



CORSO DI LAUREA MAGISTRALE IN INGEGNERIA GESTIONALE

Gestione della Produzione e della Qualità

Istogramma

Prof. Antonella Petrillo



Come facciamo a leggere ed interpretare i dati raccolti?





L'ISTOGRAMMA

Definizione

L'*istogramma* è un metodo grafico per rappresentare la *dispersione statistica* di un insieme di dati.

E' la "*fotografia*" di come i valori reali di una variabile misurata si distribuiscono intorno ad un *valore nominale*.

ORIGINE STORICA

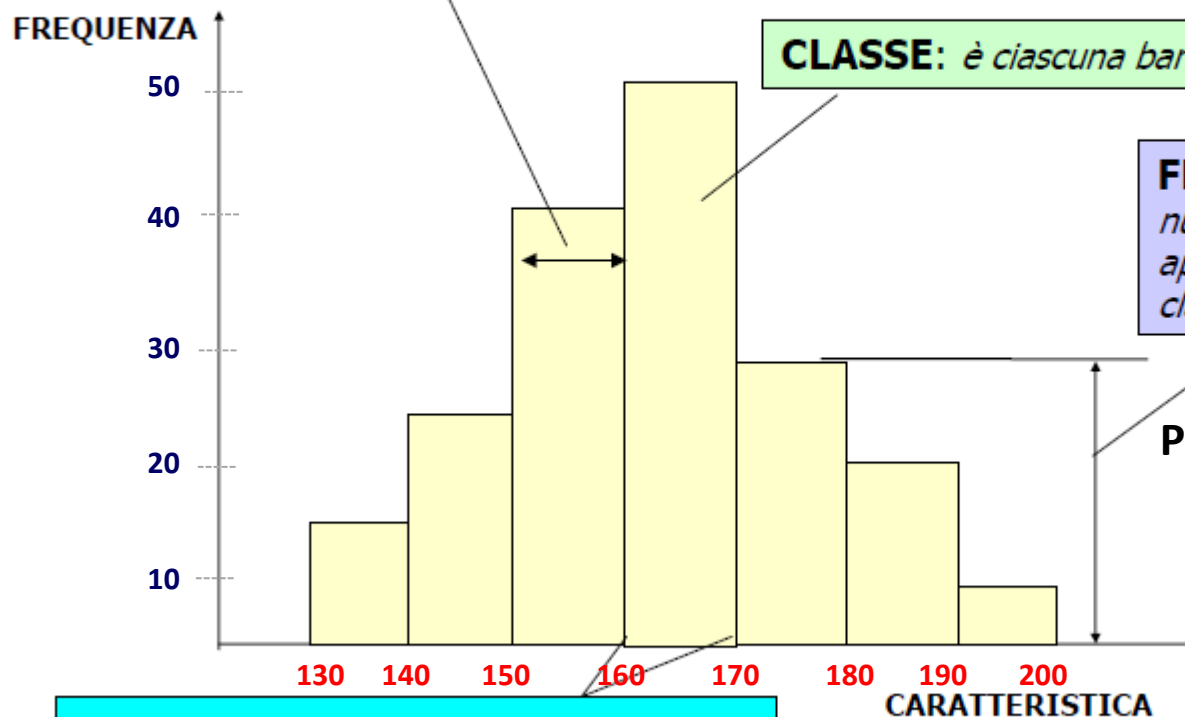
Nel **1833** *A.M. Guerry* presentò un diagramma che mostrava sull'asse verticale il *numero di crimini commessi*, sull'asse orizzontale il *numero di crimini raggruppati per fasce di età (0-5; 6-10; 11-15; 16-20; ecc.)*

Istogramma



Nella figura la dimensione delle **classi è 10** e si possono contare **7 intervalli**

AMPIEZZA DELLA CLASSE:
intervallo compreso tra il valore massimo ed il valore minimo della classe



CLASSE: è ciascuna barra

FREQUENZA:
numero di osservazioni appartenenti a ciascuna classe

Per es. nella classe da 161 a 170 sono compresi 50 elementi

LIMITI DELLA CLASSE: *valori minimi e massimi di ogni classe*

DISPERSIONE o ESCURSIONE: dimensione dell'intervallo tra il massimo dei valori e quello minimo. Nella figura la dispersione va da **130 a 200** e quindi vale **70**.



FREQUENZA, FREQUENZA RELATIVA E FREQUENZA CUMULATA

FREQUENZA = numero di volte che si è verificato un evento

$$\text{FREQUENZA RELATIVA} = f_i = \frac{N_i}{N_T}$$

Numero di volte che si è verificato un evento

Numero totale di eventi

$$\text{FREQUENZA CUMULATA} = F_{i^*} = \sum_{i=1}^{i^*} f(x_i)$$

Somma progressiva delle frequenze relative

COME COSTRUIRE UN ISTOGRAMMA

ESEMPIO 1.

