The importance of statistics

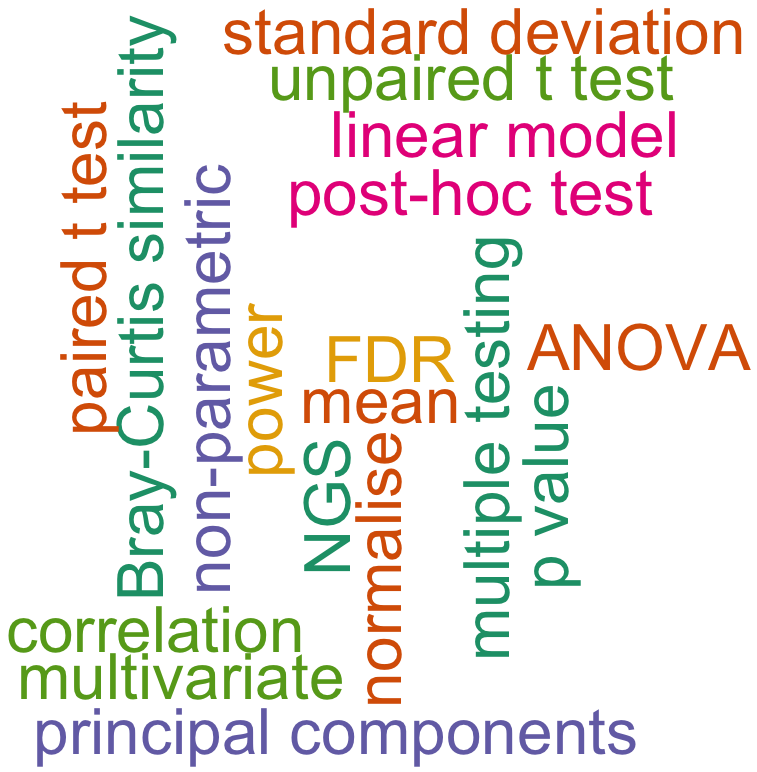
BABS Honours Induction

Peter Geelan-Small, UNSW Stats Central

14/02/2022

Statistics important?

Statistics in BABS



Outline

- Not a rundown of specific statistical methods
- Some important general principles
- Guidelines for good statistical practice
- Some hopefully useful resources
- A few random questions along the way

If you want to ask a question as we go, put your virtual hand up!

Trivia question 1

https://forms.gle/Pyw28ppJuCwtTWL59

General statistical principles

Let's find a crisis!

How about COVID-19?!

Wolkewitz and Puljak *BMC Medical Research Methodology* (2020) 20:81
https://doi.org/10.1186/s12874-020-00972-6

BMC Medical Research Methodology

EDITORIAL Open Access

Methodological challenges of analysing COVID-19 data during the pandemic

Martin Wolkewitz[1] and Livia Puljak[2*]

...

Some basic principles (Wolkewitz & Puljak 2020)

- Define research questions
  - What standards of care are most effective?
  - What is the effect of adjunct therapies?
- Define outcome and explanatory variables
- Design study using an appropriate study design
- Collect the data you need to answer your research question
- Use standardised protocols to collect and analyse data
- “Statistical expertise is needed to understand potential effects on the complexity of clinical endpoints”
- “Statistical models will play a major role in ‘fighting panic with information’ to avoid or at least minimize the risk of bias”

General statistical principles

Another crisis flagged in 2005 … or is there?

PLOS MEDICINE

OPEN ACCESS

ESSAY

Why Most Published Research Findings Are False

John P. A. Ioannidis

Published: August 30, 2005 • https://doi.org/10.1371/journal.pmed.0020124

…

General statistical principles

Is there a *reproducibility* crisis in science?

https://www.youtube.com/watch?v=j7K3s_vi_1Y&feature=emb_logo

Irreproducibility

NATURE | NEWS

Irreproducible biology research costs put at $28 billion per year

Study calculates cost of flawed biomedical research in the United States.

