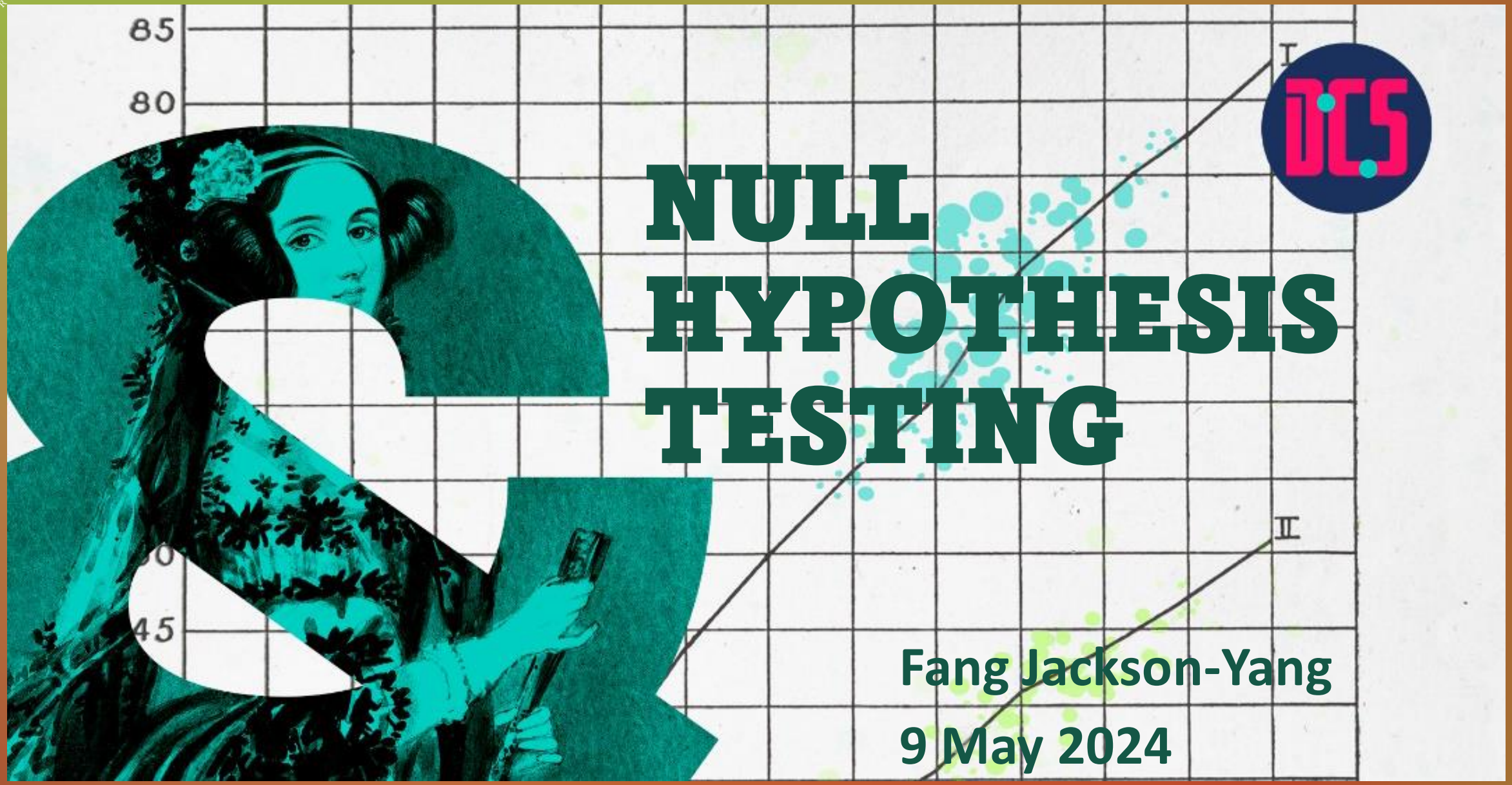
THE UNIVERSITY of EDINBURGH
Centre for Data, Culture & Society



