Statistics

and how (not) to do it right

Motivation

& what to learn

- Is statistics just black magic?
- p-value
- Pitfall of t-test
- How to torture your data (p-hacking)
- One of the good approaches.

Ethical guidelines for the appropriate use and manipulation of scientific digital images

1. Scientific digital images are data that can be compromised by inappropriate manipulations.
2. Manipulation of digital images should only be performed on a copy of the unprocessed image data file (*Always keep the original data file safe and unchanged!*).
3. Simple adjustments to the entire image are usually acceptable.
4. Cropping an image is usually acceptable.
5. Digital images that will be compared to one another should be acquired under identical conditions, and any post-acquisition image processing should also be identical.
6. Manipulations that are specific to one area of an image and are not performed on other areas are questionable.
7. Use of software filters to improve image quality is usually not recommended for biological images.
8. Cloning or copying objects into a digital image, from other parts of the same image or from a different image, is very questionable.
9. Intensity measurements should be performed on uniformly processed image data, and the data should be calibrated to a known standard.
10. Avoid the use of lossy compression.
11. Magnification and resolution are important.
12. Be careful when changing the size (in pixels) of a digital image.

These guidelines can also be found as part of the "Online Learning Tool for Research Integrity and Image Processing", the development of this website was funded by a grant from the Office of Research Integrity. See: http://www.uab.edu/researchintegrityandimages/ or: http://ori.dhhs.gov/education/products/RIandImages/ (Retrieved 12/06/2009)

Cromey, D.W. Avoiding Twisted Pixels: Ethical Guidelines for the Appropriate Use and Manipulation of Scientific Digital Images. *Sci Eng Ethics* 16, 639–667 (2010). https://doi.org/10.1007/s11948-010-9201-y

Why are we here?

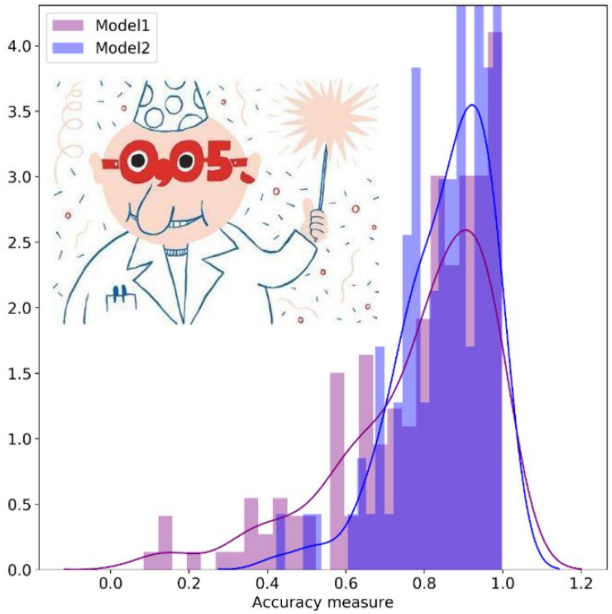
What we write

"… extremists … see the world in black and white… political moderates saw shades of grey more accurately than did either left-wing or right-wing extremists…