Monya Baker

09 June 2015

Original article: Freedman et al. doi:10.1371/journal.pbio.1002165

...

School of BABS Seminar - 20 November, 2020

Why are so many cell and molecular biology experiments and findings irreproducible?

Trivia question 2

https://forms.gle/6NfUu5x8QaSUXECx6

Reproducibility - experimental design

- Collect the right data to answer the research question and *enough* of the right data

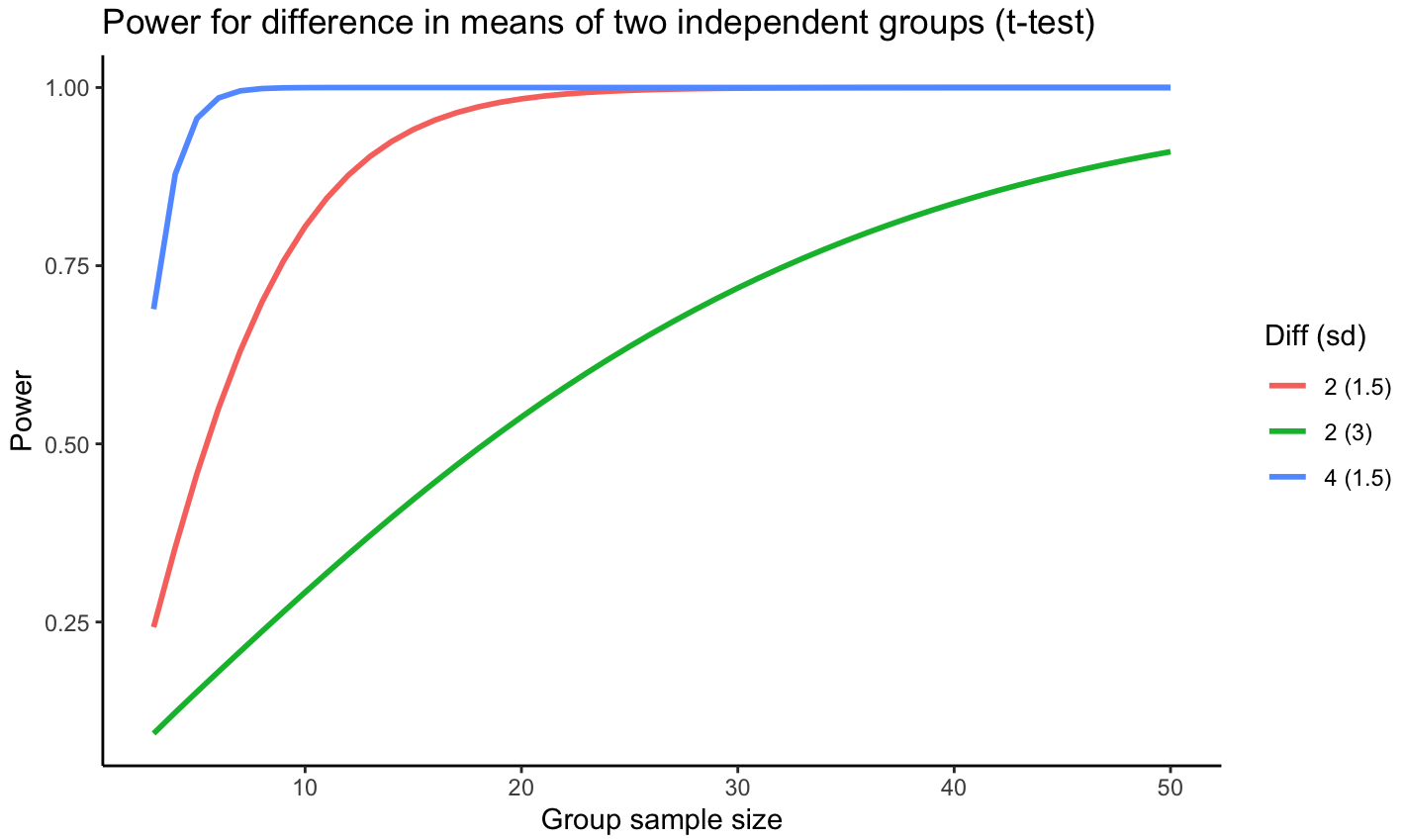
How much data do I need to be able to detect the difference I want to detect?

