Università degli Studi di Napoli "Parthenope"
Dipartimento di Scienze e Tecnologie

Corso di Topografia e Idrografia

Lezione 5

Misure ed errori

Claudio Parente

Teoria degli errori

In tutte le scienze sperimentali riveste grande importanza lo studio degli ERRORI DI OSSERVAZIONE quando:

si devono effettuare misure

In tutte le scienze sperimentali riveste grande importanza lo studio degli ERRORI DI OSSERVAZIONE quando:

si vogliono stabilire metodi per conseguire una certa approssimazione

Teoria degli errori

In tutte le scienze sperimentali riveste grande importanza lo studio degli ERRORI DI OSSERVAZIONE quando:

si vuole valutare l'entità degli errori commessi

In tutte le scienze sperimentali riveste grande importanza lo studio degli ERRORI DI OSSERVAZIONE quando:

si devono determinare VALORI NUMERICI da associare

In tutte le scienze sperimentali riveste grande importanza lo studio degli ERRORI DI OSSERVAZIONE quando:

a GRANDEZZE MISURATE

o a grandezze

collegate a quelle misurate

da relazioni analitiche

In Topografia il problema del trattamento delle misure è importante:

spesso si devono conseguire precisioni spinte che sono conseguenza dell'utilizzo di particolari strumenti e di opportune metodologie di misure

La *TEORIA DEGLI ERRORI* però si non applica solo alle *misure di precisione*,

La *TEORIA DEGLI ERRORI* si applica a tutte le *misure* quando *OCCORRE*

scegliere procedimenti di misura e di calcolo

scegliere gli strumenti

La *TEORIA DEGLI ERRORI* si applica a tutte le *misure* quando

Si vuole OTTENERE UN RISULTATO

Si vuole SAPERE se QUEL RISULTATO È STATO RAGGIUNTO

Si vuole sapere quanto ci si può FIDARE DEL RISULTATO

definizione di grandezza secondo Bertrand Russel

Esiste una certa coppia di relazioni indefinibili, *maggiore o minore*

queste relazioni sono simmetriche e transitive e sono incompatibili l'una con l'altra

$>$

$<$

definizione di <u>grandezza</u> secondo Bertrand Russel

Ognuna è inversa dell’altra nel senso che ogni volta che una è valida tra A e B e l’altra è valida tra B ed A.

$>$

$<$

I termini che risultano suscettibili di queste relazioni sono *grandezze.*

$>$

$<$

Una grandezza è un concetto che si può materializzare in forma concreta e in quantità diversa

$>$

$<$

Il concetto di **PESO**
è una **GRANDEZZA**
infatti si
concretizza in
quantità di diverse
negli oggetti fisici
che ci circondano

$>$

$<$

Possiamo esprimere un giudizio

<

maggiore

minore

>

tra due oggetti quando hanno questa grandezza in quantità diversa

Noi misuriamo *GRANDEZZE*

Cos'è una *MISURA* ?

definizione di misura secondo Bertrand Russel

"Si dice **MISURA** di una grandezza

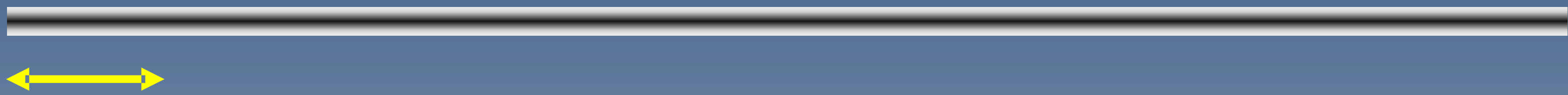
qualsiasi **metodo** con cui si stabilisca una corrispondenza univoca e **reciproca**

tra una **grandezza** di un determinato genere

e un **numero intero**."

Lo schema logico di esecuzione di una misura diretta è il seguente:

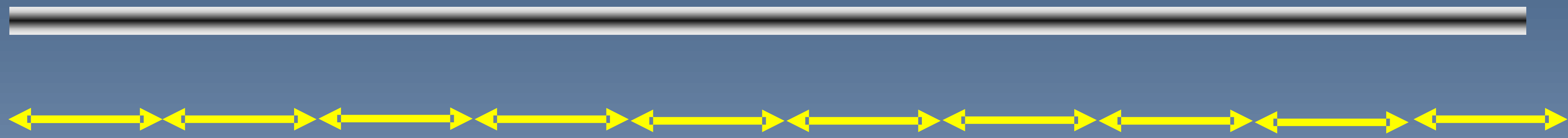
lunghezza di una barra metallica L



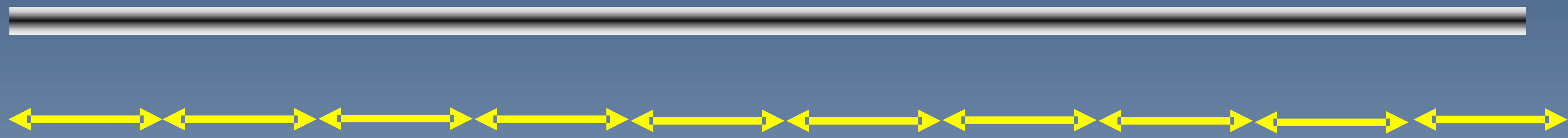
U

Stabilisco un'unità di misura U

lunghezza di una barra metallica L



Sommo tante unità U fino a formare la quantità di grandezza G che ritengo = L lunghezza di una barra metallica **L**



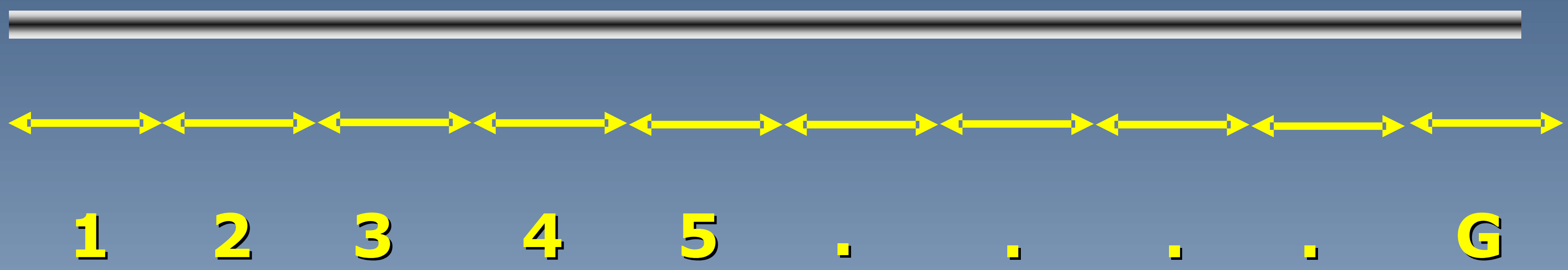
1 2 3 4 5 . . . . G

si contano quante unità di misura **U**

si sono sommate per formare **G**,

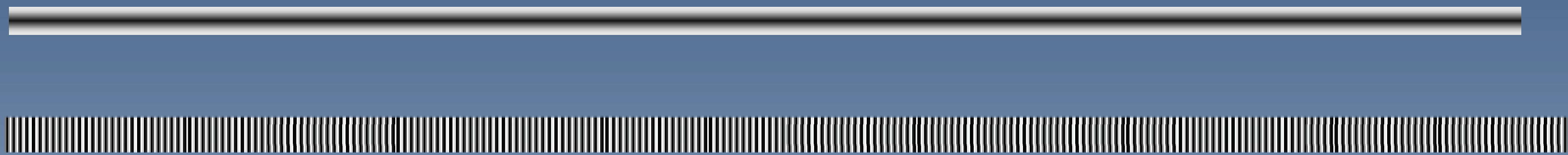
cioè si misura la numerosità di G

lunghezza di una barra metallica L



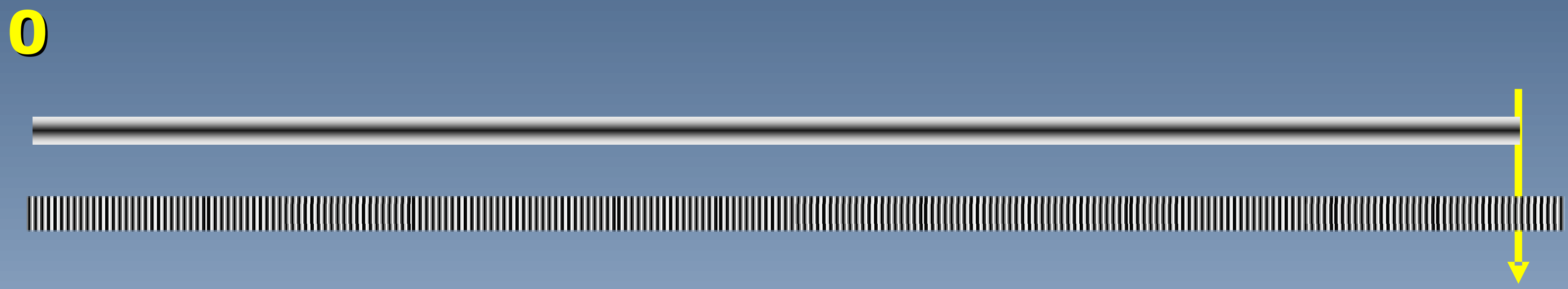
Non ci accorgiamo di applicare questa procedura.

Ma la misura diretta di una quantità di grandezza avviene PROPRIO COSÌ



Misuriamo la lunghezza della sbarra L con *una riga millimetrata* di 1 m

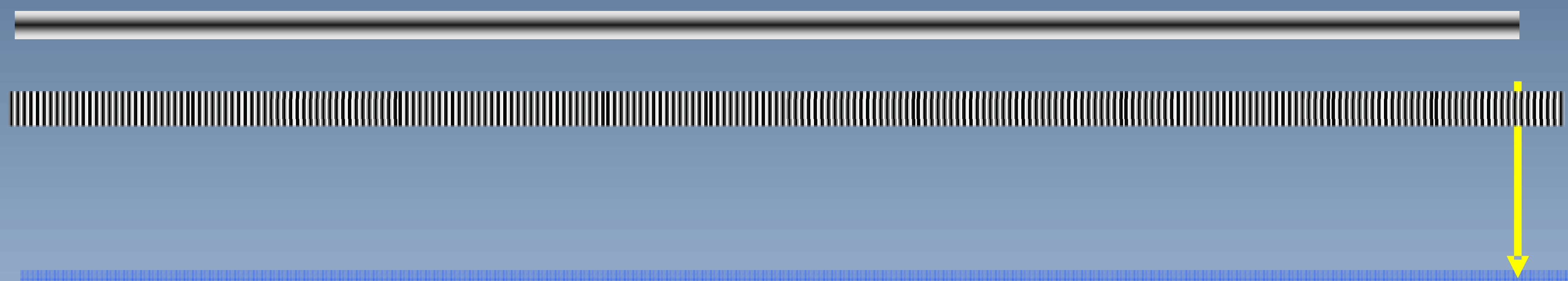
Per effettuare una misura accostiamo lo zero ad un’estremità dell’oggetto



guardiamo il punto in cui cade sulla riga l’altra estremità dell’oggetto

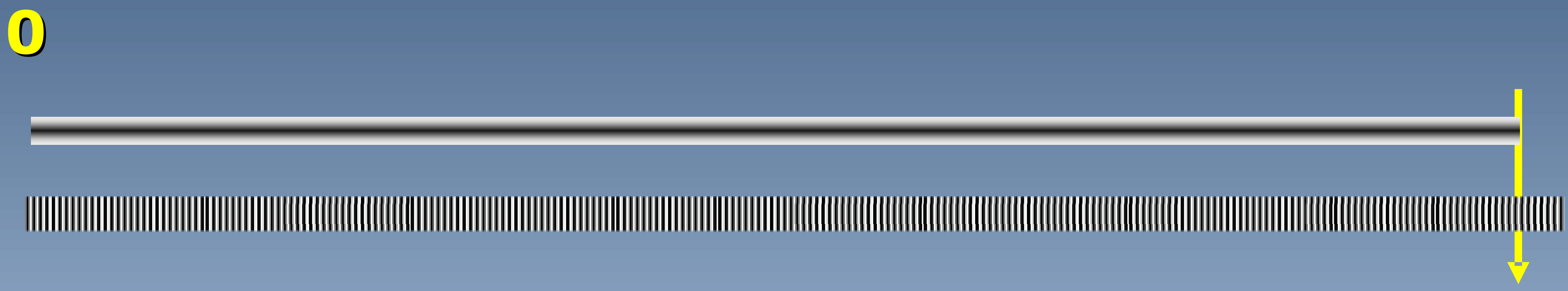
Creiamo così una quantità di GRANDEZZA DI PARAGONE

