



UNIVERSITÀ DEGLI STUDI  
**eCAMPUS**

**Corsi Pas**

# **FONDAMENTI DI ANALISI STATISTICA**

Lezione 5

# Necessità dell'analisi statistica nella scienza delle misure

L'esecuzione di una misura richiede un confronto tra una quantità incognita e una nota. Nessun risultato di una misura è esente da errore, dove per errore si intende la deviazione o scarto tra la grandezza misurata e il valore della misura atteso (o teorico, o "vero" anche se abbiamo visto che non si può parlare di valore vero). Pertanto il **valore atteso** può essere solo definito in termini **statistici** o probabilistici.

Qualsiasi misura è soggetta a limitazioni: inaccuratezze e imprecisioni che vanno indicate insieme al valore della misura. Esse sono comprese nell'incertezza che è parte integrante della misura.

# Necessità dell'analisi statistica nella scienza delle misure

## ACCURATEZZA

Grado di approssimazione della quantità misurata al valore atteso.

Definendo l'errore relativo come rapporto tra l'errore assoluto (differenza tra valore misurato  $x$  e valore atteso  $X$ ) e il valore atteso  $X$ :

$$er = \frac{x - X}{X}$$

l'accuratezza vale 
$$a = 1 - \left| \frac{x - X}{X} \right|$$

che moltiplicata per 100 dà l'accuratezza percentuale.

Si possono trovare esempi in cui viene data l'accuratezza percentuale dello strumento come 0.5%. In realtà essa rappresenta la sua inaccuratezza (il complemento a 1 dell'accuratezza). L'accuratezza pertanto è 99.5 %.

# Necessità dell'analisi statistica nella scienza delle misure

## PRECISIONE o RIPETIBILITA'

Indicazione numerica dell'approssimazione di un insieme ripetuto di misure della stessa quantità al valor medio dell'insieme delle misure. Indica pertanto il grado di scostamento dei risultati delle misure tra loro.

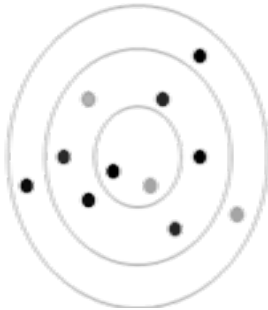
La precisione dell'i-sima misura o scarto rispetto al valor medio di tutte le misure  $\bar{x}$  vale:

$$p_i = 1 - \left| \frac{x_i - \bar{x}}{\bar{x}} \right|$$

La precisione dell'intero insieme di misure si può utilizzare il valor medio delle precisioni di ogni misura. In realtà si preferisce usare la deviazione standard che verrà definita tra poco.

# Necessità dell'analisi statistica nella scienza delle misure

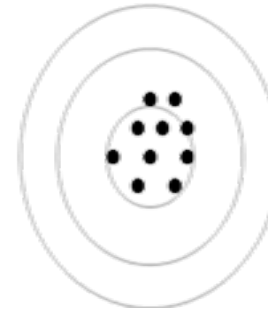
## ACCURATEZZA E PRECISIONE



Risultato accurato  
ma non preciso.



Risultato preciso  
ma non accurato.



Risultato accurato  
e preciso.

# **Necessità dell'analisi statistica nella scienza delle misure**

## **Classificazione degli errori**

Errori grossolani:

Dovuti a imperizia e distrazioni dell'operatore (letture errate, uso improprio di strumenti, errori di elaborazione). Possono essere eliminati ripetendo l'esperimento.

Errori sistematici:

Si ripetono con lo stesso segno e ampiezza ripetendo la misura con la stessa strumentazione in condizioni ambientali e operative immutate. Sono dovuti a non corretta taratura o a difetti costruttivi degli strumenti. Influenzano l'accuratezza della misura. Per evidenziare la loro entità i risultati della misura vanno confrontati con i risultati ottenuti usando strumenti o metodi più accurati.

Errori accidentali o random:

Permangono anche quando sono stati eliminati gli errori grossolani e quelli sistematici. Sono dovuti a imprevedibili fluttuazioni delle condizioni operative, strumentali e ambientali.

# **Necessità dell'analisi statistica nella scienza delle misure**

## **Classificazione degli errori**

Gli errori sistematici possono essere eliminati ripetendo la taratura dello strumento.

Gli errori accidentali possono essere quantificati mediante la taratura e rappresentati dall'incertezza calcolata durante l'operazione di taratura. Essendo di carattere aleatorio, possono essere trattati solo con metodi statistici .

La riduzione degli errori sistematici migliora l'ACCURATEZZA.

Quanto più gli errori accidentali sono piccoli tanto più la misura è PRECISA.



# Elementi di analisi statistica

## Concetti di campione e universo

- Ogni serie di valori estratta dalla totalità dei valori possibili, detta anche universo, può essere considerato un campione.
- Esistono vari metodi per estrarre un campione statistico di  $n$  elementi, che sia rappresentativo dell'universo di  $N$  elementi, uno di essi consiste nell'estrarre gli elementi del campione in maniera casuale.
- La migliore approssimazione o il valore più probabile del campione è rappresentata dalla **media** aritmetica dei valori assunti dagli  $n$  elementi:

$$m = \bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$



# Elementi di analisi statistica

## Media e dispersione intorno alla media

- La **media** rappresenta la stima del valore atteso e la differenza tra il valor medio e il valore atteso è detta **bias**, o indice di inaccuratezza della misura.
- Per determinare la precisione della misura si considera la dispersione dei dati intorno al valor medio. La deviazione dell'i-sima misura  $x_i$  dal valor medio vale:

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

Essa è detta anche residuo.

Per quantificare la dispersione dell'insieme delle misure si potrebbe valutare la media delle deviazioni, che però, in base alla sua definizione è sempre nulla e quindi non è un indice valido:

$$\frac{1}{n} \cdot \sum_{i=1}^n e_i = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x}) = \frac{1}{n} \cdot \sum_{i=1}^n x_i - \bar{x} = \bar{x} - \bar{x} = 0$$

# Elementi di analisi statistica

## Deviazione standard

- La **deviazione standard** rappresenta la migliore misura della dispersione. E' definita in termini dei quadrati delle deviazioni dalla media:

$$s = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n e_i^2} = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$$

La deviazione standard è indice dell'imprecisione della misura.

- La **varianza** è il quadrato della deviazione standard o scarto quadratico medio

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2$$

- Si divide per  $(n-1)$  perché, in riferimento agli  $n$  valori del campione, si può pensare che gli  $n$  gradi di libertà del sistema si riducono di una unità perché ne è già stata calcolata la media.

## Elementi di analisi statistica

### Proprietà della media e della deviazione standard

Definito il concetto di valor medio la somma algebrica degli scarti rispetto al valore medio è nulla.

La media gode della proprietà di rendere minima la somma dei quadrati degli scarti (varianza).

La radice quadrata della varianza  $s$ , costituisce una stima della dispersione delle misure intorno al valore medio, al pari di  $s^2$ , ma ha il pregio di avere le stesse dimensioni delle misure  $x$ .

# Elementi di analisi statistica

## Media e varianza dell'universo

I due parametri precedenti nel caso della popolazione, o universo composto di  $N$  elementi, si indicano con i simboli:

$$\mu = \frac{1}{N} \cdot \sum_{i=1}^N x_i$$

$$\sigma = \frac{1}{N} \cdot \sum_{i=1}^N (x_i - \mu)^2$$

## Elementi di analisi statistica

### Rappresentazione dei dati appartenenti al campione

- Il metodo più rappresentativo per rappresentare graficamente gli elementi appartenenti a un campione si basa sull'utilizzo della frequenza  $F_i$  delle misure, cioè il numero di volte che la generica misura  $x_i$  si ripete all'interno delle  $n$  misure. Si definisce frequenza relativa:

$$f_i = \frac{F_i}{n}$$

Nel caso in cui tutte le misure fossero diverse le une dalle altre, la frequenza relativa sarebbe uguale per tutte le misure, ovvero  $1/n$ .

- Per evitare questo rischio, conviene raggruppare sottoinsiemi di misure in classi di appartenenza. Si definisce un certo numero di intervalli tra il valore minimo e il valore massimo misurati (non necessariamente della stessa ampiezza) che definiscono le diverse classi di appartenenza. La frequenza relativa è allora rappresentata dal numero delle misure che cadono in ogni classe diviso  $n$ .

# Elementi di analisi statistica

## Suddivisione in classi di appartenenza

All'interno d'ogni campione s'individuano delle classi d'appartenenza, definibili in differenti modi; tra questi si propone il seguente ad ampiezza di classe costante.

Siano  $a$  e  $b$  il valore massimo e minimo all'interno del campione e si vogliano creare  $K$  classi di appartenenza. In tal caso l'ampiezza di ogni classe  $G_j$ , è definita da  $\Delta x$ :

$$\Delta x = \frac{(a - b)}{K}$$

Il termine generico  $x_i$  appartiene alla classe  $G_j$ , se:

$$a + (j - 1) \cdot \Delta x \leq x_i \leq a + j \cdot \Delta x$$

# Elementi di analisi statistica

## Suddivisione in classi di appartenenza - Frequenza assoluta e relativa

Il numero dei dati che appartengono a una determinata classe  $j$  si chiama frequenza e viene indicato con  $f_j$ .

Si distingue poi in frequenza assoluta e relativa.

La prima è pari al numero di elementi appartenenti ad una classe, mentre la seconda divide il numero d'elementi appartenenti per il numero totale di soggetti componenti il campione  $f_j = f_j / n$ .

Quest'ultima può anche essere espressa sotto forma percentuale  $f_{p,j} = f_j / n \cdot 100$ .

Si può calcolare la densità di frequenza come  $d_j = f_{p,j} / \Delta x$ , con  $\Delta x$  l'ampiezza della classe.

La frequenza cumulata all'aumentare delle classi si può calcolare come:

$$f_{c,j} = \sum_{j=1}^k f_{p,j}$$



# Elementi di analisi statistica

**Suddivisione in classi di appartenenza  
- Frequenza assoluta e relativa –  
Esempio: misura dimensionale**

Dato un campione di 100 elementi, esso può essere suddiviso in classi di appartenenza ( $K=16$ ) come nella tabella successiva.

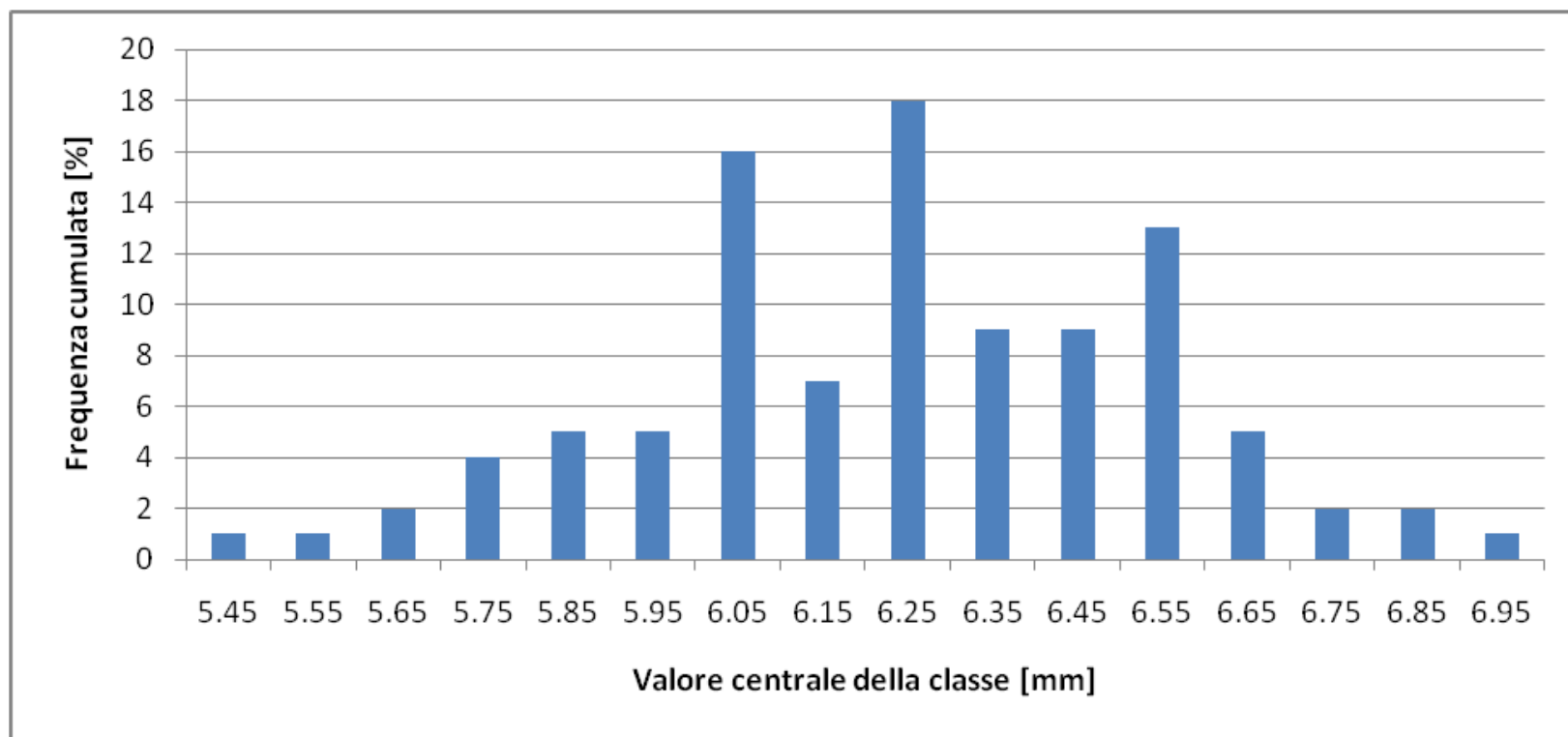
5.80	6.50	6.08	6.34	6.36
5.46	6.11	6.07	6.03	6.66
6.50	6.54	6.85	6.26	6.22
5.62	6.10	6.51	5.69	6.19
6.23	6.53	6.07	5.97	6.75
6.22	6.41	6.42	5.95	6.05
6.92	6.14	6.09	6.38	6.21
6.19	5.93	6.44	6.21	6.50
6.05	5.88	6.66	5.84	6.26
6.29	6.12	5.70	6.21	6.66
6.29	6.07	6.54	5.99	6.54
6.23	6.08	6.68	6.00	6.30
6.02	6.52	6.23	6.49	5.96
5.82	6.11	6.77	6.25	6.39
6.31	6.58	6.26	6.34	5.77
5.78	6.04	5.81	6.49	6.29
5.88	6.52	5.51	6.27	6.47
6.64	6.04	6.10	6.39	6.28
6.08	6.27	6.44	6.43	6.49
6.17	6.52	6.31	6.58	6.86