- Power - chance of detecting a true difference (true positive)
- Difference to detect - what size is biologically/clinically important?
- Variability - how variable (noisy) is what I'm studying?

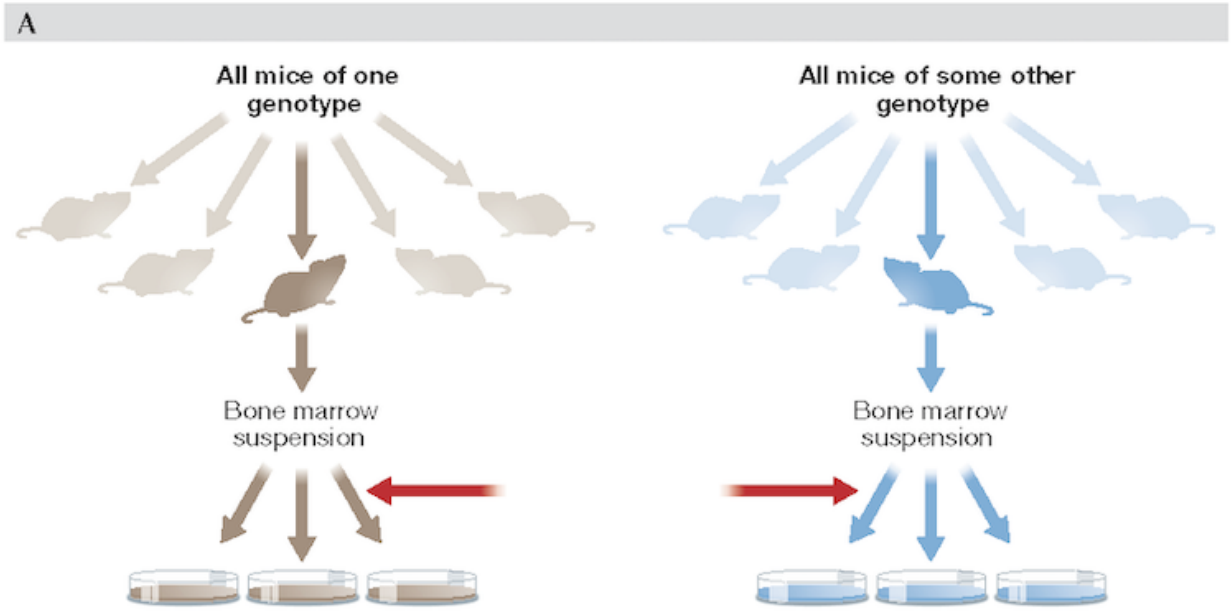
Sample size calculator: https://vbiostatps.app.vumc.org/ps/

Reproducibility - experimental design

Power, size of effect, variability, sample size …



Sample size = … - but what is the "replicate"?



