



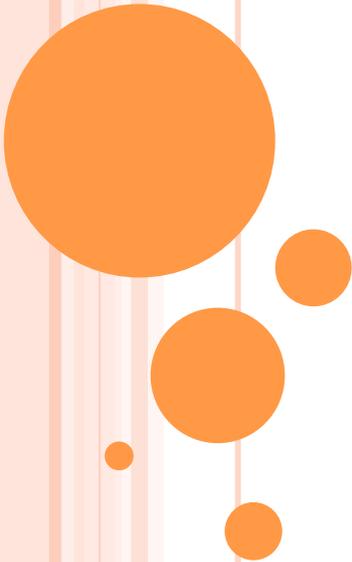
UNIVERSITÀ
DEGLI STUDI
DI MILANO



FONDAZIONE ROMEO
ED ENRICA INVERNIZZI

LA STATALE

PEDIATRIC CLINICAL
RESEARCH CENTER



BIOSTATISTICA: UN'INTRODUZIONE (M3- BIOSTAT-TEORIA)

Giorgio Bedogni

PROGRAMMA

- Sessione teorica
 - Martedì 16 maggio 2017
- Sessione pratica
 - Martedì 23 maggio 2017

PROGRAMMA

- *Sessione teorica*
 - Martedì 16 maggio 2017
- *Sessione pratica*
 - Martedì 23 maggio 2017

OBIETTIVO PRIMARIO

- Migliorare la comprensione dell'analisi statistica applicata agli studi clinici

OBIETTIVO SECONDARIO

- Mostrare come la statistica sia nata per risolvere problemi pratici (e renderla forse più amabile)

OBIETTIVO GENERALE



CAVEAT

- 4 ore di corso su un argomento così vasto rischiano di violare un principio sacrosanto:

“Make everything as simple as possible, but not simpler” (Albert Einstein)

<http://www.quotesdaddy.com/>

OK, l’ho violato, inutile negare...

PROGRAMMA

- Cosa è la biostatistica?
- Cosa è la statistica descrittiva?
- Cosa è la statistica inferenziale?

PROGRAMMA

- *Cosa è la biostatistica?*
- Cosa è la statistica descrittiva?
- Cosa è la statistica inferenziale?

COSA È LA BIO-STATISTICA?

- L'applicazione della statistica allo studio del mondo della vita*

*Ma non c'era bisogno di un corso per capirlo, vero?

È PENSIERO CRITICO

- “The study of statistics should serve to develop critical analytical thought and common sense as well to introduce specific tools and methods of processing data”

Van Belle G et al. Biostatistics. Wiley: 2004, p. 10.

NON È SOLO TECNICA

- “The need for lessons in scientific common sense is great in epidemiology, which is taught largely as an engineering discipline and practiced largely as technical tasks, making attention to core principles of scientific inquiry woefully rare”

Phillips CV et al. Emerging Themes in Epidemiology 2006; 3:5.

NON È SOLO TECNICA

- “If useful statistical methods could be implemented in software, surely would not the need for statisticians diminish?”
- “In fact, the reverse happened, for at least three reasons”

Diggle P. JRSS-A 2015;178:793 (RSS Presidential Address on Data Science).

NON È SOLO TECNICA

- “Firstly, if something is impossible it is easy to convince yourself that you can get by without it. Packages enabled scholars of many disciplines who might previously have considered statistics irrelevant to their subject to begin to appreciate its power”

Diggle P. JRSS-A 2015;178:793.

NON È SOLO TECNICA

- “Secondly, packages enabled statisticians to do more things routinely, again increasing the reach of statistics to other disciplines”
- “Thirdly, packages could not design studies”

Diggle P. JRSS-A 2015;178:793.

NON È MAGIA

- “Many people seem to believe that statistical calculations are magical and can reach conclusions that are much stronger than is actually possible. The phrase ‘statistically significant’ is seductive and often misinterpreted.”

Motulsky H. Intuitive Biostatistics. Oxford University Press: 1995, p. 7.

NON È IL CULTO DELLA “SIGNIFICATIVITÀ”

- “The emphasis given to formal tests of significance... has caused scientific research workers to pay undue attention to the results of the tests of significance they perform on their data, particularly data derived from experiments, and too little to the estimates of the magnitude of the effects they are investigating.” (Frank Yates)

<http://www.indiana.edu/~stigtsts/quotsagn.html>

NON È IL CULTO DELLA “SIGNIFICATIVITÀ”

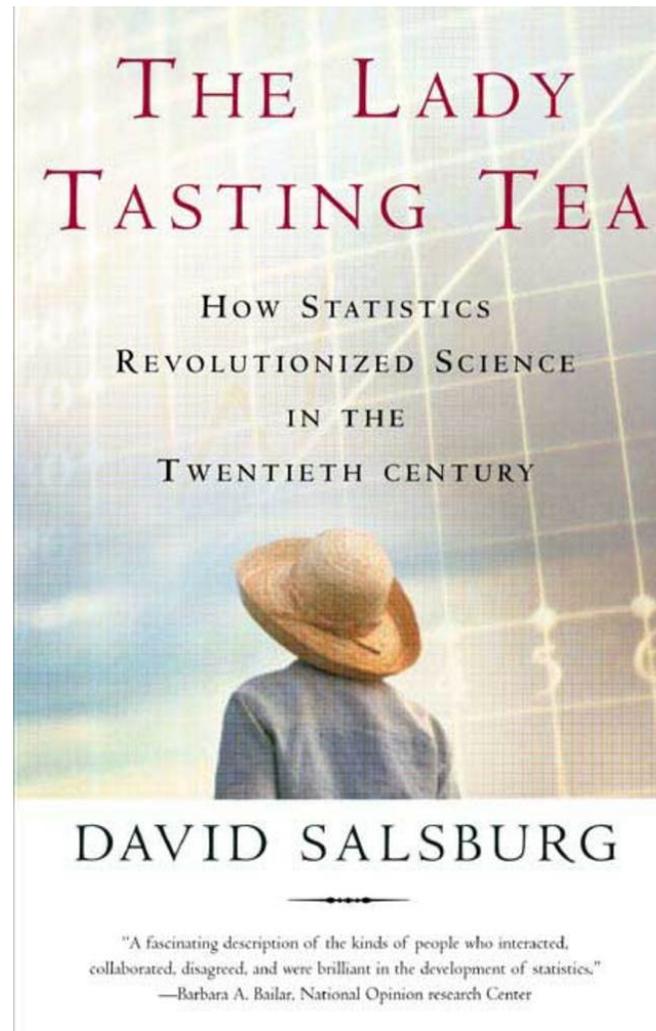
- Questo dovremmo averlo già chiaro (M1 e M2)
- Comunque, ci torneremo sopra ed approfondiremo (M3)

MA NON CI SONO I REFEREE PER QUESTO?

- “In an ideal world, editors of medical journals would do such an excellent job of ensuring the quality and accuracy of statistical methods of the papers they publish that readers with no personal interest in this aspect of research work could take it for granted that anything published was correct. If past history is any guide, however, we will probably never even approach that ideal.”

Glantz SA. Biostatistics. Mc-Graw Hill: 1981, p. 1 (il primo libro su cui GB ha studiato la statistica!).

QUAL È IL RUOLO DELLA STATISTICA NELLA SCIENZA?



www.giorgiobedogni.it

QUAL È IL RUOLO DELLA STATISTICA NELLA SCIENZA?

- “As we enter the twenty-first century the statistical revolution stands triumphant. It has vanquished determinism from all but a few obscure corners of science. It has become so widely used that its underlying assumptions have become part of the unspoken popular culture of the Western world”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 309.

OROLOGI VS. NUVOLE



QUAL È IL RUOLO DELLA STATISTICA NELLA SCIENZA?

- “It [statistics] stands triumphant on feet of clay. Somewhere, in the hidden corners of the future, another scientific revolution is waiting to overthrow it, and the men and women who will create that revolution may already be living among us”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 309.

QUAL È IL RUOLO DELLA STATISTICA NELLA SCIENZA?

A world without statistics

How would science, society and technology have evolved if we did not have statistics?
Andrew Gelman imagines what might have been.

october2014 | significance | **47**

QUAL È IL RUOLO DELLA STATISTICA NELLA SCIENZA?

- “But statistics is central to how we think about the world. I still think that it is much less significant to our lives than, say, chemistry. But it ain’t nothing”.

Gelman, A. Significance. 2014;11.

QUAL È IL RUOLO DELLA STATISTICA
NELLA SCIENZA?

An Address

ENTITLED

IS THE STATISTICAL METHOD OF ANY
VALUE IN MEDICAL RESEARCH ?

*Delivered in the Institute of Pathology and Research,
St. Mary's Hospital, on May 22nd, 1924,*

BY MAJOR GREENWOOD, F.R.C.P. LOND.,
MEDICAL OFFICER (MEDICAL STATISTICS), MINISTRY
OF HEALTH.

Greenwood, M. Lancet. 1924;204:153.

QUAL È IL RUOLO DELLA STATISTICA NELLA SCIENZA?

- “As the immediate collaborator of the experimenter, collaboration implying that he will take pains to understand the work of the experimenter, even to do some of it itself, the medical statistician has, I submit, an important part to play in modern research”
- “The statistician must be the equal not the predominant partner”

Greenwood M. Lancet. 1924;204:153.

COME POSSO IMPARARE LA BIO- STATISTICA?

- “To paraphrase Pythagoras, there is no royal road to biostatistics. You need to be involved. You need to work hard. You need to think. You need to analyze actual data”

Van Belle G et al. Biostatistics. Wiley: 2004, p. 10.

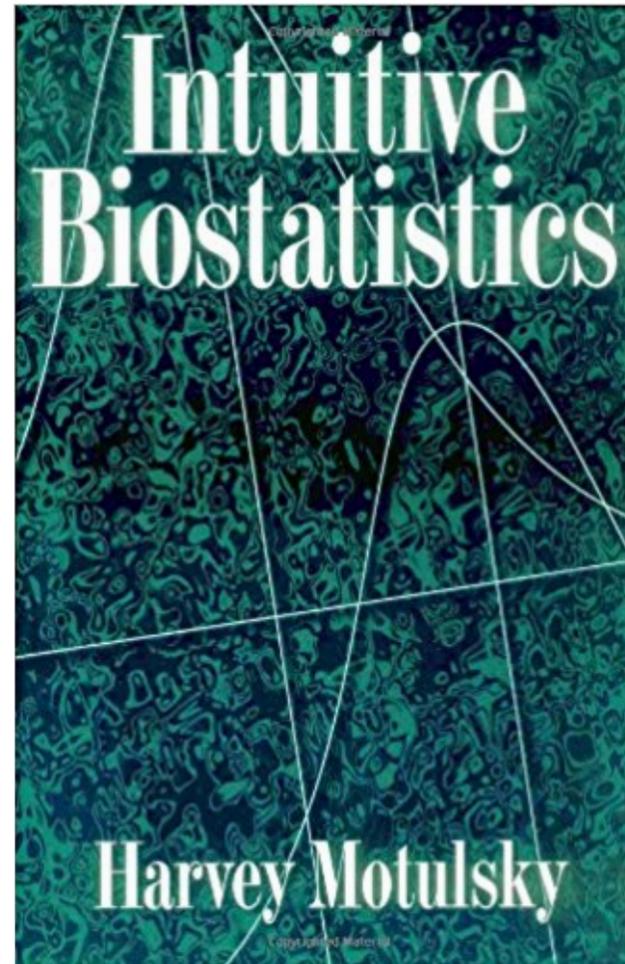
COME POSSO IMPARARE LA BIO- STATISTICA?



TESTO CONSIGLIATO 1

- Ottimo per iniziare

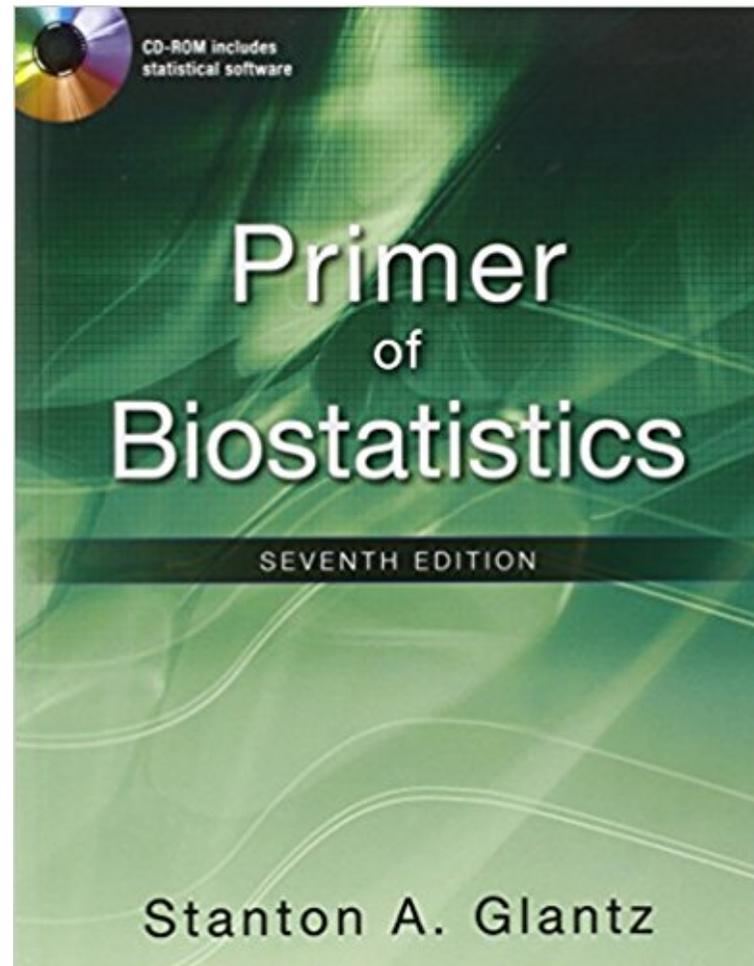
Motulsky H. Intuitive Biostatistics. Oxford University Press: 1995.