"… our method improves 9% the accuracy of the state-of-the-art method…

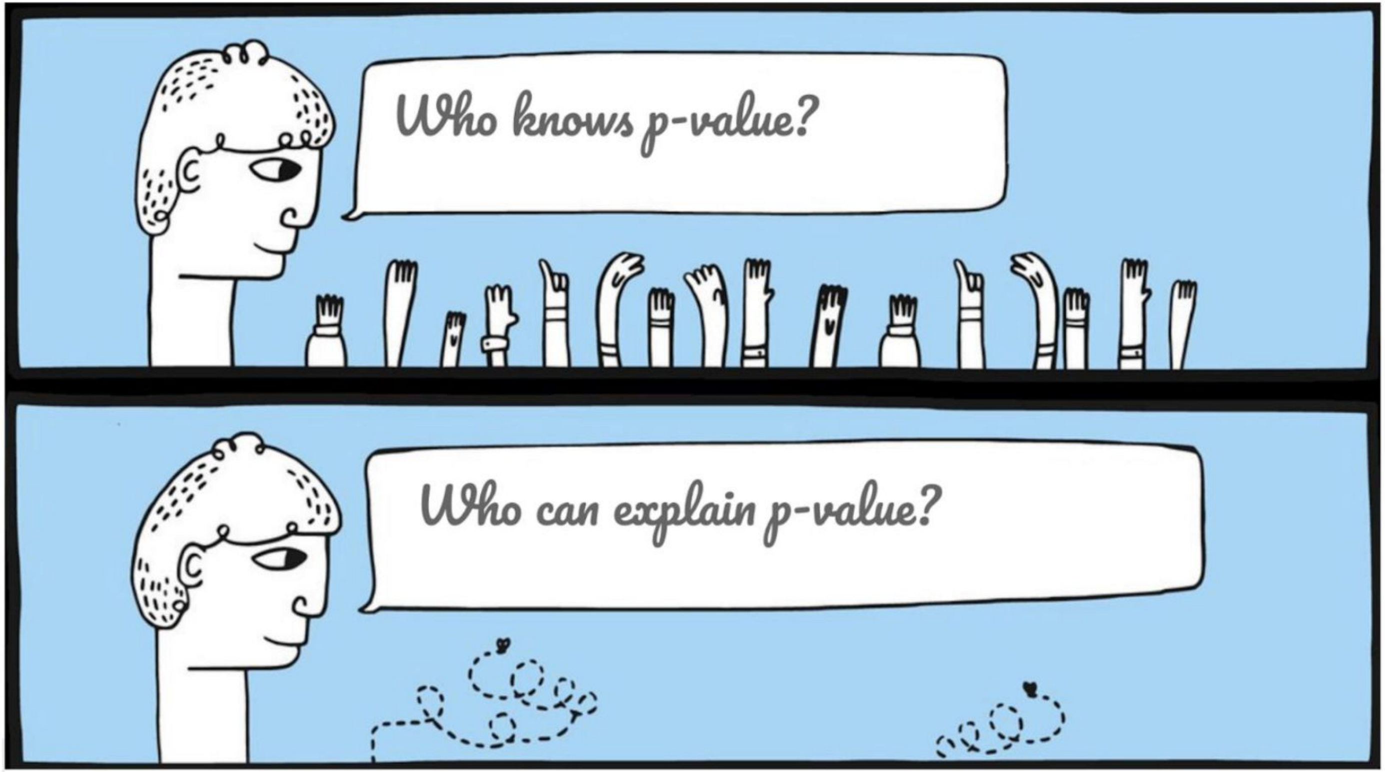
… it has an accuracy of 85% against 78%…"

p-value< 0.05 *



- Nosek, B.A., Spies, J.R.&Motyl, M. "**ScientificUtopia:II. Restructuring Incentives and Practices to Promote Truth Over Publishability**" *Perspect. Psychol.Sci.,2012*
- Nuzzo, R., "**Scientific method: Statistical errors**", *Nature, 2014*

Going back to statistics class

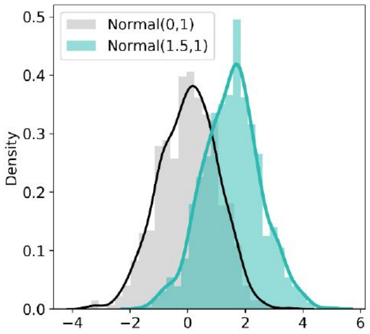


Why are we here?

Null hypothesis statistical tests

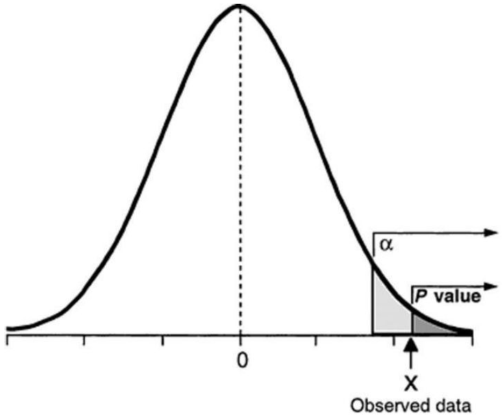
Given two normal datasets, we run a Student's t-test:

Null hypothesis H0: the means of both distributions are equal



Distribution of t under the null hypothesis

p-value: the probability that we would observe a result as extreme or more than our result IF H0 were true.



Why are we here?

