Cogs 9
Discussion Section

FA22 Week 6
Will McCarthy

Upcoming due dates

Thursday, Nov 3rd Reading quiz 4 (infovis)

Friday, Nov 11th Assignment 3

This week's content: statistical significance and EDA demo

p-values and p-hacking

EDA demo

p-values

Used to quantify the *significance* of a result.

Used in the context of *null hypothesis significance testing*.

   (Which ask: is this a *result* or did nothing happen?)

Most statistical tests output a *test statistic* (eg. *t*=-21.328).

A p-value is the probability of obtaining test statistic at least as extreme as the result actually observed, assuming the null hypothesis is true.

Small p-values => null hypothesis is likely false

   => alternative hypothesis is likely true
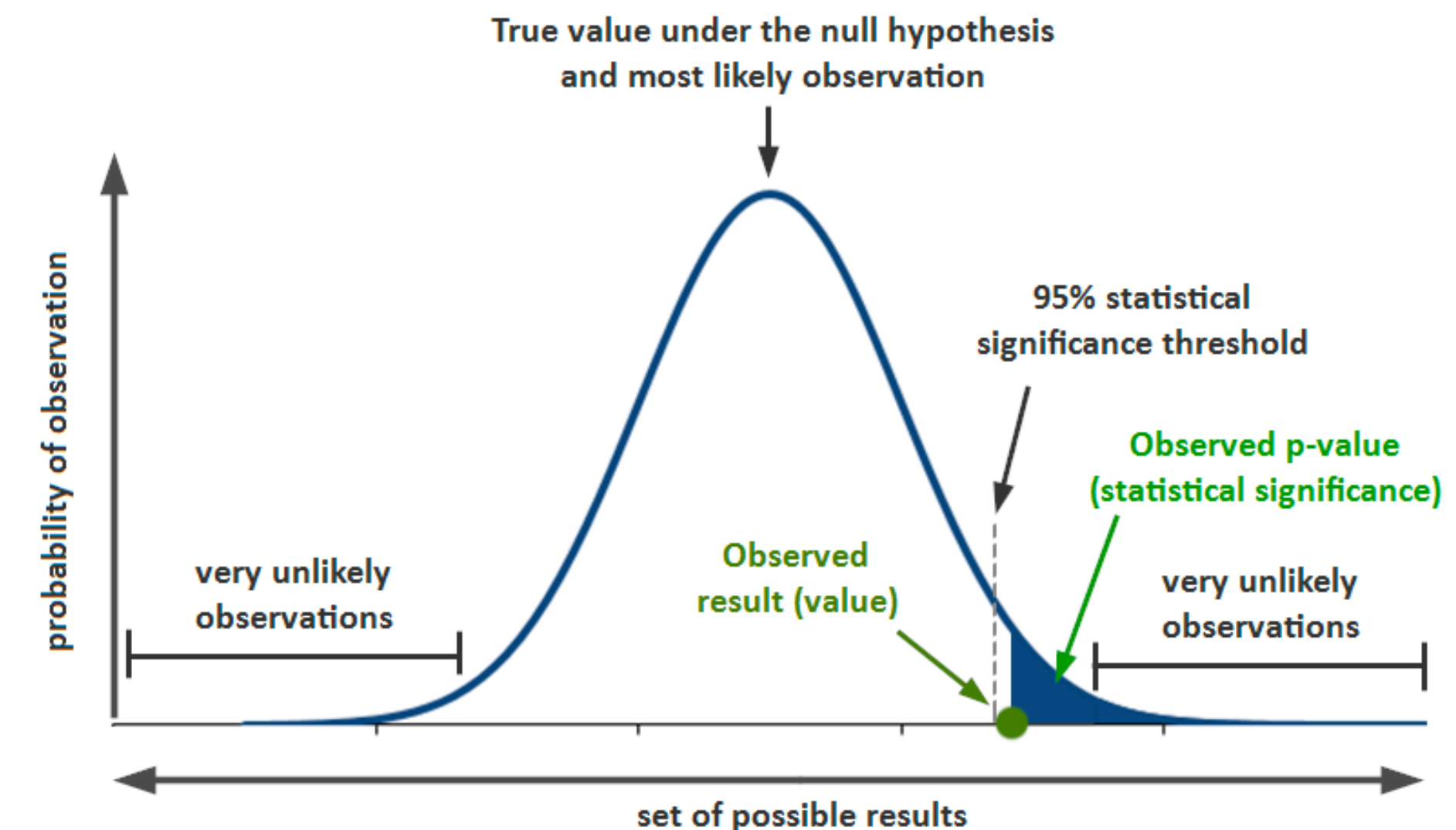
The role of p-values in science

Generally agreed upon threshold for significance is $p < 0.05$, which means:

There is a less than 5% chance of observing a result at least extreme as the one observed.

Used as a heuristic by scientists: this results deserves to be considered.