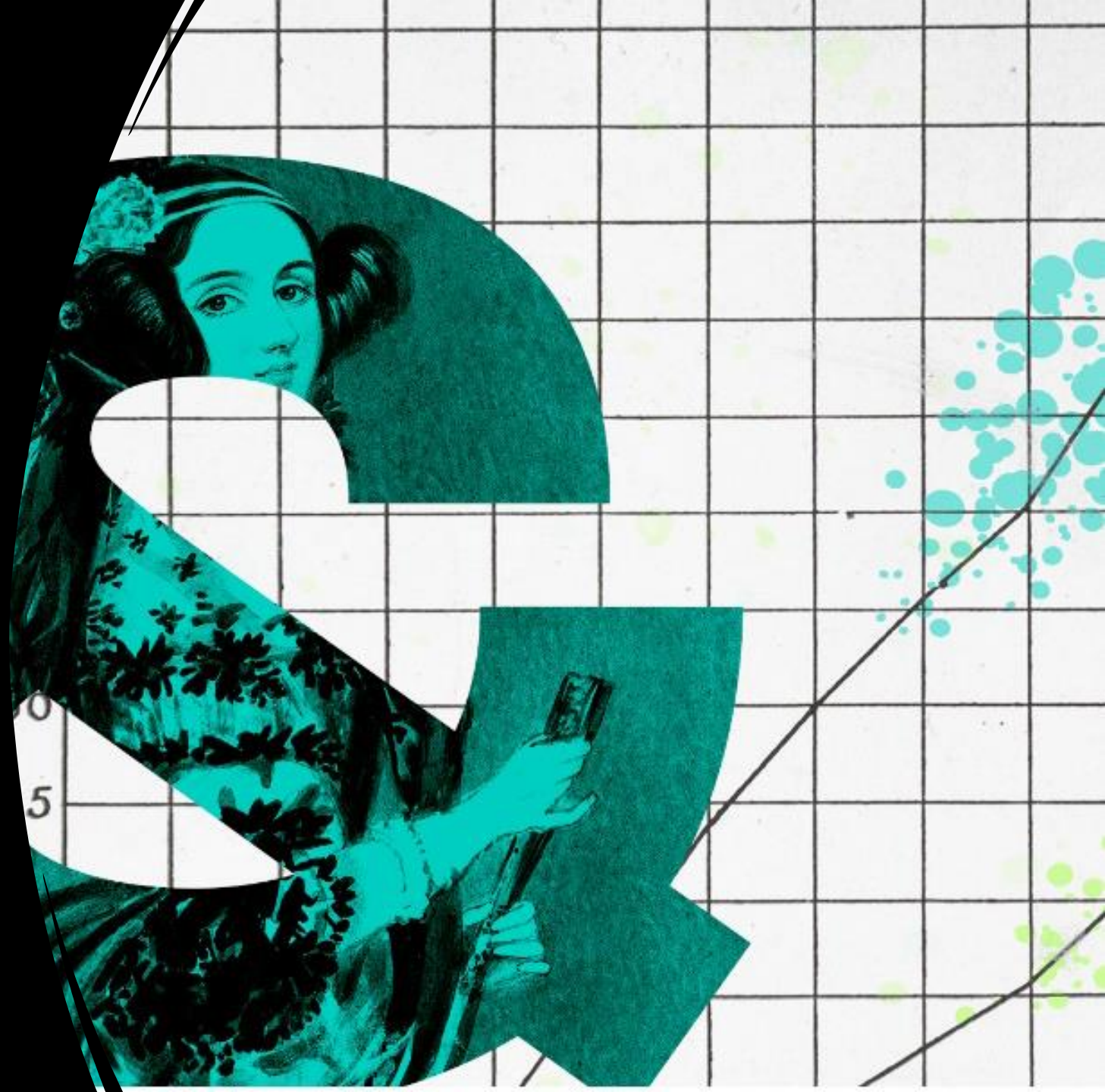
NULL HYPOTHESIS TESTING

Fang Jackson-Yang

9 May 2024

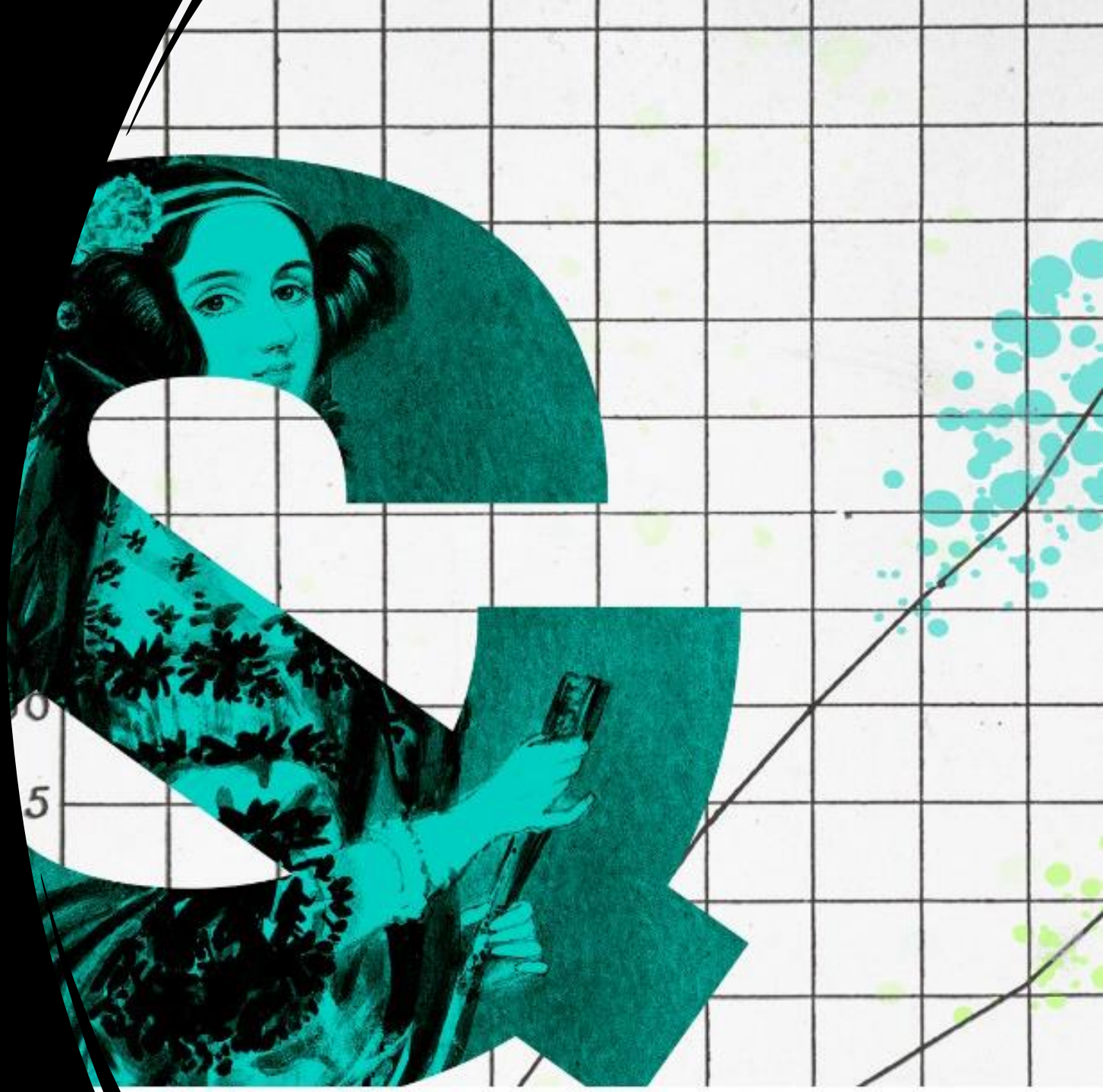
Course outline

- Session 1. Concepts & two-sample t-test
- Session 2. Power, effect size & other example tests for NHT



Session 1 Roadmap (today)

- Quantitative Research
- Null Hypothesis Testing (NHT)
- P-value & Significance
- Errors in Decision Making