TESTO CONSIGLIATO 2

- Buon compromesso tra semplicità e rigore

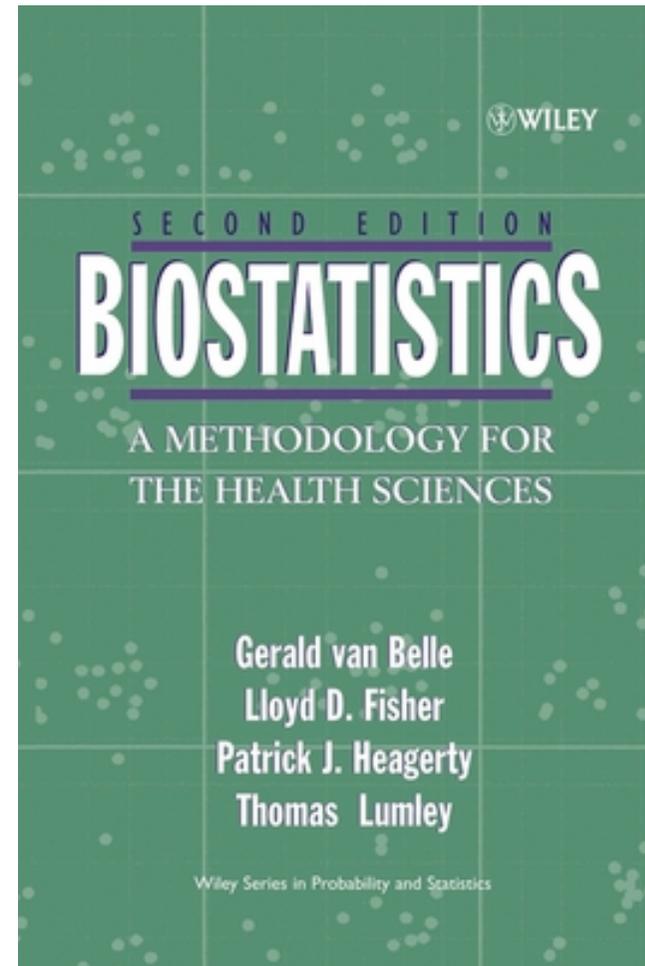
Glantz SA. Primer of biostatistics. Mc-Graw Hill: 2011.



TESTO CONSIGLIATO 3

- Rigoroso e ottimo per consultazione

Van Belle G et al.
Biostatistics. Wiley: 2004.



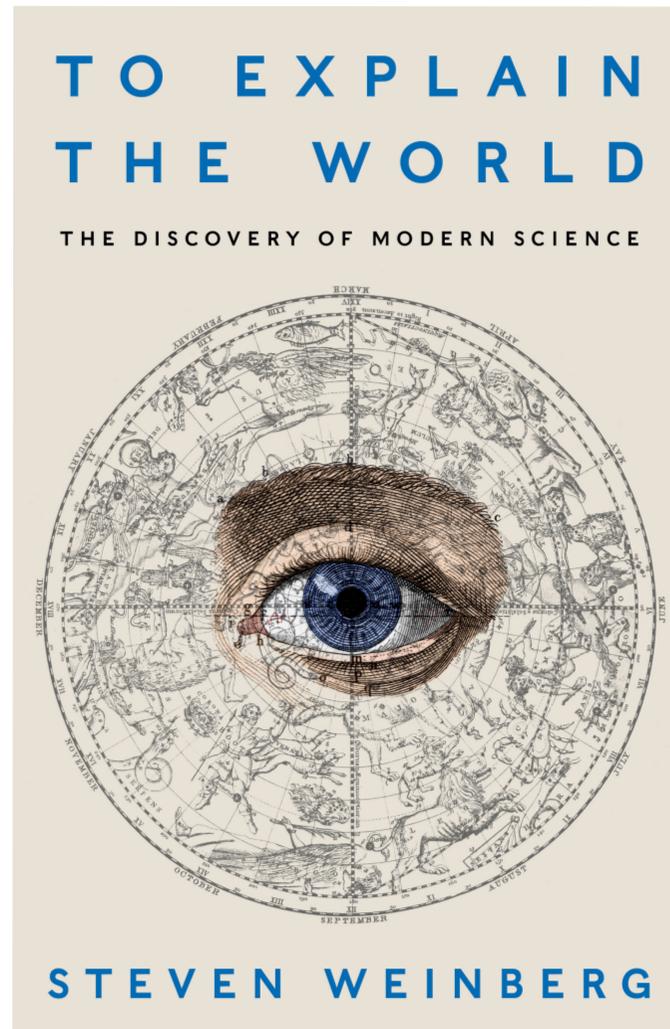
IL CONSIGLIO PIÙ IMPORTANTE

- Leggete i “classici” della biostatistica e dell’epidemiologia
- C’è una ragione per cui sono diventati dei “classici”!
- Vedremo qualche “classico” nel corso di questo modulo (uno lo abbiamo già visto)...

IL MONDO IN NUMERI

- Lo studio scientifico del mondo implica (?) la sua “riduzione” a numeri
- Tale riduzione è una parte importante del cosiddetto approccio “riduzionista”
- GB è il primo a riconoscere i limiti del modello riduzionista, che ha dei benefici, se (e solo se) lo si impiega criticamente

UN ESEMPIO DI RIDUZIONE CRITICO



Weinberg S. To Explain the World. Penguin: 2016.

36

UN ESEMPIO DI RIDUZIONISMO CRITICO

- “So the world like on us as a teaching machine, reinforcing our good ideas with moments of satisfaction”
- “It is all unplanned and unpredictable and gives us joy along the way”
- “We do not know how far we will continue on this reductive path”

Weinberg S. To Explain the World. Penguin: 2016, p. 255 ff.

IL MONDO IN NUMERI

- La (bio)statistica mira innanzitutto a fornire una descrizione sintetica delle caratteristiche numerabili del mondo (della vita)

SULL'USO E L'ABUSO DEI NUMERI

- “Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted” (Albert Einstein)*

<http://www.quotesdaddy.com/>

*Detto e ridetto sia in M1 sia in M2 ma ripeterlo è utile.

SULL'USO E L'ABUSO DEI MODELLI

- “The value of a model is that often it suggests a simple summary of the data in terms of the major systematic effects together with a summary of the nature or magnitude of random variations”

McCullagh P & Nelder, JA. Generalized linear models. Chapman and Hall: 1989.

SULL'USO E L'ABUSO DEI MODELLI

- “The most that can be expected from any model is that it can supply a useful approximation to reality. All models are wrong; some models are useful”.

Box G. Statistics for experimenters. Wiley: 2005.

SULL'USO E L'ABUSO DEI MODELLI

- “Analysis interpretation depends on contextual judgements about how reality is mapped onto the model and how the formal analysis results are mapped back to reality”

Greenland S. Eur J Epidemiol. 2017;32:3.

SULL'USO E L'ABUSO DEI MODELLI

- “Much benefit can accrue from thinking a problem through within these models as long as the formal logic is recognized as an allegory for a largely unknown reality”

Greenland S. Eur J Epidemiol. 2017;32:3.

SULL'USO E L'ABUSO DEI MODELLI

- “A tragedy of statistical theory is that it pretends as if mathematical solutions are not only sufficient but ‘optimal’ for dealing with analysis problems when the claimed optimality itself is derived from dubious assumptions”

Greenland S. Eur J Epidemiol. 2017;32:3.

SULL'USO E L'ABUSO DEI MODELLI

- “There is sometimes a tendency to look at where we are now, look back at how we got here, and to see ourselves as at the culmination of a process of development. As far as statistical ideas and methods go, this would be a mistake: the evolution is continuing”

Hand DJ. Phil Trans R Soc A. 2015;373:2014025.

SULL'USO E L'ABUSO DEI MODELLI

- “When we study logic, we learn that a valid logical argument can lead to incorrect conclusions if the premises are wrong. In an analogous way, a coherent argument based on incorrect premises is of limited value. We statisticians must always remember that our aim is to draw conclusions about the real world, based on models. Our models are not the reality”

Hand DJ. Phil Trans R Soc A. 2015;373:2014025.

INTRODUCIAMO ALCUNI TERMINI

- Variabile: quantità che varia da (s)oggetto a (s)oggetto
- Campione: raccolta di valori di una o più variabili
- Popolazione (“spazio campionario”): l’insieme di tutti i valori possibili di una o più variabili

INTRODUCIAMO ALCUNI TERMINI

- Statistica: caratteristica numerica di un campione
- Parametro: caratteristica numerica di una popolazione

PROGRAMMA

- Cosa è la biostatistica?
- *Cosa è la statistica descrittiva?*
- Cosa è la statistica inferenziale?

STATISTICA DESCRITTIVA

- Basata sui percentili
- Basata sui momenti

STATISTICA DESCRITTIVA

- *Basata sui percentili*
- Basata sui momenti

DEFINIZIONE INTUITIVA DI PERCENTILE

- Il 3° percentile è il valore di una variabile tale che il 3% delle sue osservazioni sono inferiori e il 97% superiori a quel valore
- Cosa significa che un bambino è al 3° percentile della statura per (sesso ed) età?
- E al 97° percentile dell'indice di massa corporea per (sesso ed) età?

DEFINIZIONE RIGOROSA (MA SEMPLIFICATA) DI PERCENTILE

- Il p -esimo percentile di un campione di n osservazioni è il valore della variabile con rango R :

$$R = \frac{P}{100} \cdot (1 + n)$$

DEFINIZIONE RIGOROSA (MA SEMPLIFICATA) DI PERCENTILE

- Qual è il 50° percentile di questi 5 valori di alanina-aminotransferasi (U/L): 30, 35, 20, 15, 40?
- Ordino per ranghi: 15 (1), 20 (2), 30 (3), 35 (4), 40 (5) e calcolo il rango corrispondente al 50° percentile:

