

# Statistica descrittiva per l'Estimo

Paolo Rosato

Dipartimento di Ingegneria Civile e Architettura

Piazzale Europa 1 - 34127 Trieste. Italia

Tel: +39-040-5583569. Fax: +39-040-55835 80

E-mail: [paolo.rosato@dia.units.it](mailto:paolo.rosato@dia.units.it)

# A cosa serve la statistica nell'estimo

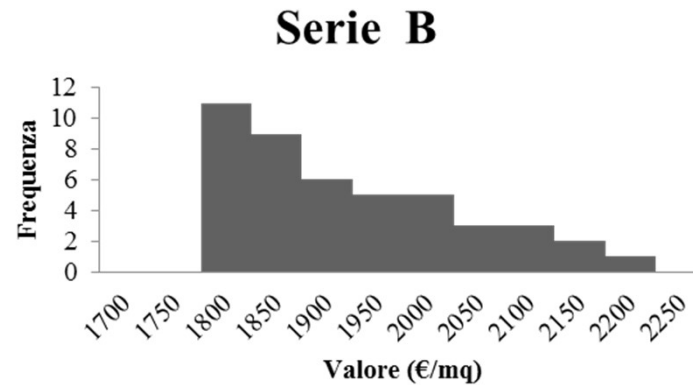
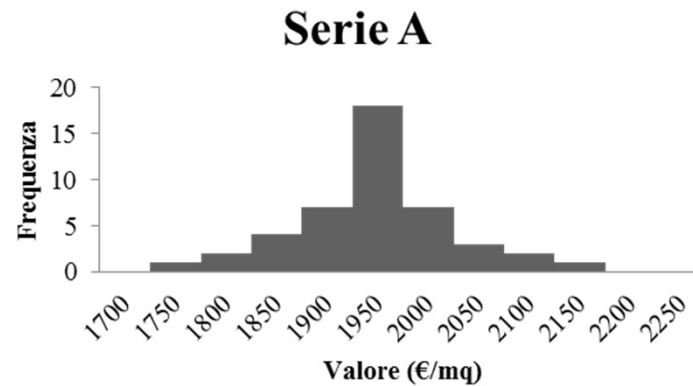
La statistica è uno strumento utile/indispensabile per studiare e sintetizzare fenomeni che si manifestano in modo incerto e/o sui quali non è possibile avere una completa conoscenza.

Il prezzo degli immobili è uno di questi fenomeni, poiché:

1. La conoscenza dei prezzi è incompleta (mercato opaco)
2. Le caratteristiche degli immobili che influiscono sui prezzi non sono perfettamente note
3. Le transazioni sono poche
4. Le fonti di dati sono scarse e poco omogenee
5. I prezzi che si formano hanno una forte componente specifica e casuale

# A cosa serve la statistica nell'estimo

Immobile	Serie	
	A	B
1	1.950	2.000
2	1.800	1.800
3	1.900	1.850
4	1.950	1.900
5	1.850	1.850
6	2.100	1.800
7	1.910	1.910
8	1.930	1.930
9	2.000	1.850
10	1.974	1.900
11	1.860	1.860
12	1.930	1.800
13	1.950	1.850
14	2.050	2.050
15	1.980	1.980
16	1.950	1.850
17	1.870	2.050
18	1.900	1.800
19	1.950	2.150
20	1.850	1.800
21	1.850	2.100
22	1.910	1.910
23	1.930	2.100
24	2.050	2.100
25	1.974	1.800
26	1.860	1.800
27	1.930	1.800
28	1.950	2.000
29	2.150	2.150
30	1.740	2.200
Media	1.933	1.931



?

# La statistica si articola in due ambiti principali

**Statistica descrittiva:** sintetizza e rappresenta i dati osservati (p.e. caratteristiche degli immobili e prezzi) mediante grafici e indici che descrivono tendenze e variabilità

- Indicatori di tendenza centrale
- Indicatori di dispersione
- Distribuzioni
- Indicatori di forma

# La statistica di articola in due ambiti principali

**Statistica inferenziale:** tenta di stabilire delle relazioni fra i dati osservati (p.e. caratteristiche dell'immobile e prezzo di mercato), spesso apparentemente disordinati, fornendone una valutazione probabilistica

- La regressione semplice
- La regressione multipla

# La statistica descrittiva

Gli indicatori di tendenza centrale:

- Media semplice
- Media geometrica
- Media armonica
- Media ponderata
- Mediana
- Moda

# Alcuni semplici indicatori statistici di tendenza centrale

**Media semplice ( $M_s$ ):** Rapporto fra la somma dei valori ( $V_i$ ) ed il loro numero ( $n$ ).

$$M_s = \frac{\sum_{i=1}^n V_i}{n}$$

# Alcuni semplici indicatori statistici di tendenza centrale

**Media armonica ( $M_h$ ):** Reciproco della media aritmetica dei reciproci dei valori ( $V_i$ ). Si usa quando è utile calcolare il reciproco dei dati: il potere di acquisto medio della moneta è il reciproco della media armonica dei prezzi.