0



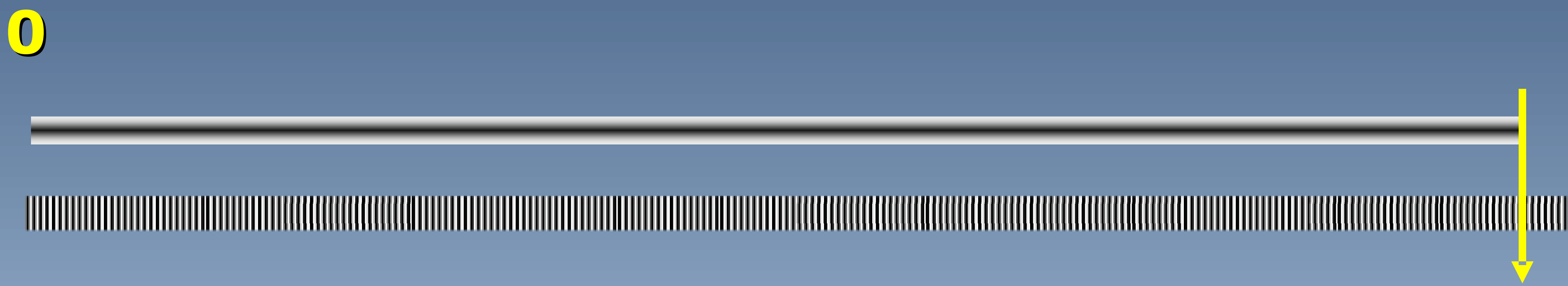
e stimiamo la sua lunghezza uguale a quella dell'oggetto

LEGGIAMO IL VALORE DELLA MISURA SULLA RIGA:



è come se avessimo contato tutti mm di cui è composta la GRANDEZZA DI PARAGONE

Assumiamo come **MISURA** dell’og-getto **il VALORE LETTO** sulla riga



Esso corrisponde
alla **NUMEROSITÀ della**
GRANDEZZA DI PARAGONE
che abbiamo creato sulla riga

La sola differenza fra il modello concettuale descritto e l'esempio fatto

0 10 20 30 40 50 60 70 80 90 100

è che gli strumenti di misura offrono già la somma delle unità di cui sono composti

La massima quantità misurabile con uno strumento è un parametro *molto importante* perché concorre a definire la PRECISIONE

Le misure delle grandezze di tipo quantitativo effettuate secondo lo schema descritto sono

0 10 20 30 40 50 60 70 80 90 100

MISURE DIRETTE

Origine della dispersione delle misure

Sensibilità e precisione degli strumenti di misura

Si definisce

sensibilità

di uno strumento

la più piccola quantità di grandezza misurabile univocamente con esso

sensibilità **di uno strumento**

per un righello millimetrato
la sensibilità è 1 mm

per una bilancia con scala graduata in grammi
la sensibilità è 1 grammo

Si definisce *precisione* di uno strumento il rapporto tra la *sensibilità* dello strumento e la *massima quantità* di grandezza che lo strumento può misurare.

$$\frac{\textit{sensibilità}}{\textit{massima quantità misurabile}}$$

una riga di 1 metro con suddivisione in millimetri

ha una precisione di

$$\frac{1mm}{1000mm} = 10^{-3}$$

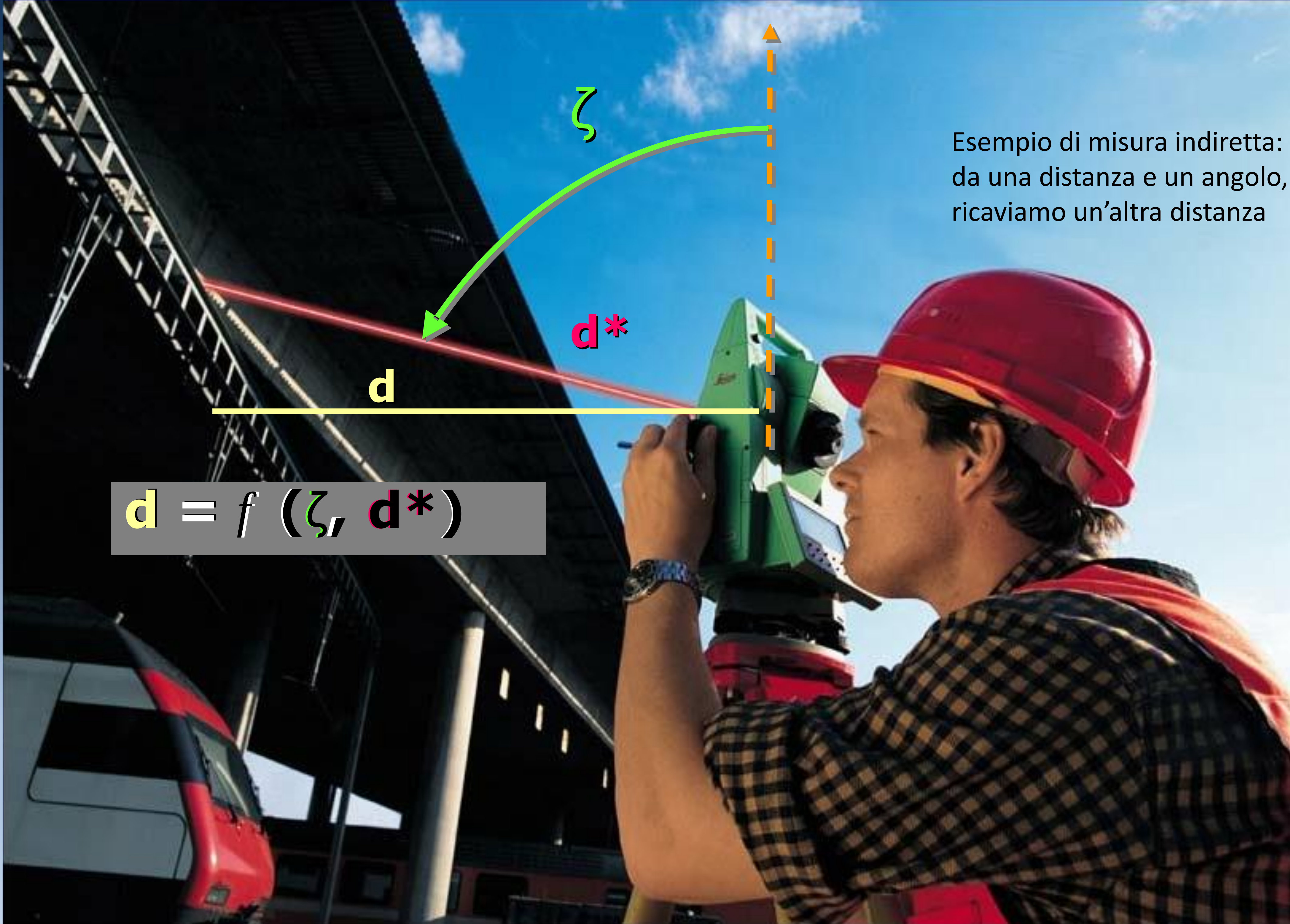
una bilancia che può pesare una massa di entità massima di 10 kg e avente una graduazione in grammi

ha un precisione di

$$\frac{1g}{10000g} = 10^{-4}$$

La precisione è quindi
ADIMENSIONALE.

Per il fatto di essere ADIMENSIONALE,
la precisione ci permette di confrontare
l'accuratezza di misure di diverso tipo
che intervengono nella determinazione
di una grandezza misurata
indirettamente



Esempio di misura indiretta:
da una distanza e un angolo,
ricaviamo un'altra distanza

Ma perché
non c'è univocità sui valori ottenuti
nel ripetere la misura
di una stessa quantità di grandezza?

Perché generalmente noi usiamo gli strumenti
pretendendo di aumentare:

***la sensibilità** con operazioni di stima

***la precisione** con operazioni ripetitive.

Ma perché

non c'è univocità sui valori ottenuti
nel ripetere la misura

di una stessa quantità di grandezza?

Influenza dell'ambiente

la dispersione dei valori numerici che si ha ripetendo le misure di una stessa quantità di grandezza,

è anche determinata dall'influenza dell'ambiente nell'operazione di misura

La misura infatti avviene in un **ambiente**

caratterizzato da parametri (temperatura, umidità, pressione atmosferica ecc.)

che non hanno un valore costante,

ma **oscillano in un certo campo**

Il numero X che rappresenta
la misura
eseguita in un certo istante

può essere considerato
un particolare valore
di una funzione f che dipende:

influenza dell'ambiente

dalla quantità di grandezza G che si misura

dall'unità di misura U che si adotta

da parametri u, v, w,...t che caratterizzano l'ambiente

$$X = f\,(G/U,\ u,\ v,\ w, \ldots t)$$

Se durante un intervallo di tempo ripetiamo la misura,

i valori dei parametri ambientali varieranno da misura a misura, e quindi si avranno diversi valori di X

Nell'effettuare un'operazione di misura, ad un certo istante i, non dobbiamo pensare ad un solo risultato possibile,

$$X_i = f(G/U, u_i, v_i, w_i, \ldots t_i)$$

ma ad una molteplicità di risultati possibili

non vi è univocità nel valore della misura

E tra tutti i possibili valori che si potrebbero registrare,

quale potremmo assumere come *misura vera* della quantità di grandezza G?

Si conviene di assumere come *misura vera* della quantità di grandezza G,

$$\overline{X} = f\,(G/U,\ u_m,\ v_m,\ w_m, \dots\ t_m)$$

e cioè il valore di $\bar{X}$ che si avrebbe se effettuassimo la **misura**
quando tutti i **parametri ambientali assumono il loro valore medio**

$$\bar{X} = f\,(G/U,\ u_m,\ v_m,\ w_m,\ldots t_m)$$

Si definisce inoltre come **errore**

ε_i

di una generica **misura**

$$\varepsilon_i = x_i - x_m$$

la differenza tra il **valore** x_i **che** corrisponde a quella misura

e **il valore** x_m che si avrebbe effettuando la misura in una situazione in cui i **parameri ambientali** assumessero il **loro valore medio**

$$\varepsilon_i = X_i - X_m$$

$$\varepsilon_i = f\,(G/U\ ,\ u,\ v,\ w, ..t) - f\,(G/U,\ u_m,\ v_m,\ w_m, ..t_m)$$

Misure dirette ed indirette

Le misure vengono distinte in *dirette* e *indirette*, a seconda che sia possibile o meno misurare direttamente ciò che vogliamo conoscere.