Limiti della classe (K=16)		Valore centrale	Frequenza assoluta	Frequenza percentuale	Densità di frequenza	Frequenza cumulata percentuale
$x_{\min}$ [mm]	$x_{\max}$ [mm]	$x$ [mm]	$f_j$	$fp,j$ [%]	$D$ [%/mm]	$fc,j$ [%]
5.4	5.5	5.45	1	1	10	1
5.5	5.6	5.55	1	1	10	2
5.6	5.7	5.65	2	2	20	4
5.7	5.8	5.75	4	4	40	8
5.8	5.9	5.85	5	5	50	13
5.9	6	5.95	5	5	50	18
6	6.1	6.05	16	16	160	34
6.1	6.2	6.15	7	7	70	41
6.2	6.3	6.25	18	18	180	59
6.3	6.4	6.35	9	9	90	68
6.4	6.5	6.45	9	9	90	77
6.5	6.6	6.55	13	13	130	90
6.6	6.7	6.65	5	5	50	95
6.7	6.8	6.75	2	2	20	97
6.8	6.9	6.85	2	2	20	99
6.9	7	6.95	1	1	10	100

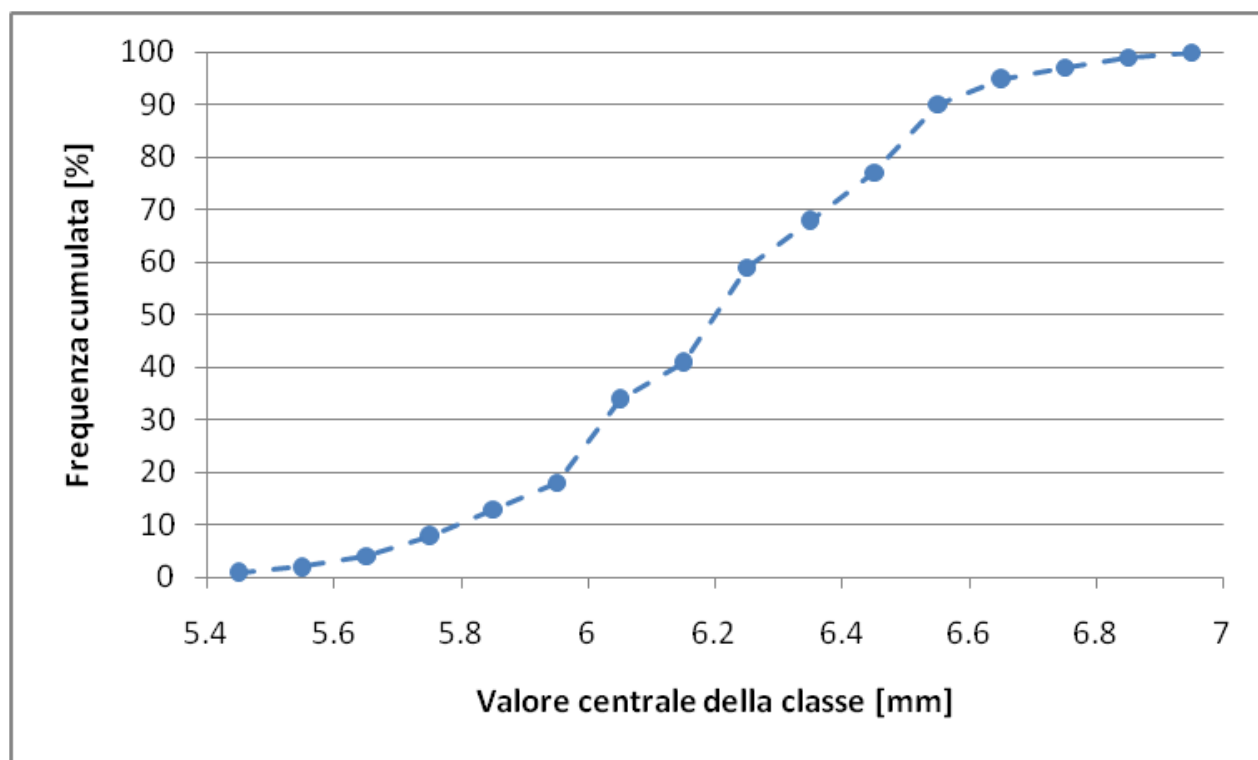
## Elementi di analisi statistica

**Suddivisione in classi di appartenenza - Frequenza assoluta e relativa – Esempio: misura dimensionale: Istogramma delle frequenze della distribuzione data nella tabella precedente**



## Elementi di analisi statistica

**Suddivisione in classi di appartenenza - Frequenza assoluta e relativa – Esempio: misura dimensionale: Diagramma delle frequenze cumulate date nella tabella precedente**



# Elementi di analisi statistica

## Definizione di probabilità

La probabilità  $p_i$  che un elemento del campione si trovi all'interno dell'intervallo  $i$ -esimo che definisce la classe  $G_j$ , è definita come:

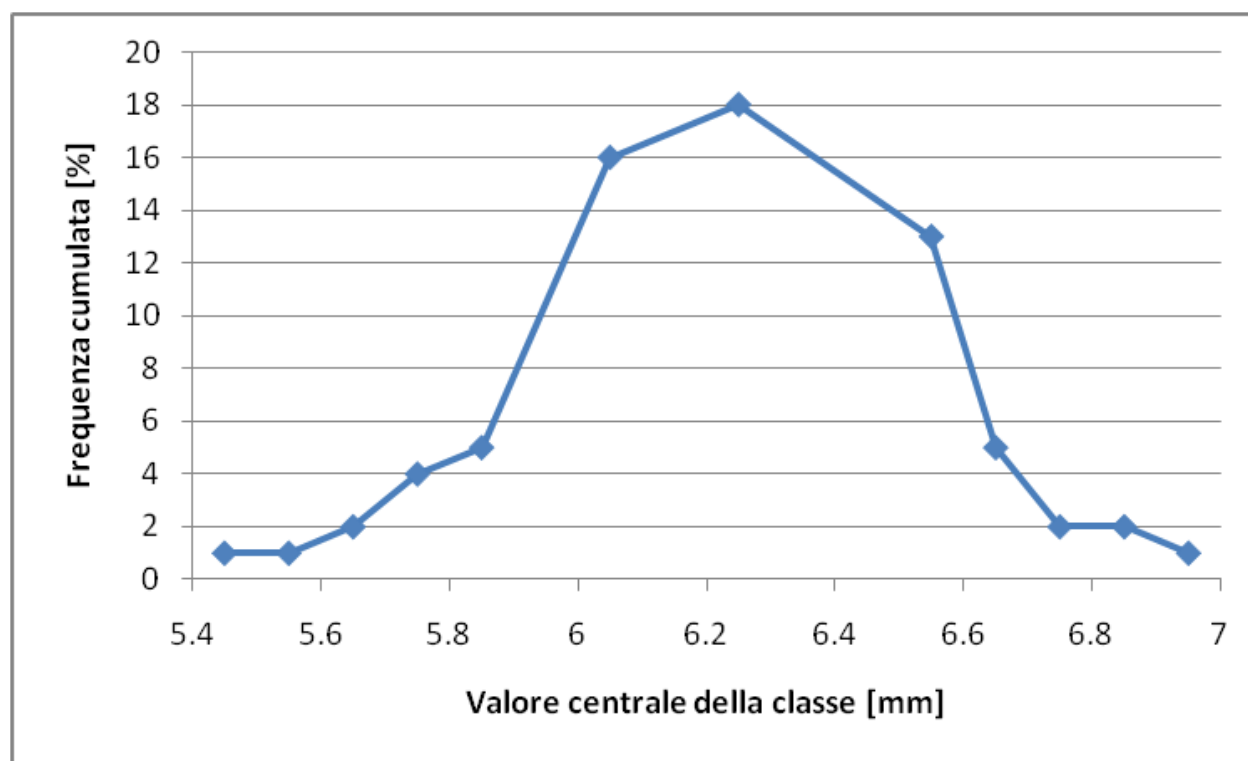
$$p_j = \lim_{n \rightarrow \infty} \frac{f_j}{n}$$

## Definizione di Distribuzione di Probabilità

- La distribuzione di probabilità è l'andamento delle probabilità  $p_j$  in funzione delle  $K$  classi d'appartenenza.
- La rappresentazione della distribuzione di probabilità può essere fatta o con l'istogramma delle frequenze o con il poligono delle frequenze, cioè mediante  $K$  punti discreti.

## Elementi di analisi statistica

**Suddivisione in classi di appartenenza - Frequenza assoluta e relativa – Esempio: misura dimensionale: Poligono delle frequenze (interpolante dei punti sperimentali)**



# Elementi di analisi statistica

## Funzione densità di probabilità

Nei diagrammi mostrati, è stata riportata come ordinata la frequenza percentuale; tuttavia spesso è più utile ricorrere ad una nuova funzione detta densità di probabilità. Nel caso di un numero infinito di dati, come classi di appartenenza, si potrebbero scegliere intervalli infinitesimi. In tal modo si avrebbe un numero infinito di classi aventi ciascuna un numero finito di dati  $z$ . I dati pertanto non saranno più rappresentati da un istogramma o da un poligono (spezzata) ma da una curva continua detta funzione densità di probabilità. definita come il rapporto fra la frequenza percentuale e l'ampiezza della classi d'appartenenza:

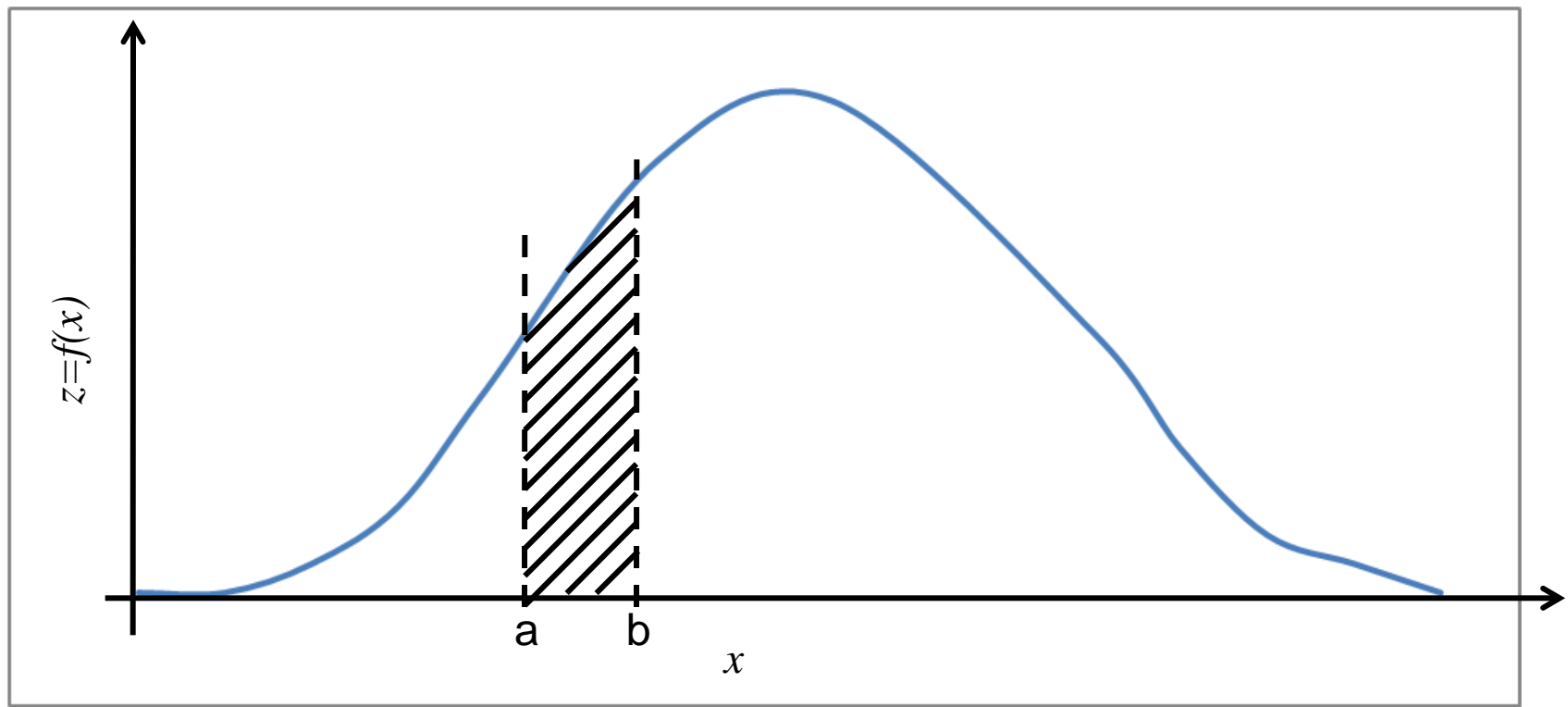
$$z = f(x) = \lim_{\Delta x \rightarrow 0} \frac{f_p}{\Delta x}$$



