0.1 Conclusions in hypothesis testing

0.1.1 Hypothesis Testing

The inferential step to conclude that the null hypothesis is false goes as follows: The data (or data more extreme) are very unlikely given that the null hypothesis is true. This means that:

(1) a very unlikely event occurred or

(2) the null hypothesis is false.

The inference usually made is that the null hypothesis is false. Importantly it doesn't prove the null hypothesis to be false.

0.1.2 Conclusions in hypothesis testing

- We always test the null hypothesis.
- We reject the null hypothesis, or
- We *fail to reject* the null hypothesis.

0.1.3 Types of Error

- The probability of a Type I error is designated by the Greek letter alpha (α) and is called the Type I error rate.
- The probability of a Type II error (the Type II error rate) is designated by the Greek letter beta (β).
- A Type II error is only an error in the sense that an opportunity to reject the null hypothesis correctly was lost.
- It is not an error in the sense that an incorrect conclusion was drawn since no conclusion is drawn when the null hypothesis is not rejected.

0.1.4 Type I and Type II Error

- A type II error would occur if it was concluded that the two drugs produced the same effect, i.e. there is no difference between the two drugs on average, when in fact they produced different ones.
- A type II error is frequently due to sample sizes being too small.

0.1.5 Type I and II errors

There are two kinds of errors that can be made in hypothesis testing:

(1) a true null hypothesis can be incorrectly rejected

(2) a false null hypothesis can fail to be rejected.

The former error is called a ***Type I error*** and the latter error is called a ***Type II error***.
The probability of Type I error is always equal to the level of significance α (alpha) that is used as the standard

for rejecting the null hypothesis .
These two types of errors are defined in the table below.

	True State: H0 True	True State: H0 False
Decision: Reject H0	Type I error	Correct
Decision: Do not Reject H0	Correct	Type II error

0.1.6 Type II Error

- The probability of a Type II error is designated by the Greek letter beta (β).
- A Type II error is only an error in the sense that an opportunity to reject the null hypothesis correctly was lost.
- It is not an error in the sense that an incorrect conclusion was drawn since no conclusion is drawn when the null hypothesis is not rejected.

0.1.7 Types of Error

- Requiring very strong evidence to reject the null hypothesis makes it very unlikely that a true null hypothesis will be rejected.
- However, it increases the chance that a false null hypothesis will not be rejected, thus lowering the likelihood of Type II error.
- The Type I error rate is almost always set at .05 or at .01, the latter being more conservative since it requires stronger evidence to reject the null hypothesis at the .01 level then at the .05 level.

0.1.8 Type I and Type II errors - Die Example

- Recall our die throw experiment example.
- Suppose we perform the experiment twice with two different dice.
- We don't not know for sure whether or not either of the dice is fair or crooked (favouring high values).
- Suppose we get a sum of 401 from one die, and 360 from the other.

0.1.9 Type I and Type II errors - Die Example

- For our first dice (sum 401), we feel that it is likely that the die is crooked.
- A Type I error describes the case when in fact that dice was fair, and what happened was just an unusual result.
- For our second dice (sum 360), we feel that it is likely that the die is fair.
- A Type II error describes the case when in fact that dice was crooked , favouring high values, and what happened was ,again, just an unusual result.

0.1.10 Type II Error

- The probability of a Type II error is designated by the Greek letter beta (β).
- A Type II error is only an error in the sense that an opportunity to reject the null hypothesis correctly was lost.
- It is not an error in the sense that an incorrect conclusion was drawn since no conclusion is drawn when the null hypothesis is not rejected.

0.1.11 Types of Error

- A Type I error, on the other hand, is an error in every sense of the word. A conclusion is drawn that the null hypothesis is false when, in fact, it is true. Therefore, Type I errors are generally considered more serious than Type II errors.
- The probability of a Type I error ( ) is called the significance level and is set by the experimenter. There is a trade-off between Type I and Type II errors. The more an experimenter protects himself or herself against Type I errors by choosing a low level, the greater the chance of a Type II error.
- Requiring very strong evidence to reject the null hypothesis makes it very unlikely that a true null hypothesis will be rejected. However, it increases the chance that a false null hypothesis will not be rejected, thus lowering the likelihood of Type II error.
- The Type I error rate is almost always set at .05 or at .01, the latter being more conservative since it requires stronger evidence to reject the null hypothesis at the .01 level then at the .05 level.

0.1.12 Type I and Type II errors - Die Example

- Recall our die throw experiment example.
- Suppose we perform the experiment twice with two different dice.
- We don't not know for sure whether or not either of the dice is fair or crooked (favouring high values).
- Suppose we get a sum of 401 from one die, and 360 from the other.