- Misure dirette: è possibile misurare direttamente la grandezza X, che si desidera conoscere, contando il numero di unità campione di cui è costituita (bilancia a piatti, misura col metro di un oggetto).

- Misure indirette: sono definite da un legame funzionale a misure dirette; ad esempio la misura indiretta della superficie del rettangolo nota la lunghezza dei due lati misurati direttamente.

L'operazione di misura, diretta o meno, ha in comune il fatto può essere considerata un'estrazione da una variabile casuale.

Esempi di misure

a) Dato un corpo rigido di lunghezza poco maggiore di 3 m ed un metro campione suddiviso in mm, si desidera misurare il corpo con il metodo del riporto.

b) La misura dell'area di una superficie.

c) Si misurano le coordinate x, y del punto ove cade un proiettile su un bersaglio sparato da uno stesso tiratore.

Errori di misura

Questi esperimenti hanno in comune il fatto che, a priori, è impossibile predire il risultato dell'esperimento: se si ripete infatti, si otterranno diversi risultati.

Nell'esempio a. il fatto che ripetendo l'operazione di misura si ottengano diversi risultati, porta a dire che in questa operazione si commettono degli «errori».
Che tipo di errori? Ne esistono diversi?

Gli errori possono essere classificati in:

1. ***errori grossolani***
2. ***errori sistematici***
3. ***errori accidentali***

Errori grossolani

Sono i più banali anche se spesso i più difficili a individuare. Sono causati da fattori esterni alla misura vera e propria.

Possono essere ad esempio il mancato conteggio di una alzata, la trascrizione errata di una misura, la codifica errata di un punto, ecc.

I rimedi per evitarli sono l'acquisizione e il trattamento automatici, il controllo e la ripetizione delle misure possibilmente indipendenti.
Tali errori non devono essere trattati con metodi statistici ma individuati ed eliminati.

Non sono questi gli «errori» a cui intendiamo riferirci nell'esempio a.

Errori sistematici

Sono dovuti ad esempio all'imperfetta taratura dello strumento di misura o legati ad errori di modello

Gli errori sistematici sono difficilmente individuabili mediante ripetizione in quanto presenti costantemente ad ogni nuova misura: hanno la caratteristica di conservare valore e segno.

Sono eliminabili con tarature o con opportune procedure operative.

Non sono questi gli «errori» a cui intendiamo riferirci nell'esempio a.

Errori accidentali

Sono a priori imprevedibili, sono di segno alterno e dipendono in senso lato dall'ambiente.

Sono questi gli «errori» commessi negli esperimenti descritti.

La scienza che studia questi fenomeni è la statistica matematica, perciò ne forniremo i concetti di base utili al trattamento delle misure geodetiche e topografiche.

Torniamo all'esempio

Nell'esempio del metro, notiamo che, se avessimo preteso di stimare la lunghezza del corpo al mm, avremmo ottenuto numeri apparentemente più variabili, mentre, chiedendo la misura al cm, il risultato sarebbe stato sempre uguale.

Ne segue che, per la misura di una grandezza, l'indeterminazione si presenta solo con procedure di misura che spingono l'approssimazione ai confini delle capacità di misura dell'apparato usato.

Che cos'è una misura di precisione?

Con il termine ***misura di precisione*** si intende una misura in cui si spinge lo strumento usato al limite della propria sensibilità.

Con ***sensibilità*** si intende la più piccola grandezza necessaria a causare uno spostamento apprezzabile della scala dello strumento.

Ad esempio, in una bilancia, l'applicazione di un peso inferiore alla sensibilità minima non comporta il movimento della lancetta dallo zero.

Fenomeni deterministici ed aleatori

Vi sono fenomeni deterministici, il cui comportamento è prevedibile:

- il risultato di un calcolo matematico;
- i fenomeni fisici in generale, come ad esempio l'elongazione di una molla.

Vi sono fenomeni non deterministici. Essi sono detti aleatori, casuali o stocastici.

Esempi di fenomeni aleatori sono, fra gli altri:

- la durata di una lampadina ad incandescenza
- il lancio di una moneta o di un dado
- la misura di precisione di una grandezza fisica, distanza, lunghezza, peso, angolo, temperatura o velocità

Fenomeni aleatori - 1

Lampadina: nessuno può dire quante ore durerà

Dado: nessuno può prevedere quale faccia uscirà dopo il lancio

Misura di precisione: misure ripetute forniscono risultati differenti

I fenomeni aleatori hanno tuttavia una regolarità di fondo che si manifesta ripetendo l'esperimento, che consente di studiarli con metodi scientifici.

Che cosa significa ripetere l'esperimento? Dipende dal contesto.

Lampadina: verificare la durata di diverse lampadine

Dado: ripetere il lancio

Misura di precisione: effettuare misure ripetute

Fenomeni aleatori - 2

Come si manifesta la regolarità di fondo.

- *Lampadina*: diversi esemplari dello stesso tipo di lampadina (aventi le stesse caratteristiche tecniche e costruttive) hanno durate diverse, ma tuttavia concentrate in un intervallo ragionevolmente stretto. In altri termini: non si verifica che una lampadina duri 5 anni e un'altra 5 giorni, a meno di difetti di fabbricazione.
- *Dado*: effettuando molti lanci, le singole facce si presentano circa $1/6$ delle volte. In altri termini non si verifica che, effettuando 1000 lanci, esca sempre la faccia 1 (più precisamente, un tale evento ha una bassissima probabilità di accadere).
- *Misura di precisione*: misurando ripetutamente la distanza fra due tratti disegnati su un foglio con una riga millimetrata, si otterranno risultati diversi, ma concentrati in un intervallo abbastanza ristretto.

Teoria della Probabilità e Statistica

I fenomeni aleatori sono descritti dalla Teoria della Probabilità e dalla Statistica.

Probabilità: permette di prevedere il comportamento di un fenomeno aleatorio noto, come ad esempio calcolare la probabilità che lanciando tre monete simmetriche (modello noto) si ottengano tre teste.

Statistica: procede nel verso contrario e cerca di conoscere un fenomeno aleatorio a partire dal suo comportamento:

- studio la vita di 100 lampadine identiche per stabilirne la vita media;
- effettuo misure angolari ripetute per stabilire la precisione con cui lo strumento e l'operatore lavorano.

Misure di precisione in Geomatica

Le misure di precisione sono un fenomeno aleatorio con cui i rilevatori si confrontano continuamente: le misure di angoli e distanze fatte da un teodolite sono, ad esempio, misure di precisione.

Per poter svolgere la propria attività in modo appropriato, il topografo deve saper trattare le misure in modo adeguato e, a volte, piuttosto sofisticato. Questo spiega perché tutti i testi riguardanti argomenti di rilevamento abbiano una significativa sezione dedicata al calcolo delle probabilità e alla statistica.

Topografia e Trattamento delle Osservazioni

Un topografo chiede alla statistica di essere guidato:

- per effettuare le migliori misure possibili con una certa strumentazione
- per conoscere comunque la bontà e la affidabilità delle misure fatte
- per capire come fare a raggiungere una certa precisione
- per confrontare efficacemente e razionalmente misure prese in momenti diversi e stabilire se la loro differenza sia attribuibile ai soli errori di misura o abbia altra causa

Questi problemi appartengono alla parte della Statistica chiamata ***Trattamento delle Osservazioni***.

Un esempio

Si vuole controllare periodicamente i movimenti di una frana misurando la distanza fra due punti, uno materializzato in una zona stabile ed uno in frana.

Misure effettuate in giorni diversi inevitabilmente differiscono per gli errori accidentali di misura.

Come capire se la differenza fra misure prese a una settimana di distanza abbia una natura puramente accidentale o indichi invece un movimento reale del terreno?

Applicando tecniche di analisi statistica.

Precisione ed accuratezza - 1

La qualità di una misura si esprime generalmente mediante due termini: ***precisione*** e ***accuratezza***:

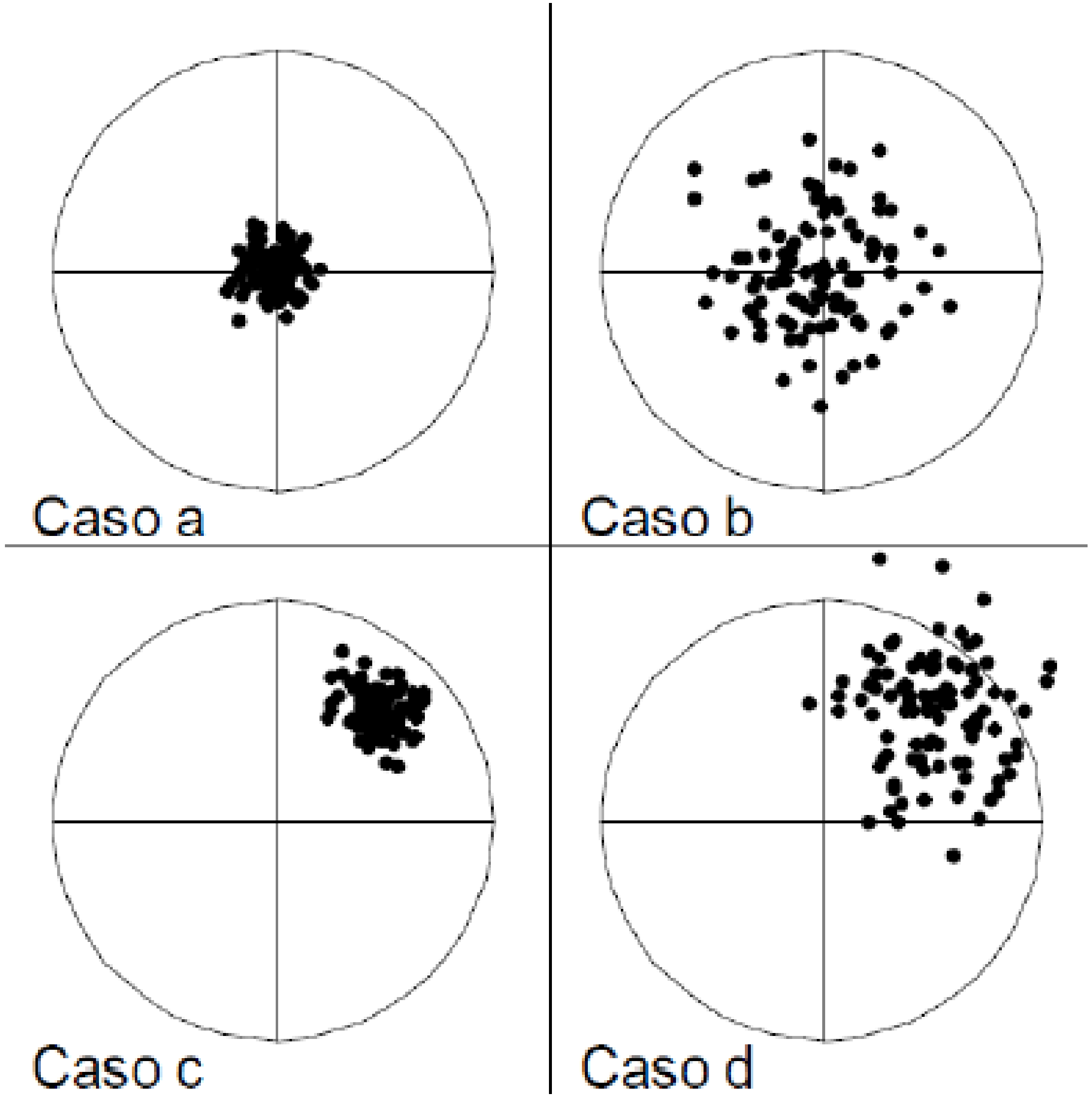
1. la precisione descrive la concentrazione di misure ripetute;
2. l'accuratezza esprime la distanza fra le misure ed il valore vero.