# Elementi di analisi statistica

## Funzione densità di probabilità

$$z = f(x) = \lim_{\Delta x \rightarrow 0} \frac{f_p}{\Delta x}$$



# Elementi di analisi statistica

## Funzione di distribuzione cumulata

Integrando fra due punti  $a$  e  $b$  del diagramma si ottiene la probabilità che il valore  $x$  si trovi all'interno dell'intervallo considerato.

$$F(x) = \int_a^b f(x)dx$$

La funzione di distribuzione cumulata è definita come:

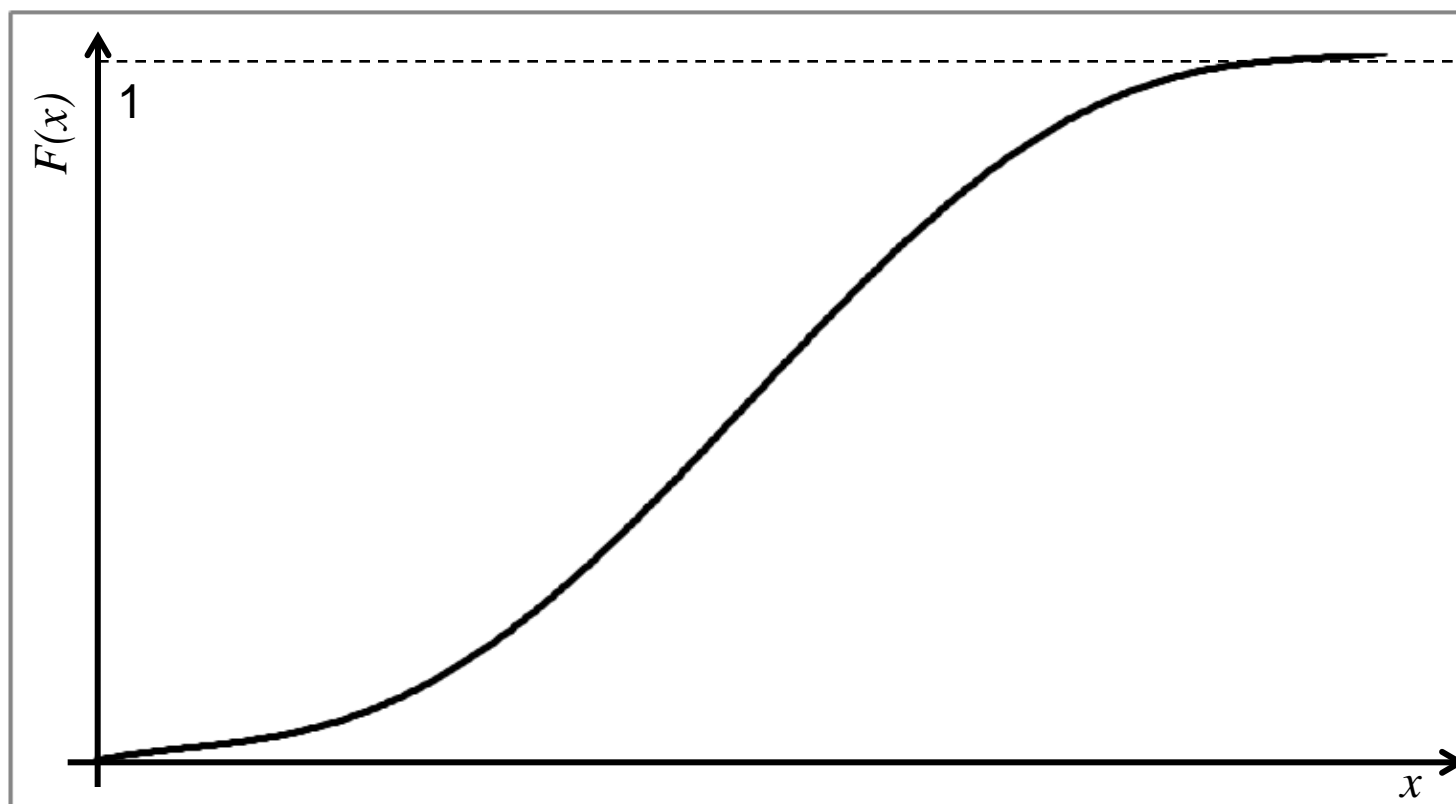
$$F(x) = \int_{-\infty}^x f(x)dx$$

ed esprime la probabilità che la lettura sia minore o uguale a un certo valore  $x$ . Naturalmente la probabilità di trovare il valore  $x$  su tutto il campo di esistenza di  $x$  stesso è 1:

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

# Elementi di analisi statistica

## Funzione di distribuzione cumulata



## Elementi di analisi statistica

### Funzione di densità di distribuzione normale o gaussiana

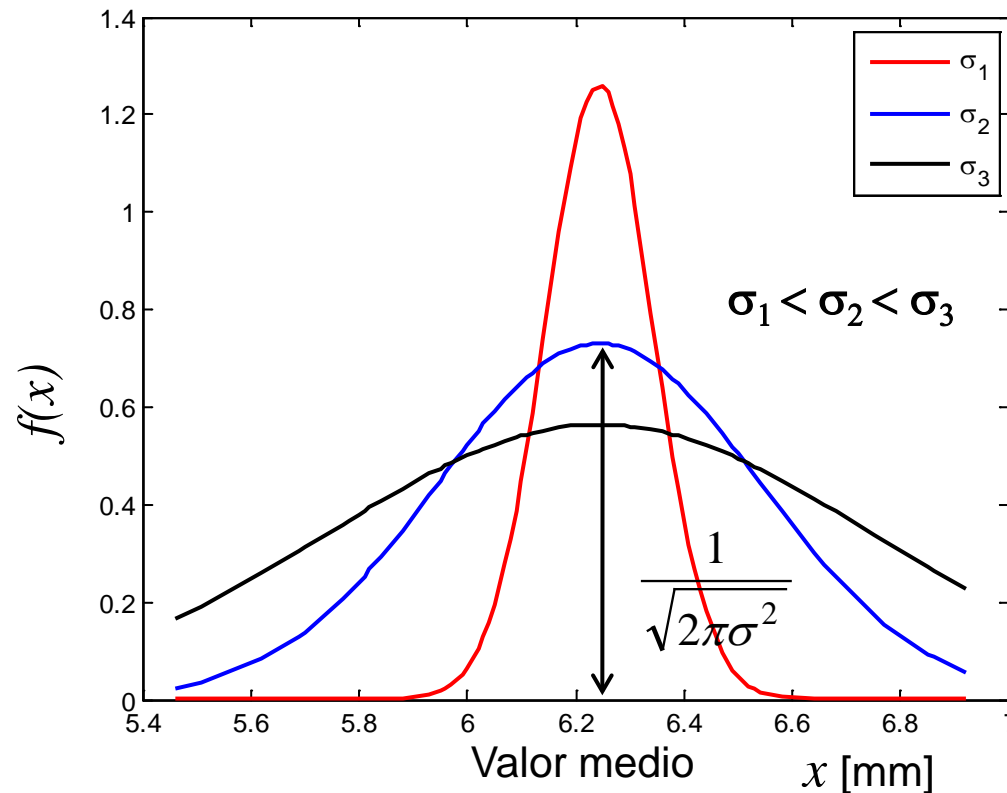
Se le misure sono sufficientemente numerose e tendono ad un valore centrale finito e se la variabilità delle misure dipende da molte cause ciascuna di esse avente una tendenza debole a scostare il risultato dal suo valore atteso di quantità positive o negative con uguale probabilità la funzione di densità di distribuzione di probabilità  $z=f(x)$  può essere rappresentata dalla legge di Gauss:

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

Essa è definita dai parametri caratterizzanti il campione statistico: il suo valor medio ( $\mu$ ) e la sua deviazione standard ( $\sigma$ ), indice dell'incertezza della misura.

## Elementi di analisi statistica

### Funzione di densità di distribuzione normale o gaussiana



## Elementi di analisi statistica

### Funzione di densità di distribuzione normale o gaussiana

La forma della curva è controllata da  $\sigma$  che definisce la larghezza della campana, mentre  $\mu$  fissa la posizione della curva lungo l'asse delle ascisse. L'area sottesa dalla curva è la probabilità di trovare un elemento  $x$  in tutto il campo di esistenza del campione e pertanto vale 1.

Integrando la curva tra  $\mu - \sigma$  e  $\mu + \sigma$  si ottiene il valore 0.683, ovvero la probabilità che la lettura cada tra  $\mu \pm \sigma$  è del 68%.

La probabilità che la lettura cada tra  $\mu \pm 2\sigma$  è del 95%.

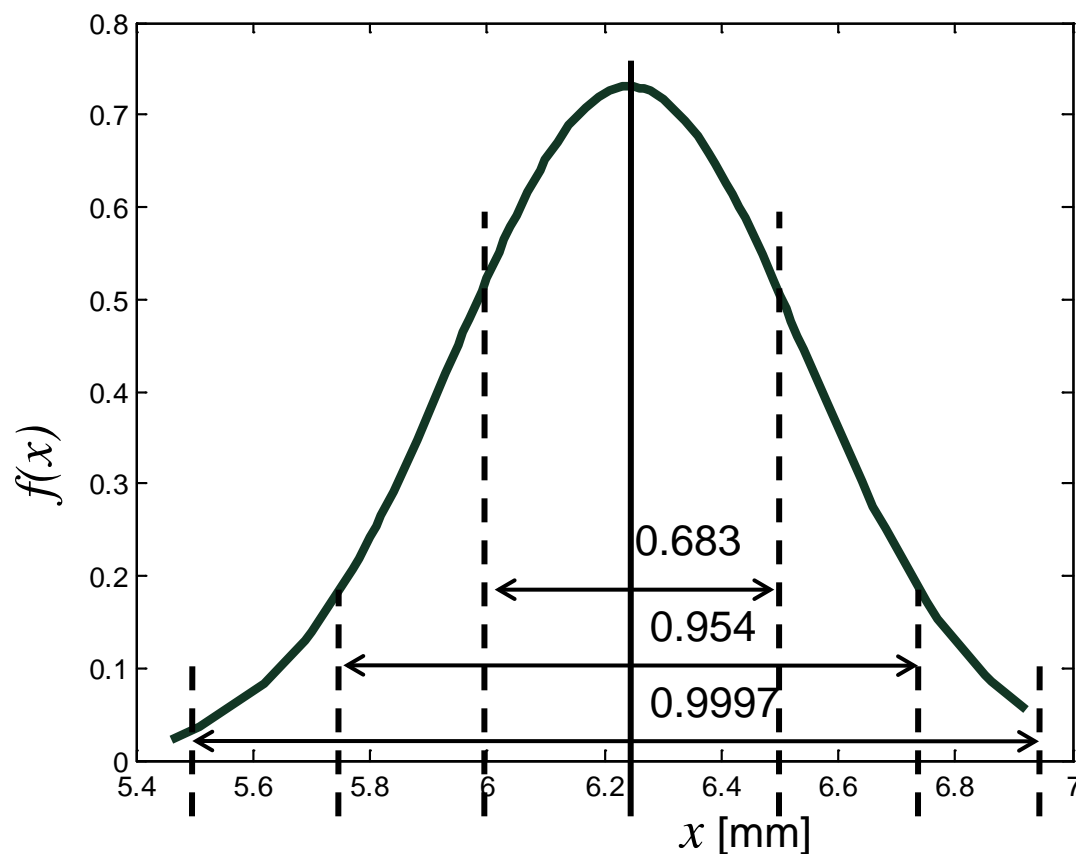
La probabilità che la lettura cada tra  $\mu \pm 3\sigma$  è del 99.7%.

I numeri reali 1, 2, 3 si chiamano livelli di confidenza o fattori copertura.