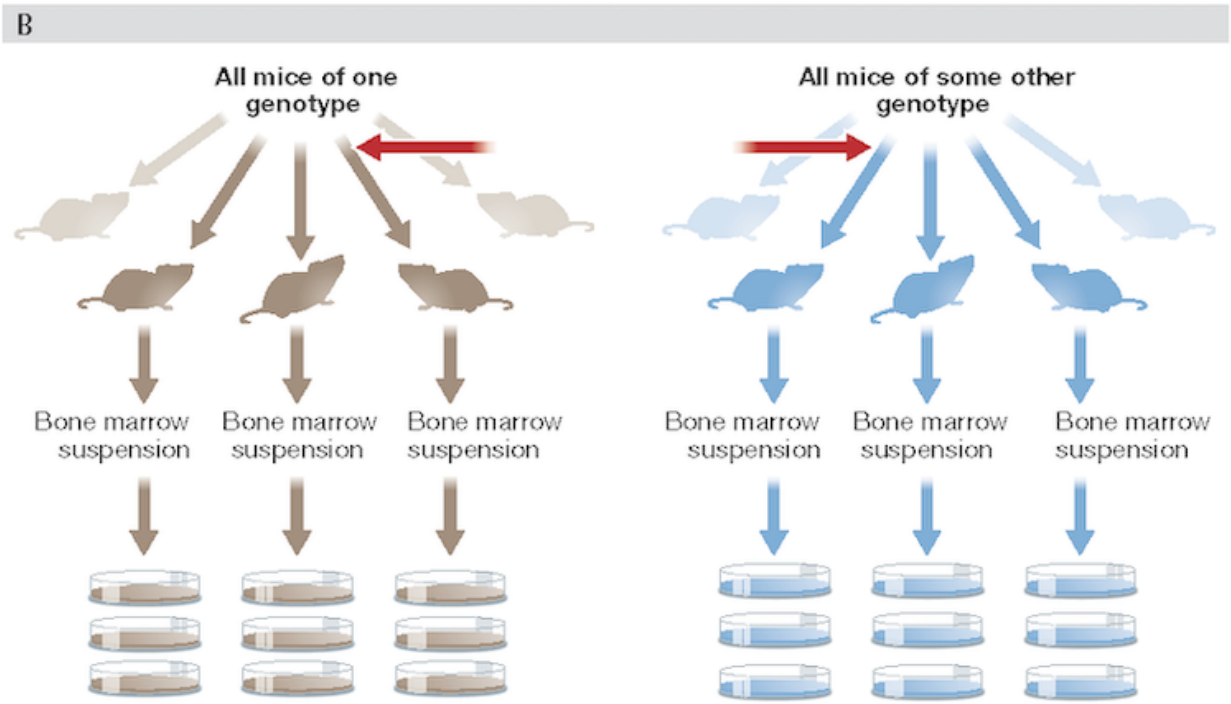
…

Technical replication only

- Replicates within each mouse are not independent measurements[1]

[1]: Vaux 2012 doi: 10.1038/embor.2012.36

Sample size = … - but what is the "replicate"?



…

Biological and technical replication[2]

[2]: Vaux 2012 doi: 10.1038/embor.2012.36

Trivia question 3

https://forms.gle/9v92pg9vs6toDZeC7

Sample size = … - but what is the "replicate"?

- Technical replication is useful to increase the precision of your measurements.
- Biological replication (this is what "sample size" means) is what counts towards power.
- There is no statistical justification for a sample size of n = 3!

Power of the test, p-values, publication bias and statistical evidence

Life as a researcher can have a few surprises!

https://www.youtube.com/watch?v=kMYxd6QeAss

- Variance ($= (\text{std. dev.})^2$) measures variation, which introduces uncertainty into estimates
- Statistics is the science of learning from data, and of measuring, controlling and communicating uncertainty[3]

[3]: Davidian & Louis 2012 doi: 10.1126/science.1218685

P values - a problem? Who would've thought?!

Not another crisis! … Let's call it an issue instead.

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 & Douglas G. Altman

European Journal of Epidemiology **31**, 337–350(2016) | Cite this article

…

The ASA 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 Statement on *p*-Values: Context, Process, and Purpose, The American Statistician, 70:2, 129-133, DOI: 10.1080/00031305.2016.1154108

…

P values - a problem? Who would've thought?!