Una misura ottimale deve essere precisa ed accurata, ma capita a volte di effettuare misure precise ma non accurate, oppure accurate ma non precise, oppure anche, nel caso peggiore, non accurate e non precise.

Per comprendere meglio è utile utilizzare una analogia fra le misure e gli spari ripetuti di un tiratore a un bersaglio.

Precisione ed accuratezza - 2

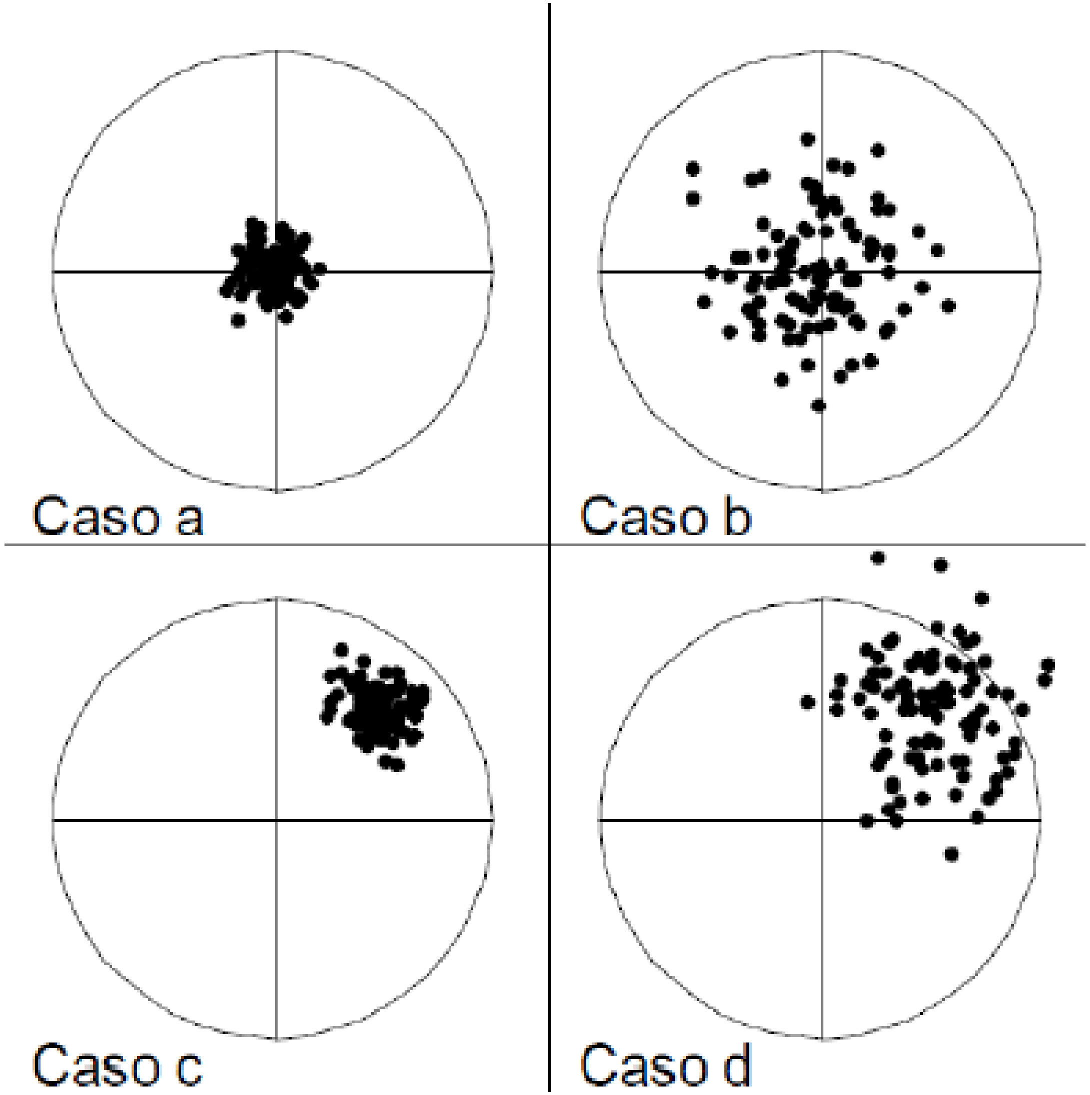
- vicini gli uni agli altri e concentrati attorno al centro del bersaglio: ***misura precisa e accurata*** (caso a)
- vicini gli uni agli altri e concentrati attorno ad un punto lontano dal centro del bersaglio, magari per una srettifica del mirino: ***misura precisa ma non accurata*** (caso c)



Precisione ed accuratezza - 3

- piuttosto dispersi attorno al centro del bersaglio: ***misura poco precisa ma accurata*** (caso b)

- dispersi attorno a un punto lontano dal centro del bersaglio: ***misura poco precisa e poco accurata*** (caso d)



Precisione ed accuratezza - 4

La terminologia usata per parlare della qualità della misura non è sempre coerente e per questo può trarre in inganno.

Spesso infatti si usa il termine precisione come indicatore onnicomprensivo della qualità di una misura, mentre quando si vogliono enucleare i vari aspetti che concorrono a determinarla, si distingue fra precisione e accuratezza.

Sta all'attenzione del lettore o dell'ascoltatore capire quale sia il reale significato dei termini usati valutando il contesto nel quale vengono impiegati.

Accuratezza assoluta e relativa - 1

E’ utile distinguere fra accuratezza assoluta e accuratezza relativa.

Quando si dice che la distanza fra i punti A e B è stata misurata con un’incertezza di 2 mm, si indica l’accuratezza assoluta delle misure. Ma la quantificazione in termini assoluti dell’accuratezza non è sempre la più opportuna.

Sbagliare di un millimetro la misura della distanza Terra-Luna è un risultato straordinario, mentre sbagliare della stessa quantità la misura con il calibro del diametro di un piccolo cilindro metallico costituisce un errore marchiano.

Accuratezza assoluta e relativa - 2

L'***accuratezza assoluta*** è un numero avente la stessa dimensione fisica della quantità misurata:

$$\tilde{x} = \bar{x} + \varepsilon_a$$

la quantità misurata $\tilde{x}$ è uguale al valore vero $\bar{x}$ più l'errore commesso a ε_a .

L'***accuratezza relativa*** è un numero puro uguale al rapporto fra l'errore commesso e l'entità della misura

$$\varepsilon_a = \frac{\varepsilon_a}{\tilde{x}}$$

Quando usare la relativa e l'assoluta? Dipende del contesto.

Esempi di accuratezza relativa

1. Nel caso del GPS le basi vengono determinate con un errore che dipende dalla loro lunghezza e che corrisponde, per lavori di qualità, a 1~mm per Km; la precisione relativa del GPS è dunque di 10^{-6} oppure, come si suole dire, 1 ppm (una parte per milione)

2. Specifiche dello strumento Leica TPS1200

Programma di misurazione EDM	Deviazione standard, ISO 17123-4, prisma standard	Deviazione standard, ISO 17123-4, Target	Durata della misurazione, tipica [s]
Standard	$\boxed{2\ \text{mm} + 2\ \text{ppm}}$	5 mm + 2 ppm	1.5
Veloce	5 mm + 2 ppm	5 mm + 2 ppm	0.8

L'errore è espresso come somma di una parte costante (assoluta) di 2 mm e di una parte che è funzione della distanza, quantifica in termini relativi, 2 ppm.

Le misure di precisione come fenomeni aleatori - 1

In teoria della probabilità, una variabile aleatoria (o variabile casuale) può essere pensata come il risultato numerico di un esperimento.

In particolare, si definisce variabile casuale continua una variabile casuale che può assumere **tutti** valori compresi in un intervallo .

Il lancio di un dado o di una moneta è un esempio di **variabile casuale discreta.**

Le misure di precisione come fenomeni aleatori - 2

Le misure di precisione sono certamente fenomeni aleatori.

Consideriamo ad esempio la misura di un angolo con un teodolite elettronico: l'insieme degli eventi elementari è finito ma davvero molto grande: i numeri con 4 cifre decimali compresi fra 0 e 400.

In tale quadro non ha senso parlare di probabilità che una ripetizione di una misura dia un certo numero; ciò che interessa è la probabilità che una misura cada in un intervallo $[a, b]$.

Se per un certo fenomeno aleatorio si conosce la probabilità di ogni intervallo $[a, b]$, quel fenomeno è descrivibile nei termini di una **variabile casuale continua** X.

Distribuzione di probabilità continue

La Funzione Densità di Probabilità (FDP) continua definisce analiticamente come si distribuiscono i valori assunti da una variabile aleatoria continua.

Se continuo a misurare le coordinate di un vertice e riporto i risultati in un grafico qual è l'aspetto della figura che ottengo?

Una distribuzione normale (o curva di Gauss); essa è la distribuzione continua più utilizzata in statistica.

Quando si dispone di un'espressione matematica adatta alla rappresentazione di un fenomeno continuo, siamo in grado di calcolare la probabilità che la variabile aleatoria assuma valori compresi in intervalli.

La curva di Gauss

E' caratterizzata dai parametri μ e σ.

La **media** μ rappresenta il punto di simmetria e anche il punto in cui assume il valore massimo.

La **deviazione standard** σ rappresenta la larghezza della curva: σ piccolo significa curva piccata, *stretta* (indicata anche come STD - STandard Deviation).