Quantitative Research

Theory Testing

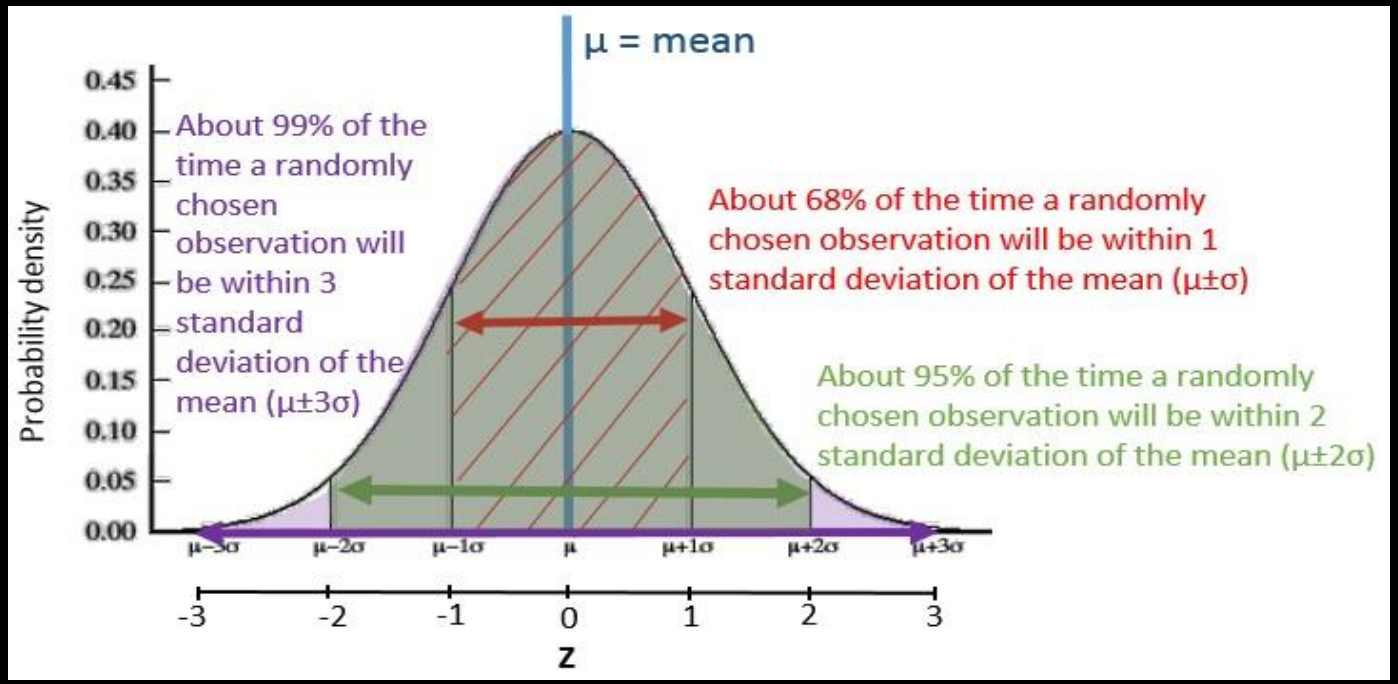
- **Testing** a claim about a population parameter, population mean μ
- Unfeasible to collect data for the full population
- Collected for a random sample of size n

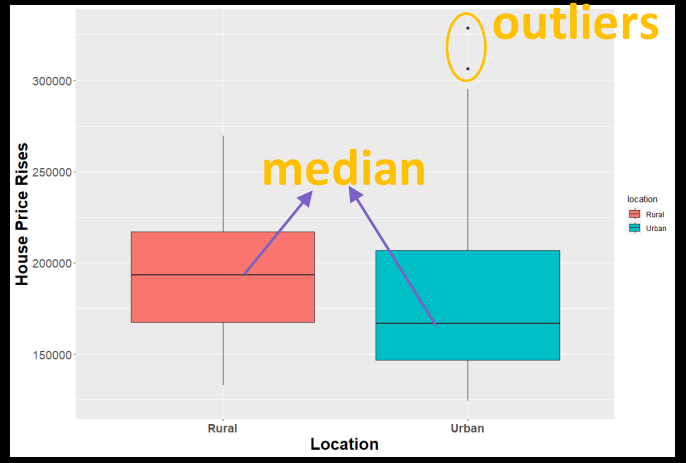
Then:

- obtain an estimate for the population mean (the **Sample Mean μ**)
- obtain a measure of precision of our estimate (the **Standard Error SE**)
- report the estimate along with the precision (the **Confidence Interval CI**)

Descriptive Statistics

- Distribution & central tendency
- Mean, median, standard deviation (sd)





Source: http://dev1.ed-projects.nyu.edu/statistics/normal-distribution/

Inferential Statistics

- Collect relevant sample data
- Perform a hypothesis test
- Report the test results correctly
- Interpret the results

Null Hypothesis Testing

Research Hypothesis

- Hypotheses are claims researchers make about the world based on some data.

- Null Hypothesis ($H0$)
  - a skeptical claim that nothing is different / nothing is happening
  - "no effect", "no difference", "no change", "no relationship"

- Alternative Hypothesis ($H1$)
  - hypothesis of your research interest
  - "Variable x has an effect on variable y", "group A differs from group B", etc

Goal: Test against H0; make inference about H1

Null Hypothesis Testing (NHT)

- A formal method for validating hypotheses
- Comparing a theoretically non-null hypothesis ($H1$) with a null hypothesis ($H0$).
- Examples:

$H0$: **No** panda has colours other than black & white

$H1$: Some pandas are not black & white

Null Hypothesis Testing (NHT)

- More examples:
  - $H0$: Improved performance in SQA exams not associated with participation in extra curriculum activities.
  - $H1$: Improved performance in SQA exams associated with participation in extra curriculum activities.

  - $H0$: No difference between university students and full-time employees on time spent on Instagram daily
  - $H1$: University students spend more time than full-time employees on Instagram daily

Null Hypothesis Testing (NHT)

- More examples from you

What is the Alternative Hypothesis $H1$ of your own research ?

What is the Null $H0$?

Two Sample T-test

- $H0$: No difference between university students and full-time employees on time spent on Instagram daily
- $H1$: University students spend more time than full-time employees on Instagram daily