$$R = \frac{50}{100} \cdot (1 + 5) = ?$$

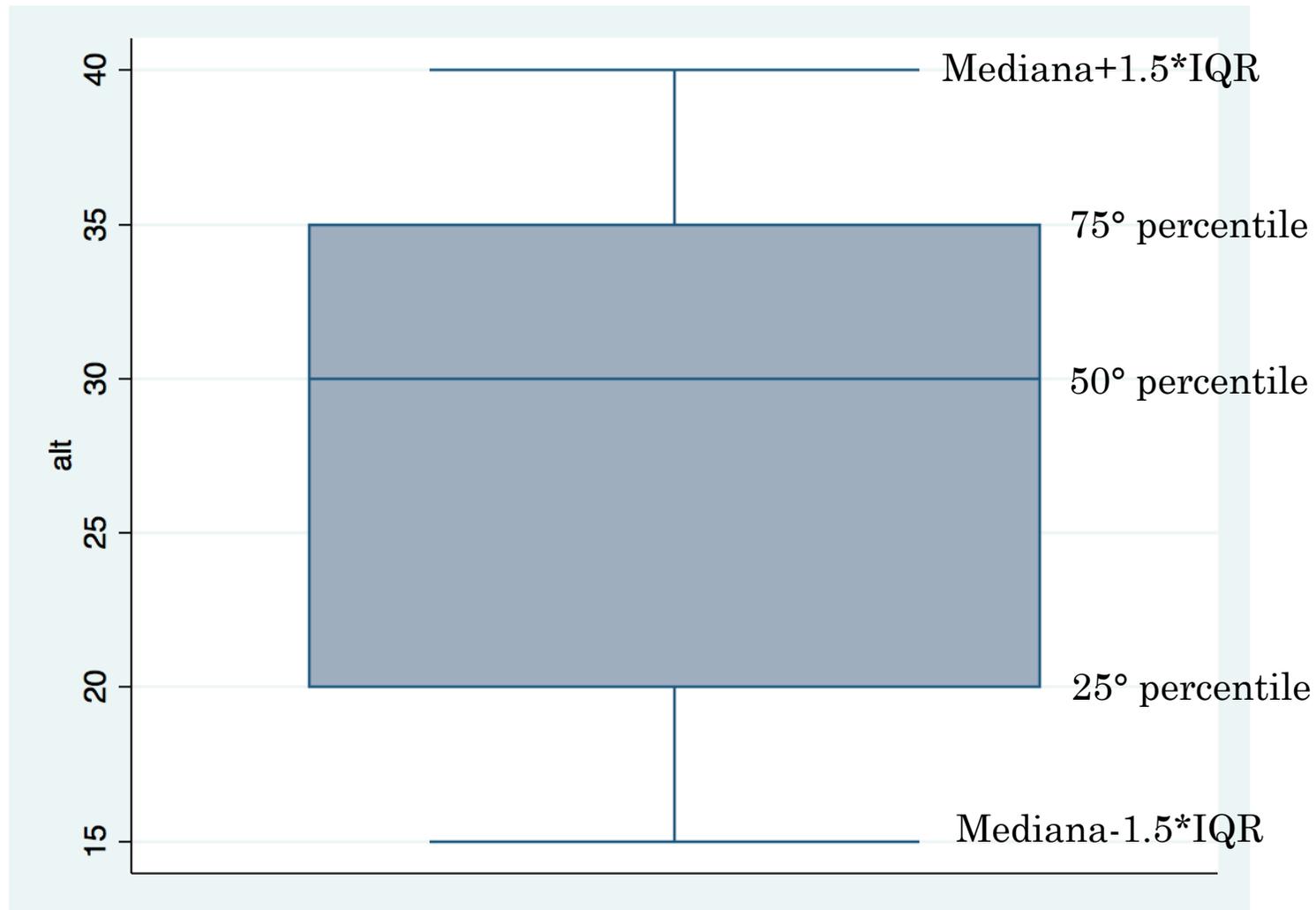
ALTRI NOMI DI ALCUNI PERCENTILI

50°	Mediana
25°	Quartile inferiore
75°	Quartile superiore

STATISTICHE BASATE SUI PERCENTILI

- Misure di locazione (“location”)
 - Mediana
 - Valore minimo
 - Valore massimo
- Misure di dispersione (“dispersion”)
 - Range interquartile (IQR)
 - = 75° percentile - 25° percentile
 - Range
 - = valore massimo - valore minimo

IL BOX-PLOT DI TUKEY



UN “CLASSICO”

THE FUTURE OF DATA ANALYSIS¹

BY JOHN W. TUKEY

Princeton University and Bell Telephone Laboratories

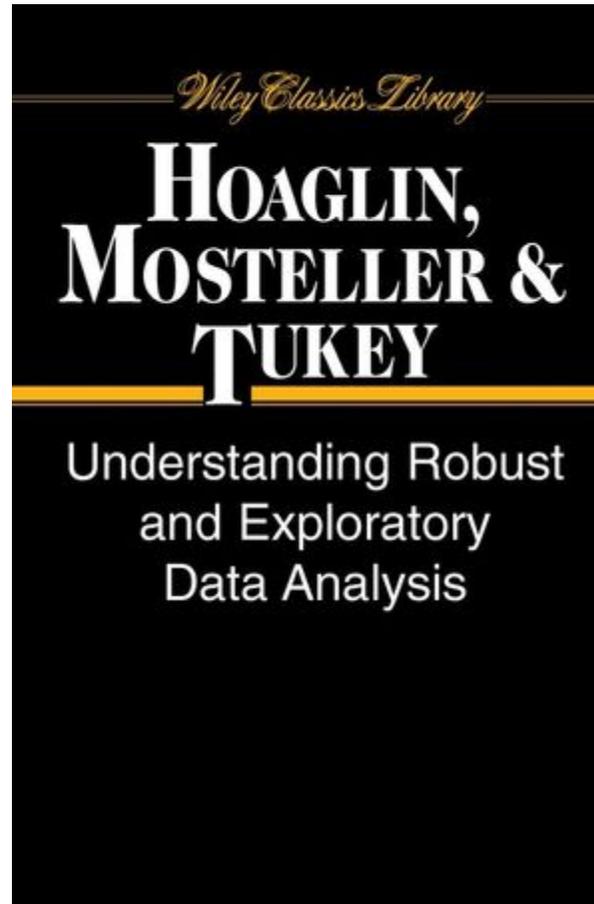
Tukey, J. Ann Math Statist. 1962;33:1.

UN “CLASSICO”

- “We need to give up the vain hope that data analysis can be found on a logico-deductive system like Euclidean plane and to face up the fact that data analysis is intrinsically an empirical science”

Tukey J. Ann Math Statist. 1962;33:1.

UN “CLASSICO”



Hoaglin DC et al. Understanding robust and exploratory data analysis.
New York: Wiley 1982.

www.giorgiobedogni.it

STATISTICA DESCRITTIVA

- Basata sui percentili
- *Basata sui momenti*

STATISTICHE BASATE SUI MOMENTI

- Misure di locazione
 - Media (“primo momento”)
 - Valore minimo
 - Valore massimo
- Misure di dispersione
 - Deviazione standard (radice quadrata della varianza ovvero il “secondo momento”)
 - Range

MEDIA (ARITMETICA)

$$x \text{ (ALT, U/L)} = (30, 35, 20, 15, 40)$$

$$\bar{x} = \frac{\sum x}{n} = \frac{30 + 35 + 20 + 15 + 40}{5} = 28$$

DEVIAZIONE STANDARD (FORMULA PER CAMPIONE)

$$x \text{ (ALT, U/L)} = (30, 35, 20, 15, 40)$$

$$\bar{x} = 28$$

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} =$$

$$s = \sqrt{\frac{(30 - 28)^2 + (35 - 28)^2 + (20 - 28)^2 + (15 - 28)^2 + (40 - 28)^2}{5 - 1}} =$$

$$= \sqrt{\frac{4 + 49 + 64 + 169 + 144}{4}} = \sqrt{\frac{430}{4}} = 10$$

COSA COMPONE LA DEVIAZIONE STANDARD?
(FORMULA PER CAMPIONE)

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

SCARTO (IN ROSSO)
(FORMULA PER CAMPIONE)

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

SCARTO QUADRATICO (IN ROSSO) (FORMULA PER CAMPIONE)

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

DEVIANZA (IN ROSSO)
(FORMULA PER CAMPIONE)

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

VARIANZA (IN ROSSO)
(FORMULA PER CAMPIONE)

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

DEVIAZIONE STANDARD (IN ROSSO)
(FORMULA PER CAMPIONE)

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

FERMIAMOCI A RIFLETTERE

- Perché elevo gli “scarti” al quadrato per poi estrarre la radice quadrata?

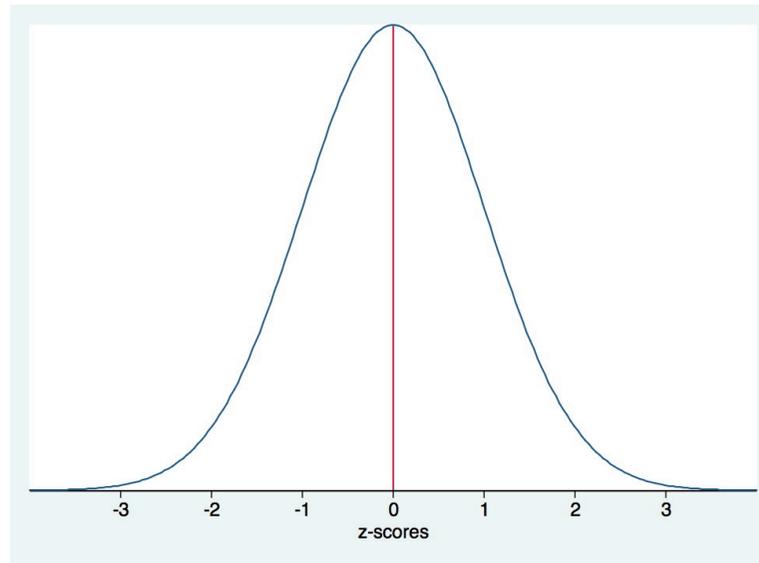
<https://stats.stackexchange.com/questions/118/why-square-the-difference-instead-of-taking-the-absolute-value-in-standard-devia>

- Perché c'è $(n-1)$ anziché n al denominatore?

<https://stats.stackexchange.com/questions/16008/what-does-unbiasedness-mean/16009#16009>

QUANDO POSSO UTILIZZARE MEDIA E DEVIAZIONE STANDARD?

- Quando la variabile di interesse ha una distribuzione Gaussiana (o “normale”)



LA FUNZIONE GAUSSIANA
(RESPIRATE PROFONDAMENTE E
CONCENTRATEVI SULLE PARTI IN ROSSO)



$$G(x) = \frac{1}{s\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{(x - \bar{x})}{s} \right)^2 \right]$$

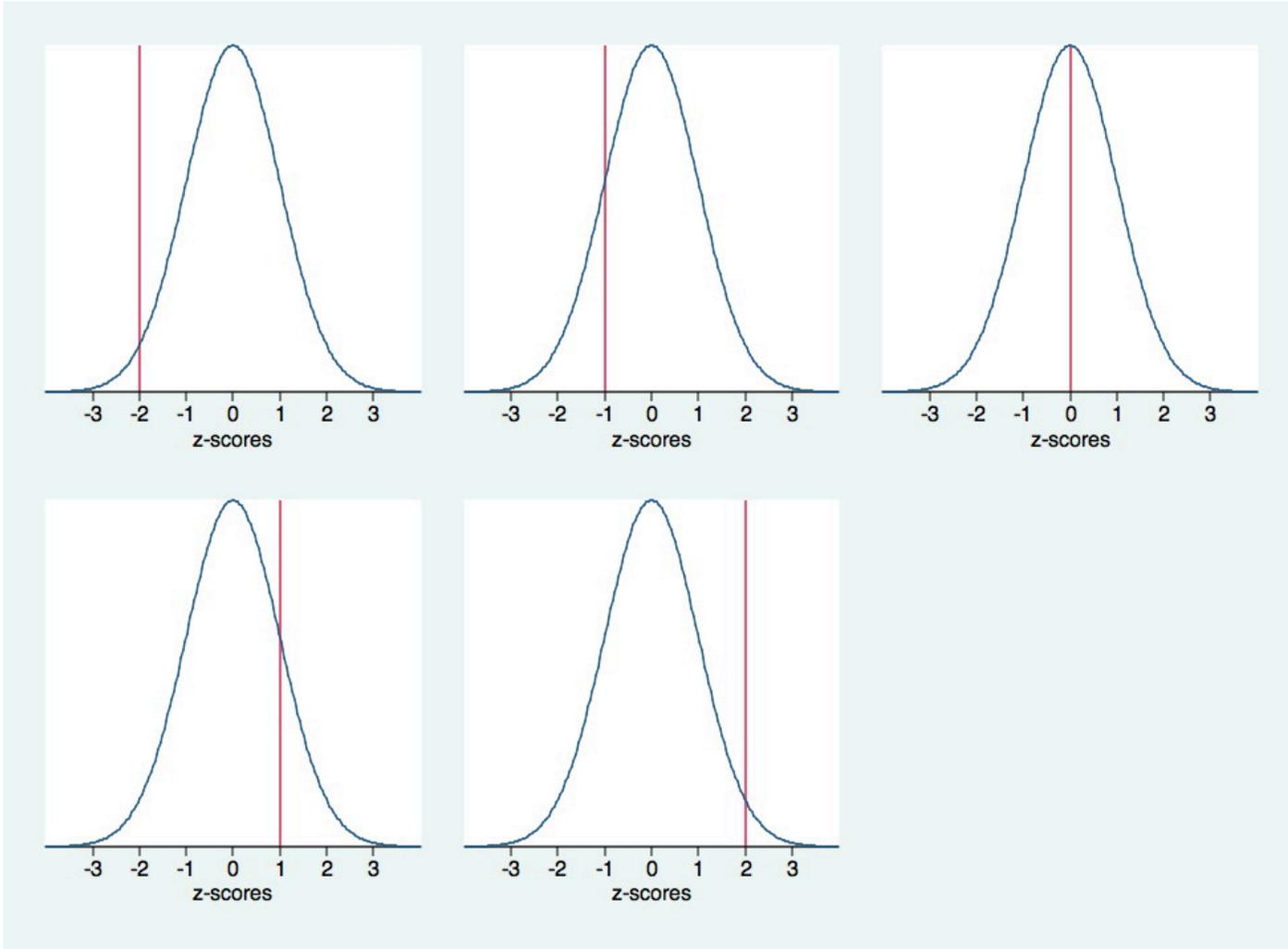
LO “Z-SCORE” (IN ROSSO)

$$G(x) = \frac{1}{s\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{(x - \bar{x})}{s} \right)^2 \right]$$

DA Z-SCORE A PERCENTILE

Z-score	Percentile
-2.00	0.0227
-1.96	0.0249
-1.00	0.1587
0.00	0.5000
1.00	0.8413
1.96	0.9750
2.00	0.9772

DOVE SIAMO?



