

Statistica medica - Metodologia della ricerca e Evidence Based Medicine

C.I. SCIENZE PROPEDEUTICHE E BASI DELLA METODOLOGIA DELLA RICERCA [418ME]

cds in Logopedia A.A. 2022/23

parte 4

Prof. Lucio Torelli

Dipartimento Clinico di Scienze mediche, chirurgiche e della salute

Università degli Studi di Trieste

torelli@units.it

Attenzione: queste slide sono solo alcune note per le lezioni, non sono pertanto
un riassunto del seminario

9/1/2023

Contenuti del corso

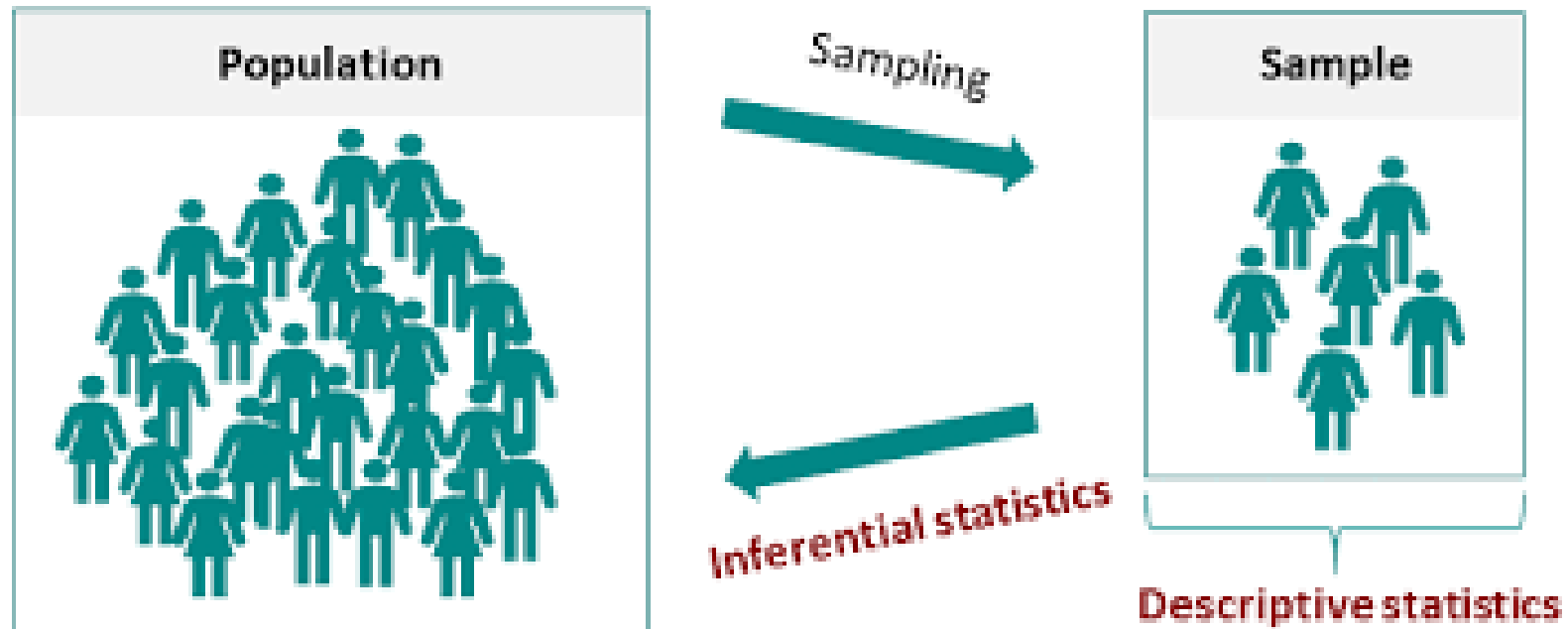
Elementi di **statistica descrittiva**

Elementi di **calcolo delle probabilità**

Cenni di statistica inferenziale



Inferenza statistica



il passaggio dal campione alla popolazione...