Si considerino i valori dello spessore (*in mm*) di **100 blocchetti metallici**, componenti di strumenti ottici.

Si disegni un istogramma dei dati.

DATI										XL	XS
3.56	3.46	3.48	3.50	3.42	3.43	3.52	3.49	3.44	3.50	3.56	3.42
3.48	3.56	3.50	3.52	3.47	3.48	3.46	3.50	3.56	3.38	3.56	3.38
3.41	3.37	3.47	3.49	3.45	3.44	3.50	3.49	3.46	3.46	3.50	3.37
3.55	3.52	3.44	3.50	3.45	3.44	3.48	3.46	3.52	3.46	3.55	3.44
3.48	3.48	3.32	3.40	3.52	3.34	3.46	3.43	3.30	3.46	3.52	3.30
3.59	3.63	3.59	3.47	3.38	3.52	3.45	3.48	3.31	3.46	3.63	3.31
3.40	3.54	3.46	3.51	3.48	3.50	3.68	3.60	3.46	3.52	3.68	3.40
3.48	3.50	3.56	3.50	3.52	3.46	3.48	3.46	3.52	3.56	3.56	3.46
3.52	3.48	3.46	3.45	3.46	3.54	3.54	3.48	3.49	3.41	3.54	3.41
3.41	3.45	3.34	3.44	3.47	3.47	3.41	3.48	3.54	3.47	3.54	3.34

Spessore dei pezzi in mm

$N=100$

$X_{Lmax}=3.68$

$X_{Smin}=3.30$



FASI PER LA COSTRUZIONE DI UN ISTOGRAMMA

Procedura:

1. **Contare** il numero dei dati (**$N=100$**).
2. **Dividere** i dati in gruppi. Per ciascun gruppo segnare il valore maggiore X_L e quello minore X_S .
3. Successivamente segnare il **più grande** dei valori X_L ed il **più piccolo** dei valori X_S , individuando così la massima **escursione** dei dati **R** .

$$R = X_{L_{\max}} - X_{S_{\min}} \quad ; \quad R = 3,68 - 3,30 = 0,38$$

4. Definire il **numero delle classi** $K = \sqrt{N}$



Il valore di K può essere definito anche in base alla Tabella

Numero di dati (N)	Numero di classi (K)
< 50	5 - 7
50 - 100	6 - 10
100 - 250	7 - 12
sopra 250	10 - 20

5. Definire **l'ampiezza della classe h**, eventualmente arrotondando il valore:

$$h = R/K \quad h = 0,38/10 = 0,038 \approx 0,05$$

pertanto **K** (*numero delle classi*) = $0,38/0,05 = 8$



6. Definire i limiti di ciascuna classe. Il valore minimo assoluto costituisce il limite inferiore della prima classe. I successivi limiti si ottengono sommando di volta in volta il valore dell'ampiezza della classe.

Es.:

classe 1:

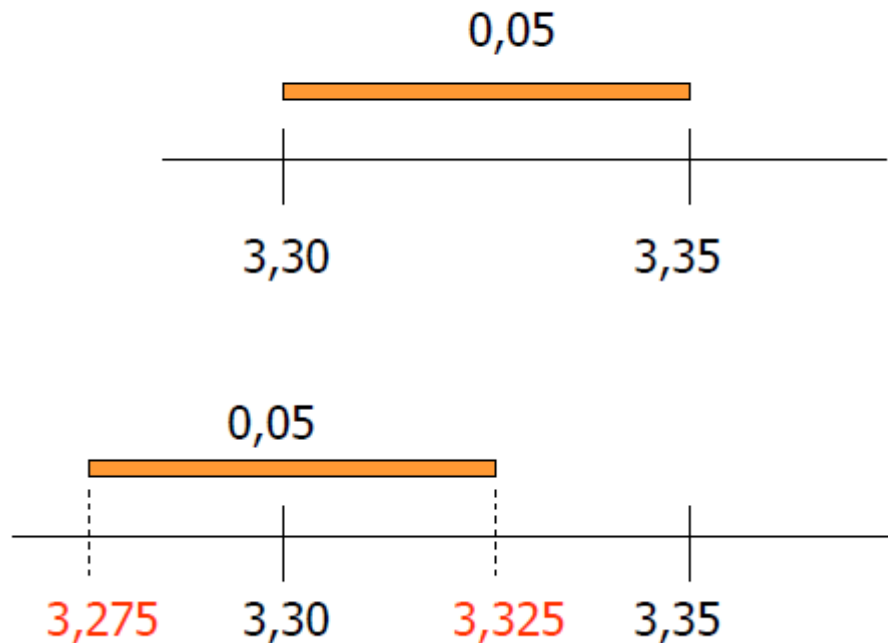
$$x_{\text{min}} ; x_{\text{min}} + h = 3,30 ; 3,30 + 0,05$$

classe 2:

$$3,35 ; 3,40$$

.....