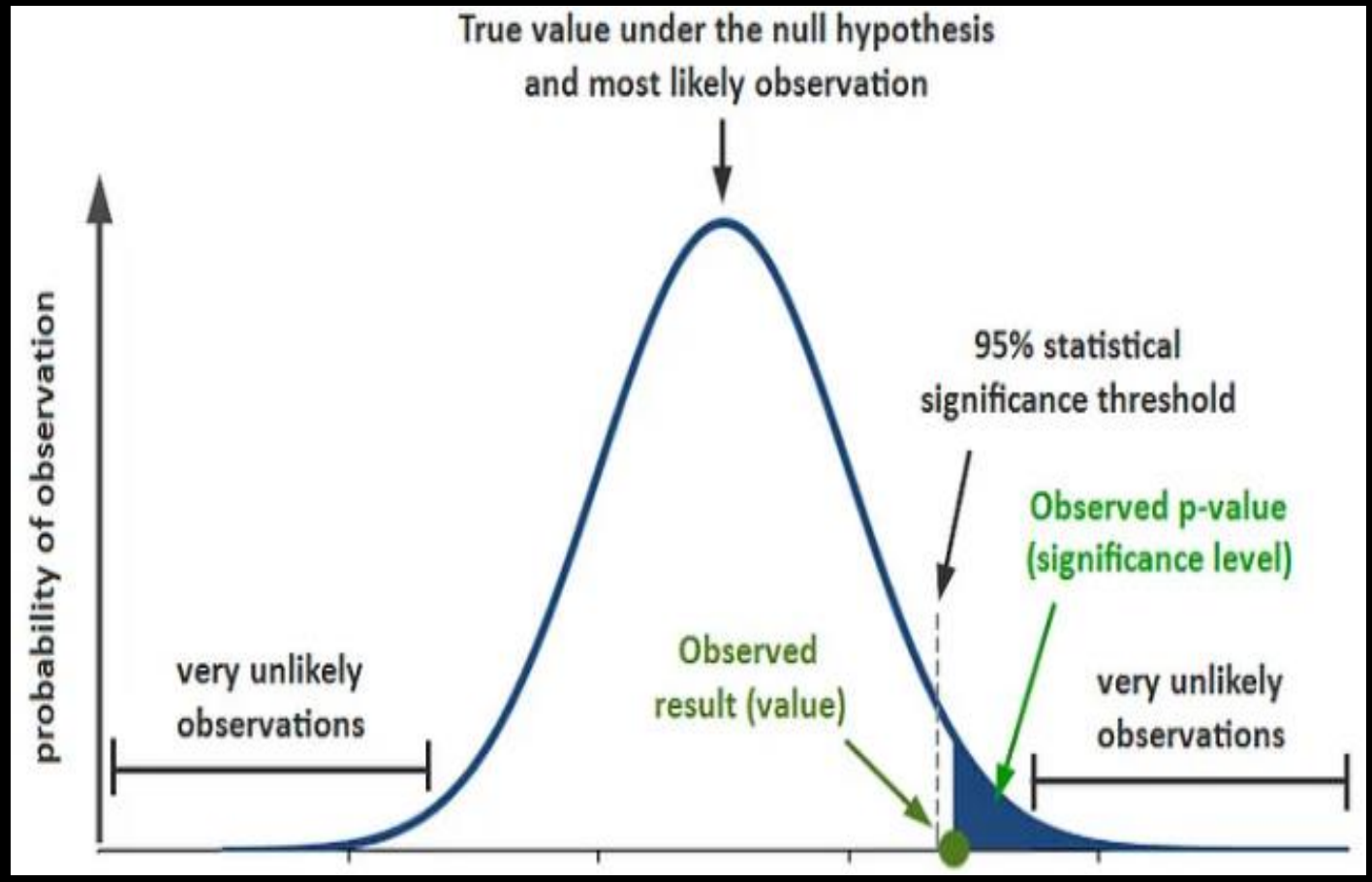
- $H0$: $\mu_{\text{uni\_student}} = \mu_{FT\_employee}$
- $H1$: $\mu_{\text{uni\_student}} > \mu_{FT\_employee}$

P-value and Significance

P-value

- The probability that measures the strength of the evidence against a null hypothesis.

- Assuming that $H0$ is true, the probability of obtaining a value of the t-statistic at least as extreme as that observed.

