WEEK 8

INTRODUCTION TO HYPOTHESIS TESTING

BASIC CONCEPTS

- Hypothesis testing is a process that is concerned with making inferences about the population using the information obtained from a sample

- Because we take a sample, there is an element of uncertainty involved and, therefore, we should accompany the conclusions we draw about the population with a **probability!!**
  - This gives an indication of the chance of getting the observed results if the hypothesis is true.

THE NULL HYPOTHESIS, H0

- H_0 = Null hypothesis ⟷ No difference in means (or no treatment effect)
- H_1 or H_A = Alternative hypothesis ⟷ A difference exists in means (or there is a treatment effect)

➢ In statistical hypothesis testing, we attempt to disprove the null hypothesis

TEST STATISTICS AND THE P VALUE

Research Question

Search Literature

Logical?

Original?

Reliability

Specify the hypothesis

Define the experimental approach

We need to measure **Variable** → *a characteristic that can take values which vary from individual or group to group*

Collect sample data

Calculate test statistics
(an algebraic expression particular to hypothesis we are testing)

Calculate the Probability
(P- value)

Cause & **Effect / Outcome**

Independent variable (Predictor)

Dependent variable (Outcome)

Logic behind the test statistics

Systematic variation

- Variation that can be explained by the model

Unsystematic variation

- Variation that can not be explained by the fitted model

$$test\ statistics = \frac{amount\ of\ variance\ explained\ by\ the\ model}{amount\ of\ variance\ not\ explained\ by\ the\ model} = \frac{effect}{error}$$

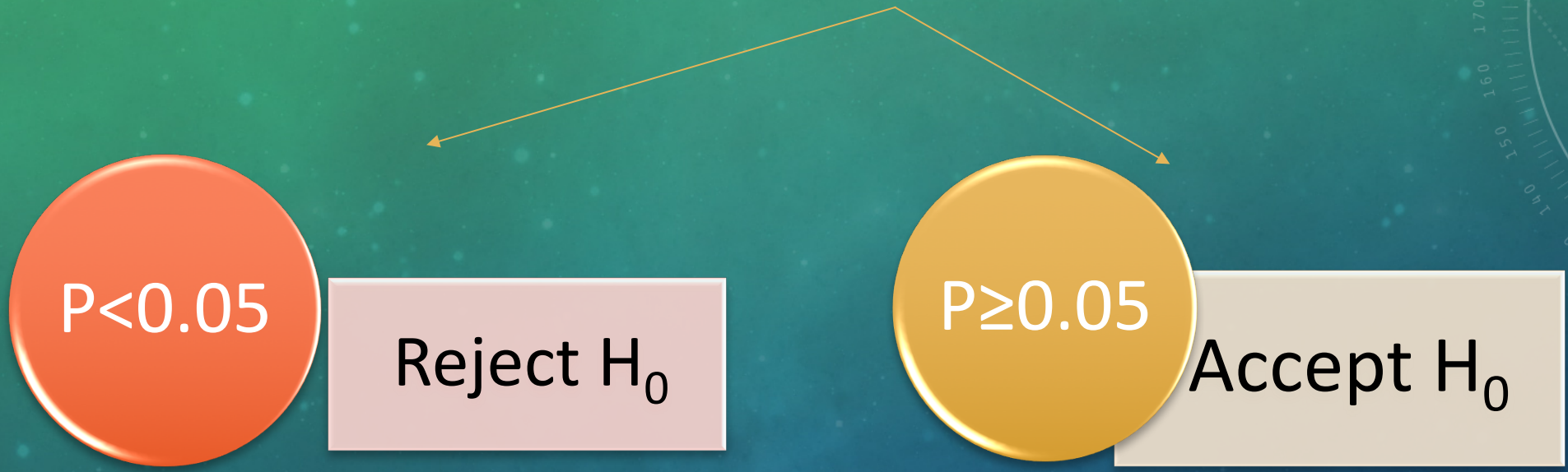
THE P VALUE

- According to the evidence obtained from our sample, we make a judgement about whether the data are inconsistent with the null hypothesis; this leads to a **decision** whether or not to reject the null hypothesis.

- P value describes how well the sample data support the argument that *the null hypothesis is true*.
  - It allows us to determine whether we have enough evidence to reject the null hypothesis in favour of the alternative hypothesis.

In technical terms, a P value is the probability of obtaining an effect at least as extreme as the one in your sample data, assuming the truth of the null hypothesis

Cut off value for P Value



If the *P*-value is small (<0.05), then it is unlikely that we could have obtained the observed results if the null hypothesis were true, so we reject *H*0.

If the *P*-value is large (≥0.05), then there is a high chance that we could have obtained the observed results if the null hypothesis were true, and we do not reject *H*0.

DEGREES OF FREEDOM

- The degrees of freedom of a statistic are the number of independent observations contributing to that statistic, i.e. the number of observations available to evaluate that statistic minus the number of restrictions on those observations.

- It relates to the number of observations that are free to vary.

