MS4S08 – Applied Statistics for Data Science

Introduction to Hypothesis Testing

Dr Penny Holborn
penny.holborn@southwales.ac.uk

University of South Wales
Prifysgol De Cymru

Contents of this lecture

- Introduction to hypothesis testing
- Principle ideas of the technique
- Example
- Resources

Learning objectives

At the end of this tutorial you will be able to:

- Demonstrate an understanding of principle ideas of hypothesis testing
- Apply your new skills to an example

Introduction

- Scientific claims about reality are called *hypotheses*

- Any hypothesis should be supported by evidence

- If they are verified by observations, hypotheses are accepted as true, at least until proved otherwise

Coin toss

- We are given a coin and asked to determine whether or not it is “fair”.

- Therefore we accept that it is fair.
  - We call this the null hypothesis, H_0 or H_N

- We attempt to provide enough evidence to reject this and accept an alternative.
  - We call this the alternative hypothesis, H_1 or H_A

Coin toss

- Let's say we toss the coin 10 times and a total of 9 heads are observed.
- Now, if H_0 is true, getting 9 heads from 10 seems unlikely.
- However, we need to determine if the probability of this occurring by chance is low enough to allow us to reject the null hypothesis and ultimately conclude that the coin is not fair.

One-tailed or two-tailed test

- Tossing a coin can be modelled as a binomial variable (set of trials with 2 outcomes)
  - Toss the coin n times
  - Proportion p of heads

- The hypothesis that a coin is fair is to test the hypothesis that $\mathrm{p} = {}^{1}/_{2}$, $H_0\colon \mathrm{p} = {}^{1}/_{2}$

- For a one-tailed test, $H_1\colon p < {}^{1}/_{2}$ or $H_1\colon p > {}^{1}/_{2}$
- For a two tailed test, $H_1\colon p \neq {}^{1}/_{2}$

Choose a significance value

- First we need to decide a significance level for the test. This is denoted by α (alpha).
- This is the probability of rejecting the null hypothesis when it is actually true.
- So for our coin toss that would mean:
  - Determining the coin was not fair when it actually is.
- In general, this is usually around 5%.

University of South Wales
Prifysgol De Cymru

Type I and Type II errors

- Type I errors – False positives
  - If H_0 is rejected when it is actually true

- Type II errors – False negatives
  - If H_0 is accepted when it was actually false

Decision	**H_0 is true**	**H_0 is false**
Reject H_0	Type I error (α)	Correct (1-β)
Do not reject H_0	Correct (1-α)	Type II error (β)

Type I errors

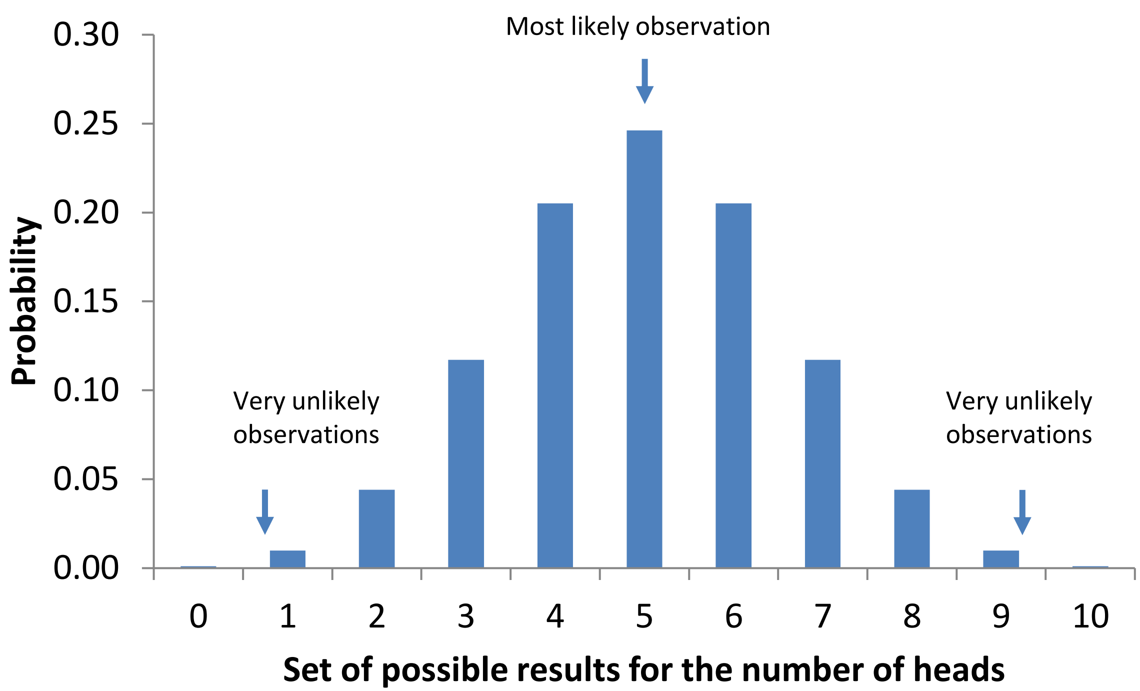
- The probability of committing a Type I error is denoted by α (the significance level).
- If we set the significance level at $\alpha = 0.05$ there is a 5% probability of rejecting H_0, when H_0 is actually correct.