7. Classificare la frequenza dei dati. Costruire le tabelle di frequenza. I dati che cadono in corrispondenza dei limiti delle classi, vengono registrati convenzionalmente (ad esempio) nella classe superiore.



OSSERVAZIONE:

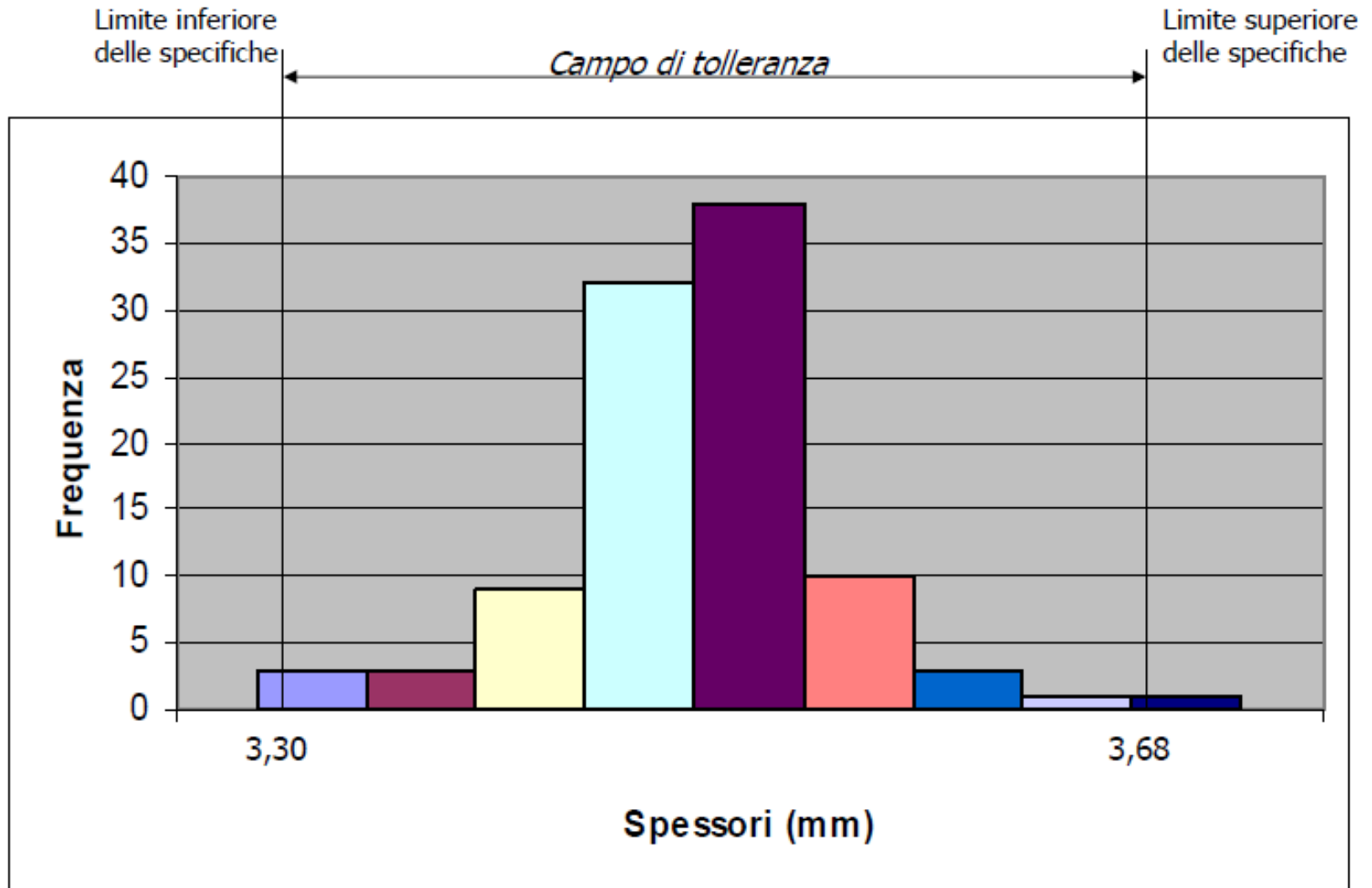
Talvolta, per evitare che qualche dato cada ai confini di una classe (*in tal caso non si saprebbe se attribuirlo all'una o all'altra delle classi contigue*), si fissano i limiti di una classe in corrispondenza della metà dell'unità di misura.

