LE MEDIE

ESEGUITA LA RILEVAZIONE STATISTICA SI PASSA ALL' ELABORAZIONE DEI DATI SECONDA FASE DI UN INDAGINE STATISTICA OSSIA, RILEVAZIONE-ELABORAZIONE-INTERPRETAZIONE ED PRESENTAZIONE. ELABORARE I DATI SIGNIFICA RICAVARE DALLA MASSA DELLE INFORMAZIONI GLI ELEMENTI CARATTERISTICI, TRATTI ESSENZIALI DEI FENOMENI OGGETTO DI STUDIO

MEDIE

1. **MEDIA ARITMETICA**
2. **MEDIA GEOMETRICA**
3. **MEDIA ARMONICA**
4. **MEDIA DI POTENZA**
5. **MEDIA QUDRATICA**

MEDIA ARITMETICA SEMPLICE

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

La media aritmetica è data dal rapporto tra la somma delle singole modalità e il numero complessivo della popolazione

CALCOLO MEDIA ARITMETICA SEMPLICE

CONSIDERIAMO UNA SEMPLICE DISTRIBUZIONE

VOTO DI STATISTICA
20-21-22-23-24-25-26-27=188

COLLETTIVO N=8
MEDIA ARITMETICA SEMPLICE =188/8=23,5

MEDIA ARITMETICA PONDERATA

$$\overline{X} = \mu = \frac{\sum_{i=1}^{S} x_i n_i}{N}$$

La media aritmetica ponderata si ottiene facendo il rapporto tra la somma del prodotto delle singole modalità con le rispettive frequenze ed il numero complessivo della popolazione

ESEMPIO MEDIA ARITMETICA PONDERATA

VOTO ESAME STATISTICA X_i	FREQUENZE ASSOLUTE n_i	PRODOTTO $x_i n_i$
24	10	240
25	13	325
26	18	468
27	20	540
totale	61	1573

CALCOLO MEDIA ARTMETICA PONDERATA

$$\overline{X} = \mu = \frac{\sum_{i=1}^{S} x_i n_i}{N} = \frac{1573}{61} = 25$$

MEDIA ARITMETICA PONDERATA CON VARIABILI IN CLASSI MODALI

CLASSI REDDITO	n_i	Valore centrale $x_i^{'} = \frac{x_i + x_{i+1}}{2}$	$x_i^{'} n_i$
10-20	10	15	150
21-30	20	25	500
31-40	30	35	1050
41-50	10	45	450
TOTALE	70		2150

CALCOLO MEDIA ARITMETICA

$$\overline{X} = \mu = \frac{\sum_{i=1}^{S} x'_i n_i}{N} = \frac{2150}{70} = 30$$

PROPRIETA DELLA MEDIA ARITMETICA

SOMMA ALGEBRICA DEGLI SCARTI E UGUALE A ZERO

$$(x_i - \mu)$$

ESEMPIO PROPRIETA' MEDIA ARITMETICA

Voto	$(x_i - \mu)$
26	26-27= -1
27	27-27= 0
28	28-27= 1
totale	0

PROPRIETA DELLA MEDIA ARITMETICA

LA SOMMA DEGLI SCARTI AL QUADRATO E' UN MINIMO

$$(x_i - \mu)^2 = \min$$

CONSIDERIAMO SOLO TRE VALORI 20,21,22

MEDIA 21

$(20\text{-}21)^2+(21\text{-}21)^2+(22\text{-}21)^2=(-1)^2+(0)^2+(1)^2=2$

PROVIAMO A SOSTITUIRE UN VALORE DIVERSO DAL VALORE MEDIO OSSIA IL PIU BASSO VALORE ZERO

$(20\text{-}0)^2+(21\text{-}0)^2+(22\text{-}0)^2=400+441+484=1325$

PROPRIETA DELLA MEDIA ARITMETICA

LA MEDIA ARITMETICA E' INTERNA SIGNIFICA CHE E' MAGGIORE DEL PIU PICCOLOE MINORE DEL PIU GRANDE

$$X_1 < \mu < X_N$$

PROPRIETA DELLA MEDIA ARITMETICA

LA MEDIA ARITMETICA E ASSOCIATIVA, IN QUANTO,SUDDIVIDENDO IN DUE O PIU' GRUPPI I VALORI DELLA VARIABILE X,LA MEDIA ARITMETICA DELLA VARIABILE E' UGUALE ALLA MEDIA ARITMETICA DELLE MEDIE PARZIALI DEI DIVERSI GRUPPI PONDERATI CON IL NUMERO DI ELEMENTI DI CIASCUNO

ESEMPIO PROPRIETA ASSOCIATIVA

MEDIA GENERALE

20+21+22+23+24+25+26+27+28=24

MEDIA PARZIALE Primo gruppo

20+21+22+23=86/4=21,5

MEDIA PARZIALE Secondo Gruppo

24+25+26+27+28=130/5=26

Verifica Proprietà associativa

21,5X4+26x5=86+130=216/9=24

PROPRIETA DELLA MEDIA ARITMETICA

PROPRIETA' TRASLATIVA O UNIFORME SIGNIFICA CHE SE SI AGGIUNGE UNA QUANTITA' h A CIASCUNO DEI VALORI ANCHE LA MEDIA RISULTERA' AUMENTATA DI h

PROPRIETA DELLA MEDIA ARITMETICA

PROPRIETA' OMOGENEA SE SI MOLTIPLICA CIASCUNO DEI VALORI PER UNA QUANTITA k ANCHE LA MEDIA RISULTERA' MOLTIPLICATA PER k.

PROPRIETA DELLA MEDIA ARITMETICA

SE LE x_i SONO IN PROGRESSIONE ARITMETICA E SE N E' UN NUMERO DISPARI LA MEDIA ARITMETICA COINCIDE CON IL TERMINE CHE OCCUPA LA POSIZIONE CENTRALE NELLA GRADUATORIA DEI VALORI ORDINATI

ES PROGRESSIONE ARITMETICA

5-7-9-11-13-15-17-19

MEDIA GEOMETRICA

LA MEDIA GEOMETRICA SI APPLICA QUANDO I DATI SONO IN PROGRESSIONE GEOMETRICA

ES.1-2-4-8-16-32-64

MEDIA GEOMETRICA

LA MEDIA GEOMETRICA DI UN INSIEME DI N VALORI POSITIVI $x_1, x_2, x_3, \ldots, x_n$ DI UN CARATTERE QUANTITATIVO X E' PARI ALLA RADICE n-ESIMA DEL PRODOTTO DEI SINGOLI VALORI

$$M_g = \sqrt[n]{x_1 \bullet x_2 \bullet x_3 \ldots\ldots x_n}$$

ASSENZA FREQUENZE

$$\overline{X} = M_g = \sqrt[N]{\prod_{i=1}^{N} x_i}$$

CON LE FREQUENZE

$$\overline{X} = M_g = \sqrt[N]{\prod_{i=1}^{N} x_i^{n_i}}$$

CALCOLO MEDIA GEOMETRICA

Voto 20, 21, 22, 23, 24

$$\sqrt[5]{20x21x22x23x24} = 22$$

CALCOLO MEDIA GEOMETRICA CON FREQUENZE

X_i	n_i
20	2
21	3
22	4
23	2
24	1
TOTALE	12

$$\sqrt[12]{20^2 * 21^3 * 22^4 * 23^2 * 24^1}$$

MEDIA ARMONICA

La media armonica si usa quando i reciproci dei termini sono in progressione aritmetica

MEDIA ARMONICA SEMPLICE

$$M_{ar} = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}}$$

Voto
20,21,22,23,24

$$Marm = \frac{5}{\frac{1}{20} + \frac{1}{21} + \frac{1}{22} + \frac{1}{23} + \frac{1}{24}} = 22$$

MEDIA ARMONICA CON FREQUENZE

$$Marm = \frac{N}{\sum_{i=1}^{s} \frac{n_i}{x_i}}$$

Esempio ....

Voto	n_i
20	2
21	3
22	6
23	2
24	1

Esempio media armonica …

$$Marm = \frac{N}{\sum_{i=1}^{s} \frac{n_i}{x_i}} = \frac{14}{\frac{2}{20}+\frac{3}{21}+\frac{6}{22}+\frac{2}{23}+\frac{1}{24}} = 22$$

Esempio media armonica…

Durata Quaderno Giorni X	Alunni n_i
8	26
10	35
12	15
15	17
Totale	93

Esempio media armonica….

$$Marm = \frac{N}{\sum_{i=1}^{s} \frac{n_i}{x_i}} = \frac{93}{\frac{26}{8} + \frac{35}{10} + \frac{15}{12} + \frac{17}{15}} = \frac{93}{9{,}13} = 10$$

MEDIA QUADRATICA

LA MEDIA QUADRATICA È USATA IN PRESENZA DI VALORI POSITIVI E NEGATIVI E SI VOGLIONO ELIMINARE I SEGNI.

MEDIA QUADRATICA SEMPLICE E PONDERATA

$$M_q = \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N}} \qquad M_q = \sqrt{\frac{\sum_{i=1}^{s} x_i^2 n_i}{N}}$$

RICORDA

$$x_1 \leq \ldots \leq Mar \leq Mgeom \leq \mu \leq Mq \leq \ldots X_N$$

MEDIE LASCHE

1. VALORE CENTRALE
2. MEDIANA
3. MODA

VALORE CENTRALE

$$V.C. = \frac{x_{(1)} + x_{(N)}}{2}$$

MODA

La Moda o valore normale è la modalità del carattere cui corrisponde la massima frequenza o densità di frequenza.

ESEMPI MODA

VARIABILI STATISTICHE DISCRETE

VOTO	FREQUENZA n_i
25	3
26	2
27	8
28	1

MODA

ESEMPI MODA – V.S CONTINUE

AMPIEZZE CLASSI UGUALI

Voto	N_i
18-20	3
21-23	5
24-26	10 moda
27-29	4

AMPIEZZE CLASSI DIFFERENTI

Voto	n_i	d_i	$H_i=n_i/d_i$
18-21	5	3	5/3=1,6
21-23	4	2	4/2=2
24-28	6	4	6/4=1,5
29-30	3	1	3/1=3 Moda

MODA

Per CARATTERI DISCRETI la moda si individua facilmente scorrendo lungo la colonna delle frequenze.

Per CARATTERI CONTINUI se le classi di modalità hanno tutti uguale ampiezza, la moda cade nella classe con maggiore frequenza.

Se le classi di modalità hanno diversa ampiezza, la moda cade nella classe con maggiore densità di frequenza.

MEDIANA

LA MEDIANA E' UN VALORE CHE BIPARTISCE LA DISTRIBUZIONE ORDINATA IN SENSO CRESCENTE DELLE MODALITA' DI UN CARATTERE, OSSIA LA MEDIANA E' QUEL VALORE CHE ASSUME LA POSIZIONE CENTRALE DELLA DISTRIBUZIONE.

