

Modelli della qualità del processo

Viene data una breve introduzione alla statistica descrittiva e alle distribuzioni di probabilità.

L'utilizzo dei semplici strumenti della statistica descrittiva serve per valutare quantitativamente le variazioni che una caratteristica o indicatore di qualità presenta tra i valori considerati in un campione.

Le distribuzioni di probabilità forniscono uno strumento per la realizzazione del modello o la descrizione delle caratteristiche di qualità di un processo produttivo.

Le principali rappresentazioni grafiche

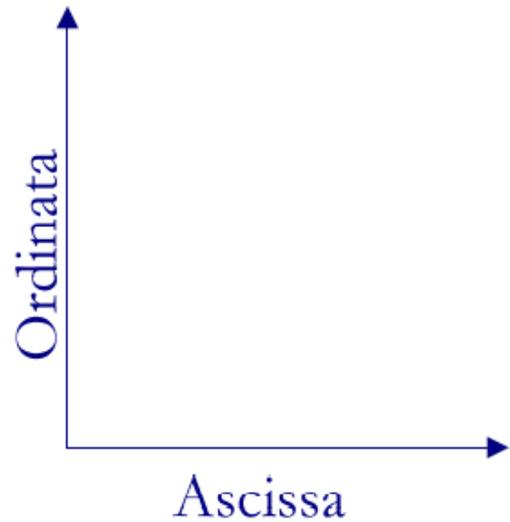


Diagramma cartesiano

I diagrammi cartesiani sono utili per:

- la rappresentazione delle variabili statistiche per singoli valori
- la rappresentazione di serie storiche (fenomeni che dipendono dal tempo)

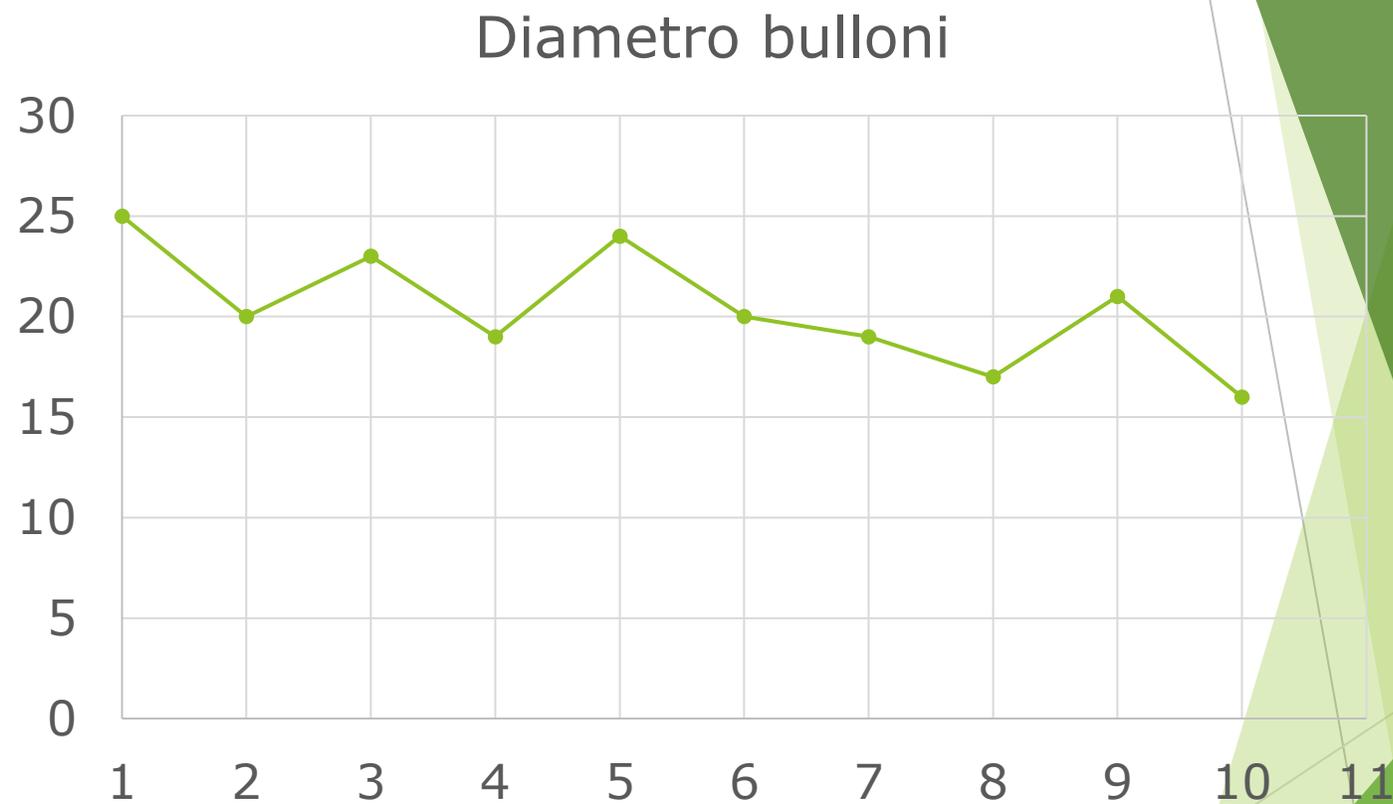
Diagramma cartesiano



In ascissa sono rappresentate le modalità del carattere (x_i), in ordinata le modalità del carattere (y_i) oppure le frequenze del carattere x_i (f_i).