Null hypothesis statistical tests

Given two normal datasets, we run a Student's t-test:

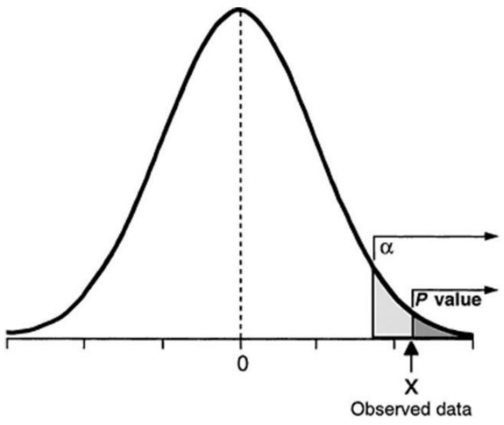
<u>Null hypothesis H0</u>: the means of both distributions are equal

Small p-value -> the result is very unlikely under H0, then there are two chances:

I. We observed a low probability event.
II. H0 is not true.

Distribution of t under the null hypothesis

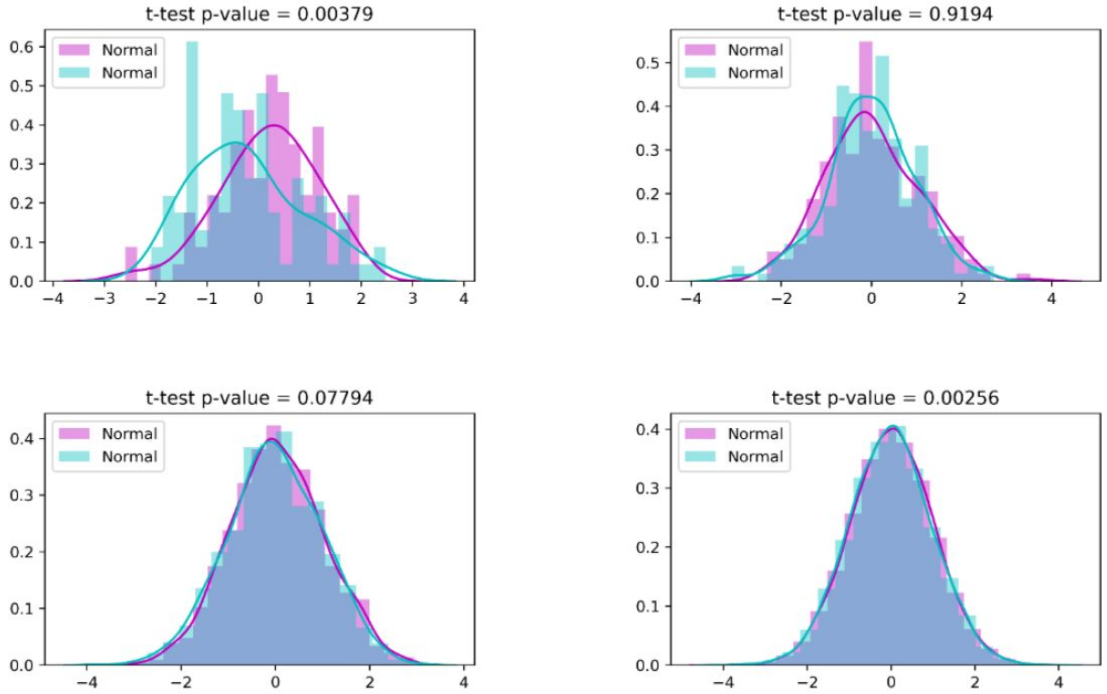
p-value:the probability that we would observe a result as extreme or more than our result IF H0 were true.