MEDIANA PER DATI DISCRETI N = DISPARI

$$Me = x_{\left(\frac{N+1}{2}\right)}$$

Il valore a pedice della x indica la posizione della modalità che divide in due parti uguali la distribuzione

20 21 22 23 24

N=5 $Me = x_{(5+1/2)} = x_3 = \mathbf{22}$

MEDIANA PER DATI DISCRETI
N = PARI

$$Me = \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2}$$

20 21 22 23 24 25 N=6

Me =x $_{(6/2)}$ + x $_{(6/2+1)}$ /2=x $_{(3)}$ + x $_{(4)}$ /2=22+23/2=22,5

Mediana per dati discreti

Carattere X_i	Famiglie n_i	Frequenza cumulata N_i
1	10	10
2	20	30
3	50	80
4	10	90
5	10	100

Mediana per dati discreti

La frequenza cumulata è data dalla somma cumulata delle frequenze assolute.

N=Pari allora

N/2 Vi aiuta ad orientarvi sul valore modale che mi rappresenta la media ma dovete poi calcolare esattamente il valore mediano con la formula

$$Me = \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2} \qquad \mathrm{N} = PARI$$

$$Me = \frac{x_{\left(\frac{100}{2}\right)} + x_{\left(\frac{100}{2}+1\right)}}{2} = \frac{x_{(50)} + x_{(51)}}{2} = \frac{3+3}{2} = 3$$

Mediana per dati continui

Classi Modali $X_i - X_{i+1}$	Frequenze assolute n_i	Frequenze cumulate N_i
100-200	10	10
201-300	20	30
301-400	30	60
401-500	10	70
Totale	70	

Mediana per dati continui

La classe che contiene il valore mediano è quella a cui corrisponde una frequenza cumulata superiore ad N/2

N/2 vi aiuta ad individuare la classe che contiene il valore mediano

$$Med = X_i + \frac{X_{i+1} - X_i}{n_i}\left(\frac{N}{2} - N_{i-1}\right) =$$

$$301 + \frac{400-301}{30}\left(\frac{70}{2} - 30\right) = 316{,}5$$

Mediana per dati continui

X_i= estremo inferiore della classe modale che contiene il valore mediano

X_{i+1}=estremo superiore della classe modale che contiene il valore mediano

ni = frequenza assoluta della classe modale che contiene il valore mediano

N = frequenze complessive

N_{i-1}=frequenza cumulata in corrispondenza della classe modale che precede quella che contiene il valore mediano

Quartili e quantili

IL PRIMO QUARTILE Q1 DI UNA V.S. è QUEL VALORE AL DI SOTTO DEL QUALE STANNO ¼ DEI VALORI (25%) DELLA X E AL DI SOPRA DEL QUALE STANNO I ¾ (75%) DEI VALORI DELLA X
N/4 per trovare la classe modale che contiene il primo quartile

IL SECONDO QUARTILE COINCIDE CON LA MEDIANA N/2 (50%)

IL TERZO QUARTILE Q3 E' QUEL VALORE AL DI SOTTO DEL QUALE STANNO I ¾ (75%)E AL DI SOPRA DEL QUALE STANNO ¼ (25%) DEI VALORI DELLA X.

Quartili e quantili

$$Q_1 = x_i + \frac{x_{i+1} - x_i}{n_i} \left(\frac{N}{4} - N_{i-1}\right)$$

$$Q_2 = x_i + \frac{x_{i+1} - x_i}{n_i} \left(\frac{N}{2} - N_{i-1}\right)$$

$$Q_3 = x_i + \frac{x_{i+1} - x_i}{n_i} \left(\frac{3N}{4} - N_{i-1}\right)$$

ESEMPIO PRIMO QUARTILE

Classi di altezza	Numero atleti	Frequenza cumulata
171-175	12	12
176-180	16	28
181-185	25	53
186-190	23	76
191-195	20	96
TOTALE	96	

N/4=24

VEDI COLONNA Ni

$$Q_1 = x_i + \frac{x_{i+1} - x_i}{n_i}\left(\frac{N}{4} - N_{i-1}\right)$$

$$Q_1 = 171 + \frac{175 - 171}{12}(24 - 12) = 174{,}9$$

ESEMPIO TERZO QUARTILE

Classi di altezza	Numero atleti	Frequenza cumulata
171-175	12	12
176-180	16	28
181-185	25	53
186-190	23	76
191-195	20	96
TOTALE	96	

3N/4=72

VEDI COLONNA Ni

$$Q_3 = x_i + \frac{x_{i+1} - x_i}{n_i}\left(\frac{3N}{4} - N_{i-1}\right)$$

$$Q_3 = 185 + \frac{185 - 181}{25}(72 - 28) = 189{,}4$$

CALCOLO MEDIA, MODA E MEDIANA QUARTILI CON EXCELL ED SPSS

LA VARIABILITA DEFINIZIONE

LA VARIABILITA' INDICA LA TENDENZA DI OGNI SINGOLA OSSERVAZIONE AD ASSUMERE VALORI DIFFERENTI RISPETTO AL VALORE MEDIO.

ESEMPIO:

$$x_i - \mu$$

LA VARIABILITA'

OSSERVATE LE SEGUENTI DISTRIBUZIONI. SECONDO VOI VI E' VARIABILITA' O MEGLIO DISPERSIONE O DISUGUAGLIANZA?

VOTO = 22-22-22-22-22-22-22-22-22-

MEDIA = 22 MODA=22 MEDIANA =22

VARIABILITA'=?

VOTO = 20-21-22-23-24-25-26-27-28-

MEDIA =24 MODA=? MEDIANA= 24

VARIABILITA= OCCORRE CALCOLARLA

LA VARIABILITA'

INDICI DI VARIABILITA

Gli indici si distinguono in:

1. **Indici di variabilità assoluta, che sono espressi nella stessa unità di misura del fenomeno osservato**
2. **Indici di variabilità relativa, che prescindono dall'unità di misura.**

LA VARIABILITA’

INDICI DI VARIABILITA’ ASSOLUTA

1. CAMPO DI VARIAZIONE
2. SCARTO SEMPLICE MEDIO
3. SCARTO QUADRATICO MEDIO
4. VARIANZA
5. DEVIANZA

LA VARIABILITA’

INDICI DI VARIABILITA’ RELATIVA

1. Indice di variabilità relativo alla media

2. Concentrazione

LA VARIABILITA'
CAMPO DI VARIAZIONE

$$W = x_{\max} - x_{\min}$$

VOTO 20-21-22-23-24-25-26

W=26-20=6

LA VARIABILITA'
Scarto semplice medio

DISTRIBUZIONE SEMPLICE

$$\delta = \frac{\sum_{i=1}^{N} |x_i - \mu|}{N}$$

DISTRIBUZIONE PO NDERATA

$$\delta = \frac{\sum_{i=1}^{s} |x_i - \mu| n_i}{N}$$

LA VARIABILITA'
Scarto semplice medio- Semplice distribuzione

X_i VOTO	$\lvert x_i-\mu\rvert$
20	2
21	1
22	0
23	1
24	2
Totale	6

$$\delta=\frac{\sum_{i=1}^{N}\lvert x_i-\mu\rvert}{N}=\frac{6}{5}=1{,}2$$

Scarto semplice medio dato dal rapporto tra la somma degli scarti in valore assoluto ed il numero complessivo delle osservazioni

LA VARIABILITA'
Scarto semplice medio con le frequenze

X_i	n_i	$\lvert x_i-\mu\rvert$	$\lvert x_i-\mu\rvert n_i$
20	10	2	20
21	20	1	20
22	30	0	0
23	20	1	20
24	10	2	20
	90		80

$$\delta=\frac{\sum_{i=1}^{s}\lvert x_i-\mu\rvert n_i}{N}=\frac{80}{90}=1$$

Somma del prodotto degli scarti per le rispettive frequenze fratto il numero complessivo delle osservazioni

SCARTO SEMPLICE MEDIO PER VARIABILI IN CLASSI MODALI

CLASSI REDDITO	n_i	Valore centrale $x_i' = \frac{x_i + x_{i+1}}{2}$	$\left\|(x_i' - \mu)\right\|$	$\left\|(x_i' - \mu)\right\| n_i$
10-20	10	15	15-30=15	15*10=150
21-30	20	25	25-30=5	5*20=100
31-40	30	35	35-30=5	5*30=150
41-50	10	45	45-30=15	150
TOTALE	70			550

SCARTO SEMPLICE MEDIO

$$\delta = \frac{\sum_{i=1}^{s} \left|x_i - \mu\right| n_i}{N} = \frac{850}{70} = 12{,}1$$

DEFINIZIONE DI SCARTO QUADRATICO MEDIO

Lo scarto quadratico medio indica la media degli scarti di tutti gli intervistati in relazione ad un carattere statistico (voto , altezza ecc)

VARIABILITA'
SCARTO QUADRATICO MEDIO

DISTRIBUZIONE SEMPLICE

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

DISTRIBUZIONE CON FREQUENZE

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

VARIABILITA'
SCARTO QUADRATICO MEDIO

X_i	$(x_i - \mu)$	$(x_i - \mu)^2$
20	**-2**	**4**
21	**-1**	**1**
22	**0**	**0**
23	**1**	**1**
24	**2**	**4**
Totale		**10**

$$\sigma = \sqrt{\frac{\sum_{i=1}^{S}(x_i-\mu)^2}{N}} = \sqrt{\frac{10}{5}} = \sqrt{2} = 1{,}4$$

Radice quadrata della somma degli scarti al quadrato diviso il numero complessivo delle osservazioni

VARIABILITA'
SCARTO QUADRATICO MEDIO

X_i	n_i	$(x_i - \mu)$	$(x_i - \mu)^2$	$(x_i - \mu)^2 n_i$
20	1	-2	4	4
21	2	-1	1	2
22	3	0	0	0
23	2	1	1	2
24	1	2	4	4
tot	9			12

$$\sigma = \sqrt{\frac{\sum_{i=1}^{S}(x_i-\mu)^2 n_i}{N}} = \sqrt{\frac{12}{9}} = \sqrt{1{,}2} = 1$$

SCARTO SEMPLICE MEDIO PER VARIABILI IN CLASSI MODALI

CLASSI REDDITO	n_i	Valore centrale $x_i' = \frac{x_i + x_{i+1}}{2}$	$(x_i' - \mu)$	$(x_i' - \mu)^2$	$(x_i' - \mu)^2 n_i$
10-20	10	15	-16	256	2560
21-30	20	25,5	-5,5	30	600
31-40	30	35,5	4,5	20	600
41-50	10	45,5	14,5	210	2100
TOTALE	70				5865

Scarto quadratico medio

$$\sigma = \sqrt{\frac{\sum_{i=1}^{s}(x_i - \mu)^2 n_i}{N}} = \sqrt{\frac{5865}{70}} = \sqrt{83,78} = 9,1$$

VARIABILITA'
VARIANZA

Elevando al quadrato lo scarto quadratico medio si ottiene la varianza

La varianza indica la misura della grandezza della variabilità del carattere osservato ma di soliti si usa lo scarto quadratico medio

VARIABILITA'
VARIANZA

DISTRIBUZIONE SEMPLICE

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

DISTRIBUZIONE PONDERATA

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

La varianza

La varianza presenta un notevole inconveniente: è espressa nel quadrato dell'unità di misura delle osservazioni, ad esempio, se le osservazioni sono espresse in metri, la varianza è espressa in metri al quadrato e per ovviare a ciò si utilizza lo scarto quadratico medio che rappresenta la radice quadrata della varianza.

