

# Medical Statistics: anàlisi estadístic

Cristian Tebé Cordero, PhD

ctebe@igtp.cat

Unitat de Suport i Recerca en Bioestadística

Germans Trias i Pujol Research Institute and Hospital

# De què parlarem?

- Pla d'anàlisi estadístic
- Variabilitat
- Inferència estadística
  - ▶ Intervals de confiança
  - ▶ Contrastos d'hipòtesis
- Modelització

“Statistics is the science of making decisions under uncertainty”

Leonard Savage

# Pla d'anàlisi estadístic

Document que escriu com es recopilaran, analitzaran i interpretaran les dades.

- Objectius
- Disseny
- Variables
- Metodologia estadística
  - ▶ Descriptius
  - ▶ Contrastos
  - ▶ Modelització
  - ▶ Subgrups
- Estimació de la grandària mostral
  - ▶ Interim
- Gestió de les dades
  - ▶ Missings
- Presentació de resultats

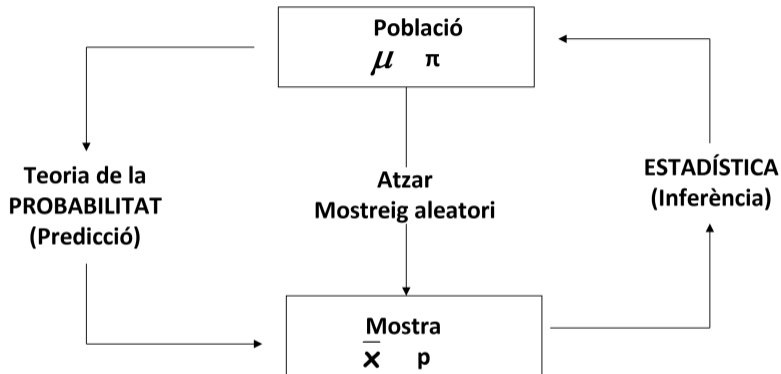
# Variabilitat

La variança és una mesura de la dispersió o extensió del conjunt d'observacions o valors d'una variable.

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

$x_i$  : cada valor de la mostra  $\bar{x}$  : mitjana de la mostra  $n$  : grandària mostral

# Inferència estadística



La inferència estadística és el procés d'arribar a conclusions a partir de dades que estan subjectes a variació aleatòria.

## Estimador puntual

Quin és el dolor no controlat en una mostra de pacients amb metàstasis òssies?

$$\bar{x} = \frac{\sum_{i=1}^n (x_i)}{n}$$

$$p = \frac{\sum_{i=1}^n (\text{favorables})}{n}$$

## Interval de confiança

Determina un rang de valors on podem confiar, amb una nivell fixat a priori, que el valor del paràmetre de la població que volem conèixer es troba dins del rang.

$$\text{Interval de confiança} = \text{estimador puntual} \pm Z_{\alpha/2} \left( \frac{\text{desviació estàndard}}{\sqrt{n}} \right)$$

On  $Z$  és la distribució de probabilitat



## Interval de confiança

La prevalença del dolor mal controlat és del 22% amb un IC95% 18% a 26%.

Tenim un 95% de confiança que el valor veritable de la prevalença estigui entre el 18% i 26%.

Si prenem 100 mostres de mida  $n$  i calculem intervals de confiança de la prevalença en cada mostra, esperaríem que el 95% d'aquestes l'intervals incloguin el veritable valor (desconegut) de la prevalença poblacional.

# Contrastos d'hipòtesis

Els contrastos d'hipòtesis són un mètode de presa de decisions on se sotmeten a comprovació dues hipòtesis complementàries sobre el valor d'un paràmetre.

$$H_0 : \mu_{\text{control}} = \mu_{\text{experimental}}$$

$$H_a : \mu_{\text{control}} \neq \mu_{\text{experimental}}$$

# Contrastos d'hipòtesis

**P-valor:** És la probabilitat d'obtenir un valor igual o més extrem que l'observat a la mostra sota la hipòtesi nul·la.

**Significació estadística:** Si  $p\text{-valor} > \text{nivell de significació}$ : no podem rebutjar la hipòtesi nul·la i direm que el resultat no és estadísticament significatiu.

Si  $p\text{-valor} \leq \text{nivell de significació}$ : podem rebutjar la hipòtesi nul·la i direm que el resultat és estadísticament significatiu.

## EL RESULTAT ÉS CLÍNICAMENT RELLEVANT?



- Els valors  $p$  no mesuren la probabilitat que la hipòtesi estudiada sigui veritable, o la probabilitat que les dades hagin estat produïdes només per atzar.
- Les conclusions científiques i les decisions empresarials o polítiques no s'haurien de fonamentar únicament en el fet que el valor  $p$  sobrepassi un llindar específic.
- Un valor  $p$ , o la significació estadística, no mesura la mida d'un efecte o la importància d'un resultat

Ronald L. et al (2016): The ASA's statement on  $p$ -values: context, process, and purpose. The American Statistician

# Modelització

La modelització estadística és el procés de construir, ajustar i validar models matemàtics per descriure i entendre les dades i com s'interrelacionen.

## **Quin és el meu objectiu?**

Vull descriure una relació.

Vull estimar una relació causal.

Vull predir un resultat.

Vull calibrar un model.

## Validar, validar, validar

- Calibració
- Discriminació
- Consistència en subgrups coneguts
- Validesa externa
- ...
- Manteniment

# Missatges finals

- Pla d'anàlisi estadístic
- Descripció de les dades
- Significació clínica
- Modelització
- Consulteu un bioestadístic!

## Referències

- Martínez Gonzalez, M.A.; De Irala, J; Faulin, FJ. Bioestadística amigable. 2ª Edició. Madrid: Díaz de Santos, 2006
- Argimon Pallás JM, Jiménez Villa J. Métodos de investigación clínica y epidemiológica. 3ª ed. Madrid: Elsevier; 2005.
- Daniel, WW. Bioestadística. Base para el análisis de las ciencias de la salud. 4ª Edició Mexico: Limusa Wiley, 2002.

### Recursos on-line:

- Pita Fernández S. Metodología de la investigación:  
<https://www.fisterra.com/formacion/metodologia-investigacion/>
- Statistics at Square One (BMJ Publishing Group 1997):  
<http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one>



Moltes gràcies!



**Germans Trias i Pujol**  
Hospital