$$M_h = \frac{n}{\sum_{i=1}^n \frac{1}{V_i}}$$



# Alcuni semplici indicatori statistici di tendenza centrale

**Media geometrica ( $M_g$ ):** Radice  $n$ -esima del prodotto degli ( $n$ ) valori ( $V_i$ ). Si usa quando ha senso moltiplicare fra loro i dati statistici: determinare il tasso d'incremento medio o di decremento di prezzi.

$$M_s = \sqrt[n]{\prod_{i=1}^n V_i}$$

## Alcuni semplici indicatori statistici di tendenza centrale

**Media quadratica ( $M_g$ ):** Radice quadrata della media semplice del quadrato degli ( $n$ ) valori ( $V_i$ ). Si usa per mettere in evidenza l'esistenza di valori anomali, che si scostano molto dai valori centrali.

$$M_q = \sqrt{\frac{\sum_{i=1}^n V_i^2}{n}}$$

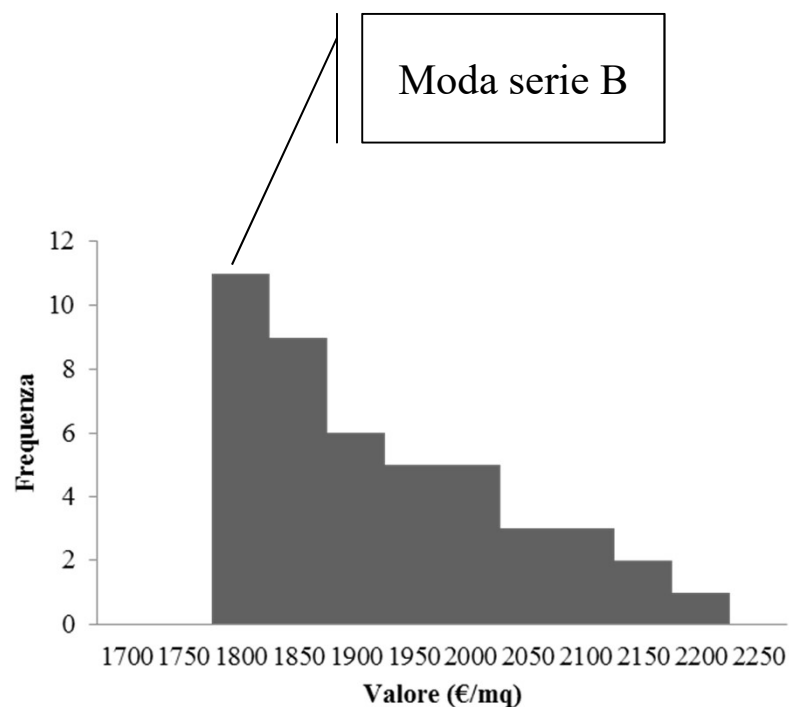
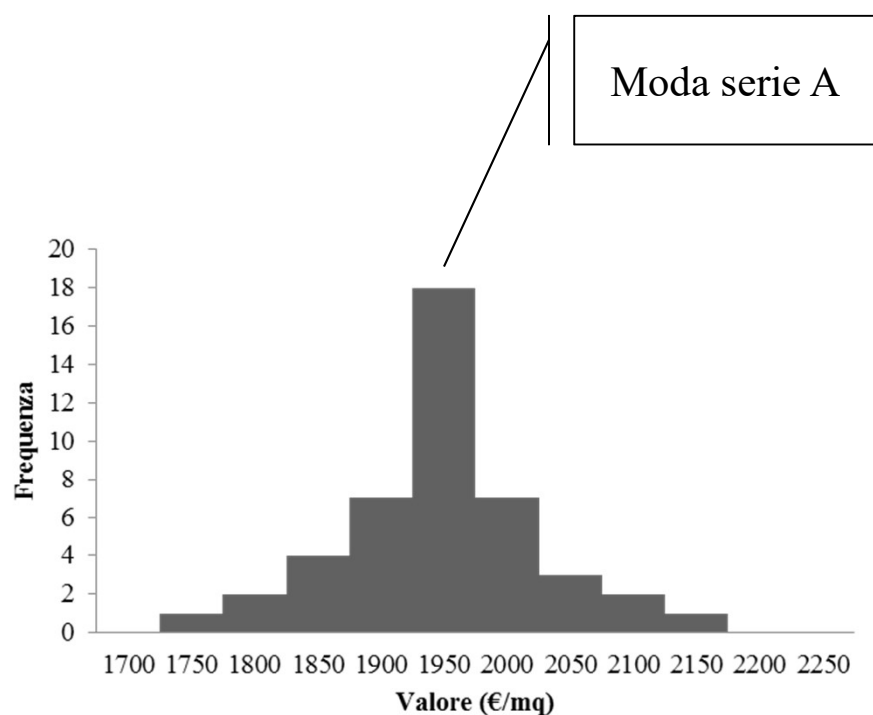
# Alcuni semplici indicatori statistici di tendenza centrale

