

# p-valors, hem d'enterrar-los o hem d'aprendre com usar-los?

## Contertulians:

**Anabel Blasco (UAB)**  
**Malu Calle (UVIC)**  
**Erik Cobo (UPC)**  
**Joan Guàrdia (UB)**  
**Miquel Porta (IMIM)**



Institut  
d'Estudis  
Catalans



**#pvalueSoCE**

**Moderadores: Lupe Gómez (UPC) i Rosa Lamarca (AstraZeneca)**

*8 de maig 2019 a les 18:00h. Sala Pi i Sunyer, Institut d'Estudis Catalans*

# ANABEL BLASCO

- Coordinadora i consultora estadística. Servei d'Estadística Aplicada de la UAB.
- Professora associada. Dept. d'Economia i Història Econòmica (UAB)
- Membre grup de recerca "Advanced Stochastic Modelling"
- **Línees de recerca:**
  - Modelització estadística
  - Models zero-inflats
  - Models captura-recaptura aplicats a ecologia, biologia, ...
- **Membre fundacional de la Societat Catalana d'Estadística**



# MALU CALLE

- Catedràtica i Directora departament Biociències. Facultat de Ciències i Tecnologia (UVic-UCC)
- Coordinadora màster en Anàlisi de Dades Òmiques (UVic-UCC)
- Presidenta (2012-2013) de la Sociedad Española de Biometría, regió de la *International Biometrics Society*
- **Línees de recerca:**
  - Estadística genètica
  - Anàlisi de dades òmiques i microbiòmiques
  - Anàlisi de supervivència
- **Membre fundacional de la Societat Catalana d'Estadística**



# ERIK COBO

- Professor titular. Dept d'Estadística i Investigació Operativa (UPC)
- IP UPC Methods in Research on Research (MiRoR)
- Membre grup de recerca Bioestadística i Bioinformàtica (GRBIO UPC-UB)
- Editor metodològic de Medicina Clínica (2000-18) i de Trials (2012-18)
- **Línees de recerca:**
  - Bioestadística
  - Assajos clínics
  - Causalitat i reproducibilitat
  - Guies
- **Membre fundacional de la Societat Catalana d'Estadística**



# JOAN GUÀRDIA

- Catedràtic Metodologia de Ciències del Comportament. Facultat de Psicologia (UB)
- IP Grup de Recerca 'Quantitative Psychology'
- Coordinador Grup d'Innovació Docent
- Membre Consell de Direcció de l'Institut de Neurociències
- Membre Grup Interdisciplinari de Reflexió i de Solucions Matemàtiques per a Entitats
- Membre de l'Institut de Sistemes Complexes de la UB
- **Línees de recerca:**
  - Tècniques Multivariants
  - Structural Equation Models
  - Computational Neuroscience
- **Membre fundacional, Vocal i VicePresident de la Societat Catalana d'Estadística (2014 al 2016)**

