

I Mestieri dello Statistico

La Statistica è bella perché è varia

Andrea Ghiglietti

Genova, 4 Aprile 2022

La mia storia (lavorativa)



POLITECNICO
MILANO 1863

- (2005-2013) Laurea e Dottorato in Ingegneria Matematica
Politecnico di Milano



Università
della
Svizzera
italiana

- (2014-2021) Assegnista e Ricercatore Universitario (tipo a)
*Università degli Studi di Milano,
Università Cattolica di Milano,
Università della Svizzera Italiana*



LivaNova
Health innovation that matters

- (2018-2019) Quality Statistician
*Valvole cardiache e dispositivi per la chirurgia
cardio-polmonare*



swissblock

- (2019-2021) Consulente Statistico
*Strategie di Trading per crypto-valute
(es. Bitcoin)*



- (2022-oggi) Ricercatore Universitario (tipo b)
Università degli Studi di Torino

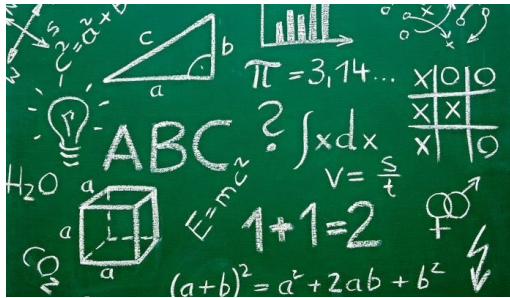
Quale Corso di Laurea scegliere?

Liceo scientifico

G.B. Vico di Corsico

settembre 2005

Matematica



Ingegneria matematica

Politecnico di Milano



Ingegneria



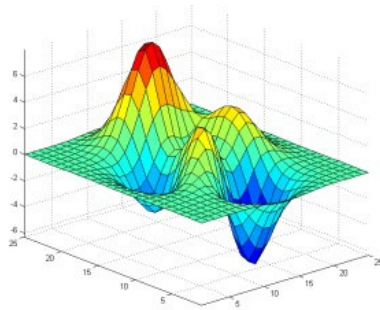
Cosa fare dopo la Laurea Triennale?

Laurea triennale

Politecnico di Milano

settembre 2008

Calcolo Scientifico

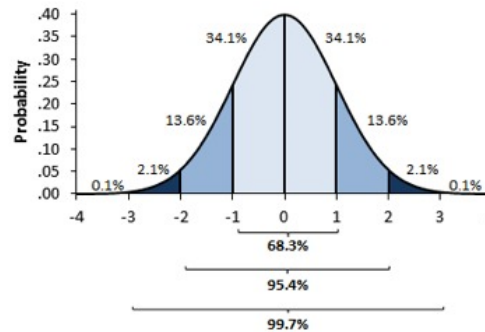


- Struttura matematica
- Evoluzione di processi nel tempo
- Caratterizzazione formale del «caso»

Finanza



Statistica



- Progetti analisi dati
- programmazione con software
- Casi studio reali

Teoria - Applicazioni

Cosa fare dopo la Laurea Specialistica?



Cosa fare dopo il dottorato?

**Dottorato in modelli e metodi
matematici per l'ingegneria**

Politecnico di Milano

febbraio 2014



Lavoro in azienda



**Assegnista / Ricercatore (a)
Universitario**



Lo Statistico in Università



Lo Statistico in Università



Didattica

- Lezioni in classe
- Preparazione Esami
- Relatore Tesi di laurea

Ricerca

- Scrittura di articoli, libri, ...
- Partecipazione a conferenze

...ma cosa vuol dire fare ricerca???

Da cosa parto?

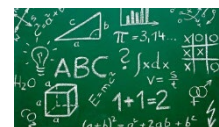
Argomento che **interessa** e **mi appassiona**

Cosa è già stato fatto?

- Libri
- Articoli su riviste scientifiche
- Conferenze

Quale può essere il mio contributo?

Cosa non è ancora stato fatto (che **interessa** e **mi appassiona**)?