**Media ponderata ( $M_p$ ):** Rapporto fra la somma dei valori ( $V_i$ ) moltiplicati per il loro peso ( $w_i$ ) e la somma dei pesi ( $w_i$ ). Serve per ponderare il dato in funzione di una specifica caratteristica.

$$M_p = \frac{\sum_{i=1}^n V_i \cdot w_i}{\sum_{i=1}^n w_i}$$

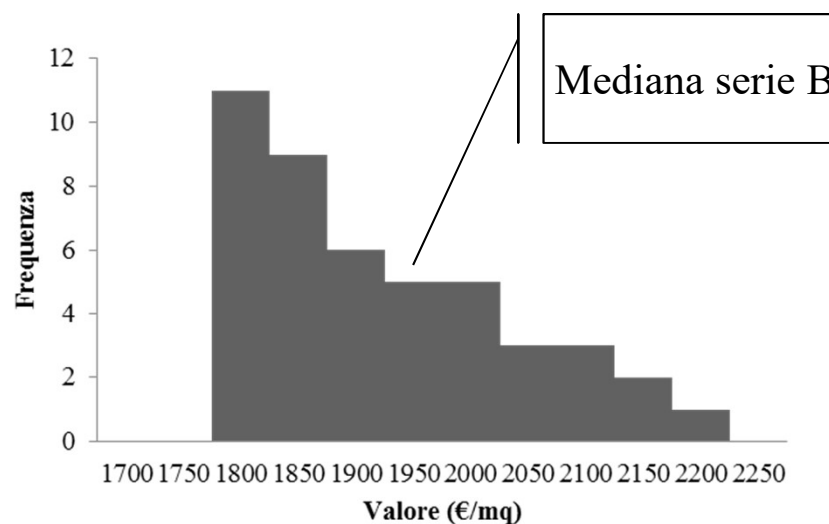
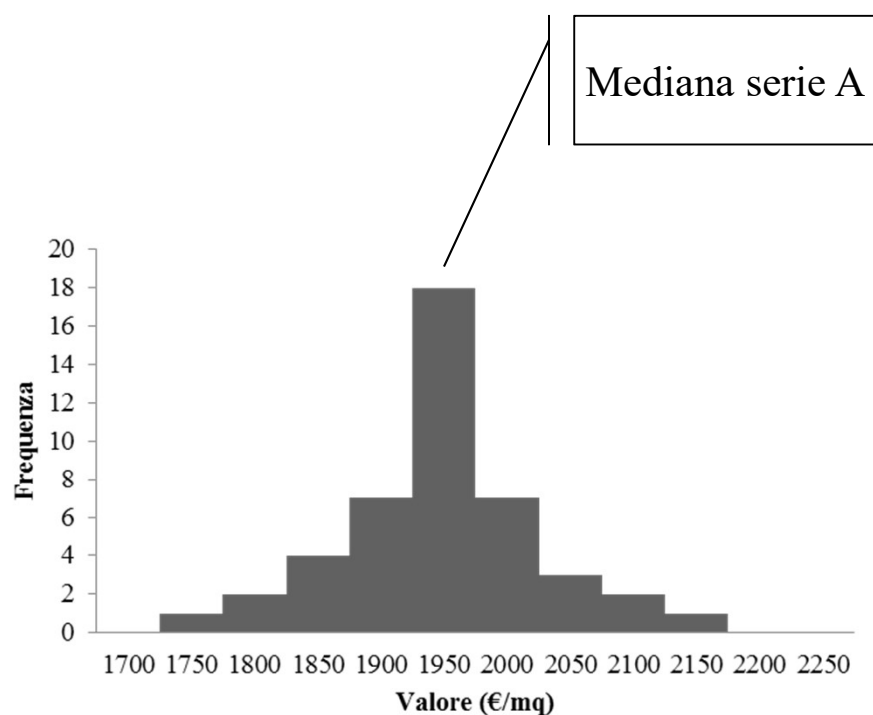
# Alcuni semplici indicatori statistici di tendenza centrale

**Moda o norma ( $M_d$ ):** Data una distribuzione è il valore ( $V$ ), o classe di valori più frequente



# Alcuni semplici indicatori statistici di tendenza centrale

**Mediana ( $M_e$ ):** Data una serie di valori ( $V_i$ ), il valore mediano è quel valore che divide a metà la serie



## Alcuni semplici indicatori statistici di tendenza centrale

Indicatore	Distribuzione	
	Simmetrica	Asimmetrica
Media semplice	1.933,27	1.931,33
Media armonica	1.929,72	1.923,49
Media geometrica	1.931,49	1.927,36
Media quadratica	1.935,06	1.935,40
Moda	1.950,00	1.800,00
Mediana	1.930,00	1.900,00