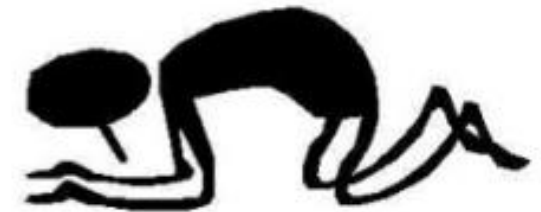
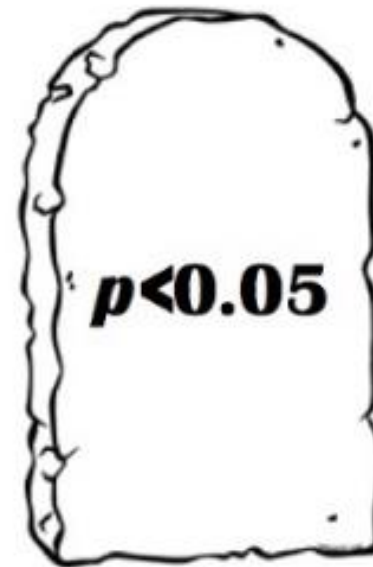
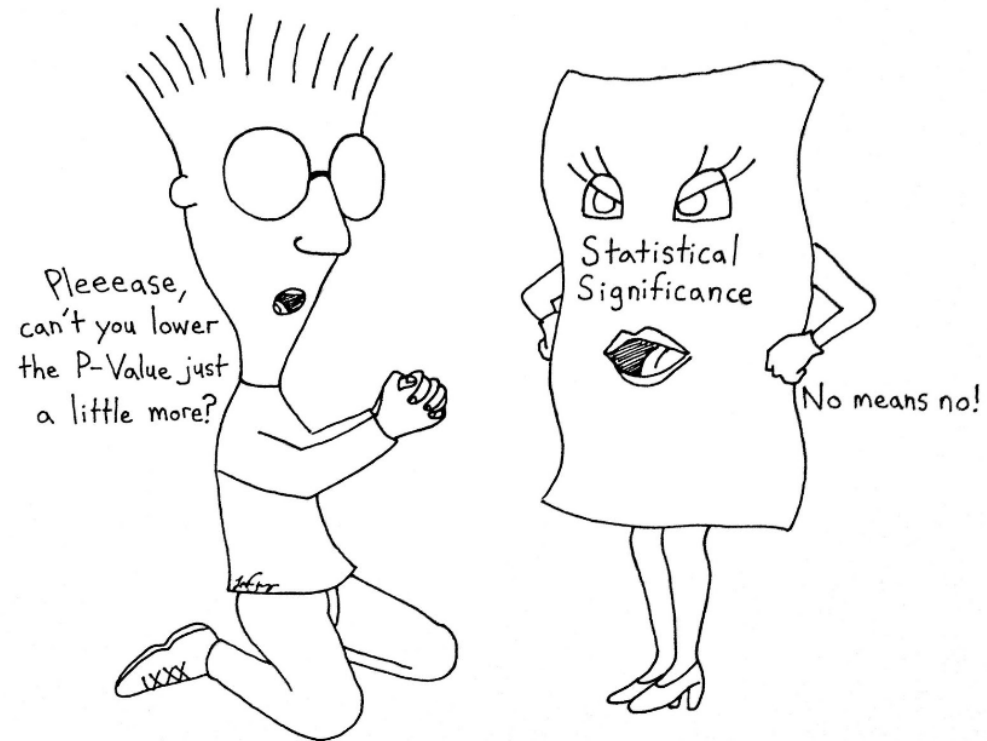


# MIQUEL PORTA

- Catedràtic de Salut Pública (UAB)
- Catedràtic adjunt d'Epidemiologia, Universitat de Carolina del Nord
- Investigador, Institut Hospital del Mar d'Investigacions Mèdiques
  
- **Línees de recerca:**
  - Epidemiologia molecular, clínica i ambiental del càncer de pàncrees
  - Biomonitorització i avaluació de l'impacte sobre la salut humana dels agents químics ambientals
  - Inferències causals en medicina i salut pública



# INTRODUCCIÓ



Font: <https://towardsdatascience.com/a-case-study-of-the-p-value-f0d708861334>

Font: <https://towardsdatascience.com/the-ambiguity-of-p-value-what-is-it-3a8d552e6632>

# QUAN/COM ES VA ORIGINAR LA CONTROVÈRSIA I QUÈ S'HA FET AL RESPECTE?

## ONCE UPON A TIME in UK...

### Edgeworth (1885) + Fisher's (1925) t-test per la diferència de mitjanes:

“... we can calculate the standard deviation of the mean of a random sample of any size, and so test whether or not *it differs significantly* from any fixed value. If the difference is many times greater than the standard error, *it is certainly significant*, and it is a convenient convention to take twice the standard error as the limit of significance ; *this is roughly equivalent to the corresponding limit  $P=.05$* ”

### Fisher: proves de significació

## I quasi al mateix temps a Paris i Varsòvia

**Neyman i Pearson (1928-1934).** Teoria matemàtica rigorosa i conceptualment clara

- hipòtesi alternativa
- errors de tipus I i II,
- nivell de significació
- potència d'un test

“Importància de situar en una seqüència lògica les etapes del raonament en cercar la solució inferèncial”

“El seu punt de partida va ser sempre un problema pràctic, *sense perdre de vista les aplicacions.*”

### N-P: proves d'hipòtesis



## 2014: Editorial a *BASP: Basic and Applied Social Psychology*

**NHSTP és invàlid. *From now on, BASP is banning the NHSTP.***

Null hypothesis significance testing procedure (NHSTP) aproximació híbrida (incongruent) de Fisher i Neyman-Pearson.

