Luca Evangelisti
(Matricola: 674HHHCLDIPSI)

Sintetizzano un insieme di misure con un unico valore rappresentativo

Tabella aggiornata al: 26 novembre 2011

					Scale			
					Dati qualitativi	Dati considerabili sia qualitativi che quantitativi	Dati quantitativi	
Indice (Indicatore)	**Definizione**	**Dati**	**Formula**	**Nota**	**Nominale**	**Ordinale**	**Intervalli equivalenti**	**Rapporti equivalenti**
Gamma	Ampiezza dell'insieme dei dati grezzi	Grezzi	$G = x_{max} - x_{min}$		No	Sì	Sì	Sì
	Ampiezza delle classi	Raggruppati in classi	$A = \frac{G}{k}$	k= numero delle classi	No	Sì	Sì	Sì
Moda (o Valore modale)	Valore più frequente nella distribuzione osservata	Grezzi	$Mo = V_{Max}$	La distribuzione di frequenza si dice unimodale quando la Moda è unica, oppure bimodale quando la Moda è definita da due valori	Sì	Sì	Sì	Sì
Moda (o Classe modale)		Raggruppati in classi						
Media aritmetica (**m** o **x soprassegnato** quando ci si riferisce al campione, **μ** quando ci si riferisce alla popolazione)	Somma delle misure osservate diviso il numero delle osservazioni fatte	Grezzi	$m, \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$	* x = valore i-esimo * n = numero dei casi Proprietà della Media: 1) la somma degli scarti dei singoli valori dalla media è sempre uguale a zero 2) la somma dei quadrati degli scarti di ciascun valore dalla media è minore della somma degli scarti degli stessi valori da un qualsiasi altro numero (proprietà dei minimi quadrati)	No	No	Sì	Sì
	Media aritmetica per dati raggruppati	1) si moltiplica la frequenza di ogni classe per il valore "tipico" corrispondente al **Punto medio** della classe 2) si somma e poi si divide per il numero dei casi	$m, \bar{x} = \frac{\sum_{c=1}^{k} x_c \cdot f_c}{n}$	* x = valore "tipico" di ciascuna classe * f = frequenza associata alla X c-esima * k = numero dei diversi valori * n = numero dei casi	No	No	Sì	Sì
Mediana (Me o Mdn)	Valore che occupa la posizione centrale in una distribuzione ordinata	Grezzi: 1) si ordinano le categorie 2) si calcolano le frequenze cumulate 3) si identifica la posizione mediana 4) si trova, tramite la distribuzione delle frequenze cumulate, il "valore" corrispondente	$POSMe = \frac{n+1}{2}$ $Me = x_{POSMe}$	* POSMe = caso individuato * n = numero dei casi * n dispari: valore corrispondente a POSMe * n pari: Media aritmetica dei valori attorno a POSMe Nota: se n>30 al numeratore si può scrivere soltanto n invece che n+1	No	Non esiste, ma si possono considerare i valori subito prima e subito dopo	Sì	Sì
		Raggruppati in classi: Formula generale quando si ha a che fare con dati continui	$POSMe = \frac{n+1}{2}$ $Me = X_{li} + \frac{POSMe - f_{licum}}{f_i} \cdot Amp$	Procedura: 1) si individua la posizione mediana POSMe 2) tra le frequenze cumulate si individua quella che "contiene" POSMe 3) si trova il valore x, o la classe, che corrisponde alla frequenza cumulata 4) infine si calcola il valore esatto della Mediana con la seconda formula, dove: Xli = limite reale inferiore della classe che contiene la Mediana flicum = frequenza cumulata della classe precedente a quella che contiene la Mediana Amp = ampiezza della classe che contiene la Mediana fi = frequenza della classe che contiene la Mediana **Nota**: le considerazioni fatte per la Mediana sono valide per qualsiasi Quantile	No		Sì	Sì

Luca Evangelisti
(Matricola: 674HHHCLDIPSI)