The probability of Type I error is always equal to the level of significance that is used as the standard for rejecting the null hypothesis; it is designated by the lowercase Greek α (alpha).

Types of Error

- The probability of a Type I error is designated by the Greek letter alpha (α) and is called the Type I error rate.

0.1.13 Type I and II errors

Type I and Type II Error

- A type II error would occur if it was concluded that the two drugs produced the same effect, i.e. there is no difference between the two drugs on average, when in fact they produced different ones.
- A type II error is frequently due to sample sizes being too small.

The probability of Type I error is always equal to the level of significance that is used as the standard for rejecting the null hypothesis; it is designated by the lowercase Greek α (alpha).

Types of Error

- The probability of a Type I error is designated by the Greek letter alpha (α) and is called the Type I error rate.

ImportantIn this module, the significance level α can be assumed to be 0.05, unless explicitly stated otherwise.

	H0 is true	H0 is false
Accept H0	correct decision P 1-alpha	type II error P alpha (significance)
Reject H0	type I error P beta	correct decision P 1-beta (power)

If you are interested in further details of probability and sampling theory at this point then please refer to one of the general texts listed in the reference section.

You must understand confidence intervals if you intend to quote P values in reports and papers. Statistical referees of scientific journals expect authors to quote confidence intervals with greater prominence than P values.

0.1.14 Notes about Type I error:

is the incorrect rejection of the null hypothesis maximum probability is set in advance as alpha is not affected by sample size as it is set in advance increases with the number of tests or end points (i.e. do 20 tests and 1 is likely to be wrongly significant)

0.1.15 Notes about Type II error:

is the incorrect acceptance of the null hypothesis probability is beta beta depends upon sample size and alpha can't be estimated except as a function of the true population effect beta gets smaller as the sample size gets larger beta gets smaller as the number of tests or end points increases

0.2 Type I and II errors

1. The experiment has been carried out in an attempt to disprove or reject a particular hypothesis, the null hypothesis, thus we give that one priority so it cannot be rejected unless the evidence against it is sufficiently strong. For example,

   H0: there is no difference in taste between coke and diet coke

   H1: there is a difference.

2. . If one of the two hypotheses is 'simpler' we give it priority so that a more 'complicated' theory is not adopted unless there is sufficient evidence against the simpler one. For example, it is 'simpler' to claim that there is no difference in flavour between coke and diet coke than it is to say that there is a difference.

- The hypotheses are often statements about population parameters like expected value and variance; for example H0 might be that the expected value of the height of ten year old boys in the Scottish population is not different from that of ten year old girls.
- A hypothesis might also be a statement about the distributional form of a characteristic of interest, for example that the height of ten year old boys is normally distributed within the Scottish population.
- This is called **Type I error**. The probability of Type I error is always equal to the level of significance that is used as the standard for rejecting the null hypothesis; it is designated by the lowercase Greek (α), and thus a also designates the level of significance.
- The most frequently used levels of significance in hypothesis testing are the 5 percent and 1 percent levels.
- A Type II error occurs if the null hypothesis is not rejected, and therefore accepted, when it is in fact false.
- There are two kinds of errors that can be made in significance testing: (1) a true null hypothesis can be incorrectly rejected and (2) a false null hypothesis can fail to be rejected.
- The former error is called a Type I error and the latter error is called a Type II error. These two types of errors are defined in the table below.

At this point, a word about error. Type I error is the false rejection of the null hypothesis and type II error is the false acceptance of the null hypothesis. As an aid memoir: think that our cynical society rejects before it accepts.

The significance level (alpha) is the probability of type I error. The power of a test is one minus the probability of type II error (beta). Power should be maximised when selecting statistical methods. If you want to estimate sample sizes then you must understand all of the terms mentioned here.

The following table shows the relationship between power and error in hypothesis testing:

Table 1: Summary of Type I and Type II Errors

	when H0 is true	when H1 is true
Do not Reject H0	correct decision $p = 1 - \alpha$	Type II error $p = \beta$
Reject H0	Type I error $p = \alpha$	correct decision $p = 1 - \beta$

0.2.1 Finite Population Correction Factor

If the sample size is more than 5% of the population size and the sampling is done without replacement, then a correction needs to be made to the standard error of the means.

In the following, N is the population size and n is the sample size. The adjustment is to multiply the standard error by the square root of the quotient of the difference between the population and sample sizes and one less than the population size.

For the most part, we will be ignoring this in class