

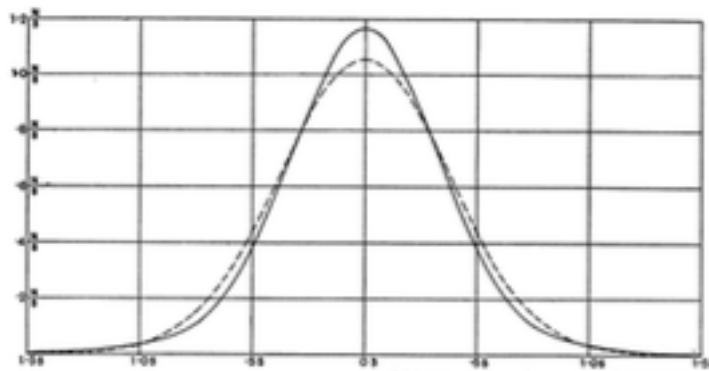


# Metodologie di Analisi dei dati biomedici

Corso di Laurea Magistrale in Medicina e Chirurgia A.A. 2015-16

Massimo Borelli, Ph.D.

La Statistica è una scienza in continua (e pervasiva) evoluzione: le tecnologie informatiche consentono oggi di fare in pochi minuti quello che cento anni fa avrebbe richiesto mesi di lavoro (nell'immagine, il disegno del 1908 pubblicato da Student per illustrare le differenze tra la distribuzione gaussiana e la distribuzione  $t$ ).



Ad esempio, le tecnologie informatiche al servizio della Medicina Molecolare generano *big data* con estrema immediatezza, rovesciando persino il processo classico di 'scoperta scientifica' (domanda di ricerca → esperimento → dati → conclusione). Di riflesso, oggi, una decisione *evidence-based* sempre di più richiede di discostarsi dall'approccio novecentesco del test statistico, privilegiando l'approccio legato alla simulazione ed all'interpretazione dei **modelli inferenziali**.

## metodologia biostatistica

### design cross-section e design longitudinali

lun 11 aprile **introduzione**, *metodologia*, fonti di letteratura, il ruolo dello 'Statistico' nel design dello studio (esempio: sample size), il ruolo delle tecnologie, gli strumenti in rete, il concetto di test statistico ed il concetto di modello statistico. Frequentisti vs. Bayesiani.

**la tecnologia**, *fogli elettronici e linguaggi dedicati*, il linguaggio R, l'interfaccia R Commander, la suite R Studio.

**modellare i dati 'evidenced based'**, *le variabili aleatorie*, la normale, lognormale, binomiale, poisson, weibull (time-to-event) nei design cross-section.

## ■ i modelli statistici lineari

### design cross-section

lun 11 aprile **la regressione multipla**, *interpretare il modello lineare*, la bontà del modello minimale adeguato e la sua diagnostica: outlier e punti di leva.

**la Anova**, *interpretare il modello lineare*, l'importante questione metodologica dei confronti multipli.

**la Ancova**, *interpretare il modello lineare*, criteri di selezione del modello minimale adeguato.

## ■ i modelli statistici lineari generalizzati

### design cross-section

gio 14 aprile **la regressione logistica**, *i modelli lineari generalizzati binomiali*, il predittore lineare e la funzione di collegamento logit.

**count data**, *i modelli lineari generalizzati di Poisson*, il predittore lineare e la funzione di collegamento.

**sopravvivenza**, *il modello di Cox*, failure e dati censurati, hazard.

## ■ i modelli ad effetti misti

### design longitudinali (misure ripetute)

gio 21 aprile **effetti fissi ed effetti casuali**, *eliminare il problema della pseudoreplica*, dati correlati in oculistica, design longitudinali: criteri di selezione del modello minimale adeguato, selezione Monte Carlo.