Diagramma cartesiano

Giorni	Diametro bulloni
1	25
2	20
3	23
4	19
5	24
6	20
7	19
8	17
9	21
10	16



Istogramma

L'istogramma è uno strumento grafico che consente di avere una visione completa e sintetica dei dati raccolti.

Esso fornisce un'impressione visiva della forma della distribuzione dei dati sia informazioni sulla variabilità intrinseca degli stessi.

Istogramma

L'istogramma è la rappresentazione di una variabile statistica suddivisa in classi.

L'istogramma, che non è altro che un diagramma a colonne, presenta in ordinata il numero di osservazioni in ciascuna classe e in ascissa le classi (il centro di ogni colonna coincide con il valore centrale della classe).

Per costruire un istogramma per dati continui, occorre dividere il campo di variazione dei dati in intervalli (classi).

Istogramma

Preferibilmente, le classi dovrebbero essere di uguale ampiezza, per non distorcere l'informazione visiva fornita dall'istogramma.

Per ottenere un grafico corretto devono essere usati alcuni criteri per la scelta del numero ottimale di classi, in quanto un numero eccessivo o troppo limitato risulterà poco informativo.

Istogramma

Grafici simili agli istogrammi possono essere usati con alcuni tipi di dati qualitativi oppure dati di conteggio.

Per costruire un grafico per dati discreti, prima si determina la frequenza di ogni valore di x .

Ognuno dei valori di x corrisponde ad una classe.

Istogramma

L'istogramma è tracciato rappresentando graficamente le frequenze sull'asse delle ordinate e i valori della x sull'asse delle ascisse.

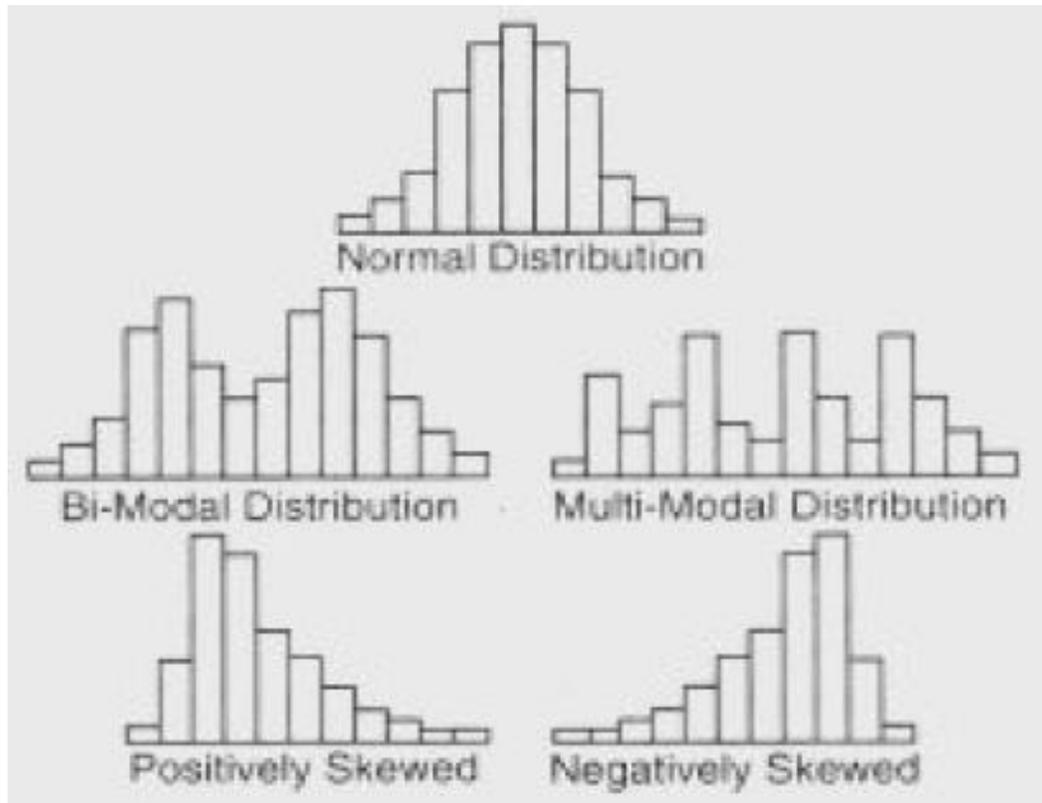
Sopra ogni valore di x , si traccia un rettangolo la cui altezza è la frequenza corrispondente a quel valore.