**Il 99.7 % delle letture cade ad una distanza dal valor medio di  $\pm 3\sigma$**   
**INCERTEZZA ESTESA (deviazione standard x fattore di copertura)**

## Elementi di analisi statistica

### Funzione di densità di distribuzione normale o gaussiana





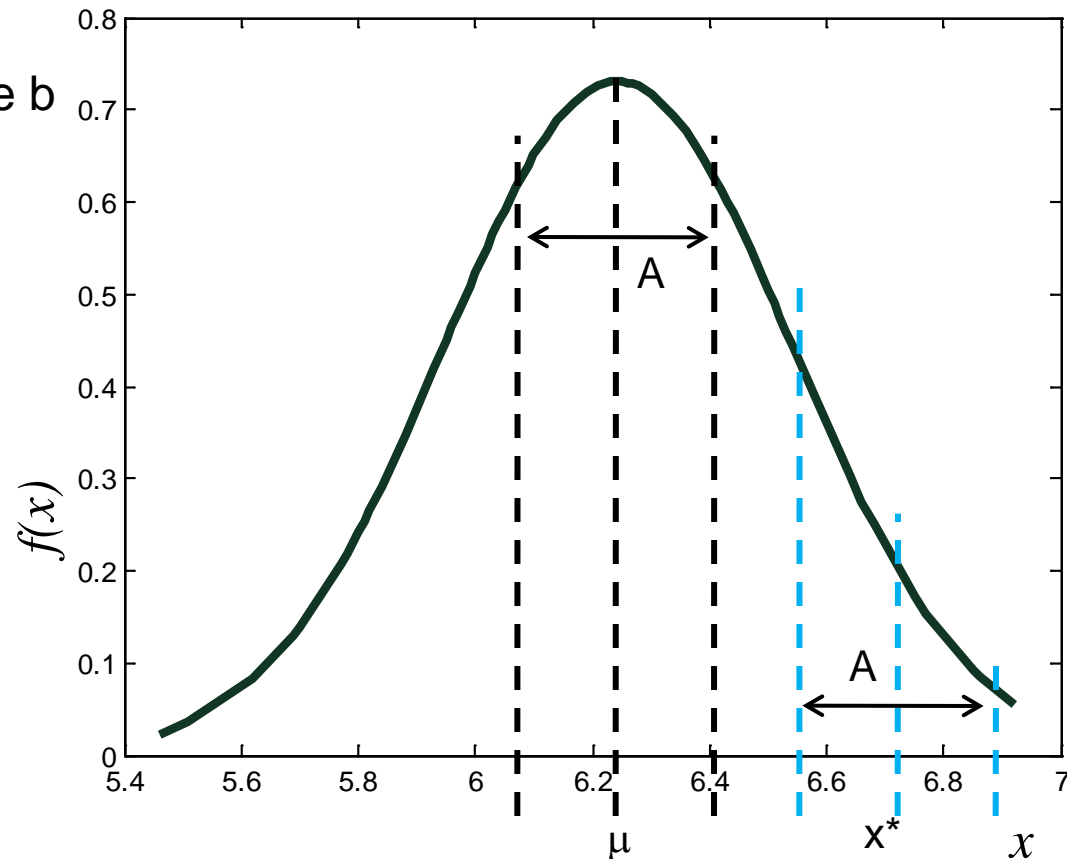
# Elementi di analisi statistica

## Distribuzione di Gauss e probabilità di ottenere un certo valore

Probabilità di avere un valore tra a e b

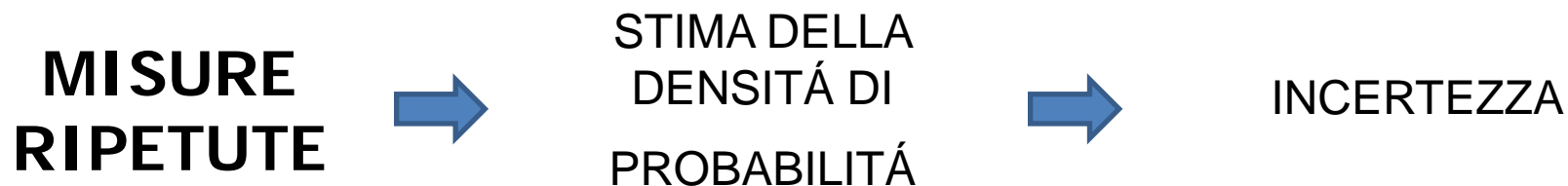
$$\int_a^b f(x)dx \quad \Rightarrow \quad \int_{-\infty}^{\infty} f(x)dx = 1$$

È più probabile avere un valore attorno a  $\mu$  che attorno a  $x^*$



## Elementi di analisi statistica

### Incertezza



IL NUMERO DA ASSOCIARE ALLA MISURA SARA' IL **VALOR MEDIO**  $\mu$  DELLA DISTRIBUZIONE DI DENSITA' DI PROBABILITA' E L'**INCERTEZZA** RELATIVA CHE E' LA DEVIATION STANDARD DELLA DISTRIBUZIONE DI DENSITA' DI PROBABILITA'  $\sigma$ . Se moltiplicata per 3 si parla di **INCERTEZZA ESTESA**.

## Elementi di analisi statistica

### Incertezza tipo di categoria A

Per l'esempio precedente sarà

$$\sigma \text{ stimato con } s = \sqrt{\frac{\sum_{i=1}^n (x_i - m)^2}{n-1}} = 0.29$$

$$\mu \text{ stimato con } m = \frac{\sum_{i=1}^n x_i}{n} = 6.24$$

In realtà l'incertezza deve essere data come incertezza sulla media che rappresenta il valore stimato e non come dispersione dei dati intorno alla media. Si può dimostrare che la deviazione standard della media vale:

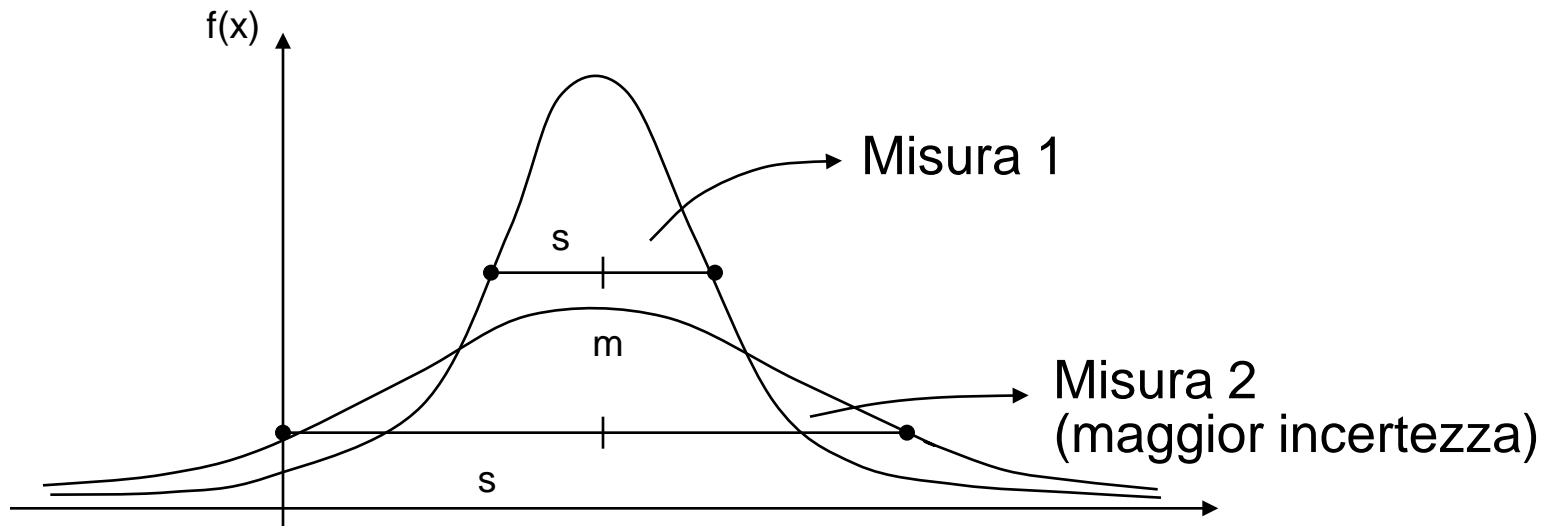
$$\sigma(m) = \frac{\sigma}{\sqrt{n}} = 0.029 \quad (n = 100)$$

Pertanto l'incertezza diminuisce all'aumentare del numero di dati esaminati, ovvero tanto più il campione ( $n$ ) si avvicina all'universo ( $N$ ).

# Elementi di analisi statistica

## Incertezza tipo di categoria A

$$\sigma(m) = \frac{\sigma}{\sqrt{n}}$$



E' possibile ridurre l'incertezza "casuale" riducendo la "dispersione"  $s$  o aumentando il numero di misure  $n$ : ciò non influenza eventuali cause di incertezza di tipo sistematico.

# Elementi di analisi statistica

## Incertezza estesa

Nelle applicazioni commerciali e industriali si preferisce dare l'incertezza come intervallo intorno al risultato della misura nel quale ci si aspetta sia compresa una gran parte della distribuzione dei valori che possono essere assunti dal misurando.

Pertanto essa vale  $\Sigma = k \cdot \sigma(m) = 0.087 \quad (k = 3)$

Con  $k$  fattore di copertura che dipende dal grado di confidenza della distribuzione. Come si è visto, per una distribuzione gaussiana

- il grado di confidenza del 68.3% vale  $k=1$ ,
- il grado di confidenza del 95.5% vale  $k=2$ ,
- il grado di confidenza del 99.7% vale  $k=3$

**IL VALORE MISURATO SARA'  $6.24 \pm 0.087$  mm**

## Elementi di analisi statistica

### **Incertezza tipo di categoria B se non è possibile ripetere la misura**

L'incertezza può essere stimata in base a tutte le informazioni disponibili:

- misure precedenti
- esperienza o conoscenze generali delle caratteristiche e proprietà di materiali e strumenti
- specifiche tecniche del costruttore
- dati di certificati di taratura
- dati di manuali

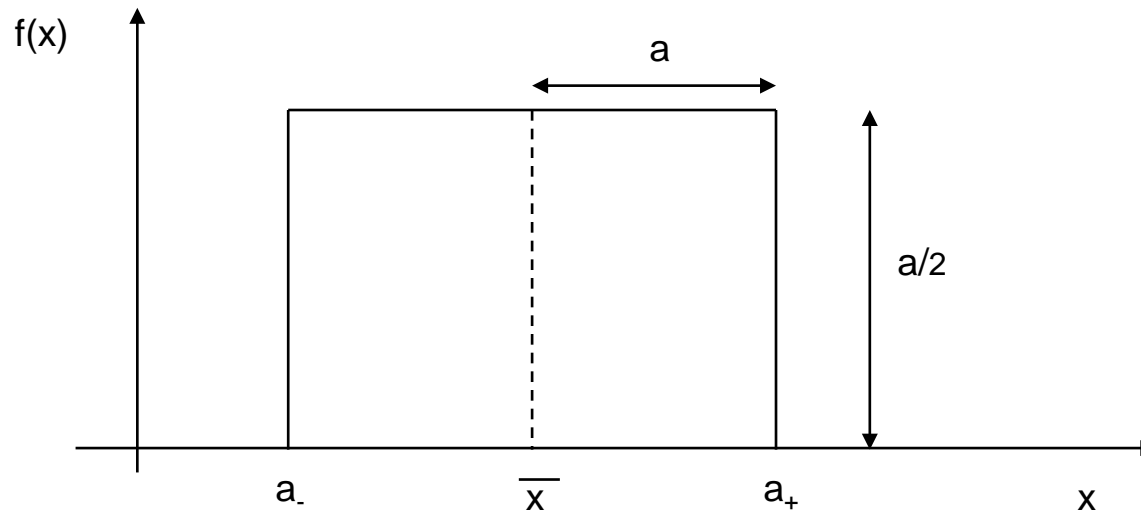


stima incertezza tipo di categoria B

In questi casi la distribuzione di ampiezza si assume "a priori"

# Elementi di analisi statistica

## Incertezza tipo di categoria B



$$s^2 = \frac{a_+ - a_-}{12}$$

$$s = \frac{a}{\sqrt{3}}$$

Es: distribuzione rettangolare:  
tutti i valori tra  $a_-$  e  $a_+$  sono  
ugualmente probabili

$s$  ➡ Incertezza tipo di categoria B





# **Sessione di studio 1**

## Sessione di studio 1

**Approfondimenti: funzione di densità di distribuzione normale standard e distribuzione di probabilità normale**

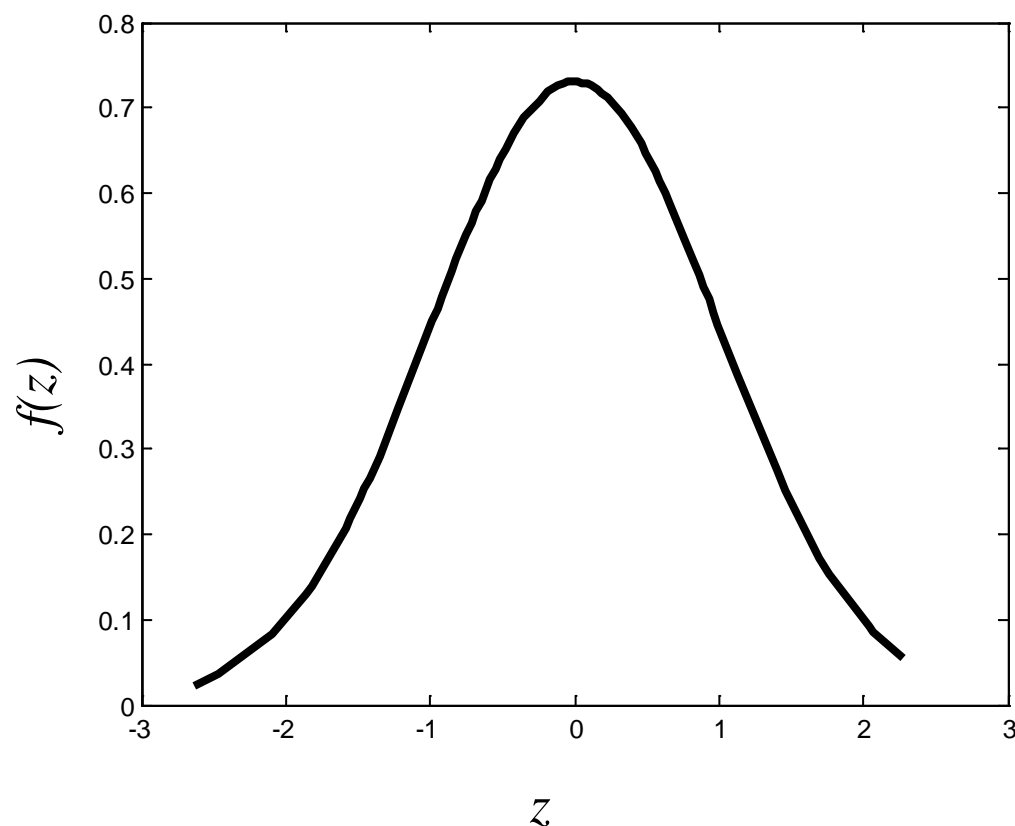
## Funzione di densità di distribuzione normale standard e teorema del limite centrale