# Alcuni semplici indicatori statistici di dispersione

**Devianza totale (DT):** sommatoria dei quadrati degli scostamenti dalla media

$$DT = \sum_{i=1}^n (V_i - M_s)^2$$

# Alcuni semplici indicatori statistici di dispersione

**Scostamento quadratico medio ( $S_{mq}$ ):** Rapporto fra la sommatoria dei quadrati degli scostamenti dalla media e il numero di osservazioni ( $n$ )

$$S_{mq} = \frac{\sum_{i=1}^n (V_i - M_s)^2}{n}$$



## Alcuni semplici indicatori statistici di dispersione

**Varianza ( $\sigma^2$ ):** Rapporto fra la sommatoria dei quadrati degli scostamenti dalla media e il numero di osservazioni ( $n$ ) meno 1

$$\sigma^2 = \frac{\sum_{i=1}^n (V_i - M_s)^2}{n - 1}$$

# Alcuni semplici indicatori statistici di dispersione

**Deviazione Standard ( $\sigma$ ):** Radice quadrata della Varianza

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (V_i - M_s)^2}{n-1}}$$

# Alcuni semplici indicatori statistici di dispersione

**Coefficiente di variazione ( $\gamma$ ):** Rapporto fra deviazione standard ( $\sigma$ ) e la media ( $M_s$ )

$$\gamma = \frac{\sqrt{\frac{\sum_{i=1}^n (V_i - M_s)^2}{n-1}}}{M_s}$$

## Alcuni semplici indicatori statistici di dispersione

Indicatore	Distribuzione	
	Simmetrica	Asimmetrica
Scost. quadratico medio	6.921,73	15.738,22
Varianza	7.160,41	16.280,92
Deviazione standard	84,62	127,60
Coefficiente di variazione	0,04	0,07

# Alcuni semplici indicatori statistici di dispersione

## I quartili

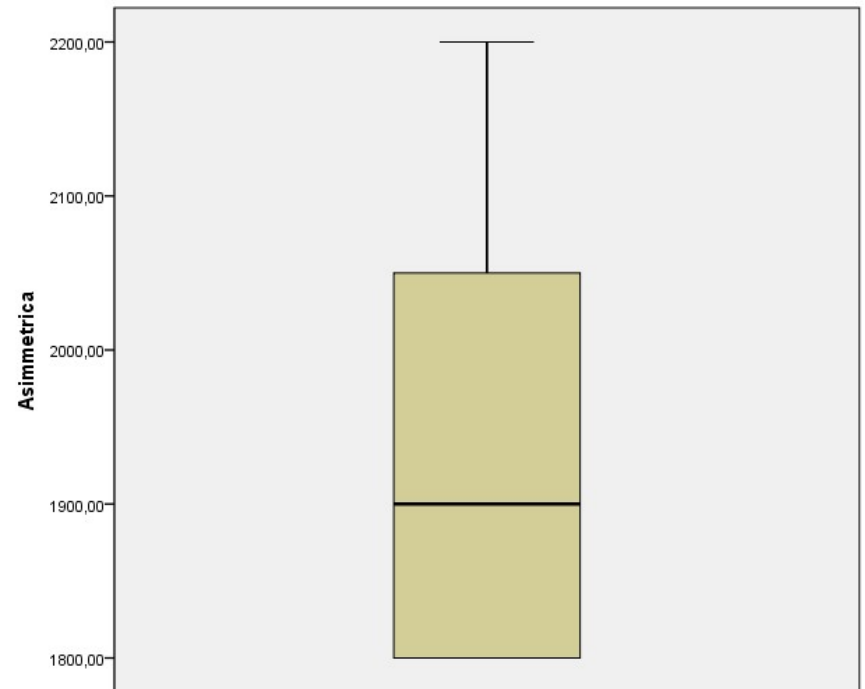
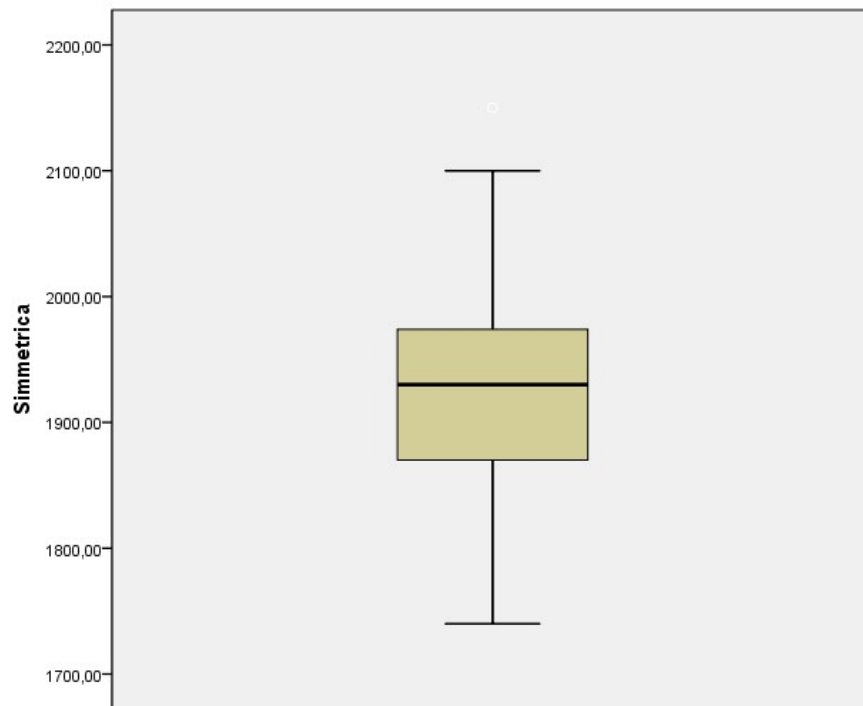
Valori che ripartiscono una serie ordinata di dati in quattro sottoinsiemi di uguale numerosità