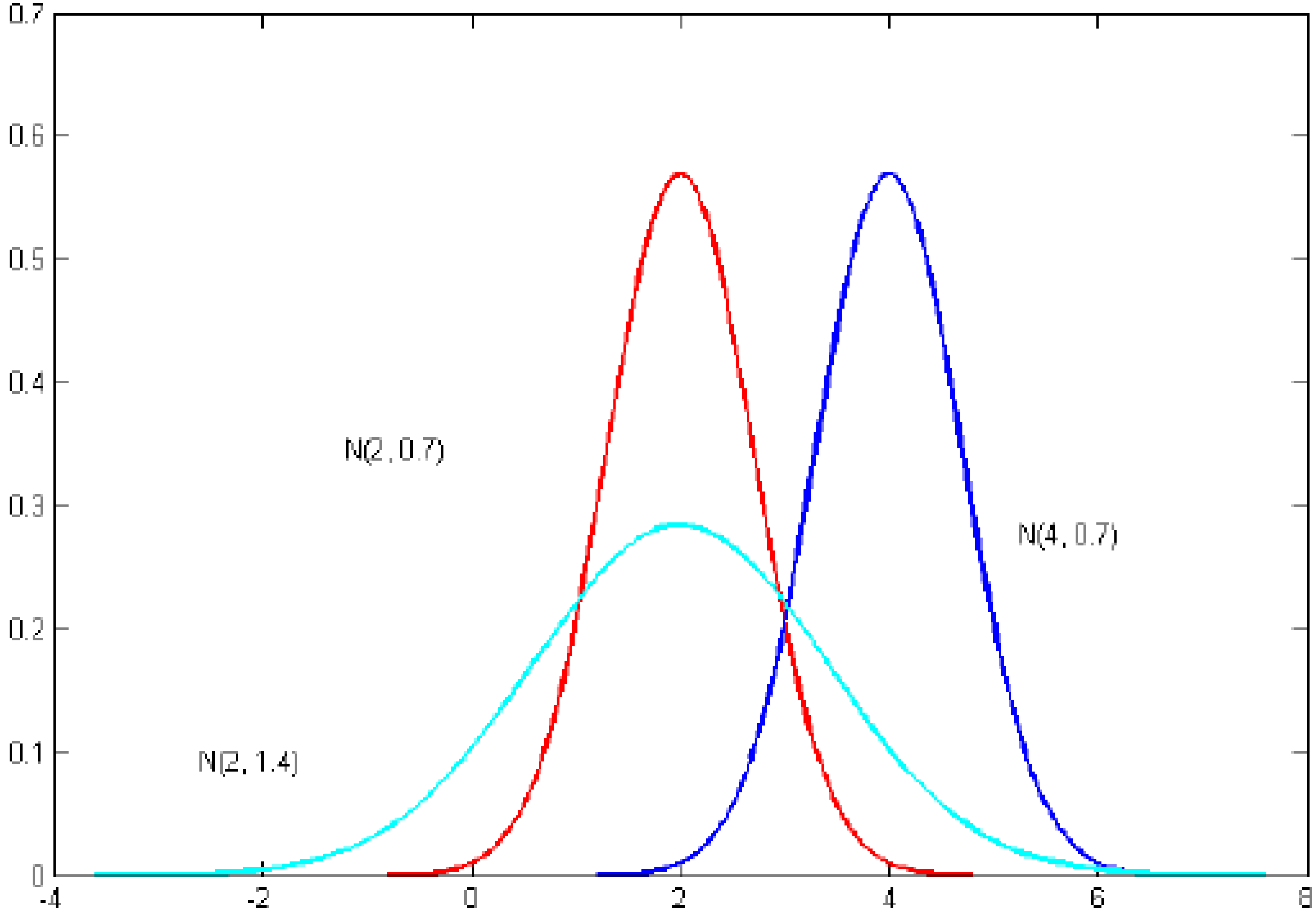
La sua espressione analitica è la seguente:

$$f_N(X;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

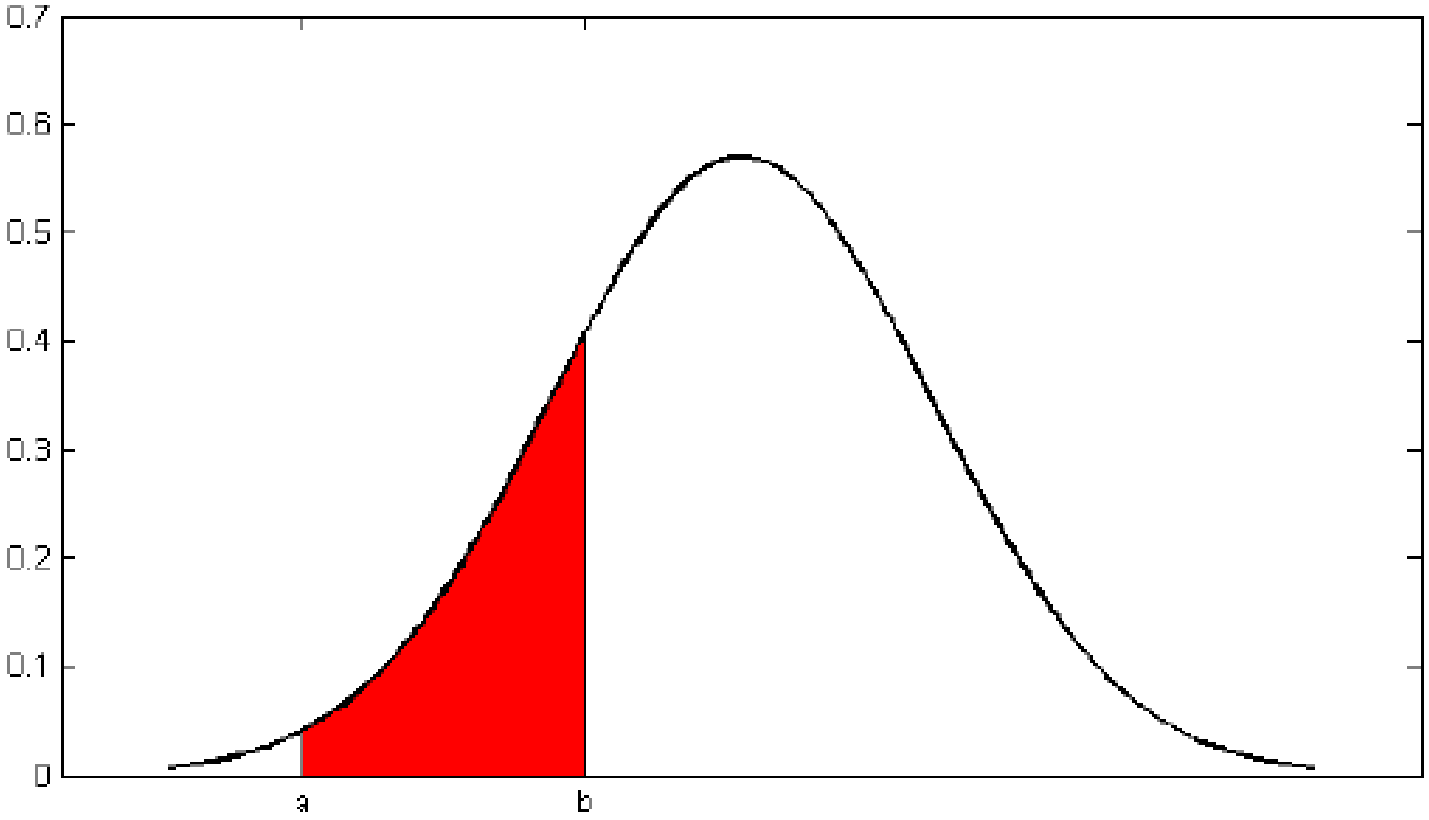
dove:

- ✓ x sono i valori assunti dalla variabile aleatoria

La curva di Gauss - 2



Interpretazione della curva di Gauss



P([a,b]) = area sottesa

La probabilità che la variabile aleatoria assuma valori compresi tra a e b è data dall'area sottesa alla curva.

Ovviamente regioni ed intervalli differenti hanno probabilità differenti.

Stima di media e deviazione standard

Fenomeno aleatorio X.

Disponibilità di *n* misure ripetute di *X*: X_i con *i = 1, 2, ..., n*

La loro media si stima con la formula

$$\mu = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

e la deviazione standard si stima

$$\sigma = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \mu)^2}$$

Un altra figura spesso utilizzata è la varianza: $\sigma^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \mu)^2$

Nella formula della σ si sostituisce ad (n-1) il valore n quando si hanno a disposizione molti dati

Fenomeni aleatori n-dimensionali

Quanto indicato fino a questo momento può essere esteso nel caso di variabili ad n-dimensioni. Esempio: la determinazione di un punto con GPS è un fenomeno aleatorio tridimensionale. Generalizzando:

- ✓ vettore *n*-dimensionale delle medie

$$\mu = (\mu_1, \mu_2, \dots, \mu_n)^t$$

- ✓ matrice di varianza-covarianza

$$C_{XX} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \cdots & \sigma_2^2 & \cdots & \sigma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \sigma_n^2 \end{bmatrix}$$

Sulla diagonale: varianze delle singole componenti.

Fuori dalla diagonale, ad esempio σ_{12}: covarianza fra componente 1 e 2.

Qual è la giusta interpretazione da dare al significato di sigma?

Un esempio:

	Lettura cerchio orizzontale su B	Lettura cerchio orizzontale su C	Angolo interno
1	210,5832	340,9302	130,3470
2	210,5832	340,9326	130,3494
3	210,5814	340,9326	130,3512
4	210,5877	341,0295	130,4418
5	210,6835	341,0273	130,3438
6	210,5127	340,8909	130,3782
			130,3686
			0,0380

Nel caso precedente, per l'angolo interno, avevamo:

- ✓ μ: 130,3686
- ✓ σ: 0,0380

quali sono i valori che può assumere la variabile casuale?

Spesso si sentono affermazioni del tipo:
il valore è 130,3686 ± 0,0380 ossia l'angolo sarà compreso al <u>massimo</u> nell'intervallo di valori [130,3306 - 130,4066].

E' corretta questa affermazione? NO!

Avevamo visto che la probabilità che la variabile aleatoria assuma valori compresi tra a e b è data dall'area sottesa alla curva.

Per la curva di Gauss si ha che:

$P([\mu - \sigma, \mu - \sigma]) = 0.683$

$P([\mu - 2\sigma, \mu - 2\sigma]) = 0.955$

$P([\mu - 3\sigma, \mu - 3\sigma]) = 0.997$

Interpretazione:

facendo 100 misure, in media

68 cadono nell'intervallo 1-sigma;

95 cadono nell'intervallo

2-sigma, ecc..