Description of the current problem

I. Let's simulate two normal distributions $N(0.01,1)$ and $N(0,1)$.

II. Compare them using Student's t-test.
H0: the means of both distributions are equal.



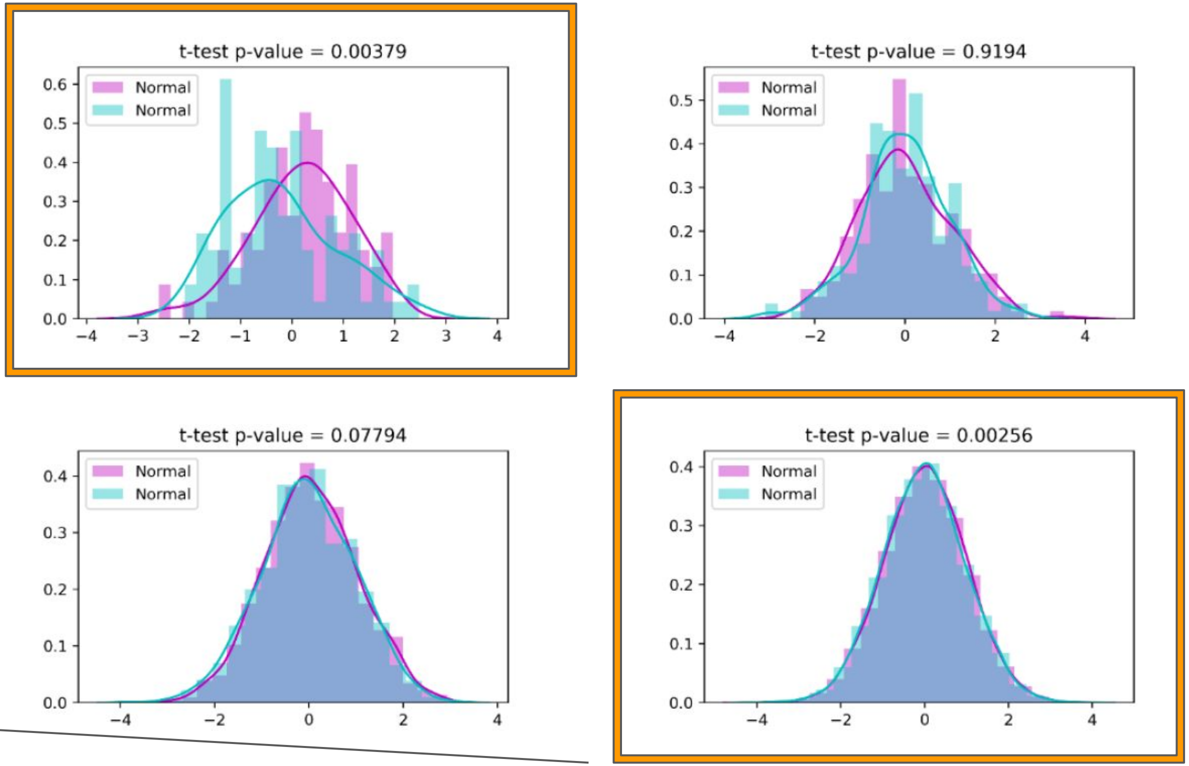
Description of the current problem

I. Let's simulate two normal distributions $N(0.01,1)$ and $N(0,1)$.

II. Compare them using Student's t-test.

H0: the means of both distributions are equal.

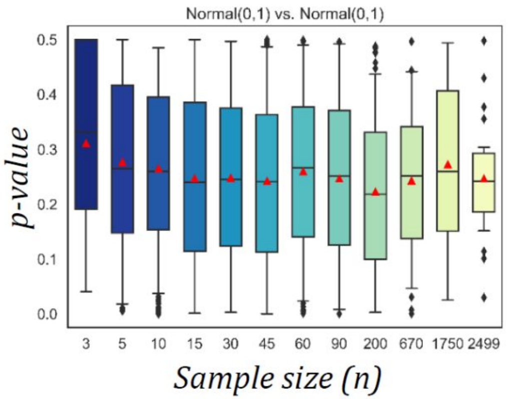
p-value< 0.05 while both distributions are almost the same.