Indicatore	Distribuzione	
	Simmetrica	Asimmetrica
Minimo	1.740,0	1.800,0
<b>Primo quartile</b>	<b>1.877,5</b>	<b>1.812,5</b>
<b>Secondo quartile (mediana)</b>	<b>1.930,0</b>	<b>1.900,0</b>
<b>Terzo quartile</b>	<b>1.740,0</b>	<b>1.800,0</b>
Massimo	2.150,0	2.200,0

# Alcuni semplici indicatori statistici di dispersione

## Il grafico Box-Plot

Rappresentazione grafica dei valori dei quartili



# Gli indicatori statistici di forma

## La distribuzione normale (gaussiana)

$$F(V_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(V_i - M_s)^2 / \sigma^2}$$

$F(V_i)$  = Frequenza con cui si rileva un certo valore  $V_i$

$M_s$  = Media;

$\sigma$  = Deviazione standard.

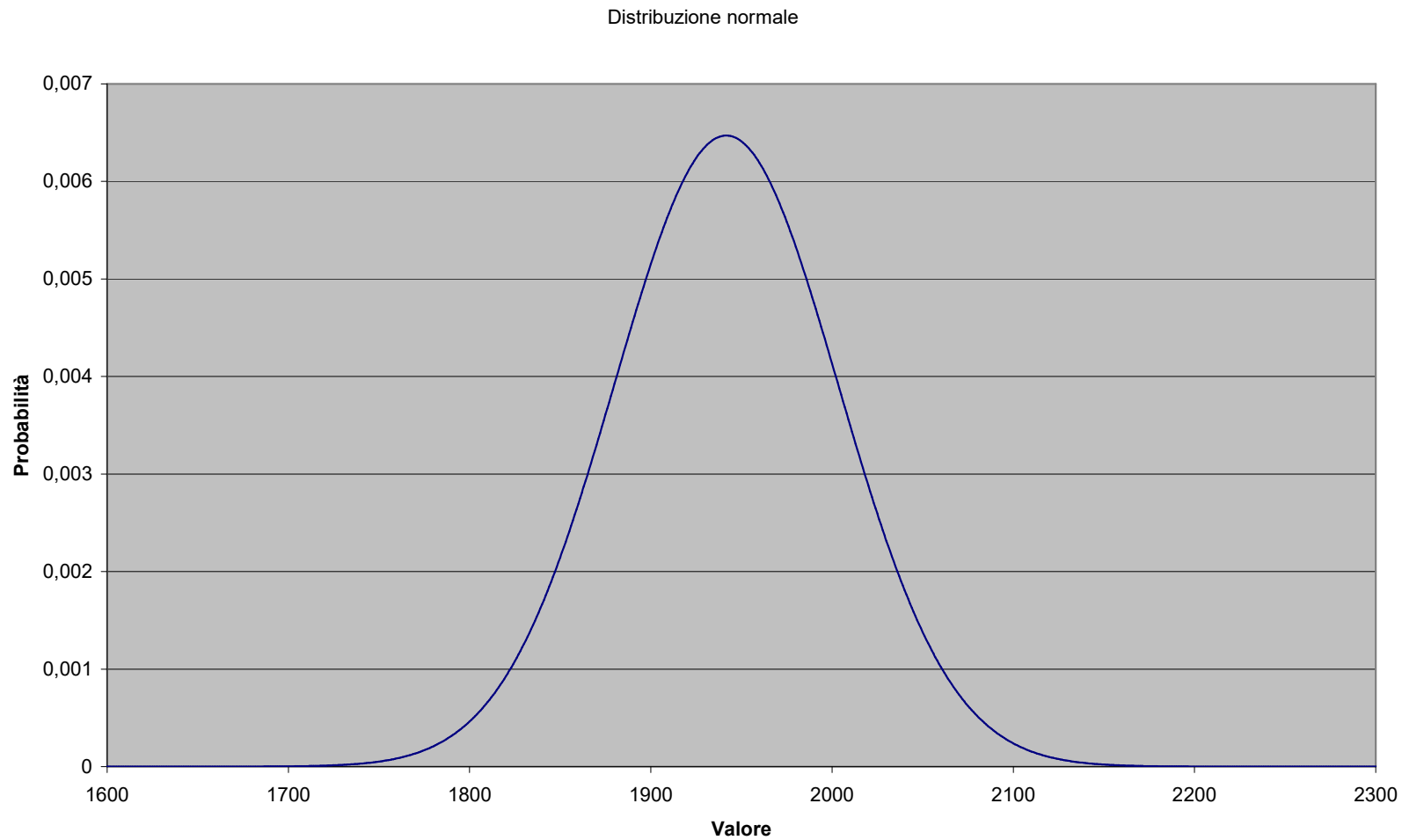
## Gli indicatori statistici di forma

La distribuzione normale (gaussiana) che approssima  