Il teorema limite centrale afferma che la distribuzione gaussiana standard permette di descrivere in maniera soddisfacente tutti quei fenomeni fisici caratterizzati dalla sovrapposizione di un elevato numero di effetti deboli indipendenti aventi loro natura statistica a media nulla. Pertanto la distribuzione di Gauss standard è espressa in funzione della variabile  $z$  o scarto ridotto che permette di avere una funzione centrata su un valor medio nullo.

$$z = \frac{x - \mu}{\sigma}$$

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

# Funzione di densità di distribuzione normale standard



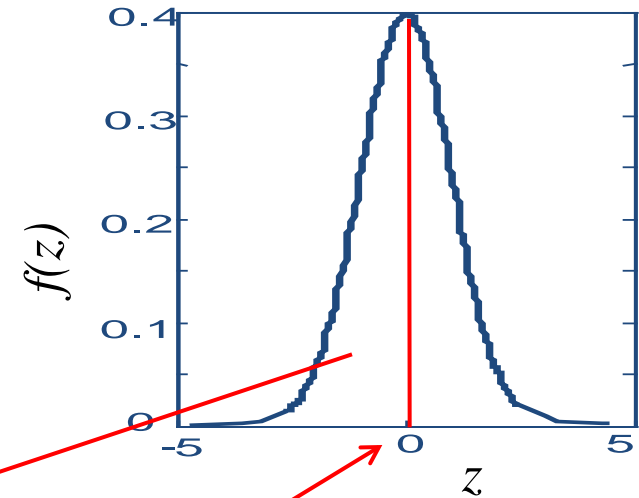
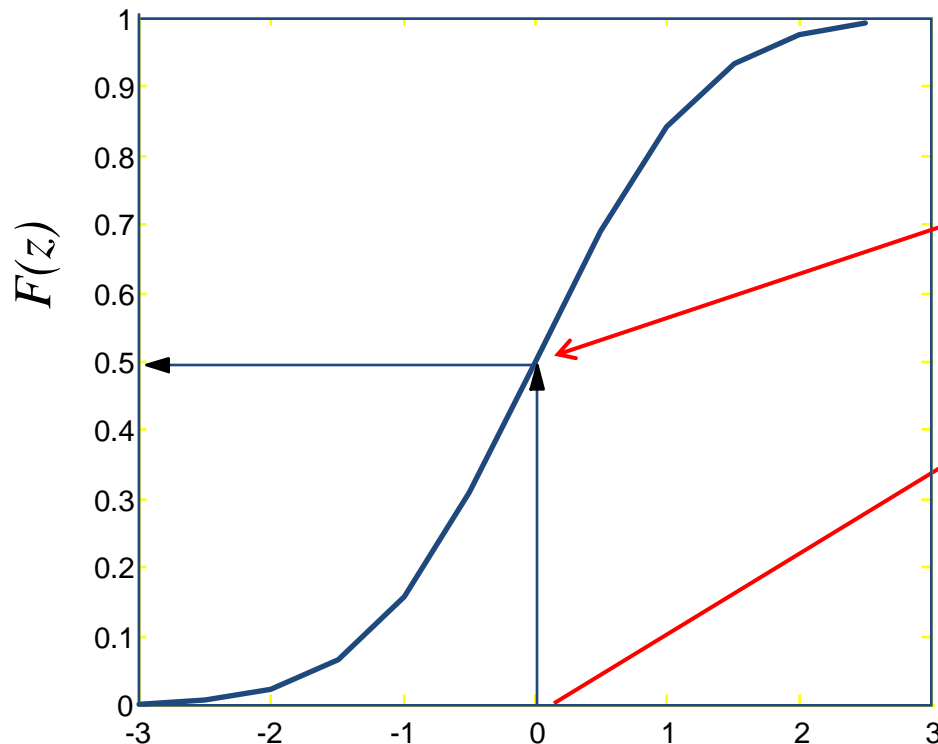
$$f(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z^2}{2}}$$

La funzione di densità di distribuzione gaussiana standard ha un massimo in corrispondenza di  $z=0$  che vale  $1/\sigma$ . Per rendere adimensionale la densità di probabilità si considera, perciò il prodotto:  $\sigma f(z)$

## Frequenza cumulata

$$F(z) = \int_{-\infty}^z f(z) dz$$

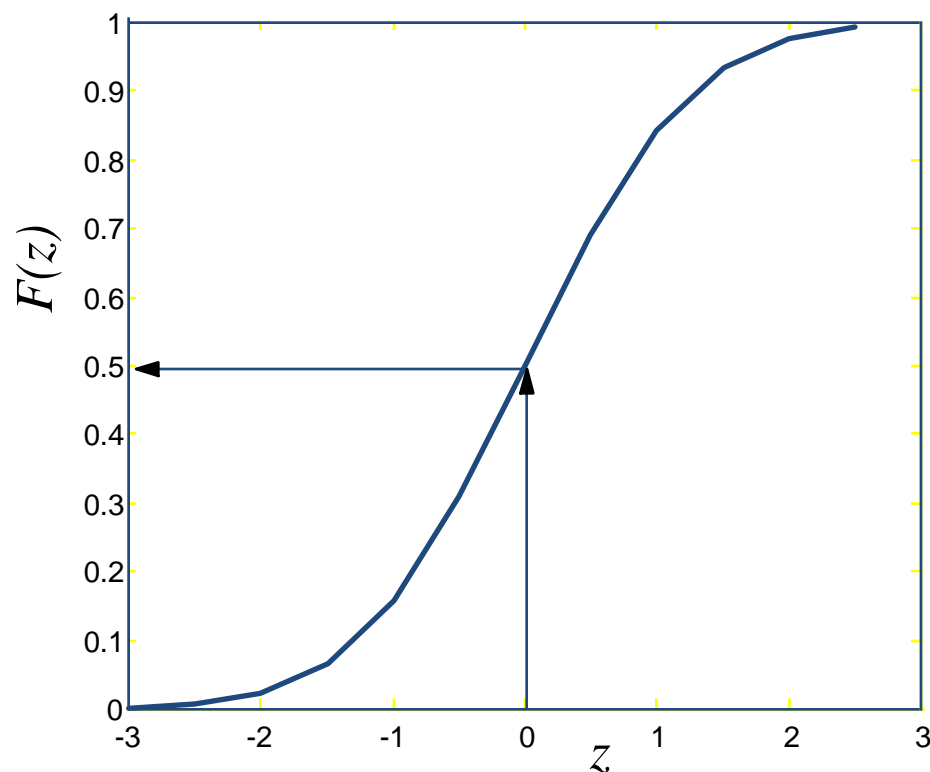
$$F(z) = P(0 < z_i \leq z)$$



La funzione  $F(z)$  è la probabilità di verificarsi di un valore  $z$  compreso tra  $-\infty$  e  $z$ . E' adimensionale, ha un flesso che vale 0.5 a 0, tende a 0 per  $z=0$  e a 1 per  $z \rightarrow \infty$ .

## Frequenza cumulata - distribuzione di probabilità normale

Il grafico della funzione  $F(z)$  è anche detto grafico della distribuzione di probabilità normale. Tale grafico viene utilizzato per valutare qualitativamente se la distribuzione dei dati in esame può essere considerata normale ovvero può essere rappresentata dalla funzione di Gauss.

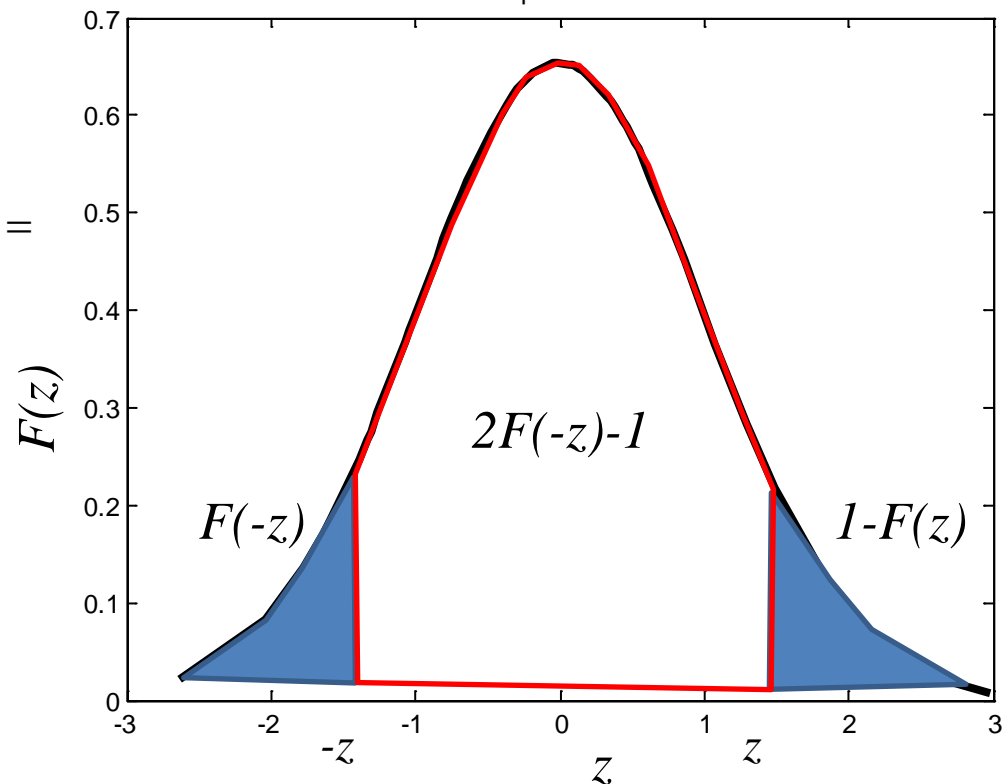


# Frequenza cumulata - distribuzione di probabilità normale

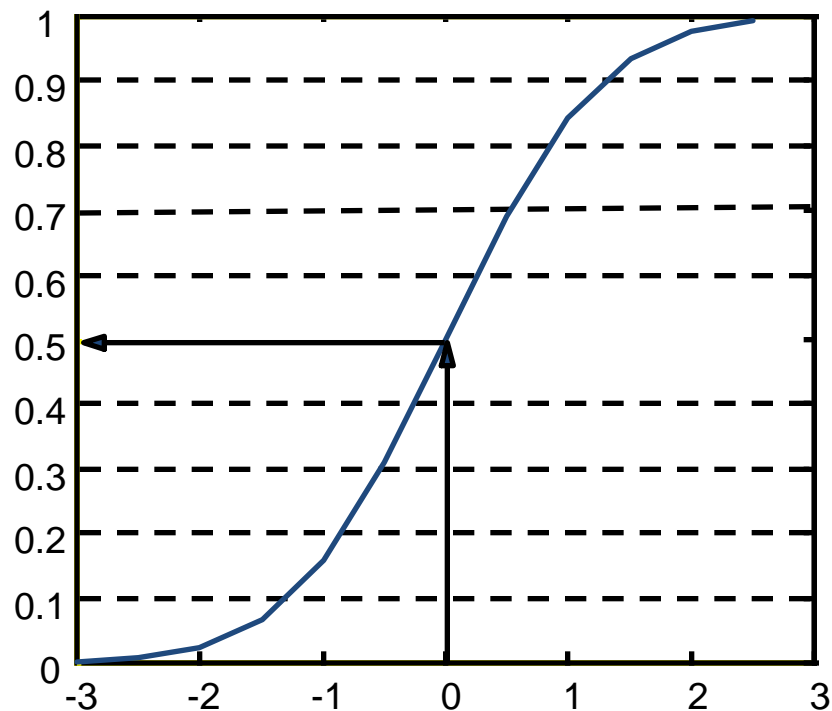
Dalle caratteristiche di simmetria della curva  $\sigma f(z)$  rispetto all'asse delle ordinate, e all'antisimmetria della curva  $F(z)$  risulta che  $F(-z) = 1 - F(z)$

Pertanto, la probabilità che un dato sia compreso fra due valori  $-z_1$  e  $z_1$  simmetrici rispetto  $z=0$  è:

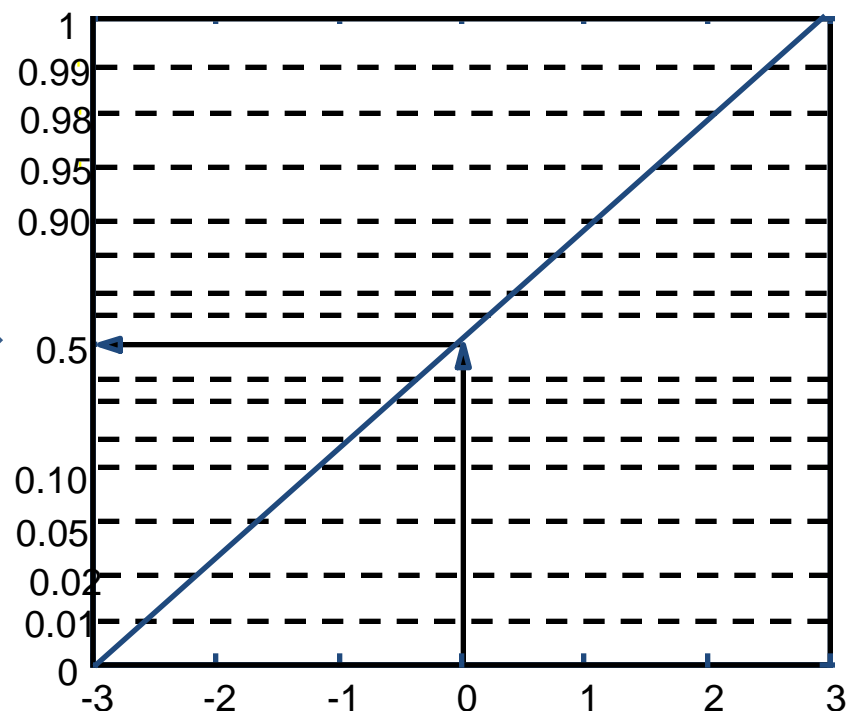
$$\begin{aligned} P_{-z_1, z_1} &= P(-z_1 < z < z_1) = F(z_1) - F(-z_1) = \\ &= F(z_1) - 1 + F(-z_1) = 2F(z) - 1 \end{aligned}$$