- Obiettivi dello studio
- Disegno dello studio
- Materiali e metodi,
criteri di inclusione/esclusione (**popolazione** oggetto dello studio)
- Scelta del **campione**
- Valutazione e descrizione del campione
- **Inferenza** sulla popolazione
- Discussione dei risultati

stime

- Stime puntuali
- Stime intervallari
- *Confidence intervals*

Cosa significa se leggiamo:

c.i. per μ al 95%: (26,4; 30,4) ?

Odds ratio – Confidence Interval

Use this calculator to determine a confidence interval for your odds ratio. An odds ratio is a measure of association between the presence or absence of two properties. For example, it could provide a measure of association between customers who are either older or younger than 25 and either have or have not claimed on their car insurance, in order to determine whether age is associated with the propensity to claim. The value of the odds ratio tells you how much more likely someone under 25 might be to make a claim, for example, and the associated confidence interval indicates the degree of uncertainty associated with that ratio.

Calculator

Contingency table		Property B	
		Presence	Absence
Property A	Presence	15	185
	Absence	14	286

What confidence level do you need? %

Typical choices are 90%, 95%, or 99%

Your odds ratio is

Your confidence interval is

Calculator

Contingency table		Property B	
		Presence	Absence
Property A	Presence	150	1850
	Absence	140	2860

What confidence level do you need? %

Typical choices are 90%, 95%, or 99%

Your odds ratio is

Your confidence interval is

Stime (per una proporzione)

Confidence interval for a proportion

Estimate the proportion with a dichotomous result or finding in a single sample. This calculator gives both binomial and normal approximation to the proportion.

Instructions: Enter parameters in the **green** cells. Answers will appear in the **blue** box below.

N = 400 Sample size
x = 20 Number in the sample with the result or finding in question
CL = 95 % Confidence level

Calculate

Cosa succede
con N=40 e x=2??

e se CL=99% o 80%

$$\hat{p} \pm z \sqrt{\frac{\hat{p} (1 - \hat{p})}{n}}$$

1. Binomial "exact" calculation

Proportion of positive results = $P = x/N = 0.0500$
Lower bound = 0.0308
Upper bound = 0.0762

stime

- Trovare il c.i. al 95% per la stima della media di una popolazione gaussiana $N(\mu, \sigma)$, avendo a disposizione i dati di un campione, opportunamente scelto, di numerosità n , media campionaria m e dev.st. campionaria s .

Risulta: $\left(m - \frac{q \cdot s}{\sqrt{n}}; m + \frac{q \cdot s}{\sqrt{n}} \right)$ q è il quantile ...

Cosa succede al variare di n , e cosa al variare di s ?

stime

- **Calcolo a priori della numerosità** n del campione per la stima della media di una $N(\mu, \sigma)$, data la lunghezza massima k del c.i.

Poiché la lunghezza del c.i. è $\frac{2qs}{\sqrt{n}}$, deve essere: $\frac{2qs}{\sqrt{n}} \leq k$,

da cui risulta: $n \geq \left(\frac{2qs}{k} \right)^2$

Ad es. se vogliamo un c.i. al 95% ($q \approx 2$) di lunghezza non superiore a 1 e se pensiamo che $s \approx 5$, risulta $n \geq 400$.