VARIABILITA'

VARIANZA

X_i	$(x_i - \mu)$	$(x_i - \mu)^2$
20	-2	4
21	-1	1
22	0	0
23	1	1
24	2	4
Totale		10

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} = \frac{10}{5} = 2$$

VARIABILITA'
VARIANZA

X_i	n_i	$(x_i - \mu)$	$(x_i - \mu)^2$	$(x_i - \mu)^2 n_i$
20	1	-2	4	4
21	2	-1	1	2
22	3	0	0	0
23	2	1	1	2
24	1	2	4	4
tot	9			12

$$\sigma^2 = \frac{\sum_{i=1}^{s} (x_i - \mu)^2 n_i}{N} = \frac{12}{9} = 1,2$$

VARIABILITA'
DEVIANZA

La devianza è data dal numeratore della varianza

DISTRIBUZIONE SEMPLICE $\rightarrow$ $Dev(x) = \sum (x_i - \mu)^2$

DISTRIBUZIONE PONDERATA $\rightarrow$ $Dev(x) = \sum (x_i - \mu)^2 n_i$

DIFFERENZE MEDIE

MISURANO LE DIFFERENZE IN VALORE ASSOLUTO TRA LE MODALITA' A DUE A DUE INOLTRE MISURANO DI QUANTO LE DIVERSE QUANTITA' RILEVATE DIFFERISCONO TRA DI LORO IN MODO DA AVERE UNA MISURA DELLA DISUGUAGLIANZA MEDIA TRA I TERMINI

$$\Delta = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \left| x_i - x_j \right|}{n^2}$$

Assume valori minimo con dati tutti uguali

INDICI DI VARIABILIT' RELATIVA

$$Vr = \frac{V_a}{\mu}$$

Indice variabilità relativo alla media

$$0 \leq V'_r \leq 1$$

$$V'r = \frac{V_a}{MaxV_a}$$

Indice variabilità relativo al massimo

$$CV = \frac{\sigma}{\mu} 100$$

MASSIMO DELLA VARIABILITA'

Si ha minimo della variabilità nel caso di distribuzione con modalità identiche. Vediamo come si calcola il massimo della variabilità

$$Max\delta = \frac{2[\mu - x_{(1)}][x_{(N)} - \mu]}{x_{(N)} - x_{(1)}}$$

$$\text{Max}\sigma = \sqrt{[\mu - x_{(1)}][x_{(N)} - \mu]}$$

$$\text{Max}\Delta = \frac{2N[\mu - x_{(1)}[x_{(N)} - \mu]]}{(N-1)[x_{(N)} - x_{(1)}]}$$

VARIABILITA'

INDICI DI VARIABILITA' RELATIVA-CONCENTRAZIONE

LA CONCENTRAZIONE SI CALCOLA PER I CARATTERI CHE GODONO DELLA PROPRIETA' DELLA TRASFERIBILITA'

PER CALCOLAE LA CONCENTRAZIONE OCCORRE ORDINARE I DATI IN SENSO CRESCENTE

LA CONCENTRAZIONE

Osservati n valori ordinati di una variabile X,

$$x_1 \le x_2 \le \ldots \le x_n$$

si è interessati a studiare come l'ammontare del carattere

$$A = \sum_{i=1}^{n} x_i$$

sia ripartito fra le diverse unità statistiche.

Si possono avere due situazioni estreme:

- **equidistribuzione**
- **massima concentrazione**

LA CONCENTRAZIONE

- **Equidistribuzione: ognuna delle n unità possiede 1/n dell'ammontare complessivo del carattere**
- **Massima concentrazione: l'intero ammontare del carattere è posseduto da una sola unità**

VARIABILITA’

INDICI DI VARIABILITA’ RELATIVA-CONCENTRAZIONE

$$R = \frac{\sum_{i=1}^{N-1} p_i - q_i}{\sum_{i=1}^{N} p_i}$$

$$R = 0 \quad Min$$

$$R = 1 \quad Max$$

p_i = Ni/N **frazione degli i redditieri**

q_i = A_i/A_N **quota di reddito posseduto da ogni singolo redditiere**

i = **numerazione dei redditieri**

N = **Numero complessivo intervistati**

A_i = **frequenza cumulata di reddito**

VARIABILITA’

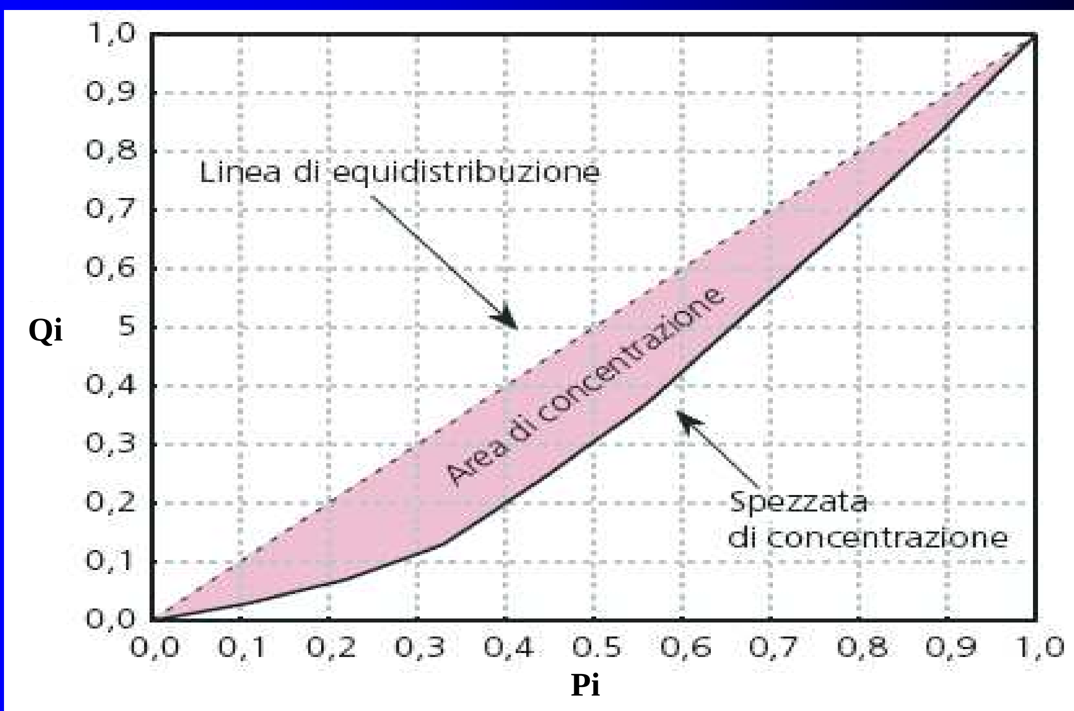
INDICI DI VARIABILITA’ RELATIVA-CONCENTRAZIONE

X_i	n_i	N_i	$p_i = N_i/N$	A_i	$q_i = A_i/A_n$	$p_i - q_i$
100	**1**	1	p_1=1/5=0,20	100	**q1=100/1500=0,06**	**0,20-0,06=0,14**
200	**1**	2	p_2=2/5=0,4	300	**q2=300/1500=0,2**	**0,4-0,2=0,2**
300	**1**	3	p_3=3/5=0,6	600	**q3=600/1500=0,4**	**0,6-0,4=0,2**
400	**1**	4	p_4=4/5=0,8	1000	**q4=1000/1500=0,6**	**0,8-0,6=0,2**
500	**1**	5	p_5=5/5=1	1500 0	**Q5=1500/1500=1**	**0**
	5		3			**0,74**

VARIABILITA'
INDICI DI VARIABILITA' RELATIVA-CONCENTRAZIONE

$$R = \frac{\sum_{i=1}^{N-1} p_i - q_i}{\sum_{i=1}^{N} p_i} = \frac{0,74}{3} = 0,2$$

VARIABILITA'
INDICI DI VARIABILITA' RELATIVA-CONCENTRAZIONE



CALCOLO INDICI DI VARIABILITA' CON EXCELL ED SPSS

ANALISI UNIVARIATA

ANALISI BIVARIATA

ANALISI MULTIVARIATA

ANALISI STATISTICA UNIVARIATA

UNA VARIABILE ANALIZZATA SINGOLARMENTE:

ETA, ALTEZZA, PESO

TECNICHE STATISTICHE:

1. **Medie**
2. **Variabilità**
3. **Numeri indici**

ANALISI STATISTICA BIVARIATA

DUE VARIABILI ANALIZZATE SIMULTANEAMENTE:
ALTEZZA E PESO

TECNICHE STATISTICHE:

1. CORRELAZIONE
2. REGRESSIONE
3. TEST CHI QUADRATO

ANALISI STATISTICA MULTIVARIATA

TECNICHE STATISTICHE:

1. Analisi fattoriale
2. Cluster analysis
3. Scaling Multidimensional
4. Regressione multipla
5. Correlazione canonica
6. Analisi corrispondenze

INDICE DI DISUGUAGLIANZA

INDICE PER MISURARE LA DIFFERENZA IN UNA SUCCESSIONE DI VALORI

$$M_{|x_1 - x_2|} = \frac{\sum_{i=1}^{N} |x_{1i} - x_{2i}|}{N}$$

INDICE DI DISUGUAGLIANZA

ESEMPIO VOTO ESAME STATISTICA E INFORMATICA

1°=(25,27) 2° =(27,24) 3°=(30,30) 4°=(26,24)
5°=(27,23) 6°=(28,25)

$$M = \frac{|25-27| + |27-24| + |30-30| + |26-24| + 2|7-23| + |28-25|}{6} =$$

$$= \frac{2+3+0+2+4+3}{6} = \frac{14}{6} = 2$$

Con questo indice si misura la differenza all'interno di ogni coppia per ogni intervistato ossia all'interno di ogni coppia vi è una differenza di 2 voti tra esame statistica ed esame informatica infatti vedete 25-27 e questo solo per ogni intervistato

INDICE DI DISSOMIGLIANZA

$$D_{|x_1 - x_2|} = \frac{\sum_{i=1}^{N} \left| x_{1(i)} - x_{2(i)} \right|}{N}$$

(i)= valori ordinati

1°=(25,27)	**25**	**23**
2° =(27,24)	**26**	**24**
3°=(30,30)	**27**	**24**
4°=(26,24)	**27**	**25**
5°=(27,23)	**28**	**27**
6°=(28,25)	**30**	**30**

$$D = \frac{|25-23| + |26-24| + |27-24| + |27-25| + |28-27| + |30-30|}{6} = \frac{10}{6} = 1{,}6$$

Definizione indice di dissomiglianza

Indice di dissomiglianza misura la differenza per tutti gli intervistati e non per ogni singolo intervistati come nel caso precedente ossia tutti gli intervistati hanno una differenza di 1 voto

CORRELAZIONE E REGRESSIONE

CORRELAZIONE → ANALISI INTERDIPENDENZA TRA DUE V.S.