## Grafico della distribuzione di Probabilità Normale (GPN)



Operando un cambiamento di scale delle ordinate, che vengono date in forma logaritmica

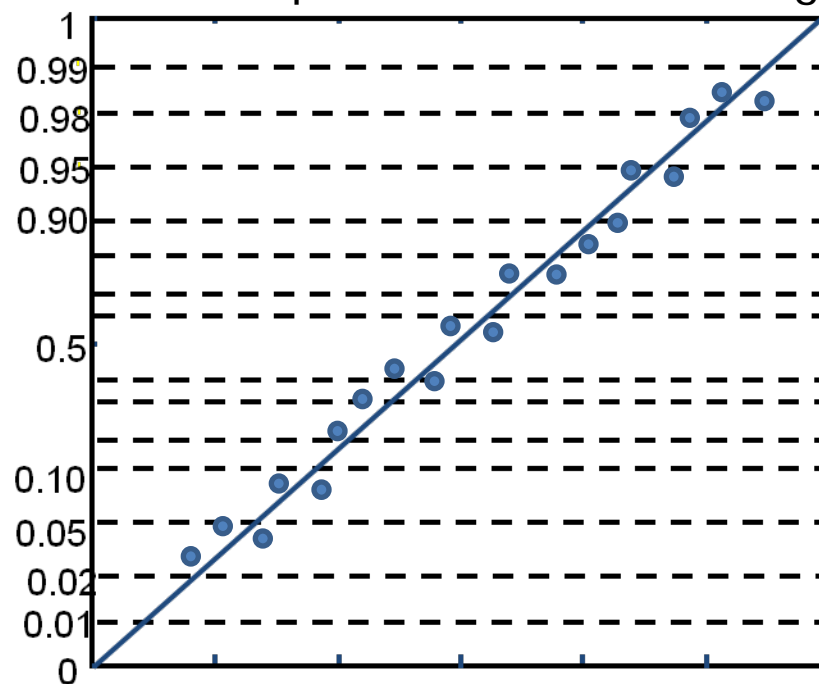


Il grafico di probabilità normale diventa una retta



## Grafico della distribuzione di Probabilità Normale (GPN) per la verifica di aderenza alla distribuzione gaussiana (I metodo)

Raggruppando i dati in esame in classi e ordinandole in ordine crescente, si riportano in ordinata, in riferimento a ciascun valore di classe, le corrispondenti frequenze cumulate. Se la curva congiungente i valori trovati (cerchi blu) può essere approssimata da una retta la distribuzione può essere considerata gaussiana.

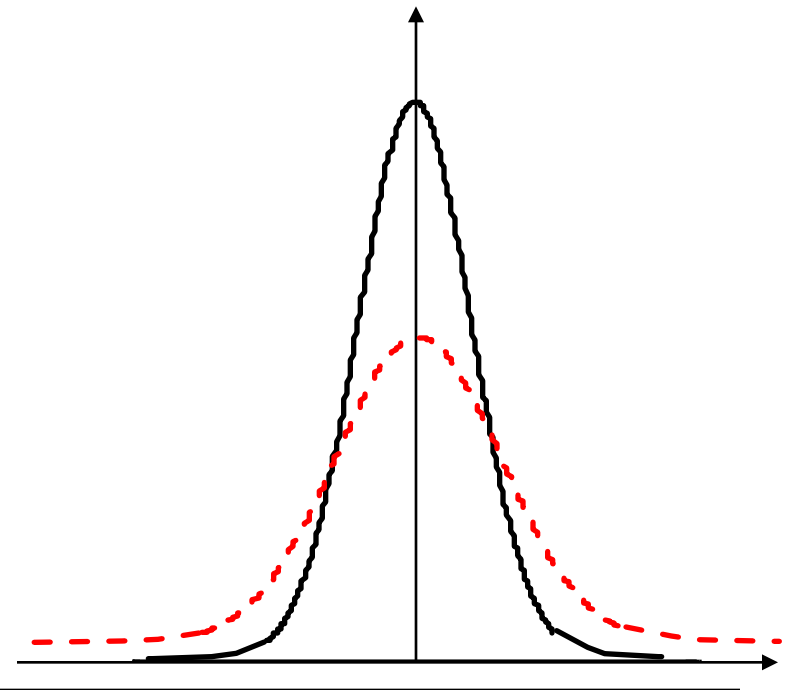
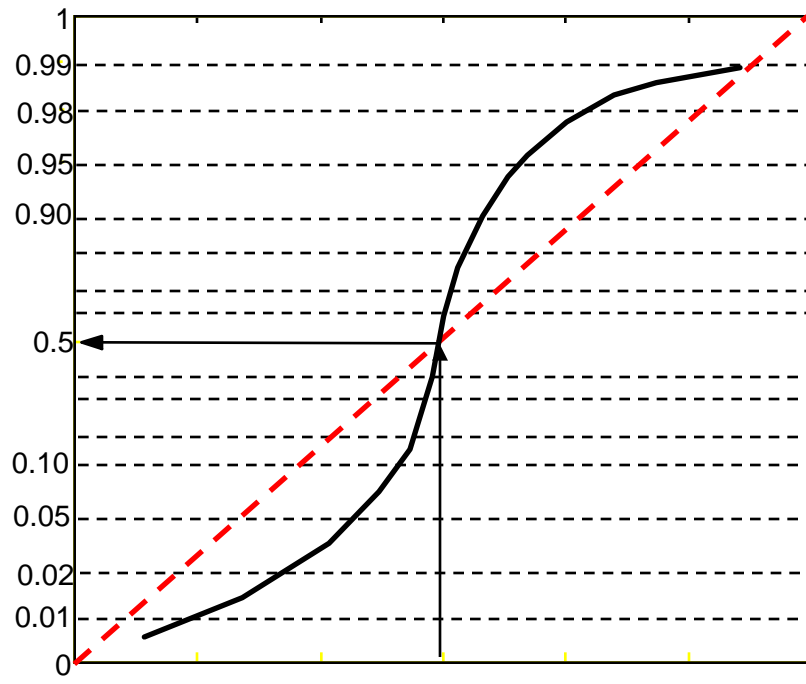


# Grafico della distribuzione di Probabilità Normale (GPN)

Il GPN è utile per capire quali sono le cause di non normalità di una distribuzione:

## Distribuzione iper-normale

Caratterizzata da una curva a 'S': è dovuta ad un'eccessiva concentrazione di dati vicino al valor medio.

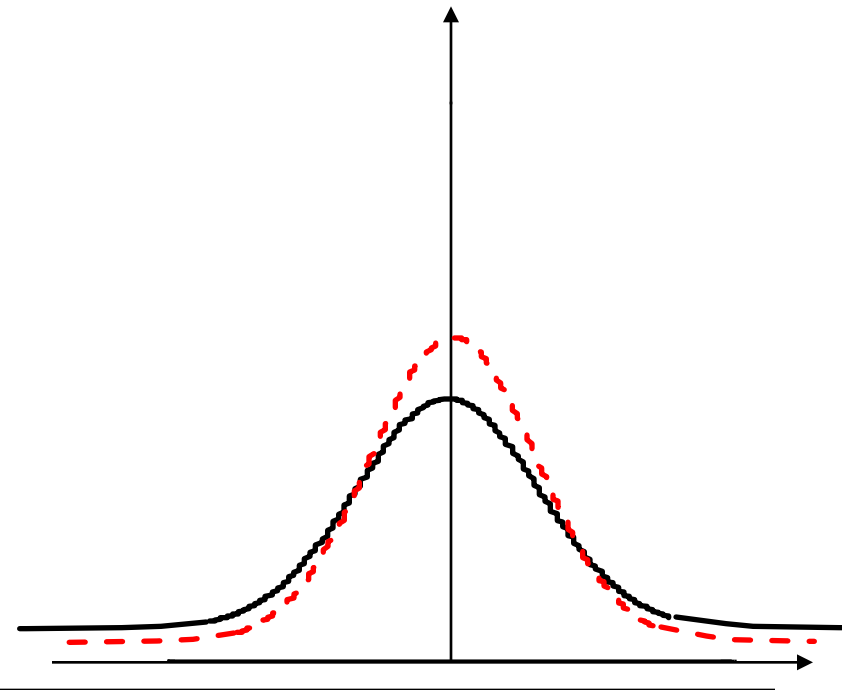
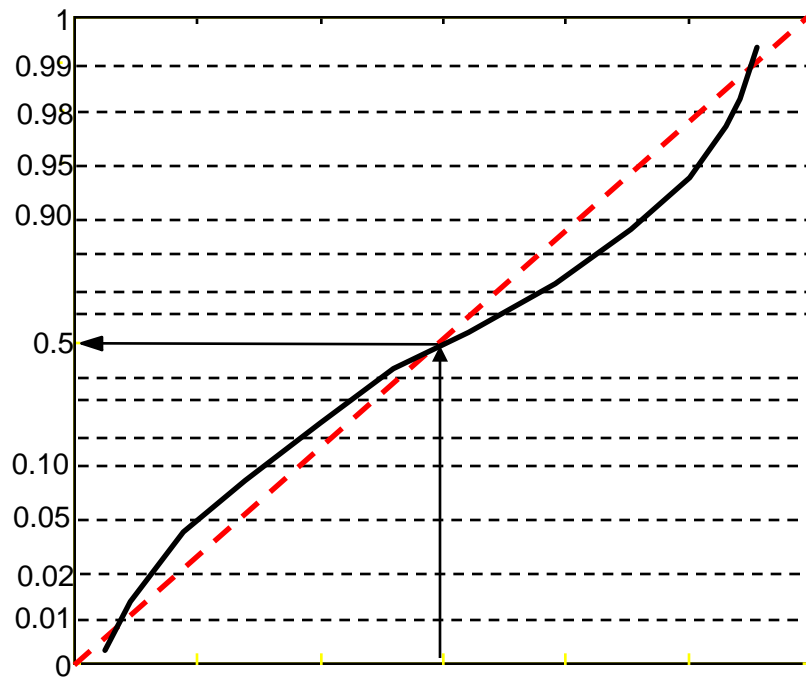


# Grafico della distribuzione di Probabilità Normale (GPN)

Il GPN è utile per capire quali sono le cause di non normalità di una distribuzione:

## Distribuzione ipo-normale

Caratterizzata da una curva a 'Z': è causata da presenza di disturbi nei dati acquisiti.

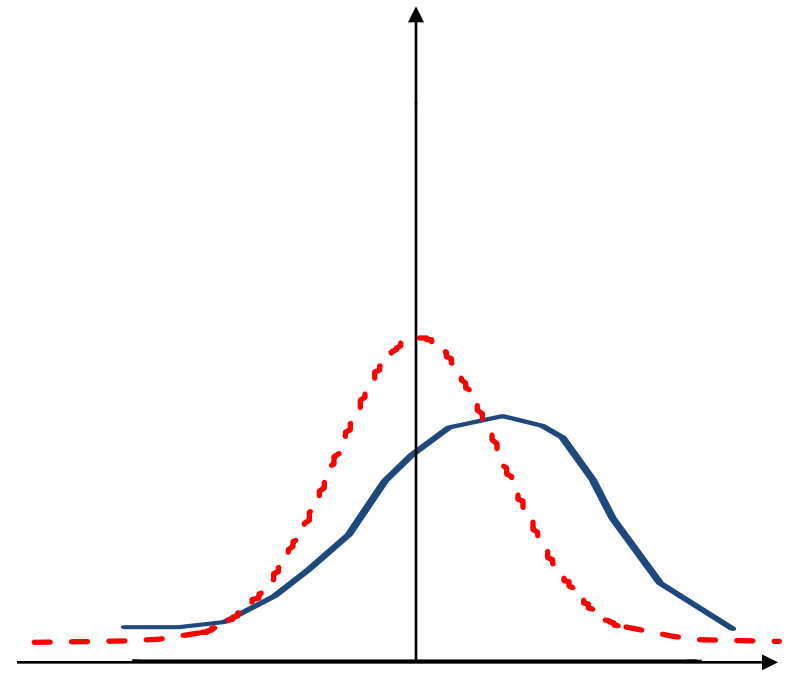
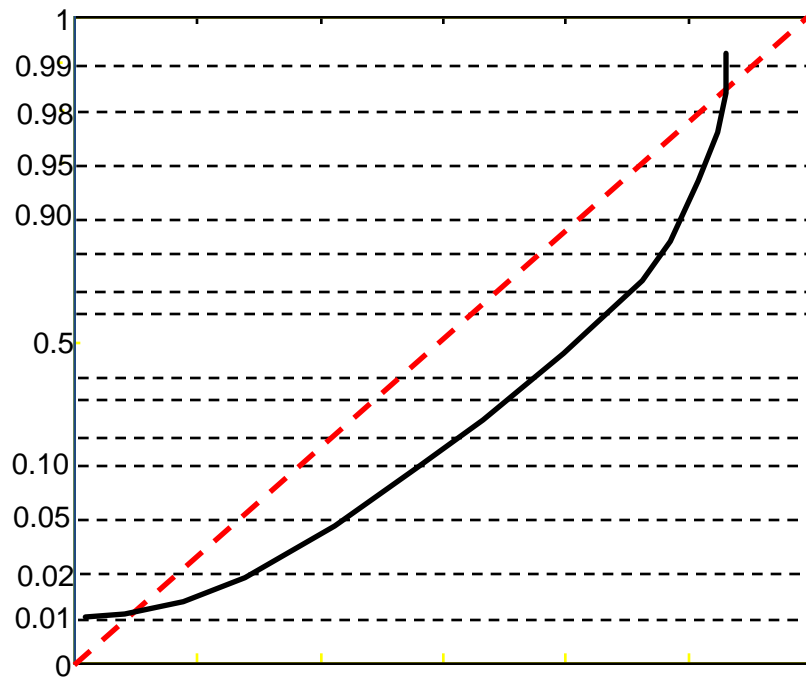


# Grafico della distribuzione di Probabilità Normale (GPN)

Il GPN è utile per capire quali sono le cause di non normalità di una distribuzione:

## Distribuzione asimmetrica

Caratterizzata da una asimmetria rispetto al valor medio.