Istogramma



Classe	Confini di classe	Valori centrali	Frequenze
1	3,275 – 3.325	3,30	3
2	3.325 – 3.375	3,35	3
3	3.375 – 3.425	3,40	9
4	3.425 – 3.475	3,45	32
5	3.475 – 3.525	3,50	38
6	3.525 – 3.575	3,55	10
7	3.575 – 3.625	3,60	3
8	3.625 – 3.675	3,65	1
9	3.675 – 3.725	3,70	1

Istogramma



In Tabella sono riportati le misure del contenuto di bottiglie contenente acqua, rilevate sistematicamente **10 volte al giorno** per **12 giorni consecutivi** ed espresse in centilitri

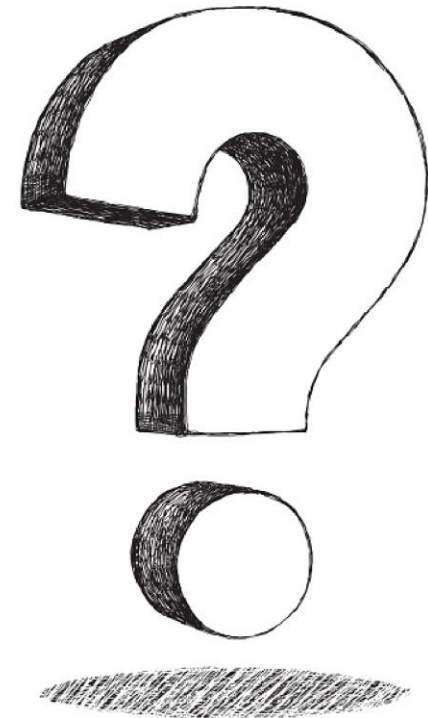
GIORNI ORA PRELIEVO	GIORNI						GIORNI					
	LUN	MAR	MER	GIO	VEN	SAB	LUN	MAR	MER	GIO	VEN	SAB
8.00	94	97	92	94	106	108	95	98	111	85	109	110
9.00	109	118	92	100	109	92	105	111	96	110	108	97
10.00	105	97	101	102	93	99	97	109	95	96	103	88
11.00	85	96	93	93	94	92	108	99	95	91	88	96
12.00	93	103	95	99	101	80	98	101	106	95	103	83
13.00	111	100	90	98	110	85	111	109	104	97	115	93
14.00	109	92	108	89	103	95	91	99	95	93	105	97
15.00	102	99	86	96	110	92	94	99	87	114	100	102
16.00	99	115	84	89	110	85	93	101	84	89	113	91
17.00	93	104	84	86	109	99	100	100	94	91	113	109



Osservando la tabella **non si riesce a giudicare quale valore** è il **più frequente**, né come i valori siano distribuiti.

Per ovviare a questo problema si possono riportare i dati su un istogramma.

ORA PRELIEVO	GIORNI											
	LUN	MAR	MER	GIO	VEN	SAB	LUN	MAR	MER	GIO	VEN	SAB
8.00	94	97	92	94	106	108	95	98	111	85	109	110
9.00	109	118	92	100	109	92	105	111	96	110	108	97
10.00	105	97	101	102	93	99	97	109	95	96	103	88
11.00	85	96	93	93	94	92	108	99	95	91	88	96
12.00	93	103	95	99	101	80	98	101	106	95	103	83
13.00	111	100	90	98	110	85	111	109	104	97	115	93
14.00	109	92	108	89	103	95	91	99	95	93	105	97
15.00	102	99	86	96	110	92	94	99	87	114	100	102
16.00	99	115	84	89	110	85	93	101	84	89	113	91
17.00	93	104	84	86	109	99	100	100	94	91	113	109





FASE 1:

Dalla tabella rileviamo il valore massimo (M) e quello minimo (m).

Nel nostro caso

- $M = 118$
- $m = 80$

Quindi calcoliamo la cosiddetta escursione R come differenza tra il valore massimo ed il valore minimo

$$R = 118 - 80 = 38$$

FASE 2

Definiamo **il numero di classi**.

Il numero di classi, indicato con K, si sceglie in funzione del numero dei dati N e può essere usato il criterio della radice quadrata $K = \sqrt{N}$

Nel nostro caso:

$$N = 120 \longrightarrow K \approx 10,95 = 10$$



FASE 3:

L'ampiezza di ogni singola classe (H), è ottenuta dividendo l'escursione di R per il numero delle classi K.

Nel nostro caso:

$$H = R/K = 38/10 \approx 3,8 = 4$$

Quindi il nostro istogramma avrà **10 classi di ampiezza 4**.

FASE 4:

A questo punto si definiscono i **limiti delle** classi iniziando dal valore minimo, nel nostro **caso 80**.

Questo valore viene assunto **come limite inferiore della prima classe**.