Istogramma

Importanti indicazioni sul comportamento del processo produttivo sono ottenibili con l'analisi di alcuni aspetti dell'istogramma in esame:

a. **FORMA DEL GRAFICO:** può essere utile verificare se la distribuzione dei dati segue un andamento a campana o al contrario siano presenti due o più picchi di frequenza (distribuzione bimodale o multimodale) dovuti generalmente alla sovrapposizione di dati di origine diversa (due macchine, due operatori).

Istogramma

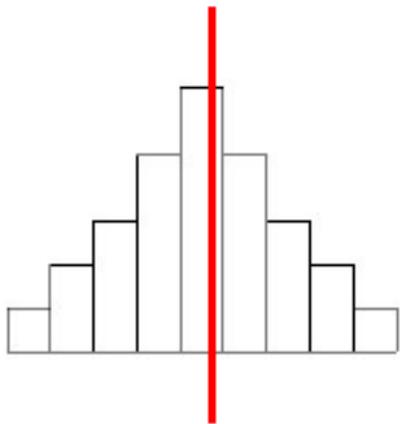


L'istogramma può rilevare una asimmetria nella distribuzione dei dati, sintomatica di qualche errore nella raccolta dati o nella misurazione o nella sovrapposizione di dati non omogenei.

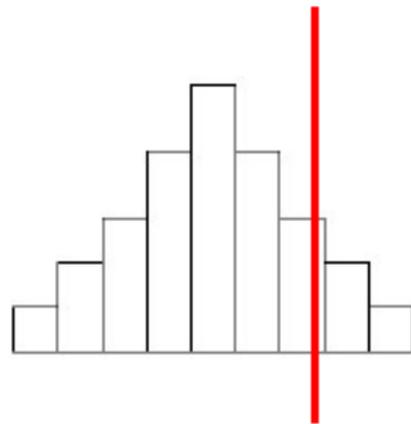
Il processo può essere sbilanciato in senso positivo (più valori a sinistra) o negativo (più valori a destra).

Istogramma

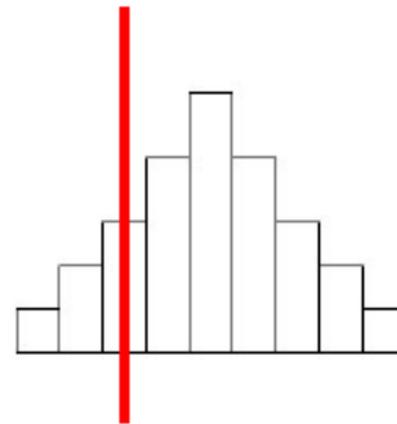
b. POSIZIONE O TENDENZA CENTRALE : l'istogramma evidenzia, anche se la distribuzione dei dati provenienti dall'output di processo è centrata sull'obiettivo (valore nominale) fornendo indicazioni sull'accuratezza del processo.



PROCESSO CENTRATO



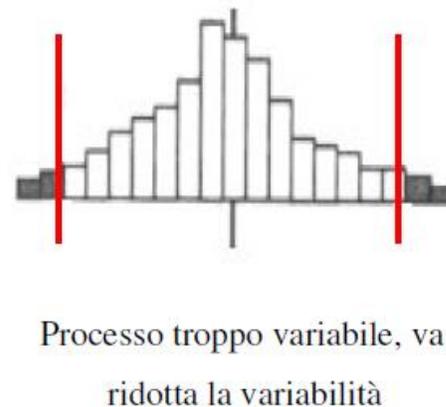
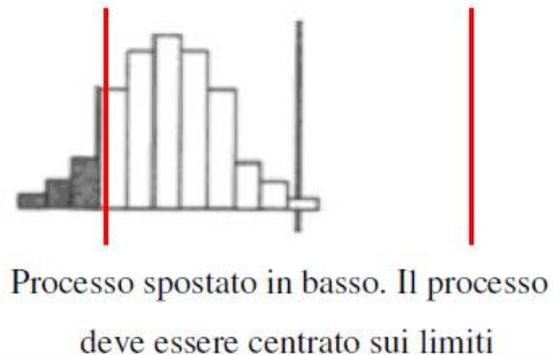
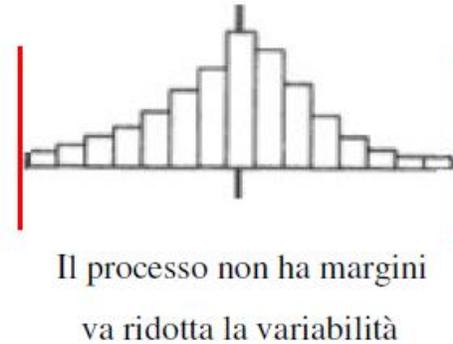
PROCESSO POSIZIONATO
TROPPO IN BASSO



PROCESSO POSIZIONATO
TROPPO IN ALTO

Istogramma

c. DISPERSIONE: un istogramma consente inoltre di valutare la precisione del processo produttivo tramite l'analisi di dispersione della distribuzione dei dati, anche in relazione ai limiti di tolleranza.



Si possono osservare diagrammi a campana fortemente appiattiti, che indicano una forte dispersione dei valori, e altri fortemente concentrati in corrispondenza del valore centrale.

