Probability and Statistics: To p, or not to p?

Module Leader: Dr James Abdey

5.1 Statistical juries

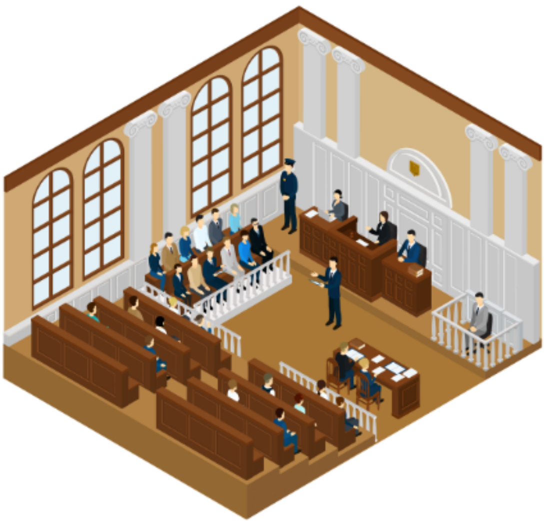
Week 5 considers **hypothesis testing**, i.e. decision theory whereby we make a binary decision between two competing hypotheses:

$$\mathbf{H_0 = \text{the null hypothesis} \quad \text{and} \quad H_1 = \text{the alternative hypothesis.}}$$

The binary decision is whether to **'reject H_0'** or **'fail to reject H_0'**. Before we consider statistical tests, we begin with a legal analogy – the decision of a jury in a court trial.

Example

In a criminal court, defendants are put on trial because the police suspect they are guilty of a crime. Of course, the police are biased due to their suspicion of guilt so determination of whether a defendant is guilty or not guilty is undertaken by an independent (and hopefully objective!) jury.



The jury has to decide between the two competing hypotheses:

$$\mathbf{H_0 : not\ guilty \quad and \quad H_1 : guilty.}$$

In most jury-based legal systems around the world there is the **'presumption of innocence until proven guilty'**. This equates to the jury initially believing H_0, which is the working hypothesis. A jury must continue to believe in the null hypothesis until they feel the evidence presented to the court proves guilt **'beyond a reasonable doubt'**, which represents the burden of proof required to establish guilt. In our statistical world of hypothesis testing, this will be known as the **significance level**, i.e. the amount of evidence needed to reject H_0.

The jury uses the following **decision rule** to make a judgement. If the evidence is:

- sufficiently inconsistent with the defendant being not guilty, then reject the null hypothesis (i.e. convict)
- not indicating guilt beyond a reasonable doubt, then fail to reject the null hypothesis – note that failing to prove guilt does not prove that the defendant is innocent![1]

Statistical hypothesis testing follows this same logical path.

Miscarriages of justice

In a perfect world juries would always convict the guilty and acquit the innocent. Sadly, it is not a perfect world and so sometimes juries reach incorrect decisions, i.e. convict the innocent and acquit the guilty. One hopes juries get it right far more often than they get it wrong, but this is an important reminder that miscarriages of justice do occur from time to time, demonstrating that the jury system is not infallible!

Statistical hypothesis testing also risks making mistakes which we will formally define as **Type I errors** and **Type II errors** in Section 5.2.

[1]Note that the jury is *not testing* whether the defendant is guilty, rather the jury is testing the hypothesis of not guilty. Failure to reject H_0 does not prove innocence, rather the jury concludes the evidence is not sufficiently inconsistent with H_0 to indicate guilt beyond a reasonable doubt. Admittedly, what constitutes a 'resonable doubt' is subjective which is why juries do not always reach a unanimous verdict!