ESERCIZIO

- Misuriamo la statura di 5 di voi scelti a caso ^{*,**}
- Calcoliamo mediana e range interquartile
- Calcoliamo media e deviazione standard

* Tra poco capiremo perché è importante la scelta casuale

** In questa sede, non ci importa l'effetto del sesso e dell'età e facciamo finta che la popolazione sia omogenea

NOTA SULLA DISTRIBUZIONE GAUSSIANA

- Quante variabili in natura sono distribuite secondo la Gaussiana? In senso stretto nessuna perché la Gaussiana comprende valori negativi.
- La statura è offerta come prototipo di variabile distribuita normalmente. Riuscite a concepire una statura negativa?
- È vero che la distribuzione Gaussiana “fitta” la statura ma non dobbiamo dimenticare l’assunzione implausibile della distribuzione!

NOTA SULLA DISTRIBUZIONE GAUSSIANA

Testing for Normality

R. C. Geary

Biometrika, Vol. 34, No. 3/4. (Dec., 1947), pp. 209-242.

Stable URL:

<http://links.jstor.org/sici?sici=0006-3444%28194712%2934%3A3%2F4%3C209%3A%3E2.0.CO%3B2-9>

Biometrika is currently published by Biometrika Trust.



NOTA SULLA DISTRIBUZIONE GAUSSIANA

- “Amends might be made in the interest of the new generation of students by printing in leaded type in future editions of existing text-books and in all new text-books:

Normality is a myth; there never was, and never will be, a normal distribution

This is an over-statement from the practical point of view, but it represents a safer initial mental attitude than any in fashion during the past two decades”

NOTA SULLA DISTRIBUZIONE GAUSSIANA

- $2017 - 1947 = 70$ anni trascorsi dal paper di Geary
- Avete mai incontrato questa “warning label”?

NOTA SULLA DISTRIBUZIONE GAUSSIANA

- Prendiamo questi valori di media e deviazione standard ottenuti da un paper: insulina = 296 ± 172 pmol/l
- Calcoliamo il 2.5° percentile: $296 - 1.96 * 172 = -41$ pmol/l. Insulina negativa?!
- Saranno anche pochi valori ma è fisiologicamente impossibile. Dobbiamo davvero descrivere i nostri dati con una distribuzione inadeguata?

NOTA SULLA DISTRIBUZIONE GAUSSIANA

- La distribuzione log-Gaussiana (“la cenerentola delle distribuzioni”) è superiore alla Gaussiana in molte circostanze (e “fitta” la statura altrettanto bene della Gaussiana)

Limpert E & Stahel WA. Significance. 2017;14:8 (leggete la bibliografia qui contenuta se volete approfondire).

NOTA SULLA DISTRIBUZIONE GAUSSIANA

- I percentili sono comunque validi
- Per quello che conta: adesso capite perché i percentili compaiono così spesso nei lavori di GB

PROGRAMMA

- Cosa è la biostatistica?
- Cosa è la statistica descrittiva?
- *Cosa è la statistica inferenziale?*

DAL CAMPIONE ALLA POPOLAZIONE

- La media e la deviazione standard ottenute da un campione casuale di una popolazione offrono una stima della media e della deviazione standard della popolazione originaria
- Tale stima varia da campione a campione. Ne siamo convinti? Tutto quello che segue dipende da questo.

DAL CAMPIONE ALLA POPOLAZIONE

	Campione ("Statistiche")	Popolazione ("Parametri")
Media	\bar{x}	μ
Deviazione standard	s	σ

DAL CAMPIONE ALLA POPOLAZIONE

$$\mu = \frac{\sum x}{n} \quad \text{da} \quad \bar{x} = \frac{\sum x}{n}$$

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \quad \text{da} \quad s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

PERCHÉ (N-1) AL DENOMINATORE DELLA DEVIAZIONE STANDARD DEL CAMPIONE?

- La ragione è correlata al concetto di “grado di libertà” (degree of freedom, d.f.)
- I gradi di libertà sono i valori che possono variare nel calcolo finale di una statistica (spiegazione matematica omessa)
- La varianza è calcolata dalla media, che ha n gradi di libertà. Poiché calcolo la varianza dalla media, “perdo” 1 grado di libertà sicchè mi restano $n-1$ gradi di libertà.

DAL CAMPIONE ALLA POPOLAZIONE

- Per quantificare l'accuratezza della stima della media si calcola l'errore standard della media
- (Si può calcolare l'errore standard di qualsiasi statistica)

ERRORE STANDARD DELLA MEDIA

- Consideriamo una popolazione di 100 soggetti (qualsiasi numero andrebbe bene)
- Campioniamo 10 di questi soggetti (qualsiasi numero andrebbe bene) in maniera casuale
- La media e la deviazione standard dei 10 soggetti scelti a caso offrono una stima della media e della deviazione standard dei 100 soggetti

ERRORE STANDARD DELLA MEDIA

- È intuitivo che i valori di un campione casuale di 10 individui differiranno almeno un po' da quelli di un altro campione casuale di 10 individui. O no?
- Adesso pensiamo di campionare tutti i campioni casuali di 10 individui di quella popolazione

ERRORE STANDARD DELLA MEDIA

- Fatto?* Bene, adesso calcoliamo la media delle medie di tutti i campioni casuali di 10 soggetti (media delle medie campionarie)
- La deviazione standard delle medie campionarie è l'errore standard della media (delle medie campionarie)

*Facciamoci i complimenti: i campioni sono $100!/(100-10)! = 62815650955529472000 = 6.28157e+19$.

ERRORE STANDARD DELLA MEDIA (POPOLAZIONE)

- L'errore standard aumenta all'aumentare della dispersione dei valori
- L'errore standard si riduce all'aumentare di n (notate che n è sotto radice quadrata)

(Spiegazione matematica omessa)

$$\sigma_{\mu} = \frac{\sigma}{\sqrt{n}}$$

ERRORE STANDARD DELLA MEDIA (CAMPIONE)

- Questa è la nostra migliore stima dell'errore standard della popolazione

(Spiegazione matematica omessa)

$$S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

IL “TEOREMA DEL LIMITE CENTRALE” (MOLTO SEMPLIFICATO...)

- La media delle medie campionarie ha una distribuzione approssimativamente Gaussiana indipendentemente dalla distribuzione dei valori nella popolazione

<https://stats.stackexchange.com/questions/643/how-do-you-convey-the-beauty-of-the-central-limit-theorem-to-a-non-statistician/11474#11474>

INTERVALLI DI CONFIDENZA DELLA MEDIA

- Gli intervalli di confidenza di una media campionaria misurano l'incertezza della stima della media della popolazione
- Un intervallo di confidenza della media al 95% ci dice che, se dovessimo ripetere la misura su 100 campioni indipendenti, la media della popolazione sarebbe inclusa in quest'intervallo nel 95% dei casi (richiede un'interpretazione frequentista della probabilità)

INTERVALLI DI CONFIDENZA AL 95% DELLA MEDIA

$$\bar{X} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

\bar{X} = media campionaria

σ = deviazione standard della popolazione

n = numerosità del campione

UN “CLASSICO”

ON THE TWO DIFFERENT ASPECTS OF THE REPRESENTATIVE METHOD :
THE METHOD OF STRATIFIED SAMPLING AND THE METHOD
OF PURPOSIVE SELECTION.

By JERZY NEYMAN

(Biometric Laboratory, Nencki Institute, Soc. Sci. Lit.
Varsoviensis, Warsaw).

[Read before the Royal Statistical Society, June 19th, 1934, the PRESIDENT,
the RT. HON. LORD MESTON of Agra and Dunottar, K.C.S.I., LL.D.,
in the Chair.]

Neyman J. JRSS 1934, 97:558.

UN “CLASSICO”

- “The form of this solution consists in determining certain intervals, which I propose to call the confidence intervals (see Note I), in which we may assume are contained the values of the estimated characters of the population, the probability of an error in a statement of this sort being equal to or less than $1 - \varepsilon$, where ε is any number $0 < \varepsilon < 1$, chosen in advance”

Neyman J. JRSS 1934, 97:558.

IMPORTANTE

- Deviazione standard e errore standard della media non sono sinonimi
- La deviazione standard è una misura della variabilità nel campione
- L'errore standard offre una misura dell'incertezza nella stima della media della popolazione

VALIDITÀ IN ASSENZA DELLA GAUSSIANA

- La statistica “parametrica” potrebbe essere comunque valida, specialmente se il campione è ampio*
- Il teorema del limite centrale autorizza a usare la statistica “parametrica” per fare inferenze sulla popolazione ma la normalità non deve essere estesa al campione
- *Nessuno statistico si sbilancerà per definire operativamente tale ampiezza...

PROBABILITÀ

- Cosa è una probabilità?
- Un oggetto matematico molto ben definito (“assiomatizzato”)
- Ma a cosa corrisponde nel mondo reale?
 - Il problema è sempre mappare l’oggetto matematico al mondo reale senza esagerare

LA RAGIONE DEI “PIEDI DI ARGILLA”

- “It [statistics] stands triumphant on feet of clay. Somewhere, in the hidden corners of the future, another scientific revolution is waiting to overthrow it, and the men and women who will create that revolution may already be living among us”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 309.

Surrogate Science: The Idol of a Universal Method for Scientific Inference

Gerd Gigerenzer

Max Planck Institute for Human Development

Julian N. Marewski

University of Lausanne

Journal of Management

Vol. 41 No. 2, February 2015 421–440

DOI: 10.1177/0149206314547522

© The Author(s) 2014

The application of statistics to science is not a neutral act. Statistical tools have shaped and were also shaped by its objects. In the social sciences, statistical methods fundamentally changed research practice, making statistical inference its centerpiece. At the same time, textbook writers in the social sciences have transformed rivaling statistical systems into an apparently monolithic method that could be used mechanically. The idol of a universal method for scientific inference has been worshipped since the “inference revolution” of the 1950s. Because no such method has ever been found, surrogates have been created, most notably the quest for significant p values. This form of surrogate science fosters delusions and borderline cheating and has done much harm, creating, for one, a flood of irreproducible results. Proponents of the “Bayesian revolution” should be wary of chasing yet another chimera: an apparently universal inference procedure. A better path would be to promote both an understanding of the various devices in the “statistical toolbox” and informed judgment to select among these.

Beyond subjective and objective in statistics

Andrew Gelman

Columbia University, New York, USA

and Christian Hennig

University College London, UK

[*Read before The Royal Statistical Society on Wednesday, April 12th, 2017, Professor P. J. Diggle in the Chair*]

Summary. Decisions in statistical data analysis are often justified, criticized, or avoided by using concepts of objectivity and subjectivity. We argue that the words ‘objective’ and ‘subjective’ in statistics discourse are used in a mostly unhelpful way, and we propose to replace each of them with broader collections of attributes, with objectivity replaced by *transparency*, *consensus*, *impartiality* and *correspondence to observable reality*, and subjectivity replaced by awareness of *multiple perspectives* and *context dependence*. Together with *stability*, these make up a collection of virtues that we think is helpful in discussions of statistical foundations and practice. The advantage of these reformulations is that the replacement terms do not oppose each other and that they give more specific guidance about what statistical science strives to achieve. **Instead of debating over whether a given statistical method is subjective or objective (or normatively debating the relative merits of subjectivity and objectivity in statistical practice), we can recognize desirable attributes such as transparency and acknowledgement of multiple perspectives as complementary goals.** We demonstrate the implications of our proposal with recent applied examples from pharmacology, election polling and socio-economic stratification. The aim of the paper is to push users and developers of statistical methods towards more effective use of diverse sources of information and more open acknowledgement of assumptions and goals.

ALCUNE INTERPRETAZIONI DELLA PROBABILITÀ

- Frequentista
 - Quella che abbiamo considerato in questo corso ma
ciò non implica che sia l'unica o la più importante
- Bayesiana soggettivista
- Bayesiana oggettivista
- Bayesiana falsificazionista

Gelman, A & Hennig, C. JRSS-B. 104:1

INTERPRETAZIONI DELLA PROBABILITÀ

- Tutte le interpretazioni della probabilità hanno dei limiti. L'importante è conoscerli e non schierarsi in maniera dogmatica

http://www.phil.vt.edu/dmayo/personal_website/

<http://andrewgelman.com/>

<https://www.ucl.ac.uk/sts/staff/gillies>

PROBABILITÀ COME FREQUENZA RELATIVA

- Qual è la probabilità di ottenere “6” con un dado? E quella di ottenere “testa” lanciando una moneta?
- Noi però non sappiamo quale sarà il risultato (outcome) al prossimo lancio.
- Secondo l'interpretazione frequentista, la probabilità è dunque caratteristica di una popolazione di outcome (esiti del lancio del dado, della moneta, ecc.)