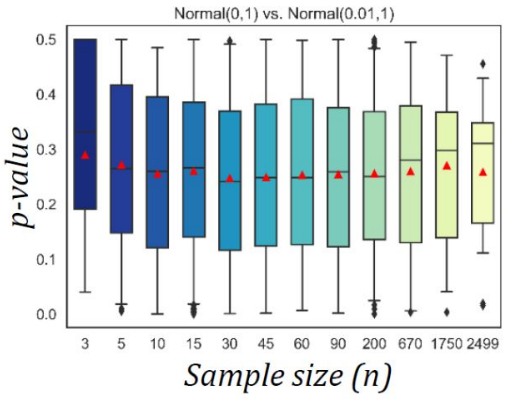


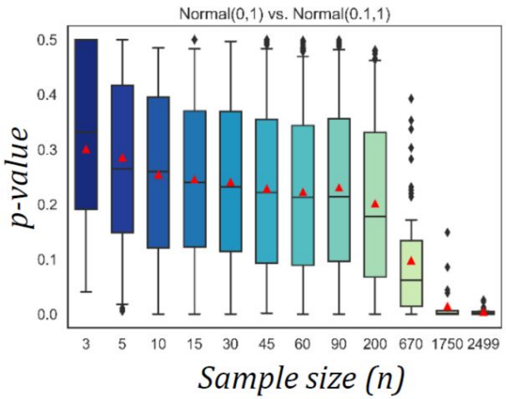
Description of the current problem

I. Let's simulate two normal distributions $\mathcal{N}(0.01,1)$ and $\mathcal{N}(0,1)$.
II. Compare them using Student's t-test. H0: the means of both distributions are equal.
III. Let's do it for different normal distributions and using different sample sizes