REGRESSIONE → ANALISI DIPENDENZA TRA DUE V.S

CORRELAZIONE

$$r = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2 \sum_{i=1}^{N}(y_i - \overline{y})^2}}$$

+1=concordanza

0=

-1=discordanza

$$+1 \le r \le -1$$

X Pane	Y Pasta	$(x_i - \overline{x})$	$(y_i - \overline{y})$	$(x_i - \overline{x})(y_i - \overline{y})$	$(x_i - \overline{x})^2$	$(y_i - \overline{y})^2$
10	30	-20	-20	400	400	400
20	40	-10	-10	100	100	100
30	50	0	0	0	0	0
40	60	10	10	100	100	100
50	70	20	20	400	400	400
				1000	1000	1000

$$r = \frac{1000}{\sqrt{1000 * 1000}} = 1$$

REGRESSIONE

ANALISI DELLA DIPENDENZA DELLA VARIABILE X IN FUNZIONE DELA VARIABILE Y

X = VARIABILE INDIPENDENTE

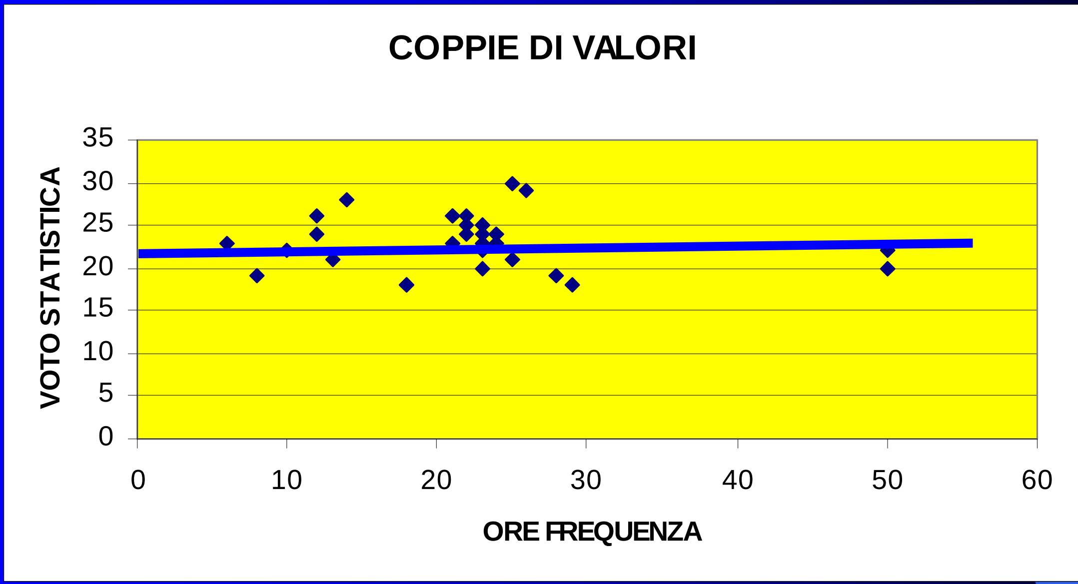
Y = VARIABILE DIPENDENTE

X = ORE FREQUENZA ALLE LEZIONI DI STATISTICA

Y = VOTO ESAME DI STATISTICA

Dati rilevati durante l'intervista

Intervistati	X Ore frequenza	Y Voto
1°	10	24
2°	20	25
3°	30	26
4°	40	27
5°	50	28
6°	60	?
7°	70	?



Individuare una relazione tra la nuvola di punti o scatter con una funzione matematica o meglio occorre individuare un modello che analizzi la relazione tra le due variabili

CERCARE UNA FUNZIONE CHE SPIEGHI Y IN FUNZIONE DI X

REGRESSIONE

La relazione tra i punti del grafico precedente la si definisce con la funzione retta di regressione e l'obiettivo è quello di stimare i due parametri incogniti della funzione retta a e b.

$$y = a + bx$$

REGRESSIONE

La stima dei parametri incogniti della funzione di regressione avviene con la tecnica dei minimi quadrati consiste nel rendere minima la differenza al quadrato tra valori teorici ed valori empirici

$$
\begin{cases}
a = \bar{y} - b\bar{x} \\
b = \dfrac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N} (x_i - \bar{x})^2}
\end{cases}
$$

B = Coefficiente di regressione

b>0

b<0

b=0

Coefficiente di regressione

- b maggiore di zero se aumenta la x aumenta la y
- Se b minore di 0 all'aumentare di X, y diminuisce
- Se b uguale a zero non vi è dipendenza

X Ore frequenza	Y Voto	$(x_i-\bar{x})$	$(y_i-\bar{y})$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$	$y=23+0{,}1x$
10	24	-20	-2	40	400	$y=23+0{,}1*10=24$
20	25	-10	-1	10	100	$y=23+0{,}1*20=25$
30	26	0	0	0	0	$y=23+0{,}1*30=26$
40	27	10	1	10	100	$y=23+0{,}1*40=27$
50	28	20	2	40	400	$y=23+0{,}1*50=28$
60	?					$y=23+0{,}1*60=29$
70	?					$y=23+0{,}1*70=30$
				100	1000	

$$y=23+0{,}1*x$$

REGRESSIONE

$$\begin{cases} a=\bar{y}-b\bar{x}=26-0{,}1*30=23 \\ b=\dfrac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{N}(x_i-\bar{x})^2}=\dfrac{100}{1000}=0{,}1 \end{cases}$$

$$y=23+0{,}1*x$$

CALCOLO REGRESSIONE E CORRELAZIONE CON EXCELL ED SPSS

TABELLE DI CONTINGENZA

Il test chi quadrato viene utilizzato per analizzare la relazione tra due variabile qualitative o una qualitativa ed una quantitativa

Dati categoriali e Tavole di contingenza
Test del chi quadrato

Istruzione	Destinazione Vacanza			Totale
	Nord America	Centro America	Sud America	
Universitaria	80	80	45	205
Non universitaria	387	161	84	632
Totale	467	241	129	837

Tavole di contingenza
Test del chi quadrato

Passo 1 Fissare Ipotesi nulla ed Ipotesi alternativa

H_0: non c'è associazione tra istruzione universitaria e la destinazione della vacanza

H_A:c'è associazione tra istruzione universitaria e la destinazione della vacanza

Tavole di contingenza
Test del chi quadrato

Passo 2 I conteggi nella tabella originale sono noti come VALORI OSSERVATI e dicono cosa è stato trovato durante l'indagine. Per il test del chi quadrato c'è la necessità di calcolare VALORI ATTESI O TEORICI. Questi sono i valori che ci aspettiamo di trovare in ciascuna casella se non ciu fosse associazione tra le due variabili.

Il test confronta i valori osservati e quelli attesi per vedere se sono differenti in modo significativo. Per calcolare i valori attesi per ciasscuna cella usimo la formula

$$\text{ValoreAtteso} = \frac{\text{Totale della colonna x totale della riga}}{\text{Totale generale}}$$

Per esempio per quelli che sono andati in nor america con istruzione universitaria il valore attes o teorico della cella è:

$$Valore\ \ atteso = \frac{467x205}{837} = 114{,}38$$

Valori attesi calcolati

Istruzione	Destinazione Vacanza			Totale
	Nord America	Centro America	Sud America	
Universitaria	467x205/887 = 114,38	241x205/837 = 59,03	129x205/837 = 31,59	205
Non Universitaria	467x632/837 = 352,62	241x632/837 = 181,97	129x632/837 = 97,41	632
Totale	467	241	129	837

Test chi quadrato Passo 3

Siamo ora pronti per calcolare il test del chi quadrato con la seguente formula:

$$\chi^2 = \sum \frac{(O-A)^2}{A}$$

O	A	(O-A)	$(O-A)^2$	$(O-A)^2/A$
80	114,38	-34,38	1181,98	10,33
80	59,03	20,97	439,74	7,45
45	31,59	13,41	179,83	5,69
387	353,62	34,38	1181,98	3,35
161	181,97	-20,97	439,74	2,42
84	97,41	-13,41	179,83	1,85
			Totale	31,09

Tavole di contingenza
Test del chi quadrato
Passo 4

Ora bisogna trovare un valore critico delle tavole chi quadrato. Per farlo si deve sapere il livello di significatività che vogliamo usare (di solito 5%) ed il numero dei gradi di libertà

Gl=(numero di righe-1)x(numero colonne-1)

In questo caso la tabella ha due righe e tre colonne per cui:

Gl=(2-1)x(3-1)=2

Consultando la tavola chi quadrato con gl=2 al livello di significatività del 5% troviamo il valore 5,99. questo è il valore critico del chia quadrato per questo test.

Tavole di contingenza
Test del chi quadrato
Passo 4

Passo 5 l'ipotesi nulla può essere scartata se il test statistico è piu grande del valore critico 5,99. Il test statistico che abbiamo trovato è di 31,09 che è molto piu grande del 5,99 per cui rifiiutima ipotesi nulla e accettiamo l'ipotesi alternativa quindi c'è relazione tra la destinazione delle vacanze e l'istruzione universitaria.

A.3 La distribuzione del χ^2

La Tavola A.3 fornisce i valori critici per la distribuzione del χ^2. La colonna di sinistra dà i valori dei gradi di libertà, mentre le restanti colonne danno i valori dei vari livelli di significatività, cioè la proporzione di area nella regione critica.

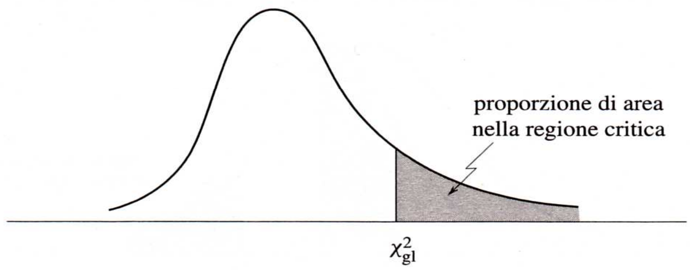


Tabella A.3 La distribuzione del χ^2.

gl	$\chi^2_{gl}(0,1)$	$\chi^2_{gl}(0,05)$	$\chi^2_{gl}(0,025)$	$\chi^2_{gl}(0,01)$	$\chi^2_{gl}(0,005)$
1	2,7055	3,8415	5,0239	6,6349	7,8794
2	4,6052	5,9915	7,3778	9,2103	10,5966
3	6,2514	7,8147	9,3484	11,3449	12,8382
4	7,7794	9,4877	11,1433	13,2767	14,8603
5	9,2364	11,0705	12,8325	15,0863	16,7496
6	10,6446	12,5916	14,4494	16,8119	18,5476
7	12,0170	14,0671	16,0128	18,4753	20,2777
8	13,3616	15,5073	17,5345	20,0902	21,9550
9	14,6837	16,9190	19,0228	21,6660	23,5894
10	15,9872	18,3070	20,4832	23,2093	25,1882
11	17,2750	19,6751	21,9200	24,7250	26,7568
12	18,5493	21,0261	23,3367	26,2170	28,2995

INDICI DI FORMA

La forma di una distribuzione di frequenza è valutabile dal confronto della curva di frequenza della distribuzione con la curva normale o gaussiana.