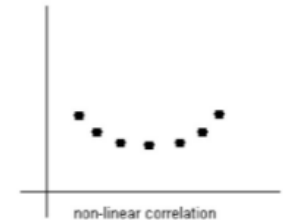
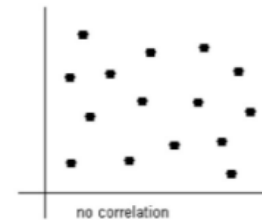
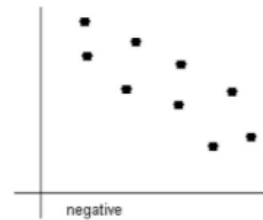
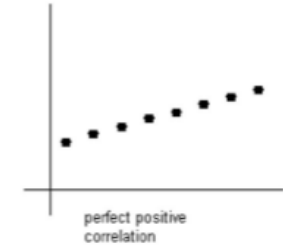
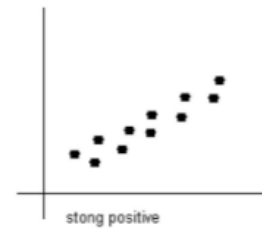
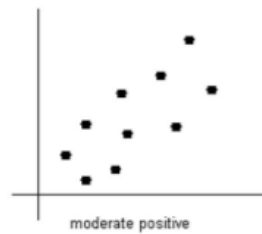
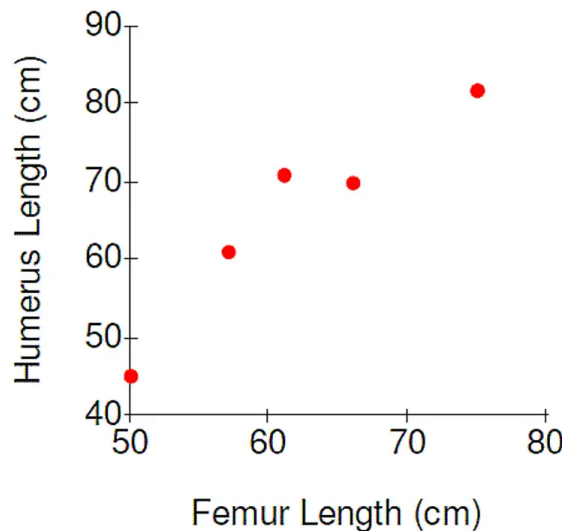
Ma se ad es. mi basta $k \leq 2$, allora $n \geq 100$.

Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study – March 11, 2020 –The Lancet

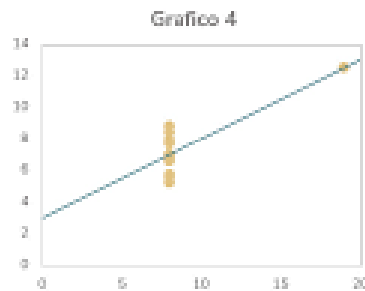
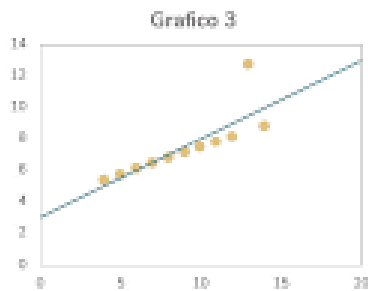
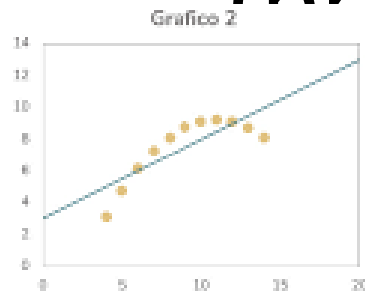
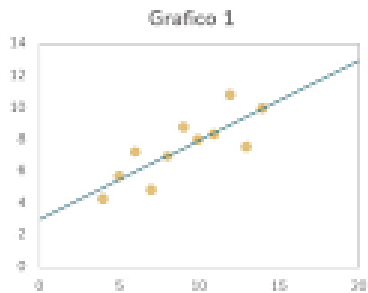
191 patients (135 from Jinyintan Hospital and 56 from Wuhan Pulmonary Hospital) were included in this study, of whom 137 were discharged and **54 died** in hospital. 91 (48%) patients had a comorbidity, with hypertension being the most common (58 [30%] patients), followed by diabetes (36 [19%] patients) and coronary heart disease (15 [8%] patients). Multivariable regression showed increasing odds of in-hospital death associated with older age (**odds ratio 1·10, 95% CI 1·03–1·17**, per year increase), higher Sequential Organ Failure Assessment (SOFA) score (**5·65, 2·61–12·23**; $p < 0·0001$), and d-dimer greater than $1 \mu\text{g/mL}$ (**18·42, 2·64–128·55**; $p = 0·0033$) on admission. ...