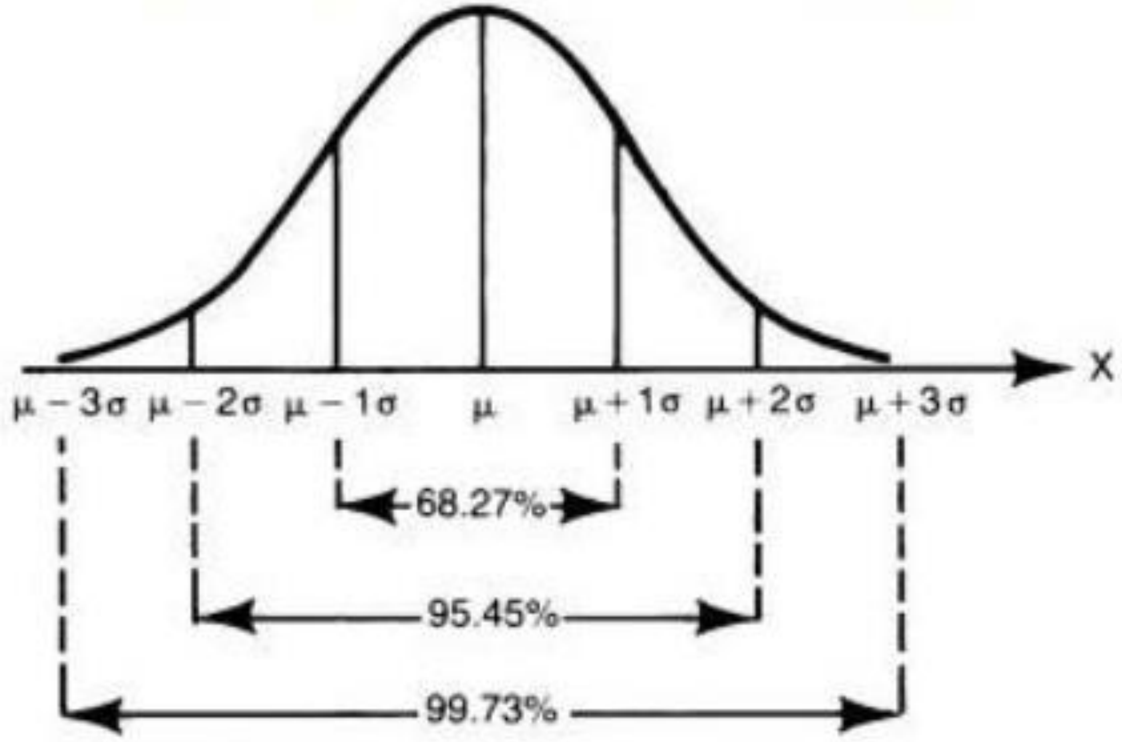


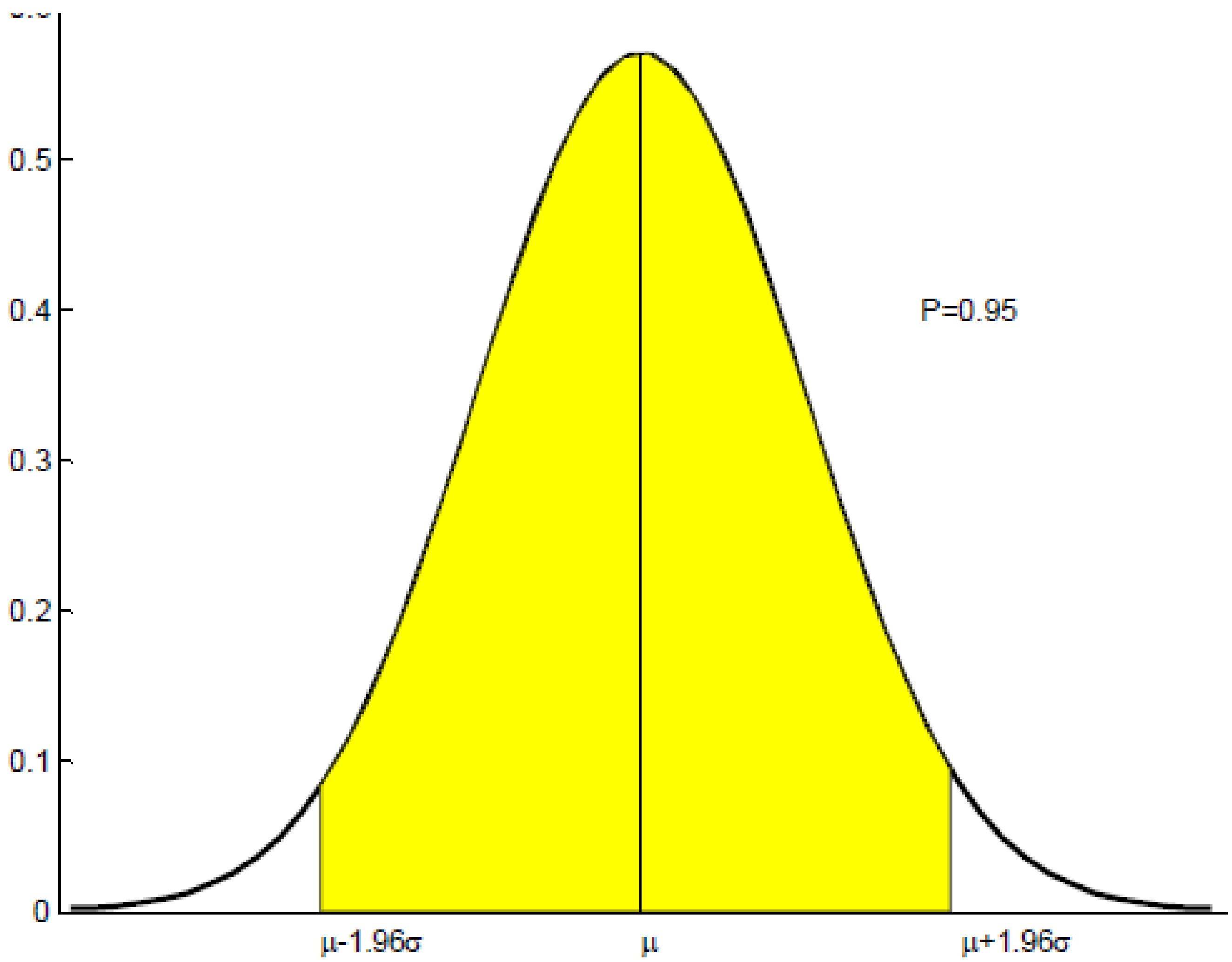
Figure 2

99.73%	99%	95.45%	95%	90%	80%
3.00	2.58	2.00	1.96	1.645	1.28

Intervalli notevoli per la curva di Gauss - 2

Intervallo 1.96-sigma, avente probabilità

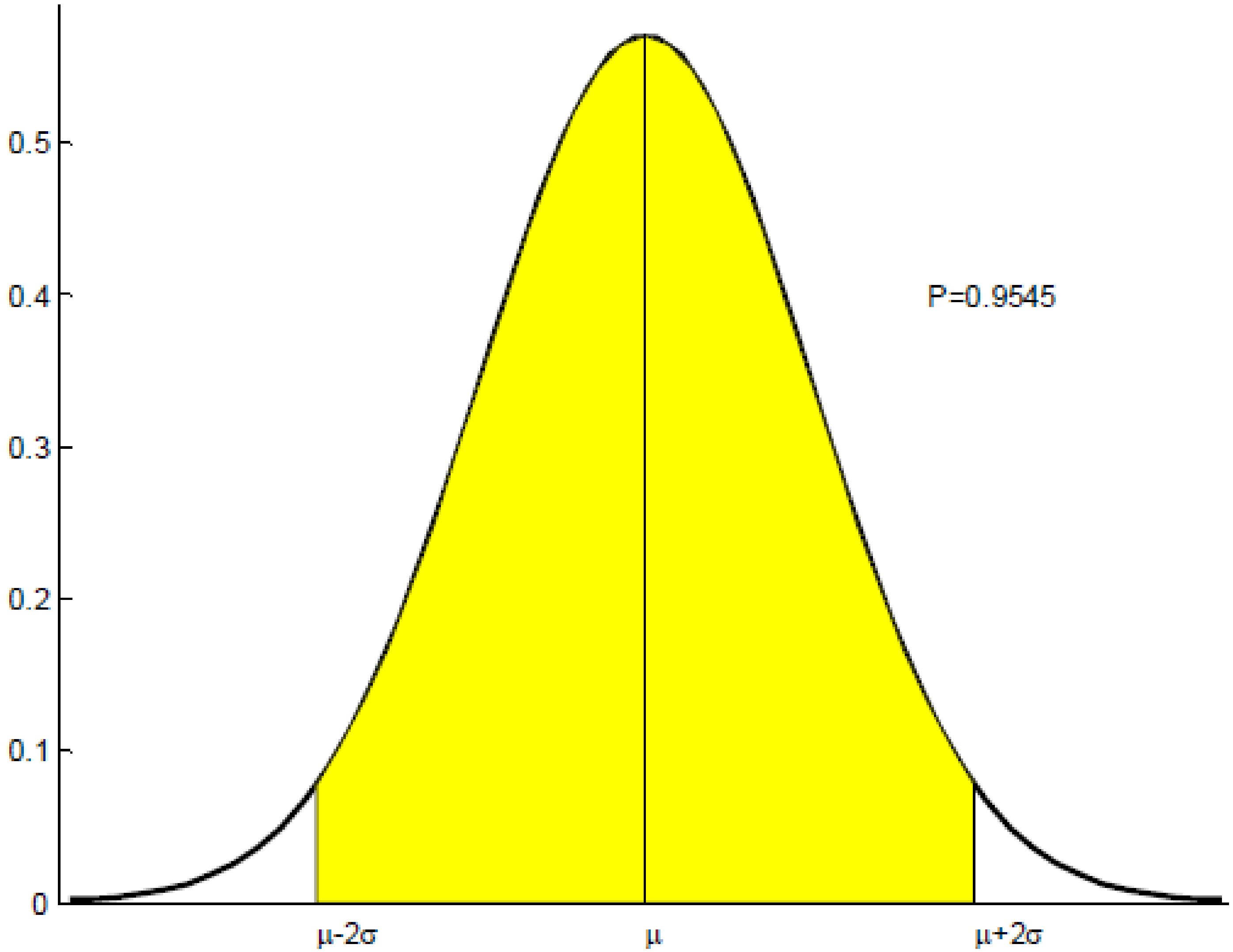
$P = 0.95$



Intervalli notevoli per la curva di Gauss - 3

Intervallo 2-sigma,
avente probabilità

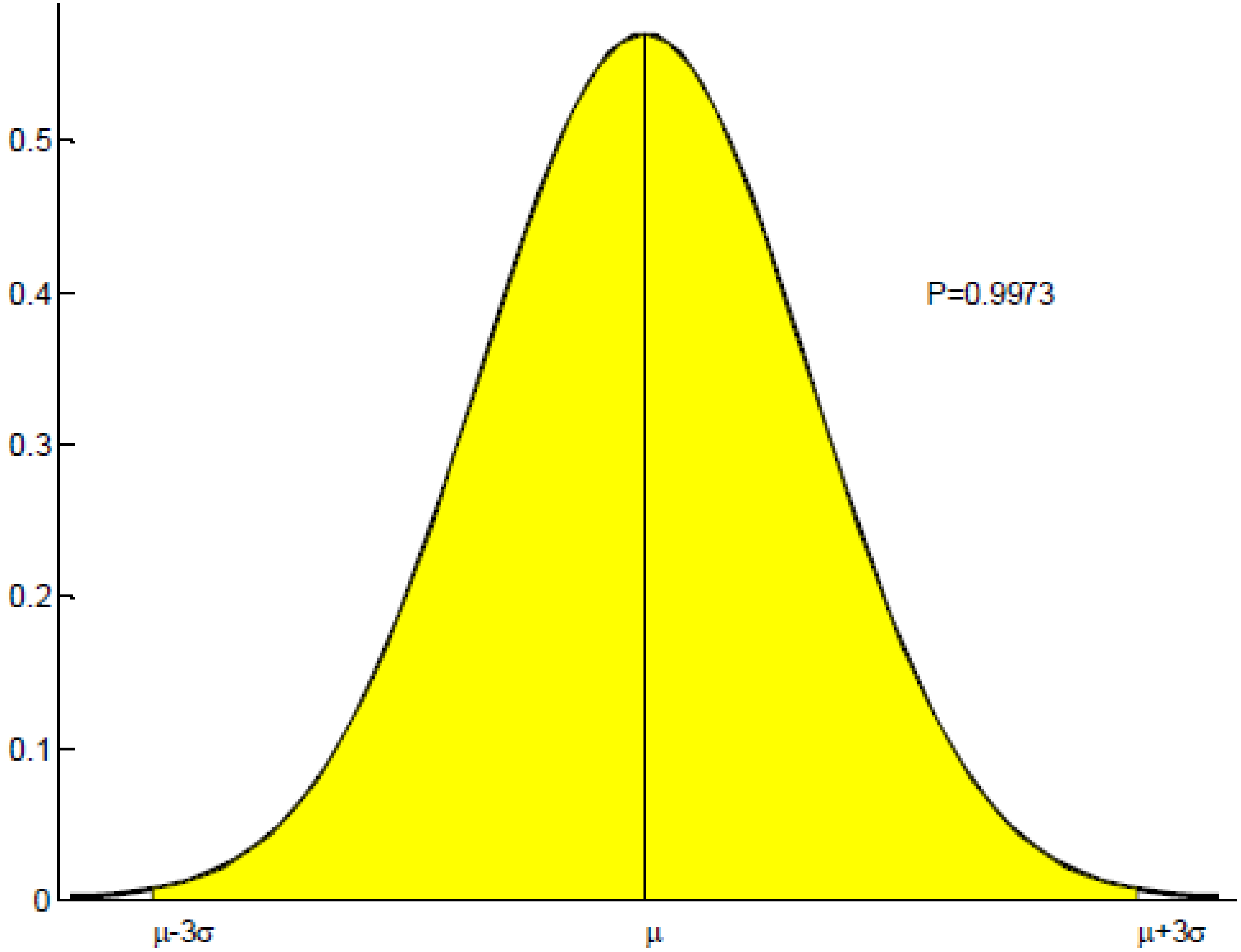
$P = 0.9545$



Intervalli notevoli per la curva di Gauss - 4

Intervallo 3-sigma,
avente probabilità

$P = 0.9973$



Esempio: monitoraggio di una frana

Si vuole monitorare una frana misurando ripetutamente, ad intervalli regolari, la distanza tra il punto A, interno alla frana, ed il punto B, localizzato su un terreno stabile.