Sintetizzano un insieme di misure con un unico valore rappresentativo

Tabella aggiornata al: 26 novembre 2011

					Scale			
					Dati qualitativi	Dati considerabili sia qualitativi che quantitativi	Dati quantitativi	
Indice (Indicatore)	**Definizione**	**Dati**	**Formula**	**Nota**	**Nominale**	**Ordinale**	**Intervalli equivalenti**	**Rapporti equivalenti**
Punto medio o **centrale** (detto anche Xcentrale, Xc)	Semisomma dei limiti inferiore e superiore	Grezzi	$p_m=\frac{l_i+l_s}{2}$	l = limite (inferiore e superiore)	No	Sì	Sì	Sì
Quartile	Detti generalmente **Quantili**: valori della distribuzione in corrispondenza dei quali la distribuzione viene suddivisa in 4, 10 o 100 parti	Grezzi: trovata la posizione dei vari Quantili, si legge il valore corrispondente V_{PosQ_i} **Nota**: per i dati raggruppati in classi si rimanda alla relativa formula della Mediana per poi naturalmente adattarla	$PosQ_i=\frac{n+1}{4}\cdot i$	Se la posizione calcolata non è un intero, per trovare il valore corrispondente si applica la formula: $V=((V_s-V_i)\cdot q)+V_i$ Vs = valore superiore Vi =valore inferiore q = la quantità che eccede la posizione del Vi	No	rappresenta soltanto la posizione	Sì	Sì
Decile			$PosD_i=\frac{n+1}{10}\cdot i$					
Centile (o Percentile)			$PosC_i=\frac{n+1}{100}\cdot i$					

Luca Evangelisti
(Matricola: 674HHHCLDIPSI)

Specificano la variabilità dei dati

Tabella aggiornata al: 26 novembre 2011

					Scale			
					Dati qualitativi	Dati considerabili sia qualitativi che quantitativi	Dati quantitativi	
Indice (Indicatore)	**Definizione**	**Dati**	**Formula**	**Nota**	**Nominale**	**Ordinale**	**Intervalli equivalenti**	**Rapporti equivalenti**
Campo di Variazione	Esprime l'ampiezza della variazione dei valori misurati	Grezzi	$C_v = x_{max} - x_{min}$	Molto sensibile a valori aberranti. Corrisponde alla Gamma.	No	No	Sì	Sì
		Se la distribuzione di frequenza è in classi si calcola la differenza tra i punti medi delle classi estreme (sup e inf)	$C_v = p_{m(Ksup)} - p_{m(Kinf)}$		No	No	Sì	Sì
Differenza interquantilica (DI)	Indice di variabilità delle misure	Raggruppati in classi	$DI = Q_3 - Q_1$		No	non molto significativa	Sì	Sì
Scarto Semplice Medio (SSM) o Scostamento	Media delle differenze, in valore assoluto, della Media della distribuzione	Grezzi	$SSM = \frac{\sum_{i=1}^{n} \lvert x_i - \bar{x} \rvert}{n}$	Questo indice è molto meno usato rispetto alla varianza	No	No	Sì	Sì
Varianza (indicata con **s^2** o **$V_{(x)}$** quando ci si riferisce al campione, **σ^2** (sigma quadro) se ci si riferisce alla popolazione)	Media del quadrato degli scostamenti dalla Media. La parte al numeratore viene definita **devianza** e può essere scomposta	Grezzi	$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$	Poiché la somma degli scarti dalla media è zero, si sommano gli scarti elevati al quadrato. Se n>30 al denominatore si può mettere soltanto n al posto di n-1. Il fatto che la Varianza è un valore quadratico e che quindi non ha la stessa unità di misura della Media, la rende poco descrittiva nell'osservazione immediata dei dati; per questo si usa spesso la **Deviazione standard**	No	No	Sì	Sì
		Raggruppati in classi	$s^2 = \frac{\sum_{c=1}^{k} f_c \cdot (x_c - \bar{x})^2}{\sum f_c}$	k = numero delle classi	No	No	Sì	Sì
		Formula abbreviata (utile se si dispone di una calcolatrice)	$s^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n} \right)^2$	In pratica è sufficiente calcolare le somme delle x_i e delle x_i^2.	No	No	Sì	Sì
Deviazione standard (o scarto quadratico medio)	Indica di quanto, mediamente, i dati osservati si discostano dalla loro Media ed *è la radice quadrata della Varianza*	Grezzi	$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$	Si indica con **s** se si tratta di dati osservati su campioni, mentre si indica con **σ** (sigma) se si tratta di distribuzioni teoriche riferite alla popolazione. Spesso si indica anche come DS(x)	No	No	Sì	Sì
		Raggruppati in classi	$s = \sqrt{\frac{\sum_{j=1}^{k} f_j \cdot (x_{c_j} - \bar{x})^2}{n}}$	Rappresenta un indicatore di dispersione relativa	No	No	Sì	Sì
Coefficiente di Variazione (CV)	Sintetizza il rapporto tra la **Deviazione standard** e la **Media**		$CV = \frac{s}{\bar{x}} \cdot 100$	Indica la dispersione dei punteggi usando come unità di misura la media dei punteggi stessi: è quindi un indicatore di variabilità relativa	No	No	Sì	Sì

