfettamente rigido ed elastico, considerato (per una errata analogia con le onde elastiche) indispensabile per la trasmissione attraverso l'Universo delle onde luminose, ebbe fine quando Einstein, rinunciando ad esso, interpretò correttamente il risultato dell'esperimento di Michelson e Moreley. L'esperimento si proponeva di mostrare con una effettiva misura che "l'Etere" esisteva.
Infine il principio di indeterminazione di Heisenberg sul quale si erige la meccanica delle particelle elementari ebbe origine da una revisione critica del concetto classico di "orbita", valido per la meccanica classica, ma operativamente non sostenibile per es. nel caso del moto di un elettrone; troppo leggero per essere traguardato anche da un fascio di luce.
In ognuno dei casi ora citati, ossia per il "calorico", l'"Etere cosmico" e la "traiettoria di un elettrone" fu dunque constatato che non era possibile rivelare, *attraverso misure*, quelle proprietà che si erano inventate per giustificarne l'esistenza. Brevemente né il calorico, nè l'etere, né la traiettoria elettronica, erano degli "*osservabili*"(Bernardini, 1974, 9)

Pur non mettendo assolutamente in discussione l'estrema utilità di questo modo di definire i concetti, voglio qui sottolineare che tale approccio, da un punto di vista teoretico, non può sempre risolvere tutti i problemi e le aporie che possono comunque sorgere alla luce dei cambiamenti culturali che ogni linguaggio(come anche quello scientifico) , essendo per definizione materia viva in continua evoluzione, cerca di recepire.
Vediamo, solamente per accennare al problema, il caso di una grandezza fisica fondamentale come la lunghezza (*un esempio assolutamente analogo si potrebbe fare per altre grandezze*).
Lo sviluppo scientifico e tecnologico ha portato ad una progressiva estensione dell'utilizzazione delle grandezze fisiche al di fuori del campo della comune esperienza quotidiana. I valori delle grandezze fisiche possono ora coprire molti ordini di grandezza.
E' ovvio che né le lunghezze su scala atomica né le lunghezze di interesse astronomico possono essere misurate per confronto diretto con un campione di unità di misura. Generalizzando, una stessa grandezza fisica può richiedere tecniche di misura diverse per differenti ordini di grandezza dei suoi valori.
A questo punto si pone, a mio avviso, un primo problema di ordine teoretico.
Tecniche di misura diverse corrispondono a definizioni operative diverse: stiamo quindi, in questo caso, parlando della stessa grandezza fisica *o di due grandezze fisiche differenti*?
Cerco di spiegare meglio con alcuni esempi. Ricordando che si chiama "*anno-luce*" la distanza percorsa da un segnale luminoso in un anno si può dire che la stella a noi più vicina dista circa 4.3 anni luce (a parte il sole che è, in media, a 149.5 milioni di chilometri).
Le stelle aventi una parallasse troppo piccola per esser misurata(vedi anche Bernardini, 1974), hanno imposto agli scienziati una generalizzazione del concetto di misura di una distanza. I corrispondenti risultati hanno prodotto alcune fra le maggiori scoperte della astronomia.

Le stelle osservate con mezzi (spettroscopici) che ne analizzino in colori la luce emessa hanno colori diversi e si possono raggruppare in categorie a seconda del colore. Il Sole per es. appartiene alla categoria bianco-giallo. Osservando, fra le 300 stelle la cui parallasse è misurabile, quelle che hanno presso a poco lo stesso colore, gli astronomi hanno fatto la seguente scoperta: il loro splendore apparente, ossia il flusso luminoso da esse inviato nei grandi telescopi, varia abbastanza regolarmente come l'inverso del quadrato della loro distanza da noi.
Gli astronomi hanno considerato vera la regolarità della relazione esistente fra splendore apparente e distanza per qualsiasi stella e hanno misurato le distanze delle stelle, la cui parallasse è inapprezzabile, dal loro splendore apparente. Più questo è tenue, più lontana è la stella.(Bernardini,1974,11)

In questo modo per esempio si è trovato che le nebulose più lontane sono a distanze dell'ordine di $10^{25}\ m$.
Consideriamo ora distanze atomiche o nucleari. G1i atomi sono sistemi assai complessi. Hanno cioè una *struttura*. Sono (per quello che gli strumenti apposta escogitati hanno permesso capire) atmosfere di elettroni distribuite in volumi dell'ordine di $10^{-30}m^3$. Le loro dimensioni lineari sono dell'ordine di $10^{-10}m$.
In ogni atomo l'atmosfera elettronica elettricamente negativa gravita attorno al nucleo, che e' al centro, che è carico positivamente, ed è costituito dalle particelle elementari pesanti più note: il

protone, che ha una carica positiva uguale a quella dell'elettrone (ma di segno opposto) e il *neutrone*.

Le dimensioni dei nuclei, e dei neutroni e protoni che li costituiscono, sono molto più' piccole di quelle atomiche. C'è qui nella scala delle lunghezze un salto ulteriore. Le dimensioni lineari (i diametri per così dire) dei nuclei sono dell'ordine di $10^{-14}/10^{-15}$m. Sono stati misurati e si misurano lanciando contro di essi fasci paralleli di elettroni, protoni od altre particelle elementari aventi velocità (e corrispondentemente energie) estremamente elevate. Spesso la velocità di tali particelle differisce da quella della luce per meno di una parte su un milione.
Alcune delle particelle dei fasci incidenti urtano contro i nuclei e vengono bruscamente deviate dalla loro direzione iniziale. Contando qual è la percentuale di queste deviazioni si può con calcoli più o meno diretti risalire alle dimensioni dei centri urtati.(Ivi,13)

Riflettiamo quindi ancora sul fatto che un concetto fisico, come ampiamente sottolineato, è strettamente legato al metodo, al procedimento con cui si misura. Quando siamo in presenza di brevissimi spazi o enormi lunghezze (come gli anni luce) che non sono semplicemente un multiplo di una quantità omogenea -o anche di tempi, di azioni di secondo, di tempi astronomici o di tempi cosmici- siamo in presenza di *concetti diversi*?
Un fisico risponderebbe più o meno così: definite due classi di misure differenti, che in linea di principio definiscono due grandezze differenti, se queste classi hanno intersezione non vuota almeno una misura definisce simultaneamente le due grandezze per il medesimo sistema fisico. Allora le unità di misura (locali) possono essere raccordate e le due grandezze considerate le medesime. La mia domanda potrebbe apparire quindi poco più di una provocazione. Personalmente non sono affatto convinto che il dubbio, da un punto di vista teoretico, sia così completamente risolto. Credo a questo riguardo, ed è ciò che mi interessa per ora mettere in evidenza, che lo studio della storia e della pratica recente *possano veramente cercare d'illuminare i concetti della scienza e della tecnologia d'oggi.*
E ciò vale anche per i concetti relativi alla Fisica: il discutere di essi in chiave storica *può essere di aiuto inoltre a cogliere lo spessore delle stesse teorie scientifiche* (l'esempio della teoria della relatività è davvero sin troppo noto). Ciò che voglio qui affermare è che la concettualizzazione non dovrebbe essere fondata solo su una teoria dell'oggetto, ma anche sulle condizioni storiche che motivano la nostra concettualizzazione: abbiamo bisogno della consapevolezza storica della nostra situazione attuale. Non dovremmo mai dimenticare, d'altro canto, che i concetti sono innanzitutto degli strumenti cognitivi che ci permettono di mettere ordine nella realtà che percepiamo.

[…]sembra[…]sempre più plausibile l'ipotesi che i significati dei termini teorici non siano mai stabiliti una volta per tutte e dipendano dalle premesse teoriche che di volta in volta si assumono. Ovviamente esistono dei limiti alle trasformazioni estreme dei significati, ma sono limiti non regolamentati[…]che si possono stabilire solo in senso storico[…].(Continenza, Gagliasso, 1996, 17)

E quindi, anche se Popper nella Prefazione alla edizione italiana della *Logica della scoperta scientifica* del 1970 affermava:

dobbiamo smetterla di preoccuparci delle parole e dei loro significati, per preoccuparci invece delle teorie criticabili, dei ragionamenti e delle loro validità. (Popper, 1990, XLV)

credo non possiamo non riflettere sulle parole usate dell'epistemologo francese Georges Canguilhem nel rispondere:

Ironizzare sull'importanza accordata ai concetti è più facile che comprendere perché senza di essi non si dia scienza. (Canguihem, cit. in Continenza, Gagliasso, 1996, 25)

Mi piace ricordare che lo stesso Einstein nel 1916, nello spiegare in che modo arrivò a formulare il concetto di tempo "relativo", riconosca l'importanza di un tale atteggiamento e i contributi che possono venire dalla riflessione filosofica.

Perché mai è necessario trascinare giù dalle sfere olimpiche di Platone i concetti fondamentali del pensiero scientifico, e sforzarsi di svelare il loro linguaggio terrestre? Risposta: allo scopo di liberare questi concetti dai *tabù* loro annessi, e pervenire così a una maggiore libertà nella formazione dei concetti. Costituisce il merito imperituro di D.Humee di E. Mach quello di avere, più di tutti gli altri, introdotto questa mentalità critica.(Einstein, 1967, 300)

Einstein prosegue inoltre con un primo, ma davvero decisivo esempio, di come i concetti possano essere, come alcuni sostengono (vedi Continenza, Gagliasso, 1996 e Gagliasso, 2001) dei veri e propri *generatori di teorie*.

La scienza ha attinto dal pensiero prescientifico i concetti di spazio, di tempo e di oggetto corporeo[…]e li ha modificati e resi più precisi. Il primo risultato significativo da essa conseguito è stato lo sviluppo della geometria euclidea, la cui formulazione assiomatica non deve renderci ciechi circa la sua origine empirica (la possibilità di giustapporre corpi solidi).(Ibidem)

Ricordo qui che un certo tipo di empirismo postulerebbe la possibilità di un linguaggio in cui i concetti hanno un significato immutato nel tempo, e debbo ricordare che in effetti, come abbiamo visto nel caso di Popper, alcuni ritengono che, almeno nel campo delle scienze fisiche, le analisi riguardanti il formarsi e lo svilupparsi dei concetti utilizzati nella formulazione di teorie non siano da considerare attività di grande interesse. Presentare una compiuta discussione di questo dibattito sarebbe lungo ed al di fuori delle intenzioni di questo scritto. Sicuramente si può sostenere, però, che differente è stato il caso, per esempio, della Biologia, proprio per il carattere storico che questa scienza ha da subito assunto.

Nelle scienze biologiche , […] i progressi più importanti furono ottenuti , nella maggior parte dei casi, con l'introduzione di nuovi concetti, o con l'affinamento di concetti esistenti. Comprendiamo il mondo più efficacemente attraverso miglioramenti concettuali che attraverso la scoperta di fatti nuovi, benché le due cose non si escludano reciprocamente. […]i fenomeni che sono ora spiegati dalla selezione naturale erano largamente noti molto prima di Darwin, ma ebbero un senso solo quando fu introdotto il concetto di popolazioni consistenti di individui unici. […] Non sono troppo lontani dal vero coloro i quali sostengono che il progresso della scienza consiste principalmente nel progresso dei concetti scientifici(Mayr, 1999, 25)

Con queste premesse nell'ultimo capitolo tenterò di chiarire, lo accennavo nell'Introduzione, come la *discussione storica e critica sul concetto convenzionale di probabilità* stia *generando* i nuovi sviluppi della teoria degli errori.

Una seconda riflessione di ordine epistemologico si pone a proposito del fatto che, come si ricordava all'inizio del capitolo, per affrontare in modo scientifico lo studio di un determinato fenomeno, occorre dapprima *selezionare le grandezze significative(*nella guida ISO definite *variabili di influenza)*. Una tale procedura è affetta almeno da un problema: identificare un pezzo di mondo per darne una descrizione *come sistema chiuso* è un'operazione mentale che solo in casi limitati ha un corrispettivo nella realtà. Ogni pezzo di mondo riceve attraverso il proprio contorno degli *influssi ambientali che lo influenzano nella sua evoluzione*. Vediamo quello che dovrebbe essere l'esempio più semplice, l'atomo di idrogeno:

l'elettrone di un atomo di idrogeno interagisce col protone del nucleo e la dinamica corrispondente ammette una soluzione esatta. Ma poi l'elettrone interagisce anche con i campi di radiazione di tutte le altre cariche dell'universo, e ciò implica delle correzioni sensibili che vanno calcolate con procedure ben più elaborate che la soluzione dell'equazione di Schrödinger o di Dirac.

Questa è la base del dibattito attuale sulla cosiddetta *complessità*: l'immagine del mondo circoscritto in una stanza, in cui gli oggetti hanno una collocazione e funzione precisa, non è spesso sufficiente; occorrerebbe in teoria includere tutto quello che viene dal di fuori (Arecchi, 2000,9).

Infatti si innestano scambi con l'ambiente che impongono una struttura "aperta" alla stessa descrizione scientifica. L'irruzione, o *emergenza*, di aspetti nuovi non inclusi nel pacchetto dei dati di partenza rappresenta la parte più innovativa di questa nuova avventura interdisciplinare:

non limitarsi cioè alla scienza delle interazioni fondamentali, da cui poi estrarre, con una procedura pitagorico-platonica, tutti i possibili comportamenti e invece adeguare la descrizione scientifica del mondo agli elementi di realtà che intervengono in ciascuna situazione(ivi, 10).

Ciò significa rendersi innanzi tutto conto che qualunque scienza, che opera nell'ambito di precise formulazioni linguistiche, è sempre un taglio che operiamo sui possibili modi di leggere la realtà. Qualunque teoria seleziona un numero limitato di aspetti da spiegare: in pratica nel rappresentare la realtà facciamo uso di una pluralità di moduli che corrisponde ad una pluralità di *punti di vista*. Consideriamo, per fare un esempio, l'irreversibilità dei fenomeni naturali che *emerge* quando utilizziamo gli strumenti teorici(punto di vista) della termodinamica nei fenomeni macroscopici. Ebbene, a livello della Fisica atomica la freccia del tempo non esiste, non c'è nessuna freccia del tempo, le leggi della Fisica sono perfettamente reversibili; i sistemi fisici tendono verso uno stato di equilibrio, però tutto è simmetrico, sia che il tempo tenda a più infinito che a meno infinito.
Tutto ciò ha un'influenza diretta sulla metodologia della misura:

i protocolli di misura non sono fissati una volta per tutte, non fanno parte integrante del linguaggio formale in cui si è congelata una teoria scientifica, ma in ogni atto di misura c'è un aggiustamento della sensibilità e risoluzione degli apparati per adeguarsi alla realtà locale e questo aggiustamento precede la stessa formulazione linguistica dell'esperimento. Già sappiamo che ciò avviene a livello percettivo. Ad esempio, vedere non è registrare un campo di segnali su un ricevitore passivo, come fa una macchina fotografica, ma è un dialogare con la realtà in quanto la memoria archiviata di situazioni precedenti modifica le soglie neuronali in modo da esaltare certi dettagli e trascurarne altri. Il vedere è perciò un dialogo attivo col mondo che esclude, da una parte la ricezione passiva legata al dualismo fra res cogitans e res extensa, e dall'altra esclude una "autopoiesi" per cui non si esce mai da se stessi e si continua a prendere coscienza solo di proprie rappresentazioni interne. In effetti, i meccanismi di sincronizzazione neuronale di cui ci parlano i neurofisiologi implicano delle correlazioni fra flussi di segnali provenienti dall'esterno (bottom-up: dal basso verso l'alto) e flussi di congetture provenienti dalla memoria (top-down: dall'alto verso il basso). La percezione vista come armonizzazione di questi due flussi implica una sintesi complessa fra due punti di vista che - se isolati - appaiono semplicistici, quali l'empirismo della tabula rasa da una parte e il solipsismo dell'autopoiesi dall'altra.(ivi, 11).

Un'ulteriore conseguenza di ciò può essere schematizzata dicendo che con le leggi che interpretano il comportamento degli atomi non si possono spiegare tutte le proprietà delle molecole, con le leggi che spiegano le proprietà delle molecole non si possono interpretare tutte le proprietà della cellule, con le leggi che regolano le proprietà delle cellule non si possono spiegare tutte le proprietà degli organi e con le leggi che spiegano le proprietà degli organi non si possono spiegare tutte le proprietà dell'uomo.

Su questa linea di una conoscenza che implichi un ruolo attivo dell'intelletto ed allo stesso tempo rivaluti i solidi aspetti della realtà, si possono leggere i temi della ricerca scientifica del futuro. Nell'ultimo capitolo di questo scritto cercherò di mostrare come un nuovo paradigma teorico emergente nel campo dell'inferenza scientifica possa favorire contributi allo studio dei sistemi complessi, oltre che proporre nuovi metodi statistici per determinare l'incertezza di misura.

2. ERRORI DI MISURA

Neppure in un laboratorio le cose si mostrano come dovrebbero essere. Si scostano dalla regola in tutte le direzioni, ed è in fondo un'ipocrisia la nostra di attribuire questo fatto ad un errore di esecuzione e riconoscere all'esperimento un valore medio reale.

R. MUSIL, *L'uomo senza qualità*

In questo capitolo cercherò di riassumere brevemente, anche con alcuni cenni storici, gli elementi fondamentali della *teoria convenzionale degli errori di misura*, così come è oggi ancora presentata nella maggior parte dei corsi universitari delle facoltà scientifiche e tecniche, e che è alla base della moderna metodologia della misura. Per approfondimenti vedi per esempio Taylor, 2000, Caporaloni et al., 1987, Bevington e Robinson, 1994 e Severi, 1986. In particolare si potrà fare riferimento a questi testi se si è interessati alle dimostrazioni di alcuni risultati tecnici che in questo ambito verranno, per lo più, solamente riassunti. Per gli aspetti più propriamente storici si può consultare Stingler, 2000, Hacking, 1975, Roche, 1998 e Birnbaum 1983.
La *teoria convenzionale degli errori ha iniziato a svilupparsi pienamente agli inizi dell'Ottocento e si consolidò pienamente in periodo neopositivista*(ed almeno l'impostazione di base è rimasta sostanzialmente invariata sino ai nostri giorni). Ovviamente in questo contesto la parola *errore* non implica il solito significato di *sbaglio*. Errore significa l'inevitabile incertezza che è presente in tutte le misure. Come tali, gli errori non sono sbagli e non si possono eliminare operando con molta cura.
Per realizzare gli obiettivi della scienza le misure devono essere suscettibili di confronto. Il confronto delle misure richiede un riconoscimento comune della loro precisione, un modo per misurare ed esprimere l'incertezza dei loro valori e le affermazioni inferenziali che ne risultano.
Tutto ciò avvenne compiutamente per la prima volta , come ho appena sottolineato, agli inizi dal XIX secolo.
Ma la consapevolezza del problema dell'accuratezza di una misura era presente dall'antichità.
La mancanza di strumenti matematici adeguati, con cui poter trattare la propagazione degli errori e la rappresentazione dell'incertezza nel caso di misure di tipo statistico, non aveva permesso l'elaborazione di una teoria completa.