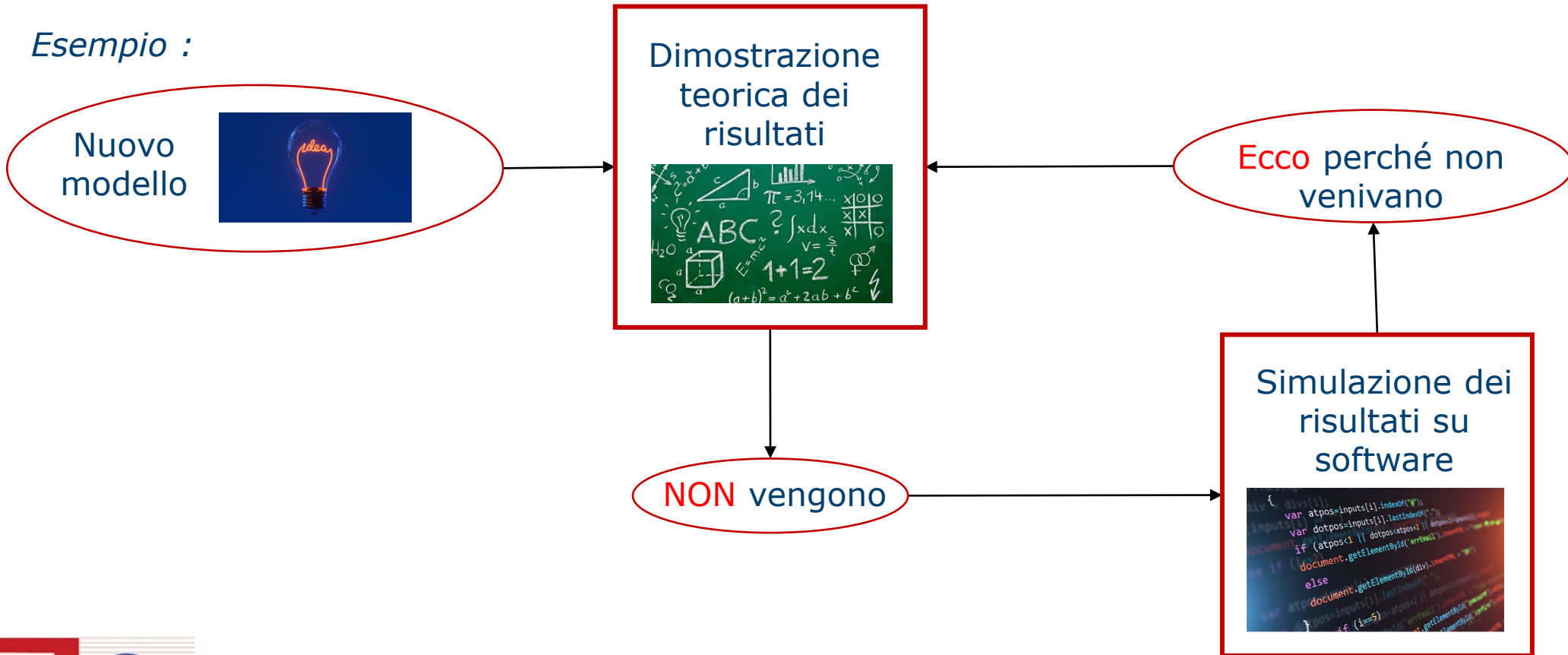
- Posso inventare un **nuovo** modello?
- Posso implementare un **nuovo** algoritmo su software?
- Dimostrare un **nuovo** risultato teorico?

Lo Statistico in Università



Come procedere?ognuno trova la sua strada...