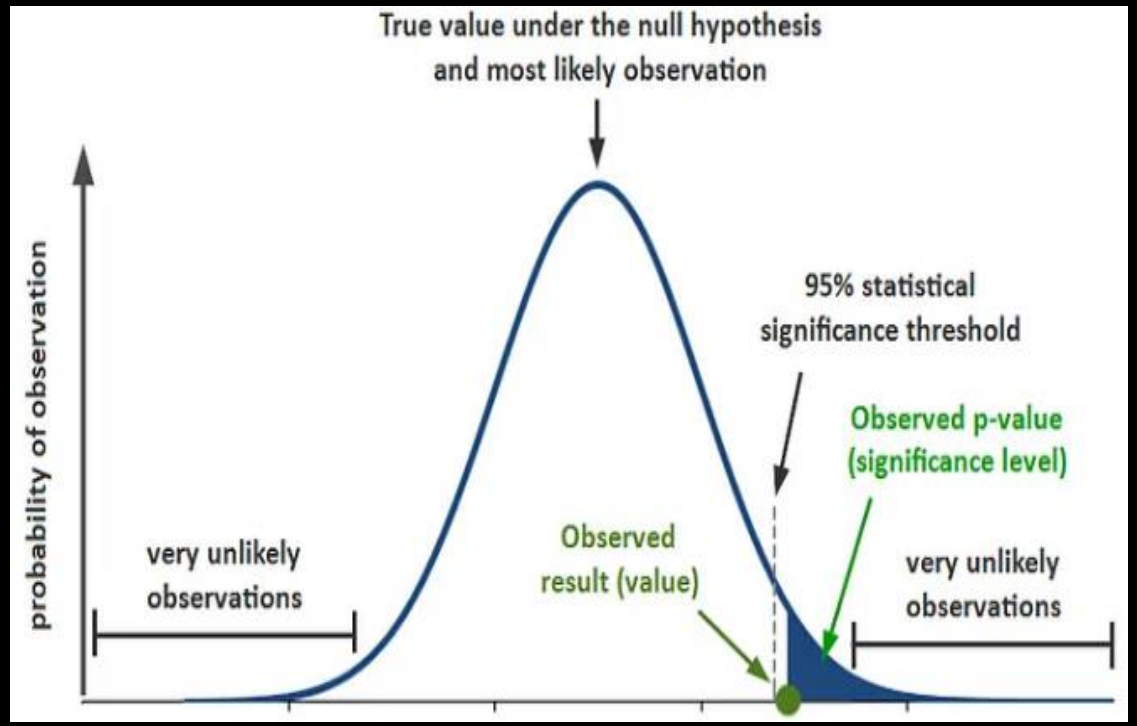


Source: https://medium.com/datadriveninvestor/p-value-t-test-chi-square-test-anova-when-to-use-which-strategy-32907734aa0e; https://www.sagepub.com/sites/default/files/upm-binaries/40007_Chapter8.pdf; https://learningstatisticswithr.com/lsr-0.6.pdf

P-value

- The **smaller** the p-value, the stronger the evidence against $H0$.
- Large p-values fail to provide sufficient evidence against $H0$, suggesting your $H1$ is likely to be untrue.

How small is small?



Significance Level (α)

- Chosen by you, the researcher
- Depends on the field
- Most commonly : α = 0.05, α = 0.01

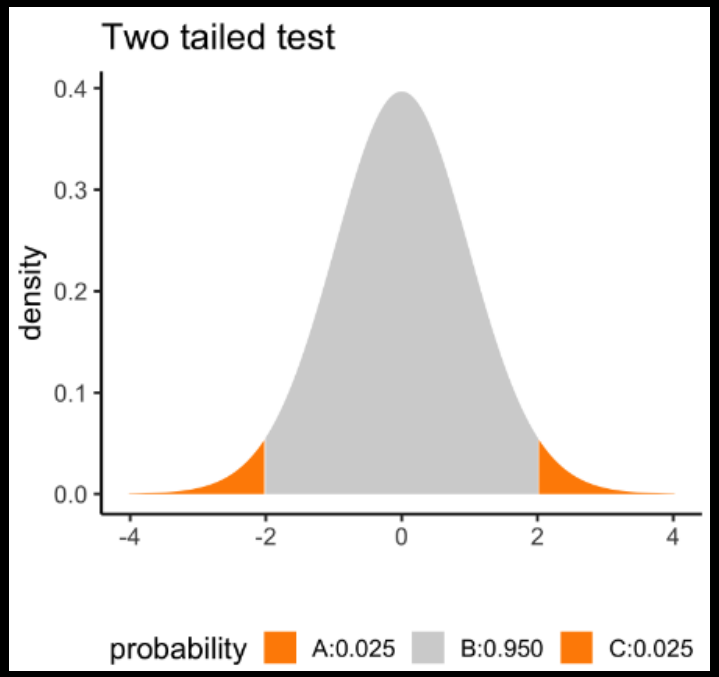
- $p \leq .05$: Given $H0$ is true, the chance to observe the data is no more than 1 in 20 (5%)