There has always been an awareness of accuracy and errors in exact science. Ptomely has frequent references to observational accuracy, error and approximation. There was a corresponding awareness of errors of various sorts during the Renaissance and subsequently, and increasing efforts were made to allow for them. Thomas Digges discussed at length errors made by his predecessors in neglecting to put a sight at the eye end of the *radius astronomicus*. Tycho Brahe drew up tables to compensate for refraction errors in measuring low solar and stellar altitudes. The most important publication in the sixteenth century on practical problems of accuracy and error was Edward Wright's *Certaine Errors in Navigation*, published in 1599. Henry Briggs (1561-1630) in 1616 is fully aware of the number of significant figures which are reliable in Napier's logarithmic tables. Advances in the theory of statistics and probability were required before errors of measurement could be treated mathematically.(Roche, 1998, 57)

Ma fu innanzi tutto la mancanza di un *ambiente culturale* (intendo sociale, economico e politico) come quello creatosi alla nascita della società moderna che impedì un pieno sviluppo della problematica relativa alla teoria degli errori. Il problema del grado di precisione di una misura e della relativa rappresentazione dell'incertezza inizia ad uscire *dagli ambiti della semplice curiosità di uno studioso*(interesse che è rintracciabile certo anche in tanti "scienziati" dell'antichità, vedi Russo, 2001) *per diventare terreno comune di ricerca solo con la nascita della società moderna, con i suoi ideali di razionalità, universalità e precisione.*
Gli storici sono d'accordo sul fatto che sia stato Galileo (vedi per esempio Birnbaum 1983 I, Erto 1999) nel 1632 il primo a formulare dei concetti chiari, anche se non matematici, relativi agli errori di misura. Egli usava questi concetti per guidare le sue analisi delle osservazioni astronomiche.
Uno degli scopi principali della teoria degli errori è di valutare l'affidabilità dei metodi di stima.
Nel 1755 il matematico inglese Thomas Simpson provò che la media aritmetica di n misure ripetute (vedi par.2.2), che era già in uso limitato, è superiore ad una singola misura in un senso precisamente specificabile. Va notato il fatto che il primo esempio di trattamento dei dati che usa il concetto di media aritmetica si riscontra tra gli astronomi di Baghdad nel IX secolo (vedi per esempio Grasshoff, 1990). Ma questa tecnica ha cominciato ad essere consolidata e ad essere universalmente accettata solo molti secoli dopo, durante la rivoluzione scientifica.

Problemi relativi alla teoria degli errori erano importanti nello sviluppo classico della teoria della probabilità iniziata da Laplace, Gauss e Legendre nel primo '800. Quello più importante riguardava la distribuzione di probabilità della media aritmetica di un gruppo di misure, sotto varie possibili assunzioni matematiche circa le misure (o gli errori) individuali.

Laplace and later writers proved that if the individual measurements independent and approximately equally variable, the distribution of the arithmetic mean approaches the normal, or Gaussian, distribution as *n* increases. This provides an approximation to the distribution of errors incurred when the arithmetic mean is used as an estimate. The same mathematical results make it plausible that if, as Galileo, Gauss, and others suggested, a measurement error is the net result of numerous independent, largely uncontrollable, and unknown disturbances, then in many circumstances the distribution of errors should be approximately normal.(Birnbaum, 1983 I, 560)

Vorrei qui accennare che fra i principali problemi della teoria degli errori possono annoverarsi i seguenti seguenti: (1) La determinazione delle prove empiriche e teoriche delle varie assunzioni riguardo la forma di una distribuzione degli errori. (2) Sulla base di tali assunzioni, *l'individuazione delle migliori stime basate sulle osservazioni disponibili*, e la determinazione della precisione di tali stime. (3) *Il chiarimento dei problemi concettuali coinvolti nell'inferenza statistica.* (4) L'estensione di tutte le considerazioni menzionate a problemi più complessi d'importanza pratica, in particolare il problema della valutazione simultanea di molte costanti sconosciute partendo da un singolo gruppo di dati.
Come accennavo è solo in pieno periodo neopositivista che abbiamo il consolidamento di una teoria completa. *I risultati*, per motivi diversi che discuterò più avanti, *sono però tutt'altro che soddisfacenti*. Voglio sottolineare al riguardo che la stessa esigenza di produrre la più volte citata guida dell'ISO *Guide to the Expression of Uncertainty in Measurement* è nata e si è sviluppata per contribuire a ridurre *la mancanza di consenso internazionale oggi ancora esistente* nel problema della espressione di un risultato di misura. A tale riguardo riprendo qualche parola dal "Foreword" della guida stessa.

In 1978, recognizing the lack of international consensus on the expression of uncertainty in measurement, the world's highest authority in metrology, the Comité International des Poids et Mesures (CIPM), requested the Bureau international des Poids et Mesures (BIPM) to address the problem in conjunction with the national standards laboratories and to make a recommendation.[…]
The BIPM then convened a meeting for the purpose of arriving at a uniform and generally acceptable procedure for the specification of uncertainty; it was attended by experts from 11 national standards laboratories. This Working Group on the Statement of Uncertainties developed Recommendation INC-I (1980), Expression of Experimental Uncertainties. The CIPM approved the Recommendation in 1981 and reaffirmed it in 1986.
The task of developing a detailed guide based on the Working- Group Recommendation (which is a brief outline rather than a detailed prescription) was referred by the CIPM to the International Organization for Standardization (ISO), since ISO could better reflect the needs arising from the broad interests of industry and commerce.(ISO,1993,v)

Lo scopo della Guida ISO (detta in maniera colloquiale GUM) è la creazione di "*general rules for evaluating and expressing uncertainty in measurement that can be followed at various levels of accuracy and in many fields--from the shop floor to fundamental research.*" Di conseguenza, l'intenzione è quella di rendere i principi applicabili ad un ampio spettro di misure, comprese quelle per:

il mantenimento dei controlli di qualità e dell'assicurazione della qualità nella produzione;

il rispetto e l'attuazione delle leggi e dei regolamenti;

la conduzione di ricerca di base, di ricerca applicata e di sviluppo, nella scienze e nell'ingegneria;

la taratura degli strumenti e le specifiche per lo svolgimento dei test attraverso un sistema nazionale di misura per realizzare una tracciabilità secondo standard nazionali;

lo sviluppo, il mantenimento, e il confronto degli standard di riferimento a livello nazionale e internazionale , compresi tutti i materiali di riferimento.

E l'obiettivo finale è palese:

Just as the nearly universal use of the International System of Units (S.I.) has brought coherence to all scientific and technological measurements, a worldwide consensus on the evaluation and expression of uncertainty in measurement would permit the significance of a vast spectrum of measurement results in science, engineering, commerce, industry, and regulation to be readily understood and properly interpreted. In this era of the global marketplace, it is imperative that the method for evaluating and expressing uncertainty be uniform throughout the world so that measurements performed in different countries can be easily compared.(ivi, VII)

Più recentemente è stata creata una nuova organizzazione internazionale per assumere la responsabilità del mantenimento e la revisione della GUM. Il nome dell'organizzazione è *Joint Committee for Guides in Metrology (JCGM)* ed i suoi membri sono:

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

Cominciamo con alcune definizioni. Come ho più volte sottolineato a causa dell'inevitabile incertezza il risultato di una misura è espresso da un intervallo di valori. *Si tratta ora di decidere quale è il modo più conveniente per indicare tale intervallo.* La sola indicazione degli estremi dell'intervallo è molto scomoda nel caso in cui ci si debba servire del risultato della misura per fare dei calcoli, poiché si avrebbe a che fare con due valori. Si suole allora indicare il centro dell'intervallo e la sua semiampiezza. Così, se dobbiamo misurare uno spigolo e se la fine dello spigolo è compresa tra la linea dei 528 cm e quella dei 529 cm, si scrive $1 = 528{,}5 \pm 0{,}5$ cm.
In questo modo, accanto al valore della grandezza, si è indicato l' *errore* con cui tale valore può ritenersi noto. La guida ISO definisce l'errore di misura la *``differenza fra il risultato di una misura e il valore vero del misurando''*. Diremo che la misura è tanto più precisa quanto più è piccolo l'errore.
L'inevitabilità degli errori di misura e l'ignoranza della loro entità fa sì che non è dato di conoscere con esattezza il valore del misurando.

When reporting the result of a measurement of a physical quantity, it is obligatory that some quantitative indication of the quality of the result be given so that those who use it can assess its reliability. Without such an indication, measurement results cannot be compared, either among themselves or with reference values given in a specification or standard. It is therefore necessary that there be a readily implemented, easily understood, and generally accepted procedure for characterizing the quality of a result of a measurement, that is, for evaluating and expressing its *uncertainty.*(ISO, 1993, VII)

Il concetto di *incertezza* quale caratteristica quantificabile è relativamente nuovo nella storia della misura. È ormai ampiamente riconosciuto che when all of the known or suspected components of error have been evaluated and the appropriate corrections have been applied, there still remains an uncertainty about the correctness of the stated result, that is, a doubt about how well the result of the measurement represents the value of the quantity being measured.(ibid, VII)

Ne segue allora che ad ogni risultato di misura è associato un certo grado di una incertezza (o indeterminazione), ove con questo termine si intende (guida ISO) qualitativamente,

> ``*un parametro che caratterizza la dispersione dei valori che possono essere ragionevolmente attribuiti al misurando".*

Come ho già accennato, per errore di misura non va inteso un eventuale sbaglio, ad esempio di lettura. Peraltro tali sbagli possono succedere per cui è sempre un buon metodo quello di fare due volte la misura per mettersi al riparo da simile eventualità. L'errore di misura ora illustrato rientra nella categoria dei cosiddetti *errori di sensibilità* (come introdotto analiticamente nel par. precedente), in quanto riferiti alla più piccola quantità apprezzabile con quel campione, o strumento. E' possibile in teoria ridurre tale errore fin che si vuole, ma oltre un certo limite non ha più senso usare campioni, o strumenti, in cui siano stati segnati sottomultipli dell'unità sempre più piccoli. Non sempre però l'errore di sensibilità è la grandezza più significativa per esprimere la precisione di una misura.
Una importante considerazione è che non è ben chiaro che cosa debba intendersi per *valore vero*. Supponiamo di provare a misurare la lunghezza di un bastone rotto, le cui estremità siano frastagliate e scheggiate. Possiamo stabilire che la lunghezza è compresa tra certi limiti, per esempio tra 14 e 15 cm. Ma, se cerchiamo di essere più precisi, dobbiamo decidere dove si trovano le estremità. Se vogliamo misurare la lunghezza del bastone con una precisione dell'ordine di 0,01 cm, non possiamo affermare che il bastone ha una lunghezza ben definita a questo ordine di grandezza. E' chiaro che ciò, anche a diversi ordini di scala, è vero per qualsiasi misura macroscopica .
La definizione ufficiale ISO del valore vero è la seguente:

> ``*un valore compatibile con la definizione di una data grandezza particolare".*

Va sottolineato l'utilizzo dell'articolo *un*, in quanto ci possono essere più valori consistenti con la definizione di una data grandezza particolare. *Questa definizione, che può sembrare un po' vaga, è in effetti quella più pragmatica*.
Nei paragrafi che seguiranno, assumeremo che stiamo effettuando misure su quantità per le quali il valore vero esista. Pur riconoscendo che il concetto di valore vero è, di fatto, una idealizzazione, ha senso parlare di valore *convenzionalmente vero* se con esso si intende quello attribuito a una grandezza particolare e accettato, per convenzione, in quanto avente un'incertezza appropriata all'uso. Dovremo comunque riflettere che non è corretto affermare che una ben determinata quantità osservabile abbia un valore completamente definibile. Dobbiamo accontentarci di constatare che la media di un gran numero di osservazioni ha un determinato valore.

2.1 Le cause di errore

Se chiudiamo la porta di fronte all'errore,
come farà la verità ad entrare?
R.Tagore

Riprendiamo ora il nostro discorso sugli *errori ad una scala macroscopica*.
Ogni misura diretta comporta necessariamente un errore, per quanto piccolo possa essere. Nella *teoria convenzionale* gli errori, dovuti al fatto che non esistono né strumenti perfetti né condizioni ideali in cui effettuare la misura, possono essere di natura sia *sistematica che accidentale:*

- gli *errori sistematici* sono dovuti a imperfezioni strumentali o a una procedura sperimentale non corretta**.** Essi tendono ad alterare la misura sempre nello stesso modo e possono essere individuati e corretti utilizzando strumenti e tecniche differenti;

- gli *errori casuali o accidentali* sono invece dovuti a tutte le piccole variazioni delle condizioni ambientali che non si riescono a tenere completamente sotto controllo. Essi influenzano la misura in modo casuale e non possono mai essere eliminati.

Vediamo meglio il concetto di errore sistematico.
Consideriamo la misura dello spigolo di un tavolo. Se il nostro metro campione è fatto di un materiale con un grosso coefficiente di dilatazione termica è possibile trovare due risultati diversi a seconda della temperatura a cui si opera. Se non si tiene conto di tale fenomeno si commette un errore sistematico. Ancora può succedere che il campione da noi usato non corrisponda alla sua definizione per cui ripetendo la misura con un altro campione troviamo un altro risultato.
Errori di questo tipo sono detti sistematici perché, a parità delle condizioni in cui si svolge la misura, essi influenzano il risultato della misura sempre nello stesso senso. Essi sono quindi ineliminabili con ripetizione di misure e possono essere "scoperti" solo cambiando metodo, strumento o ambedue (in tal caso ci si aspetterebbe di trovare infatti lo stesso risultato: l'eventuale differenza tra i due risultati da un'idea dell'importanza degli errori sistematici presenti). Non è possibile fare misure senza errori, anche di questo tipo: tuttavia ci si può porre in condizioni in cui essi siano resi ragionevolmente piccoli.
Ancora un esempio. Se misuriamo il tempo impiegato da un oggetto a cadere da un'altezza nota utilizzando un cronometro manuale, la misura risulterà alterata da un grosso errore sistematico, dovuto ai nostri tempi di reazione nell'utilizzo del cronometro, e non sarà quindi accurata, indipendentemente dal fatto che il cronometro sia preciso o meno.
Perché la misura sia accurata, occorre collegare il nostro cronometro a un sistema di fotocellule. Se usiamo un cronometro al decimo di secondo, la precisione del risultato non potrà mai essere migliore di un decimo di secondo, poiché il nostro strumento di misura non è in grado di risolvere intervalli di tempo inferiori e, se ripetiamo molte volte la misura, otterremo sempre lo stesso risultato.
Se utilizziamo ora un cronometro al millesimo di secondo, la precisione aumenta, ma se ripetiamo la misura molte volte troveremo risultati leggermente diversi tra loro. Ora la precisione della misura è limitata non dallo strumento ma dagli errori casuali (dovuti p.e. a variazioni di temperatura o di pressione, o a correnti d'aria), il cui effetto è di alterare in maniera percettibile il fenomeno fisico che stiamo misurando. Poiché le variazioni agiscono in modo casuale, causando a volte errori per eccesso e altre volte per difetto, la precisione della misura può essere aumentata ripetendola molte volte e prendendo poi il *valore medio* dei risultati ottenuti.

Anche solo da questi brevi cenni appare chiaro che la discussione degli errori sistematici, così fondamentale nell'acquisizione dell'informazione sperimentale, è molto delicata, anche perché, per la loro stessa natura, non si può mai essere sicuri di aver tenuto conto completamente di tale tipo di errori; generalmente ci si accontenta di essere sicuri che il loro valore sia molto minore del contemporaneo errore casuale.

Le cause degli errori sono molteplici. La guida ISO introduce le possibili sorgenti di incertezza di misura seguendo questo elenco:

Incompleta definizione del misurando.
Ad esempio l'"accelerazione di gravità al livello del mare" non definisce completamente il misurando in quanto essa dipende anche dalla latitudine.

Imperfetta realizzazione della definizione del misurando.
Molto spesso non si riesce a realizzare sperimentalmente un dato fenomeno fisico così come lo

si è delineato concettualmente. Consideriamo la caduta di un grave: se si vuole verificare la legge del moto di un corpo sottoposto alla forza di gravità occorre eliminare l'effetto della presenza dell'aria che introduce un'altra forza i cui effetti sulle misure sono ovviamente di natura sistematica.

Campione non rappresentativo, ovvero il campione misurato non rappresenta il misurando definito .
Consideriamo per esempio i sondaggi. Talvolta, in mancanza di un modello teorico per la scelta del campione, si rischia di effettuare un sondaggio su un campione appunto non rappresentativo.

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.

Errore di lettura di uno strumento.
La lettura delle scale analogiche dipende dall'acuità visiva e dall'abilità di stima dello sperimentatore. Ad esempio, un errore grossolano è quello di *parallasse.*

Risoluzione finita o soglia di discriminazione dello strumento.
Ad esempio, se la lettura avviene con uno strumento digitale si è limitati alla cifra meno significatica 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.

Valori inesatti dei campioni e dei materiali di riferimento.
Un caso particolare è quello relativo ad un difetto di taratura (l'impiego di campioni costituisce, nella maggioranza dei casi, una difficoltà e nella pratica giornaliera si fa sempre uso di apparecchi già tarati): anche se il funzionamento dello strumento può apparire normale, in realtà non è verificata la corrispondenza prevista tra la sua risposta ed il valore della grandezza. Sarebbe come misurare delle distanze con un regolo che riteniamo essere lungo un metro , ma che in realtà è per esempio 99.5 centimetri. Il modo di ovviare a questo tipo di errore è quello di controllare la taratura dello strumento prima, ed eventualmente anche dopo, l'esecuzione della misura. Inoltre sono sempre possibili errate condizioni di impiego. Anche se lo strumento in sé non presenta difetti, esso può essere impiegato in modo non corretto, come per esempio una bilancia tenuta in posizione non orizzontale, ed in particolare non nelle condizioni in cui è previsto essa abbia la taratura giusta.

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.

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.

Variazioni in osservazioni ripetute del misurando sotto condizioni di misura apparentemente identiche.
Queste variazioni sono legate agli errori casuali, dei quali parlerò nel paragrafo 6.6.

Va sottolineato che ovviamente 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, e in particolare nella lettura di strumenti analogici, che per la capacità di vagliare i vari contributi all'errore di misura. Ed inoltre bisogna tener conto dell'interazione tra strumento e sistema in misura. Ogni strumento, per operare, deve essere messo in relazione col fenomeno fisico di cui si vuole misurare una grandezza. Da ciò derivano due diversi tipi di errori sistematici:

-uno stesso strumento può essere messo in relazione col fenomeno per mezzo di dispositivi diversi, ciascuno dei quali può a sua volta introdurre un errore sistematico. Per es. misure di spessore eseguite con un calibro risentono di eventuali corpi estranei, come polvere, vernice, ecc.,

-quale che sia il sistema usato per farlo, la parte sensibile dello strumento viene posta in relazione col sistema fisico. Come conseguenza quest'ultimo, almeno in linea di principio, ne risulta sempre modificato, nel senso che il sistema fisico su cui si esegue la misura non è più quello in studio, ma uno nuovo che comprende, come parte non sempre trascurabile, anche lo strumento stesso con la sua fenomenologia. In altri termini l'inserzione dello strumento modifica il fenomeno iniziale introducendo in esso variazioni sistematiche che possono essere apprezzabili.