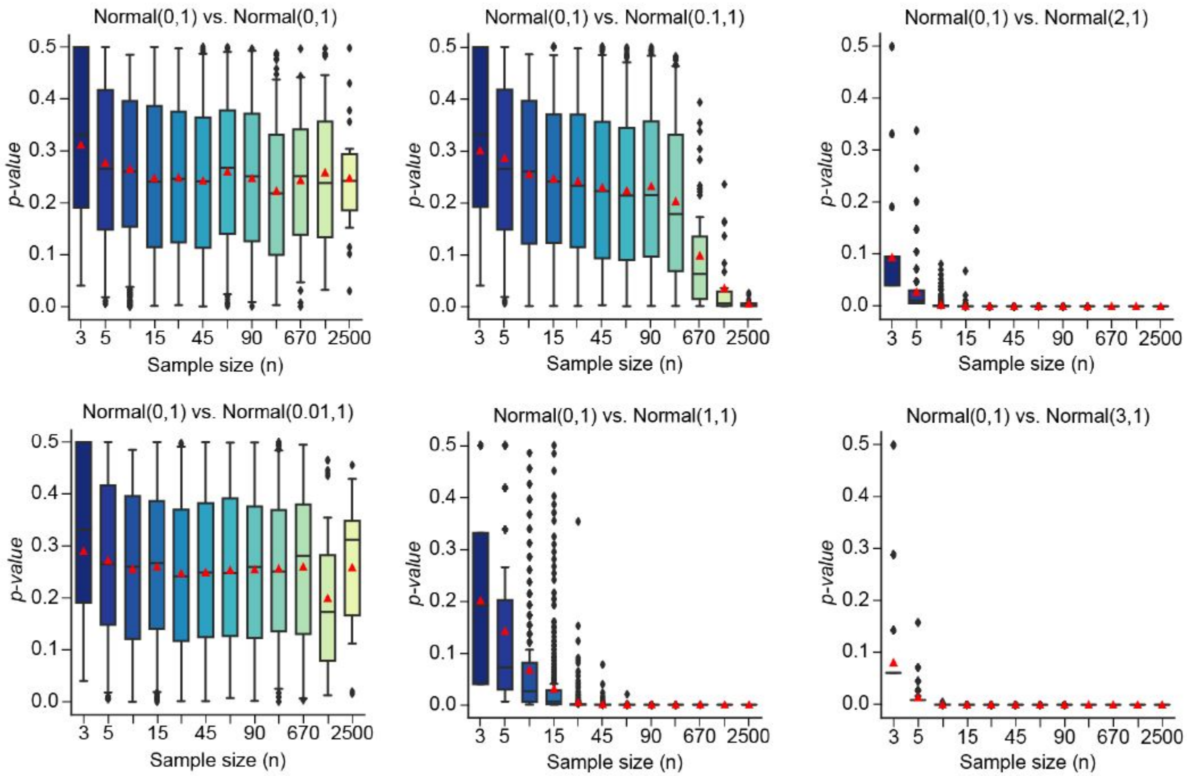
Fact: p-values follow a distribution







Description of the current problem



Description of the current problem

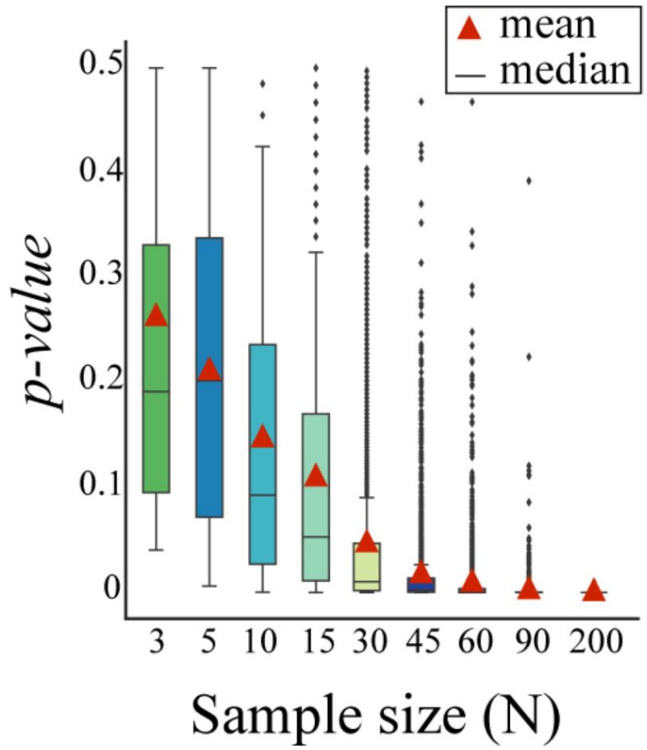
I. Let's simulate two normal distributions $N(0.01,1)$ and $N(0,1)$.
II. Compare them using Student's t-test. H0: the means of both distributions are equal.
III. Let's do it for different normal distributions and using different sample sizes

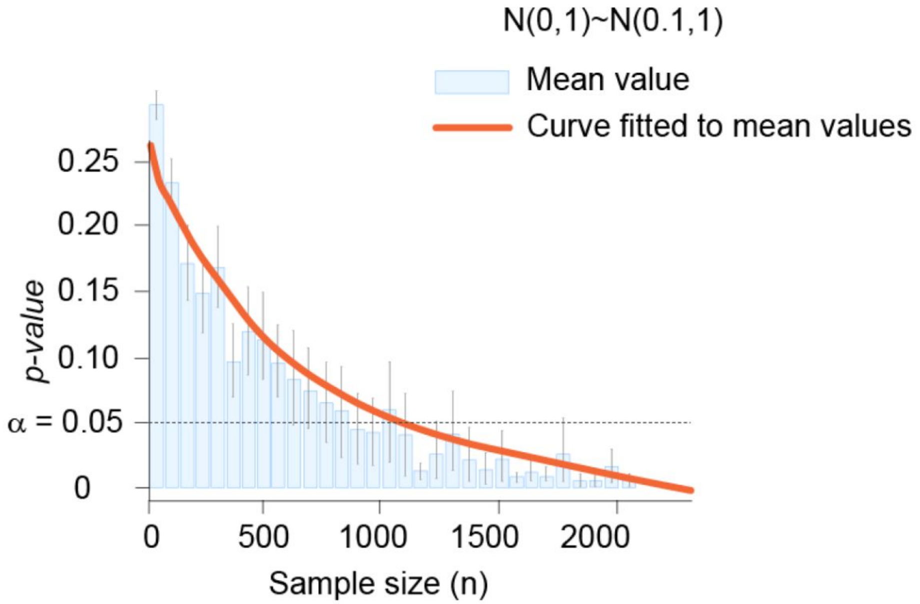
Fact: p-values follow a distribution

Problem I: In similar datasets, the distribution is uniform but not every p-value is always > 0.05

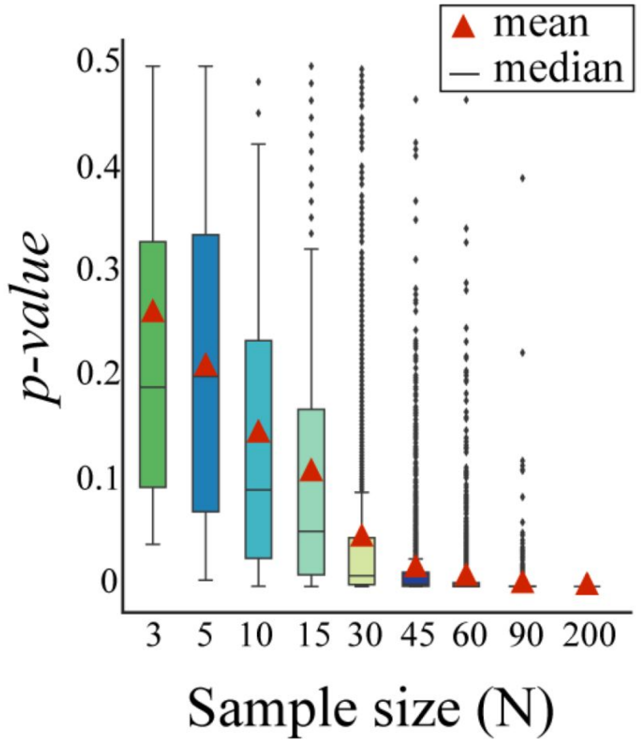
Problem II: If datasets are NOT EXACTLY the same, the p-value behaves as a function that depends on n.