Esempio di istogramma

Una azienda farmaceutica decide di effettuare un controllo sul contenuto di un farmaco all'interno di appositi flaconi.

L'azienda assume come tollerabili un quantitativo minimo di medicinale nei flaconi pari a 82 ml e uno massimo di 118 ml e in fase di progetto stabilisce un quantitativo obiettivo di 95 ml.

Esempio di istogramma

Gli operatori addetti a tale compito hanno a disposizione le seguenti misure del contenuto dei flaconi del prodotto medicinale:

orario	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	108	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

Esempio di istogramma

Per un'immediata valutazione sul processo si decide di costruire un istogramma dei dati di output.

Dalla tabella dei dati sono stati ricavati il valore massimo (M) e quello minimo (m): $M=118$ e $m=80$.

Quindi si calcola l'escursione (R) come differenza tra il valore massimo e quello minimo: $R=M-m=38$.

Esempio di istogramma

Il numero delle classi, indicato con K , si sceglie in funzione del numero dei dati e viene usato il criterio della radice quadrata di N .

Nel qual caso, poiché $N=120$ si ottiene $K=10.95$ e per approssimazione si ottiene $K=11$.

L'ampiezza di ogni singola classe (h) è ottenuta dividendo l'escursione di R per il numero delle classi $K= 3,8$ arrotondato a 4 per comodità.

Esempio di istogramma

A questo punto si definiscono i limiti delle classi iniziando dal valore minimo, che viene assunto come limite inferiore della prima classe.

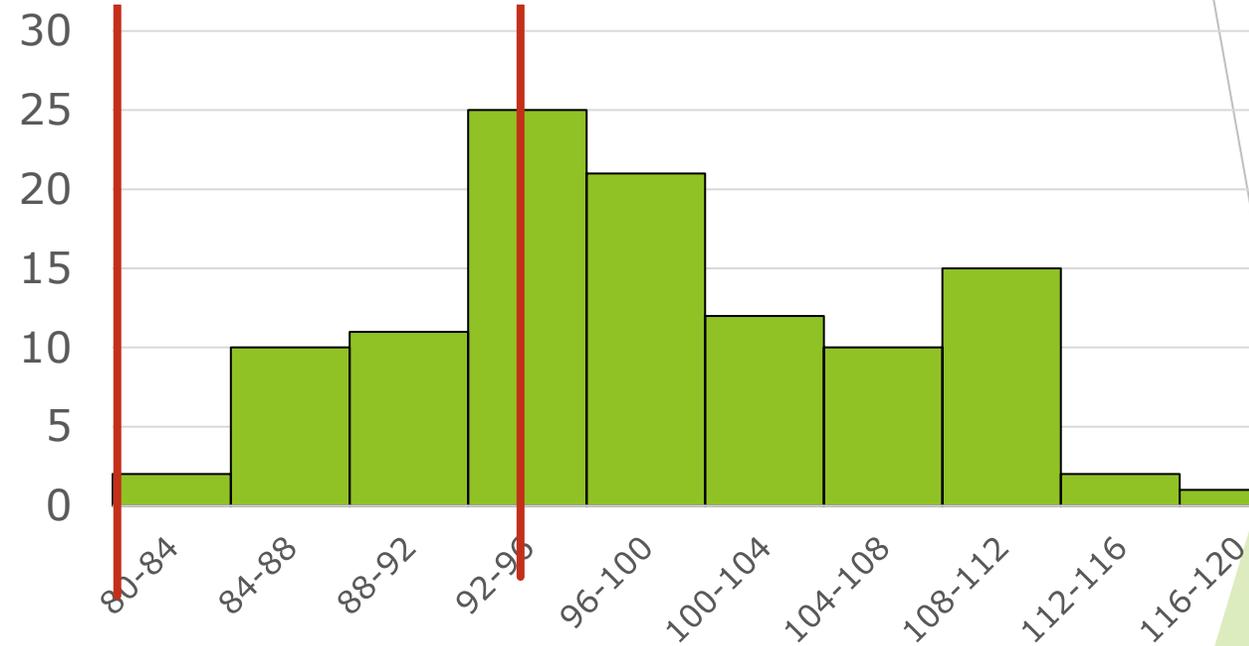
Il limite superiore sarà dato da quello inferiore più l'ampiezza di classe (4).

I limiti delle classi successive si individuano sommando di volta in volta l'ampiezza di classe.