Utilizzati per adeguare la misura ad una scala standard con Media e Varianza note

Tabella aggiornata al: 27 novembre 2011

					Scale			
					Dati qualitativi	Dati considerabili sia qualitativi che quantitativi	Dati quantitativi	
Indice (Indicatore)	**Definizione**	**Dati**	**Formula**	**Nota**	**Nominale**	**Ordinale**	**Intervalli equivalenti**	**Rapporti equivalenti**
punto z	Indica la posizione del dato in termini di distanza dalla media	Grezzi	$z_i = \frac{x_i - \bar{x}}{s}$	Se il punto z è negativo, significa che xi è al di sotto della Media, viceversa, se il punto z è positivo, xi è un valore sopra alla Media. * * * È possibile procedere a trasformazioni lineari dei punti z attraverso la formula: $Y = a + bz \rightarrow Y = \bar{x} + sz$ a = Media b = Deviazione standard	No	No	Sì	Sì
Standardizzazione dei valori della x attraverso i punti z	La Distribuzione Normale (DN) viene trasformata in Distribuzione Normale Standard (DNS) con μ=0 e σ=1 trasformando i valori della x in punti z			Dopodiché si possono standardizzare i valori della x con la formula dei punti z. In tal modo diventa possibile utilizzare un'unica Tavola (Tav. A) dove sono riportate le aree della curva in corrispondenza dei diversi valori di z				
scala T	Questa trasformazione lineare evita i valori negativi e decimali dei punti z	punti z	$T = 50 + 10 \cdot z_i$	Valori di a e b per la formula: M = 50 S = 10	No	No	Sì	Sì
scala sten (standard ten)	Permette di ottenere una nuova scala con 10 categorie standardizzate	punti z	$sten = 5{,}5 + 2 \cdot z_i$	Valori di a e b per la formula: M = 5,5 S = 2	No	No	Sì	Sì
scala stanine (standard nine)	Permette di ottenere una nuova scala con 9 categorie standardizzate	punti z	$stanine = 5 + 2 \cdot z_i$	Valori di a e b per la formula: M = 5 S = 2	No	No	Sì	Sì
QI	Quoziente di intelligenza	punti z	$Q.I. = \frac{età\ mentale}{età\ cronologica} \cdot 100$ $QI = 100 + 15 \cdot z_{Q.I.}$	età mentale = punteggio ottenuto attraverso dei test specifici --- Procedura 1) si calcola il punto z del Q.I. 2) si fa la trasformazione lineare con i valori di a e b: M=100 e s=15	No	No	Sì	No
punti St		punti z	$St = 5 + 1 \cdot z_i$	Valori di a e b per la formula: M = 5 S = 1	No	No	Sì	Sì
		Grezzi	$RP(x_i) = \frac{POS_{x_i} \cdot 100}{n+1}$	POSx è la posizione che il valore di x occupa nella distribuzione dei valori che deve essere preventivamente ordinata in maniera crescente --- Nota: naturalmente è possibile calcolare anche i Ranghi Quartili o Decili sostituendo il 100 al numeratore con, rispettivamente, 4 o 10	No	Sì	Sì	Sì