Type II error

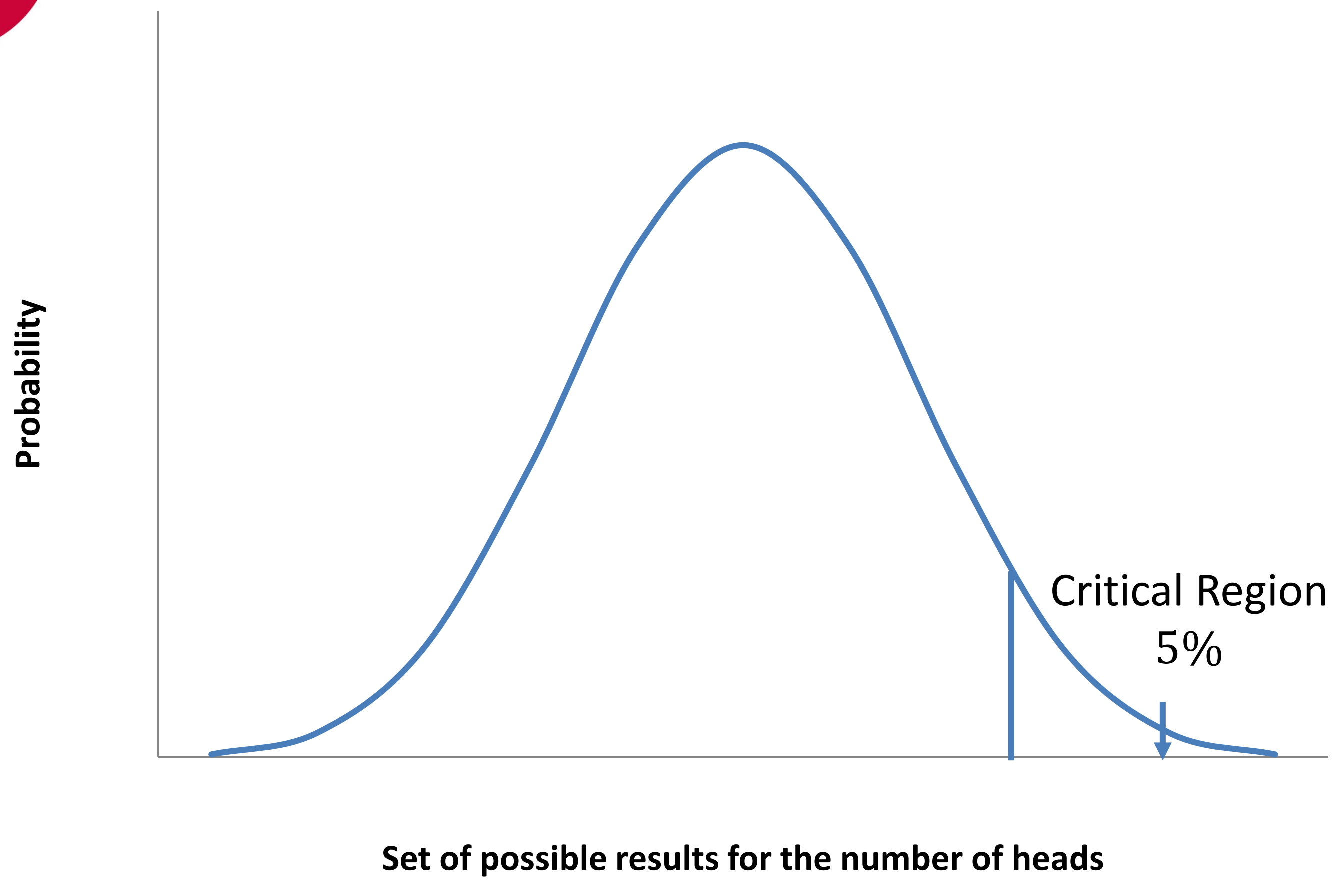
- The probability of committing a Type II error is denoted by β
- $1 - \beta$ *is known as the statistical power of the test*