PROBABILITÀ COME FREQUENZA RELATIVA

- Il caso del dado e della moneta potrebbe non avervi convinto. In fondo hanno una propensione “intrinseca” a fornire outcome equiprobabili.
- D'accordo ma come fate a verificare se un dado o una moneta è truccata?

PROBABILITÀ COME FREQUENZA RELATIVA

- Ma c'è di meglio: pensate di lanciare un “ossicino” come si faceva con gli astragali di quadrupedi prima che venissero inventati i dadi.
- Come fate a valutare la probabilità degli outcome? Non potete certo contare sulla natura fisica del materiale!

PROBABILITÀ COME FREQUENZA RELATIVA



- “Usiamo un ossicino di 6000 anni con 4 facce (non è quello rappresentato a fianco)
- Distribuzione degli outcome rilevata dagli studenti di Ian Hacking: faccia non marcata = 110, faccia blu = 88, faccia = nera: 52, faccia rossa = 50.

Hacking I. An introduction to probability and inductive logic. Cambridge University Press: 2001.

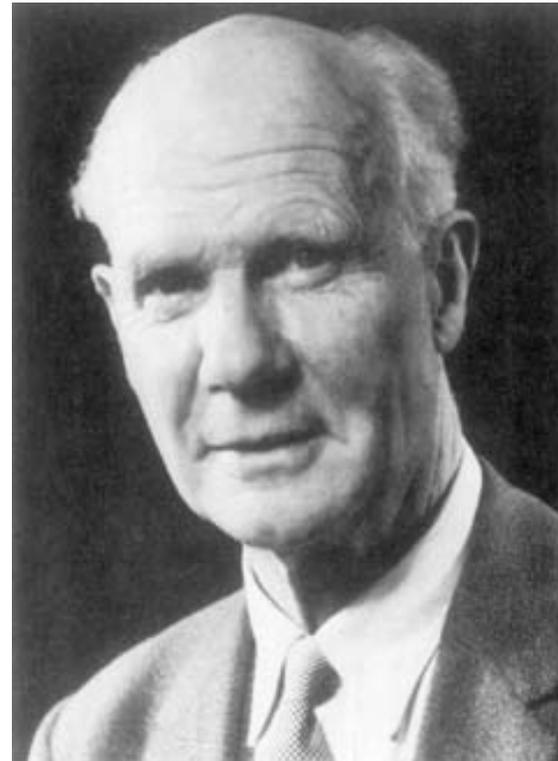
INGREDIENTI DELLA PROBABILITÀ IN CHIAVE FREQUENTISTA

- Un esperimento (trial) ha un set predefinito di outcome
- L'outcome di un trial non influenza quello di un altro trial
- I trial sono identici
- La probabilità di un outcome è il limite della sua frequenza relativa quando il numero di trial tende all'infinito

TEST DI IPOTESI – IL PARADIGMA NEYMAN-PEARSON



Jerzy Neyman
(1894-1981)



Egon Sharpe Pearson
(1895-1980)

JERZY NEYMAN
(1894-1981)



www.giorgiobedogni.it

JERZY NEYMAN – IL MATEMATICO

- “It is a pleasure to read his papers. The ideas evolve naturally, the notation is deceptively simple, and the conclusions appear so natural that you find it hard to see why no one has produced these results before”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 104.

JERZY NEYMAN – L'UOMO

- “Neyman was always attentive to his students and fellow faculty members. They describe the pleasures of the afternoon department teas... He would gently prod someone, student or faculty, to describe some recent research and then generally work his way around the room, getting comments and aiding the discussion”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013.

THIS IS THE PLACE!



JERZY NEYMAN – L’UOMO

- “In particular he was helpful and encouraging to many women supervising their PhD thesis, publishing papers jointly with them, and finding prominent places for them in the academic community”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 198.

TEST DI IPOTESI – IL PARADIGMA NEYMAN-PEARSON

- “A simplified version of the Neyman-Pearson formulation of hypothesis testing can now be found in all elementary statistics textbooks. This is how it must be done, the text implies, and this is the only way it can be done”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 109.

TEST DI IPOTESI – IL PARADIGMA NEYMAN-PEARSON

- “This rigid approach to hypothesis testing has been accepted by regulatory agencies such as FDA and the EPA and is taught in medical schools to future generation of medical researchers”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 109.

TEST DI IPOTESI – IL PARADIGMA DI FISHER



Ronald Aylmer Fisher
(1890-1962)

www.giorgiobedogni.it

LA RISPOSTA DI FISHER A NEYMAN

- “Neyman, thinking he was correcting and improving my own early work on tests of significance, as a means to the ‘improvement of natural knowledge’, in fact reinterpreted them in terms of that technological and commercial apparatus which is known as acceptance procedure”

(Fisher che commenta Neyman)

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013, p. 113.

LA RISPOSTA DI FISHER A NEYMAN

- “For the logical fallacy of believing that a hypothesis has been proved to be true, merely because it is not contradicted by available facts, has no more right to insinuate itself in statistical than in scientific reasoning”

(Fisher che commenta Neyman)

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company: 2013. p. 108.

DALLA PENNA DI JERZY NEYMAN (1976)

COMMUN. STATIST.-THEOR. METH., A5(8), 737-751 (1976)

TESTS OF STATISTICAL HYPOTHESES AND
THEIR USE IN STUDIES OF NATURAL PHENOMENA

Jerzy Neyman

Statistical Laboratory
University of California, Berkeley, CA 94720

125

DALLA PENNA DI JERZY NEYMAN (1976)

ABSTRACT

Contrary to ideas suggested by the title of the conference at which the present paper was presented, the author is not aware of a conceptual difference between a "test of a statistical hypothesis" and a "test of significance" and uses these terms interchangeably. A study of any serious substantive problem involves a sequence of incidents at which one is forced to pause and consider what to do next. **In an effort to reduce the frequency of misdirected activities one uses statistical tests.** The procedure is illustrated on two examples: (i) Le Cam's (and associates') study of immunotherapy of cancer and (ii) a socio-economic experiment relating to low-income homeownership problems.

DALLA PENNA DI JERZY NEYMAN (1976)

(i) In the post Abraham Wald terminology, the problem of testing a statistical hypothesis H is a "two-decision problem": (a) either reject H or (b) do not reject H . The phrase "do not reject H " is longish and cumbersome and, therefore, there are several alternatives in common use. One is "accept" \bar{H} , which is the negation of H or its "alternative". My own preferred substitute for "do not reject H " is "no evidence against H is found".

DALLA PENNA DI JERZY NEYMAN (1976)

(iii) It seems to me that the distinction between "tests of significance" and "tests of statistical hypotheses" mentioned in the title of the Session is not really one applicable to some two conceptual entities but to subjective attitudes of the practicing statistician. Thus, the same statistic, say the "Student"-Fisher t-statistic, may be used in two different capacities determined by the attitude of the user. One seems to be the capacity of a "test of significance" and the other that of a "test of a statistical hypothesis". In my own empirical work on a "substantive" problem such distinctions do not appear necessary. My use of the t-statistic is limited to certain familiar conditions in which its frequency properties have been proved to be, in a sense, optimal.

IPOTESI NULLA

- Come è attualmente insegnata mette insieme due concezioni diverse della scienza (Fisher e Neyman-Pearson)
- Sulla liceità o meno dell' "ibrido" F-NP:

<https://stats.stackexchange.com/questions/23142/when-to-use-fisher-and-neyman-pearson-framework>

FISHER VS. NEYMAN

- Fisher ha attaccato Neyman per tutta la vita accusandolo anche di plagio.
- Da giovane Fisher aveva altri ideali, come scopriremo tra poco ...

FISHER SU “RAGIONE E SENTIMENTO”

- “A scientific career is peculiar in some ways. Its raison d’être is the increase of natural knowledge. Occasionally, therefore an increase in natural knowledge occurs. But this is tactless and feelings are hurt... Some undoubtedly take it hard as a blow to the amount of their amour propre and even an invasion of their territory... I do not think that anything can be done about it”

(Fisher alla BBC, 1947)

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company, 2013: p. 51.

L'IPOTESI NULLA (H_0)

- Riprendiamo il concetto da M1

L'IPOTESI NULLA (H_0)

- L'ipotesi nulla asserisce che non vi sono differenze dell'outcome all'interno del predittore/ dei predittori di interesse

H_0 NEL NOSTRO STUDIO - M1

- “A 5 anni non ci attendiamo una differenza nel numero di fratture (outcome = O) nelle donne supplementate col calcio rispetto a quelle supplementate col placebo (predittore = P)”

H_0 NEL NOSTRO STUDIO

- Nel caso di un trial randomizzato controllato c'è un solo predittore: il trattamento
- La randomizzazione dovrebbe infatti proteggere dai fattori confondenti noti e ignoti

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		
	Accetto $H_0 \Rightarrow$ associazione PO assente		

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente	CORRETTO	
	Accetto $H_0 \Rightarrow$ associazione PO assente		

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		
	Accetto $H_0 \Rightarrow$ associazione PO assente		CORRETTO

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		ERRORE TIPO 1
	Accetto $H_0 \Rightarrow$ associazione PO assente		

QUATTRO POSSIBILITÀ

		“Verità” nella popolazione	
		Associazione PO presente	Associazione PO assente
Risultati nel campione	Rifiuto $H_0 \Rightarrow$ associazione PO presente		
	Accetto $H_0 \Rightarrow$ associazione PO assente	ERRORE TIPO 2	

ERRORE DI TIPO 1

- Rifiuto l'ipotesi nulla (= concludo che c'è associazione) quando non c'è associazione tra predittore e outcome
- La probabilità di commettere un errore di tipo 1 è detta “alpha” (α)
- (Metafora del “falso positivo”)

ERRORE DI TIPO 2

- Accetto l'ipotesi nulla (= concludo che non c'è associazione) quando c'è associazione tra predittore e outcome
- La probabilità di commettere un errore di tipo 2 è detta "beta" (β)
- (Metafora del "falso negativo")

ERRORE DI TIPO 3

- Confondere l'errore di tipo 1 con l'errore di tipo 2 (scherzo!)

COSA È P ?

- Il valore p calcolato dai test di significatività statistica rappresenta la probabilità di osservare per caso un effetto uguale o maggiore a quello osservato se l'ipotesi nulla è vera

COSA È LA SIGNIFICATIVITÀ STATISTICA?

- p è detto “statisticamente significativo” se $< \alpha$
- α è 0.05 (5%) nelle scienze biologiche solo per convenzione

COSA È LA POTENZA?

- β è 0.20 (20%) o 0.10 (10%) nelle scienze biologiche solo per convenzione
- Potenza = $1 - \beta$
 - $1 - 0.20 = 0.80 = 80\%$
 - $1 - 0.10 = 0.90 = 90\%$

L'IPOTESI ALTERNATIVA

- È la negazione dell'ipotesi nulla
- È accettata quando l'ipotesi nulla è rifiutata
- È stata la grande idea (o meglio una delle grandi idee) di Neyman, per quanto poi abusata...

COSA È P ? (TORNIAMO INDIETRO)

- Il valore p calcolato dai test di significatività statistica rappresenta la probabilità di osservare per caso un effetto uguale o maggiore a quello osservato se l'ipotesi nulla è vera

DIREZIONE DELL'IPOTESI

- Ipotesi “a una direzione” (“one-sided”)
 - Specifica la direzione dell’associazione PO
 - Esempio: ci attendiamo meno fratture nel gruppo delle donne supplementate con calcio.
- Ipotesi “a due direzioni” (“two-sided”)
 - Non specifica la direzione dell’associazione PO
 - Esempio: ci attendiamo una frequenza diversa di fratture nei due gruppi.

DIREZIONE DELL'IPOTESI

- L'ipotesi uni-direzionale aumenta la potenza dello studio a parità di soggetti ma riduce la possibilità di testare la differenza nei due sensi:
 - E se le donne supplementate col calcio avessero più fratture?
 - Ricordate che non sappiamo ancora se il calcio protegge, altrimenti non avremmo fatto lo studio !

DIREZIONE DELL'IPOTESI

- Nella grande maggioranza dei casi, vi troverete ad utilizzare ipotesi bi-direzionali

STIMA DELLA NUMEROSITÀ CAMPIONARIA

○ Procedura

- Definisci l'ipotesi nulla (una o due direzioni)
- Scegli il test statistico appropriato in base alla variabile di outcome e al predittore/ai predittori (ricordiamo la critica fatta a Stevens - M1!)
- Scegli una dimensione e una variabilità dell'effetto
- (Adesso e solo adesso) fissa α e β

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- Non partire* col calcolo della numerosità campionaria finchè non hai risposto a questa domanda:

“Qual è la differenza minima clinicamente (o biologicamente) rilevante”?

*L'imperativo è appropriato perché questo è un comandamento.