2.2 Ripetizione delle misure

Discutiamo ora il problema della ripetizione della misura e di che cosa ci si attende da essa. Esistono due possibilità :

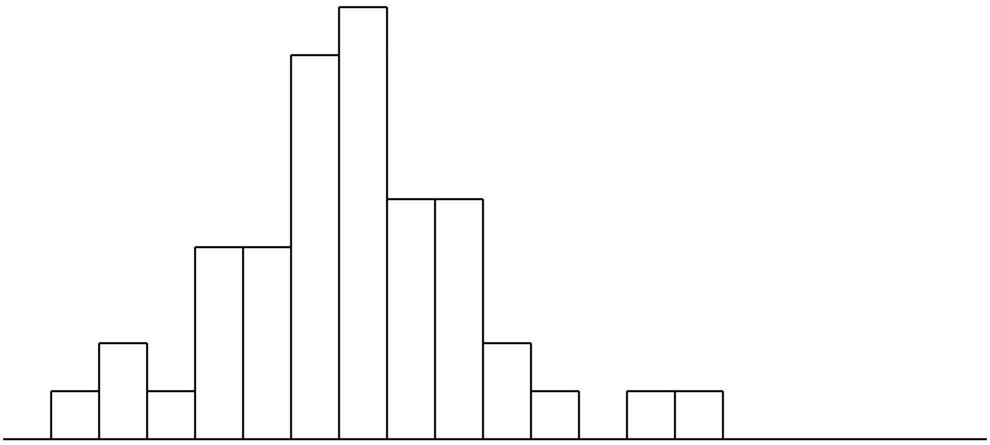
a) si ottiene sempre lo stesso risultato entro i limiti di sensibilità del campione;

b) si ottengono risultati differenti, cioè con differenze non comprese nella sensibilità.

Abbiamo già visto che possiamo definire come casuale quel tipo di errore (vedi anche paragrafo successivo), dovuto all'impossibilità di avere sotto controllo infinitamente accurato tutte le variabili in gioco, che può far variare il risultato di una misura con eguali probabilità nei due sensi. I ragionamenti che seguono servono per indicare come si fa, in un caso come questo, quando cioè l'elevata sensibilità del campione mette in evidenza la presenza degli errori casuali, ad indicare un valore ed un errore per un primo risultato di una misura.
E' conveniente rappresentare i risultati delle misure in un grafico (istogramma, vedi figura), in cui si divida l'asse x in intervalli di ampiezza uguale o multipla del valore della sensibilità (non avrebbe senso usare intervalli minori della sensibilità), riportando poi sull'asse y , in corrispondenza ad ogni intervallo, un segmento di lunghezza proporzionale al numero di misure che hanno dato un risultato compreso nell'intervallo stesso. Il modo di distribuirsi dei risultati è suscettibile, come vedremo più avanti, di interpretazioni quantitative; tuttavia la semplice osservazione di tali distribuzioni permette conclusioni e stime qualitative. E' ragionevole ad esempio supporre che, purché le misure siano eseguite con cura, i "grandi" valori degli errori siano

meno probabili dei “piccoli” valori. Di conseguenza ci si può aspettare una distribuzione di valori caratterizzata da un addensamento di risultati in alcuni intervalli ed approssimativamente simmetrica. Segnando con un crocetta per ogni misura, per l’esempio prima citato, si otterrebbe l’istogramma in figura. Risulta ragionevole, data una distribuzione circa simmetrica e dotata di massimo:

a) adottare come stima migliore del misurando la media aritmetica dei valori ottenuti. Essa è infatti il valore da cui si discosta meno il maggior numero dei risultati(questa grandezza viene anche spesso chiamata *media dei campioni*;
b) assegnare come errore a tale valore la semilarghezza della distribuzione (cioè il modulo della semidifferenza tra il valore massimo e il valore minimo riscontrati). Questo errore, cui si dà nome di *semidispersione massima*, ha il significato che se si ripetesse ancora una volta la misura si troverebbe “quasi certamente” un valore compreso tra i limiti definiti dalla media aritmetica più o meno la semidispersione.



Vediamo ora però di definire meglio alcuni concetti appena introdotti, insieme ad altri che ci saranno utili in tutta la nostra trattazione.

Supponiamo di voler misurare con accuratezza il diametro di un tubo cilindrico con un micrometro. Supponendo che il tubo abbia un «diametro vero», noi probabilmente otterremo risultati diversi se ripeteremo più volte la misura. Può darsi infatti che alcune volte la vite micrometrica venga stretta più che in altre, vi possono essere piccole particelle di polvere, può avvenire che si compiano piccoli errori nell’apprezzare i decimi della minima divisione sulla scala, etc…Comunque, si intuisce che dovrebbe potersi ottenere un risultato più attendibile per il diametro utilizzando dieci misure piuttosto che una soltanto.

Che uso faremo dunque ora delle dieci misure? Il primo passo, come prima introdotto (e che spontaneamente ci viene di fare) è calcolare la media aritmetica. La media di una serie di N numeri è definita come la somma di tutti i numeri, divisa per N. Se abbiamo dieci misure, ne facciamo la somma e la dividiamo per dieci. In termini più generali chiamiamo con x_i il risultato di una generica misura. Se si hanno dieci osservazioni, l’indice i può assumere un qualunque valore compreso fra 1 e 10. Se le misure sono N, i varia da 1 a N. In generale, possiamo definire la media di una serie di numeri x_i così:

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

Nel seguito una sbarretta sulla lettera significherà sempre un «valore medio».

Ci sono buone ragioni per ritenere la media aritmetica di una serie di misure come la migliore stima del valore vero della grandezza misurata(vedi bibliografia consigliata all'inizio del capitolo).

Osserviamo comunque, come prima ragione, che calcolare la media *appare intuitivamente come un procedimento logico.* Va a questo riguardo notato che *ciò non significa che sia affatto un procedimento scontato.*

Come già accennato questa tecnica ha iniziato ad essere riconosciuta universalmente solo con l'inizio della rivoluzione scientifica.

Una volta ottenuta una serie di misure x_i , potremo voler disporre di un criterio per stabilire quantitativamente in che modo le singole osservazioni sono distribuite intorno alla media. Una descrizione quantitativa dalla dispersione delle misure ci darà un'idea della precisione di queste misure.

Per ottenere una simile descrizione quantitativa, definiamo dapprima per ogni misura x_i lo *scarto* d_i. Lo scarto d_i è definito come la differenza tra ogni misura x_i e la media della serie di numeri, vale a dire:

$$d_i = x_i - \bar{x}$$

(Alcuni autori chiamano d_i «deviazione» invece che scarto: i due termini sono equivalenti).

Si noti che non sarebbe corretto chiamare d_i l'«errore» sulla misura x_i, perché la media non è affatto il valore vero della grandezza osservata. Si può solo dimostrare che in molti casi, se si esegue un gran numero di osservazioni (*ammettendo che non intervengano errori sistematici*) allora gli scarti d_i approssimano i veri errori di cui le misure x_i sono affette.

Questo è il caso, per esempio, in cui gli errori siano distribuiti secondo la legge di Gauss, o «funzione normale», che discuteremo oltre.

Per il modo in cui abbiamo definito la media e gli scarti, la «media degli scarti» è sempre zero. Ciò significa che, la media degli scarti non serve in pratica per una caratterizzazione della dispersione. Forse un'idea migliore sarebbe quella di prendere il valore assoluto di ogni scarto, e calcolare la media dei valori assoluti. Otteniamo così la cosiddetta «deviazione media», che si indica con α, cioè:

$$\alpha = \frac{1}{N}\sum_{i=1}^{N} |d_i| = \frac{1}{N}\sum_{i=1}^{N} |x_i - \bar{x}|$$

Questa grandezza si indica spesso come deviazione media, ma si tratta di una denominazione imprecisa, come anche l'altra: «scarto medio». Essa non è infatti lo scarto medio, bensì la media dei valori assoluti degli scarti. Questa grandezza è talora usata per caratterizzare la dispersione delle misure. Per vari motivi però (vedi bibliografia consigliata all'inizio del capitolo), essa non è così utile come l'altra come stiamo per definire, chiamata «deviazione standard». Nel definire la deviazione standard aggiriamo il problema di trattare con degli scarti negativi, quadrando ogni deviazione e ottenendo così una quantità sempre positiva. Allora calcoliamo la media dei quadrati e poi estraiamo la radice quadrata del risultato ottenuto.

Così la deviazione standard risulta essere la radice quadrata della media dei quadrati degli scarti. La deviazione standard si indica generalmente con σ e la relazione che la definisce è:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

Il quadrato della deviazione standard, σ^2, è detto «varianza».

Si noti che σ ha sempre le stesse dimensioni di x_i e che è sempre positiva. Vedremo nel prossimo paragrafo che uso fare di queste definizioni.

2.3 Errori casuali

Riprendiamo ora, in maniera più approfondita, il concetto di errore casuale.
Per il postulato di invarianza spazio-temporale ci si aspetterebbe che la stessa operazione di misura, fatta con uno strumento infinitamente preciso ed infinitamente sensibile, desse sempre lo stesso risultato a parità di condizioni. In generale però questo non è vero: ripetendo più volte la misura si troverebbero valori diversi anche operando rigorosamente nelle stesse condizioni.
Qualitativamente si può dare ragione di questo fatto, per fenomeni macroscopici, considerando che quando si parla di condizioni si intende qualcosa di generale, mentre *il valore assunto da una grandezza è poi dovuto all'effetto concomitante di tanti contributi* di cui non si riesce a tener conto individualmente nell'analisi delle condizioni generali. Per esempio, le forze di attrito sono date dall'effetto delle tante azioni, non valutabili, di ciascuna delle parti dei due corpi in contatto; la corrente ed il potenziamento elettrico sono dovuti in realtà al moto ed alla posizione di moltissimi piccoli corpi elettricamente carichi, i quali però sono soggetti anche a moti disordinati, per es. dovuti alla agitazione termica (per fenomeni microscopici, invece, per i quali non ha senso parlare di condizioni di carattere generale, vale come visto il principio di indeterminazione che pone un limite superiore alla quantità di informazione che si può ottenere su di essi).

Quanto detto si riferisce anche *all'operazione di misura*, dato che *il funzionamento di ogni strumento è determinato da fenomeni fisici*. Ciò fa si che ad una non riproducibilità intrinseca del valore di una grandezza(con G indico qui di nuovo una grandezza fisica e con V(G) il *valore vero*, vedi oltre, sconosciuto) si sovrapponga una non-riproducibilità del funzionamento dello strumento, ed è importante notare che per l'osservatore, la cui unica informazione è la risposta dello strumento, i due effetti sono sovrapposti e possono manifestarsi o no a seconda della loro entità confrontata con l'errore di sensibilità dello strumento stesso, $\Delta V(G)$.

Da queste considerazioni segue che V(G) non può essere conosciuto che con una certa indeterminazione, che è chiamata appunto *errore casuale*.

Riassumiamo quindi le possibilità che si presentano nel caso in cui le misure siano ripetibili (tale condizione, assunta implicitamente finora, non è sempre verificata; il caso di misure non ripetibili verrà affrontato subito dopo).

- Errori massimi. Se ripetendo la misura si ottiene sempre lo stesso risultato significa che $\Delta V(G)$ è molto maggiore della deviazione standard σ della distribuzione di valori prodotta dalla fluttuazione intrinseca di G e dalle fluttuazioni introdotte dall'apparato di misura a causa della sua limitata precisione. In questo caso, pertanto, $\Delta V(G)$ assume il significato di *errore massimo*.
- Errori statistici. Nel caso opposto, $\Delta V(G) \ll \sigma$, ripetendo l'operazione di misura si ottengono effettivamente valori diversi, il cui istogramma approssima tanto meglio la distribuzione che esso rappresenta sperimentalmente, generalmente una gaussiana (vedi par 7.5), quanto più $\Delta V(G)$ è minore di σ e quanto maggiore è il numero delle misure effettuate. In questo caso si usa rappresentare l'indeterminazione sulla conoscenza di V(G) con σ, che è in effetti la stima di σ ottenuta sperimentalmente dai dati, e questa indeterminazione è detta *errore statistico*.

2.4 Misure non ripetibili.

Può capitare di riuscire ad eseguire una sola operazione di misura, o perché si è nell'impossibilità, per varie ragioni, di effettuarne più di una, oppure perché è il fenomeno ad essere unico, per lo meno in determinate condizioni, come capita per es. in molti fenomeni naturali (eclissi, fenomeni meteorologici, ecc).

In quest'ultimo caso, naturalmente, non ha senso parlare di fluttuazione intrinseca della grandezza, per cui l'indeterminazione sulla conoscenza della V(G) è dovuta solo alle caratteristiche dell'apparato. Queste vanno valutate separatamente eseguendo misure ripetibili su una grandezza omogenea a quella in misura e preferibilmente nelle stesse condizioni, anche per tenere conto di eventuali effetti sistematici.

2.5 Errori massimi su funzioni di grandezze

In una misura indiretta, il valore di una grandezza fisica viene determinato utilizzando una formula (che viene spesso denominata *equazione di misura*) che lega tale grandezza ad altre quantità già misurate. L'errore nasce in questo caso dagli errori associati a queste quantità; si parla a tale proposito di *propagazione degli errori*.
In particolare, se una grandezza è funzione di altre, misurate effettivamente, ha interesse un calcolo a priori del massimo errore da cui può essere affetta : cioè dell'intervallo, intorno al suo valore medio, entro cui varieranno i risultati della misura, quando le grandezze misurate assumano valori compresi entro i rispettivi errori.
Vogliamo in altre parole vedere come si calcola in generale l'errore massimo su una grandezza y che sia funzione di altre grandezze misurabili direttamente (x_1,x_2 ,…..).

Sappiamo che il differenziale totale della funzione $dy = \left(\frac{\partial y}{\partial x_1}\right)dx_1 + \left(\frac{\partial y}{\partial x_2}\right)dx_2 + \ldots$ rappresenta la variazione infinitesima che il valore $y(x_1,x_2$,…..) subisce per variazioni infinitesime delle singole variabili. Facendo l'ipotesi che tale espressione *sia sufficientemente approssimata anche per variazioni finite* ma piccole delle variabili, considerando dx_1, dx_2, ….. come gli errori da cui sono affette le grandezze misurate, otteniamo l'errore risultante su y:

$$\Delta y \approx \left|\frac{\partial y}{\partial x_1}\right|\Delta x_1 + \left|\frac{\partial y}{\partial x_2}\right|\Delta x_2 + \ldots$$

ad esempio

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

Il valore assoluto delle derivate sta a significare che dovremmo calcolare l'errore massimo nel modo più "pessimista", che è quello che si ottiene quando gli errori sulle singole misure non si compensano, ma si combinano in modo da far variare y sempre nello stesso senso.

2.6 Errore relativo ed errore percentuale

Definiamo *errore relativo* il rapporto tra l'errore e il valore della grandezza

$$\in_r = \frac{\Delta y}{y}$$

L'*errore percentuale* si ottiene semplicemente moltiplicando per 100 l'errore relativo.
L'errore relativo e l'errore percentuale permettono di confrontare la precisione di misura di più grandezze molto meglio che non l'errore assoluto. Se si tratta poi di grandezze non omogenee l'errore assoluto non permette affatto tale confronto. Diciamo che una misura è tanto più precisa quanto più è piccolo l'errore relativo.

2.7 Errori statistici su funzioni di grandezze

Per una misura indiretta di una grandezza fisica l'errore statistico si ricava dagli errori sulle misure dirette delle grandezze a cui la prima è legata da una relazione analitica che sarà qui indicata con $y=f(x_1,x_2 ,\ldots..)$. Come già visto per gli errori statistici generati da misure dirette, anche in questo caso si usa rappresentare l'indeterminazione con σ, che è in effetti la stima di σ ottenuta sperimentalmente dai dati (vedi par 3.3) ed è anch'essa detta errore statistico. Con considerazioni di una certa complessità, di carattere statistico appunto (vedi bibliografia consigliata all'inizio del capitolo), si può mostrare che la σ (y) si può ricavare in buona approssimazione dalle $\sigma(x_i)$ mediante l'espressione

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

per esempio

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

Nel prossimo capitolo, si definirà, una volta introdotti i concetti di probabilità e funzione di distribuzione di probabilità, come meglio stimare le $\sigma(x_i)$ e come meglio utilizzarle per stimare le incertezze.

3. UNO SGARDO ALLA "COMBINAZIONE DELLE OSSERVAZIONI"

The rules of clockwork might apply to familiar objects such as snookerballs, but when it comes to atoms, the rules are those of roulette.
Paul Davies

Come sappiamo il nostro cervello elabora informazioni in condizioni di incertezza: i dati che provengono dagli organi di senso forniscono infatti informazioni su una realtà incompleta, variabile

ed incerta e *gli oggetti di questa realtà sono percepiti in maniera dinamica, mutevole e parziale. L'incertezza relativa al risultato di una operazione di misura è l'esempio più evidente di questa situazione*. Esiste un importante capitolo della matematica moderna che si occupa di dati e variabili aleatori: la matematica dell'incertezza. *Questo ramo della matematica è costituito principalmente dalla teoria del calcolo probabilistico.*
Nei capitoli precedenti ho discusso il significato da attribuire al valore di una misura. L'errore massimo (di sensibilità o di semidispersione per le misure dirette, calcolato tramite la relazione di propagazione per le misure indirette) delimita un intervallo entro cui si è *praticamente certi* debba trovarsi il valore della grandezza. Ciò significa che si può prevedere che, ove si ripeta la misura, il risultato di essa sarà compreso entro l'intervallo definito dall'errore.
Abbiamo anche già visto come attribuire un errore alla misura eseguita se il valore della grandezza varia in modo statistico. In questo capitolo voglio introdurre brevemente alcuni elementi di calcolo delle probabilità che possono essere di stimolo per ulteriori approfondimenti e fornire soprattutto lo spunto per alcune considerazioni critiche. Debbo sottolineare che i risultati tecnici presentati in questo capitolo sono stati in massima parte tratti da Young, 1970 e Kempthorne e Folks, 1971(vedi anche Giolanella, 2001, Staford e Vardeman, 1994, e Baldi, 1998).

La conoscenza di alcuni elementari metodi statistici per l'elaborazione degli errori di misura e per l'analisi delle osservazioni sperimentali è indispensabile agli scienziati ed a chi si interessa di misure tecniche. I concetti fondamentali di probabilità, di distribuzione e propagazione degli errori sono parte importante della conoscenza di chiunque si trovi a trattare dei numeri connessi a delle misure sperimentali.
Come ho già ampiamente sottolineato *il problema è trovare un numero che esprima il risultato della misura e un numero che ne esprima l'errore*, in modo che sia consentito fare previsioni circa il risultato di ulteriori misure. *Il calcolo delle probabilità permette appunto di fare previsioni su una base matematica*. Ciò non significa ovviamente che le osservazioni fatte in seguito debbano corrispondere rigorosamente alle nostre previsioni; significa semplicemente che si possono fare le migliori previsioni consistenti con i dati a disposizione. Il linguaggio del calcolo delle probabilità permette di esprimersi quantitativamente circa un fenomeno che può essere grandemente variabile.

3.1 La probabilità

Un concetto fondamentale definito nella teoria del calcolo delle probabilità è appunto quello di probabilità. Esistono principalmente due definizioni convenzionali accettate del concetto di probabilità. La prima è la cosiddetta *definizione classica* di probabilità e la si può enunciare nel modo seguente:

> *la probabilità, P(A), di un evento A è il rapporto tra il numero N di casi "favorevoli" (cioè il manifestarsi di A) e il numero totale M di risultati ugualmente possibili e mutuamente escludentesi:*

$$P(A) = \frac{N}{M}$$

Questa probabilità è talvolta detta probabilità *oggettiva* o anche probabilità *a priori*.

La seconda definizione è quella *frequentista* ed è la definizione sperimentale di probabilità come *limite della frequenza* misurabile in una serie di esperimenti. Essa ricalca lo schema della definizione classica, introducendo però un'importante variazione: sostituisce al rapporto numero casi favorevoli / numero di casi possibili il rapporto numero di esperimenti effettuati con esito favorevole / numero complessivo di esperimenti effettuati. Vediamo allora come viene definita questa probabilità:

la probabilità di un evento è il limite cui tende la frequenza relativa di successo all'aumentare del numero di prove.