La forma riguarda il modo in cui i valori osservati per un carattere quantitativo si dispongono attorno alla media

INDICI DI FORMA

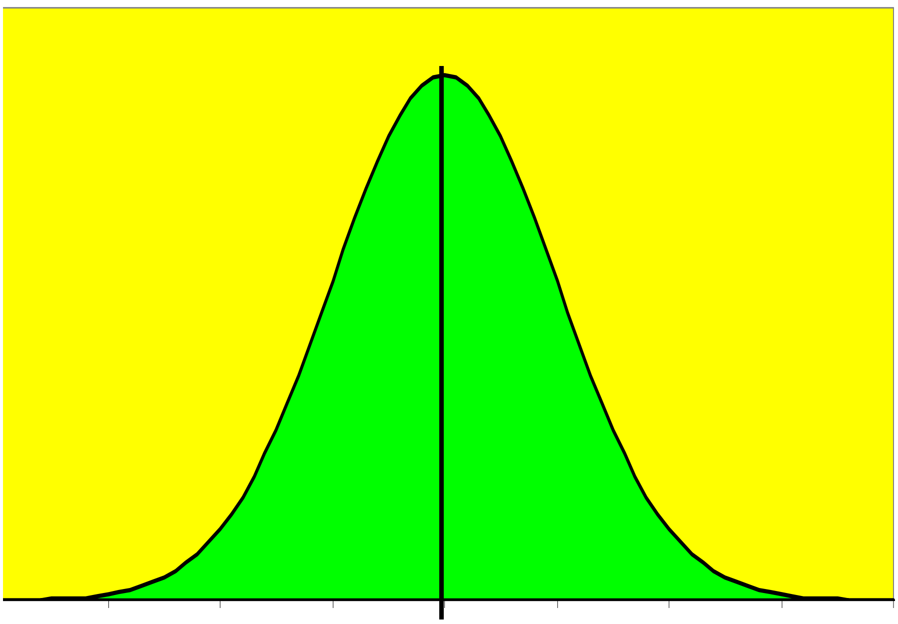
La relazione esistente tra media moda e mediana DI UNA VARIABILE O CARATTERE STATISTICO DI TIPO CONTINUO consentono di verificare se una distribuzione si presenta simmetrica o asimmetrica

SIMMETRICA SE MEDIA =MEDIANA=MODA

ASIMMETRICA POSITIVA SE MODA<MEDIANA<MEDIA

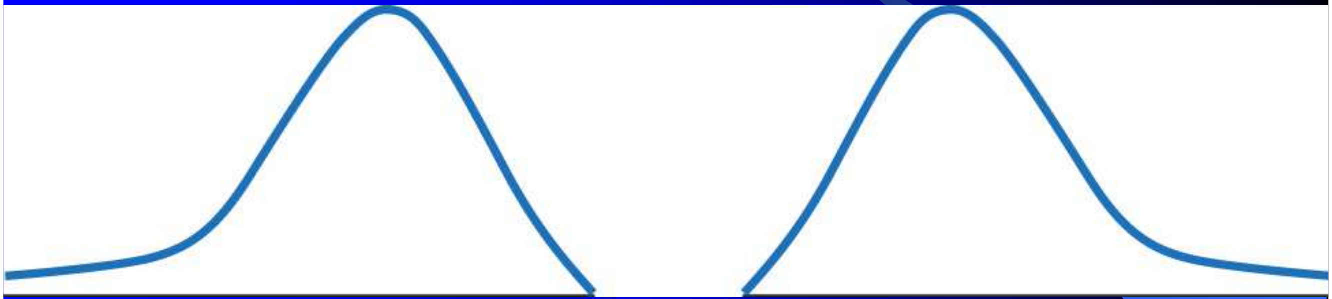
ASIMMETRICA NEGATIVA SE MEDIA<MEDIANA<MODA

DISTRIBUZIONE SIMMERICA



MEDIA=MODA=MEDIANA

DISTRIBUZIONI ASIMMETRICHE



ASIMMETRICA NEGATIVA SE MEDIA<MEDIANA<MODA

ASIMMETRICA POSITIVA SE MODA<MEDIANA<MEDIA

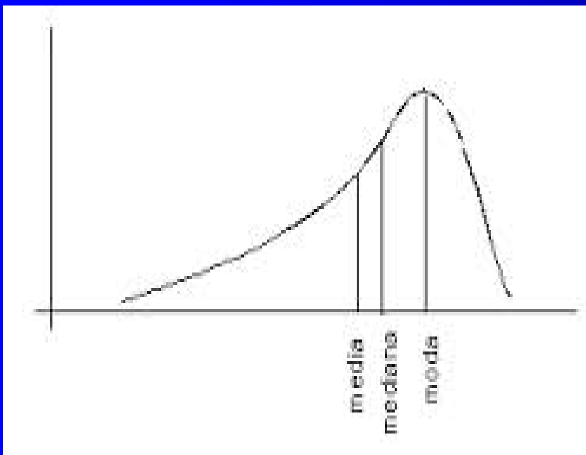
INDICE DI ASIMMETRIA

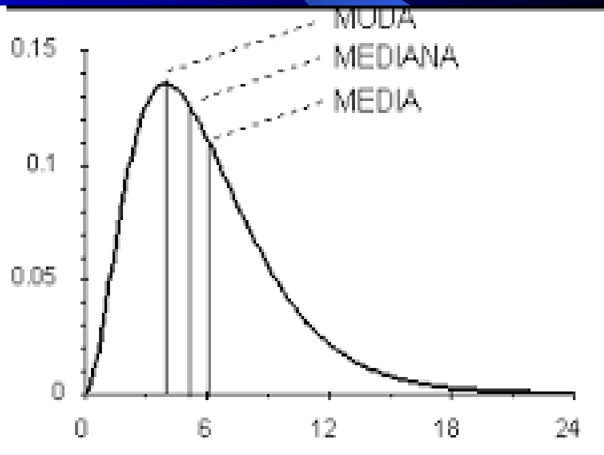
$$\gamma_1 = \frac{\sum_{i=1}^{N} (x_i - \mu)^3}{N\sigma^3}$$

$$\gamma_1 = \frac{\sum_{i=1}^{N} (x_i - \mu)^3 n_i}{N\sigma^3}$$

INDICE DI ASIMMETRIA Esempio…

SE la distribuzione si espande maggiormente a destra della media allora γ_1 è positivo prevalgono gli scarti positivi su quelli negativi entrambi al cubo mentre l'indice è negativo nel caso in cui la distribuzione si espande maggiormente a sinistra





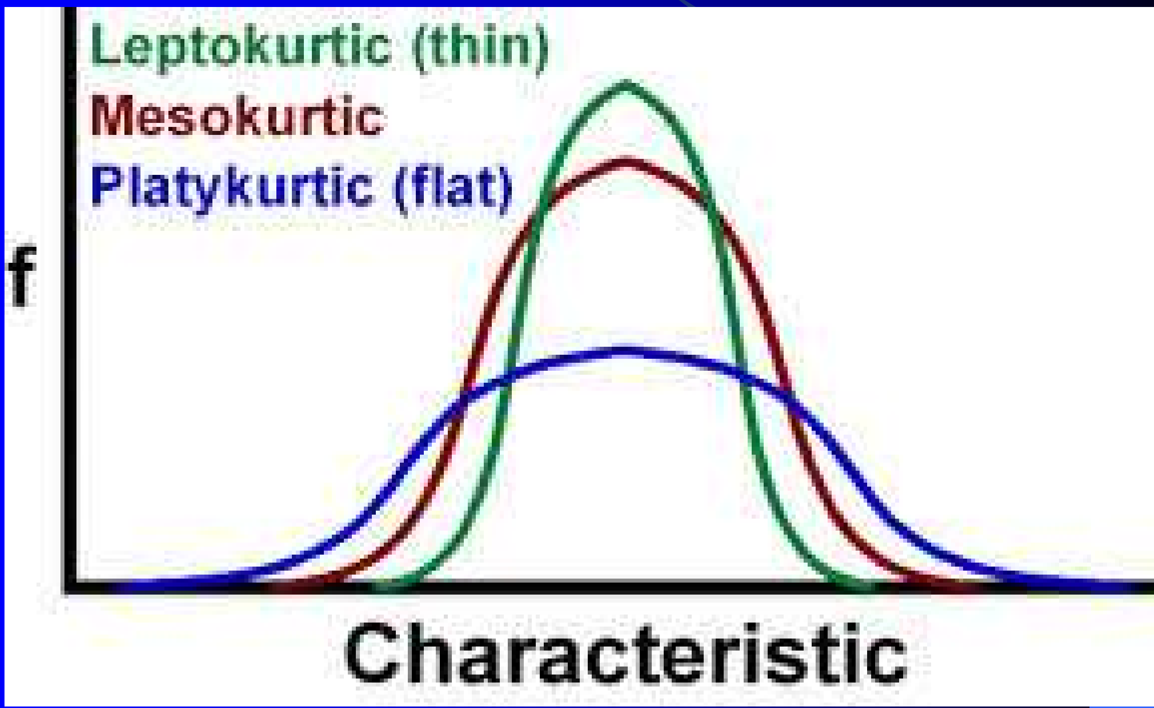
Coefficiente di disnormalità o di Curtosi

$$\gamma_2 = \frac{\sum (x_i - \mu)^4}{N\sigma^4} - 3$$

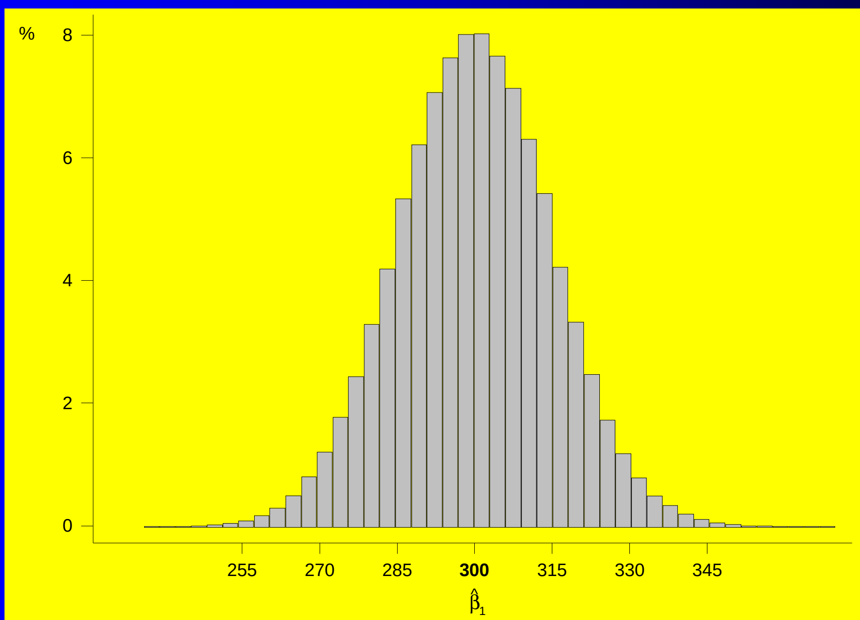
$$\gamma_2 = \frac{\sum (x_i - \mu)^4 n_i}{N\sigma^4} - 3$$