Il limite superiore della prima classe si calcola sommando ad **80** l'ampiezza di classe, che è pari a **4**.

FASE 4:

I limiti delle classi successive si individuano sommando di volta in volta.

CLASSI	FREQUENZE
80-84	
84-88	
88-92	
92-96	
96-100	
...	
116-120	

FASE 5:

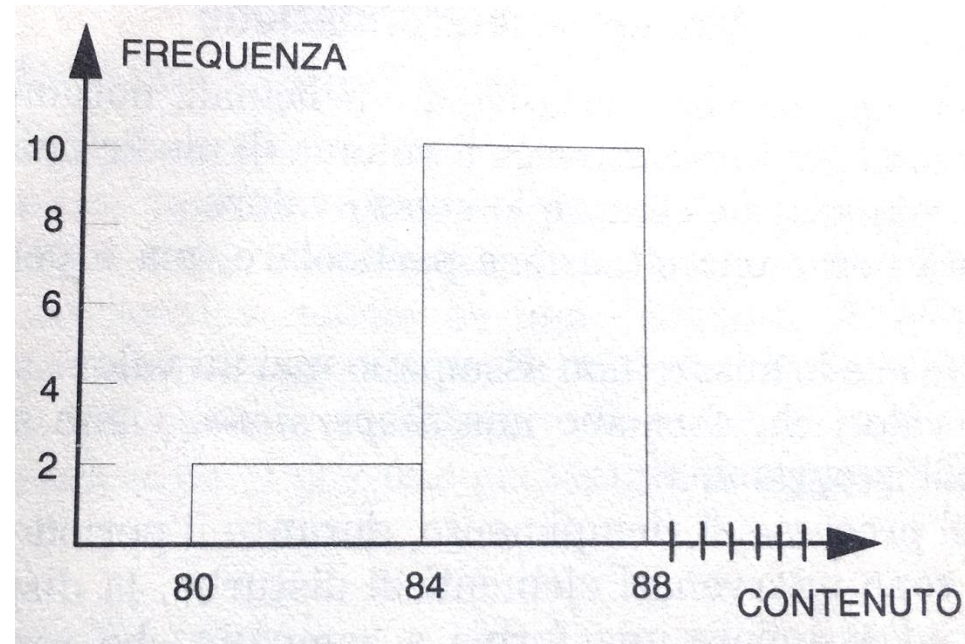
Si prepara la tabella delle frequenze e si registrano i dati

CLASSI	FREQUENZE
80-84	//
84-88	/// - ///
88-92	
92-96	
96-100	
...	
116-120	

FASE 6:

Si disegna l'istogramma.

- Sulla retta orizzontale delle ascisse metteremo i limiti delle classi.
- Sulla retta verticale delle ordinate segneremo la scala della frequenza con cui compaiono i valori di ogni classe.
- Cominceremo a disegnare le varie classi, partendo da quella che comprende i valori più bassi.



COME INTERPRETARE GLI ISTOGRAMMI

L'interpretazione di un istogramma può avvenire sulla base di due aspetti:

A. Forma della distribuzione

C'è un valore di tendenza centrale ?

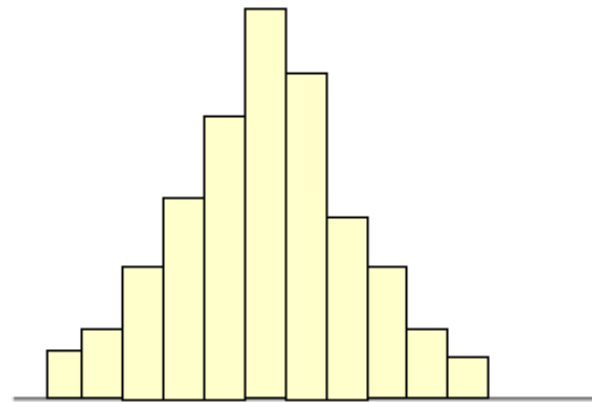
Le frequenze delle classi diminuiscono regolarmente a partire dalle classi con maggiore frequenza ?

Vi è simmetria ?

Quanto è grande la dispersione ?

Emergono configurazioni anomale ?

E' necessaria una stratificazione ?