In pratica, se abbiamo un esperimento ripetuto *m* volte ed un certo risultato *A* che accade *n* volte, la probabilità di *A* è data dal limite della *frequenza* (*n/m*) quando m tende all'infinito

$$P(A) = \lim_{m \to \infty} \frac{n}{m}$$

Gli *m* tentativi possono essere effettuati sia ripetendo in sequenza *m* volte lo stesso esperimento sia misurando simultaneamente *m* esperimenti identici. Se lanciamo una moneta in aria, sappiamo intuitivamente che la probabilità che venga testa è il 50%. Se gettiamo un ordinario dado, sappiamo ancora che la probabilità che esca 5 è 1/6.

Che cosa significa questo in realtà? Ad ogni lancio della moneta esce o testa o croce (non è concepibile l'eventualità che in un lancio esca per metà testa e per metà croce). Ciò che voglio dire, in effetti, è che, se lanciamo la moneta un numero molto grande di volte, il numero di volte che uscirà testa sarà all'incirca uguale alla metà del numero totale dei lanci. E se gettiamo un dado un gran numero di volte, il 5 uscirà press'a poco in un sesto delle prove.
Per gli scopi di questo capitolo utilizzeremo la definizione frequentista della probabilità. In altre parole, noi ci chiediamo quante volte un certo evento avviene, se facciamo un numero elevato di prove.
Si dovrebbe ben precisare che nel problema del lancio della moneta abbiamo stabilito soltanto che il rapporto fra il numero delle volte in cui esce testa e il numero complessivo dei tentativi si avvicina al valore di un mezzo, appena il numero dei tentativi *diventa elevato*.
Se sappiamo calcolare le probabilità per alcuni semplici eventi, come il lancio di una moneta o di un dado, possiamo calcolare probabilità per eventi più complicati, che risultino composti da questi semplici eventi: si noti che queste probabilità sono sempre numeri minori di 1. Se sommiamo le probabilità di tutti gli eventi che possono avvenire, otteniamo la probabilità che si verifiche una qualsiasi delle cose che possono avvenire, probabilità che naturalmente è uguale ad uno.

3.2 Il gioco delle tre scatole

Il concetto di probabilità è certamente uno dei più problematici nella storia della scienza. Nell'ultimo capitolo di questo scritto affronteremo ancora il problema. Voglio intanto far notare che anche la più utilizzata e famosa delle definizioni, quella frequentista, è tutt'altro che semplice ed ovvia da concepire. Moltissime persone, anche di spiccata intelligenza, che sono convinte di averla compresa, di fronte ad un primo problema pratico hanno difficoltà ad applicarla. Non mi riferisco alla eventuale complessità di calcolo di eventi composti, ma al concetto stesso appena introdotto. Per spiegarmi meglio descrivo brevemente il test che va sotto il nome di "gioco delle tre scatole".
Ci sono tre scatole sul tavolo ed io (diciamo) che conduco il gioco ho messo *precedentemente* un premio in una delle tre. Il vostro compito è quello di trovare il premio, ed il gioco avviene in due passi successivi. Vi chiedo di fare una prima scelta. Una volta che avete effettuato la vostra scelta su una delle tre scatole(che rimane chiusa), vi dico che in una tra le due rimanenti scatole sicuramente non c'è il premio, e la tolgo dal tavolo. A questo punto vi do la possibilità di conservare la vostra prima scelta o di spostarla sulla seconda scatola rimasta.
Il problema è il seguente: vi converrà mantenere la prima scelta o vi converrà cambiare?
La maggioranza delle persone (provare per credere) è convinta che non c'è nessun vantaggio nel cambiare ora la scelta precedentemente effettuata e, di norma, mantiene la scelta iniziale.
La risposta giusta è che invece bisogna cambiare: nella scatola rimasta, sulla quale vi do la possibilità di cambiare, vi è una probabilità di 2/3 di trovare il premio, e non ½ come erroneamente

molti ritengono. L'errore che si compie è quello di ritenere che, essendo la scelta che vi si propone quella tra due sole possibilità, entrambe debbano avere necessariamente la stessa probabilità.
Presento una delle possibili spiegazioni.
Se per caso la vostra prima scelta è quella giusta, dopo che io avrò aperto la scatola il cambio sarà *certamente* (non solo probabilmente) penalizzante. Al contrario, se la prima scelta è caduta sulla scatola vuota, sarete *certamente* (non solo probabilmente) avvantaggiati dal cambio. Quanto spesso vi capiterà di scegliere la scatola giusta (ed essere così *necessariamente penalizzati* dal cambio)? *Una* volta su tre. Quante volte vi capiterà di scegliere una scatola vuota (essendo così necessariamente favoriti dal cambio)? *Due* volte su tre.
Sono sicuro che molti di voi rimarranno perplessi. Comunque per una trattazione completa del gioco potete vedere Piattelli Palmarini, 1995 e Agnoli, Di Nicola, 2001.

3.3 Distribuzioni di probabilità

Per l'analisi di osservazioni complesse è spesso necessario trattare le probabilità di intere classi di eventi. A questo proposito introduciamo il concetto di *distribuzione di probabilità*.

Allo scopo di introdurre l'idea di distribuzione di probabilità, supponiamo di lanciare dieci monete simultaneamente. Con procedimento elementare si può determinare la probabilità che quattro monete presentino testa e le rimanenti sei croce. Orbene, supponiamo di chiederci: qual è la probabilità di ottenere 5 volte testa e 5 croce, o 7 volte testa e 3 croce, o, più in generale, di ottenere n volte testa e $(10 - n)$ croce, dove n è un intero compreso tra 0 e 10? La risposta a questa domanda è una serie di numeri, uno per ogni valore di n. Questi numeri possono essere visti come funzione di n, $f(n)$; cioè, per ogni n si ottiene un valore di $f(n)$ che fornisce la probabilità dell'evento caratterizzato dal numero n. Una funzione di questo tipo è chiamata «distribuzione di probabilità».

Una distribuzione di probabilità è sempre definita per un campo definito di valori dell'indice n. Nell'esempio sopra citato, n è un intero compreso tra 0 e 10. Se il campo di variabilità dell'indice include tutti i possibili eventi, e questo sarà generalmente il caso che si presenta nei nostri problemi, allora la somma di tutte le probabilità deve essere l'unità (la certezza). In questo caso,

$$\sum_{n} f(n) = 1$$

dove la somma si estende a tutto il campo dei valori n, proprio dello specifico problema che si considera.

Fino a qui ho trattato le distribuzioni di probabilità in base alla definizione di probabilità che consiste nel fare il rapporto tra il numero degli eventi favorevoli ed il numero, molto grande, di prove eseguite.

Anche le discussioni fatte sulla media e sulla deviazione standard si basano sull'assunzione che siano state eseguite molte prove.

Può tuttavia non essere immediatamente chiaro quale relazione intercorra tra queste quantità e i risultati che si sarebbero ottenuti se avessimo fatto un esperimento con un numero relativamente piccolo di prove. In poche prove, per esempio, la media molto probabilmente non è uguale alla media di un numero infinito di prove. La stessa cosa dicasi per la deviazione standard calcolata per un modesto numero di prove.

Per descrivere la differenza tra il numero infinitamente grande di prove pensato per definire $f(n)$ e un numero piccolo di prove qualunque in un esperimento reale, chiamiamo $f(n)$ la *distribuzione teorica infinita*, e i risultati di ogni gruppo di prove un *campione* di quella data distribuzione. E' chiaro che la media di un campione è solo una stima della media della distribuzione teorica infinita. Per alcuni tipi di distribuzione si può mostrare che la precisione di questa stima aumenta con le dimensioni del campione, ma è importante ricordare che essa non è mai niente più che una stima. *Ugualmente, la deviazione standard di un campione è una stima della deviazione standard della distribuzione teorica infinita.*

Inoltre, vi sono buone ragioni teoriche (vedi bibliografia consigliata all'inizio del capitolo), per stabilire che la formula di σ precedentemente introdotta non da ancora la migliore stima della deviazione standard della distribuzione teorica, che si può ottenere da un campione dato.

Risulta che una stima alquanto migliore è fornita da:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N - 1}}$$

che differisce dalla precedente per il fatto che la somma dei d_i^2 è divisa per *N*-1 invece che per N. Per dirla in breve, la ragione di questo fatto è che le deviazioni non sono tutte indipendenti; gli stessi dati sono stati prima adoperati per calcolar la media del campione, a sua volta usata per calcolare i d_i^2, e così il numero di deviazioni indipendenti è solo *N*-1. Sebbene questa modifica abbia un suo significato teorico, non è di solito di alcuna importanza pratica. Ordinariamente *N* è abbastanza grande, sicchè la deviazione standard del campione è influenzata molto poco dalla scelta di *N* o di *N*-1.

3.4 La distribuzione di Gauss, o "legge normale degli errori"

Passiamo ora a considerare una particolare distribuzione di probabilità (il cui nome abbiamo già introdotto più volte) e che è di grande importanza, la distribuzione di Gauss. Essa è importante per molte ragioni(vedi bibliografia introdotta all'inizio del capitolo), ma soprattutto, per quanto riguarda la nostra trattazione:

1) *sotto condizioni molto generali descrive la distribuzione degli errori casuali in differenti tipi di misure*;
2) *è possibile dimostrare che, anche se i singoli errori non seguono questa distribuzione, le medie di gruppi di siffatti errori sono distribuite in maniera da approssimare la legge di Gauss*, con la sola ipotesi che i gruppi siano abbastanza numerosi. Possiamo avere, per esempio, un gruppo di osservazioni che seguono la distribuzione *f(x)*; questa può essere una funzione qualsiasi, del tutto generica. Se prendiamo gruppi di *N* osservazioni e ne facciamo la media, allora, purchè *N* sia un numero molto grande, le medie seguiranno la distribuzione di Gauss. L'unica condizione da imporre è che sia finita la varianza della distribuzione *f(x)*. Questo risultato è conosciuto come il teorema del *limite centrale*, ed è molto importante negli sviluppi più avanzati della statistica matematica.

La distribuzione di Gauss può essere riguardata come un risultato analitico derivato da elementari considerazioni matematiche, o come una formula empirica che si è trovato essere in accordo con la distribuzione degli errori casuali che realmente intervengono in un data misura.

Qualcuno ha osservato a questo proposito, non senza una certa malizia, che tutti sono convinti che la distribuzione di Gauss descriva fedelmente il comportamento degli errori casuali, i matematici perché credono che i fisici l'abbiano verificata sperimentalmente, e i fisici perché credono che i matematici l'abbiano dimostrata su rigorose basi teoriche!(Young, 1970, 75)

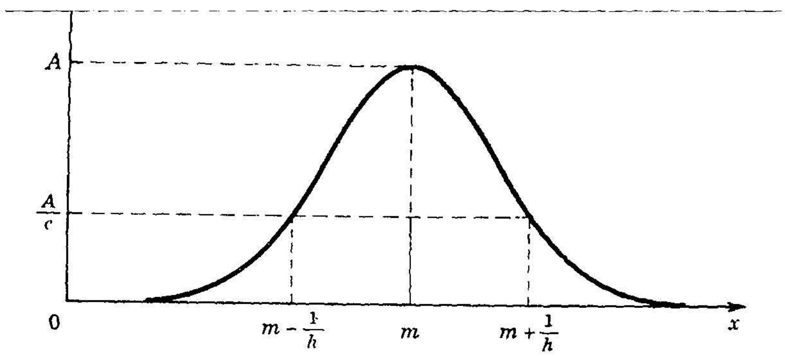
Da un punto di vista teorico si può fare la ragionevole affermazione che ogni errore accidentale si può pensare come il risultato di un gran numero di errori elementari, tutti di eguale entità, e ciascuno con uguale probabilità di produrre una variazione in eccesso o in difetto. In questo caso si può dimostrare che, se il numero degli eventi indipendenti(corrispondente appunto agli errori elementari) diventa elevato la distribuzione tende ad una gaussiana .

Molti sono tuttavia dell'avviso che *la vera giustificazione della validità della legge di Gauss e della*

sua applicazione nel descrivere il comportamento degli errori casuali è che molti gruppi di osservazioni sperimentali finiscono per obbedire ad essa. Questa argomentazione è più convincente di qualsiasi dimostrazione matematica. Per questo è valido il punto di vista secondo il quale questa distribuzione va trattata come un fatto sperimentale, nel senso che si stabilisce la sua espressione assiomaticamente e si esaminano quindi il suo significato e le sue applicazioni.
La funzione di distribuzione gaussiana è spesso chiamata la «legge normale degli errori» e gli errori che seguono tale distribuzione si dice che sono «distribuiti normalmente».
La forma della legge di Gauss è:

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

ove A, h e m sono delle costanti e x è il valore che si ottiene in una qualsiasi misura. Innanzi tutto, raffiguriamo $f(x)$ in un grafico per avere un'idea del suo andamento.



La figura è un diagramma della funzione di distribuzione di Gauss, la cui forma è data dalla f(x). Il massimo valore della funzione è *A, m* rappresenta il valore di x in corrispondenza del quale la funzione assume il massimo valore, ed h ha qualcosa a che vedere con il tipo di configurazione, piatta o slanciata, della curva a foggia di campana. Un valore elevato di h corrisponde ad una curva agile e appuntita, mentre un valore basso di h corrisponde ad una curva piatta e schiacciata.

Ci chiediamo, qual è il significato della funzione $f(x)$? Siamo tentati di dire che la $f(x)$ rappresenta la probabilità di osservare il valore x della quantità misurata. Ma ciò non è corretto. Ricordando che x è una variabile continua, facciamo notare che la probabilità che x assuma esattamente un qualsiasi particolare valore è zero. Dobbiamo invece cercare la probabilità che x assuma un valore compreso in un certo intervallo, ad esempio tra x e $x+\Delta x$.

Così l'esatta interpretazione della funzione $f(x)$ è che per un piccolo intervallo dx, $f(x)dx$ rappresenta la probabilità che una misura cada nell'intervallo compreso tra x e $x+dx$.

Ciò significa in altre parole che l'integrale $\int_a^b f(x)dx$ è la probabilità che una misura dia un risultato nell'intervallo $a \leq x \leq b$. Questo integrale può essere calcolato. In particolare, possiamo calcolare la probabilità che una misura cada entro una deviazione standard σ dal valore vero(potremmo ugualmente trovare la probabilità per un risultato per esempio entro 2σ etc.).
Si può dimostrare (vedi bibliografia introdotta all'inizio del capitolo) che la probabilità che una misura cada entro una deviazione standard dal valore vero è il 68%(si può anche dimostrare che per 2σ è 95.4% e per 3σ è 99.7%).

4. NUOVI SVILUPPI

Se volete apprendere qualcosa [dagli scienziati] sui metodi che essi usano, vi consiglio di tener fermo un solo principio: non ascoltate le loro parole in proposito ma fissate la vostra attenzione su ciò che fanno.
A. Einstein

If we were not able or did not desire to look in any new direction, if we did not have a doubt or recognise ignorance, we would not get any new ideas
R. Feynman

Nei capitoli precedenti ho cercato di introdurre quella che può essere definita la *teoria convenzionale degli errori*: come detto è quella che, di norma, si studia ancora oggi ed è riportata in dettaglio in almeno parte dei testi universitari.
In questo capitolo voglio brevemente delineare alcuni interessanti nuovi sviluppi che sono emersi in diversi campi scientifici e tecnici negli ultimi 15-20 anni e che stanno proponendo un nuovo approccio nella determinazione delle incertezze di misura, un approccio che non è esagerato definire un nuovo *paradigma,* una nuovo approccio alla teoria della misura. Debbo ricordare che molte delle informazioni di questo capitolo sono state raccolte soprattutto per la disponibilità di Giulio D'Agostini, professore di "Laboratorio di Fisica" alla Sapienza di Roma e convinto protagonista del nuovo approccio. In particolare ho intrattenuto con lui più di una conversazione sui suoi lavori e le sue pubblicazioni(da cui ho tratto parte delle mie considerazioni, insieme alle indicazioni di una aggiornata bibliografia al riguardo): vedi D'Agostini, Gennaio 1999, Luglio 1999, Dicembre 1999, Ottobre 2001 e 2002. Su gli aspetti discussi in questo capitolo si può anche vedere "*The International Society for Bayesan Analysis*", all'indirizzo http://www.bayesan.org/, Howson, Urban, 1993 e Lad, 1996.
E' utile presentare innanzi tutto alcuni problemi che scaturiscono dalla applicazione della teoria convenzionale.
Nei miei anni di lavoro come tecnico sperimentale ho potuto verificare personalmente che la situazione in cui ci si viene a trovare sulla base della teoria convenzionale *può portare a diversi motivi di insoddisfazione*. Lascio a D'Agostini fare un commento di carattere generale sullo stato attuale:

[…]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[…]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.(D'Agostini, Gennaio 1999, 5)

Sarebbero molti gli argomenti che si potrebbero analizzare per spiegare questa situazione. Mi limiterò qui a discutere criticamente due importanti risultati tecnici presentati nella prima parte di questo scritto, quelli che si riferiscono alle tecniche per la valutazione delle incertezze usate a proposito della propagazione degli errori casuali, sia massimi che statistici.

4.1 Critica della teoria degli errori casuali.

Measurement has meaning only if we can transmit the information without ambiguity to others

Russell Fox, *The Science of Science*

Riassumiamo le nozioni tipiche riportate nel par 2.5 e nel par 2.7.

Propagazione degli errori massimi

$$\Delta y \approx \left|\frac{\partial y}{\partial x_1}\right|\Delta x_1 + \left|\frac{\partial y}{\partial x_2}\right|\Delta x_2 + \ldots$$

ad esempio

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

Propagazione degli errori statistici

$$\sigma^2(y) = (\frac{\partial y}{\partial x_1})^2 \sigma^2(x_1) + (\frac{\partial y}{\partial x_2})^2 \sigma^2(x_2) + \ldots$$

per esempio

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

Cerchiamo di capire meglio su cosa sono fondate queste tecniche. Vediamo prima il caso degli errori massimi. Ho ricordato che

$$\Delta y = \sum_i \left|\frac{\partial y}{\partial x_1}\right|\Delta x_i$$

Questa espressione starebbe a significare che

> se siamo "praticamente certi" che il valore vero x_{vi} è compreso nell'intervallo dato da $x_i \pm \Delta x_i$, ne segue che siamo "praticamente certi" che il valore vero y_v, è compreso nell'intervallo dato da $y \pm \Delta y$.

Se accettiamo per buona tale espressione di "propagazione lineare degli errori massimi" e i presupposti sui quali essa si basa andiamo incontro però ad incongruenze. Riporto solo due semplici esempi.

1. 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 ± 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, giudicato generalmente *oggettivo*, in netta contraddizione con quanto si crede: *una conclusione paradossale!*

2. 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…±…°C
T_2 = 23.00…±…°C
T_2 - T_1 = ...±...°C

La risposta usuale a questo quesito è, che Δ_{T1} e Δ_{T2} sono pari a 0.6°C, mentre

$$T_2 - T_1 = 1.0 \pm 1.2\ °C.$$

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. 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.*

Riassumendo, si può affermare che l'uso della cosiddetta *teoria degli errori massimi* conduce, almeno, a *una tendenza a sovrastimare le incertezze*. (Ho qui presentato due davvero semplici esempi; ma le precedenti considerazioni sono ovviamente ancora valide per esperimenti molto complessi! Vedi per esempio D'Agostini, Gennaio 1999).
Qualcuno potrebbe pensare che la sovrastima delle incertezze sia da ritenere un pregio. *Ciò non è assolutamente vero.*
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 o di scoprire addirittura una nuova fenomenologia .
Vediamo ora il caso degli errori statistici. Riportiamo ancora il risultato introdotto al par 2.7.