La prima serie ripetuta di misure, effettuate il primo giorno, mostra:

- ✓ $\mu = 121.780$ m
- ✓ $\sigma = 0.012$ m $= 1.2$ cm

Una settimana dopo, con lo stesso strumento, viene effettuata una nuova misura:

- ✓ $d = 121.800$ m

La frana si sta muovendo? Occorre evacuare il paese sotto la frana?

Esempio: monitoraggio di una frana - 2

Se la frana non si è mossa, la nuova misura effettuata è frutto della variabilità figlia degli errori accidentali di misura.

Supponiamo di analizzare l'accaduto in termini di livello di significatività al 95%.

Sotto questa premessa ci aspettiamo che le misure di distanza cadano in un intervallo 2σ pari a:

$[121.780 - (2 * 0.012), 121.780 + (2 * 0.012)] = [121.756, 121.804]$

La nuova misura è all'interno dell'intervallo?

SI, statisticamente non c'è evidenza di movimento.

La nuova misura è al di fuori?

SI, è compatibile con un possibile movimento.

Vantaggi delle osservazioni ripetute

Perché occorre effettuate numerose ripetizioni?

1. Per aver la possibilità di individuare errori grossolani
2. Per avere una migliore stima del valore vero della variabile casuale: la media di misure ripetute è più vicina al valore vero delle singole misure
3. Per avere la possibilità di stimare la dispersione delle osservazioni

Riassumendo

Tre tipologie di errori: grossolani, sistematici ed accidentali

Due terminologie per la qualità: precisione ed accuratezza

Due figure statistiche: media e deviazione standard

Esiste un legame tra tutte queste grandezze? SI!

L'esempio del bersaglio

Riutilizzando l'esempio del bersaglio, possiamo considerare la serie di tiri effettuati come una variabile casuale e determinarne il comportamento attraverso una curva di Gauss.

In questo caso, l'ipotetico valore vero che stiamo ricercando è il centro del bersaglio.

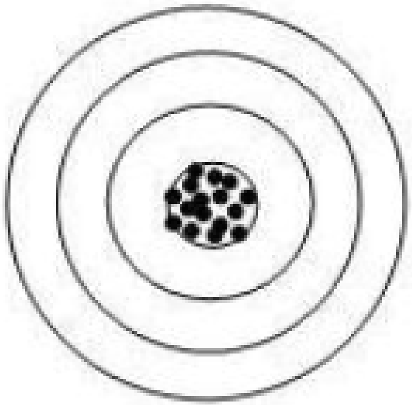
Il tiratore effettuerà una serie di spari e, dall'esisto di queste operazioni ripetute, si cercherà di trarre alcune conclusioni sulla sua abilità.

La **media** dei tiri fornirà un'indicazione sull'**accuratezza** del tiratore mentre la **deviazione standard** sul suo livello di **precisione**.

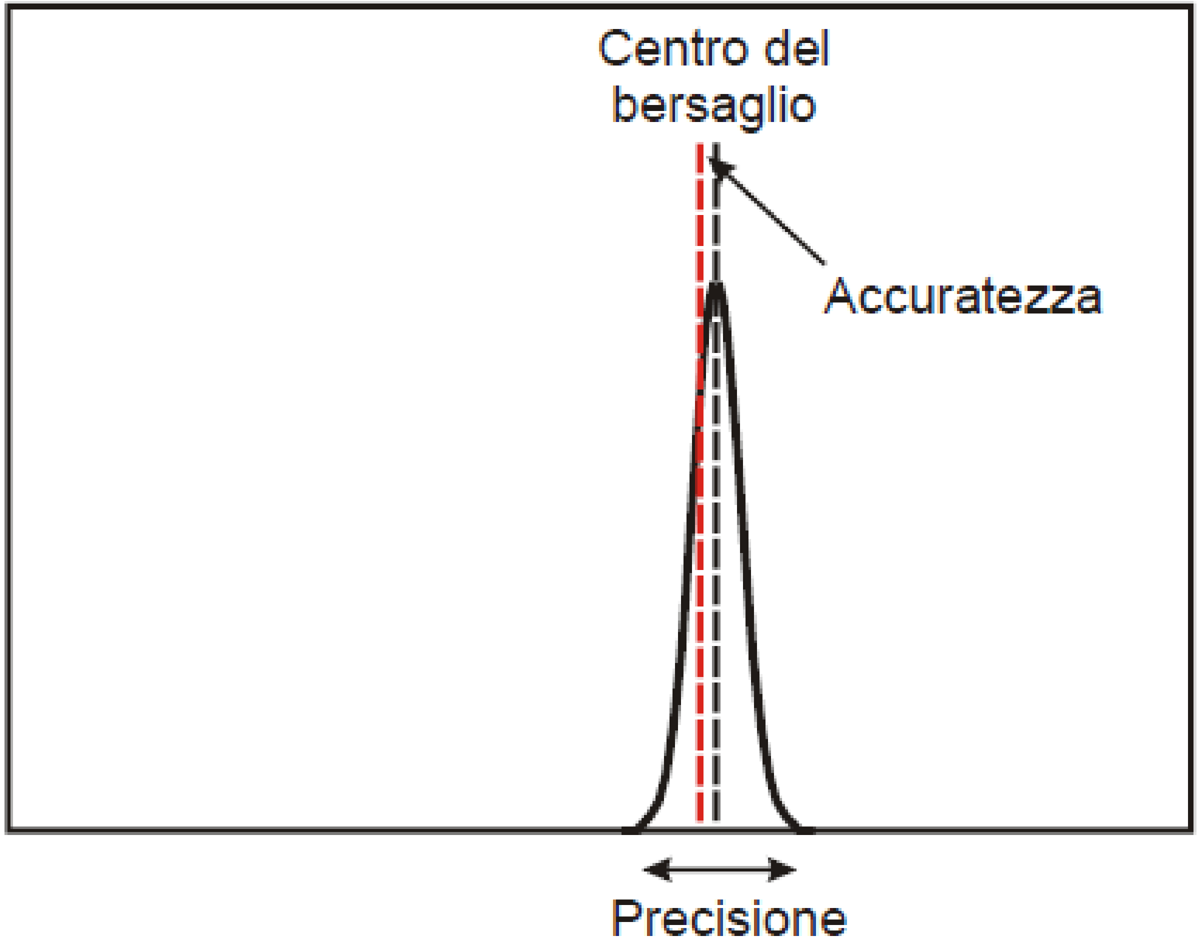
L'esempio del bersaglio - caso d

Caso ideale: il tiratore è molto abile e mostra grande accuratezza e precisione.

La media dei tiri è molto vicina al valore vero (il centro del bersaglio) e la deviazione standard contenuta mostra come la dispersione dei tiri sia limitata.



(d)



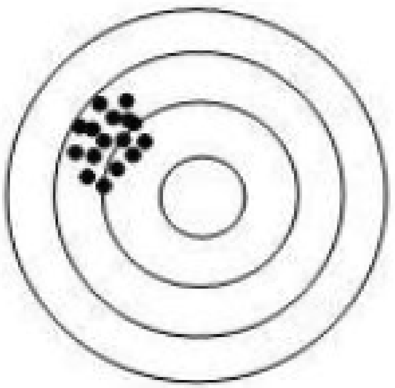
L'esempio del bersaglio - caso c

Tiratore molto preciso ma poco accurato.

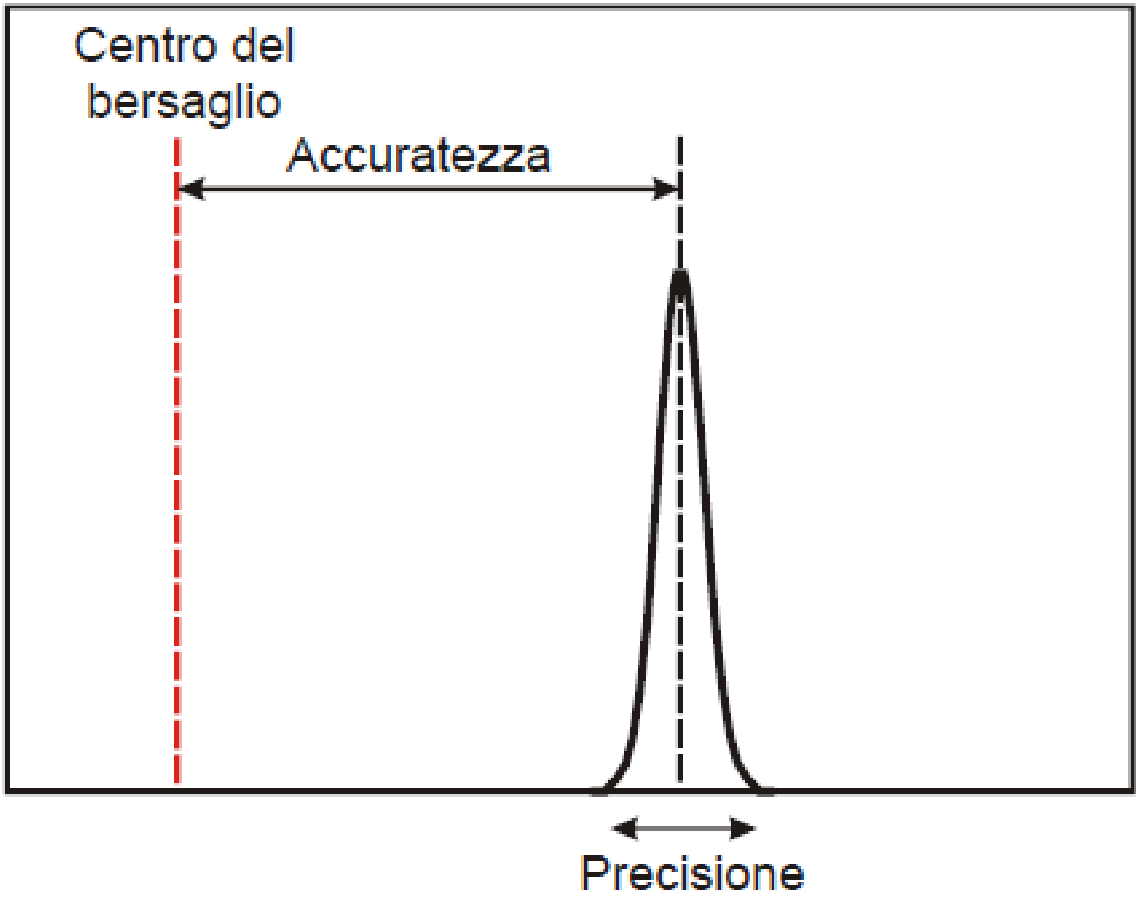
La media dei tiri è lontana dal valore vero (il centro del bersaglio) ma la deviazione standard è contenuta.

Qual è l'origine di tale fenomeno?

Tipicamente la presenza di un **errore sistematico** (minino srettificato, occhio sifulo, etc.)