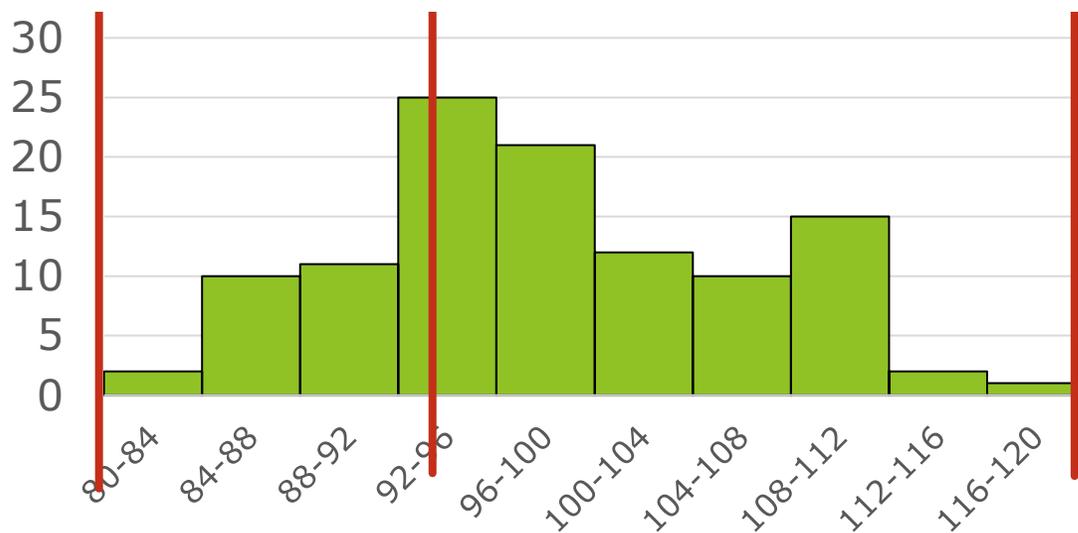
CLASSI	FREQUENZE
80-84	2
84-88	10
88-92	11
92-96	25
96-100	21
100-104	12
104-108	10
108-112	15
112-116	2
116-120	1

Esempio di istogramma

CLASSI	FREQUENZE
80-84	2
84-88	10
88-92	11
92-96	25
96-100	21
100-104	12
104-108	10
108-112	15
112-116	2
116-120	1



Esempio di istogramma



Dall'istogramma si può subito notare come i dati seguano approssimativamente una distribuzione normale, con una piuttosto accentuata variabilità dei dati.

Rispetto al target aziendale il processo è abbastanza centrato, mentre in termini dei limiti di tolleranza il processo sembra non avere margini per cui potrebbe essere necessaria una azione correttiva sulla variabilità del processo.

Misure di Tendenza Centrale

Le misure di tendenza centrale forniscono una visione di insieme e sintetizzano la posizione del centro ideale intorno al quale si spargono gli elementi della distribuzione.

Esistono tanti tipi di fenomeni, quindi sono stati introdotti tanti indicatori di tendenza centrale.

Media

La media (aritmetica) è il punto di equilibrio o baricentro dell'insieme di valori.

È un valore analitico, funzione di tutti i valori della distribuzione.

Dati n valori osservati x_1, x_2, \dots, x_n di un carattere quantitativo X , la media si calcola facendo il rapporto tra l'ammontare del carattere e il numero dei casi osservati:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Esempio media

Misura del peso di 10 bulloni:

0,485

0,444

0,442

0,446

0,447

0,480

0,448

0,419

0,420

0,460

$$\bar{x} = \frac{0.485+0.444+0.442+0.446+0.447+0.480+0.448+0.419+0.420+0.460}{10} = 0.449\text{kg}$$

Mediana

La mediana è una media di posizione: è il valore che occupa la posizione centrale nell'insieme ordinato di tutti i valori.

$$X_{\min} = X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)} = X_{\max}$$

Tra $x(1)$ e Me è contenuto il 50% dei valori

Tra Me e $x(n)$ è contenuto il restante 50% dei valori



Mediana

1. Ordinare le n unità in senso crescente rispetto alle modalità del carattere

2. Individuare la posizione in graduatoria dell'unità centrale:

- se n è dispari, la posizione è $\frac{(n+1)}{2}$

- se n è pari, si hanno due unità centrali, con posizione $\frac{n}{2}$ e $\left(\frac{n}{2}\right)+1$

3. Osservare la modalità presentata dall'unità centrale:

- se n è dispari la mediana è $Me = x_{(n+1)/2}$

- se n è pari abbiamo due modalità corrispondenti alle due unità centrali: $x_{n/2}$ e $x_{(n/2)+1}$.

Se il carattere è quantitativo, è possibile considerare come mediana la semisomma dei valori delle due modalità centrali.