i dati dell'esempio

$$F(V_i) = 0,004715 \cdot e^{-\frac{1}{2}(V_i - 1.933,27)^2 / 7169,41}$$

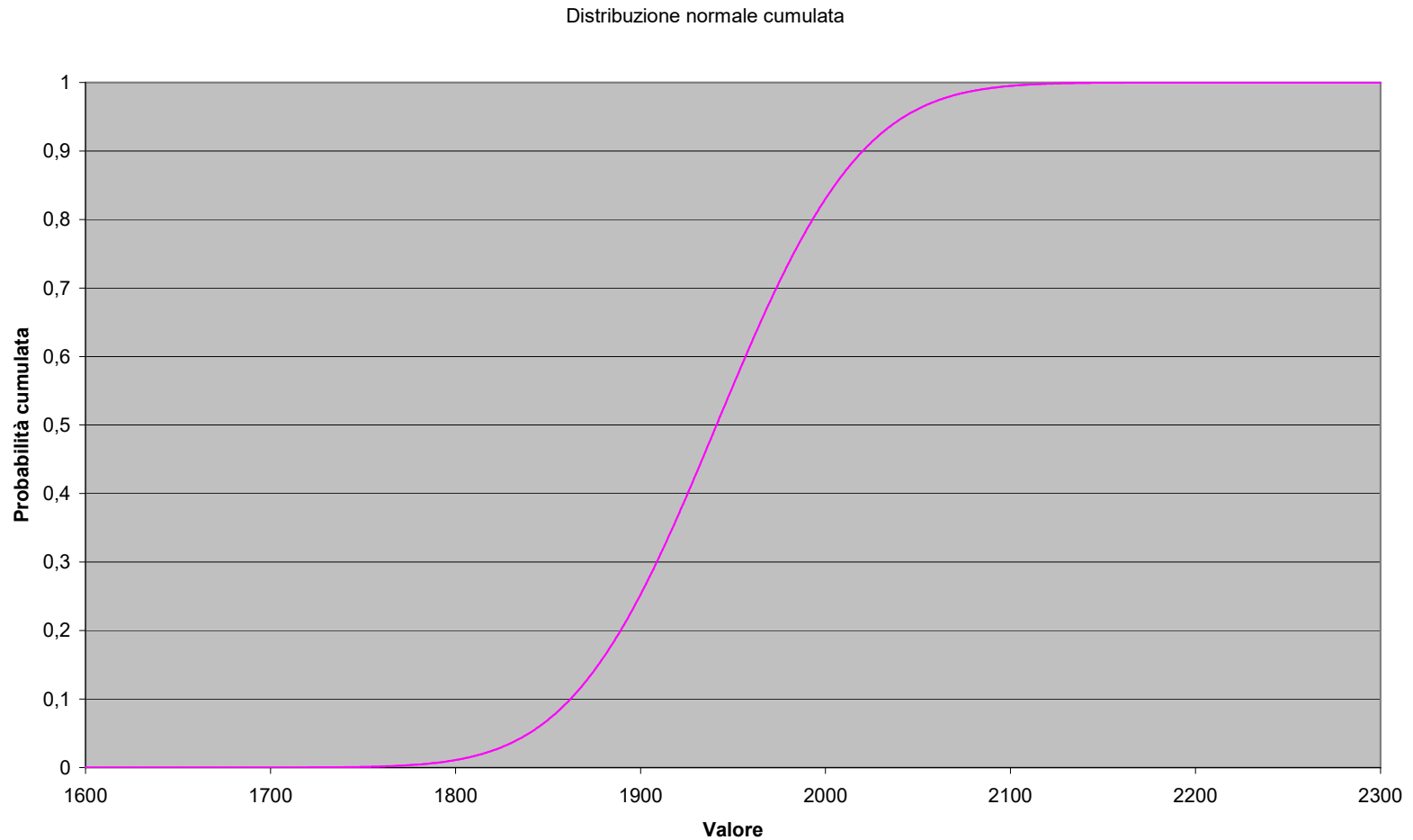


# La distribuzione normale (gaussiana)

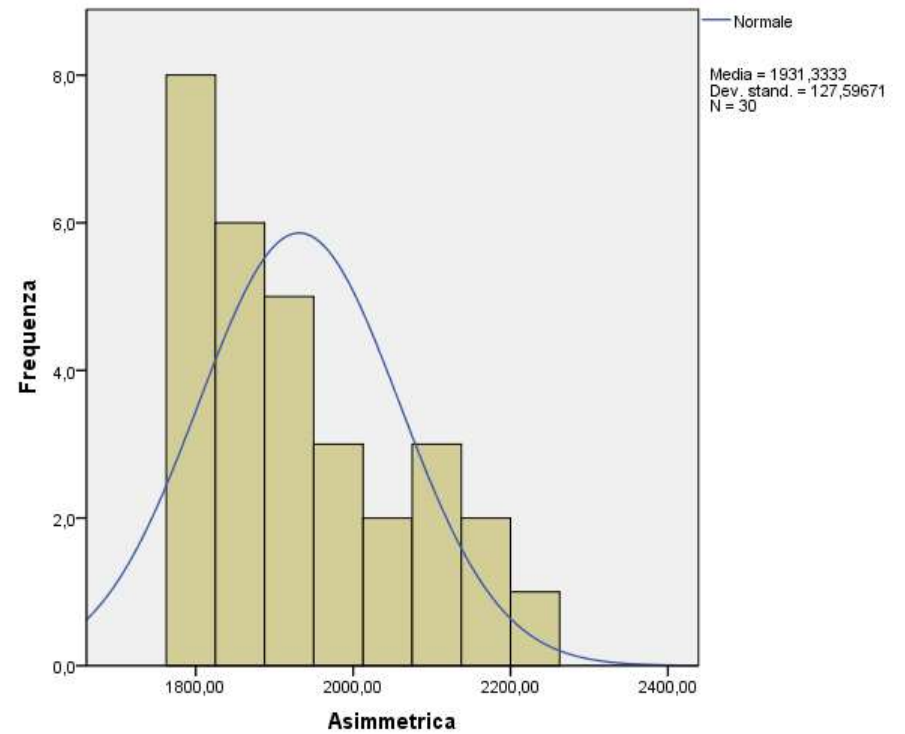
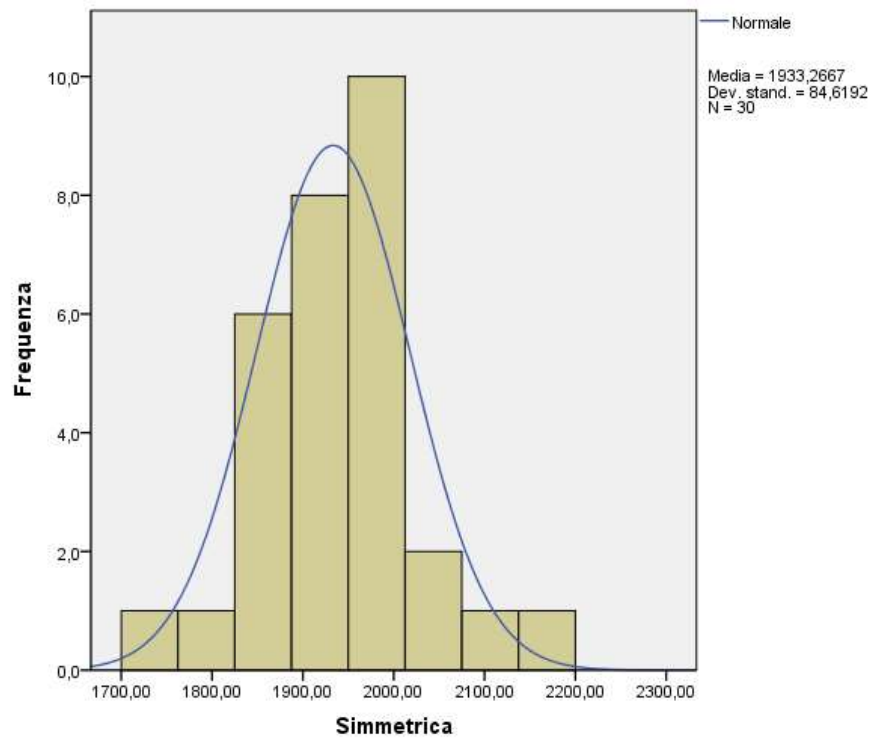


	Da	A	% casi
$M_s \pm \sigma$	1803,74	2058,93	68,27
$M_s \pm 2\sigma$	1676,14	2186,53	95,45
$M_s \pm 3\sigma$	1548,54	2314,12	99,73