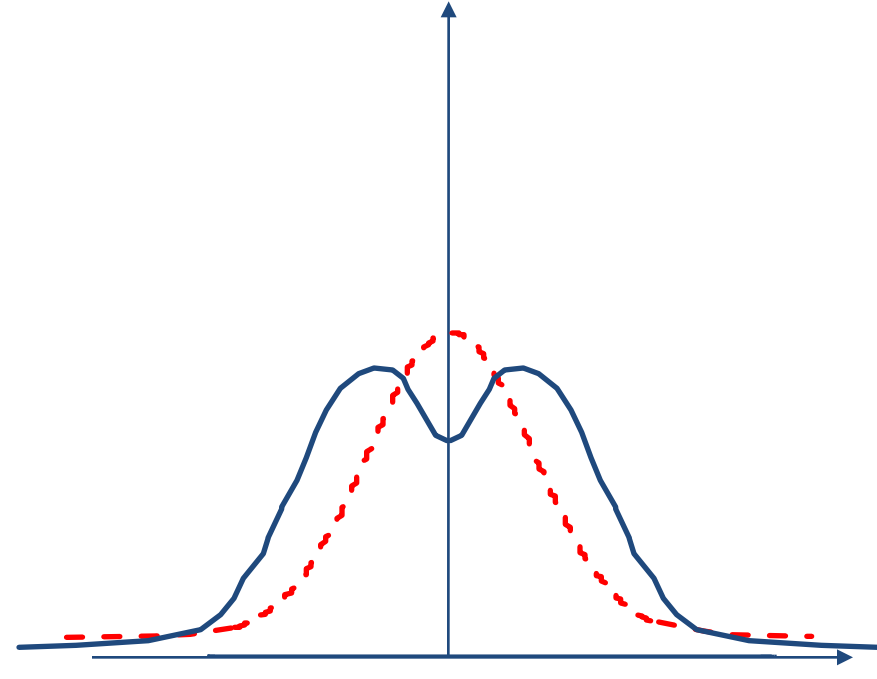
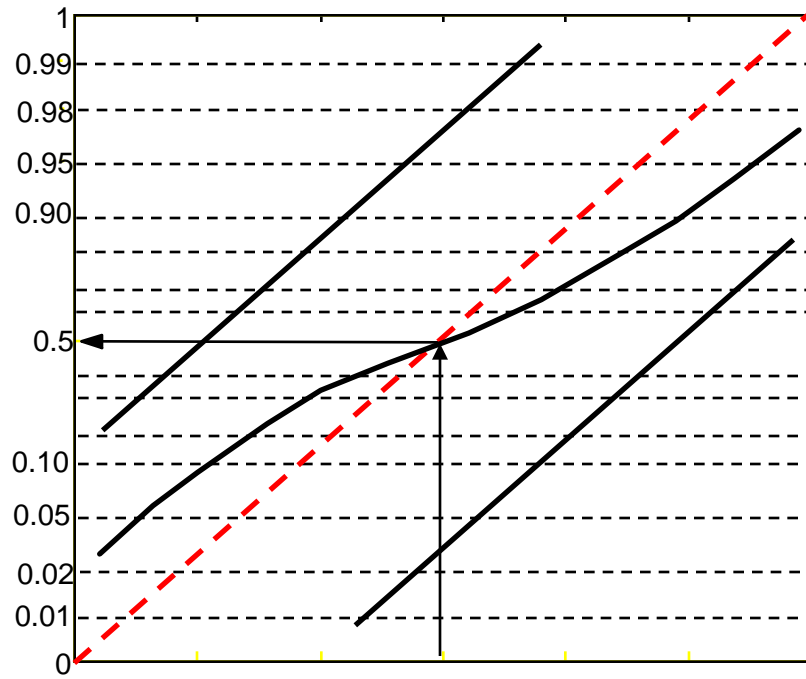


# Grafico della distribuzione di Probabilità Normale (GPN)

Il GPN è utile per capire quali sono le cause di non normalità di una distribuzione:

## Distribuzione bi-modale

Caratterizzata da una curva a 'S' ma è data dalla composizione di due rette: è causata dalla combinazione di due o più distribuzioni gaussiane di media diversa.





# **Sessione di studio 2**

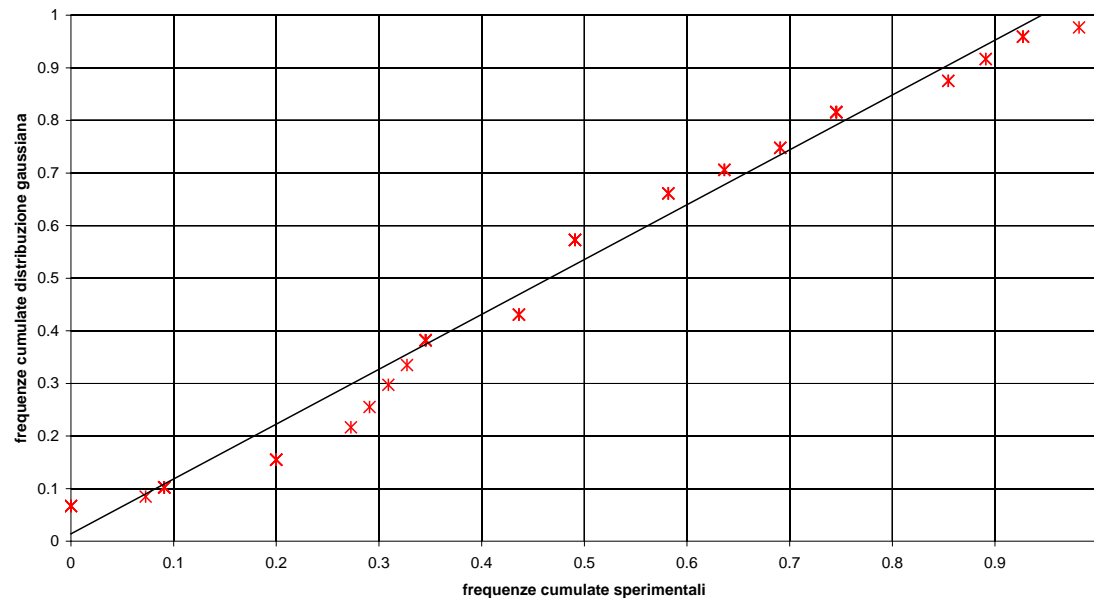
## Sessione di studio 2

**Il metodo per la verifica di “aderenza” di una distribuzione a quella normale**

## Il metodo per la verifica di “aderenza” di una distribuzione a quella normale

- In **ascissa** riportiamo la probabilità cumulata dei valori sperimentali normalizzati = (numero di valori  $< z_i$ )/N
- In **ordinata** riportiamo la probabilità cumulata della gaussiana  $F(z_i)$ , dedotta dalle tabelle o calcolata dalla formula (in EXCEL = **DISTRIB.NORM.ST()** )

Se la distribuzione è pressoché normale i dati di disporranno lungo una retta (saranno linearmente correlati).





## Analisi dei dati: il criterio di Chauvenet

Nel campo dell'analisi sperimentale è frequente trovare, in una serie di misure, qualche dato che non concorda con gli altri.

In una serie di  $n$  dati sperimentali, se alcuni valori presentano uno scostamento dal valore medio che ha probabilità di verificarsi inferiore di  $1/(2n)$ , allora quei valori devono essere scartati.

## Analisi dei dati: il criterio di Chauvenet

1. Si calcola la media, varianza e lo scarto ridotto dei dati  $z_i = \frac{x_i - \mu}{\sigma}$
2. Si calcola la probabilità di Chauvenet  $p_c = 1 - \frac{1}{2n}$
3. Si calcola la frequenza cumulata, ricordando che  $p_c = 2F(z) - 1$ . Per cui

$$F(z_{\text{lim}}) = \frac{p_c + 1}{2}$$

4. Si determina il valore dello scarto ridotto corrispondente al valore limite  $z_{\text{max}}$
5. Si escludono i valori per i quali il valore assoluto dello scarto ridotto è superiore a  $z_{\text{max}}$
6. Si ricalcola la media e lo scarto quadratico con i dati restanti, ma non è più lecito applicare di nuovo il criterio ai dati rimasti.



# **Sessione di studio 3**

## **Sessione di studio 3**

### **Esercizio:**

**Analizzare i dati di temperatura misurati da un termometro digitale durante una campagna di misura per la rilevazione di un campione statistico di 100 elementi.**

**Stimare la misura effettuata come valor medio del campione e deviazione standard della media. Identificare pertanto l'incertezza estesa associata alla misura.**

**Aiutarsi con il foglio di 432.xls e 433.xls (per determinare la frequenza di appartenenza alle classi)**



## Lecture dal termometro digitale (°C)

22.38	22.42	22.18	21.86	22.66
22.80	21.82	22.13	22.22	22.30
21.49	22.44	22.31	22.39	22.27
22.49	22.73	22.31	22.56	22.72
22.31	22.37	21.93	22.70	21.95
21.79	22.54	22.20	22.24	22.43
22.07	22.44	22.16	21.73	22.48
22.32	22.11	22.41	21.97	22.13
23.36	22.30	22.56	21.87	22.28
23.10	21.96	22.56	22.96	21.84
21.78	22.49	21.93	22.01	21.84
23.18	21.84	22.23	22.45	22.24
22.44	21.87	21.82	22.15	22.44
22.19	21.95	21.85	22.49	23.04
22.44	21.27	22.21	21.97	22.00
22.14	22.67	22.70	21.76	22.27
22.17	22.31	21.96	21.75	22.18
22.69	21.97	22.33	22.37	21.59
22.66	22.65	22.14	22.15	22.07
22.66	21.66	22.57	22.15	21.64

dati (mm)	21,20	21,30	21,40	21,50	21,60	21,70	21,80
22,38	0	0	0	0	0	0	0
22,80	0	0	0	0	0	0	0
21,49	0	0	0	1	0	0	0
22,49	0	0	0	0	0	0	0
22,31	0	0	0	0	0	0	0
21,79	0	0	0	0	0	0	1
22,07	0	0	0	0	0	0	0
22,32	0	0	0	0	0	0	0
23,36	0	0	0	0	0	0	0
23,10	0	0	0	0	0	0	0
21,78	0	0	0	0	0	0	1
23,18	0	0	0	0	0	0	0
22,44	0	0	0	0	0	0	0
22,19	0	0	0	0	0	0	0
22,44	0	0	0	0	0	0	0
22,14	0	0	0	0	0	0	0
22,17	0	0	0	0	0	0	0
22,69	0	0	0	0	0	0	0
22,66	0	0	0	0	0	0	0
22,66	0	0	0	0	0	0	0
22,42	0	0	0	0	0	0	0
21,82	0	0	0	0	0	0	0
22,44	0	0	0	0	0	0	0
22,73	0	0	0	0	0	0	0
22,37	0	0	0	0	0	0	0
22,54	0	0	0	0	0	0	0
22,44	0	0	0	0	0	0	0
22,11	0	0	0	0	0	0	0
22,30	0	0	0	0	0	0	0
21,96	0	0	0	0	0	0	0
22,49	0	0	0	0	0	0	0
21,84	0	0	0	0	0	0	0
21,87	0	0	0	0	0	0	0
21,95	0	0	0	0	0	0	0
21,27	0	1	0	0	0	0	0
22,67	0	0	0	0	0	0	0
22,31	0	0	0	0	0	0	0
21,97	0	0	0	0	0	0	0
22,65	0	0	0	0	0	0	0
21,66	0	0	0	0	0	1	0
22,18	0	0	0	0	0	0	0
22,13	0	0	0	0	0	0	0
22,31	0	0	0	0	0	0	0
22,31	0	0	0	0	0	0	0
21,93	0	0	0	0	0	0	0
22,20	0	0	0	0	0	0	0
22,16	0	0	0	0	0	0	0
22,41	0	0	0	0	0	0	0
22,56	0	0	0	0	0	0	0

22,56	0	0	0	0	0	0	0
21,93	0	0	0	0	0	0	0
22,23	0	0	0	0	0	0	0
21,82	0	0	0	0	0	0	0
21,85	0	0	0	0	0	0	0
22,21	0	0	0	0	0	0	0
22,70	0	0	0	0	0	0	0
21,96	0	0	0	0	0	0	0
22,33	0	0	0	0	0	0	0
22,14	0	0	0	0	0	0	0
22,57	0	0	0	0	0	0	0
21,86	0	0	0	0	0	0	0
22,22	0	0	0	0	0	0	0
22,39	0	0	0	0	0	0	0
22,56	0	0	0	0	0	0	0
22,70	0	0	0	0	0	0	0
22,24	0	0	0	0	0	0	0
21,73	0	0	0	0	0	0	1
21,97	0	0	0	0	0	0	0
21,87	0	0	0	0	0	0	0
22,96	0	0	0	0	0	0	0
22,01	0	0	0	0	0	0	0
22,45	0	0	0	0	0	0	0
22,15	0	0	0	0	0	0	0
22,49	0	0	0	0	0	0	0
21,97	0	0	0	0	0	0	0
21,76	0	0	0	0	0	0	1
21,75	0	0	0	0	0	0	1
22,37	0	0	0	0	0	0	0
22,15	0	0	0	0	0	0	0
22,15	0	0	0	0	0	0	0
22,66	0	0	0	0	0	0	0
22,30	0	0	0	0	0	0	0
22,27	0	0	0	0	0	0	0
22,72	0	0	0	0	0	0	0
21,95	0	0	0	0	0	0	0
22,43	0	0	0	0	0	0	0
22,48	0	0	0	0	0	0	0
22,13	0	0	0	0	0	0	0
22,28	0	0	0	0	0	0	0
21,84	0	0	0	0	0	0	0
21,84	0	0	0	0	0	0	0
22,24	0	0	0	0	0	0	0
22,44	0	0	0	0	0	0	0
23,04	0	0	0	0	0	0	0
22,00	0	0	0	0	0	0	0
22,27	0	0	0	0	0	0	0
22,18	0	0	0	0	0	0	0
21,59	0	0	0	0	1	0	0
22,07	0	0	0	0	0	0	0