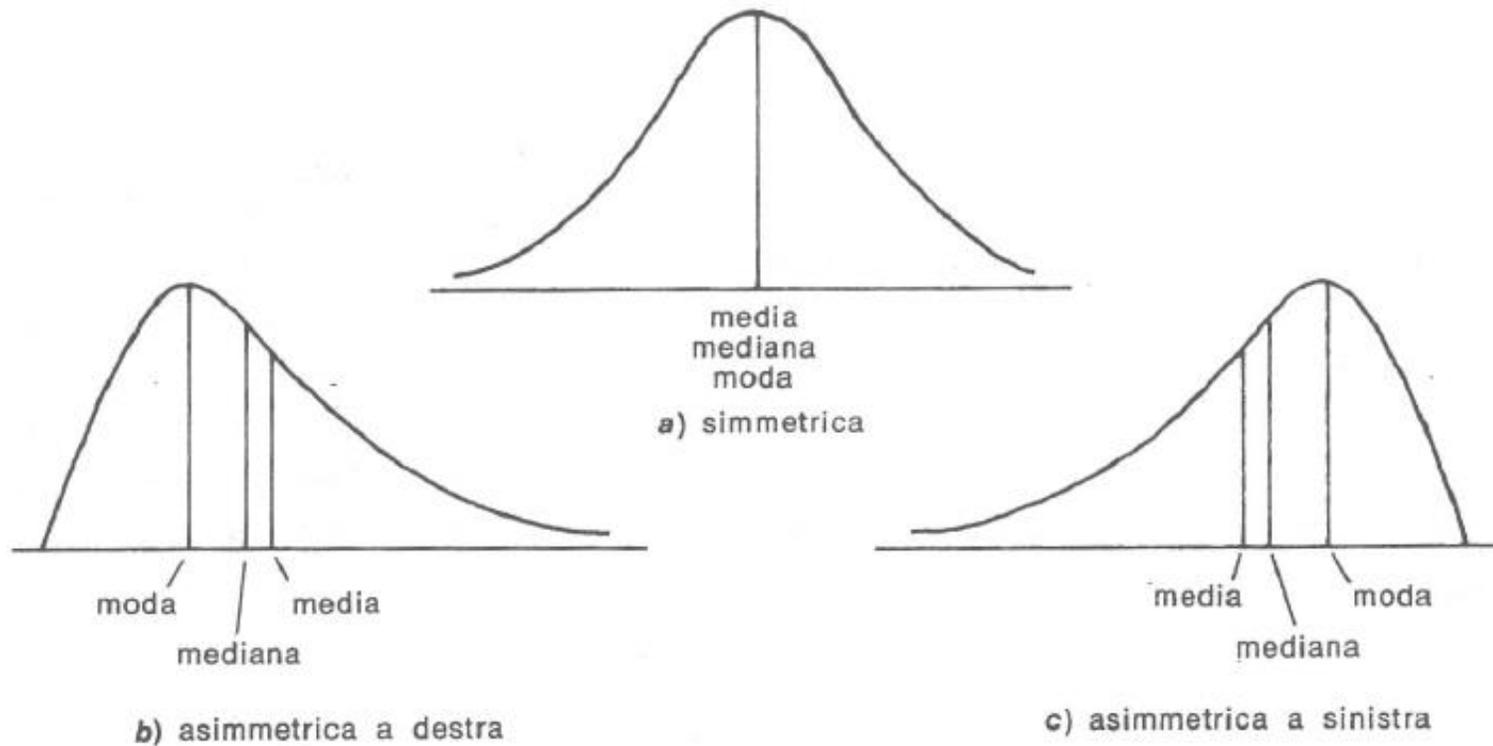
Moda

La moda indica il valor più ricorrente.

Può non essere unica (distribuzione plurimodale).

Non ha grande significato quando i dati sono pochi e molto dispersi.

Esempi di come si dispongono la media, la mediana e la moda



Varianza

Il calcolo di una media non esaurisce la descrizione sintetica di un fenomeno osservato in un collettivo.

Due insiemi di valori, pur avendo lo stesso valore medio, possono essere molto differenti tra di loro.

Gli indici di variabilità forniscono informazioni complementari a quelle degli indici medi.

Varianza

La varianza campionaria S^2 è funzione delle differenze (scarti) tra ogni valore x_i e la media:

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

Varianza

L'unità di misura della varianza campionaria è data dal quadrato delle unità di misura della grandezza osservata.

Questo è motivo di qualche difficoltà di interpretazione, per cui di solito si preferisce impiegare correntemente la radice quadrata della varianza campionaria, che viene detta deviazione standard campionaria (o scarto quadratico medio), la quale è espressa nella stessa unità di misura dei dati osservati:

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

Campo di variazione

Si definisce campo di variazione o campo di variabilità di un insieme di valori la differenza tra il massimo e il minimo di tali valori:

$$\text{Range} = x_{\max} - x_{\min}$$

Il Box Plot

L'istogramma fornisce una rappresentazione visiva di un set di dati, mentre la media e la deviazione standard forniscono informazioni quantitative su caratteristiche specifiche dei dati.

Il box-plot è una rappresentazione grafica che permette di visualizzare contemporaneamente diverse caratteristiche importanti dei dati, come posizione, variabilità, simmetria e identificazione di osservazioni insolitamente distanti dalla maggior parte dei dati (queste osservazioni sono spesso chiamate "valori anomali").

Il Box Plot

Il box-plot visualizza i tre quartili, il minimo e il massimo dei dati su un riquadro rettangolare, allineati in orizzontale o in verticale.

Un riquadro racchiude l'intervallo interquartile con la linea sinistra (o inferiore) nel primo quartile Q_1 e la linea destra (o superiore) in il terzo quartile Q_3 .

Una linea viene disegnata attraverso il riquadro nel secondo quartile (che è il cinquantesimo percentile o la mediana).