Results for which $p > 0.05$ are virtually ignored (not publishable on their own)
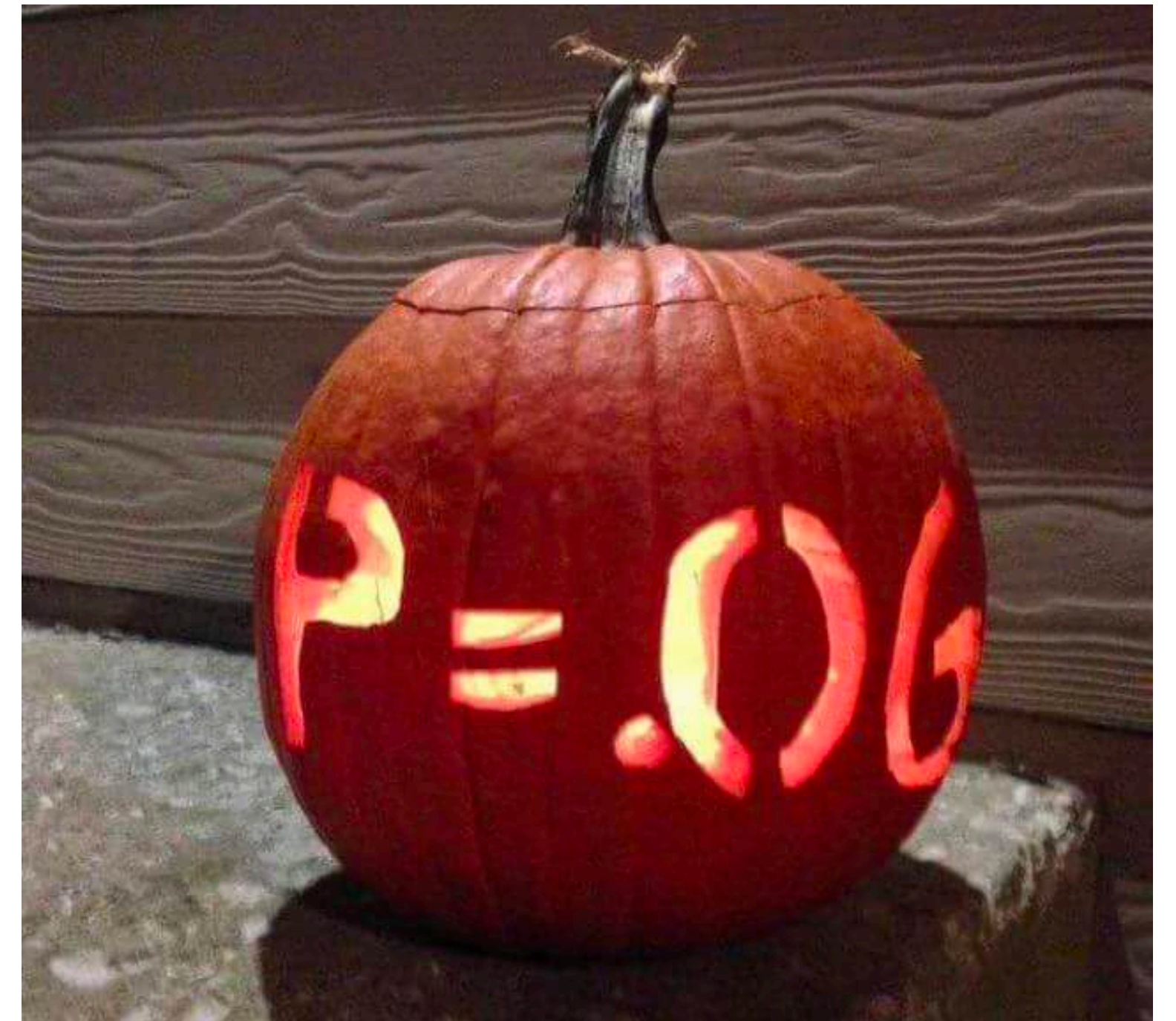
p-hacking

P-values < 0.05 are valuable…

… so people cheat

P-hacking: messing around with data until statistical significance ($p < 0.05$) is reached

You can cheat more or less explicitly

- Stop collecting data once $p < 0.05$

- Report only measure with $p < 0.05$

- Exclude participants/ data to reach $p < 0.05$

- …

There are ways to prevent p-hacking: e.g. pre-registration

Assignment 3

In A3 you'll do some p-hacking in a website designed to show you how easy it is.

You will learn:

- how easy it is to p-hack

- how tempting it is to p-hack

- how 'preregistering' your analyses avoids p-hacking

Before you play with the website, write a specific answer for question 1.

Demo: EDA

Group work / questions

Future Readings

5(a): Nicholas Diakopoulos, 2016, Accountability in Algorithmic Decision Making

5(b): Julia Angwin, et al., 2016, Machine Bias