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

Nella sua applicazione anche questa formula presenta i suoi problemi. Il più importante è quello di *interpretazione*. Questo è un punto cruciale. Cosa significa $y \pm \sigma(y)$? Come introdotto nel par…, assunto un modello gaussiano, ciò starebbe ad indicare

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

Cioè, c'è il 68 % di *probabilità* che il valor vero di y si trova nell'intervallo $y \pm \sigma(y)$.

Il problema nasce quando ci si chiede *cosa sia la probabilità*. Se utilizziamo le definizioni convenzionali otteniamo le risposte tipiche ("casi favorevoli su casi possibili" e "limite della frequenza") che *non contemplano affermazioni probabilistiche sui valori veri*, così come sono espresse dalla espressione introdotta. Vedrò di chiarire meglio questa critica nel prossimo paragrafo.

4.2 Il concetto di probabilità soggettiva.

If you toss a penny 10,000 times, it will not be heads 5,000 times, but more like 4,950. The heads picture weighs more, so it ends upon the bottom.
Da *The Herald Tribune*

If we were not ignorant there would be no probability,
there could only be certainty. But our ignorance cannot
be absolute, for then there would be no longer any probability
at all. Thus the problems of probability may be classed
according to the greater or less depth of our ignorance."
Henri Poincaré

Per ricostruire una teoria delle incertezza di misura che non soffra del tipo dei problemi mostrati, partiamo da due considerazioni.
Innanzi tutto non si può non sottolineare che l'unica affermazione sulla quale è difficile non essere d'accordo è quanto detto sin dall'inizio di questo scritto, ovvero che *il processo di induzione dalle osservazioni ai valori di grandezze fisiche conduce ad affermazioni che, inevitabilmente, sono affette da un certo grado di incertezza.* Il secondo punto è che 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.(D'Agostini,Gennaio 1999, 27)

Possiamo subito vedere che il secondo punto 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.
Quindi il primo concetto da discutere di nuovo è proprio quello di probabilità.
Prima di introdurre un nuovo concetto di probabilità, riflettiamo sul fatto che le "definizioni" standard non possono *da un punto di vista teorico* definire il concetto di probabilità in quanto esse *assumono il concetto di probabilità*. Prese alla lettera, esse sono infatti definizioni circolari:

- (*definizione classica)*

$$p \approx \frac{\text{numero dei casi favorevoli}}{\text{numero dei casi possibili(se ugualmente probabili)}}$$

- *(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).

La definizioni mostrano una evidente circolarità nella definizione *in quanto utilizzano il concetto di equiprobabilità per definire la probabilità*. La definizione frequentista inoltre è applicabile solamente quando sia possibile ripetere numerose volte l'esperimento, ma non è affatto chiaro che cosa debba intendersi per *numerose*.

Una critica famosa alla nozione di equiprobabilità così come utilizzata per esempio nella definizione classica è quella di De Finetti:

Tale giudizio di equiprobabilità [...] rispecchia una situazione di simmetria che viene spesso precisata obiettivamente dicendo che le palline debbono essere uguali, la moneta e il dado perfetti simmetrici fisicamente) ecc.; il criterio rimane tuttavia fondamentalmente soggettivo perché la scelta più o meno vaga di requisiti più o meno oggettivi da includere o no in tale concetto di "uguaglianza" non può che riflettere la distinzione soggettiva di ciascuno tra le circostanze che influiscono o non influiscono sulla sua opinione.(De Finetti, 1970, 235)

Questa critica è alla base dell'atteggiamento dei cosiddetti "soggettivisti"(vedi oltre) come qui riassunto.

La valutazione di equiprobabilità di due individui riflette quindi, secondo gli autori in esame, la distinzione puramente soggettiva delle circostanze che influiscono o non influiscono sugli eventi considerati. Pertanto se due individui pervengono alla medesima valutazione ciò dipenderebbe da ragioni psicologiche la cui validità è più o meno fortuita e che comunque non possono in alcun modo imporsi a livello razionale. Il fatto è che, sempre secondo gli anzidetti autori, non esisterebbe alcuna ragione, né logica né metafisica, in grado di costringere due individui a considerare allo stesso modo le circostanze e quindi ad arrivare alla medesima valutazione di equiprobabilità.(Costantini,Geymonat,1982, 65)

Oltre queste considerazioni di carattere generale, voglio subito sottolineare che, dal punto di vista storico, le due "definizioni" di probabilità appena incontrate, hanno indotto molti

a confondere il concetto di probabilità con il suo metodo di valutazione, come si può verificare in vari libri di testo e voci di enciclopedia;
ritenere che il valore di probabilità sia oggettivo, cioè che sia insito nella natura delle cose e non dipenda da chi lo valuta;
credere che si possa parlare di probabilità solo in due casi, molto particolari e riduttivi, rispetto alla complessità del mondo reale. Sarebbero quindi esclusi da argomentazioni probabilistiche tutti quegli eventi (la stragrande maggioranza di quelli di interesse pratico e scientifico) per i quali è impossibile eseguire l'inventario dei casi possibili e di quelli favorevoli o per i quali non è possibile ripetere "infinite volte" l'esperimento nelle identiche condizioni.(D'Agostini, Ottobre 2001,23)

Che succede, in altre parole, se tutti i risultati possibili non sono ugualmente probabili? Con alcuni dadi progettati per certi scopi, alcuni risultati sono più probabili di altri: è più probabile, infatti, che un dado truccato cada su di un lato con maggiore frequenza rispetto agli altri.

Supponiamo che, in una coppia di dadi, uno sia truccato - bilanciato, cioè, in modo che un numero esca con maggiore frequenza degli altri. In questo caso le regole di calcolo più semplici non possono essere utilizzare per accertare le probabilità. Immaginiamo, ad esempio, di sapere che un dado è stato truccato in modo che il "sei" esca con maggiore probabilità rispetto agli altri numeri. In tale situazione, se tutti e due i dadi vengono lanciati, sappiamo che un doppio sei uscirà probabilmente con maggiore frequenza di 1 su 36, che, invece, è la probabilità corretta con cui dovrebbe uscire lanciando due dadi ben equilibrati. Ma di quanto è più probabile?

One way that we can find out is by tossing the dice repeatedly, counting the number of times we get double sixes, and expressing that count as a proportion of the number of tosses. This is called the *long-range relative frequency* definition of probability. It is clear that when the dice are not loaded, the probability of double sixes we get by the method of counting sides will be duplicated by a long-range relative frequency, also equal to 1/36 (assuming that the dice are cast a very large number of times). The long-range relative frequency definition of probability relates the important construct of mathematical probability to something that can be observed empirically, and thereby marries it to empirical science.

For all its advantages, the long-range relative frequency definition of probability made the term (and the laws for manipulating probabilities) applicable only to situations that could, conceptually at least, be repeated many times. But there are many other sorts of situations, ones that are unique in that they cannot, even conceptually, be repeated many times, to which we often want to apply the notion of probability. For example, we may be interested in the probability that this scientific theory predicts future observations well[...]the probability that it will rain during the next hour, or the probability that our candidate will win in the upcoming election. These events will occur only once. It became clear that the definition of probability needed to be expanded to one that that would cover all such situations and would include the earlier definitions as special cases.(Press, Tanur, 2001, 201-202)

Vediamo di introdurre una nuova definizione, assumendo in sostanza 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";(D'Agostini, Gennaio 1999, 28)

Una definizione che vada oltre il concetto intuitivo e che non sia limitata ai casi *limite* degli eventi equiprobabili o delle prove effettuate in condizioni di equiprobabilità può far ricorso al concetto di scommessa, percepibile a livello intuitivo da tutte le persone razionali.

Una formulazione accettata, detta appunto *probabilità soggettiva*, è per esempio la seguente:

la probabilità di un evento A è la misura del grado di fiducia che un individuo coerente attribuisce, secondo le sue informazioni e opinioni, all'avverarsi di A.

Il campo di applicabità di questa definizione è molto vasto; occorre aggiungere che la "coerenza" significa *la corretta applicazione delle norme di calcolo*. Vediamo meglio che cosa si intende con ciò. Il concetto di probabilità soggettiva potrebbe suggerire che ognuno possa assegnare una probabilità qualsiasi a un evento. Non è così. Infatti si introduce la seguente "regola": il soggetto è tenuto ad accettare tutte le scommesse che si riferiscono agli eventi stessi, alle loro negazioni e alle loro composizioni (una scommessa su, cioè, più di un evento); le probabilità associate agli eventi non devono essere modificate di volta in volta se le informazioni in possesso non mutano. Questa "regola" viene detta *assioma di coerenza* e serve a far sì che una persona non cambi la propria probabilità soggettiva per tornaconto personale. In altre parole possiamo dire che 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. Per un approfondimento del concetto di probabilità soggettiva e del concetto di coerenza vedi Scozzafava, 2001, Ross, 2000, De Finetti, 1989 e 1995.

Soggettiva sta ad indicare che la valutazione di probabilità dipende dallo stato di informazione del soggetto che la esegue, e, anche se basata su una credenza specifica, *non è affatto arbitraria*: il ruolo normativo della scommessa coerente obbliga a tener conto di tutte le informazioni a disposizione.

The term *subjective* used here implies that this construct is based on a belief specific to an individual. The other definitions of probability given above are special cases of subjective probability. The belief could be that the mathematical probability (an enumerative proportion) is what is called for in a given instance. Or it might be that the long-range relative frequency (the proportion of times some event occurs within many trials) is what is called for in the belief system of the person making the probability statement. Or it could be that neither definition is appropriate, and the subjective probability statement instead implies a belief of the individual based on his or her scientific knowledge and deep understanding of some underlying biological, physical, or social phenomenon. (Press, Tanur, 2001, 202)

Ed inoltre è importante notare che

La conoscenza che altri valutano diversamente la probabilità dello stesso evento è una informazione che andrebbe utilizzata. Ma non tanto per mediare le due stime quanto piuttosto per capire meglio il problema e comportarsi alla fine nel modo più coerente possibile come se veramente si dovesse scommettere del denaro su o contro quell'evento. La così detta "oggettività", come è percepita da chi è al di fuori della ricerca scientifica, viene recuperata quando una comunità di esseri razionali condivide lo stesso stato di informazione. Ma anche in questo caso, si dovrebbe parlare più propriamente di *"intersoggettività"*.(D'Agostini, Ottobre 2001, 31)

Si noti inoltre che questa definizione *non è affatto in contrasto con le due precedenti*, nel senso che le due "pseudo-definizioni" sono prontamente recuperate come regole di *valutazione* della probabilità, qualora colui che effettua la valutazione ritenga che le clausole presupposte dalle definizioni convenzionali siano soddisfatte.
Il concetto di probabilità soggettiva dipende quindi dallo stato di informazione del soggetto che effettua la valutazione. Questa formulazione è anche riconosciuta dalla 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.(ISO, 1993, 35)

E in effetti la principale differenza tra questo concetto di probabilità ed una presunta definizione oggettiva è che essa non è una caratteristica intrinseca di un evento, ma dipende dallo stato di informazione disponibile a chi esegue la valutazione. In una affermazione provocatoria De Finetti (De Finetti, 1970) afferma: *La probabilità non esiste*.

The objectivist formulation of probability presumes the existence of a "real" world", which is beyond the pale of our ability to experience it and to measure it directly[…]The subjectivist formulation of probability and statistics is based upon a radical break with this conception of empirical science as yielding approximate results about a real world beyond our experience. We presume that scientific investigation constitute our studies about the world of human experience, not some supposed world outside of experience. Empirical measurements are merely the recording of human activity in the world, conducted according to operationally described procedures.(Lad, 1996, 11)

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").

We could make an analogous argument about any quantity unknown to you, such as the height of the next person you might sce accidentally in the street or whether a particular scientific theory is true. Your personal belief about the height of the person in the street would most certainly not be that the height is greater than 10 feet, or as small as 5 inches; you know in advance that people are not that tall or small. So you could say that your subjective or personal probability that the height of the person is less than 5 inches or greater than 10 feet is zero. But your subjective probability might be that there's a 50% chance that the person's height is within an inch or so of 5 feet 8 inches, that there's a 25% chance that the person's height is less than 5 feet 4 inches, and a 15 % chance that the person's height will exceed 6 feet. These subjective probabilities are based on your own previous observations, experiences, and understandings about human growth and development. Indeed, if you thought that the next person you will see is, in some sense, randomly chosen from the population, you might use all the published information about the distribution of heights in the population to form your personal probability.(Press, Tanur, 2001, 203)

Questa interpretazione, va notato, negli ultimi anni si sta lentamente imponendo in molti ambienti in base alla sua pragmaticità (si ispira a metodi che pongono l'accento sulla decisione piuttosto che sulle argomentazioni in base alle quali la decisione stessa è presa) e differenti ragioni.

Una di queste fu l'uso sempre più ampio di nozioni probabilistiche nella gestione delle aziende industriali e la conseguente posizione privilegiata che, almeno nei paesi occidentali, segnatamente in quelli di lingua inglese, vi assumono le decisioni personali. Va ricordato infatti che nonostante i "padri" europei, il moderno soggettivismo prese vigore e si irradiò in modo quasi travolgente dalla nazione guida del mondo capitalistico occidentale. L'altra fu il richiudersi delle concezioni oggettivistiche - e ciò vale in particolar modo per il frequentismo, la concezione più seguita da fisici e statistici - in un dogmatismo sempre più sordo ai progressi scientifici, in particolare ai sempre nuovi usi delle nozioni probabilistiche.(Costantini,Geymonat,1982,62)

Il concetto di probabilità soggettiva è basato sull'idea che la probabilità è legata allo stato di incertezza e non soltanto al risultato di esperimenti ripetuti.
Questo punto di vista, che corrisponde al significato originale di "probabile", è quello che in effetti sostenevano (vedi per esempio Hacking, 1975 and Stigler, 2000) scienziati e filosofi come Bernoulli, Hume, Laplace e Gauss, prima di essere abbandonato in periodo neopositivista.
Si può notare che anche prima, nel 1662, la *Logica* di Port Royal, scritta da Antoine Arnauld ed altri, finisce con una discussione sulle *credenze ragionevoli* e la credibilità (o grado di fiducia).
Inoltre, è interessante sottolineare che nel punto di vista di Hume sul concetto della probabilità il concetto stesso e le valutazioni sono chiaramente separati.

Though there be no such thing as Chance in the world; our ignorance of the real cause of any event lias the same influence on the understanding, and begets a like species o£ belief or opinion.There is certainly a probability, which arises from a superiority of chances on any side; and according as this superiority increases, and surpasses the opposite chances, the probability receives a proportionable increase, and begets still a higher degree of belief or assent to that side, in which we discover the superiority. If a dye were marked with one figure or number of spots on four sides, and with another figure or number of spots on the two remaining sides, it would be more probable, that the former would turn up than the latter; though, if it had a thousand sides marked in the same manner, and only one side different the probability would be much higher, and our belief or expectation of the event more steady and secure. This process of the thought or reasoning may seem trivial and obvious; but to those who consider it more narrowly, it may, perhaps, afford matter for curious speculation.(Hume, 1957, 78; vedi anche http://www.utm.edu/research/hume/wri/1enq/1enq-6.htm)

La "dualità"della probabilità è stata di fatto riconosciuta da molto tempo.

Carnap said we ought to distinguish a 'probability$_1$' from a 'probability$_2$'; later he spoke of inductive and statistical probabilities. Poisson and Cournot said we should use the ready-made French words *chance* and *probabilité* to mark the same distinction. Before that Condorcet suggested *facilité* for the aleatory concept and *motif de croire* for the epistemic one […] Bertrand Russell uses 'credibility' for the latter […]. There have been many other words. We have had *Zuverlässigkeit*, 'propensity', 'proclivity', as well as a host of adjectival modifiers of the word 'probability', all used to indicate different kinds of probability.(Hacking, 1975, 13)

Concludendo voglio sottolineare che i diversi approcci, pur perseguendo degli scopi prettamente scientifici, si sono anche impegnati *nel problema filosofico di delineare una concezione generale di probabilità*. L'indirizzo classico e quello frequentista hanno sostenuto che la probabilità possiede una natura oggettiva; al contrario, l'indirizzo soggettivista ha sostenuto che la probabilità possiede una natura esclusivamente soggettiva.
Queste seppur brevi considerazioni giustificano quindi davvero l'affermazione che *il problema della natura della probabilità, e la relativa scelta tra oggettivismo e soggettivismo, sono genuinamente filosofici.*

4.3 Il teorema di Bayes

An effect may be produced by the cause a or by the cause b
The effect has just been observed. We ask the probability that
it is due to the cause a. This is an à posteriori
probability of cause. But I could not calculate it, if

a convention more or less justified did not tell me in advance
what is the à priori probability for the cause a
to come into play. I mean the probability of this event to
some one who had not observed the effect.
Henri Poincaré

Our knowledge is the amassed thought and experience of innumerable minds.
Ralph Waldo Emerson

I nuovi sviluppi della teoria degli errori sono basati sul concetto di probabilità soggettiva e sul *teorema di Bayes*. Per introdurre quest'ultimo cercherò di evitare il più possibile i formalismi, nondimeno dovremo familiarizzare con un minimo di 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, che ora indicherò con 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 la deviazione standard σ 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

$$f(x) \rightarrow f\,(x|\mu = 1000.00)$$
$$f(\mu) \rightarrow f\,(\mu|x = 2.315),$$

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.
Ogni misura è eseguita allo scopo di accrescere la conoscenza di chi la esegue, si tratti di uno scienziato, un ingegnere che deve valutare/collaudare un prodotto o un apparato, un medico che ha prescritto una certa analisi. 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, l'affidabilità di una apparecchiatura per trasmissione dati, lo stato di salute di un paziente, etc…* 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 termometro 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'aggiornamento è nullo.(D'Agostini, Gennaio 1999, 32)

Molto spesso si pensa che l'unico metodo scientifico valido sia quello della falsificazione. Ma non è sempre così.

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 alle incertezze di misura. Infatti, prendendo alla lettera tale metodo, si sarebbe autorizzati soltanto a verificare se il valore osservato sullo strumento è compatibile o no con un valore vero, niente di più. Si capisce come, con queste premesse, non si possa fare molta strada.(Ibidem)

In generale, non si può applicare il metodo della falsificazione alle ipotesi statistiche, l'area in cui peraltro la metodologia è di maggior importanza per gli scienziati praticanti.

Many scientific theories are explicity probabilistic and, for this reason, have no logical consequences of a verifiable character. An example is Mendel's theory of inheritance. This states the probabilities with which certain combinations of genes occur during reproduction; but, strictly speaking, the theory does not categorically rule out, nor predict, any particular genetic configuration. Nevertheless, Mendel obtained impressive confirmation from the results of his plantgrowing trials, results which this theory did not entail but stated to be relatively probable.(Howson, Urbach, 1993, 7)

In altre parole, i principi sui quali si basa Popper sono troppo deboli per ristringere sufficientemente il campo di ipotesi alternative, laddove, in pratica, gli scienziati hanno un sistema di classificazione delle ipotesi secondo la loro eleggibilità per una seria considerazione Tale sistema è stato tradizionalmente caratterizzato come uno spettro di probabilità.