"P values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume"[4]

"Misinterpretation and abuse of statistical tests, confidence intervals, and statistical power … remain rampant [and] dominate much of the scientific literature"[5]

"Report P values sparingly."[6]

Let's look at P values …

[4]: Nuzzo 2014 Nature doi: 10.1038/506150a
[5]: [illegible] et al. 2016 doi: 10.1007/s10654-016-0149-3
[6]: Michel et al. 2020 doi: 10.1124/dmd.119.090027

P values - What's the idea behind them?

The FAMuSS study examined the percentage change in strength of the non-dominant arm as a result of exercise in two human genotypes.[7]

Question: Is there a difference between the two genotypes in the change in mean percentage strength?

A glimpse at the data

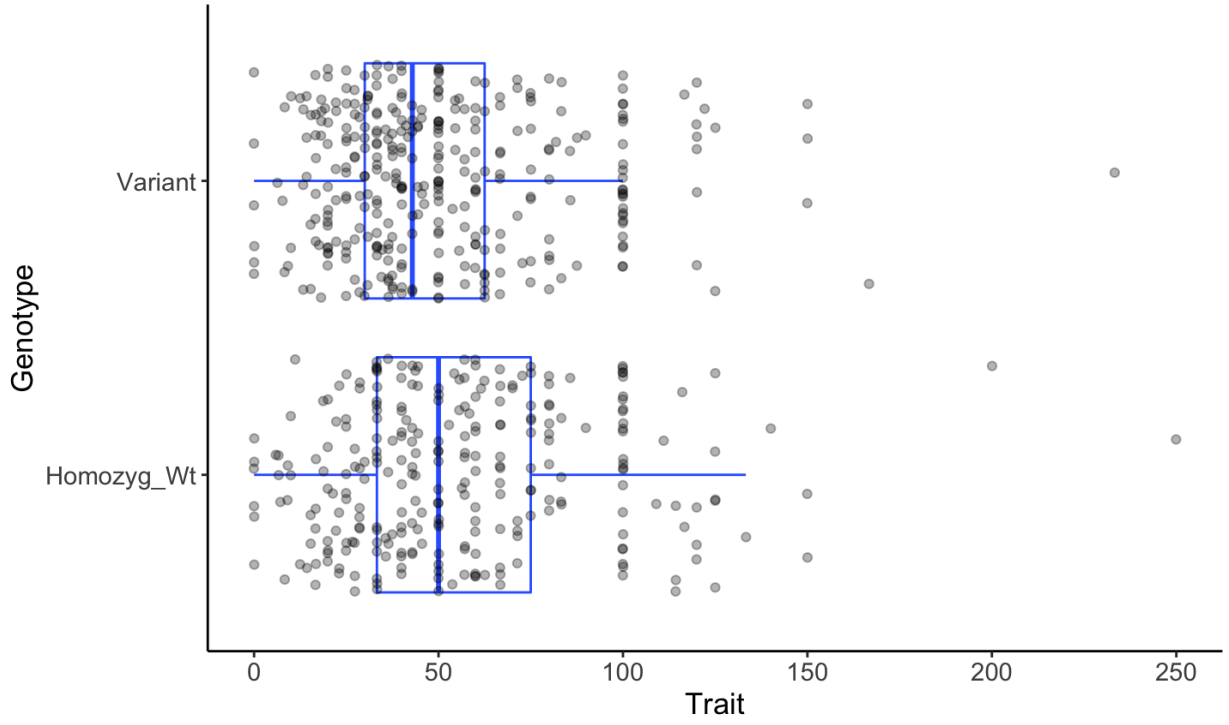
- 606 independent measurements of change in percentage strength.

```
## 'data.frame':    606 obs. of  2 variables:
##  $ Trait   : num  40 25 40 125 40 75 100 57.1 33.3 20 ...
##  $ Genotype: Factor w/ 2 levels "Homozyg_Wt","Variant": 1 2 1 2 1 1 2 1 1 2 ...
##  - attr(*, "na.action")= 'omit' Named int [1:133] 8 11 16 17 25 28 32 37 39 41 ...
##   ..- attr(*, "names")= chr [1:133] "8" "11" "16" "17" ...
```

[7] Ahmetov I. S., 2009, Applied Statistical Genetics with R: For Population-Based Association Studies, Springer, New York

P values - What's the idea behind them?