### **NHSTP:**

- interpretar un valor P més petit com a evidència més forta que la hipòtesi nul·la és falsa.
- acceptar les conclusions com a "significatives" si el valor P és inferior a 0,05.

**5 anys després:** els autors tendeixen a exagerar les conclusions (Fricker et al)

## 2016: "The ASA's Statement on p-Values: Context, Process and Purpose". Wasserstein & Lazar

Declaració de ASA: 6 principis per millorar la pràctica i la interpretació de la ciència quantitativa.

L'ús del p-valor incideix en:

- la recerca i el seu finançament
- l'acceptació d'articles a les revistes
- l'avenç professional i l'educació
- les polítiques públiques i el periodisme

### **Què és el valor p?**

Un valor p és la probabilitat d'obtenir un valor igual o més extrem que el valor observat a partir de les dades calculat sota un model estadístic especificat ( $H_0$ )

## Declaració de l'ASA sobre la significació estadística i els valors p

5 de febrer de 2016

Editat per Ronald L. Wasserstein, director executiu

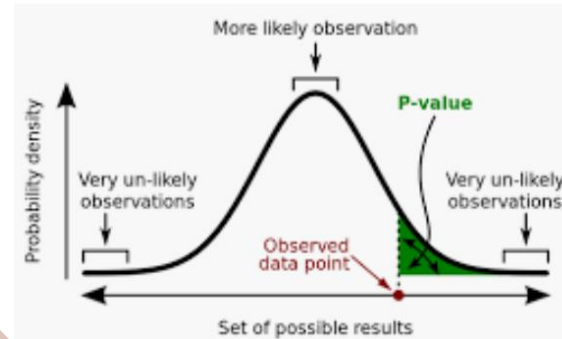
en nom de la Junta Directiva de l'*American Statistical Association*[1]

### Introducció

Els darrers anys, l'augment de la quantificació en la recerca científica i la proliferació de grans i complexos conjunts de dades han ampliat l'abast de les aplicacions dels mètodes estadístics. Aquest fet ha creat noves possibilitats per al progrés científic, però també ha portat certa preocupació sobre les conclusions obtingudes a partir de les dades. La validesa de les conclusions científiques i la possibilitat de reproduir-les no depèn únicament dels mètodes estadístics en sí mateixos. L'elecció apropiada de les tècniques, la correcció de les anàlisis efectuades i la interpretació adequada dels resultats estadístics juguen també un paper essencial a l'hora de garantir conclusions sòlides i per garantir que la incertesa que les envolta estigui ben representada.

El fonament de moltes conclusions científiques publicades és el concepte de "significació estadística", normalment avaluada mitjançant un índex anomenat valor p (*p-value*). Ara bé, tot i que el valor p pot ser una mesura estadística útil, sovint s'usa de forma incorrecta i també és mal interpretada. Això ha dut a que algunes revistes científiques dissuadeixin del seu ús i a que alguns científics i estadístics en recomanin el seu abandonament, basant-se en arguments que essencialment són els mateixos des de que els valors p van ser introduïts per primera vegada.