Have we any right, for istance, to enunciate Newton's Law? No doubt numerous observations are in agreement with it, but is not that a simple fact of chance? And how do we know, besides, that this law which has been true for so many generations will not be untrue in the next? To this objection the only answer you can give is: It is very improbable[…] From this point of view all the sciences would only be unconscious applications of the calculus of probabilities. And if this calculus be condemned, then the whole of the sciences must also be condemned.(Poincaré, 1952, 186; cited in Howson, Urbach, 1993, 9)

Per formalizzare il discorso intrapreso, 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.
Come detto,

$f(x|\mu)$

sta per la funzione densità di probabilità (x è una variabile continua, dal punto di vista pratico) di osservare 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 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.
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'osservazione:

$$f(\mu \mid x) \propto f(x \mid \mu) \cdot f_0(\mu)$$

Questo è uno dei modi di scrivere il teorema di Bayes, che ha peraltro un ruolo centrale in tutte le inferenze probabilistiche (per approfondimenti vedi per esempio Box, 1992). Il teorema prende il nome da un ministro della chiesa Presbiteriana Riformista che nel 1763 pubblicò un documento(*Essay toward solving a Problem in the Doctrine of Chances*) nei *Philosophical Transactions of the Royal Society* in cui ci si serve di un metodo atto ad elaborare operazioni statistiche basandosi su una prima comprensione di un fenomeno e combinando formalmente quella prima visione con dati misurati al momento in modo da aggiornare l'opinione scientifica di colui che conduce l'esperimento. Il punto centrale dell'articolo (vedi Lad, 1996) è una teorema che identifica come le prove miste di una sequenza di esperimenti possano essere utilizzate per derivare le probabilità circa il risultato del prossimo esperimento, dettagliando precisamente come le prove di 100 osservazioni possano essere riconosciute come diverse dalla prova di una singola di esse.
Come già sottolineato *questo tipo di approccio all'inferenza statistica ha quindi una storia lunga*, e cominciò ad essere messo in discussione solo con l'avvento del neopositivismo per essere messo in disparte all'inizio del XX secolo.

A large international school of scientists preceded, supported, expanded, and developed Bayesian thinking about science. These included such famous scientists as James Bernoulli in 1713 (even before the paper by Bayes was published), Pierre Simon de Laplace in 1774, and many twentieth century scientists such as Bruno de Finetti, Harold Jeffreys, L.J. Savage, Dennis V. Lindley, and Arnold Zellner. Bayesian methodology was the method of statistical inference generally used from the time of Bayes until the early part of the twentieth century, when Sir Ronald A. Fisher and others introduced the *frequentist approach* to statistical inference.(Press, Tanur, 2001, 204)

Ed è facile notare che il successo dell'"approccio frequentista" sull'onda dell'affermazione della cultura neopositivista rappresentò, in questo campo, uno sviluppo originale di quella ricerca di "precisione" e razionalità iniziata con la società moderna e sulla quale a lungo mi sono soffermato nella prima parte di questo scritto. In quest'ultimo capitolo sto poi cercando di mostrare in particolare che questo però non è certo l'unico sviluppo possibile che può scaturire da quell'atteggiamento di fondo, e che i risultati che ne scaturiscono per quanto riguarda la stima delle incertezze di misura non sono i più affidabili.
Forse è di un certo interesse sapere che uno dei primi fisici del XX secolo a riflettere sull'approccio Bayesiano alle misure fu Enrico Fermi. Ricorda un suo allievo alla University of Chicago nel 1947.

I have one more personal example of how Fermi left his mark on the entire international physics community[…]most physicist were not very knowledgeable about statistical inference. In my thesis I had to find the best 3-parameter fit to my data and the errors of those parameters in order to get the 3 phase shifts and their errors. Fermi showed me a simple analytical method. At the same time other physicists were using and publishing other cumbersome methods. Also Fermi taught me a general method, which he called Bayes Theorem, where one could easily derive the best-fit parameters and their errors as a special case of the maximum-likelihood method. I remember asking Fermi how and where he learned this. I expected him to answer R.A. Fischer or some textbook on mathematical statistics. Instead he said «perhaps it was Gauss.»[…]Frank Solmitz, a fellow grad student, and I felt we should get down on paper what we were learning from Fermi. So with help from Fermi and Frank Solmitz I pulled all this together a few years later in a 1958 UCRL report called «Notes on Statistics for Physicists».(Orear, 2001, 32-33)

Dal teorema di Bayes nasce la statistica Bayesiana, *basata appunto sull'idea intuitiva che la probabilità quantifica il grado di fiducia nell'occorrenza di un evento*. E la statistica diventa allora *uno strumento logico* per aggiornare la probabilità *alla luce di tutti i differenti tipi di informazione. Ciò permette, per esempio, di trattare tutti i differenti contributi all' incertezza di una misura alla stessa maniera (* vedi paragrafo 4.5)

Forse è utile mostrare un ragionamento intuitivo che aiuta a capire meglio 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 posizione 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. Spesso, in semplici misure di routine il modello cacciatore → cane e cane → cacciatore funziona secondo l'inversione intuitiva che abbiamo descritto. Ci sono però dei problemi (specialmente in fisica) in cui questo è tutt'altro che vero.

Molti possono rimanere perplessi al pensiero che le conclusioni scientifiche possano dipendere dal "pregiudizio" sulla grandezza fisica. Ma anche se pregiudizio ha nel linguaggio corrente un significato prevalentemente negativo, in realtà significa semplicemente un giudizio basato su una esperienza acquisita in precedenza. Secondo il punto di vista Bayesiano, il ragionamento scientifico, e almeno in parte quello quotidiano, vengono condotti in termini probabilistici. In altre parole, quando una persona valuta un affermazione incerta, lo fa calcolando la probabilità dell'ipotesi alla luce delle informazioni conosciute.

Notiamo ora, come esempio che può aiutare a comprendere in che modo *i pregiudizi siano comunque sempre presenti nelle nostre valutazioni*, che la stessa esperienza visiva che l'osservatore ha quando guarda un oggetto dipende dalla sua esperienza passata, dallo stato delle sue conoscenze e dalle sue aspettative.

Ecco due esempi che illustrano bene questo punto.

In un esperimento molto noto, a dei soggetti venivano mostrate in rapida successione delle carte da gioco e veniva loro chiesto di identificarle. Finché si utilizzava un normale mazzo di carte, tali individui erano capaci di soddisfare questa richiesta. Non appena, però, venivano introdotte carte irregolari, ad esempio un asso di picche rosso, allora sulle prime quasi tutti i soggetti identificavano da principio quelle carte irregolari come carte normali: vedevano così un asso di picche rosso come se fosse un normale asso di quadri o un normale asso di picche. Le impressioni soggettive esperite dagli osservatori erano influenzate dalle loro attese. Quando, dopo un periodo di confusione, i soggetti cominciavano a rendersi conto oppure venivano avvisati della presenza nel mazzo di carte anomale, allora non avevano più alcuna difficoltà a identificare correttamente tutte le carte che venivano loro mostrate, sia che fossero anomale sia che non lo fossero. Un mutamento nelle loro cognizioni e nelle loro aspettative era accompagnato da un mutamento in ciò che vedevano, benché guardassero sempre gli stessi oggetti fisici.

Un altro esempio è fornito da un gioco infantile che consiste nel trovare i caratteri di un volto umano confusi tra il fogliame nel disegno di un albero. In questo caso ciò che si vede, vale a dire l'impressione soggettiva esperita da una persona che osserva il disegno, corrisponde sulle prime a un albero con relativi fusto, foglie e rami. Ma tutto questo cambia non appena viene rintracciato il volto. Ciò che prima veniva visto come fogliame e parti di rami è visto adesso come volto umano. Ancora una volta, lo stesso oggetto fisico è stato osservato prima e dopo la soluzione dell'indovinello ed è presumibile che l'immagine sulla retina dell'osservatore non cambi quando si è trovata la soluzione e il volto è stato scoperto. Inoltre, se la figura viene guardata dopo qualche tempo, un osservatore che abbia già sciolto una volta l'indovinello riconoscerà di nuovo il volto facilmente. In questo esempio, quel che uno spettatore vede è condizionato dallo stato delle sue conoscenze e dalla sua esperienza.(Chalmers, 1986, 34)

Non sarebbe difficile produrre esempi tratti dalla pratica scientifica e dalla storia della scienza che mettono in chiaro questo stesso punto, vale a dire che ciò che gli osservatori vedono, le esperienze soggettive che hanno quando osservano un oggetto o una scena, non è determinato soltanto dalle immagini sulle loro retine, ma dipende anche dall'esperienza, dalle cognizioni, dalle attese e dallo stato interiore complessivo di chi osserva.

Un ricercatore che sperava di sapere qualcosa circa l'opinione degli scienziati a proposito della natura della teoria atomica, chiese a un famoso fisico e a un eminente chimico se un singolo atomo di elio fosse una molecola o no. Entrambi risposero senza esitare, ma le loro risposte non furono identiche. Per il chimico l'atomo di elio era una molecola, perché si comportava come tale rispetto alla teoria cinetica dei gas. Per il fisico, d'altra parte, l'atomo di elio non era una molecola, perché non presentava nessuno spettro molecolare. Presumibilmente entrambi parlavano della stessa particella, ma essi la consideravano secondo le differenti prospettive delle loro rispettive formazioni ed attività di ricerca. Le loro esperienze nel risolvere problemi avevano insegnato loro che cosa doveva essere una molecola. Senza dubbio le loro esperienze avevano avuto molti punti in comune, ma in questo caso, portavano i due specialisti a conclusioni diverse.(Kuhn, 1969, 73-74)

Ci volle più di un secolo e mezzo dopo la scoperta degli spermatozoi perché si arrivasse a capirne il ruolo nella fecondazione, e dovettero passare duecento anni dalla scoperta dei gameti maschile e femminile perché il concetto di fecondazione venisse formulato correttamente. Si vede qui bene quale parte giocassero le rappresentazioni mentali dei vari studiosi nell'interpretazione delle loro osservazioni, *cioè, in ultima analisi, nell'edificazione del sapere scientifico.*

[…]pure observation does not create, or increase, knowledge without personal inputs which are needed to elaborate the information. In fact, there is nothing really objective in physics, if by objective we mean that something follows necessarily from observation, like the proof of a theorem. There are, instead, beliefs everywhere. Nevertheless, physics is objective, or at least that part of it that is at present well established, if we mean by 'objective', that a rational individual cannot avoid believing it. This is the reason why we can talk in a relaxed way about beliefs in physics without even remotely thinking that it is at the same level as the stock exchange, betting on football scores, or... New Age.(D'Agostini, July 1999, 124)

Gli scienziati più famosi del mondo hanno condizionato le loro metodologie di ricerca con le loro credenze, le loro opinioni e le loro intuizioni personali in diversi modi, spesso senza rendersene conto o senza darne esplicita e formale comunicazione. Tale soggettività comprendeva idee, intuizioni, credenze forti, interpretazioni scientifiche di principi di base, che essi avevano elaborato e acquisito durante i propri studi o su cui avevano ricevuto adeguate informazioni: si fondavano, cioè, su osservazioni precedentemente elaborate da loro stessi o da altri.
Le conclusioni scientifiche sono, pertanto, *il risultato della fusione delle credenze e delle*

interpretazioni presenti prima di effettuare esperimenti con l'analisi ben più oggettiva di dati sperimentali.

E' necessario imparare a guardare con occhio esperto nel telescopio o al microscopio: i raggruppamenti non strutturati di macchie scure e luminose che vede il principiante sono altra cosa dalle forme e spettacoli minuziosi che l'osservatore competente è in grado di discernere. Qualcosa di simile deve essere accaduto quando Galileo per primo introdusse il telescopio come strumento utile all'esplorazione dei cieli. Le riserve avanzate dai rivali di Galileo nell'accettare fenomeni quali le lune di Giove da lui scoperte devono essere attribuite in parte non a pregiudizi, ma a positive difficoltà nel vedere attraverso quello che, dopo tutto, era un telescopio alquanto rudimentale.(Chalmers, 1986, 35)

Quest'ultima citazione mi permette inoltre di sottolineare che la stessa scelta della strumentazione di laboratorio è ovviamente influenzata dalle nostre aspettative.

In breve, la decisione di usare una particolare apparecchiatura e di usarla in un particolare modo indica che si dà per scontato in forma più o meno cosciente che si dovranno verificare solo circostanze di un certo tipo. Vi sono, oltre alle aspettative teoriche, anche quelle strumentali, e queste hanno spesso svolto un ruolo decisivo nello sviluppo scientifico.(Kuhn, 1969, 83)

Per dirla con una espressione di Koyré che mi piace molto(vedi per esempio Koyrè, 1950, 39), gli strumenti sono davvero "*incarnazioni" della teoria*. Con ciò ovviamente non si vuole affermare che la scienza debba abbandonare la strumentazione convenzionale, ma mettere in luce che le relative scelte inevitabilmente restringono il campo dei fenomeni accessibili all'indagine scientifica. Come già accennato ora si sta proponendo un nuovo paradigma scientifico, all'interno del quale si vuole tener conto delle credenze pregresse *in maniera per quanto possibile formale*. Il relativo formalismo parte proprio dal teorema di Bayes e da alcune conseguenze teoriche che da esso scaturiscono. Nei prossimi due paragrafi, pur evitando i dettagli tecnici, cercherò di presentare meglio la portata e le possibili conseguenze scientifiche di questo cambiamento. Voglio qui sottolineare quindi, anche se questa tesi credo sia ormai chiara per quanto detto sinora in questo lavoro, che questo nuovo paradigma è anch'esso figlio legittimo di quegli ideali di razionalità e precisione che hanno caratterizzato la nascita della società moderna. E' questo però un approccio che non cerca rifugio in un concetto ideale di "oggettività", ma cerca di fare i conti con il modo concreto con cui scienziati o ricercatori operano o possono operare al meglio nel valutare i risultati dei loro esperimenti e delle loro osservazioni.
E non è forse inutile sottolineare che l'effettivo successo di tale paradigma sarà legato anche a considerazioni economiche, sociali e politiche sulle quali mi è prematuro riportare ora considerazioni affidabili(anche se sono da tempo impegnato personalmente a studiare e riflettere in merito). Posso però ancora ricordare, come accennavo nel paragrafo precedente, che ponendo accento soprattutto sul *momento decisionale* questo approccio sembra poter offrire *una maggiore pragmaticità* nel risolvere problemi in differenti ambienti tecnici, industriali ed aziendali.

Many scientists now believe that a paradigm shift in the sense of Kuhn[…] has been taking place in the way that scientific inference is, and will be, carried out. Many scientists now recognize the advantages of bringing prior beliefs into the inference process in a formal way from the start. The alternative is attempting to achieve total objectivity, or pretending to have done so, even though prior information is often brought to bear on the problem anyway, in surreptitious or even unconscious ways. Subjectivity may enter the scientific process surreptitiously in the form of seemingly arbitrarily imposed constraints, in the introduction of initial and boundary conditions, in the arbitrary levels of what should be called a statistically significant result, in the deemphasizing of certain data points which represent suspicious observations[…], and so on.(Press, Tanur, 2001, 204)

Le conoscenze pregresse e le relative aspettative solo apparentemente non hanno una parte rilevante nel comporre ed analizzare un esperimento. Anche i cosiddetti "dati grezzi" sono per loro natura interpretati. E per quanto riguarda la successiva valutazione di questi dati, anche se spesso tacitamente si suppone che le appropriate procedure siano neutrali rispetto alla teoria ciò *non è evidentemente vero.* Nella complessa valutazione delle prove esse sono applicate sensatamente solo da coloro che comprendono tutti i dettagli dell'esperimento.

Gli esperimenti iniziano e finiscono in una matrice di credenze[…]credenze sul tipo di strumentazione, sui programmi di ricerca, sui giudizi individuali relativi ad ogni comportamento locale di pezzi di apparato.(Galison, 1987, 45)

Nel laboratorio nulla è semplicemente dato. *Le misure sono prese, non date.*
La scienza senza credenze pregresse non ha senso.

4.4 Possibili campi di applicazione

[…]much qualitative research, both empirical and theoretical, is normally prerequisite to fruitful quantification of a given research field. In the absence of such prior work, the methodological directive, "go ye forth and measure," may well prove only an invitation to waste time.
Thomas S. Kuhn

Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.
H. G. Wells

I possibili campi di applicazione dell'inferenza bayesiana sono molteplici, nelle scienze come in tutti i processi decisionali. Per avere una interessante panoramica di differenti esempi si può senz'altro vedere Press, Tanur, 2001 e Bernardo, Smith, 2000 e Bradley, Thomas 2000.

The allocation of billions of U.S. federal dollars now depends on Bayesian estimates of population characteristics (e.g. poverty) in small geographic areas. The FDA now not only permits but encourages Bayesian designs and analyses in applications seeking approval for new medical devices such as pacemakers, an area where large clinical trials are infeasible but directly relevant historical information is often available. The Microsoft Windows Office Assistant is based on a Bayesian artificial intelligence algorithm- similarly, a prominent movie-rental chain now offers a web-based Bayesian expert system to help you choose which film to rent, based on your general preferences and the names of other films you have enjoyed.(Bradley, Thomas, 2000, viii)

La statistica Bayesiana si può di fatto applicare ogniqualvolta si voglia valutare la probabilità (grado di fiducia) che un particolare evento possa avverarsi, denotata per esempio con P{event}, quando questo evento è generato da un meccanismo *random*. Il *frequentista* valuta le procedure basandosi sull'immaginazione della raccolta ripetuta di campioni da un particolare modello che definisce la distribuzione della probabilità dei dati osservati condizionata da parametri sconosciuti. La statistica Bayesiana richiede un modello e, in più, una distribuzione "a priori" di tutte le quantità sconosciute nel modello (parametri e dati mancanti). Il modello e la distribuzione vengono utilizzati per calcolare la distribuzione degli elementi sconosciuti utilizzando anche i dati osservati (la distribuzione *posteriore*).
Una prima formula di carattere generale, che deriva direttamente da una delle forme con cui può essere scritto il teorema di Bayes, e che per essere compresa non richiede particolari conoscenze specifiche, può essere ben introdotta come segue.

Bayesian statistical analysis depends fundamentally on a relationship derivable from the axioms of probability theory that enables us to convert the chance that an event *A* will happen, conditional on our having information about event *B* into the chance that event *B* will occur once we know that event *A* has already occurred. Then, using this symbolism, Bayes' formula asserts that

P{(event *B* will occur) *given* (event *A* has occurred)}

$$= P\{B \text{ given } A\} = \frac{P\{A \text{ given } B\} \times P\{B\}}{P\{A \text{ given } B\} \times P\{B\} + P\{A \text{ given not } B\} \times P\{\text{not } B\}}$$

In this formula, *P{A given B}* is generally referred to as the *model;* it depends on the observational data, as we will see. *P{B}* is generally referred to as the *prior probability* (since this probability depends only on information known prior to observing the data). *P{B given* A} is referred to as the *updated or posterior probability* of *B* since it is conditioned on having information about event *A*. (The prior probability of *B* was not so conditioned).(Press, Tanur, 2001, 206-207)

È forse utile ricordare che, nonostante il teorema di Bayes sia stato formulato in varie maniere diverse durante gli anni, una delle prime formulazioni, presentate da Laplace all'inizio delle sue memorie del 1774 (*Mémoire sur la probabilité des causes par les évènements. Mémoires de l'Académie royale des sciences présentés par divers savants*; vedi Stigler, 2000), è molto simile a quella contemporanea che ho appena presentato.