Assume valori zero quando è normale, valori positivi per distribuzioni ipernormali o aguzze e valori negativi per distribuzioni iponormali

INDICE DI ASIMMETRIA Esempio…



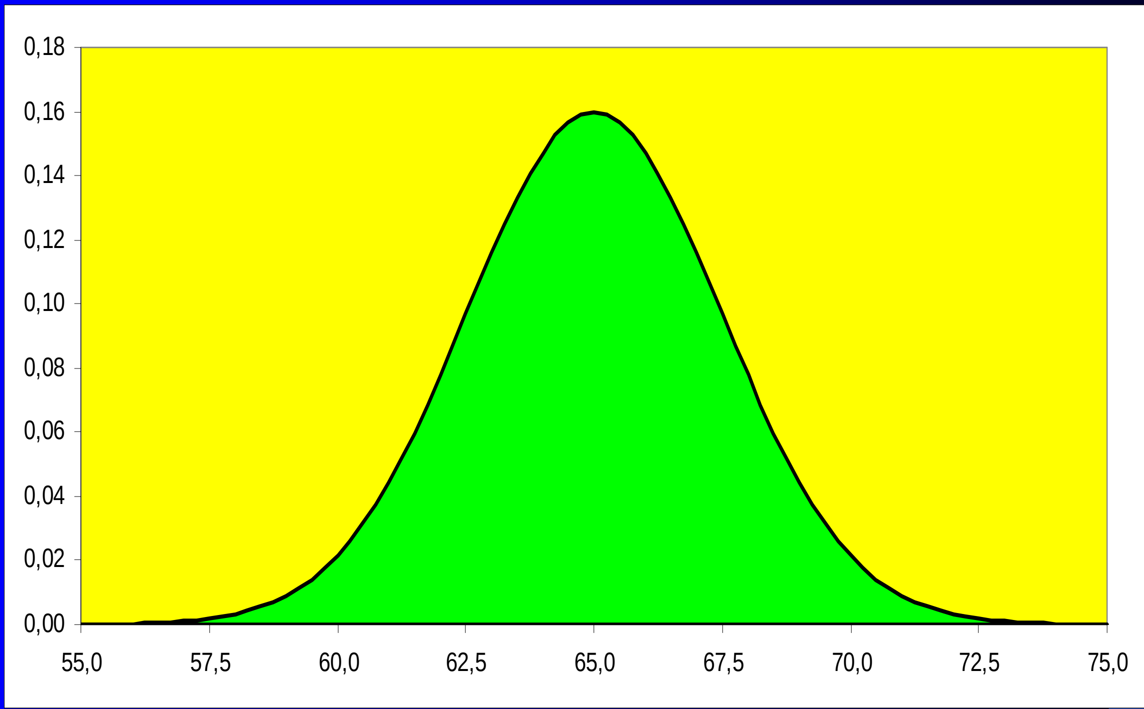
CURVA NORMALE



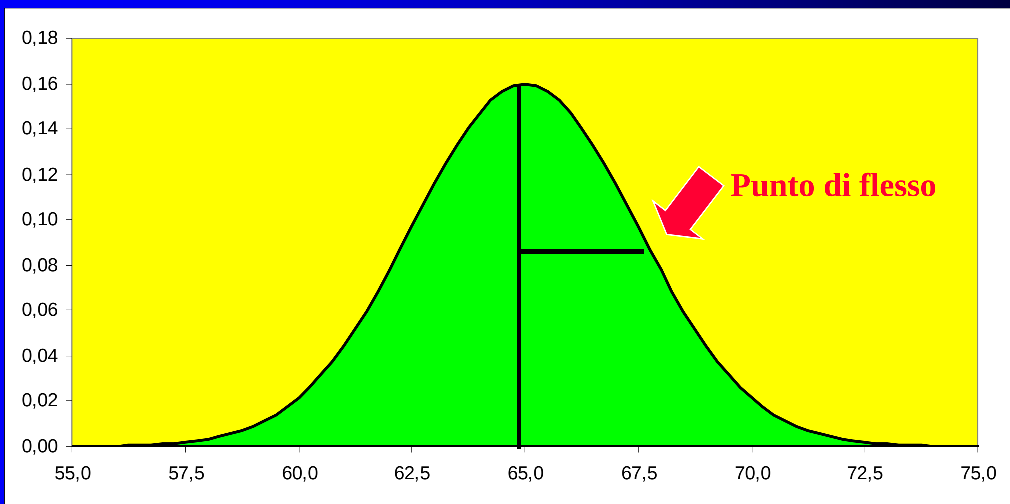
V.S. DIVISA IN INTERVALLI → ISTOGRAMMA → RIDUZIONE INTERVALLI

LINEA PASSANTE PER I PUNTI CENTRALI DI OGNI RETTANGOLO ORIGINA UNA CURVA NORMALE O GAUSSIANA

GRAFICO CURVA NORMALE



CURVA NORMALE V.S. ETA'
MEDIA=65 DEVIANZA STANDARD 2,5



LA PUNTA PIU' ALTA INDICA IL VALORE MEDIO

MEDIA=MEDIANA=MODA

ANDAMENTO PERFETTAMENTE SIMMETRICO ASINTOTICA ASSE X OSSIA NON TOCCA ASSE X

CARATTERISTICHE CURVA NORMALE

- Asintotica all'asse delle x, cioè quanto più si allontana dalla dalla media tanto più la curva si avvicina all'asse delle x
- L'ascissa del punto di massimo è paria lla media e coincide con moda e mediana
- Forma campanulare ed è immetrica
- Punti di flesso da concava verso l'alto diventa convessa verso il basso

CURVA NORMALE

LA MAGGIOR PARTE DELLE DISTRIBUZIONI O VARIABILI STATISTICHE QUANTITATIVE E DI TIPO CONTINUO TENDONO A DISTRIBUIRSI SECONDO UNA CURVA NORMALE OSSIA LE SINGOLE OSSERVAZIONI DI UN FENOMENO COLLETIVO TENDONO AD ADDENSARSI INTORNO AL VALORE MEDIO DELLA OSSERVAZIONE STESSA

LA CURVA NORMALE E PERFETTAMENTE SIMMETRICA OSSIA LA CODA SINISTRA E' UGUALE ALLA CODA DESTRA

CURVA NORMALE

PARTENDO DALLE RICERCHE DI GAUSS SULLA DISTRIBUZIONE DEGLI ERRORI DI OSSERVAZIONI DI UNA STESSA GRANDEZZA MISURATA PIU' VOLTE,(variabile statistica continua) SI DIMOSTRA CHE L'EPRESSIONE ALGEBRICA DI QUESTA CURVA DIPENDE SOLTANTO DAL NUMERO DELLE OSSREVAZIONI N, DALLA MEDIA E DALLO SCARTO QUADRATICO MEDIO ,PRECISAMENTE LA FUNZIONE DELLA CURVA NORMALE RISULTA ESSERE LA SEGUENTE

FUNZIONE DELLA CURVA NORMALE

$$Y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

Π=3,14 e=2,71 N=osservazioni σ=scarto quadratico medio μ=media aritmetica

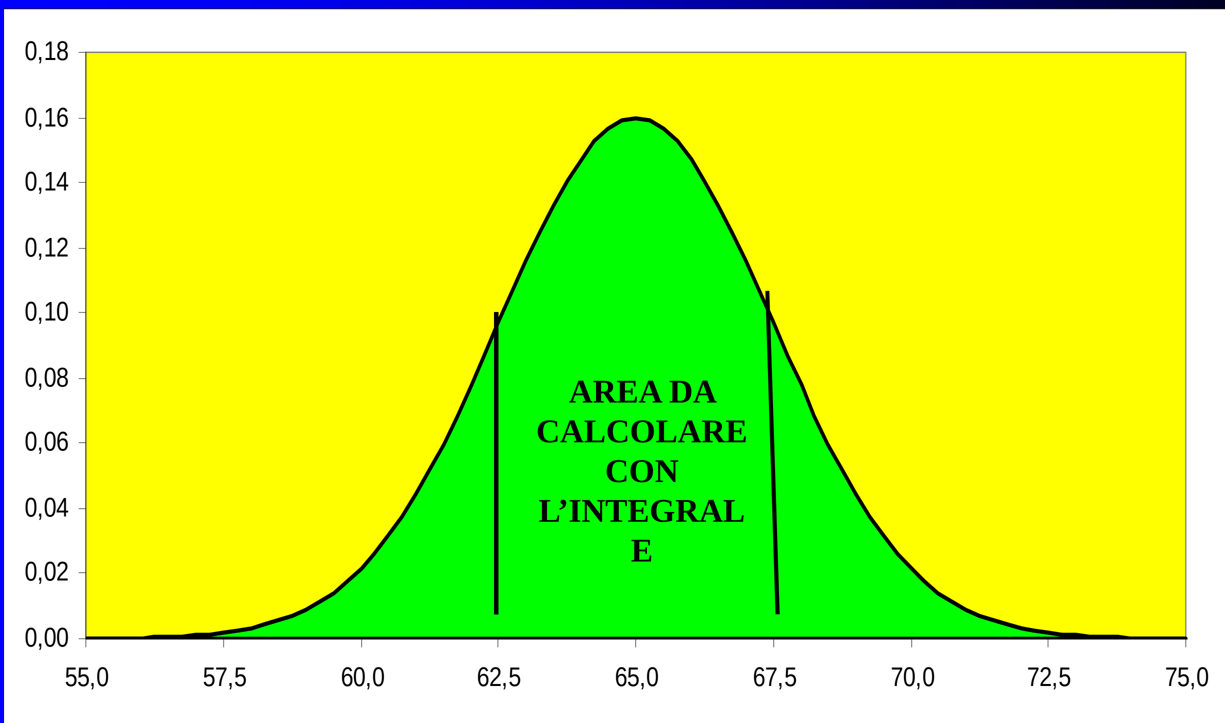
CURVA NORMALE

LA FUNZIONE PRECEDENTE CI PERMETTE DI CALCOLARE TUTTA L'AREA SOTTESA DALLA CURVA NORMALE, MA PER CALCOLARE UNA PORZIONE DI AREA TRA A E B DEVO UTILIZZARE IL SEGUENTE INTEGRALE