Ranghi percentili (o quartili o decili)	Il Rango percentile di un punteggio x, RP(x), è la percentuale di dati che assumono un valore uguale o minore di x	Raggruppati in classi	$POScl_{x_i} = f_{licum} + \frac{x - x_{li}}{ai} \cdot f_i$ $RP(x_i) = \frac{POScl_{x_i} \cdot 100}{n}$	1) si ordinano le classi (inserendo anche quelle mancanti) 2) si calcolano le frequenze cumulate 3) si identifica la classe che contiene il valore x 4) si determina la POScl(xi) 5) si trasforma la posizione in RP --- flicum = frequenza cumulata fino alla classe precedente xli = limite reale inferiore della classe che include x ai = ampiezza della classe che include x fi = frequenza della classe	No	Sì	Sì	Sì
punto t (distribuzione "t di Student")	La distribuzione **t di Student** viene utilizzata per definire degli intervalli di confidenza per la media di una popolazione		$t_i = \frac{X_i - M}{\sigma}$ $\sigma = \frac{s}{\sqrt{(n-1)}}$ dove	Se la distribuzione campionaria della media (DCM) non ha forma normale (tipicamente quando n<=30), se la varianza della popolazione (**σ^2**) non è nota, allora si fa riferimento alla distribuzione "t di Student" con (n-1) gradi di libertà (GDL). Non conoscendo la varianza della popolazione (**σ^2**), la verifica dell'ipotesi sulla media della popolazione (**μ**) si effettua sostituendo alla Deviazione standard (errore standard) **σ** la sua stima	No	No	Sì	Sì

Valori (Xi)	Xi-Media
4	0,250
7	2,750
2	2,250
4	0,250
4,25	5,500
Media	∑

n=	4
δ=	1,375

Valori (Xi)	Xi-Media	(Xi-Media)^2
4	-0,250	0,06
7	2,750	7,56
2	-2,250	5,06
4	-0,250	0,06
4,25		12,75
Media		∑

n=	4	
s^2=	4,250	Varianza
s=	2,062	Deviazione standard

	Proprietà della Distribuzione Campionaria della Media (DCM) rispetto alla Distribuzione della popolazione							
	Parametri della Distribuzione della popolazione			Parametri della Distribuzione Campionaria della Media (DCM)				
Caso	**Media**	**Varianza**	**Forma**	**Ampiezza (n)**	**Media**	**Varianza**	**Deviazione standard (errore standard)**	**Forma**
1	μ	Nota	Qualsiasi	>30	$\mu_{\bar{x}}=\mu$	$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}$	$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$	Normale
2	μ	Nota	Normale	Qualsiasi	$\mu_{\bar{x}}=\mu$	$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}$	$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$	Normale
3	μ	Ignota	Normale	>30	$\mu_{\bar{x}}=\mu$	$\sigma_{\bar{x}}^2=\frac{s_{campione}^2}{n-1}$	$\sigma_{\bar{x}}=\frac{s_{campione}}{\sqrt{n-1}}$	Normale
4	μ	Ignota	Normale	≤30	$\mu_{\bar{x}}=\mu$	$\sigma_{\bar{x}}^2=\frac{s_{campione}^2}{n-1}$	$\sigma_{\bar{x}}=\frac{s_{campione}}{\sqrt{n-1}}$	t di Student

Luca Evangelisti
(Matricola: 674HHHCLDIPSI)

chi quadrato

Tipo formaggio	Preferenze (fo)
F1	20
F2	39
F3	25
3	84
k	n

fo: frequenze osservate
ft: frequenze teoriche

Tipo formaggio	Preferenze (fo)	ft=n/k	fo-ft	$(fo-ft)^2$	$(fo-ft)^2/ft$	
F1	20	28	-8	64	2,29	
F2	39	28	11	121	4,32	
F3	25	28	-3	9	0,32	
					6,93	chi^2