If an event can be produced by a number *n* of different causes, then the probabilities of these causes given the event are to each other as the probabilities of the event given the causes, and the probability of the existence of each of these is equal to the probability of the event given that cause, divided by the sum of all the probabilities of the event given cach of these causes. (Laplace, quoted in Stigler, 2000, 102)

Vale la pena di sottolineare ancora, anche per la particolare importanza pratica che ciò rappresenta, che negli anni recenti risultati davvero incoraggianti sono stati ottenuti nelle *decisioni in campo medico*(vedi Parmigiani, 2002).
Ciò che segue è un esempio semplice ma appropriato in questo campo, per mostrare un'applicazione dell'approccio generale (preso da D'Agostini, 2002 dove si trovano la presentazione e discussione dettagliate). Viene scelto un cittadino italiano da sottoporre al test dell'AIDS. Supponiamo che l'analisi usata per riscontrare l'infezione da HIV possa dichiarare "positive" persone sane con una probabilità dello 0,2%. Assumiamo che l'analisi dichiari la persona "positiva". Possiamo dire che, visto che la probabilità di un errore d'analisi è solo dello 0,2%, la probabilità che la persona sia veramente infettata è del 99,8%? Certamente no se si calcola sulla base di una stima di 1.000.000 di persone infette su una popolazione di 60 milioni (*informazioni conosciute prima di osservare i dati*). Infatti, una semplice applicazione della formula di Bayes appena presentata, dimostra che la probabilità che la persona sia sana è del 55%! Una maniera semplice per arrivare a tale risultato è d'immaginare di sottoporre al test l'intera popolazione . In tal caso, il numero di persone dichiarate positive sarà uguale a quello di tutte le persone infette più lo 0,2% della popolazione restante. Il totale sarebbe di 1.000.000 persone infette e 120.000 persone sane.
Ovviamente nei differenti campi di applicazione il formalismo matematico raggiunge livelli di una certa complicazione, sui quali non è certo il caso di soffermarci ora da un punto di vista strettamente tecnico.
Per quanto riguarda il campo più prettamente scientifico si può innanzi tutto ipotizzare un contributo allo studio dei sistemi complessi. Come forse è noto(lo ho accennato anche in questo lavoro al par.1.1) negli ultimi decenni dalla ricerca condotta nel campo della chimica, della fisica, della biologia, dell'informatica e, di recente, delle scienze sociali è emerso un nuovo paradigma scientifico: il paradigma della complessità(vedi Lane, 2002).
I fenomeni complessi sono caratterizzati da una serie di entità che interagiscono l'una con l'altra.

In seguito a queste interazioni, alcune proprietà di un'entità possono cambiare, compresa la sua posizione nella struttura reticolare e le modalità di interazione. Se le interazioni sono locali, gli oggetti di interesse nei fenomeni complessi sono però in genere globali, funzione di un modello di eventi interattivi che rimane relativamente stabile su una scala temporale molto più ampia rispetto a quella degli eventi stessi dell'interazione. Di frequente questi modelli meta-stabili di interazione si auto-organizzano, e le strutture e persino la funzionalità che ne risultano possono spesso essere descritte in un linguaggio che non contiene riferimenti alle entità sottostanti e alle loro interazioni. In quel caso i modelli vengono definiti *emergenti,* e lo studio dell'auto-organizzazione e dell'emergenza costituisce l'obiettivo primario della ricerca sulla complessità.(Lane, 2002, 14-15)

Questi modelli sono in parte simili ai modelli statistici. In particolare, entrambe le classi di modelli sono costituite da famiglie parametrizzate di modelli di probabilità, ed entrambe mirano a estrapolare deduzioni sui fenomeni del mondo reale.

[…]negli ultimi dieci anni, i teorici della complessità hanno dimostrato quanti interessanti problemi di biologia, chimica, fisica e scienze sociali presupponga proprio questo tipo di dinamica. Le teorie e i metodi della deduzione sarebbero di grande aiuto nell'affrontare tali problemi.[…]i metodi data-analitici e deduttivi per studiare i fenomeni emergenti sono difficili da individuare. Le tecniche statistiche disponibili non sono adatte allo scopo e la misurazione delle strutture emergenti e la correlazione di queste misure a descrizioni adeguate degli eventi di interazione a livello locale sottostanti rappresentano problemi di difficile soluzione. Attualmente quasi tutti i ricercatori della complessità si trovano ad agire d'istinto, ed è difficile intuire in quale direzione si muovano e riconoscere quando arrivano a destinazione. Credo ci sia la reale possibilità che giovani scienziati di impostazione statistica offrano contributi reali e importanti ai gruppi di ricerca della complessità[…].(ivi, 37)

Un campo, per esempio, strettamente correlato è quello della genetica statistica.

One of the most important challenges to data analysis in modern statistical genetics and molecular biology is the problem of how to analyze gene expression data in which the number of variables or dimensions, "p", greatly exceeds the number of replicates, "n" (that is, we have p much greater than n). Such data frequently are generated by a piece of equipment called a microarrayer. The data analysis problem is opposite to the usual problem in statistical data analysis in which typically p is much less than n. In the microarrayer context, we might have p = 10,000, for example, while n = 10. In such contexts, to reduce the problem to one where ordinary statistical methods can be applied, it is necessary either to increase the number of replicates (sample size) or to reduce the number of dimensions, or both. The Bayesian approach may offer the best solution to such problems since in Bayesian analysis, as is easy to show, bringing prior information to bear on a model is equivalent to adding more replicates, or increasing the effective sample size.(Press, Tanur, 2001, 212-213).

Inoltre, come ho più volte sottolineato precedentemente, si sta tentando di arrivare ad una nuova teoria della valutazione delle incertezze nel campo strettamente fisico. L’atteggiamento di fondo è a mio avviso ben riassumibile come segue.

As far as physics applications are concerned, the importance of the subjectivist approach stems from the fact that it is the only approach which allows us to speak in the most general way about the probability of hypotheses and true values, concepts which correspond to the natural reasoning of physicists. As a consequence, it is possible to build a consistent inferential framework in which the language remains that of probability. This framework is called Bayesan statistics, because of the crucial role of Bayes’ theorem in updating probabilities in the light of new experimental facts using the rules of logics. Subjective ingredients of the inference, unavoidable because researchers do not share the same status of information, are not hidden with the hope of obtaining objective inferences, but are optimally incorporated in the inferential framework. Hence, the prior dependence of the inference should not be seen as a weak point of the theory. On the contrary, it obliges practitioners to consider and state clearly the hypotheses which enter the inference and take personal responsibility for the result. In any case, prior information and evidence provided by the data are properly balanced by Bayes’ theorem, and the result is in qualitative agreement with what we would expect rationally. Priors dominate if the data is missing or of poor quality or if the hypotheses favored by the data alone is difficult to believe. They become uninfluential for routine high accuracy measurements, or when the evidence provided by the data in favor of a new hypotheses is so strong that physicists are obliged to remove deeply rooted ideas.(D’agostini, December 1999, 11-12)

Talvolta, voglio sottolineare, *è assolutamente legittimo credere più ai propri pregiudizi che ai dati empirici.* Quando ad esempio si usa uno strumento che sappiamo potrebbe non essere totalmente affidabile, o quando si fanno per la prima volta misure in un nuovo campo o in un nuovo range di possibili valori di una specifica grandezza e così via. Per esempio è ovvio che è più facile credere che uno studente abbia fatto un banale errore piuttosto che credere che abbia fatto una nuova scoperta! La seguente riflessione di Poincaré(del 1905) può essere esemplificativa al riguardo.

The impossibility of squaring the circle was shown in 1885, but before that date all geometers considered this impossibility as so 'probable' that the Académie des Sciences rejected without examination the, alas! too numerous memoirs on this subject that a few unhappy madmen sent in every year. Was the Académie wrong? Evidently not, and it knew perfectly well that by acting in this manner it did not run the least risk of stifling a discovery of moment. The Académie could not have proved that it was right, but it knew quite well that its instinct did not deceive it. If you had asked the Academicians, they would have answered: 'We have compared the probability that an unknown scientist should have found out what has been vainly sought for so long, with the probability that there is one madman the more on the earth, and the latter has appeared to us the greater' (Poincaré, 1952, 24)

Le conclusioni hanno, a mio avviso, il pregio di essere chiare e coraggiose.

In conclusion[…], I prefer to state explicitly the naturalness and necessity of subjective priors […]. If rational people (e.g. physicists), under the guidance of coherency (i.e. they are honest), but each with unavoidable personal experience, have priors which are so different that they reach divergent conclusions, it just means that the data are still not sufficiently solid to allow a high degree of intersubjectivity (i.e. the subject is still in the area of active research rather than in that of consolidated scientific culture). On the other hand, the step from abstract objective rules to dogmatism is very short […].(D'Agostini, Luglio 1999, 30)

Ovviamente tutte queste considerazioni possono essere discusse criticamente, e infatti sono state contrastate da tutti color che *credono* in un ideale "oggettivista".
In poche parole, tutti coloro che contestano il metodo Bayesiano si appellano al concetto filosofico di "razionalità" per confermare la necessita e la possibilità di "oggettività". Ciò è per esempio caratteristico della scuola Popperiana.

The cognitive value of a theory has nothing to do with its psychological influence on people's mind. Bilief, commitment, understanding are states of the human mind. But the objective, scientific values of a theory is indipendent of the human mind which creates it or understands it, its scientific value depends only on what objective support these conjecture have in facts.(Lakatos, 1978 , 1)

Popper afferma più di una volta che è impossibile che un'ipotesi possieda informazioni di alto contenuto e contemporaneamente di alta qualità.

They[scientists]have to choose between high probability and high informative content, *since for logical reasons they cannot have both*.(Popper, 1992, 363)

Si può rispondere a questa critica, a mio avviso, in maniera relativamente semplice nel seguente modo.

Such a charge is quite baseless. There is *nothing* in logic or the probability calculus which precludes the assignment of even probability 1 to any statement, however strong, as long as it is not a contradiction, of course. The only other way in which probabilities depend on logic is in their decreasing monotonically from entailed to entailing statements. But this again does not preclude anybody from assigning any consistent statement as large a probability as they wish. Popper's thesis that a necessary concomitant of logical strength is low probability is simply incorrect.(Howson, Urbach, 1993, 390)

Nel cercare difetti nell'approccio Bayesiano, alcuni avanzano altre obiezioni (una, per esempio, si basa sulla *necessità della probabilità soggettiva a priori*). Data la natura introduttiva di questo libro, non discuterò qui queste obiezioni. Per una buona analisi delle ragioni di chi contesta l'approccio Bayesiano vedi per esempio, Moore, 1997, Mayo, 1996 ed Efron, 1986. Per risposte dettagliate vedi Howson, Urbach, 1993, in particolare l'ultimo capitolo. Vorrei concludere queste specifiche considerazioni con una citazione da Feynman che mi ha fatto riflettere per prima su questi temi.

Some years ago I had a conversation with a layman about flying saucers- because I am scientific I know all about flying saucers! I said “I don’t think there are flying saucers”. So my antagonist said, “Is it impossible that there are flying saucers? Can you prove that it’s impossible? “ No”, I said, “I can’t prove it’s impossible. It’s just very unlikely”. At that he said, “You are very unscientific. If you can’t prove it impossible then how can you say that it’s unlikely?” *But this is the way that is scientific*. It is scientific only to say what is more likely and what less likely, and not to be proving all the time the possible and the impossible. To define what I mean, I might have said to him, “Listen, I mean that *from my knowledge* of the world that I see around me, I think that it is much more likely that the reports of flying saucers are the results of the known irrational characteristics of terrestrial intelligence than of the unknown rational efforts of extra-terrestrial intelligence”. *It is just more likely*. That is all.(Feynman, 1967, 47)

Il problema dell’analisi tecnica di dettaglio della statistica Bayesiana a seconda dei differenti campi di applicazione, come appena accennavo e come si può facilmente intuire, si fa piuttosto complesso per il formalismo matematico richiesto, e va ben al di là degli scopi di questo scritto. Chi fosse interessato comunque può vedere un’ottima introduzione in Press, Tanur, 2001, ed una esauriente esposizione degli attuali risultati in Fisica delle Alte Energie in D’Agostini, Luglio 1999, dove per esempio si può trovare descritta nel dettaglio una possibile soluzione ai problemi che ho discusso nel caso della propagazione degli errori casuali. Cionondimeno nel prossimo paragrafo cercherò almeno brevemente di introdurre come il concetto di probabilità inteso come grado di fiducia dell’occorrenza di un evento porta a ridefinire il problema della valutazione delle incertezze e permetta per esempio *di trattare allo stesso modo i diversi tipi di errori* precedentemente discussi.

4.5 Nuova impostazione nel determinare l’incertezza di misura

Our assent ought to be regulated by the grounds of probability
J. Locke

Si può riassumere in questo caso il nuovo atteggiamento dicendo che *il centro del problema si è spostato dall'analisi del valer vero, limite privilegiato a cui tende il processo di misurazione, all'analisi del complesso di informazioni disponibili* (su tutti gli aspetti descritti in questo paragrafo si può vedere l’ottima trattazione fatta in Mana,1994 e nella guida ISO).
Come visto l’impostazione convenzionale descrive il risultato di una misurazione mediante i concetti di valore misurato e di valore vero, considerando il secondo non conoscibile e tuttavia rappresentativo dello stato della grandezza ed il primo una sua valutazione errata. Viene quindi introdotto il concetto di errore quale differenza (incalcolabile) tra i valori misurato e vero. Il valore vero risulta così collocato in una posizione indeterminata ed indeterminabile entro l'intervallo, intorno al valore misurato, definito dall'errore. E’ chiaro che la definizione di valore vero che implica una serie indefinita di misure con strumenti ideali dà l’illusione che il valore vero sia unico. La nuova definizione, invece, prende esplicitamente in considerazione il fatto che le misure sono compiute in condizioni reali e possono essere accompagnate da tutte le componenti di incertezza (ognuna pesata con il suo “grado di fiducia”, come subito vedremo).

[L'impostazione nuova] abbandona questa via in favore di una definizione di misura che include esplicitamente l'incertezza. Con ciò si intende che nessuna misurazione può portare ad una valutazione dello stato di una grandezza più definita di quella affermante che ogni sua ulteriore misurazione produrrà valori compresi nell'intervallo (più precisamente avrà una certa probabilità di produrre) del quale il valore e l'incertezza sono gli elementi rappresentativi. Risulta pertanto preferibile definire la misura come l'insieme dei valori possibili. Va sottolineato che il numero assunto come valore non è più vero di ogni altro elemento dell'insieme, è soltanto il più conveniente.(Mana, 1994, 3)

Sottolineo anche di riflettere sulla l'impossibilità di definire operativamente il valore vero (e di conseguenza l'errore). Infatti ogni procedimento per la sua determinazione finisce per coincidere con quello con cui si determina il valore come qui definito.

Concorrono a formare l'incertezza le differenze tra le caratteristiche del sistema e dell'apparato per la misurazione ed i modelli scelti per descrivere i medesimi e la loro interazione, lo stato non perfettamente definito (o non perfettamente definibile) dei due sistemi al momento della misurazione e le loro interazioni con l'ambiente. Un ruolo determinante ha l'incertezza con cui è realizzata l'unità di misura locale (la classe di precisione dello strumento) che costituisce il limite inferiore per l'incertezza della misura.(Ibidem)

Come ampiamente già visto, seguendo l'impostazione (che ora discuterò criticamente) della teoria della misura basata sul valore vero si usa distinguere tra le cause di errore quelle di natura casuale da quelle di natura sistematica. Ricordo che la distinzione tra i due tipi di errore è la seguente: si intendono casuali quelli i cui effetti sul valore sono attenuati ripetendo la misurazione e mediando i valori ottenuti e sistematici quelli i cui effetti si riproducono invariati ad ogni ripetizione.
L'incertezza di una misura dovrebbe essere:

- consistente: derivabile dalle componenti che vi contribuiscono indipendente da come queste sono raggruppate o decomposte;
- trasferibile: utilizzabile per valutare l'incertezza di una di un'altra misura per ottenere la quale è usata la prima.

Inoltre è necessario fornire un intervallo intorno al valore che racchiuda una "significativa" frazione della distribuzione di valori attribuita alla grandezza oggetto di misurazione. Così la valutazione dell'incertezza deve fornire tale intervallo con un livello di confidenza corrispondente a quanto richiesto dalle specifiche applicazioni.

La valutazione dell'incertezza contrasta alcuni criteri diffusi nella pratica e basati su due presupposti infondati. Il primo è che l'incertezza debba essere "sicura" o "conservativa", significando che essa non dovrebbe essere errata per difetto: se la valutazione dell'incertezza è difficile, essa dovrebbe essere sovrastimata(ho già discusso criticamente questo presupposto nel par.9.1; voglio qui notare che la sovrastima dell'incertezza ha ulteriori effetti negativi: può incoraggiare per esempio l'acquisto di strumenti più costosi di quanto necessario o scoraggiare le calibrazioni). Il secondo, che ho appena ricordato, è che le cause di incertezza siano classificabili in casuali o sistematiche a seconda della loro natura e che i contributi associati a ciascuna causa debbano essere analizzati e riportati separatamente. Oltre a ciò, va considerato che spesso occorre incorporare nell'analisi del risultato altre misure ciascuna delle quali ha la sua incertezza. Per valutare l'incertezza del risultato *è necessaria la miglior stima delle incertezze delle misure incorporate, non un valore "sicuro"*. Inoltre è necessario un metodo semplice per combinare queste incertezze ed ottenere l'incertezza finale. *Per queste ragioni è opportuno considerare i contributi casuali e sistematici allo stesso modo*. La chiave è una nuova interpretazione dell'incertezza, oltre ovviamente l'uso delle regole della statistica per la "propagazione degli errori".

Per spiegare questo punto meglio, pensiamo innanzi tutto di raggruppare le componenti dell'incertezza secondo il modo in cui il loro valore numerico è stimato:

- A- quelle valutate secondo metodi statistici,
- B - quelle valutate diversamente,

senza con ciò stabilire una corrispondenza con la usuale classificazione in casuale e sistematico.
The ISO GUM, with regards to this point, says:

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.(ISO, 1993, VIII)

It is also important to point out that, on this subject, the ISO recommendation can be summarized with the following quotation from the GUM:

> This Guide presents a widely applicable method for evaluating and expressing uncertainty in measurement. It provides a realistic rather than a 'safe' value of uncertainty based on the concept that there is no inherent difference between an uncertainty component arising from a random effect and one arising from a correction for a systematic effect. The method 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.
>
> - The second idea is that the influences that give rise to uncertainty were always recognizable as either 'random' or 'systematic' with the two beíng of different nature[…] In fact, the method of combining uncertainty was often designed to satisfy the safety requirement. (ivi,43)

Nella valutazione dell'incertezza occorre elencare tutte le componenti e specificare quale metodo è stato usato nella stima del loro valore numerico. Le componenti della classe A sono caratterizzate dalla deviazione standard ottenuta con metodi statistici, quelle della classe B da valutazioni approssimate delle corrispondenti deviazioni standard, la cui esistenza viene postulata. *L'incertezza totale è ottenuta applicando le regole statistiche per la combinazione delle deviazioni standard ed è interpretata quale deviazione standard del risultato.*

I passi per la valutazione dell'incertezza possono essere riassunti come segue. La relazione tra il misurando y e le grandezze x_i da cui esso dipende è espressa da una relazione del tipo $y = f(x_i)$ che include tutte le correzioni che contribuiscono all'incertezza dei risultato. Le stime $\overline{x}_i$ delle grandezze di ingresso sono ottenute dall'analisi statistica di serie di osservazioni o con altri mezzi, ma per ciascuna è disponibile la deviazione standard (incertezza) σ_i.