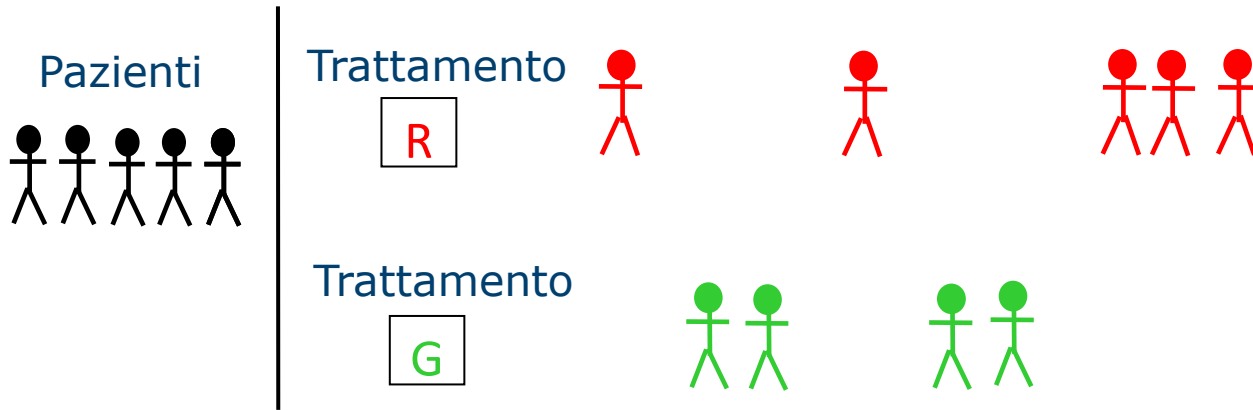
Esempio :



Esempio 1: modelli d'urna



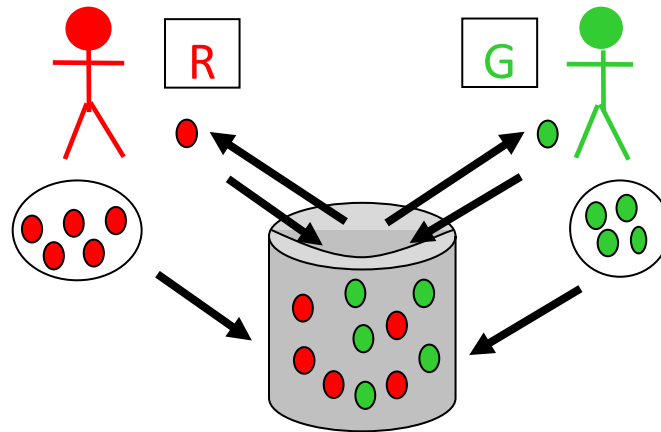
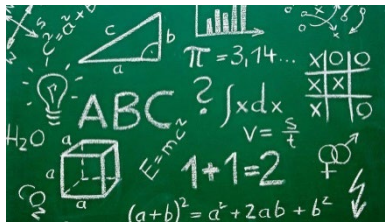
Framework: Esperimento clinico per confrontare 2 farmaci: Red vs Green.



Come assegniamo i pazienti ai trattamenti?

..in modo random e tale che «dopo un po'»
vengano assegnati solo al trattamento migliore!

Formule matematiche
complicate ma...



Obiettivi della ricerca:

- Costruire un nuovo modello d'urna
- Analizzare simulazioni su software
- Dimostrare alcuni risultati teorici

Esempio 2: FDA – Functional Data Analysis



Framework: i dati non sono più numeri ma funzioni!