MAKING THE WRONG DECISION: TYPE I AND TYPE II ERRORS

- The final decision whether or not to reject the null hypothesis may be incorrect.

The **probability of making a Type I error** is the probability of incorrectly rejecting the null hypothesis

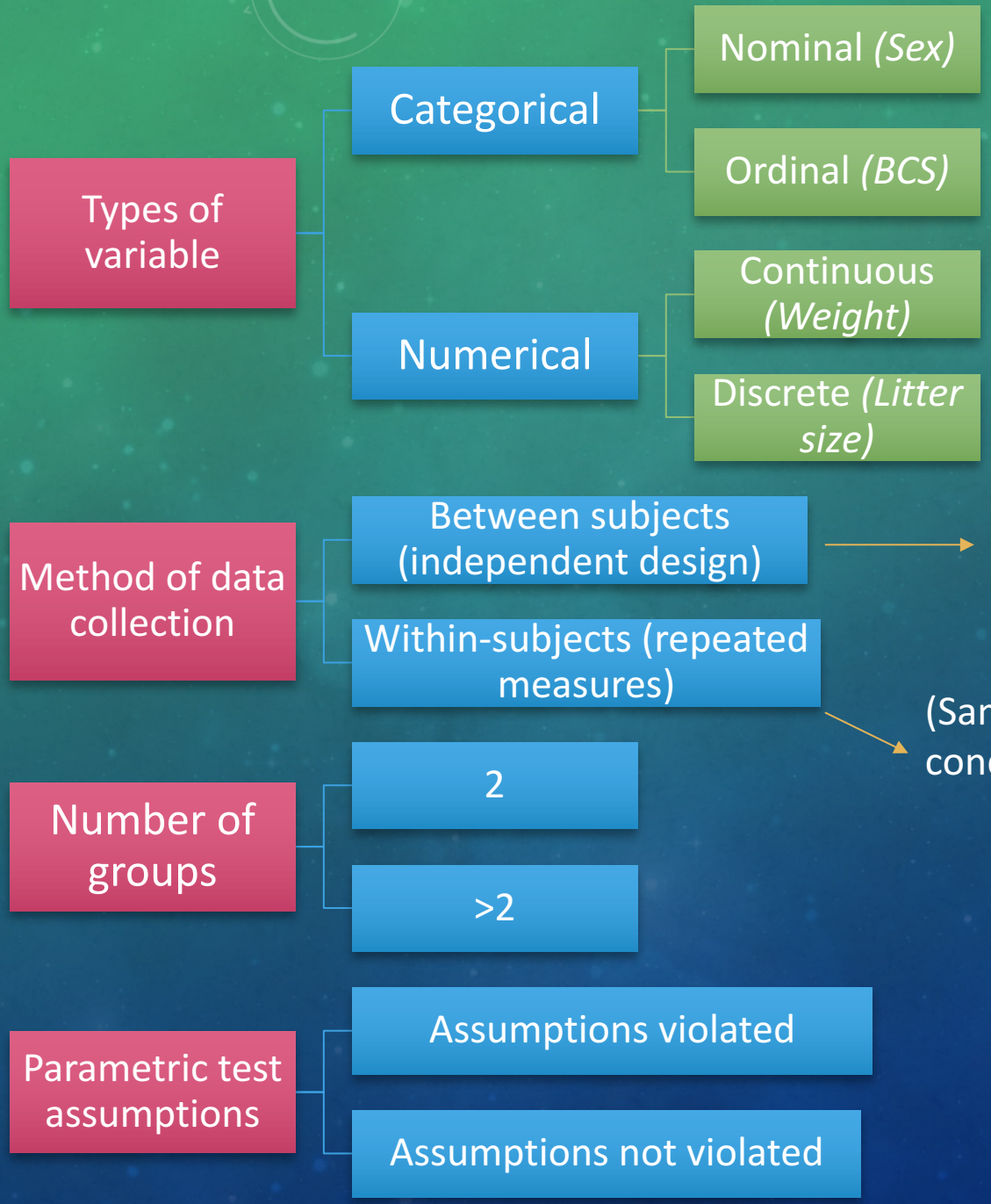
	Reject H_0	**Do not reject H_0**
H_0 True	Type 1 error (α)	Correct decision
H_0 False	Correct decision (1-β) (*Power*)	Type 2 error (β)

Power is the probability of rejecting the null hypothesis when the null hypothesis is false.

It represents the ability of the test to detect a real effect

The **probability of making a Type II error is** the probability of not rejecting the null hypothesis when the null hypothesis is false

SELECTING THE APPROPRIATE HYPOTHESIS TEST



STEPS FOR ALL HYPOTHESIS TESTS…

Step 1

- Specify the null and alternative hypothesis

Step 2

- Calculate the test statistics (an algebraic expression particular to hypothesis we are testing)

Step 3

- Obtain P- value by referring the calculated value of test statistic

Step 4

- Make a decision whether or not to reject the null hypothesis

Step 5

- Interpret the result