Residuo standardizzato

Tipo formaggio	Preferenze (fo)	ft=n/k	fo-ft	radice di ft	(fo-ft)/radice di ft	verifico se >±2
F1	20	28	-8	5,29	-1,51	No
F2	39	28	11	5,29	2,08	Si
F3	25	28	-3	5,29	-0,57	No

Legenda
input
formule
risultato

Verifica delle ipotesi sulla forma della distribuzione nel caso di due campioni

Tavola di contingenza: composizione della tabella della distribuzione dei soggetti per le varie categorie e calcolo dei marginali

	A	B	C		
1	20	6	8	=1A+1B+1C	frequenza marginale
2	8	20	22	=2A+2B+2C	di riga
	=1A+2A	=1B+2B	=1C+2C	=tot (righe+colonne)	

frequenze marginali di colonna

ovvero:

	A	B	C	
1	20	6	8	34
2	8	20	22	50
	28	26	30	84

Calcolo delle frequenze teoriche (ft): si calcolano per ogni cella moltiplicando i corrispondenti totali marginali di riga e di colonna e dividendo il prodotto per il totale generale (nell'ipotesi che H0 sia vera)

	A	B	C
1	11,33	10,52	12,14
2	16,67	15,48	17,86

Confronto tra frequenze osservate (fo) e frequenze teoriche (ft)

Cella	fo	ft	(fo-ft)	$(fo-ft)^2$	$(fo-ft)^2/ft$	
1-A	20	11,33	8,67	75,11	6,63	
1-B	6	10,52	-4,52	20,46	1,94	
1-C	8	12,14	-4,14	17,16	1,41	
2-A	8	16,67	-8,67	75,11	4,51	
2-B	20	15,48	4,52	20,46	1,32	
2-C	22	17,86	4,14	17,16	0,96	
					16,78	**Chi^2 (X^2)**

Calcolo del Residuo standardizzato per capire quale dei valori è più significativo

Cella	fo	ft	(fo-ft)	(fo-ft)/ la radice di ft	Verifica se è > 2	
1-A	20	11,33	8,67	2,57	Sì	C'è una relazione significativa
1-B	6	10,52	-4,52	-1,39	No	
1-C	8	12,14	-4,14	-1,19	No	
2-A	8	16,67	-8,67	-2,12	Sì	C'è una relazione significativa
2-B	20	15,48	4,52	1,15	No	
2-C	22	17,86	4,14	0,98	No	

Nota: calcolo semplificato del chi quadrato nel caso di tabelle di contingenza 2x2

A	B	A+B
C	D	C+D
A+B	B+D	N (totale)

$$chi^2 = \frac{N \cdot (AD - BC)^2}{(A+B)\cdot(C+D)\cdot(A+C)\cdot(B+D)}$$

Livello di misura	Indice di correlazione
Relazione tra 2 variabili misurate almeno su scale a **intervalli equivalenti**	**r di Bravais-Pearson**
Relazione tra 2 variabili **dicotomiche** e **ordinali**	$\mathbf{r_{phi}}$
Relazione tra 1 variabile dicotomica e 1 variabile misurata almeno su scala a **intervalli equivalenti**	**r punto-biseriale ($\mathbf{r_{pb}}$)**
Relazione tra 2 variabili misurate su scala **ordinale**	**ρ (rho) di Spearman (rs)**
Relazione tra 1 variabile su scala ordinale e 1 su scala a intervalli o **rapporti equivalenti**	

r di Pearson

(per variabili a livello degli intervalli equivalenti)