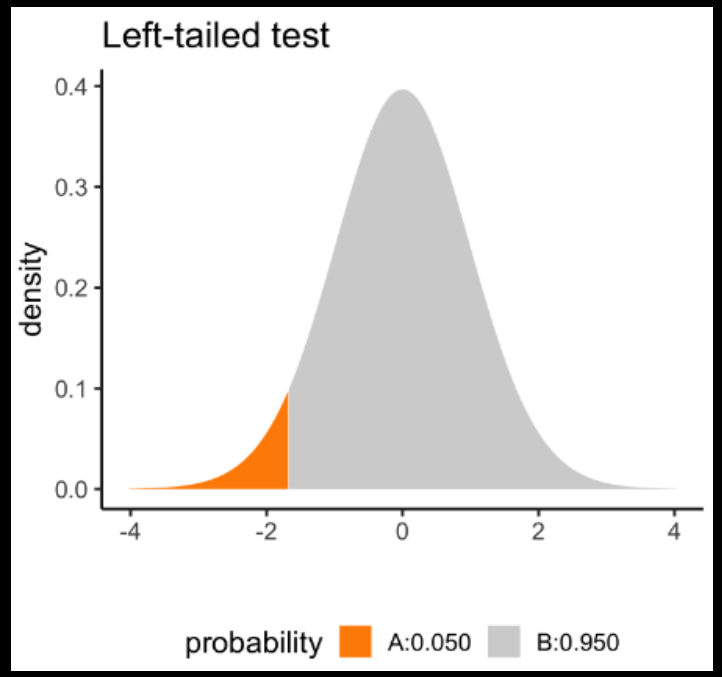
- $p \leq .01$: Given $H0$ is true, the chance to observe the data is no more than 1 in 100 (1%)

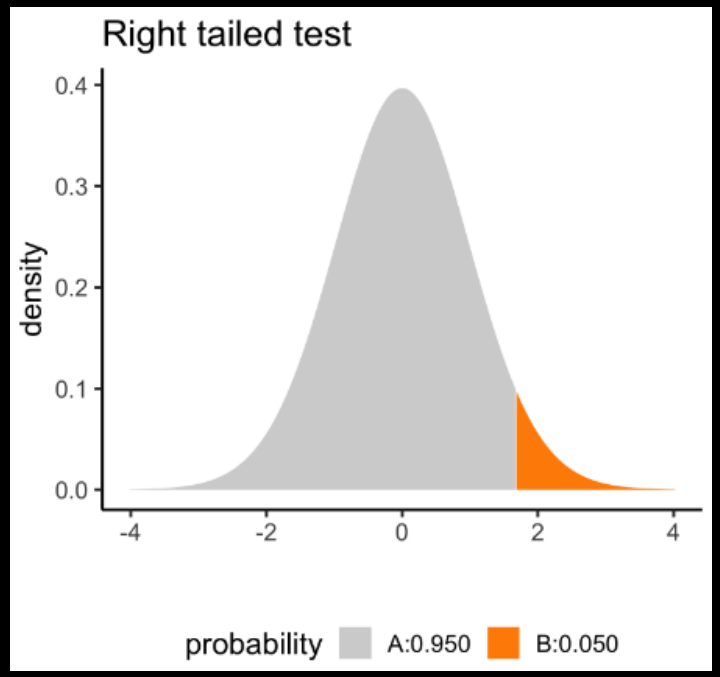
Significance

$H0$: $\mu_{\text{uni\_student}} = \mu_{FT\_employee}$

- $p \leq .05$







$H1$: $\mu_{\text{uni\_student}} \neq \mu_{FT\_employee}$

$H1$: $\mu_{\text{uni\_student}} < \mu_{FT\_employee}$

$H1$: $\mu_{\text{uni\_student}} > \mu_{FT\_employee}$

Errors in Decision Making

Two Types of Errors

- Type 1 Error : Reject a true null hypothesis
- Type 2 Error : Fail to reject a false null hypothesis

		Your Decision	
		Retain H0	**Reject H0**
$H0$ **Truth**	True	1 - α = probability of correct retention	**α (Type I error rate)**
	False	**β (Type II error rate)**	1 - β = **power** of the test

Significance vs Importance

- Errors in decision making
- Power (Probability of rejecting $H0$ given that $H0$ is false)
- Effect size ($Cohen's\ D$)

Statistical Significance $\neq$ Practical Importance

Assumption Check

For t-test:

- **Independence**: Randomly obtained sample
- **Normality**: The sampling distribution of the sample mean is normally distributed, or sample size sufficiently large ($n > 30$)

THANK YOU

THE UNIVERSITY of EDINBURGH
Centre for Data, Culture & Society