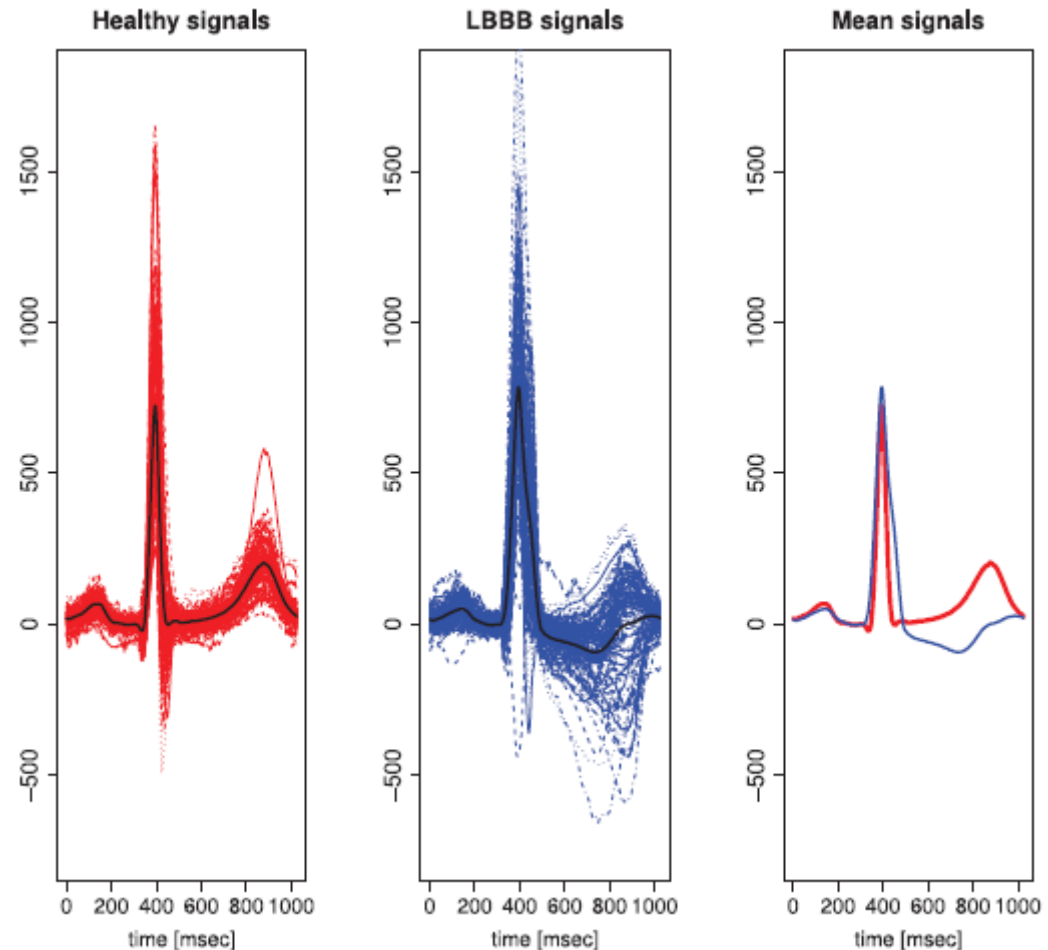
Esempi: Temperatura giornaliera, andamento del prezzo in borsa, Elettrocardiogramma (ECG)

Problemi:

- cos'è la media di un gruppo di curve?
- cos'è la differenza tra due curve?

Obiettivi della ricerca:

- **TEST:** sani e malati hanno gruppi di curve «significativamente» diversi?
- **PREVISIONE:** posso distinguere un sano da un malato in base al suo ECG?