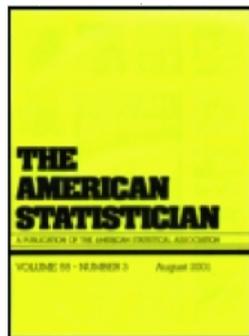
STIMA DELLA NUMEROSITÀ CAMPIONARIA

- Ad esempio: potrei pensare che per dichiarare il calcio efficace nella prevenzione delle fratture, ci debba essere una riduzione della loro incidenza relativa di almeno il 20% nel gruppo trattato con calcio rispetto a quello trattato con placebo

STIMA DELLA NUMEROSITÀ CAMPIONARIA

- Minore è il valore di α , maggiore è il numero di soggetti da arruolare
- Maggiore è il valore della potenza (= minore è il valore di β), maggiore è il numero di soggetti da arruolare
- Minore è la dimensione dell'effetto, maggiore è il numero di soggetti da arruolare
- Minore è la variabilità attorno all'effetto, maggiore è il numero di soggetti da arruolare

ASA POSITION STATEMENT ON P-VALUES



The American Statistician



ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: <http://amstat.tandfonline.com/loi/utas20>

The ASA's Statement on p-Values: Context, Process, and Purpose

Ronald L. Wasserstein & Nicole A. Lazar

To cite this article: Ronald L. Wasserstein & Nicole A. Lazar (2016) The ASA's Statement on p-Values: Context, Process, and Purpose, *The American Statistician*, 70:2, 129-133, DOI: 10.1080/00031305.2016.1154108

To link to this article: <http://dx.doi.org/10.1080/00031305.2016.1154108>

ASA POSITION STATEMENT

- “Q: Why do so many colleges and grad schools teach $p = 0.05$?

A: Because that’s still what the scientific community and journal editors use.

Q: Why do so many people still use $p = 0.05$?

A: Because that’s what they were taught in college or grad school.”

(George Cobb, 2014, on the ASA discussion forum)

ASA POSITION STATEMENT

- “The ASA Board was also stimulated by highly visible discussions over the last few years”
- “The ASA has not previously taken positions on specific matters of statistical practice”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT

- “That turned out to be relatively easy to do [to find points of agreement], but it was just as easy to find points of intense disagreement”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT

- E arrivato il momento di leggerlo!

ASA POSITION STATEMENT – WHAT IS A P-VALUE?

- “Informally, a p-value is the probability under a specified statistical model that a statistical summary of the data (e.g., the sample mean difference between two compared groups) would be equal to or more extreme than its observed value”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 1

- “P-values can indicate how incompatible the data are with a specified statistical model. A p-value provides one approach to summarizing the incompatibility between a particular set of data and a proposed model for the data”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 1

- “The most common context is a model, constructed under a set of assumptions, together with a so-called null hypothesis”
- “Often the null hypothesis postulates the absence of an effect, such as no difference between two groups, or the absence of a relationship between a factor and an outcome”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 1

- “The smaller the p-value, the greater the statistical incompatibility of the data with the null hypothesis, if the underlying assumptions used to calculate the p-value hold”
- “This incompatibility can be interpreted as casting doubt on or providing evidence against the null hypothesis or the underlying assumptions”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 2

- “P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone”
- “Researchers often wish to turn a p-value into a statement about the truth of a null hypothesis, or about the probability that random chance produced the observed data”
- “The p-value is neither. It is a statement about data in relation to a specified hypothetical explanation, and is not a statement about the explanation itself”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 3

- “Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold”
- “Practices that reduce data analysis or scientific inference to mechanical ‘bright-line’ rules (such as ‘ $p < 0.05$ ’) for justifying scientific claims or conclusions can lead to erroneous beliefs and poor decision making”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 3

- “A conclusion does not immediately become “true” on one side of the divide and “false” on the other”
- “Researchers should bring many contextual factors into play to derive scientific inferences, including the design of a study, the quality of the measurements, the external evidence for the phenomenon under study, and the validity of assumptions that underlie the data analysis”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 3

- “Pragmatic considerations often require binary, “yes-no” decisions, but this does not mean that p-values alone can ensure that a decision is correct or incorrect”
- “The widespread use of “statistical significance” (generally interpreted as ‘ $p < 0.05$ ’) as a license for making a claim of a scientific finding (or implied truth) leads to considerable distortion of the scientific process”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 4

- “Proper inference requires full reporting and transparency. P-values and related analyses should not be reported selectively”
- “Conducting multiple analyses of the data and reporting only those with certain p-values (typically those passing a significance threshold) renders the reported p-values essentially uninterpretable”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 4

- “Cherry-picking promising findings, also known by such terms as data dredging, significance chasing, significance questing, selective inference, and ‘p-hacking’, leads to a spurious excess of statistically significant results in the published literature and should be vigorously avoided”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 4

- “One need not formally carry out multiple statistical tests for this problem to arise”
- “Whenever a researcher chooses what to present based on statistical results, valid interpretation of those results is severely compromised if the reader is not informed of the choice and its basis”
- “Researchers should disclose the number of hypotheses explored during the study, all data collection decisions, all statistical analyses conducted, and all p-values computed”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 4

- “Valid scientific conclusions based on p-values and related statistics cannot be drawn without at least knowing how many and which analyses were conducted, and how those analyses (including p-values) were selected for reporting”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 5

- “A p-value, or statistical significance, does not measure the size of an effect or the importance of a result”
- “Statistical significance is not equivalent to scientific, human, or economic significance”
- “Smaller p-values do not necessarily imply the presence of larger or more important effects, and larger p-values do not imply a lack of importance or even lack of effect”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 5

- “Any effect, no matter how tiny, can produce a small p-value if the sample size or measurement precision is high enough, and large effects may produce unimpressive p-values if the sample size is small or measurements are imprecise”
- “Similarly, identical estimated effects will have different p-values if the precision of the estimates differs”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 6

- “By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis”
- “Researchers should recognize that a p-value without context or other evidence provides limited information”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – PRINCIPLE 6

- “For example, a p-value near 0.05 taken by itself offers only weak evidence against the null hypothesis”
- “Likewise, a relatively large p-value does not imply evidence in favor of the null hypothesis; many other hypotheses may be equally or more consistent with the observed data”
- “For these reasons, data analysis should not end with the calculation of a p-value when other approaches are appropriate and feasible”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – OTHER METHODS

- “In view of the prevalent misuses of and misconceptions concerning p-values, some statisticians prefer to supplement or even replace p-values with other approaches”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – OTHER METHODS

- “These include methods that emphasize estimation over testing, such as confidence, credibility, or prediction intervals; Bayesian methods; alternative measures of evidence, such as likelihood ratios or Bayes Factors; and other approaches such as decision-theoretic modeling and false discovery rates”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA – POSITION STATEMENT – OTHER METHODS

- “All these measures and approaches rely on further assumptions, but they may more directly address the size of an effect (and its associated uncertainty) or whether the hypothesis is correct”

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – CONCLUSION

- No single index should be a substitute for scientific reasoning.

Wasserstein RL. Am. Stat. 2016;70:129.

ASA POSITION STATEMENT – ALLEGATO

Statistical Tests, *P*-values, Confidence Intervals, and Power: A Guide to Misinterpretations

Sander GREENLAND, Stephen J. SENN, Kenneth J. ROTHMAN, John B. CARLIN, Charles POOLE,
Steven N. GOODMAN, and Douglas G. ALTMAN

ASA POSITION STATEMENT – UN ANNO DOPO

The ASA's p -value statement, one year on

Its aim was to stop the misuse of statistical significance testing. But **Robert Matthews** argues that little has changed in the 12 months since the ASA's intervention

38 | SIGNIFICANCE | April 2017

ASA POSITION STATEMENT – UN ANNO DOPO – ROBERT MATTHEWS

- “Yet a year on, it is not clear that the ASA’s statement has had any substantive effect at all”
- “The commentaries accompanying the statement show that even at the time some participants feared there would be no lasting impact”
- “Goodman’s frustration with this resonates with mine: exactly how [are] scientists supposed to do that?”

Matthews R et al. Significance. 2017;14:38.

ASA POSITION STATEMENT – UN ANNO DOPO – – RON WASSERSTEIN*

- “The statement did not go as far as it should go, but it went as far as it could go”
- “To address the statement’s shortcomings the ASA is convening a Symposium on Statistical Inference this October”

*Executive Director, American Statistical Association

Matthews R et al. Significance. 2017;14:38.

ASA POSITION STATEMENT – UN ANNO DOPO – – DAVID SPIGELHALTER*

- “If by this he [Matthews] means [to ban] all use of p-values than I must disagree”
- “Fortunately the Neyman-Pearson’s decision-theoretic idea of “accepting” the null has just been consigned to the overflowing dustbin of inappropriate scientific ideas”

Matthews R et al. Significance. 2017;14:38.

*President, Royal Statistical Association

ASA POSITION STATEMENT – UN ANNO DOPO – – DAVID SPIGELHALTER

- “While it would be wonderful if every analysis would be informed by someone skilled in statistical methodology, whether a nominal statistician or not, the rise of data science means that even more practitioners will be without a full professional training and continue to do their *t*-ing and *p*-ing”

Matthews R et al. Significance. 2017;14:38.

UNA SCORRIBANDA TRA ALCUNI TEST STATISTICI

Tipo di variabile	2 gruppi diversi	3+ gruppi diversi	Prima e dopo nello stesso gruppo	Trattamenti multipli nello stesso gruppo	Associazione tra variabili
Continua con distribuzione normale	t-test per dati non appaiati	Analisi della varianza (ANOVA)	t-test per dati non appaiati	ANOVA per misure ripetute	Correlazione e regressione
Nominale	Test X ²	Test X ²	Test di McNemar	Test Q di Cochran	Coefficiente di contingenza
Ordinale	Test di Mann - Whitney	Test di Kruskal - Wallis	Test di Wilcoxon	Test di Friedman	Coefficiente di correlazione di Spearman

RICORDIAMO LA CRITICA A STEVENS (M1)!

- “Statistics programs based on Stevens’s typology suggest that doing statistics is simply a matter of declaring the scale type of data and picking a model”
- “Worse, they assert that the scale type is evident from the data independent of the questions asked of the data”
- “They thus restrict the questions that may be asked of the data”
- “Such restrictions lead to bad data analysis and bad science”

Velleman, P. Am. Stat. 1993;47:65.

www.giorgiobedogni.it

TEST T DI STUDENT

Tipo di variabile	2 gruppi diversi	3+ gruppi diversi	Prima e dopo nello stesso gruppo	Trattamenti multipli nello stesso gruppo	Associazione tra variabili
Continua con distribuzione normale	t-test per dati non appaiati	Analisi della varianza (ANOVA)	t-test per dati non appaiati	ANOVA per misure ripetute	Correlazione e regressione
Nominale	Test X ²	Test X ²	Test di McNemar	Test Q di Cochran	Coefficiente di contingenza
Ordinale	Test di Mann - Whitney	Test di Kruskal - Wallis	Test di Wilcoxon	Test di Friedman	Coefficiente di correlazione di Spearman

WILLIAM SEALY GOSSET DETTO “STUDENT”
(1876-1937)



www.giorgiobedogni.it

UN CLASSICO

Statistical Science
1987, Vol. 2, No. 1, 45-52

Guinness, Gosset, Fisher, and Small Samples

Joan Fisher Box

191

WILLIAM SEALY GOSSET

- “A most appealing character-quiet, unaffectedly friendly, helpful, patient, loyal, everybody liked and trusted him. In the very quarrelsome world of statistics, he managed to be on friendly terms with everyone”

Fisher Box J. Stat. Sci. 1987;2:45.

WILLIAM SEALY GOSSET

- “Gosset asked the first question and found the answer despite his lack of mathematics. Fisher made it elegant and went on from here”

Fisher Box J. Stat. Sci. 1987;2:45.

WILLIAM SEALY GOSSET

- “They [Guinness] had decided to make brewing scientific”
- “First they began to hire a series of the brightest men they could find”
- “Now can you imagine half a dozen of energetic and bright researchers given a free hand to explore the whole subjects of brewing from a condition of almost total ignorance?”

Fisher Box J. Stat. Sci. 1987;2:45.

THIS IS THE PLACE!



WILLIAM SEALY GOSSET

- “And to them it seemed natural to talk their numerical problems to Gossett. He had studied some mathematics at Oxford and was less scared than they were”
- For him [Gossett], the main question was exactly how much wider should be the error limits to be make allowance for the error introduced by using the estimates m and s instead of the parameters μ and σ ”

Fisher Box J. Stat. Sci. 1987;2:45.

WILLIAM SEALY GOSSET

- “I couldn’t understand his stuff and wrote and said that I was going to study it when I had time. I actually took it up to the lakes with me and lost it!”

(Gossett che scrive a Karl Pearson a proposito della lettera con la prova rigorosa del t -test inviataagli da Fisher)

Box Fisher J. Stat. Sci. 1987;2:45.

IL PROBLEMA DI GOSSET

VOLUME VI

MARCH, 1908

No. 1

BIOMETRIKA.

THE PROBABLE ERROR OF A MEAN.