Nel caso più semplice sono disponibili N osservazioni (estrazioni casuali) indipendenti (ipotesi essenziale per la validità delle formule seguenti) della grandezza di ingresso x. La miglior stima di tale grandezza, la media aritmetica

$$\overline{x} = \frac{\sum_{k=1}^{N} x_k}{N}$$

è ancora una variabile casuale con deviazione standard

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum_{k=1}^{N}(x_k - \bar{x})^2}{N(N-1)}}$$

Se la grandezza di ingresso non è ottenuta attraverso osservazioni ripetute (in generale attraverso l'analisi statistica di osservazioni ripetute) *la sua deviazione standard viene stimata sulla base delle informazioni disponibili*: misurazioni precedenti, conoscenza del comportamento del materiale o dello strumento, specificazioni del costruttore, certificati di taratura eccetera.
E' chiaro quindi che deve qui essere adottato il punto di vista che considera la probabilità una misura del grado di fiducia di un evento(concetto appunto di probabilità soggettiva). In tal senso una distribuzione di probabilità a priori è postulata anche per le componenti di tipo B consentendo, in linea di principio, di stabilire il concetto di deviazione standard. Questo *consente di incorporare l'incertezza di un risultato nell'incertezza di un altro risultato* per la valutazione del quale il primo viene utilizzato, di utilizzare l'incertezza per valutare l'intervallo corrispondente ad un prestabilito grado di confidenza e di evitare confusioni conseguenti ambigue classificazioni delle cause di errore.

E' importate sottolineare che, per le premesse stabilite, l'applicazione della relazione funzionale $\bar{y} = f(\bar{x}_i)$ per ottenere la stima $\bar{y}$ del risultato a partire dalle stime dei parametri produrrà ancora una variabile casuale la cui deviazione standard è calcolabile secondo le regole della statistica. Nel caso più semplice le non linearità di f possono essere trascurate(vedi Mana, 1994), pertanto

$$\sigma_y^2 = \sum_{i=1}^{n} c_i^2 \sigma_i^2$$

dove i coefficienti di sensibilità sono le derivate parziali $c_i = \partial f / \partial x_i$ valutate per $x_i = \bar{x}_i$.

Un esempio può chiarire quanto detto. Consideriamo la misura y funzione delle grandezze a, b e x secondo il modello

y = a + bx,

dove a è una polarizzazione comune a tutte le osservazioni, b un fattore di scala ed x una variabile casuale di cui sono disponibili N estrazioni indipendenti. Per concretezza possiamo pensare x la lettura di un termometro alla quale debba essere applicata la correzione appena espressa. La polarizzazione ed il fattore di scala sono caratterizzati da una distribuzione di probabilità a priori, essendo $\bar{a}$ e $\bar{b}$ le migliori stime dei valori medi e σ_a e σ_b le rispettive incertezze (deviazioni standard). La miglior stima di x è la media aritmetica $\bar{x}$ dei campioni disponibili la cui incertezza è la deviazione standard σ_x. Quindi il valore della misura cercata è $\bar{y} = \bar{a} + \bar{b}\bar{x}$. Per valutarne l'incertezza calcoliamo i cosiddetti coefficienti di sensibilità

$$\frac{\partial f}{\partial a} = 1, \frac{\partial f}{\partial b} = \bar{x}, \frac{\partial f}{\partial x} = \bar{b}$$

Quindi

$$\sigma_y = \sqrt{\sigma_a^2 + \bar{x}^2\sigma_b^2 + \bar{b}^2\sigma_x^2}$$

è l'espressione cercata. *Secondo la terminologia tradizionale* il terzo termine della espressione precedente è un errore casuale in quanto il suo contributo decresce con il numero di campioni , mentre i primi due termini sarebbero detti sistematici in quanto indipendenti dal numero di campioni. La teoria convenzionale però, come visto, non sarebbe capace di trattare formalmente questi tipi di contributi all'incertezza di misura.
Voglio finire questo paragrafo con una ulteriore citazione dalla GUM, che bene riassume l'obiettivo ultimo di una misura.

> […] When the value of a measurand is reported, the best estimate of its value and the best estimate of the uncertainty of that estimate must be given, for if the uncertainty is to err, it is not normally possible to decide in which direction it should err safe. An understatement of uncertainties might cause too much trust to be placed in the values reported, with sometimes embarrassing and even disastrous consequences. A deliberate overstatement of uncertainty could also have undesirable repercussions.(ISO, 1993,45)

4.6 Conclusioni

In almost all circumstances, and at all times,
we find ourselves in a state of uncertainty.
Uncertainty in every sense.
Uncertainty about actual situations, past and present ...
Uncertainty in foresight: this would not be eliminated or diminished even if we accepted, in its most absolute form,
the principle of determinism; in any case, this is no longer in fashion.
Uncertainty in the face of decisions: more than ever in this case...
Even in the field of tautology (i.e of what is true or false by mere
definition, independently of any contingent circumstances) we always
find ourselves in a state of uncertainty... (for instance,
of what is the seventh, or billionth, decimal place of π ...) ...
Bruno de Finetti

Cercherò qui di riassumere i temi più importanti emersi negli sviluppi teorici e relativi aspetti tecnici delle teorie degli errori di misura, iniziando da quella convenzionale.

- Anche in assenza di errori sistematici il valore di una misura è sempre caratterizzato da una più o meno grande *indeterminazione* rappresentata dall'errore casuale.
- Nel caso degli errori casuali questi si propagano da quelli delle misure dirette a quelli delle misure indirette, ed anche se probabilmente con le tecniche usuali questi ultimi sono talvolta sovrastimati, nella propagazione la tendenza è quella di aumentare gli errori.
- Per condurre correttamente una qualsiasi operazione di misura è necessario conoscere bene le caratteristiche dell'apparato usato, che *comunque influirà* sul risultato finale.
- Non si può mai essere certi dell'assenza degli errori sistematici , i quali , per loro natura , hanno effetto di togliere validità all'informazione ottenuta con la misura.

Nella scienza, forse più che negli altri comportamenti umani, *l'errore appare inevitabile*. Ho cercato di chiarirlo con un certo dettaglio nel caso della Fisica, e la conclusione può apparire curiosa in una disciplina che tradizionalmente viene ritenuta appartenere alle cosiddette "*scienze esatte"*. Per quanto possa apparire strano, solo in tempi recenti si è iniziato a discutere la portata ed

il significato della precedente osservazione con riferimento al mondo macroscopico: si pensi per esempio agli studi sui *processi caotici*. Al contrario andrebbe invece sottolineato(vedi Casati, 2001) che mentre *in meccanica classica l'errore si propaga in modo esponenziale, in meccanica quantistica si propaga solo linearmente*. In questo senso la meccanica quantistica sembra essere più predicibile di quella classica(questa proprietà della meccanica quantistica può avere profonde implicazioni teoriche ed essere alla base di grandi applicazioni tecnologiche del futuro, quali la ipotizzata possibilità di costruire calcolatori basati sulla logica quantistica).
Si deve riflettere in particolare sul fatto che se il futuro di un sistema è determinato in modo univoco dal suo stato presente, ciò non significa che noi siamo effettivamente in grado di determinarlo.

Può stupire qualcuno il fatto che non sappiamo nemmeno risolvere il problema di tre corpi in interazione tra loro. Può apparire quasi scandaloso che, a distanza di tre secoli dalla formulazione delle equazioni di Newton, non sappiamo ancora se il nostro sistema solare è stabile e se la Terra continuerà a girare attorno al Sole oppure andrà a cadervi sopra, o altro. Siamo - è vero - in grado di prevedere in modo approssimato l'orbita di un satellite e di seguirne la posizione istante per istante grazie ai moderni calcolatori elettronici, ma non siamo in grado di ottenere la soluzione esatta e quindi fare previsioni per tempi arbitrariamente lunghi.(Casati,2001,55)

Il grande fisico James Clerk Maxwell, nel 1873, con una intuizione davvero anticipatrice ne aveva colto l'essenza.

E' una dottrina metafisica che dagli *stessi* antecedenti seguono le *stesse* conseguenze. Nessuno può dubitare di questo. Ma ciò non è di molta utilità in un mondo come il nostro in cui gli stessi antecedenti non si verificano mai e nulla capita due volte... L'assioma fisico che ha un contenuto simile afferma che da antecedenti *simili* seguono conseguenze *simili.* Ma qui siamo passati dalla uguaglianza alla somiglianza, dalla precisione assoluta alla approssimazione più o meno buona. Esistono alcune classi di fenomeni nei quali un errore piccolo nei dati produce solo un errore piccolo nei risultati. Il corso degli eventi in questi casi è stabile. Esistono altre classi di fenomeni, più complessi, nei quali possono nascere casi di instabilità; il numero di tali casi aumenta in modo estremamente rapido con l'aumentare del numero delle variabili.(Maxwell, cit. in Casati, 2001,54)

Questi fatti consentono innanzi tutto di specificare in maniera del tutto particolare il concetto di scienze esatte, *ammesso che si voglia comunque utilizzare questo termine*:

una scienza può definirsi esatta non perché basata su informazioni esatte, ma perché la sua metodologia consente di conoscere statisticamente il valore dell'indeterminazione associata ad esse, ovvero di conoscere il limite del contenuto di informazione che esse portano.

Prendendo ampliamente atto della precedente conclusione, abbiamo poi visto che nuovi sviluppi sono in atto in questi anni per migliorare l'approccio che è alla base della teoria degli errori di misura. Riformuliamo il ragionamento presentato per quanto riguarda questo possibile nuovo approccio:

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 sulla grandezza fisica (prima delle nuove osservazioni).

La verosimiglianza descrive lo stato di conoscenza su
- strumentazione;
- condizione ambientali e fattori di influenza;
- contributo dello sperimentatore;
- etc. etc.

La conoscenza 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 in effetti tutti i fisici, per esempio, già sanno:

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.
Utilizzare questo approccio significa quindi sì utilizzare nuovo tecniche, ma anche un atteggiamento *che viene da lontano.*
Quanto detto credo sia davvero ben riassunto nella 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. (ISO, 1993,8)

Come sottolineato in precedenza, l'affermazione del filosofo greco Protagora *l'uomo è misura di tutte le cose* è ancora e forse sarà per sempre del tutto attuale. E ciò, ritengo, aiuta ancora meglio di quanto ho già indicato nella premessa di questo scritto a chiarire l'importanza di una successiva analisi: quella che si impegna a discutere le reciproche influenze esistenti e/o possibili tra i modelli epistemologici ed i modelli etici.

BIBLIOGRAFIA

Agnoli P., Bosco F., Canavesio F., Spinelli M., Vecchione M., *Studio comparativo di apparati per servizi Audiotex, Rapporto tecnico 93170*, Torino, Centro Studi e Laboratori Telecomunicazioni, 1993

Agnoli P., Di Nicola P., *Why Did the Chicken Cross the Road?*, Roma, Armando Editore, 2001

Arecchi F. T., *Introduzione*, in AA.VV., *Determinismo e complessità*, a cura di F. T. Arecchi, Roma, Armando Editore, 2000, pp. 9-11

Baldi A., *Calcolo delle Probabilità e Statistica*, McGraw-Hill Libri Italia, 1998

Bernardini G., *Fisica Generale Parte I*, Roma, Veschi, 1974

Bernardo J. M., Smith A. F. M., *Bayesan Theory*, New York, John Wiley & Sons INC, 2000

Bevington P.R., Robinson D.K., *Data reduction and error analysis for the physical sciences*, London, Mc Graw Hill Int. Ed.,1994

Birnbaum A., *Theory of errors*, in AA.VV, *Encyclopedia Americana*, Danbury, Grolier, 1983 I, vol.10, p. 560

Birnbaum A., *Least squares*, in AA.VV, *Encyclopedia Americana*, Danbury, Grolier, 1983 vol. 17, pp123-124

Box G. E. P., *Bayesian Inference in Statistical Analysis*, New York, John Wiley & Sons INC, 1992

Bradley P., Thomas A. L., *Bayes and Empirical Bayes Methods for Data Analysis*, Boca Raton, CRC Press, 2000

Campbell N. R., *Foundations of Science*, Dover, New York, 1957

Caporaloni M., Caporaloni S., Ambrosiani R., *La misura e la valutazione della sua incertezza nella fisica sperimentale*, Bologna, Zanichelli, 1987.

Casati G., *L'errore: caos o determinismo?*, In: Kos n. 187, Milano, aprile 2001, pp.54-59

Chalmers B., *Che cos'è questa scienza?*, Milano, Est Mondatori, 1987

Continenza B., Gagliasso E., *Giochi aperti in Biologia*, Milano, Franco Angeli, 1996

Costantini D., Geymonat L., *Filosofia della probabilità*, Milano, Feltrinelli, 1982

Cromer A., *L'eresia della scienza*, Milano, Cortina, 1996

D'Agostini G., *Bayesian Reasoning in High-Energy Physics: Principles and Applications*, Geneve, CERN, July 1999

D'Agostini G., *Errori e incertezze di misura – rassegna critica e proposte per l'insegnamento -*, Roma, Università La Sapienza Dipartimento di Fisica, gennaio 1999

D'Agostini G., *Probabilità e Incertezza di misura (parte I)*, Roma, Università La Sapienza Dipartimento di Fisica, ottobre 2001

D'Agostini G., *Teaching Statistics in the Physics Curriculum: Unifying and Clarifying Role of Subjective Probability*, in *American Journal of Physics Special Theme Issue on Thermodynamics, Statistical Mechanics, and Statistical Physics*, Gould and Tobochnik eds, December 1999

D'Agostini G., *Bayesan Reasoning in Physics. A Critical Introduction*, Singapore, Worl Scientific, 2002

De Finetti B., *Filosofia della probabilità*, Milano, Il Saggiatore, 1995

De Finetti B., *La logica dell'incerto*, Milano, Il Saggiatore, 1989

De Finetti B., *Teoria delle probabilità*, Milano, Il Saggiatore, 1970

Dyson G. B., *L'evoluzione delle macchine*, Milano, Raffaello Cortina Editore, 2000

Efron B., *"Why isn't everyone a Bayesian?"*, Am. Stat. 40,1986

Einstein A., *Relatività: esposizione divulgativa*, Torino, Boringhieri, 1967

Erto P., *Probabilità e statistica*, Milano, McGraw – Hill Libri Italia, 1999

Feynman R., *The character of the physical law*, Cambridge MA, The MIT Press, 1967

Fornasini P., *Dispense di esperimentazione fisica*, Trento, Università Dipartimento di Fisica, 2001

Galimberti V., *Psiche e techne. L'uomo nell'età della tecnica*, Milano, Feltrinelli, 1999

Galison, P.L., *How experiments end*, Chicago, Chicago University Press, 1987

Giolanella G., *Analisi dei dati e complementi di meccanica e termodinamica*, Torino, UTET, 2001

Grasshoff G., *The History of Ptolemy's Star Catalogue (Studies in the History of Mathematics and Physical Sciences, 14)*, New York, Springer Verlag, 1990

Hacking I., *The Emergence of Probability*, Cambridge, Cambridge University Press, 1975

Hacking I., *L'autogiustificazione delle scienze di laboratorio*, in AA.VV., *La scienza come pratica e cultura,* a cura di Pickering A., Torino, Edizioni di comunità, 2001, pp 33-75

Howson C., Urbach P., *Scientific Reasoning. The Bayesian Approach*, Chicago, Open Court, 1993

Hume D., *Ricerche sull'intelletto umano e sui principi della morale*, a cura di Dal Pra M., Bari, 1957

International Organization for Standardization, *Guide to the expression of uncertainty in measurement*, Genève, 1993

Kempthorne O, Folks L., *Probability, statistics and data analysis*, London, Ames,1971

Koyré A, *Apport scientifique de la Renaissance*, in *Revue de Synthese*, XLVII , 1950

Kuhn T.S., *La struttura delle rivoluzioni scientifiche*, Torino, Einaudi, 1969

Kuhn T.S., *The Function of Measurement in Modern Physical Science*, Isis v52, part z, No 168, June 1961, pp.161-193

Kula W., *Le misure e gli uomini dall'antichità ad oggi*, Bari, Laterza, 1987

Lad F., *Operational Subjective Statistical Methods. A Mathematical, Philosophical, and Historical Introduction*, New York, John Wiley & Son INC., 1996

Lakatos I, *Philosophical Papers*, vol. 1, Cambridge, Cambridge University Press, 1978

Lane D. A., *Complessità: modelli e inferenza*, in AA.VV., *Complessità e biologia*, a cura di Biava P.M., Milano, Bruno Mondatori, 2002, pp. 13-41

Mana G., *Metrologia:l'arte della misurazione-dispense del corso di Scienza dei materiali*, Torino, Università di Torino, 1994

Mandel J., *The Statistical Analysis of Experimental Data*, Dover, New York, 1964

Markowits W., *Measurement and Determination of Time*, in AA.VV, *Encyclopedia Americana*, Danbury, Grolier, 1983, vol. 26, pp 752-758

Mayo D. G., *Error and the Growth of Experimental Knowledge. Science and its Conceptual Foundation*, Chicago, Chicago University Press, 1996

Mayr E., *Storia del pensiero biologico*, Torino, Bollati Boringhieri, 1999

Moore D. S., *"Bayes for beginners? Some reasons to hesitate",* Am. Stat. **51,** 1997

Orear J., *Enrico Fermi, The Man*, in Nuovo Saggiatore vol.17 anno 2001 no 5-6, pp. 30-38

Parmigiani G., *Modeling in Medical Decision Making. A Bayesan Approach*, New York, John Wiley & Sons INC, 2002

Piattelli Palmarini M., *L'illusione di sapere – che cosa si nasconde dietro i nostri errori*, Milano, Mondatori, 1995

Poincaré H., *Science and Hypothesys*, London, Dover Publications, 1952

Popper K., *La logica della scoperta scientifica*, Torino,Einaudi, 1990

Press S. J., Tanur J.M., *The Subjectivity of Scientists and the Bayesan Approach*, New York, John Wiley & Sons, INC, 2001

Roche J. J., *The Mathematics of Measurement. A Critical History*, The Athlone Press, London, 1998

Ross S., *Introduction to Probability Models*, London, Academic Press, 2000

Russo L., *La rivoluzione dimenticata. Il pensiero scientifico greco e la scienza moderna*, Milano, Feltrinelli, 2001

Scozzafava R., *Incertezza e Probabilità*, Bologna, Zanichelli, 2001

Severi M., *Introduzione all'esperimentazione fisica*, Bologna, Zanichelli, 1986

Stanford J.L., Vardeman S.B. (a cura di), *Statistical Methods for Physical Science* - vol. 28 della collana *Methods of Experimental Physics* , London, Academic Press,1994

Stigler S. M., *The History of Statistics. The Measurement of Uncertainty before 1900*, Cambridge, The Belknap Press of Harvard University, 2000

Taylor J.R., *Introduzione all'analisi degli errori*, Bologna, Zanichelli,1986

Tenner E.C., *Why Things Bite Back: Technology and the Revenge of Unintended Consequences*, Knopf, New York, 1996

Toulmin S., *The Philosophy of Science, An Introduction*, Harper, New York, 1960

Vineis P., *Nel crepuscolo della probabilità*, Torino, Einaudi, 1999

Young H. D., *Elaborazione statistica dei dati sperimentali*, Roma, Veschi, 1970