Lo Statistico in LivaNova

LivaNova

Health innovation that matters

Azienda
Farmaceutica /
Biomedicale



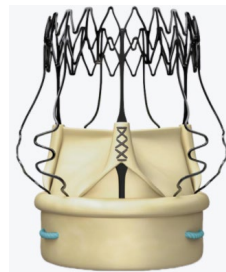
LivaNova

Health innovation that matters

Strumenti per la
chirurgia
cardiopulmonare



Valvole cardiache



Neuromodulazione

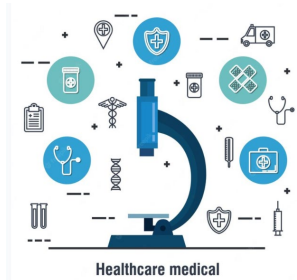


LivaNova – Dipartimenti

LivaNova

Health innovation that matters

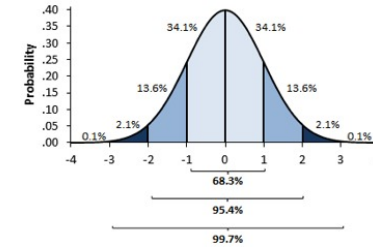
Clinicians



R&D – Research & Development



GB – Global Biometrics



LivaNova

Health innovation that matters

CuQ – Customer Quality



QE – Quality Engineering



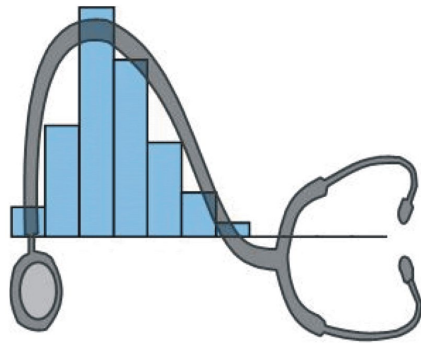
Sales



Marketing



BS - Biostatistician

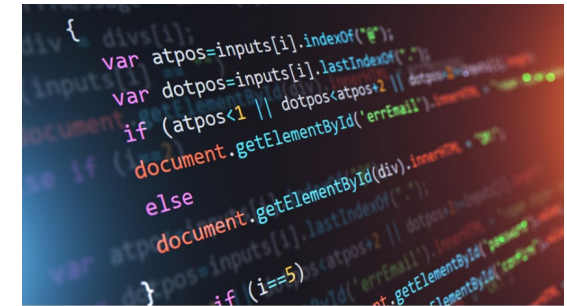


GB - Global Biometrics

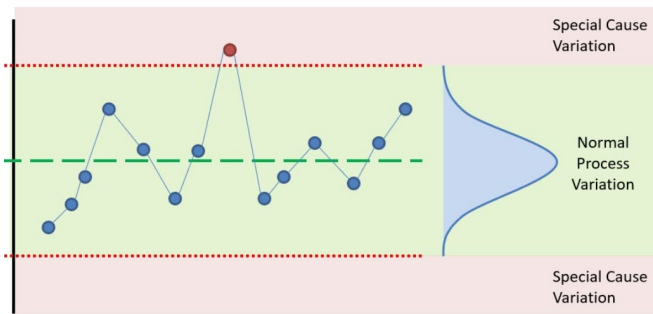
DS - Data Scientist



SP - Statistical Programmer



QS - Quality Statistician



DM - Data Manager

request_number	request_name	request_status	request_date	request_time	request_type	request_category	request_subcategory	request_priority	request_urgency	request_severity	request_status	request_date	request_time	request_type	request_category	request_subcategory	request_priority	request_urgency	request_severity	
00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000	00000000000000000000000000000000

Esempio 1: Analisi Statistica nei Clinical Trials

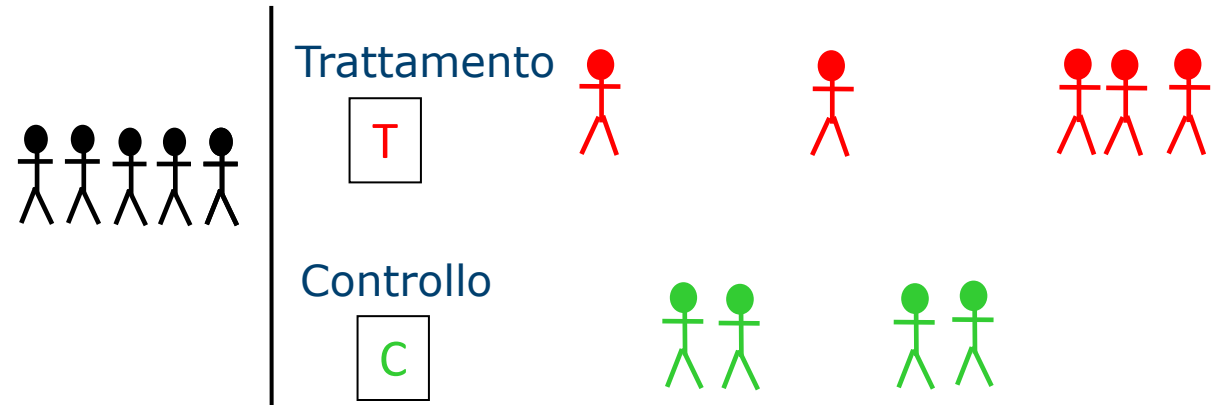
Cosa: Esperimento in cui alcuni pazienti sono curati con un prodotto biomedicale sotto studio (anche già in commercio), detto «trattamento»