By STUDENT.

IL PROBLEMA DI GOSSET

- “If the number of experiments be very large, we may have precise information as to the value of the mean, but if our sample be small, we have two sources of uncertainty”

Student. Biometrika. 1908;6:1.

IL PROBLEMA DI GOSSET

- “(1) owing to the ‘error of random sampling’ the mean of our series of experiments deviates more or less widely from the mean of the population, and”

“(2) the sample is not sufficiently large to determine what is the law of distribution of individuals”

Student. Biometrika. 1908;6:1.

TEST T DI STUDENT

Tipo di variabile	2 gruppi diversi	3+ gruppi diversi	Prima e dopo nello stesso gruppo	Trattamenti multipli nello stesso gruppo	Associazione tra variabili
Continua con distribuzione normale	t-test per dati non appaiati	Analisi della varianza (ANOVA)	t-test per dati appaiati	ANOVA per misure ripetute	Correlazione e regressione
Nominale	Test X ²	Test X ²	Test di McNemar	Test Q di Cochran	Coefficiente di contingenza
Ordinale	Test di Mann - Whitney	Test di Kruskal - Wallis	Test di Wilcoxon	Test di Friedman	Coefficiente di correlazione di Spearman

TEST T DI STUDENT PER 2 GRUPPI NON APPAIATI

- Ipotesi nulla: la differenza tra le medie di due variabili continue con distribuzione Gaussiana è uguale a 0
- Importante: è la differenza tra le medie a necessitare della distribuzione Gaussiana e non le due medie
- Più importante: le varianze dei due gruppi devono essere omogenee

TEST T DI STUDENT PER 2 GRUPPI NON APPAIATI

$$t = \frac{\text{differenza tra le medie}}{\text{errore standard della differenza tra le medie}}$$

$$\text{gradi di libertà} = n_1 + n_2 - 2$$

(dettagli matematici omissi)

TEST T DI STUDENT

- Cosa me ne faccio del valore di t ottenuto dal test?
- Lo devo associare a una distribuzione di probabilità e ottenere il valore di p ma soprattutto la dimensione dell'effetto, cosa che il t -test fa direttamente.

TEST T DI STUDENT

- Siete d'accordo che se ripetiamo lo studio su campioni diversi otterremo valori diversi di t ?
- Chiediamoci allora: qual è la distribuzione di tutti i valori di t che possono essere ottenuti se l'ipotesi nulla è vera?
- Questa distribuzione dipende dai gradi di libertà.

TEST T DI STUDENT

- Qual è la probabilità p che si abbia un valore t come quello osservato se l'ipotesi nulla è vera?
- Questa probabilità è inferiore al valore di α che ho fissato?

TEST T DI STUDENT PER 2 GRUPPI NON APPAIATI

- Supponiamo di volere confrontare la glicemia a digiuno di 23 bambini HIV+ con quella di 19 bambini HIV-

IL T-TEST AL LAVORO

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	19	81.21053	1.610599	7.020438	77.82678	84.59427
1	23	79.65217	1.603613	7.69066	76.32648	82.97786
combined	42	80.35714	1.133802	7.347877	78.06738	82.6469
diff		1.558352	2.293057		-3.076088	6.192793

diff = mean(0) - mean(1) t = 0.6796
 Ho: diff = 0 degrees of freedom = 40

Ha: diff < 0
 Pr(T < t) = 0.7497

Ha: diff != 0
 Pr(|T| > |t|) = 0.5007

Ha: diff > 0
 Pr(T > t) = 0.2503

P-VALUE E INTERVALLI DI CONFIDENZA

- C'è una connessione stretta tra gli intervalli di confidenza e il valore di p
- Se gli intervalli di confidenza al 95% contengono lo 0, la differenza non può essere statisticamente significativa a un livello di $\alpha < 0.05$. E se lo contengono al 99%?
- Però, l'intervallo di confidenza offre una misura dell'effetto e della sua variabilità che può essere “tradotta” in un effetto biologico o clinico

RONALD AYLMER FISHER (1890-1962)



www.giorgiobedogni.it

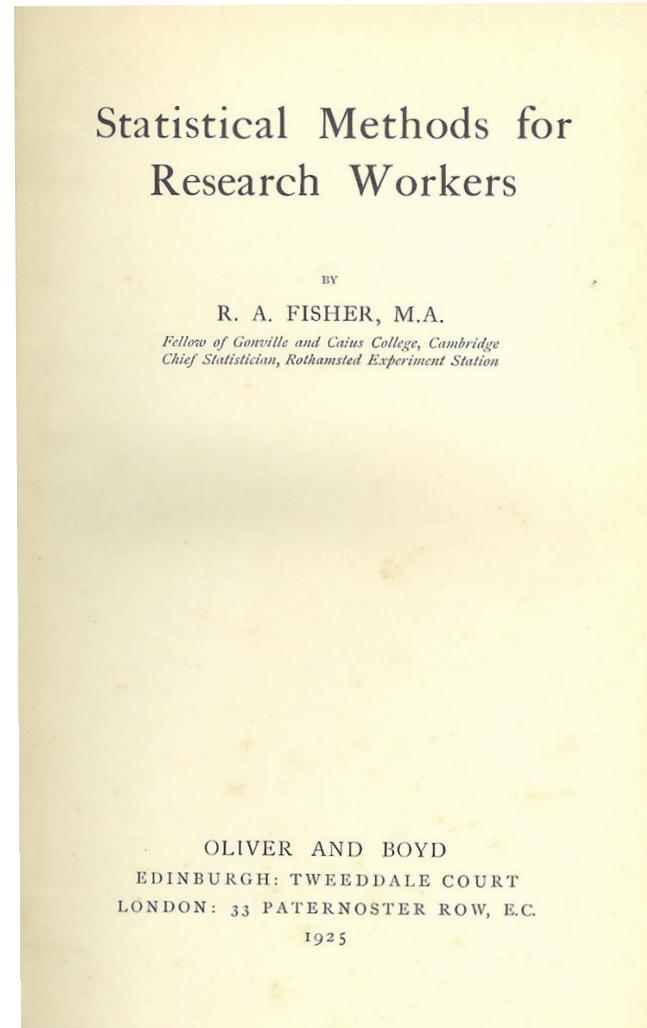
RONALD AYLMER FISHER – IL GIOVANE

- “He had a consuming desire to make himself useful, to serve his country and humankind in a significant and practical way”
- “Fisher brought its mathematical genius to prove that the inheritance of continuous variables was entirely consistent with Mendelian principles”
- “The way it proved it was important as he formulated the problem in a different way and introduced the concept of analysis of variance”

Box Fisher J. Stat. Sci. 1987;2:45.

www.giorgiobedogni.it

STATISTICAL METHODS FOR SOCIAL WORKERS (1925)



www.giorgiobedogni.it

PREFACE TO FIRST EDITION

FOR several years the author has been working in somewhat intimate co-operation with a number of biological research departments; the present book is in every sense the product of this circumstance. Daily contact with the statistical problems which present themselves to the laboratory worker has stimulated the purely mathematical researches upon which are based the methods here presented. Little experience is sufficient to show that the traditional machinery of statistical processes is wholly unsuited to the needs of practical research. Not only does it take a cannon to shoot a sparrow, but it misses the sparrow! The elaborate mechanism built on the theory of infinitely large samples is not accurate enough for simple laboratory data. Only by systematically tackling small sample problems on their merits does it seem possible to apply accurate tests to practical data. Such at least has been the aim of this book.

I owe more than I can say to Mr W. S. Gosset, Mr E. Somerfield, and Miss W. A. Mackenzie, who have read the proofs and made many valuable suggestions. Many small but none the less troublesome errors have been removed; I shall be grateful to readers who will notify me of any further errors and ambiguities they may detect.

ROTHAMSTED EXPERIMENTAL STATION,
February 1925.

STATISTICAL METHODS FOR SOCIAL WORKERS (1925)

- “Despite, or because of, its lack of theoretical mathematics, the book was rapidly taken by the scientific community. It met a serious need”
- “During the Second World War, the Swedish mathematician Harald Cramér spent days and weeks filling the missing steps of proofs and deriving proofs where none were indicated”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company, 2013, p. 39.

STATISTICAL METHODS FOR SOCIAL WORKERS (1925)

- “In the 1970s, L.J. Savage of Yale University went back to Fisher’s original papers and discovered how much Cramér had left out”
- “He was amazed to see that Fisher had anticipated much later work by others and had solved problems that were thought to be unsolved still in the 1970s”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company, 2013, p. 39.

UNA SCORRIBANDA TRA ALCUNI TEST STATISTICI

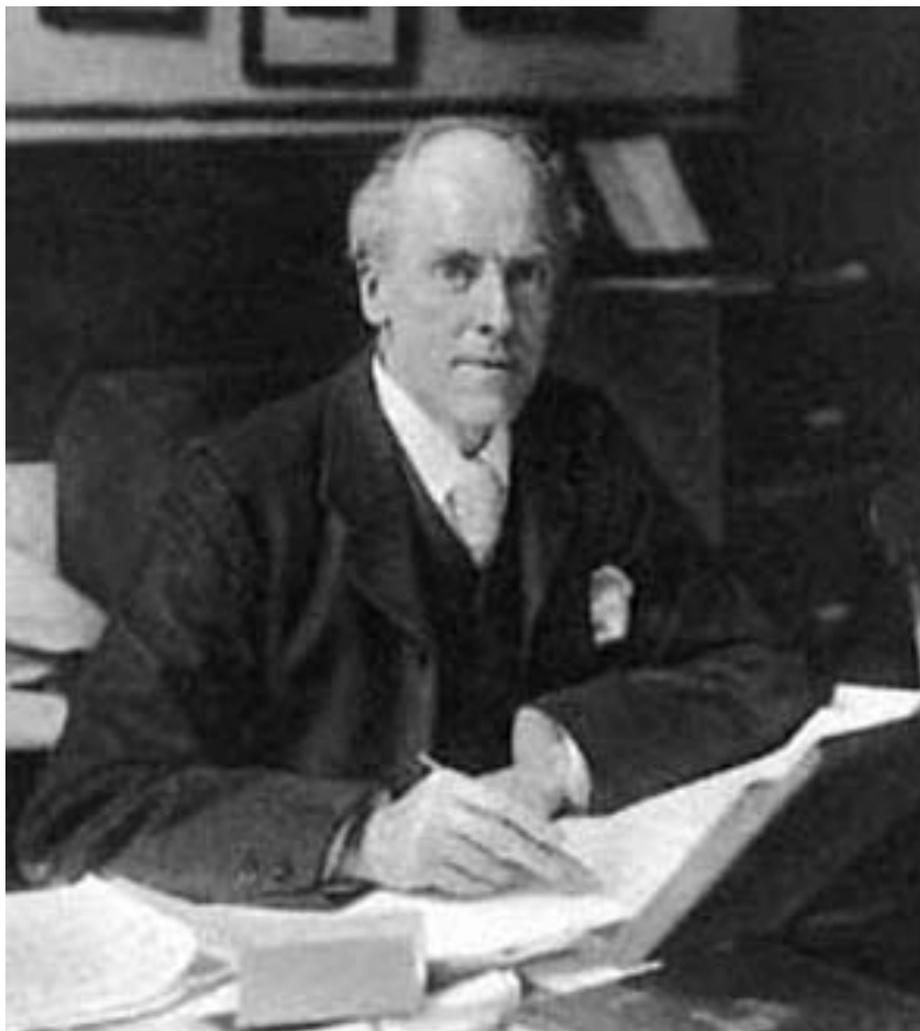
Tipo di variabile	2 gruppi diversi	3+ gruppi diversi	Prima e dopo nello stesso gruppo	Trattamenti multipli nello stesso gruppo	Associazione tra variabili
Continua con distribuzione normale	t-test per dati non appaiati	Analisi della varianza (ANOVA)	t-test per dati appaiati	ANOVA per misure ripetute	Correlazione e regressione
Nominale	Test X²	Test X²	Test di McNemar	Test Q di Cochran	Coefficiente di contingenza
Ordinale	Test di Mann - Whitney	Test di Kruskal - Wallis	Test di Wilcoxon	Test di Friedman	Coefficiente di correlazione di Spearman

UNA SCORRIBANDA TRA ALCUNI TEST STATISTICI

Tipo di variabile	2 gruppi diversi	3+ gruppi diversi	Prima e dopo nello stesso gruppo	Trattamenti multipli nello stesso gruppo	Associazione tra variabili
Continua con distribuzione normale	t-test per dati non appaiati	Analisi della varianza (ANOVA)	t-test per dati appaiati	ANOVA per misure ripetute	Correlazione e regressione
Nominale	Test X²	Test X ²	Test di McNemar	Test Q di Cochran	Coefficiente di contingenza
Ordinale	Test di Mann - Whitney	Test di Kruskal - Wallis	Test di Wilcoxon	Test di Friedman	Coefficiente di correlazione di Spearman

KARL PEARSON

(1857-1936)



www.giorgiobedogni.it

L'IDEA RIVOLUZIONARIA DI KARL PEARSON

- “Armed with Pearson’s revolutionary idea, we do not look upon experimental results as carefully measured numbers in their own right. Instead they are an example of a scatter of numbers, a distribution of numbers, to use the more accepted term”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company, 2013, p.15.