- The probability of committing a Type II error depends on several factors, including the sample size

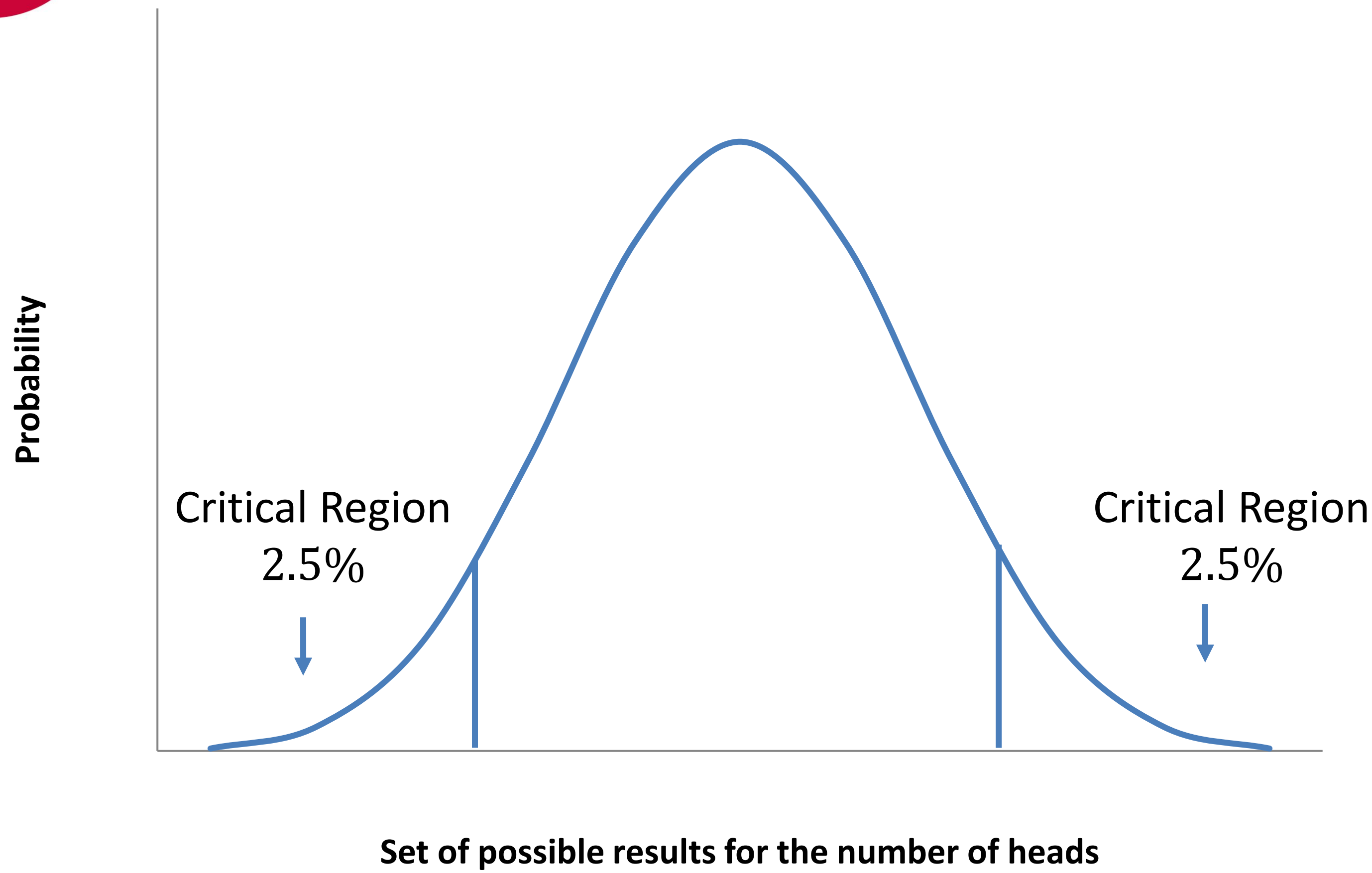
Probability mass function



$$H_1 : \mathrm{p} > {}^{1}/_{2} \ , \ \alpha = 0.05$$



$$H_1\colon \mathrm{p} \neq {}^{1}/_{2}\ ,\ \alpha = 0.05$$



p-value

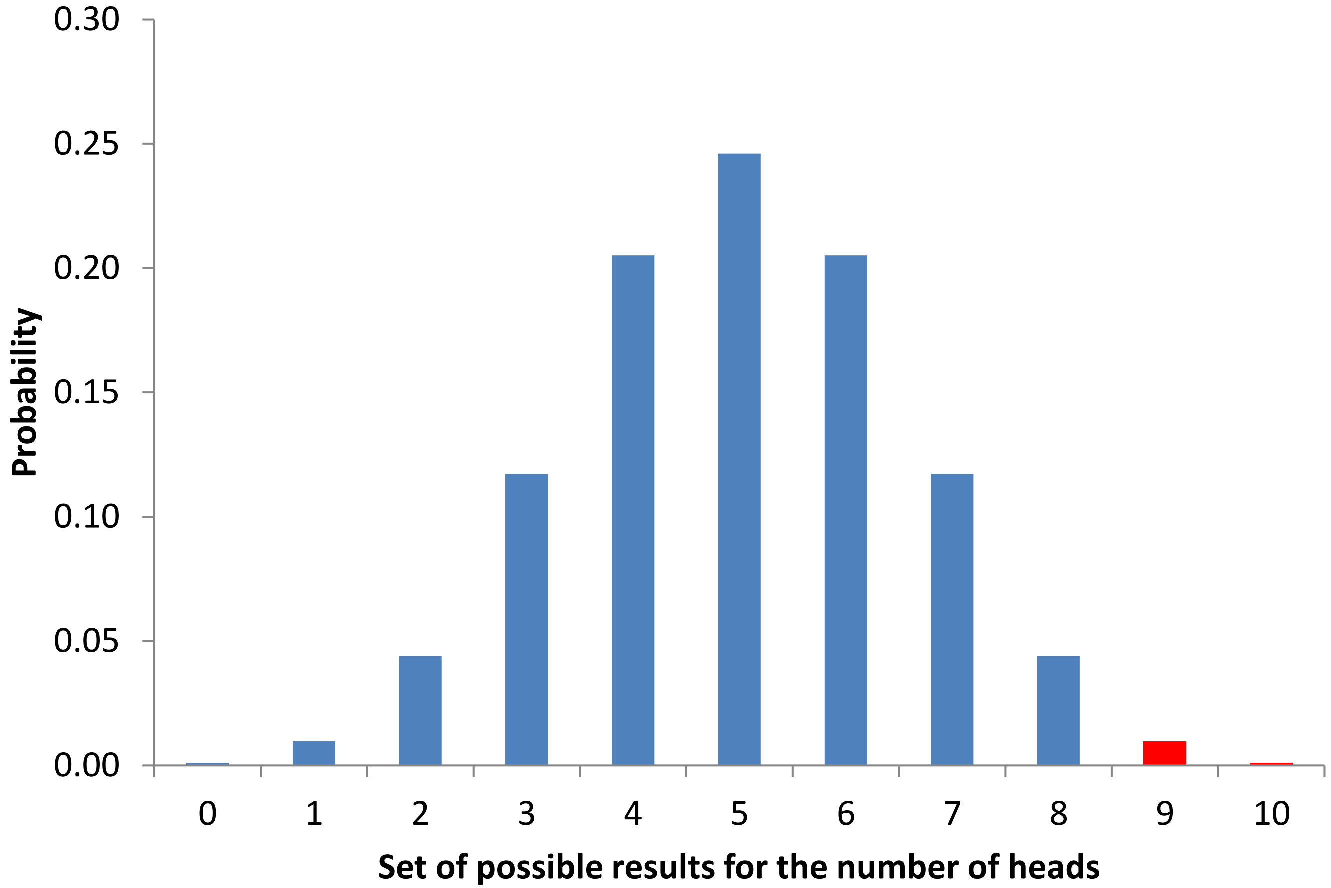
- The probability of obtaining a result at least as extreme, assuming H_0 is true.

- We toss the coin 10 times and a total of 9 heads are observed.

- Let X be the number of heads obtained, $\mathrm{X} \sim B(10,0.5)$.

- $\mathrm{P}(\mathrm{X} \geq 9) = P(X = 9) + P(X = 10) = 0.01074.$

Probability mass function



Result

- p-value $= 0.01074$
- Two tailed test, $H_1 : p \neq 1/2$, $\alpha = 0.05$
- p-value < 0.025 we reject H_0
- Can conclude, at a significance level of 5% that the coin is not fair.

Resources

<u>Online resources</u>

- http://www.statstutor.ac.uk/