21,64	0	0	0	0	0	1	0
frequenza assoluta	0	1	0	1	1	2	5





0	0	0	0	0	0	0	1	0
0	1	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	1	0
1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	1	0	0
0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	1	0	0
0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1
0	0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0				

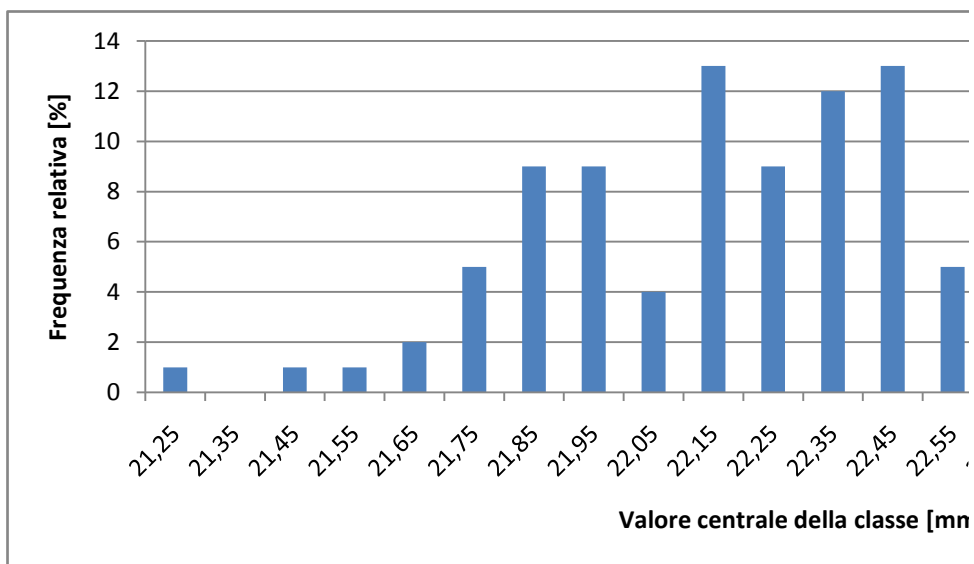
0	0	0	0	0	0	0	0	0
9	9	4	13	9	12	13	5	6

[illegible]

[illegible]

0	0	0	0	0	0	0	
4	1	1	1	2	0	1	100

dati (mm)	n dati	dx	classe min	classe max	valore centrale
22,38	100	0,1000	21,2	21,3	21,25
22,80			21,3	21,4	21,35
21,49	23,3600	23,40	21,4	21,5	21,45
22,49	21,2700	21,20	21,5	21,6	21,55
22,31			21,6	21,7	21,65
21,79	22,2488		21,7	21,8	21,75
22,07	0,3722	0,111670196	21,8	21,9	21,85
22,32	0,037223		21,9	22	21,95
23,36	m	22,2488	22	22,1	22,05
23,10	s	0,01116702	22,1	22,2	22,15
21,78	campione rappresentabile da distribuzione gaussiana anche se leggermente asimmetrica (v. GPN)		22,2	22,3	22,25
23,18			22,3	22,4	22,35
22,44			22,4	22,5	22,45
22,19			22,5	22,6	22,55
22,44			22,6	22,7	22,65
22,14			22,7	22,8	22,75
22,17			22,8	22,9	22,85
22,69			22,9	23	22,95
22,66			23	23,1	23,05
22,66			23,1	23,2	23,15
22,42			23,2	23,3	23,25
21,82			23,3	23,4	23,35



22,73  
22,37  
22,54  
22,44  
22,11  
22,30  
21,96  
22,49  
21,84  
21,87  
21,95  
21,27  
22,67  
22,31  
21,97  
22,65  
21,66  
22,18  
22,13  
22,31  
22,31  
21,93  
22,20  
22,16  
22,41

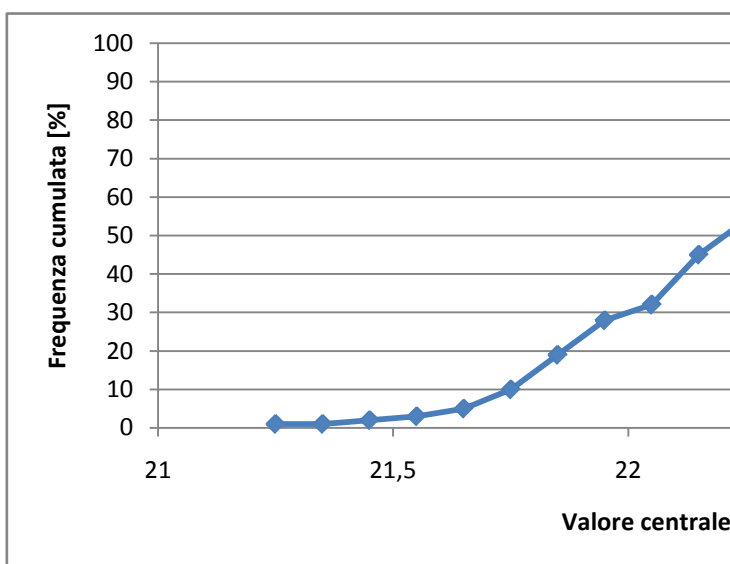
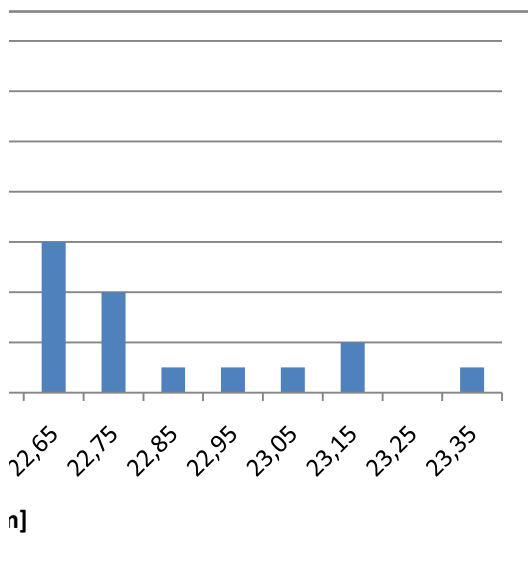
22,56  
22,56  
21,93  
22,23  
21,82  
21,85  
22,21  
22,70  
21,96  
22,33  
22,14  
22,57  
21,86  
22,22  
22,39  
22,56  
22,70  
22,24  
21,73  
21,97  
21,87  
22,96  
22,01  
22,45  
22,15  
22,49  
21,97  
21,76  
21,75  
22,37  
22,15  
22,15  
22,66  
22,30  
22,27  
22,72  
21,95  
22,43  
22,48  
22,13  
22,28  
21,84  
21,84  
22,24  
22,44  
23,04  
22,00  
22,27  
22,18  
21,59



22,07

21,64

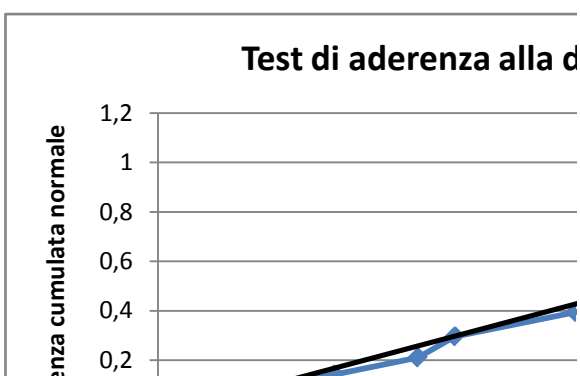
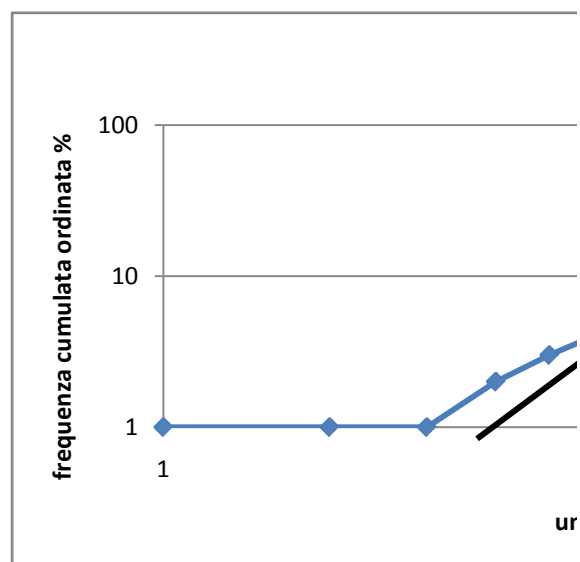
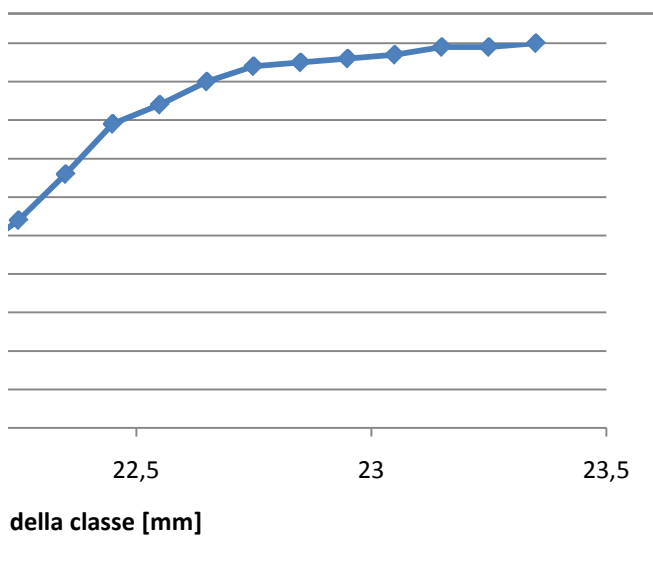
frequenza assoluta	frequenza %	fp/dx	frequenza % cumulata	scarto ridotto	Frequenza cumulata normale
1	1	1	10	1 -2,68326	0,003645431
0	0	0	0	1 -2,41461	0,007876028
1	1	1	10	2 -2,14596	0,015938004
1	1	1	10	3 -1,87731	0,030237549
2	2	2	20	5 -1,60867	0,053844744
5	5	5	50	10 -1,34002	0,090119853
9	9	9	90	19 -1,07137	0,142001746
9	9	9	90	28 -0,80272	0,211068036
4	4	4	40	32 -0,53407	0,296645628
13	13	13	130	45 -0,26542	0,395341253
9	9	9	90	54 0,003224	0,501286099
12	12	12	120	66 0,271872	0,607139781
13	13	13	130	79 0,54052	0,705580843
5	5	5	50	84 0,809168	0,790790871
6	6	6	60	90 1,077817	0,859442213
4	4	4	40	94 1,346465	0,910923684
1	1	1	10	95 1,615113	0,946856901
1	1	1	10	96 1,883761	0,970201368
1	1	1	10	97 2,15241	0,984317445
2	2	2	20	99 2,421058	0,992262292
0	0	0	0	99 2,689706	0,996424251
1	1	1	10	100 2,958354	0,998453568
			0		

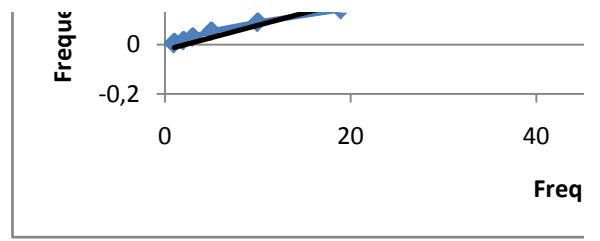






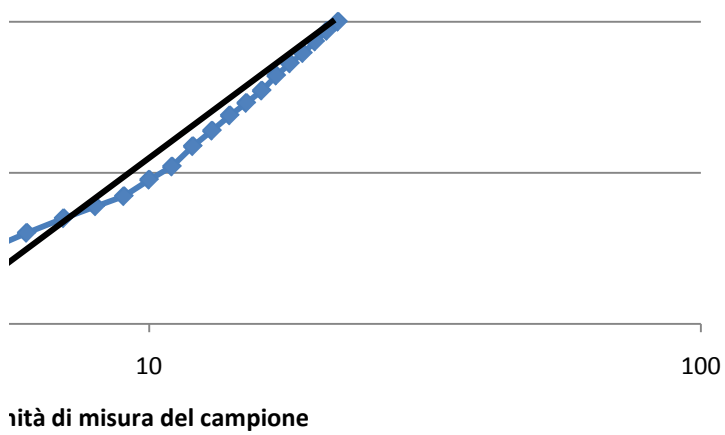
frequenza % ordinata	frequenza cumolata ordinata
0	0
0	0
1	1
1	2
1	3
1	4
1	5
1	6
1	7
2	9
2	11
4	15
4	19
5	24
5	29
6	35
9	44
9	53
9	62
12	74
13	87
13	100







## GPN



## Distribuzione normale (II metodo)