# La distribuzione normale cumulata (Funzione di ripartizione)



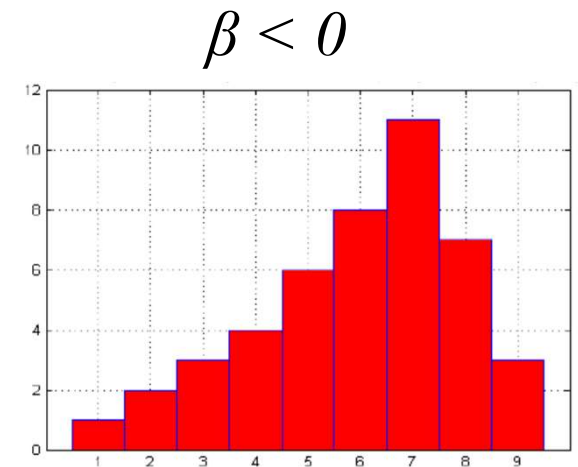
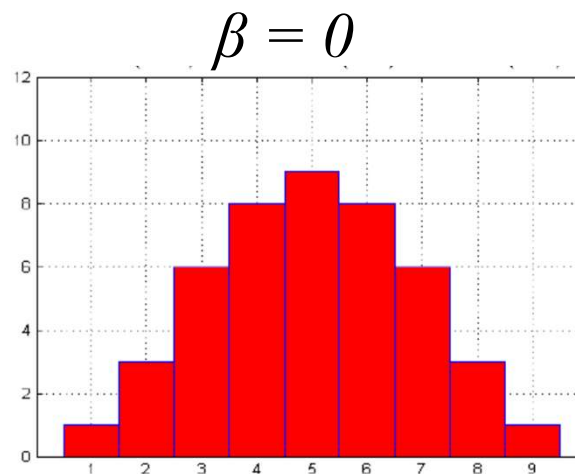
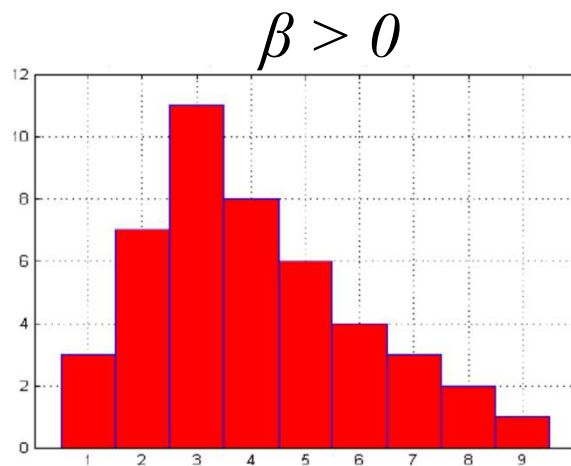
# La distribuzione normale e reale



# Alcuni semplici indicatori statistici di forma

**Indice di asimmetria (skewness):**  $\beta$  di Fisher normalizzato e corretto per la numerosità: positivo coda asimmetrica verso valori più alti, negativo: coda asimmetrica verso i valori più bassi.

$$\beta = \frac{1}{(n-1) \cdot (n-2)} \sum_i \left( \frac{V_i - M_s}{\sigma} \right)^3$$

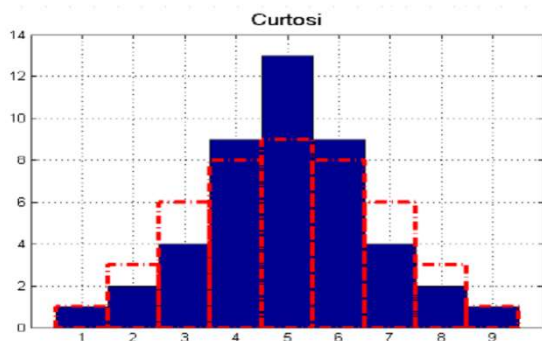


# Alcuni semplici indicatori statistici di forma

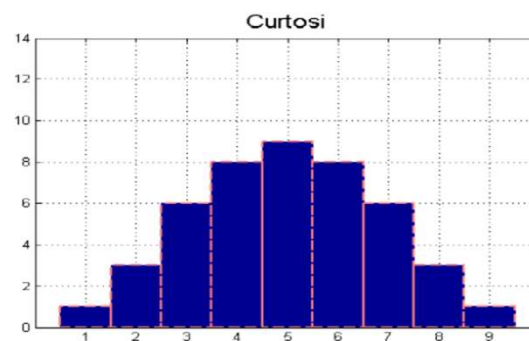
**Indice di curtosi:**  $k$  valuta il grado di adesione ad una distribuzione normale; positivo: concentrazione maggiore attorno alla media, negativo: concentrazione maggiore sulle code.

$$k = \left\{ \frac{n \cdot (n-1)}{(n-1) \cdot (n-2) \cdot (n-3)} \sum_i \left( \frac{V_i - M_s}{\sigma} \right)^4 \right\} - \frac{3 \cdot (n-1)^2}{(n-2) \cdot (n-3)}$$

$k > 0$



$k = 0$



$k < 0$