Nazione	X (Consumo pro capite di sigarette nel 1930)	Y (Tasso di mortalità per cancro ai polmoni nel 1950)	XY	X^2	Y^2
Islanda	240	60	14400	57600	3600
Norvegia	250	90	22500	62500	8100
Svezia	310	120	37200	96100	14400
Danimarca	370	160	59200	136900	25600
Australia	450	160	72000	202500	25600
Olanda	450	240	108000	202500	57600
Canada	500	150	75000	250000	22500
Svizzera	530	250	132500	280900	62500
Finlandia	1110	350	388500	1232100	122500
Gran Bretagna	1130	460	519800	1276900	211600
Stati Uniti	1280	190	243200	1638400	36100
11	6620	2230	1672300	5436400	590100
N	**∑X**	**∑Y**	**∑XY**	**$\sum X^2$**	**$\sum Y^2$**

Legenda
- input
- formule
- risultato

numeratore 330245,45 **∑XY-((∑X*∑Y)/N)**

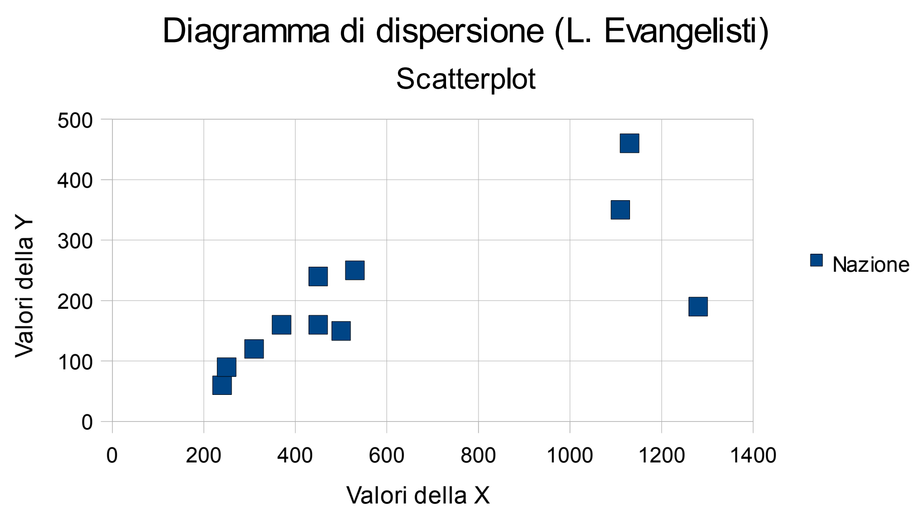
denominatore 447719,32 **Radice($\sum X^2$-($(\sum X)^2$)/N) * Radice($\sum Y^2$-($(\sum Y)^2$)/N)**

r di Pearson= 0,738

Verifica delle ipotesi

Con n-2 gdl la r di Pearson si distribuisce come la t di Student

t= 3,277 **t=r*radice((n-2)/(1-r2))**



Luca Evangelisti
(Matricola: 674HHHCLDIPSI)

r_{phi}
(per variabili dicotomiche e ordinali)

Variabili	**a**	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**
X	0	0	1	0	1	1	1	0	1	0
Y	0	0	1	0	0	1	1	1	0	0

Legenda
- input
- formule
- risultato

		Dom Y		
		1	0	
Dom X	1	fa	fb	p
	0	fc	fd	q
		p'	q'	n

		Dom Y		
		1	0	
Dom X	1	2	1	3
	0	3	2	5
		5	3	8

r_{phi} 0,07

r_{phi} = ((fa*fd) – (fb*fc)) / radice(p*p'*q*q')

chi^2 0,04

$chi^2 = n * r^2_{phi}$

Luca Evangelisti
(Matricola: 674HHHCLDIPSI)

r_{pb}
(punto-biseriale)

Test attitudinale cui si può rispondere giusto (Y=1) o errato (Y=0)
Esempio: verificare, p.e., se l'item 3 va nello stesso "verso" del totale del test
Domanda: quando un soggetto ha un punteggio alto nel test, risponde correttamente all'Item 3 e viceversa?