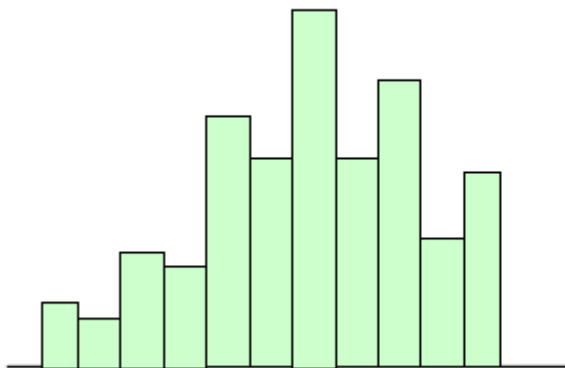
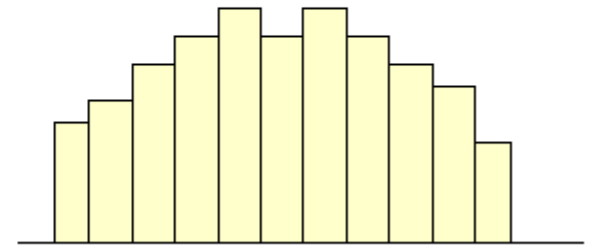


COME INTERPRETARE GLI ISTOGRAMMI

1. Piatto

Forma: la frequenza registrata in ogni classe è più o meno costante, determinando la forma di un altipiano.

Nota: la distribuzione può essere somma di distribuzioni aventi differenti valori modalì.



2. A pettine (*multimodale*)

Forma: l'istogramma assume una configurazione a denti di pettine.

Nota: Ciò può dipendere da:

- numero di classi eccessivo rispetto alla numerosità dei dati;
- arrotondamento improprio delle misure;
- errori nelle persone.

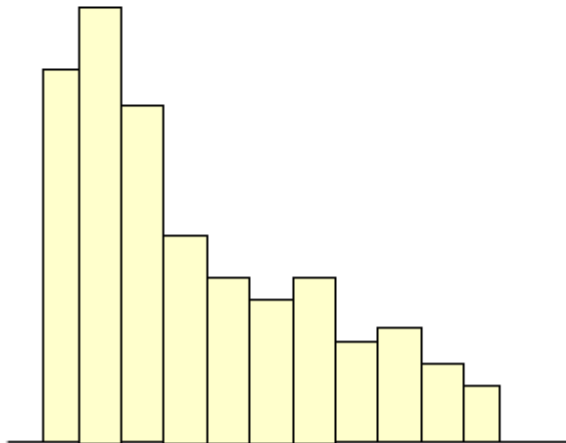
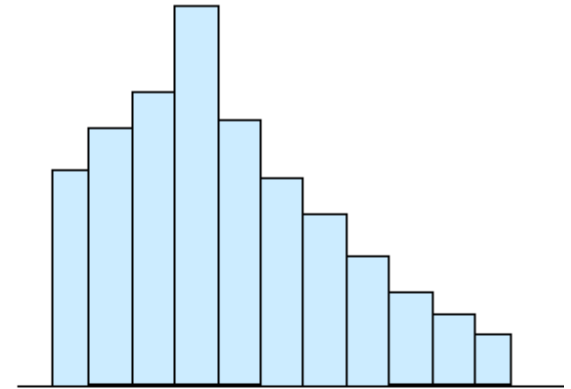


3. Asimmetrico positivo (*asimmetrico negativo*)

Forma: il valore medio dell'istogramma è posizionato sulla sinistra (destra) rispetto alla classe centrale del campo di variazione.

Nota: Non si presentano valori inferiori (superiori) ad un determinato valore. Ciò può dipendere da:

- errori di campionamento o di misura;
- sovrapposizione di dati non omogenei.

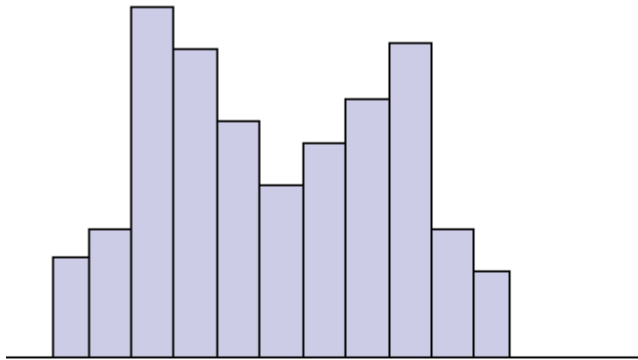


4. A precipizio sinistro (*destra*)

Forma: il valore medio dell'istogramma è posizionato molto a sinistra (destra) rispetto alla classe centrale del campo di variazione.

Nota: le principali cause possono essere:

- errori di misurazione;
- valori falsati (fondoscala di utensili e/o strumenti, esclusione arbitraria di valori).