En aquest context, l'*American Statistical Association* (ASA) creu que la comunitat científica podria beneficiar-se d'una declaració formal que aclarís alguns principis que són generalment acceptats i que són implícits en la correcta utilització i interpretació dels valors p. Els aspectes considerats aquí no afecten només la recerca, sinó també el seu finançament, les pràctiques de les revistes, l'avenç professional, l'educació científica, les polítiques públiques, el periodisme i el dret. Aquesta declaració no pretén resoldre totes les qüestions relacionades amb les bones pràctiques estadístiques, ni tampoc resoldre les controvèrsies fonamentals. Més aviat presenta en termes no tècnics una selecció breu de principis que podrien millorar la pràctica i la interpretació de la ciència quantitativa, d'acord amb un consens ampli assolit a la comunitat estadística.



### Principis (traduïts per la SoCE i ara reduïts)

1. El valor p **INDICA** una certa incompatibilitat de les dades amb un model especificat
2. El valor p **NO** mesura la probabilitat que la hipòtesi sigui certa ni que les dades hagin estat produïdes només per l'atzar
3. No basis les conclusions únicament en el fet que  $p < 0.05$
4. Realitza inferències apropiades amb informes complets i transparents
5. Un valor p **NO** mesura la mida d'un efecte o la importància d'un resultat
6. Un valor p **NO** proporciona una mesura de l'evidència en relació a un model o una hipòtesi

## 2019 “Scientists rise up against statistical significance” en Nature

Amrhein, Greenland, McShane i més de 800 signants

**MAI** hauríem de concloure que

- no hi ha "cap diferència" o "associació" **només** perquè el valor  $P > 0.05$  o, perquè un interval de confiança inclou zero.
- dos estudis es troben en conflicte perquè un té un resultat estadísticament significatiu i l'altre no.

Aquests errors desaprofiten esforços de recerca i desinformen sobre les decisions que s'han de prendre.

## 2019: “Moving to a World Beyond $p < 0.05$ ” en American Statistician

Wasserstein, Schirm & Lazar

- 43 articles innovadors i provocadors d'estadístics amb visió de futur
- Parlant sobre el que no es pot fer amb els valors-p i oferint idees sobre què fer sobre el difícil problema de
  - **separar el senyal del soroll**
  - **prendre decisions sota incertesa**

**ELLS SÓN MOLT CLARS:**

Stop using the term “statistically significant”, “significantly different”, “nonsignificant”

**I NOSALTRES QUE OPINEM?**



# BLOCS

1. Posicionament en vers el p-valor

2. Alternatives al p-valor

3. Paper de l'estadístic en un bon ús del p-valor



# BLOC 1

Vers la pregunta d'aquesta taula rodona 'p-valors, hem d'enterrar-los o hem d'aprendre com usar-los?'

Quina és la teva posició al respecte?

Creus que ve influenciada per l'àrea on desenvolupes la teva carrera professional?

**Posicionament en vers el p-valor**



## BLOC 2

Si es donés la possibilitat que la comunitat científica decidís abandonar l'ús del p-valor i el terme significació estadística

quin podria ser un mètode alternatiu per dissenyar/interpretar un estudi i extreure conclusions?

# Alternatives al p-valor



## BLOC 3

Si part del problema resideix en com s'han usat e interpretat els p-valors,

quin rol i quines iniciatives hauria de prendre la comunitat estadística per fer-ne un bon ús?

**Paper de l'estadístic en un bon ús del p-valor**



**CLOENDA**

**El minut d'or**



# ATOMA

(Moving to a World Beyond  $p < 0.05$ )

- **A**ccept uncertainty
- Be **T**houghtful
- Be **O**pen
- Be **M**odest
- **A**bandon “statistical significance”

# AROMA

- **A**cepta l'incertesa
- Sigues **R**eflexiva/iu
- Sigues **O**bert/a
- Sigues **M**odest/a
- **A**bandona la “significació estadística”

# ATOM

- **A**ccept uncertainty
- Be **T**houghtful
- Be **O**pen
- Be **M**odest

# AROM

- **A**cepta l'incertesa
- Recerca **R**eflexiva
  - Objectius clars
  - Efectes rellevants
  - Implicacions i precisió de l'estimació
  - Especificació correcte del model
  - Múltiples aproximacions
- Sigues **O**bert (i flexible)
- Sigues **M**odest (no hi ha un únic model estadístic)