Current smoker (vs non-smoker): **2·23, 0·65–7·63** ...

Female sex (vs male): **0·61, 0·31–1·20** ...



La misurazione della correlazione (primi cenni)



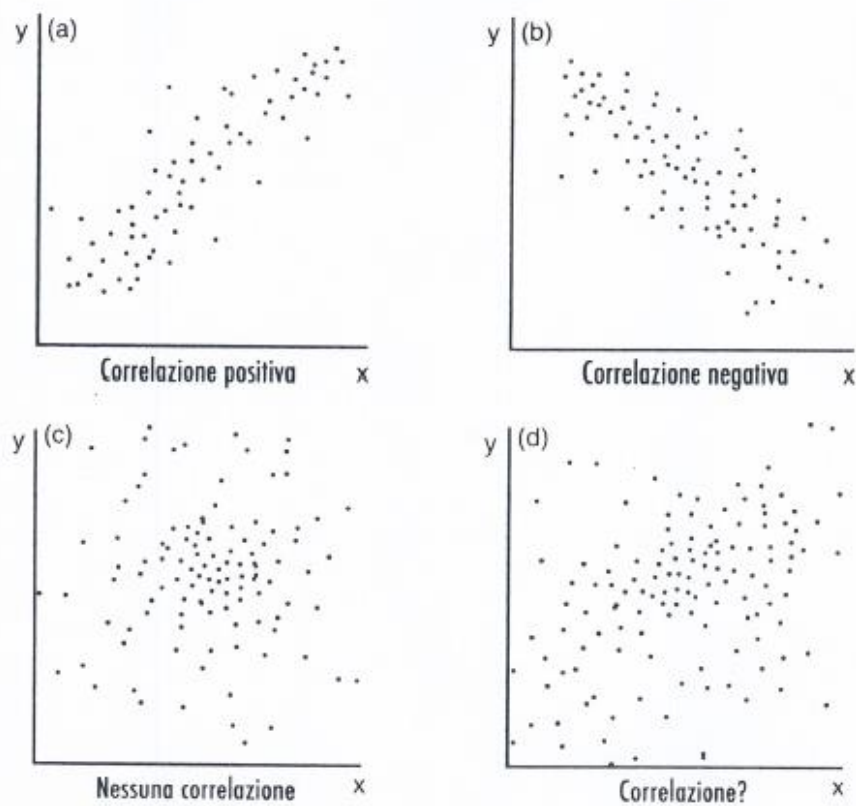


Figura 13.1 Diagramma di dispersione di dati bivariati

«... La valutazione soggettiva di un diagramma di dispersione può essere sostituita da una tecnica statistica in grado di fornire indicazioni più precise e obiettive.

La statistica che fornisce un indice di accordo su come le due variabili siano in relazione tra di loro è il **coefficiente di correlazione**.

La statistica è calcolata a partire dai dati campionari, e fornisce una stima del corrispondente parametro nella popolazione.