$$\int_b^a \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = N$$

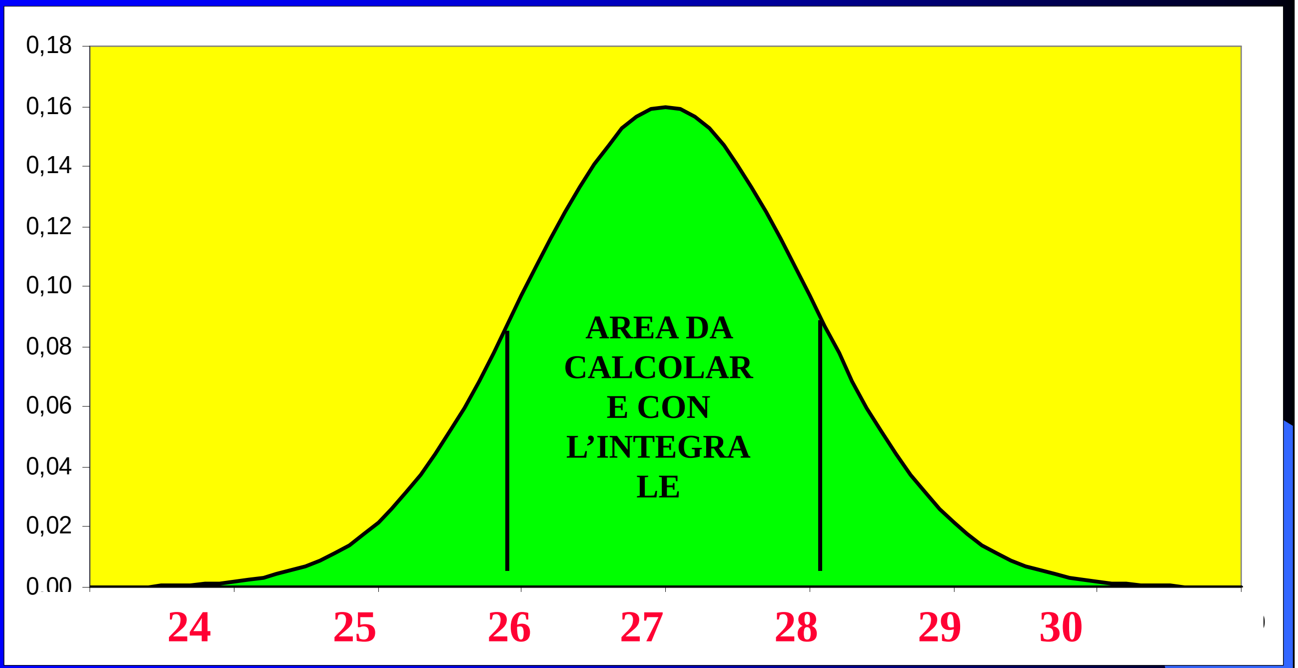
Con il calcolo dell'integrale possiamo saper in valore assoluto i casi compresi nell'intervallo a e b.

AREA CURVA NORMALE



Capire quanti sono in valore assoluto i casi compresi tra 62 e 67 anni

AREA CURVA NORMALE



Quanti sono in valore assoluti gli studenti che hanno presi da 26 a 27 all'esame di statistica

CURVA NORMALE

Per calcolare il numero dei casi compresi nell'area di interesse per il primo caso l'età da 62 a 67 anni e nel secondo caso per il voto da 26 a 28 occorre fare l'integrale, impresa è molto difficile motivo per cui si ricorre ai valori relativi e non assoluti. Ad esempio non potremmo dire abbiamo ottenuto 345 casi che hanno avuto un voto da 26 a 28 ma diremo abbiamo il 25 % dei casi compresi nell'intervallo da 26 a 28 ossia la percentuale è un valore relativo. Ciò è possibile con l'utilizzo degli scarti standardizzati

SCARTO STANDARDIZZATO

PER SEMPLIFICARE TUTTO SI INTRODUCE LO SCARTO STANDARDIZZATO OSSIA CONSIDERIAMO NON I VALORI ASSOLUTI MA RELATIVI.

Lo scarto standardizzato rappresenta una unità di misura

$$Z = \frac{X - \mu}{\sigma}$$

ESEMPI SCARTI STANDARDIZZATI

$$Z = \frac{X - \mu}{\sigma}$$

X = Voto individuale ottenuto

μ = media dei voti

$\sigma = scarto\ quadratico\ medio$

Primo studente Z=28-25/2=1,5

Secondo studente Z=30-25/2=2,5

Terzo studente Z=25-30/2=-2,5

Quarto studenete Z=28-30/2=-1

Nel primo caso uno studente ha preso 28 all'esame di statistica e la media dei voti è di 25 e lo scarto quadratico medio è di 2. Il risultato di 1,5 dista dalla media del gruppo considerato 1,5 SIGMA. Sigma rappresenta la nuova unità di misura.

ESEMPI SCARTI STANDARDIZZATI

$$Z = \frac{X - \mu}{\sigma}$$

X = Voto individuale ottenuto

$\mu = media \quad dei \quad voti$

$\sigma = scarto\, quadratico\, medio$

Primo studente Z=28-25/2=1,5

Secondo studente Z=30-25/2=2,5

Terzo studente Z=25-30/2=-2,5

Quarto studenete Z=28-30/2=-1

Nel primo caso uno studente ha preso 28 all'esame di statistica e la media dei voti è di 25 e lo scarto quadratico medio è di 2. Il risultato di 1,5 dista dalla media del gruppo considerato 1,5 SIGMA. Sigma rappresenta la nuova unità di misura.

Standardizzare la variabile x_i=1,3,4,5,7 di media 4 e s.q.m σ=$\sqrt{20/5} = 2$

1-4/2=-1,5

3-4/2=-0,5

4-4/2=0

5-4/2=0,5

7-4/2=1,5

La media (-1,5-0,5+0+0,5+1,5)/5=0

La varianza risulta unitaria

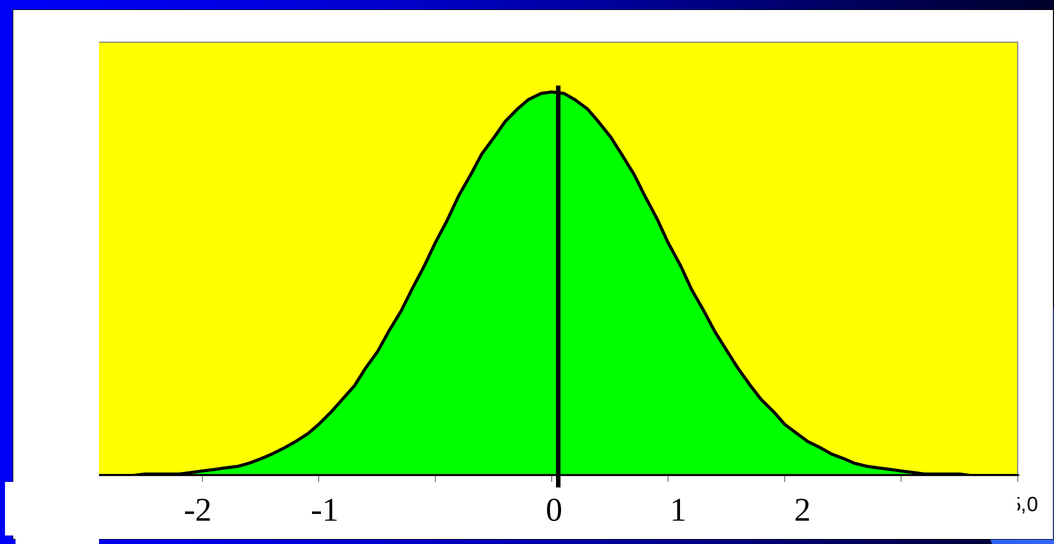
(-1,5)2+(-0,5)2+(0)2+(0,5)2+(1,5)2=5/5=1

CURVA NORMALE STANDARDIZZATA

PER FINALITA' PRATICHE E' OPORTUNO RICONDURRE LE INFINITE CURVE NORMALI CARATTERIZZATE DA DIVERSI VALORI DI N, σ, μ AD UN UNICA CURVA COM MEDIA 0 E VARIABILITA 1 ED AREA N UGUALE 1