A picture of the data



...

Observed difference in means is 5.62 (%)

P values - What's the idea behind them?

What measure (i.e. statistic) can we use to investigate the difference in percentage strength between the two genotypes?

One possibility: The difference in means (we could use any other sensible statistic - e.g. standardised difference, t statistic, ...)

P values - What's the idea behind them?

1. We assume change (%) in mean strength is the same for both genotypes (Null hypothesis).
2. Shuffle the two genotype labels on the data values to get many possible arrangements of the order.
3. Calculate the mean strength change for the two genotypes for each arrangement.
4. Find the difference between the two genotype means for each arrangement. Do steps 2, 3 and 4 many times.
5. Look at the pattern of values of all these differences.
6. See **how unusual** the difference is that we got from our sample data.

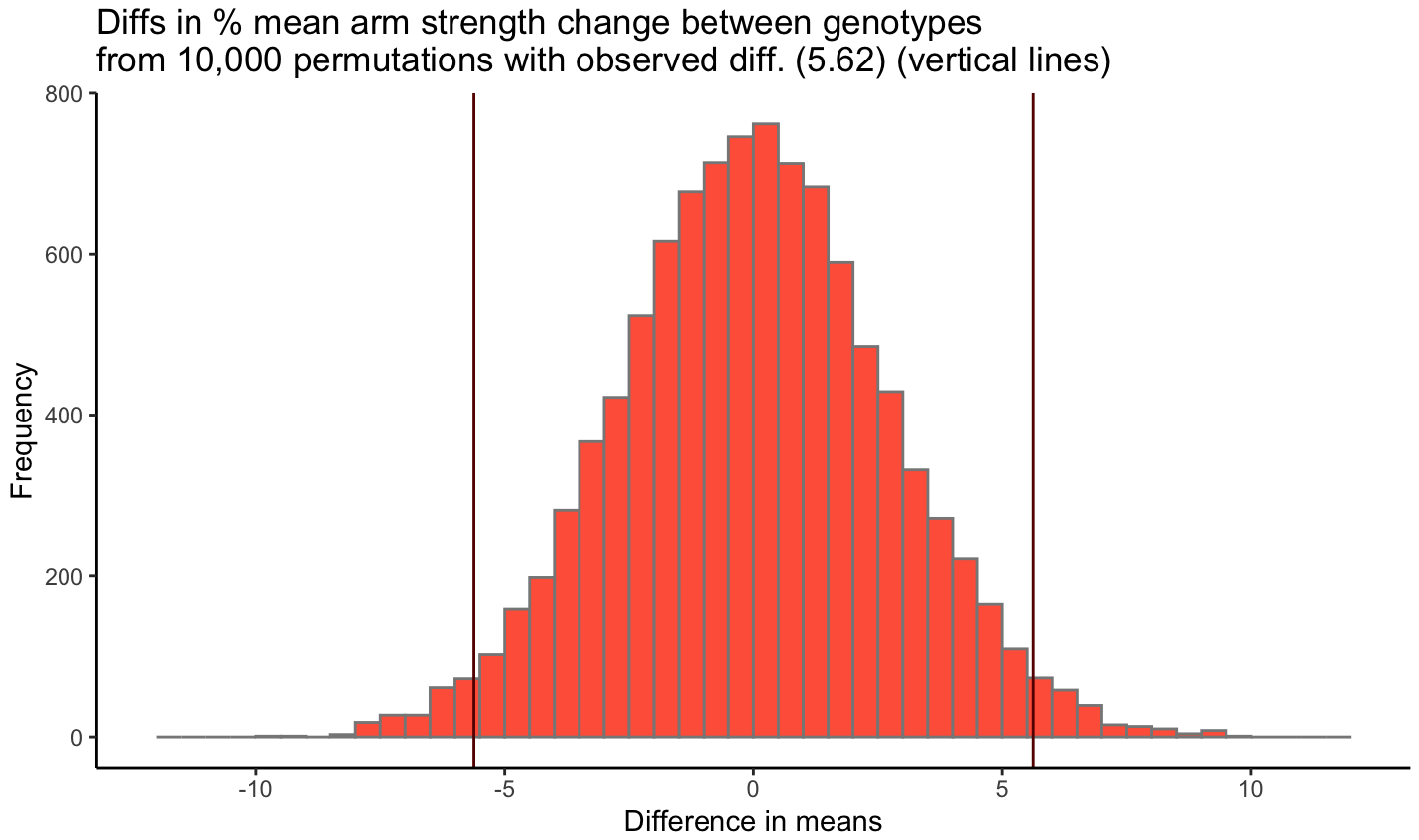
This technique makes no assumptions about the distribution of the data.