Soggetti	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Tot (X)
a	1	0	1	1	1	1	1	6
b	1	1	1	1	1	1	1	7
c	1	1	0	1	1	1	0	5
d	0	0	0	1	1	0	0	2
e	1	0	1	0	1	1	1	5
f	1	0	0	0	1	0	0	2
g	1	1	1	0	0	1	0	4
h	1	0	1	1	1	1	1	6
i	1	1	1	1	1	1	1	7
l	0	0	1	1	1	0	0	3
m	1	0	1	1	1	0	0	4
11								4,636

Calcolo della s_x della variabile continua			
Tot (X)	M_x	X-M	$(X-M)^2$
6	4,636	1,364	1,860
7	4,636	2,364	5,587
5	4,636	0,364	0,132
2	4,636	-2,636	6,950
5	4,636	0,364	0,132
2	4,636	-2,636	6,950
4	4,636	-0,636	0,405
6	4,636	1,364	1,860
7	4,636	2,364	5,587
3	4,636	-1,636	2,678
4	4,636	-0,636	0,405
		$\sum(X-M)^2$=	32,55
		$\sum(X-M)^2/n$=	2,96
Deviazione standard (s_x)		radice($\sum(X-M)^2/n$)=	1,72

Legenda

input
formule
valori calcolati manualmente
risultato

Procedura per passi

1) calcolo la media del totale quando Item 3 =1	5,250
2) calcolo la media del totale quando Item 3 =0	3,000
3) calcolo della deviazione standard (s_x)	1,72
4) conteggio degli n con valore "1"	8
5) conteggio degli n con valore "0"	3
6) n_1/n	0,73
7) n_0/n	0,27
8) radice($(n_1/n)*(n_0/n)$)	0,45
9) calcolo del punto biseriale r_{pb}	0,58
10) calcolo r_{pb}^2	0,34
11) trasformazione di r_{pb} in **t di Student**	**2,150**

rs (o rho) di Spearman

(correlazione tra ranghi)

C'è concordanza tra le due graduatorie?

Soggetti	X	Y	d = X-Y	d^2
a	1	2	-1	1
b	3	4	-1	1
c	4	5	-1	1
d	2	3	-1	1
e	5	1	4	16
5			**$\sum d^2$=**	20

Legenda

input
formule
valori calcolati manualmente
risultato

rs= $1 - (6 * \sum d^2) / (n * (n^2 - 1)$

rs= **0**

se n ≤ 30: valori tabulati di rs in funzione di α e di n
se n > 30: si trasforma rs in t di Student con (n-2) gdl

t di Student= rs * radice ((n – 2)/(1 - rs))

t di Student= **0**

Luca Evangelisti
(Matricola: 674HHHCLDIPSI)

Regressione semplice

$$Y' = a + bX$$

	VI variabile indipendente	VD variabile dipendente				Valore teorico della VD	Per il calcolo dell'errore standard della stima	
soggetti	**X**	**Y**	**XY**	**X^2**	**Y^2**	**Y'**	**(Y-Y')**	**$(Y-Y')^2$**
a	1	10	10	1	100	9,750	0,250	0,063
b	2	12	24	4	144	18,875	-6,875	47,266
c	6	50	300	36	2500	55,375	-5,375	28,891
d	3	30	90	9	900	28,000	2,000	4,000
e	6	62	372	36	3844	55,375	6,625	43,891
f	7	60	420	49	3600	64,500	-4,500	20,250
g	4	45	180	16	2025	37,125	7,875	62,016
∑= 7	29	269	1396	151	13113			206,38
Medie=	4,143	38,429						

$$b = \frac{N \cdot \sum XY - \sum X \sum Y}{N \cdot \sum X^2 - (\sum X)^2}$$

b (coefficiente angolare)= 9,125

$$a = \bar{Y} - b\bar{X}$$

a (intercetta)= 0,625

$$s_e = \sqrt{\frac{\sum (Y - Y')^2}{N - 2}}$$

s_e (errore standard della stima)= 41,275

Legenda
- input
- formule
- risultato

Retta di regressione

(interpolazione della nuvola dei punti)