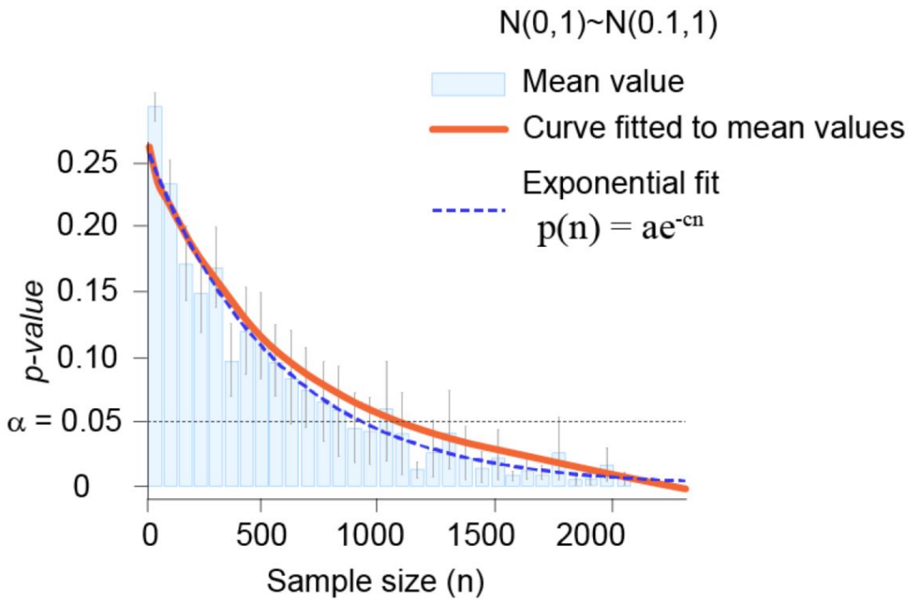
p-values change with the size of the sample