Perché: testare l'efficacia del trattamento

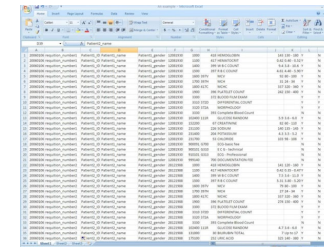
Dove: decine di ospedali/cliniche

Quanti pazienti: da 600 a 10k

Quanto tempo: a termine 5/10 anni



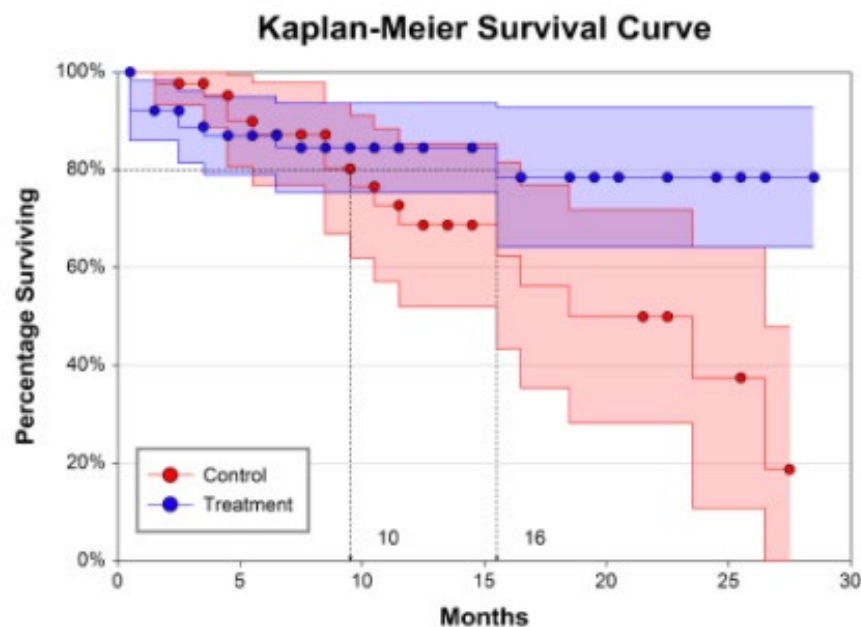
Le risposte di tali pazienti costituiscono i dati usati per le analisi statistiche



Esempio 1: Survival Analysis

Analisi di Sopravvivenza: studio del tempo di attesa di un certo **evento** (es. sintomi, morte)

Curva di Sopravvivenza: **% di pazienti** in cui l'evento deve ancora verificarsi al variare del tempo



Gradini: punti dove accade almeno un evento

Decremento % nei gradini:

$$\frac{\text{pazienti in cui l'evento si è appena verificato}}{\text{pazienti in cui l'evento deve ancora verificarsi}}$$

Test statistici sono necessari per determinare se le due curve del Trattamento T e del Controllo C sono **significativamente** diverse