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

GRAFICO CURVA NORMALE STANDARDIZZATA



Area = 1 Media=0 Varianza=1

Con la c.n.s. sappiamo in percentuale i casi compresi in un intervallo

GRAFICO CURVA NORMALE STANDARDIZZATA

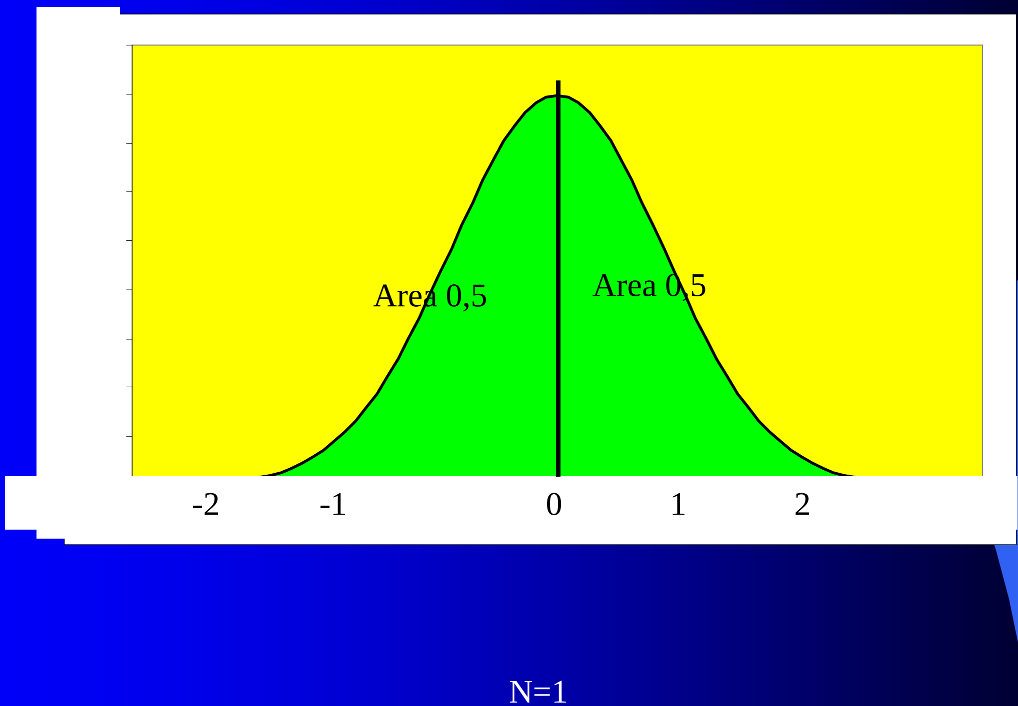
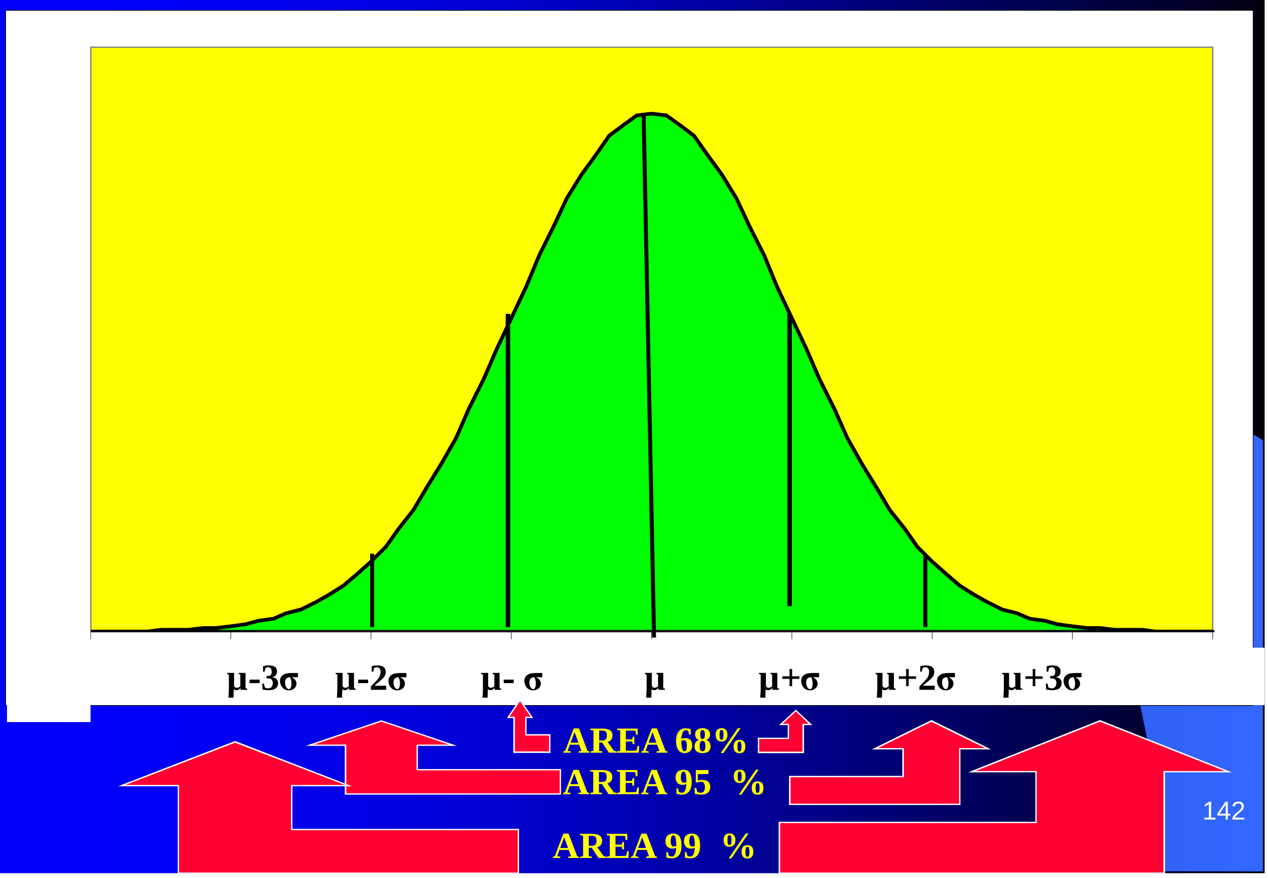


GRAFICO CURVA NORMALE STANDARDIZZATA



AREA SOTTO LA CURVA NORMALE

Voto Medio Esame statistica 25 Sigma 2

$$\left.\begin{matrix}\mu-\sigma\\ \mu+\sigma\end{matrix}\right\rangle = 68\%\ \text{area}$$

25-2=23

25+2=27

Z_1=23-25/2= -1

Z_2= 27-25/2= +1

Il 68% degli studenti consegue un voto da 23 a 27

CONSULTA TAVOLE

Val.Ass.

Val.Rel.

$$\left.\begin{matrix}\mu-2\sigma\\ \mu+2\sigma\end{matrix}\right\rangle = 95\%\ \text{area}$$

25-4=21

25+4=29

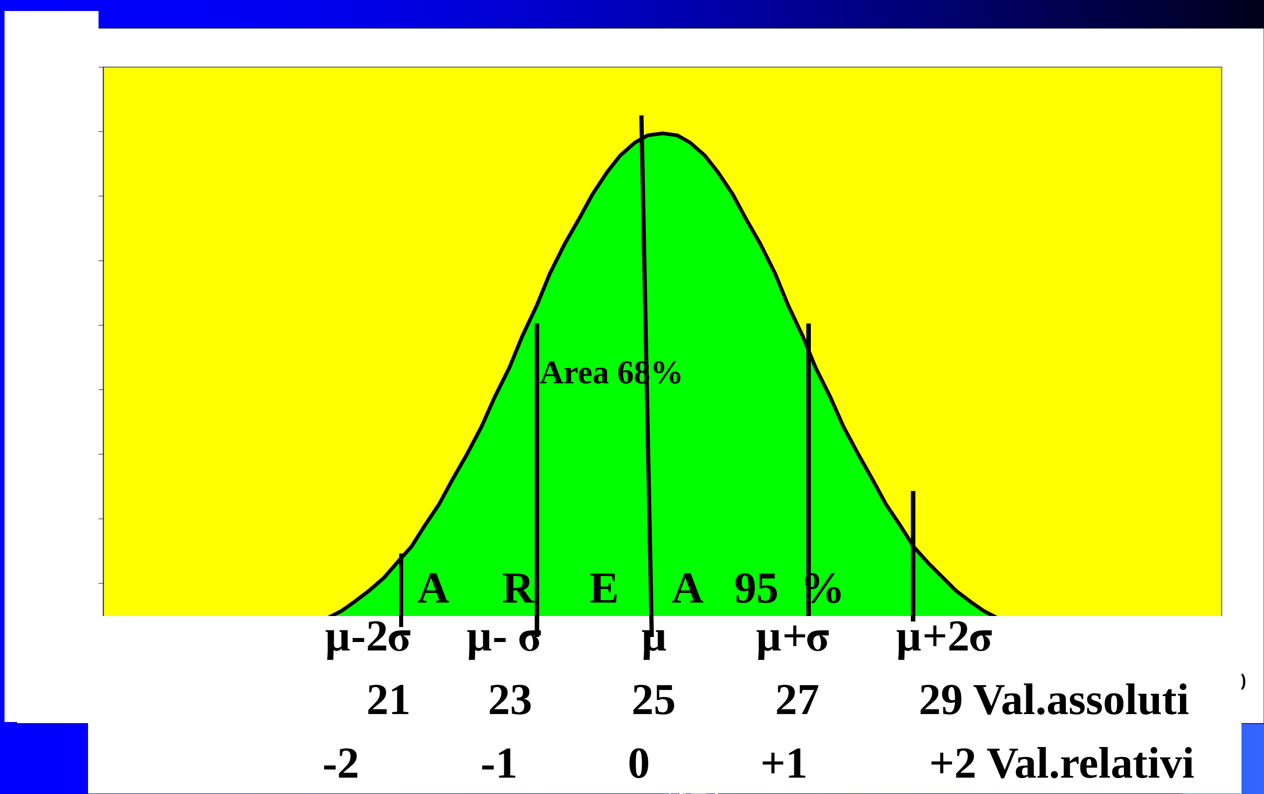
Z_1=21-25/2= - 2

Z_2=29-25/2= +2

Il 95% degli studenti consegue un voto da 21 a 29

$$\left.\begin{matrix}\mu\text{-}3\sigma\\ \mu+3\sigma\end{matrix}\right\rangle = 99\%\text{area}$$

AREA CURVA NORMALE STANDARDIZZATA



AREA SOTTO LA CURVA NORMALE

A condizione che la variabile sia normale

ALTEZZA MEDIA 1,68 Sigma 2

$$\left.\begin{matrix}\mu-\sigma \\ \mu+\sigma\end{matrix}\right\rangle = 68\%\ \text{area}$$

168-2=166

168+2=170

Il 68% degli studenti è alto da 166 a 170 cm

$$\left.\begin{matrix}\mu-2\sigma \\ \mu+2\sigma\end{matrix}\right\rangle = 95\%\ \text{area}$$

168-4=164

168+4=172

Il 95% degli studenti è alto da 164 a 172 cm

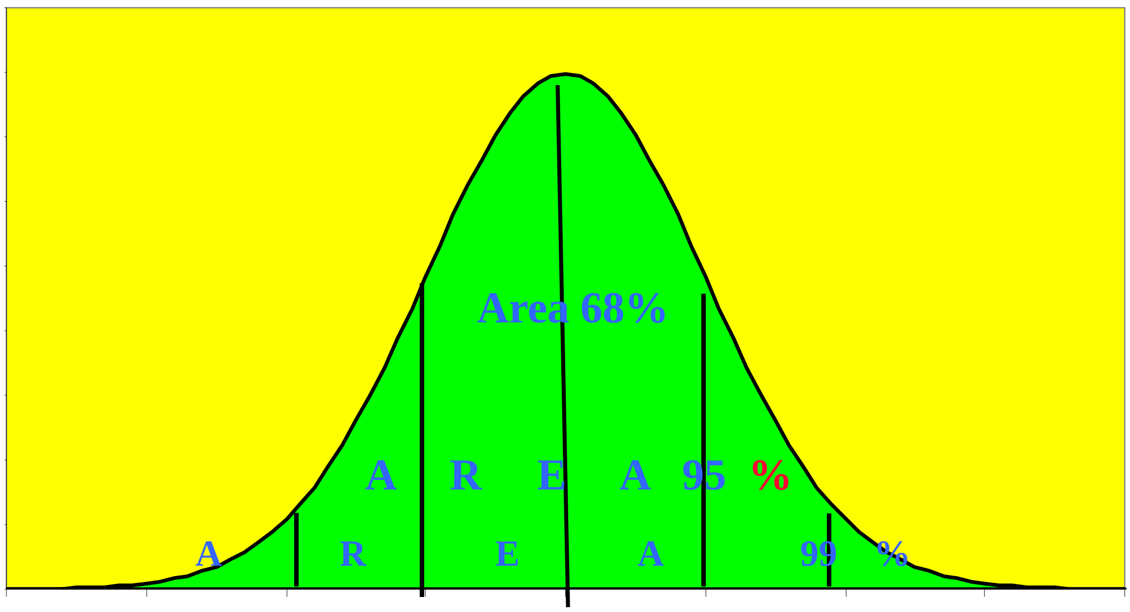
$$\left.\begin{matrix}\mu-3\sigma \\ \mu+3\sigma\end{matrix}\right\rangle = 99\%\text{area}$$

168-6=162

168+6=174

Il 99% degli studenti sono alti da 162 ad 174

AREA CURVA NORMALE STANDARDIZZATA



μ-3σ	μ-2σ	μ- σ	μ	μ+σ	μ+2σ	μ+3σ	
162	164	166	168	170	172	174	Val ass
-3	-2	-1	0	+1	+2	+3	Val rel