5. Bimodale

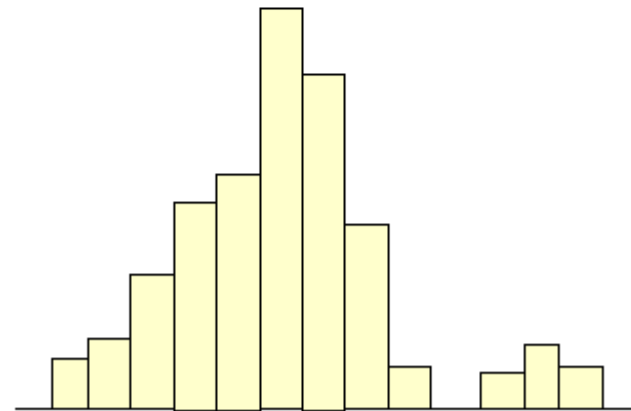
Forma: la frequenza è bassa in prossimità della classe centrale dell'intervallo dei dati con due picchi.

Nota: sovrapposizione di dati non omogenei (due distribuzioni con valori medi diversi)

6. Con picco isolato

Forma: istogramma normale con picco isolato

Nota: cercare le cause dei pochi dati isolati utilizzando la stratificazione (anomalie di processi, errori di misura o inclusione di dati da un differente processo)



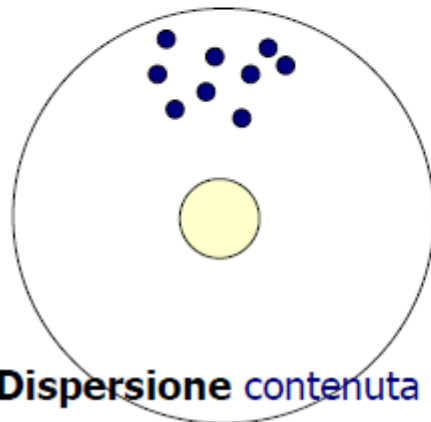
COME INTERPRETARE GLI ISTOGRAMMI

B. POSIZIONAMENTO RISPETTO ALLE SPECIFICHE

Tutte le volte che si ha una distribuzione di dati, si deve qualificare tale distribuzione rispetto a 2 caratteristiche:

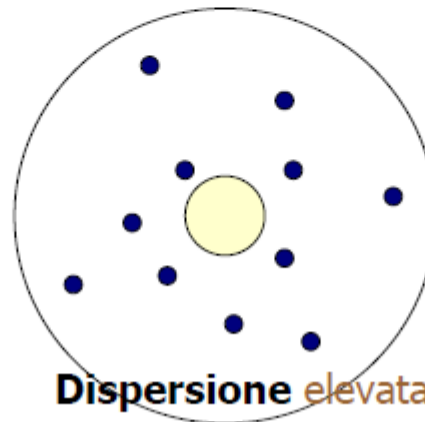
1. CENTRATURA

2. DISPERSIONE



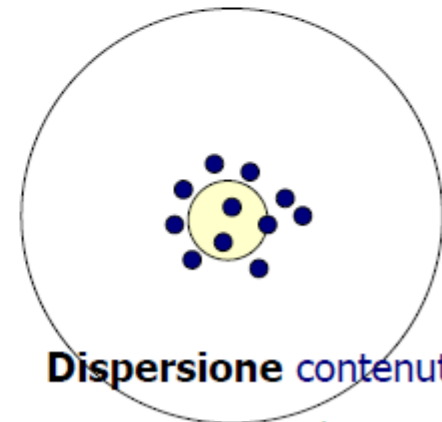
Dispersione contenuta

Centratura scarsa



Dispersione elevata

Centratura buona



Dispersione contenuta

Centratura buona



CONFRONTO DI ISTOGRAMMI CON I LIMITI DI SPECIFICA

Se esistono **limiti di specifica**, è necessario disegnare i limiti di specifica sull'istogramma e confrontare la distribuzione con la specifica.

Come si posiziona l'istogramma rispetto alle specifiche ?

E' centrato ?

E' più o meno ampio delle specifiche ?



CENTRATURA

MEDIA: Si definisce media aritmetica di un insieme di n numeri

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

$$M_e = X_{\left(\frac{n+1}{2}\right)} \quad \text{Se } n \text{ dispari}$$

$$M_e = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} \quad \text{Se } n \text{ pari}$$

MEDIANA: Si definisce valore mediano o mediana M_e quel valore che lascia tanti elementi a destra quanto a sinistra.

MODA: valore dominante indica il valore che si presenta più spesso



DISPERSIONE

VARIANZA: media quadratica degli scarti dei singoli dati dalla loro media aritmetica

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

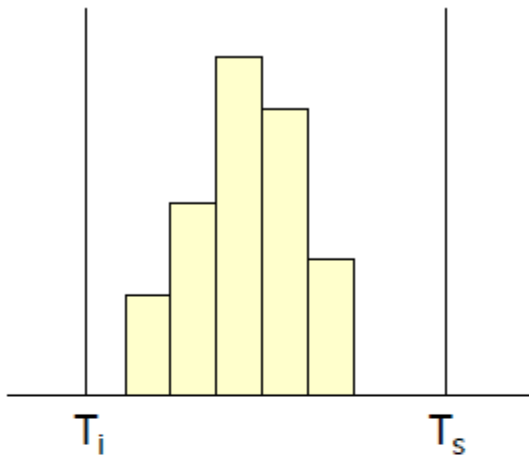
$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