SAP – Statistical Analysis Plan

1. Table of Contents

Section	Page
Statistical Analysis Plan: ISB-MC-JGDP	1
1. Table of Contents	3
2. Revision History	7
3. Study Objectives	8
3.1. Primary Objective	8
3.2. Secondary Objectives	8
3.3. Tertiary Objectives	9
4. Study Design	10
4.1. Summary of Study Design	10
4.1.1. Phase 1b	10
4.1.2. Phase 2	11
4.2. Determination of Sample Size	13
4.3. Method of Assignment to Treatment	14
4.3.1. Phase 1b	14
4.3.2. Phase 2	14
4.3.2.1. Stratification Factors	14
5. A Priori Statistical Methods	16
5.1. General Considerations	16
5.2. Handling of Dropouts or Missing Data	16
5.3. Populations for Analyses	17
5.4. Patient Disposition	19
5.5. Demographic and Baseline Characteristics	19
5.6. Treatment Compliance	19
5.7. Concomitant Therapy	20
5.8. Efficacy Analyses	20
5.8.1. Primary Outcome and Methodology	20
5.8.1.1. Overall Survival	20
5.8.2. Secondary Efficacy Analyses	21
5.8.2.1. Progression-Free Survival	21
5.8.2.2. Objective Response Rate	22
5.8.2.3. Duration of Response	23
5.9. Health Outcomes/Quality-of-Life Analyses	23
5.10. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods	27
5.11. Healthcare Resource Utilization	28
5.12. Safety Analyses	28
5.12.1. Phase 2	28
5.12.2. Extent of Exposure	29
5.12.2.1. Olaratumab	29
5.12.2.2. Nab-paclitaxel	29
5.12.2.3. Gemcitabine	29
5.12.3. Adverse Events	30
5.12.4. Clinical Laboratory Evaluation	33
5.12.5. Vital Signs and Other Physical Findings	33
5.12.6. Electrocardiograms	33
5.13. Subgroup Analyses	33
5.14. Important Protocol Deviations	33
5.15. Interim Analyses, Data Monitoring and Advisory Committee	34
5.15.1. Introduction	34
5.16. Annual Report Analyses	34
5.16.1. Clinical Investigator Brochure	34
5.16.2. Development Safety Update Report	35
5.17. Clinical Trial Registry Analyses	35
6. Unblinding Plan	36
7. References	37
8. Appendices	38

Storia del documento (SAP)

Obiettivi dello studio

Struttura generale del trial (ospedali, cliniche, trattamenti, caratteristiche dei pazienti)

Numero minimo di pazienti da includere nel trial

Come assegniamo i trattamenti ai pazienti

Come trattiamo i pazienti che lasciano lo studio o i dati mancanti?

Popolazione di interesse

Eventi avversi monitorati nel trial

Cosa si è dovuto cambiare durante il trial e perché

Lo studio va avanti «in continuo», le analisi ad-interim si fanno ogni tanto..

Analisi annuali obbligatorie

Analisi principale: modello statistico usato per dimostrare gli obiettivi del trial

Esempio 1: Analisi Statistica nei Clinical Trials

Clinicians

Richiedono un'evidenza che il trattamento «funzioni»

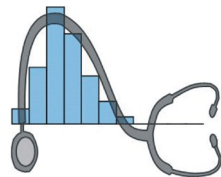
- ente governativo per la sicurezza
- articolo scientifico per «promozione»
- R&D per innovazione



BS - Biostatistician

Decidono come «misurare» tale evidenza in termini statistici

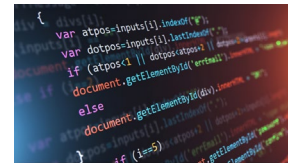
- che tipo di analisi servono
- su quali trial
- con quanti e quali pazienti



SP - Statistical Programmer

Creano le tabelle e i grafici dell'analisi richiesta

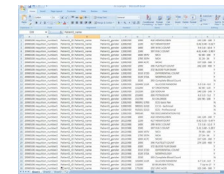
- codice nel software statistico più appropriato (per es. SAS)



DM - Data Manager

Forniscono e puliscono i dati necessari per l'analisi

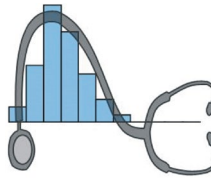
- dati mancanti
- errori



BS - Biostatistician

Interpretano le tabelle e i grafici dell'analisi

- Validazione delle ipotesi
- Comunicazione dei risultati statistici ai clinici



Clinicians

Prendono una decisione in base all'esito (buono o no) dell'analisi sulle performance del trattamento