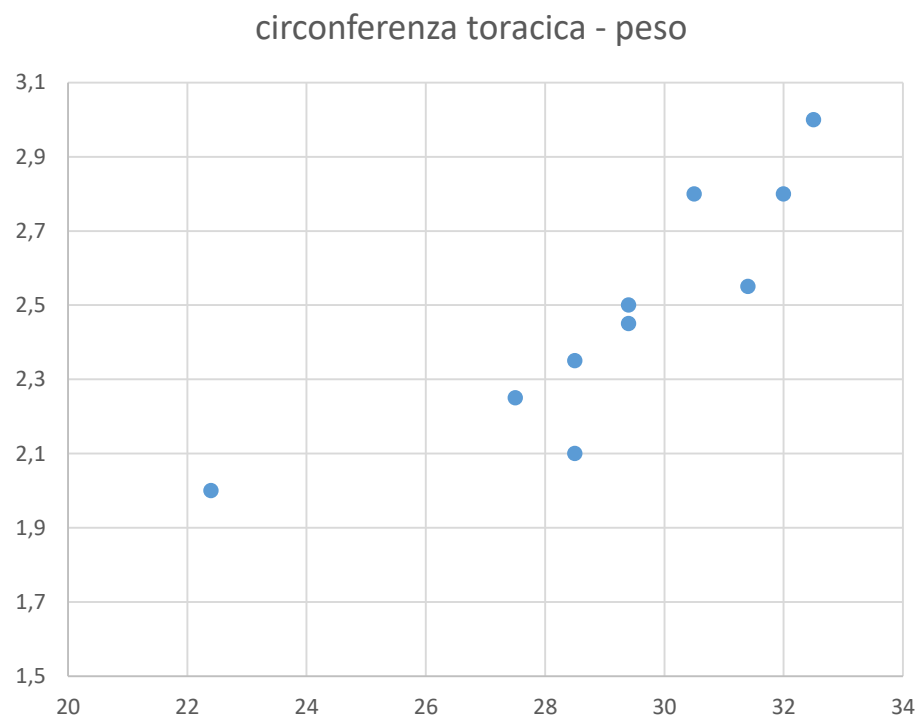
Il valore numerico del coefficiente di correlazione, r , è compreso tra -1 (perfetta correlazione negativa) e +1 (perfetta correlazione positiva) ...»

«... I coefficienti di correlazione possono essere calcolati attraverso metodi parametrici o non parametrici

Un coefficiente parametrico è il **coefficiente di correlazione di Pearson**.

Dei diversi coefficienti non parametrici, il **coefficiente di correlazione per ranghi di Spearman** è il più comunemente utilizzato ...»

circ toracica	peso
22,4	2
27,5	2,25
28,5	2,1
28,5	2,35
29,4	2,45
29,4	2,5
30,5	2,8
32	2,8
31,4	2,55
32,5	3



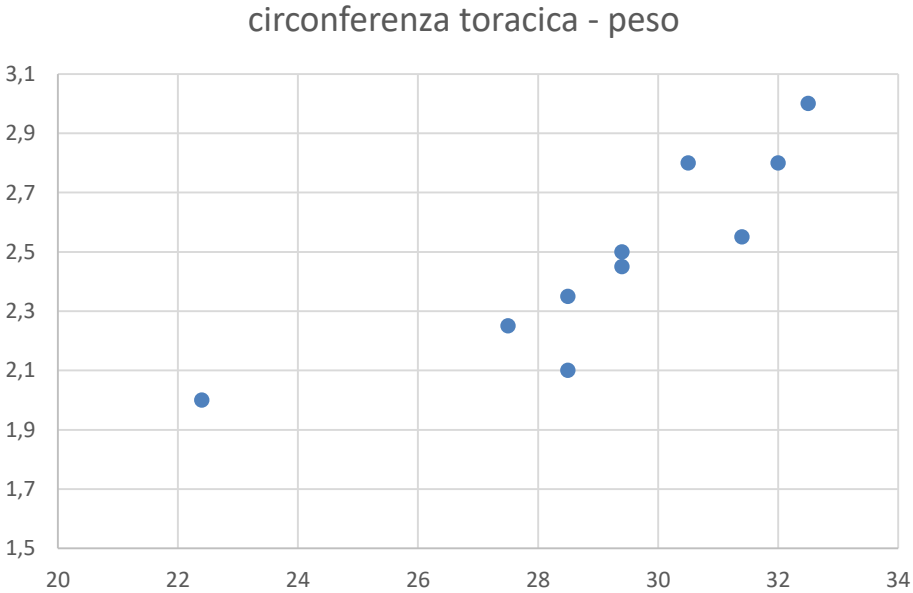
Excel: PEARSON()

r=	0,86
----	------

circ toracica	peso
22,4	2
27,5	2,25
28,5	2,1
28,5	2,35
29,4	2,45
29,4	2,5
30,5	2,8
32	2,8
31,4	2,55
32,5	3