(c)



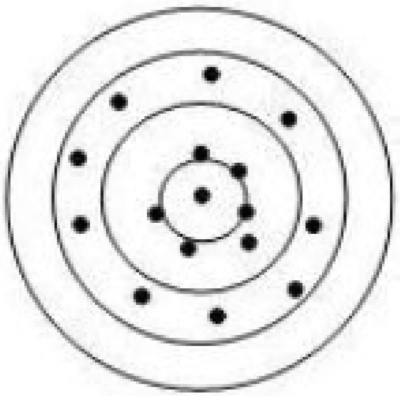
L'esempio del bersaglio - caso b

Tiratore molto accurato ma poco preciso.

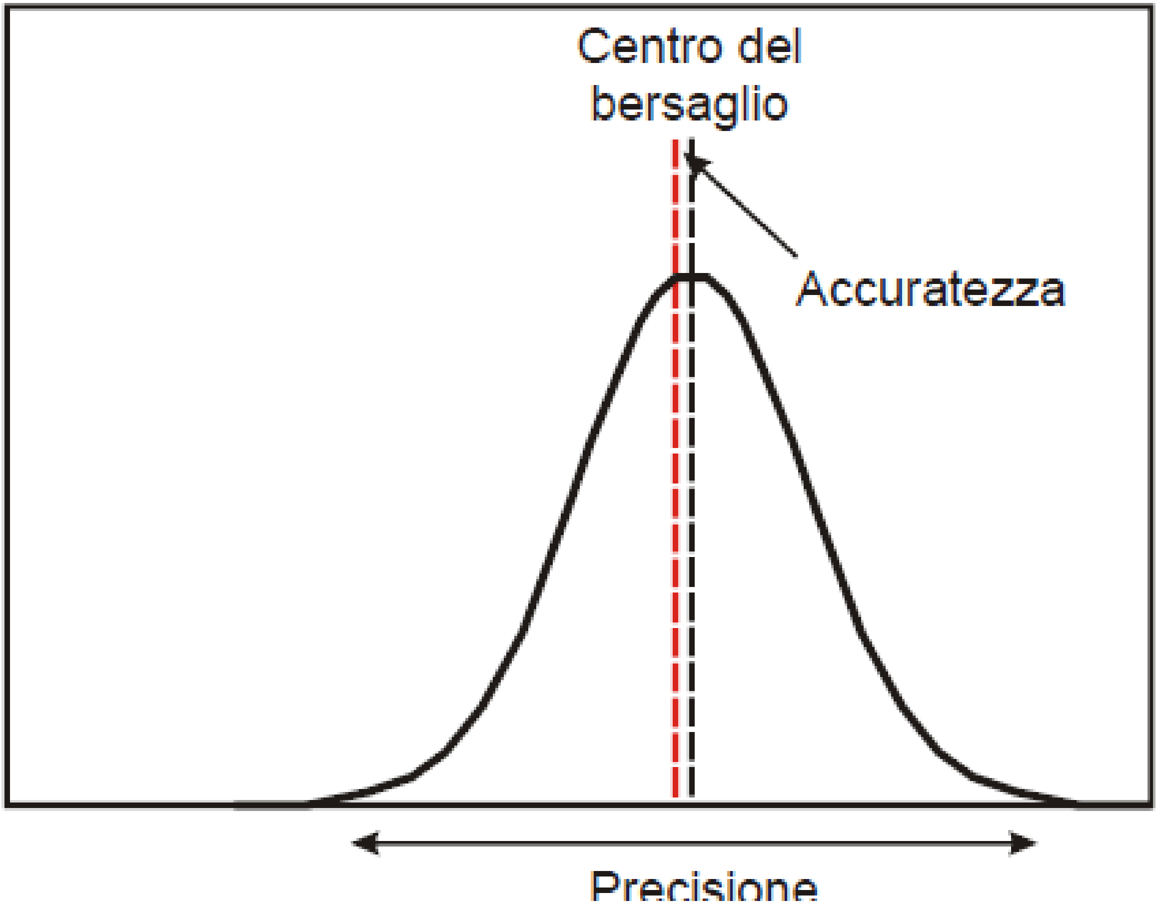
La media dei tiri è molto vicina al valore vero (il centro del bersaglio) ma la deviazione standard è grande sintomo di poca precisione.

Qual è l'origine di tale fenomeno?

Tipicamente la presenza di **errori accidentali** (condizioni ambientali diverse, variazioni nell'impugnatura dell'arma, etc.)



(b)

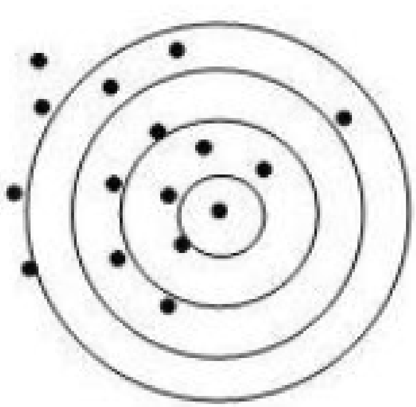


L'esempio del bersaglio - caso a

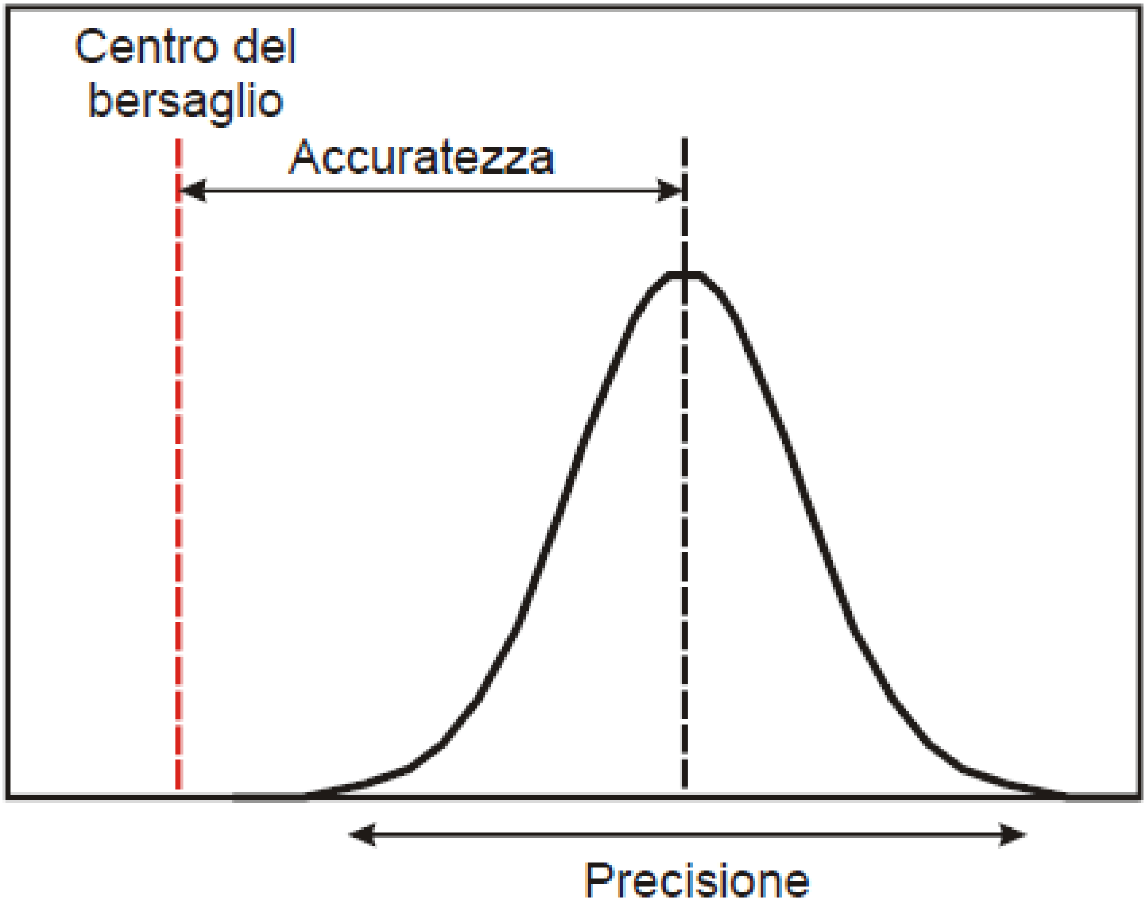
Tiratore poco accurato e preciso.

La media dei tiri è lontana dal valore vero (il centro del bersaglio) e anche la deviazione standard è grande.

Qual è l'origine di tale fenomeno? L'insieme di **errori sistematici** ed **accidentali.**



(a)



Riassumendo - 2

Abbiamo visto come le figure statistiche caratterizzanti al curva di Gauss possono essere interpretate in funzione della presenza di errori sistematici ed accidentali; l'analisi permette inoltre di esprimere un giudizio in termini di accuratezza e precisione.

Un esempio reale

Quando si effettua un rilievo topografico si misurano delle quantità come angoli e distanze. Per non incorrere in errori grossolani si effettuano misure ripetute → questo permette inoltre di analizzare la precisione delle misure effettuate tramite la deviazione standard (σ).

	Lettura cerchio orizzontale su B	Lettura cerchio orizzontale su C	Angolo interno
1	210,5832	340,9302	130,3470
2	210,5832	340,9326	130,3494
3	210,5814	340,9326	130,3512
4	210,5877	341,0295	130,4418
5	210,6835	341,0273	130,3438
6	210,5127	340,8909	130,3782
			130,3686
			0,0380

La media (μ) non fornisce tuttavia nessuna informazione sull'eventuale presenza di errori sistematici ma solo una stima del valore vero.

Un esempio reale - 2

Durante una validazione è però possibile effettuare un'analisi completa.

Validare delle misure significa confrontarle con dei valori ottenuti da una tecnica più precisa e considerati, per questo motivo, "veri" .

Alcuni esempi:

- validare una livellazione GPS su alcuni capisaldi rilevati tramite livellazione geometrica;

Un esempio reale - 2

Validazione dell'accuratezza delle immagini Google Earth su Pavia

Scarti determinati come differenze tra le coordinate derivate da collimazioni su Google Earth e quelle ottenute per via topografica.

# 60 punti	Scarti	
	E [m]	N [m]
Media	15,926	-1,162
STD	0,691	1,108
EQM	15,941	1,606

Media: errore sistematico di circa 16 metri in direzione Est-Ovest

STD: errori accidentali dell'ordine del metro

EQM: ??

Domanda: la media sulla coordinate Nord è un errore sistematico? Test statistici!

Errore quadratico medio

Abbiamo visto che la media può essere utilizzata per stanare gli errori sistematici e che la deviazione standard fornisce un'indicazione sugli errori accidentali.

All'operatore (topografo) piacerebbe aver una stima compatta delle due fonti di errore → errore quadratico medio

$$eqm = \sqrt{media^2 + std^2} = \sqrt{\mu^2 + \sigma^2}$$

L'eqm rappresenta una stima complessiva della qualità delle misura analizzate.

ATTENZIONE: in letteratura è possibile reperire diversi nomi. Spesso l'eqm è indicato come RMSE (Root Mean Squared Error) mentre la std con RMS (Root Mean Error) - bisogna sempre leggere attentamente il significato che hanno nel contesto in cui sono state inserite!

- Si ricorda che eqm è anche uguale a

$$\sqrt{\frac{1}{n}\left({x_1}^2 + {x_2}^2 + \cdots + {x_n}^2\right)}$$

Come si può facilmente dimostrare.