DEVIAZIONE STANDARD o scarto quadratico medio: radice quadrata della varianza

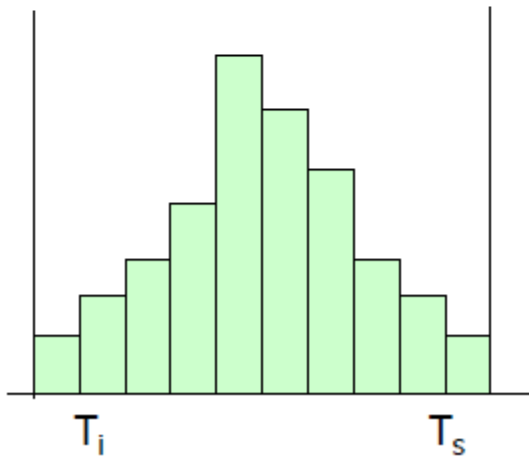
RANGE: campo di variazione o di variabilità indica la differenza tra il valore massimo e minimo

$$R = x_{max} - x_{min}$$

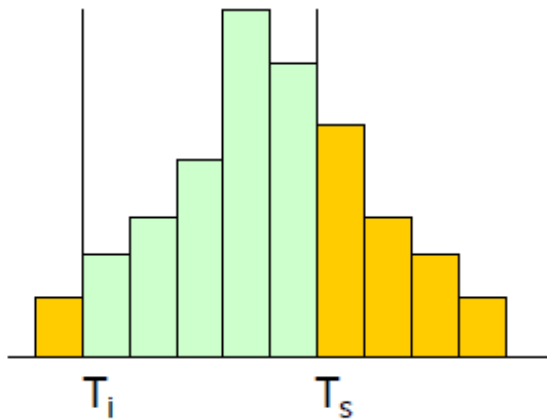
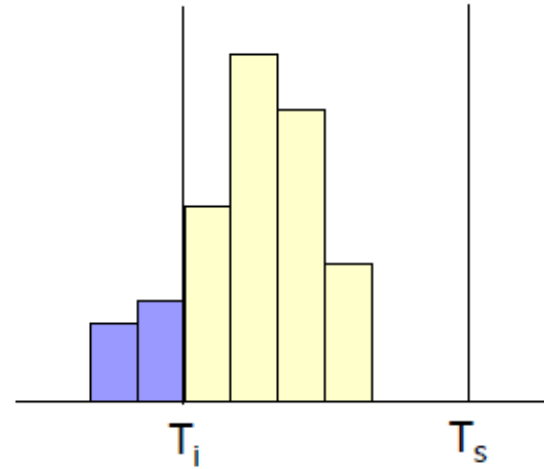
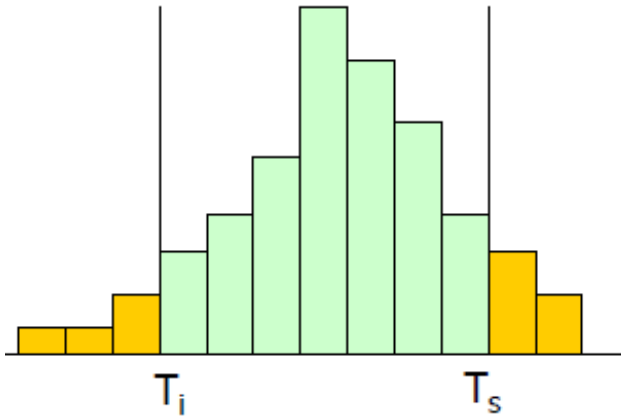
Istogramma



Casi in cui l'istogramma soddisfa
le specifiche



Istogramma



Casi in cui l'istogramma non
soddisfa le specifiche

A COSA SERVE

A fotografare la dispersione di un process
Ci aiuta a capire la variabilità di un fenomeno

COME SI APPLICA

Bisogna raccogliere i dati in condizioni omogenee
Bisogna rappresentare graficamente i dati rispettando le regole sui numeri a barre
Nell'interpretazione, basarsi sui modelli di riferimento per le forme anomale

DOVE SI APPLICA

Nel tenere sotto controllo la dispersione delle variabili più importanti e critiche

QUANDO SI APPLICA

Subito dopo la raccolta dati per dare un indirizzo all'analisi delle cause

ERRORI DA EVITARE

Accontentarsi di pochi dati
Considerare l'istogramma come uno strumento solo di rappresentazione e non di analisi
Confondere l'istogramma con il diagramma di Pareto