Una linea alle due estremità si estende ai valori estremi.

Il Box Plot

Queste linee sono generalmente chiamate baffi.

Alcuni autori definiscono questo grafico «scatola e baffi».

In alcuni programmi, i baffi estendono solo una distanza di 1,5 ($Q_3 - Q_1$) dalle estremità della scatola, al massimo, e le osservazioni oltre questi limiti sono contrassegnate come potenziali valori anomali.

Il Box Plot

Diametri (in mm) dei fori in un gruppo di 12 bulloni.

N.	Diametri
1	120.5
2	120.9
3	120.3
4	121.3
5	120.4
6	120.2
7	120.1
8	120.5
9	120.7
10	121.1
11	120.9
12	120.8

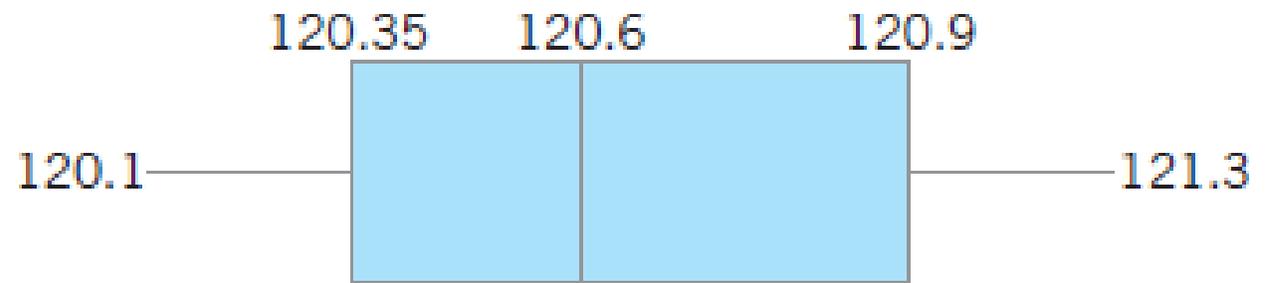
Le principali statistiche
descrittive



Media	120.64
Mediana	120.60
Moda	120.50
Deviazione standard	0.37
Varianza campionaria	0.14
Curtosi	-0.82
Asimmetria	0.25
Intervallo	1.20
Minimo	120.10
Massimo	121.30
Conteggio	12

Il Box Plot

Media	120.64
Mediana	120.60
Moda	120.50
Deviazione standard	0.37
Q_1	120.35
Q_3	120.9
Varianza campionaria	0.14
Curtosi	-0.82
Asimmetria	0.25
Intervallo	1.20
Minimo	120.10
Massimo	121.30
Conteggio	12



La trama della scatola indica che la distribuzione del diametro del foro non è esattamente simmetrica attorno al valore centrale, perché i baffi a sinistra e destra e le caselle a sinistra e a destra attorno alla mediana non sono delle stesse dimensioni.

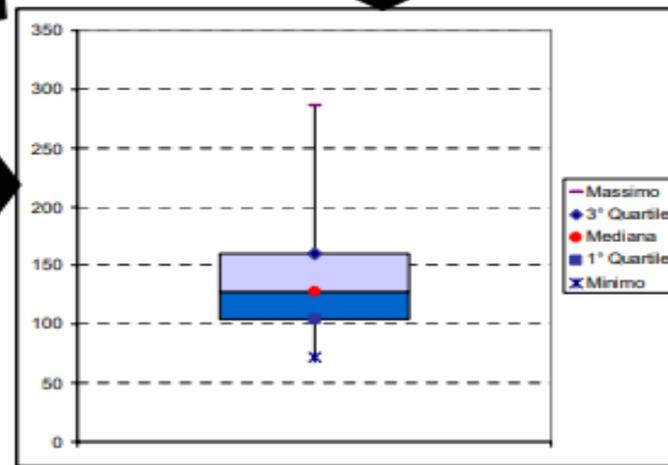
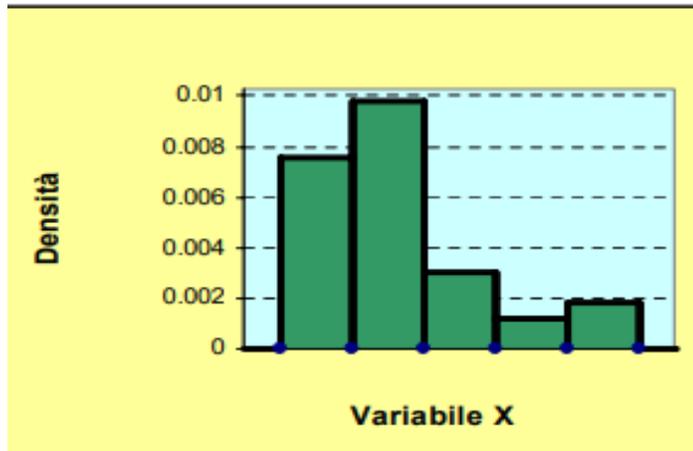
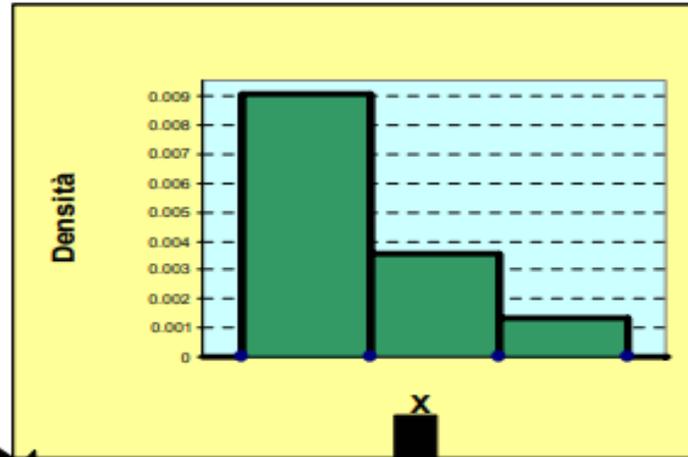
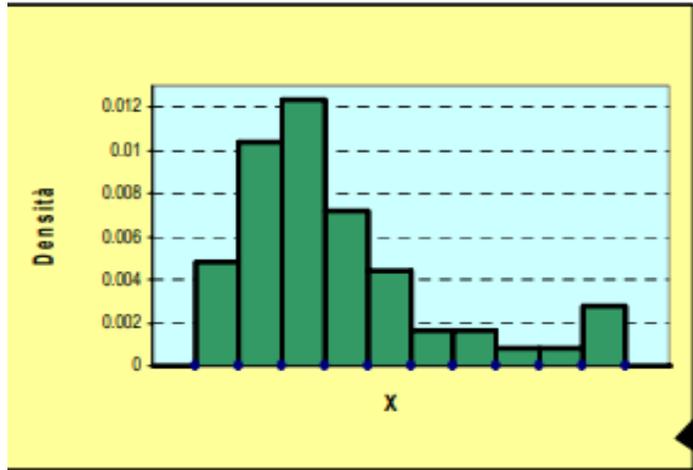
Il Box Plot