P values - What's the idea behind them?

Some examples of the first six genotype values from different shuffled arrangements. Original data order in first column.

Original	Shuffle_1	Shuffle_2	Shuffle_3	Shuffle_4
Homozyg_Wt	Homozyg_Wt	Variant	Homozyg_Wt	Variant
Variant	Variant	Homozyg_Wt	Homozyg_Wt	Variant
Homozyg_Wt	Variant	Homozyg_Wt	Variant	Homozyg_Wt
Variant	Homozyg_Wt	Homozyg_Wt	Variant	Variant
Homozyg_Wt	Homozyg_Wt	Variant	Variant	Variant
Homozyg_Wt	Variant	Variant	Homozyg_Wt	Homozyg_Wt

P values - What's the idea behind them?



P values - What's the idea behind them?

Observed difference in means is 5.62

- How unusual is our sample?
  - What is the proportion of differences out of our 10,000 values that are larger in magnitude than 5.62?
- There are 391 difference values that are at least as large in magnitude as the difference in our sample.
- The proportion of these differences out of 10,000 is 0.039
- This proportion tells us *how unusual* our sample is, *assuming* the null hypothesis of equal group means is true.

P values - What's the idea behind them?

- P value is an indication of *how unusual* a sample is, assuming the null hypothesis is true (in this case, that the two groups have the same mean).
- Proportion from permutation test is our P value for that test:
  - $p = 0.039$
- t test (two independent samples) of the null hypothesis of equal means (and check the assumptions for the test!!!):
  - $p = 0.038$

P values - What does a P value mean?

The P value is the probability of getting the data we got, *assuming the null hypothesis is true*.

- It tells us how consistent with the null hypothesis our sample data is.
- A P value should be interpreted on a *continuum*, not whether $p < 0.05$ or $p > 0.05$
- A P value:
  - is **not** the probability that the null hypothesis is true
  - is **not** the probability that the results were due to chance

Reporting results

Assuming

- you have chosen an appropriate statistical test and
- the assumptions of the test are satisfied …

you still have to *judge* whether your results are *biologically or clinically meaningful*

- P values do *not* tell the whole story - they say nothing about the size of an effect or how strong an association is[8]
- A small P value is **not** evidence of an important finding - a small P value does not guarantee scientific or clinical importance
- Report *both* the size of the effect and the P value!

STATS CENTRAL

[8]: Harrington 2019, New England Journal of Medicine, doi: 10.1056/NEJMe1906559

Trivia question 4

https://forms.gle/jfpaMggJcvxvwbFo7

Where does statistics fit into research?

To consult the statistician after an experiment is finished is often merely to ask [them] to conduct a post mortem examination. [They] can perhaps say what the experiment died of.[9]

PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS

EDITORIAL

Ten Simple Rules for Effective Statistical Practice

Robert E. Kass, Brian S. Caffo, Marie Davidian, Xiao-Li Meng, Bin Yu, Nancy Reid

Published: June 9, 2016 • https://doi.org/10.1371/journal.pcbi.1004961

...