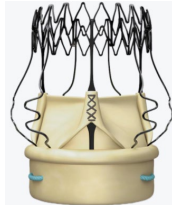


Esempio 2: Model and Sample size for V&V tests

R&D

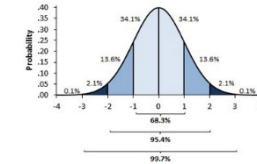


Dettagli dell'esperimento di laboratorio



Dati preliminari

Quality Statistician



1. Realizza **analisi preliminari** per determinare:

- Tipo di dati (numerici, categorici, ...)
- Qualità dei dati (valori strani, mancanti o errati)
- Campionamento (random o fissato, frequenza alta o bassa)
- Sorgenti di variabilità (pezzi, operatori, giorni, ...)
- Validità delle ipotesi (indipendenza, normalità, ...)

2. Individua e formalizza il **modello statistico** adatto per il test grazie alla:

- propria conoscenza/esperienza (se metodo noto)
- lettura di libri/articoli scientifici (se metodo non noto)

3. Calcola il **sample size** per il test attraverso:

- Formule matematiche
- Software statistici

Modello statistico



NIH Public Access

Author Manuscript

Biom J. Author manuscript; available in PMC 2014 January 08.

Published in final edited form as:

Biom J. 2014 January ; 56(1): . doi:10.1002/bimj.201300048.

Stratified Fisher's Exact Test and its Sample Size Calculation

Sin-Ho Jung¹

Department of Biotatistics and Bioinformatics Duke University Durham, NC 27710, U.S.A.

Summary

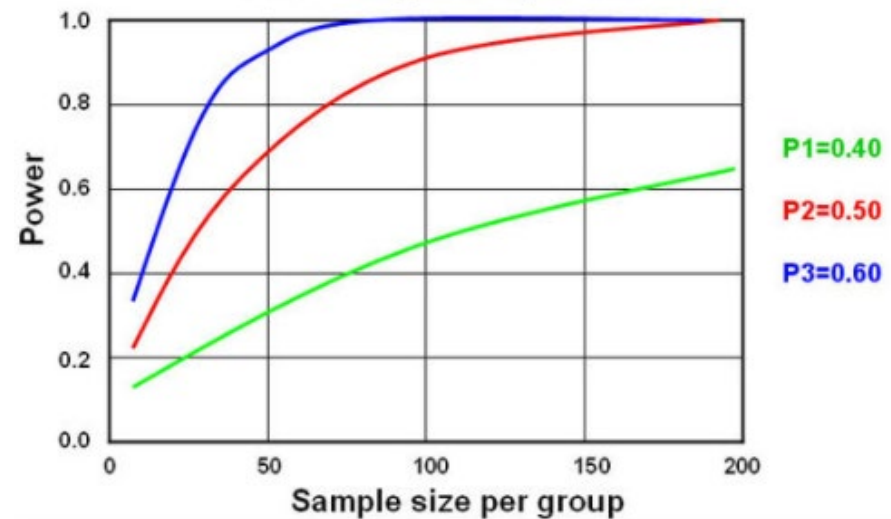
Chi-squared test has been a popular approach to the analysis of a 2×2 table when the sample sizes for the four cells are large. When the large sample assumption does not hold, however, we need an exact testing method such as Fisher's test. When the study population is heterogeneous, we often partition the subjects into multiple strata, so that each stratum consists of homogeneous subjects and hence the stratified analysis has an improved testing power. While Mantel-Haenszel test has been widely used as an extension of the chi-squared test to test on stratified 2×2 tables with a large-sample approximation, we have been lacking an extension of Fisher's test for stratified exact testing. In this paper, we discuss an exact testing method for stratified 2×2 tables which is simplified to the standard Fisher's test in single 2×2 table cases, and propose its sample size calculation method that can be useful for designing a study with rare cell frequencies.

NIH-PA Author Manuscript

NIH-PA

Calcolo del sample size

Power as a function of Effect Size (P)
and Sample Size
Two Sample Proportion



Grazie per l'attenzione

...e in bocca al lupo per il vostro futuro!