socsistatistics					
r=	0,8604	r^2=	0,740288	p=	0,0014

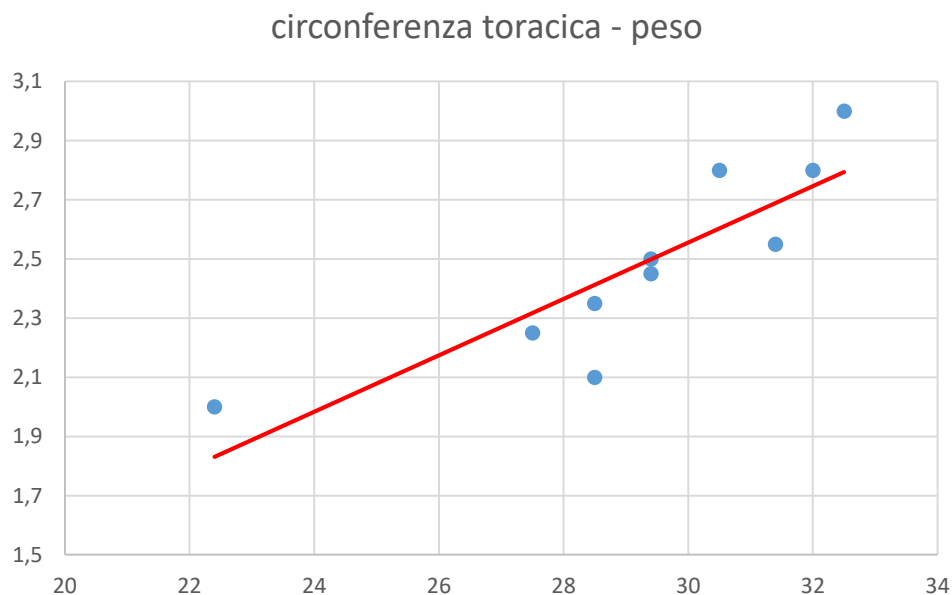
ci r=vassarstats	
0,504	0,966



Excel:PEARSON()	
r=	0,86

«... Nel presentare un diagramma a dispersione, è talvolta utile tracciare una linea retta all'interno dell'agglomerato di punti per illustrarne la relazione media.

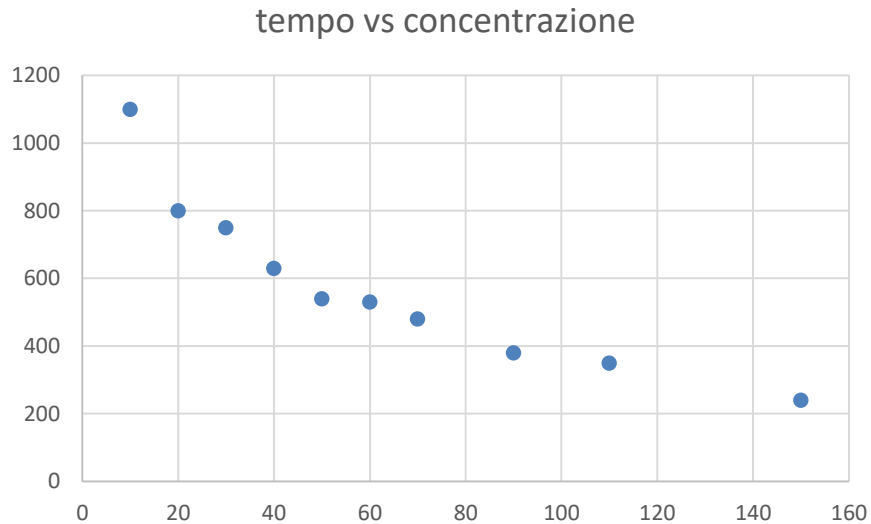
Tale retta, viene chiamata **retta di regressione** ...»



«... Molte relazioni in medicina/riabilitazione non sono però lineari...»

Come fare?

tempo	concentrazione del farmaco nel plasma
10	1100
20	800
30	750
40	630
50	540
60	530
70	480
90	380
110	350
150	240



tempo	concentrazione del farmaco nel plasma
10	1100
20	800
30	750
40	630
50	540
60	530
70	480
90	380
110	350
150	240