[9]: Fisher R, 1938. Sankhya 4: 14-17

Life as a new researcher

- Keep a good work-life balance - have a life outside study to stay sane!
- Make a research project plan early
  - plan your statistical analysis at the start, not the end!
- Find other people doing similar research - socialise with your lab group and other colleagues
- Ask for help when you need it - asking for help can be a strength not a weakness - don’t struggle on your own!
- Things will not always go to plan - think about possible alternative statistical analyses in your initial planning
- Think positive and have a good time![10]

STATS CENTRAL

[10]: Taylor 2018, Twenty things I wish I'd known ..., Nature, doi: 10.1038/d41586-018-07332-x

Stats Central

Who are we?

Stats Central is a statistical consulting unit established at UNSW in 2016 to provide university-wide support for staff and HDR students during study design and analysis.

...

Resources - Guidelines

- Kass et al. 2016, Ten simple rules for effective statistical practice https://doi.org/10.1371/journal.pcbi.1004961
- Pollard 2019, Empowering statistical methods for cellular and molecular biologists https://doi.org/10.1091/mbc.E15-02-0076
- Vaux 2014, Basic statistics in cell biology. https://doi.org/10.1146/annurev-cellbio-100913-013303
- Vaux 2012, Repeats and replicates …. https://doi.org/10.1038/embor.2012.36
- Lord 2020, SuperPlots: communicating reproducibility and variability in cell biology. https://doi.org/10.1083/jcb.202001064

- Holmes and Huber 2019, *Modern Statistics for Modern Biology* http://web.stanford.edu/class/bios221/book/introduction.html

Resources - Guidelines continued

- Michel, M. C. et al., 2020, New author guidelines for displaying data and reporting data analysis and statistical methods in experimental biology. https://doi.org/10.1124/dmd.119.090027
- Lang 2016, Never P alone https://journal.emwa.org/statistics/never-p-alone-the-value-of-estimates-and-confidence-intervals/
- Lang and Altman 2016, Statistical analyses and methods in the published literature: the SAMPL guidelines https://journal.emwa.org/statistics/statistical-analyses-and-methods-in-the-published-literature-the-sampl-guidelines/
- Harrington, D. et al., 2019, New guidelines for statistical reporting in the Journal (New England Journal of Medicine) https://doi.org/10.1056/NEJMe1906559

- Greenland et al. 2016, Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations https://doi.org/10.1007/s10654-016-0149-3
- Wasserstein et al. 2019, Moving to a world beyond “p < 0.05” https://doi.org/10.1080/00031305.2019.1583913

Resources - Guidelines continued

- Naegle 2015, Criteria for biological reproducibility: what does “n” mean? https://doi.org/10.1126/scisignal.aab1125
- Nuzzo 2018, Tips for communicating statistical significance https://www.nih.gov/about-nih/what-we-do/science-health-public-trust/perspectives/science-health-public-trust/tips-communicating-statistical-significance

- Taylor, L. A., 2018, Twenty things I wish I’d known when I started my PhD, *Nature*, https://doi.org/10.1038/d41586-018-07332-x

Resources - Guidelines continued

Free R courses

- Keep an eye out here - UNSW Research Technology Services - these courses fill up very quickly: https://research.unsw.edu.au/research-technology-training
- On the ResTech web page above, click “R” under “Other Courses” for some other R courses. The “Learning R” course from LinkedIn Learning is free for UNSW students.
- https://www.edx.org (e.g. Data Analysis for the Life Sciences, Statistics and R, Bioinformatics, Introduction to Bioconductor, Statistical Inference and Modeling for High-throughput Experiments, … - look for courses by Rafael Irizarry, who is very well regarded)