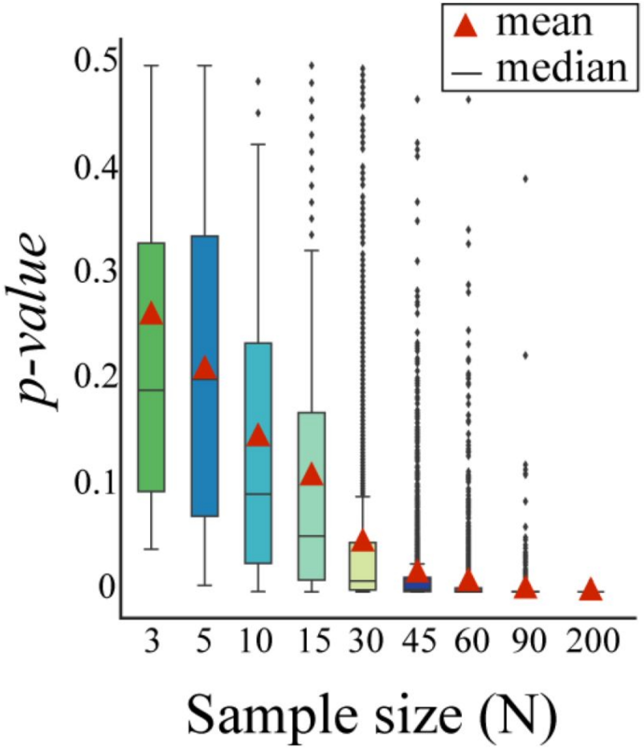


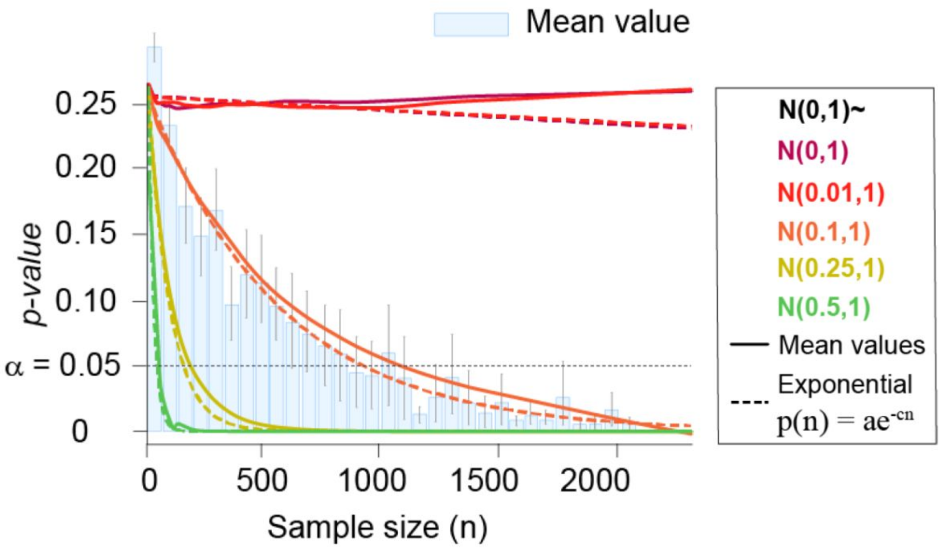
p-values change with the size of the sample



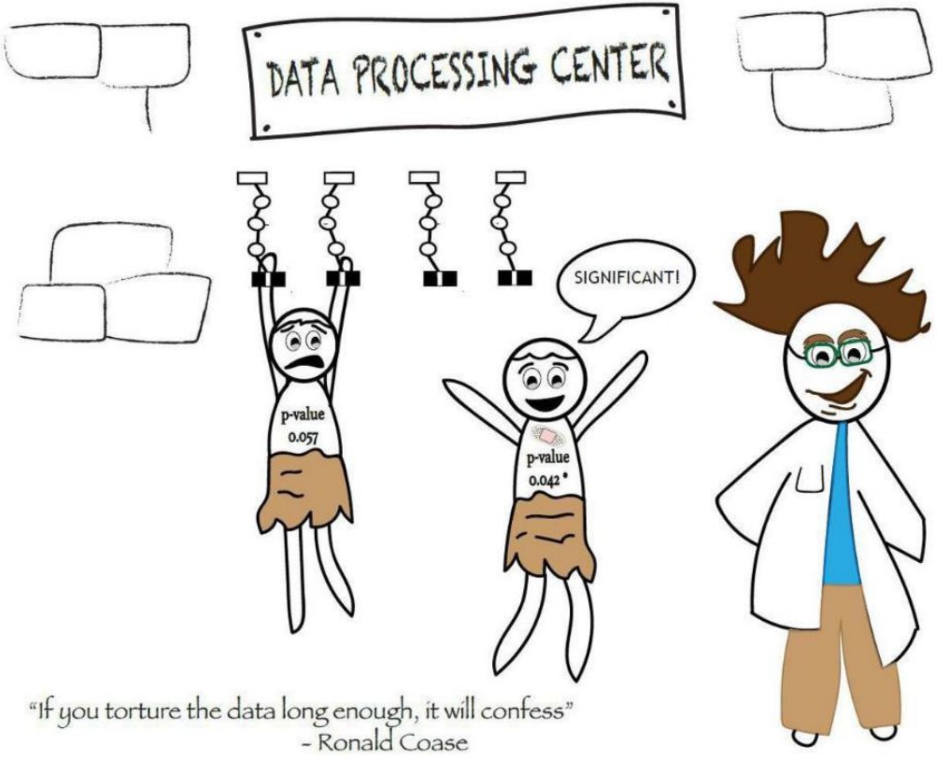


p-values change with the size of the sample





What is the connection with *p-hacking*?



What is the connection with *p-hacking*?

Under the same conditions, the p-value is different depending on the size of the sample.

It is possible to get the DESIRED p-value with a large enough dataset: p-HACKING.