OF CLOCKS AND CLOUDS



L'IDEA RIVOLUZIONARIA DI KARL PEARSON

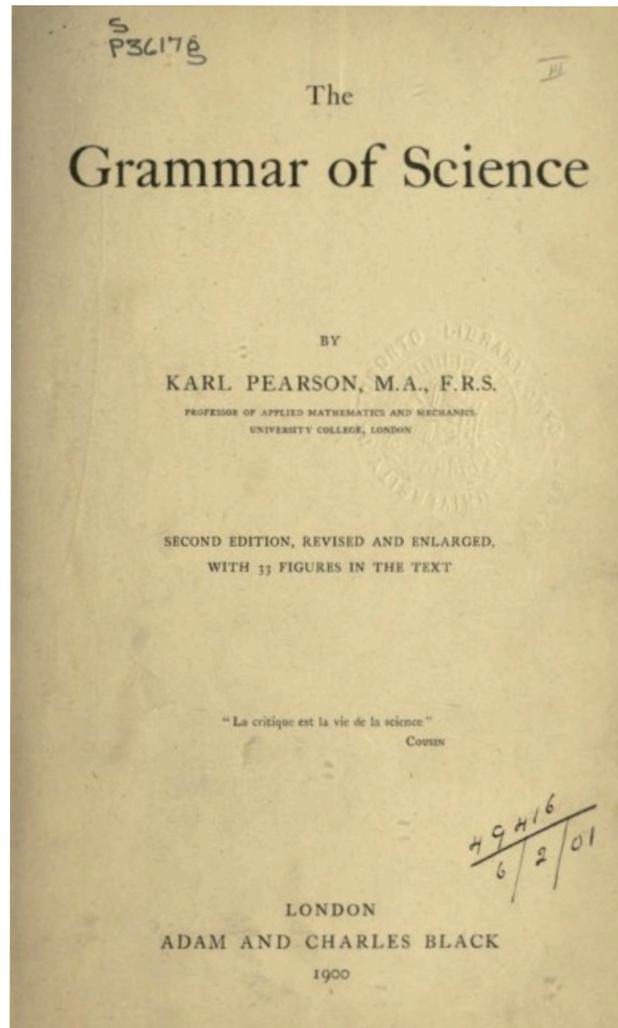
- “Regardless of one’s underlying philosophy, the fact remains that Pearson’s idea about distribution functions and parameters came to dominate the twentieth-century science and stands triumphant on the threshold of twenty-first century ”

Salsburg, D. The lady tasting tea. New York: Henry Holt and Company, 2013, p. 24.

L'IDEA RIVOLUZIONARIA DI KARL PEARSON

- Karl Pearson non usava le “statistiche” per stimare i “parametri” come facciamo oggi. Credeva che con un numero sufficiente di dati si potessero stimare direttamente i parametri.
- (Non pensiamo per questo di essere più intelligenti di Pearson. Solo Fisher poteva pensarlo e lo era davvero, ma i due non sono mai andati d'accordo... stupiti?)

THE GRAMMAR OF SCIENCE (1892)



www.giorgiobedogni.it

THE GRAMMAR OF SCIENCE (RISTAMPA 2007, P. 36)

- “There is an insatiable desire in the human breast to resume in some short formula, some brief statement, the facts of human experience”
- “It leads the savage to account for all natural phenomena to deyfying the wind and the stream and the tree”
- “It leads civilised man, on the other hand, to express his emotional experience in works of art and his physical and mental experience in the formulae or so-called law of science”

THE GRAMMAR OF SCIENCE (RISTAMPA 2007, P. 36)

- “It may seem at first strange to the reader that the laws of science should thus be associated with the creative imagination in man rather than the physical world outside him”
- “But as we shall see in the following chapters, the laws of science are the products of the human mind rather than factors of the external world”

THE GRAMMAR OF SCIENCE (RISTAMPA 2007, P. 36)

- “Science endeavors to provide a mental résumé of the universe and its last great claim to our support is the capacity it has for satisfying our craving for a brief description of the history of the world”

IL TEST X^2 DI PEARSON AL LAVORO

- Disegniamo un trial randomizzato controllato (RCT)
- Lo sappiamo fare da M1 e M2 (più o meno...)

DISEGNIAMO UN RCT

- Domanda dello studio
 - È una nuova formula (NF) superiore a una formula standard (SF) per la prevenzione della diarrea in neonati pretermine?
- Ipotesi nulla
 - L'incidenza di diarrea nel gruppo di neonati NF è la stessa del gruppo SF

DISEGNIAMO UN RCT

- Questa ipotesi può essere testata utilizzando il test X^2
 - È la scelta di default di molti studi clinici, spesso per pigrizia o abitudine...
- Comunque, qui ci serve per vedere il X^2 all'opera e quindi non stiamo (troppo) a discutere

DISEGNIAMO UN RCT

- Definiamo l'outcome
 - Diarrea: ≥ 1 episodio di diarrea in 3 mesi
- Qual è l'incidenza di diarrea nella popolazione che ci interessa studiare?
 - 25%

DISEGNIAMO UN RCT

- Qual è la differenza minima clinicamente rilevante?
 - Per dichiarare efficace la formula ci aspettiamo che essa riduca l'incidenza assoluta di diarrea di almeno il 15% ($10-25=-15\%$)
- Come sia chiama la differenza (10-25)? Come si chiama il suo inverso? (M2).

DISEGNIAMO UN RCT

- Adesso fissiamo $\alpha = 0.05$ e potenza = 90%
 - Non sono una scelta obbligata!
 - Dobbiamo giudicare l'importanza relativa di un errore di tipo I e di tipo II.
 - (Stiamo ovviamente seguendo il paradigma N-P)

DISEGNIAMO UN RCT

Estimated sample sizes for a two-sample proportions test

Pearson's chi-squared test

Ho: $p_2 = p_1$ versus Ha: $p_2 \neq p_1$

Study parameters:

alpha =	0.0500	
power =	0.9000	
delta =	-0.1500	(difference)
p1 =	0.2500	
p2 =	0.1000	

Estimated sample sizes:

N =	266
N per group =	133

RISULTATI DELLO STUDIO (CHE VELOCITÀ!)

	NF	SF	Total
Diarrhea+	4	20	24
Diarrhea-	91	73	164
Total	95	93	188

RISULTATI

- Cosa è andato storto?
- Cosa facciamo adesso?
- Come ci assicuriamo che non riaccada?
 - Molti ci si abituanano ma così non imparano. Potete fare come loro solo se non vi interessa imparare...

COSA FACCIAMO ADESSO?

- Di certo, non possiamo più testare l'ipotesi originale e questo va detto chiaramente
- Ovviamente, siamo fuori dal paradigma N-P. Perché?
- Calcolare il sample size post-hoc è una follia (se abbiamo compreso la logica del paradigma N-P). È pure un falso...

<https://homepage.divms.uiowa.edu/~rlenth/Power/> leggete la FAQ

COSA FACCIAMO ADESSO?

- Possiamo comunque stimare la differenza e vedere se è clinicamente rilevante
- Non ci interessa la significatività statistica
 - Non ci interessa comunque!

TEST X^2 DI PEARSON

$$X^2 = \sum_{cells} \frac{(O - E)^2}{E}$$

O = valore osservato (observed)

E = valore atteso (expected)

df = (rows-1)*(columns-1)

TEST X^2 DI PEARSON

$$X^2 = \sum_{cells} \frac{(O - E)^2}{E}$$

O = valore osservato (observed)

E = valore atteso (expected)

df = (rows-1)*(columns-1)

COME SI CALCOLA IL VALORE ATTESO?

	NF	SF	Total
Diarrhea+	4	20	24
Diarrhea-	91	73	164
Total	95	93	188

$20/93 = 0.21$ i casi di diarrea osservati nel gruppo SF

$95 * 0.21 = 20$ i casi che mi aspetterei nel gruppo NF se il trattamento non fosse efficace

P-VALUE DEL X^2 CONTRO PRECISIONE DELL'EFFETTO-COME È ANDATA

	NF	SF	Total
Diarrhea+	4	20	24
Diarrhea-	91	73	164
Total	95	93	188
Risk	.0421053	.2150538	.1276596
	Point estimate	[95% Conf. Interval]	
Risk difference	-.1729485	-.265704	-.080193
Risk ratio	.1957895	.0695674	.5510271
$\chi^2(1) = 12.62 \quad Pr > \chi^2 = 0.0004$			

COME SPERAVAMO CHE ANDASSE (DAVVERO?)

	NF	SF	Total
Diarrhea+	13	33	46
Diarrhea-	120	100	220
Total	133	133	266
Risk	.0977444	.2481203	.1729323
	Point estimate		[95% Conf. Interval]
Risk difference	-.1503759		-.2394577 - .0612942
Risk ratio	.3939394		.2172615 .7142924
chi2(1) = 10.51 Pr>chi2 = 0.0012			

CI SONO ALTERNATIVE?

- Certamente. Possiamo calcolare il sample size a partire dalla precisione della differenza di interesse. (Siamo fuori dal paradigma N-P).
- Solo che non è questo che ci chiedono gli Enti Regolatori, gli Editor delle Riviste, ecc.

Harrell F Jr & Slaughter JC. Biostatistics for Biomedical Research. Version 27 Jan 2017.

<http://biostat.mc.vanderbilt.edu/tmp/bbr.pdf>

CI SONO ALTERNATIVE?

- Comunque possiamo sempre calcolare l'intervallo di confidenza della differenza che corrisponde all'ipotesi nulla che ci chiedono di specificare
- Così li accontentiamo e ci assicuriamo che lo studio possa rilevare una differenza biologicamente/clinicamente importante col margine di errore che ci interessa

LA “REGOLA DEL TRE” DI WEINBERG

- The rule of three:

“If you can’t think of three things that can go wrong with your plan, then there is something wrong with your thinking”

Weinberg, G. The secrets of consulting. Weinberg, 2014.

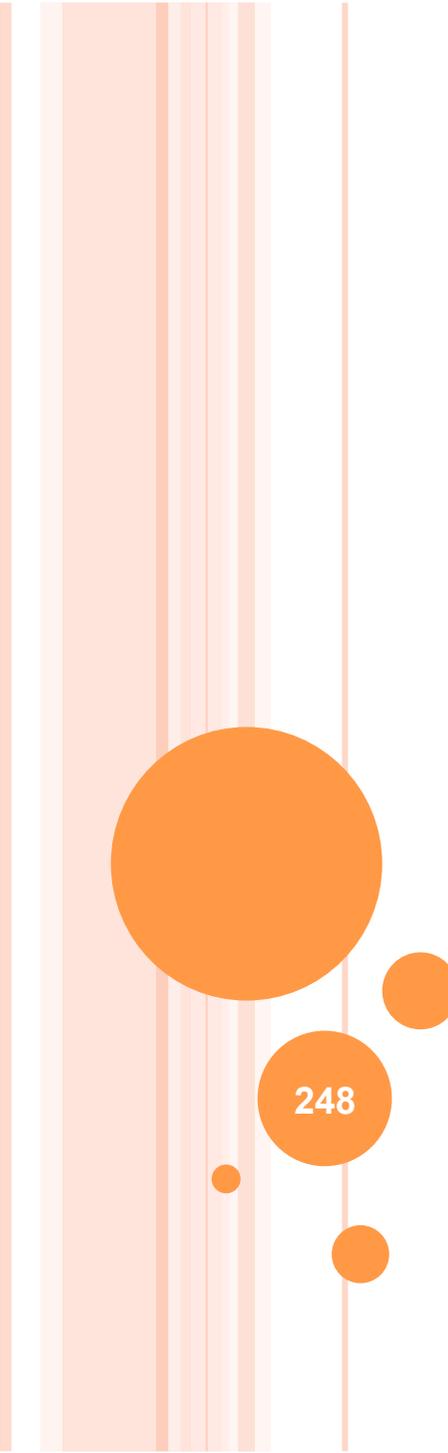
<https://leanpub.com/u/jerryweinberg> (Potete leggere tranquillamente tutto quello che ha scritto Weinberg. È una miniera di idee!)

M3-SESSIONE PRATICA 23-05-2017-COSA FACCIAMO?



M3-SESSIONE PRATICA 23-05-2017-COSA FACCIAMO?

- Leggiamo insieme il commentary al position paper ASA?
- Qualcuno di voi espone un progetto a cui sta lavorando e tutti ne discutiamo (Neyman-style ovviamente)?
- Qualcuno di voi espone un articolo e tutti ne discutiamo (Neyman-style ovviamente)?



**GRAZIE E ARRIVEDERCI AL
23-05-2017!**

248