Il box plot fornisce una rappresentazione univoca della distribuzione, a differenza dell'istogramma che può dare rappresentazioni diverse a seconda degli estremi delle classi scelte.

Ad esempio, di seguito sono riportati 3 istogrammi relativi ad uno stesso carattere, ma ottenuti scegliendo un numero differente di classi di diversa ampiezza.

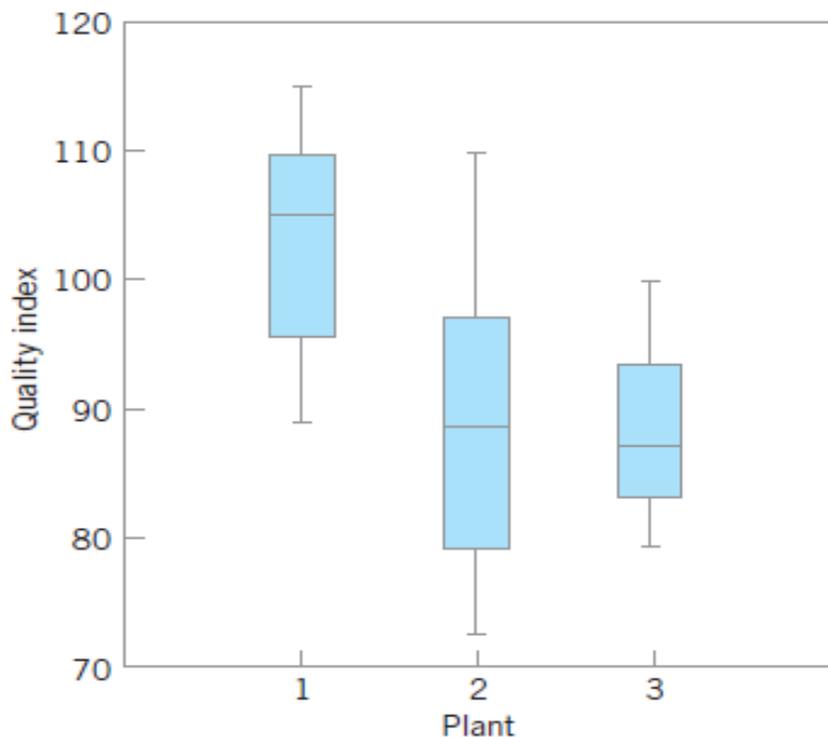
Il box plot relativo alla distribuzione, però, non varia.

Il Box Plot



Il Box Plot

I box plot sono molto utili nei confronti grafici tra set di dati, perché hanno un impatto visivo e sono facili da capire.



Ad esempio, i 3 box plot della figura mostrano la comparazione tra un indice di qualità della produzione di tre stabilimenti produttivi.

L'ispezione di questo grafico rivela che c'è troppa variabilità nell'impianto 2 e che gli impianti 2 e 3 devono aumentare le prestazioni dell'indice di qualità.