Drawing and testing assumptions

Errors and estimates

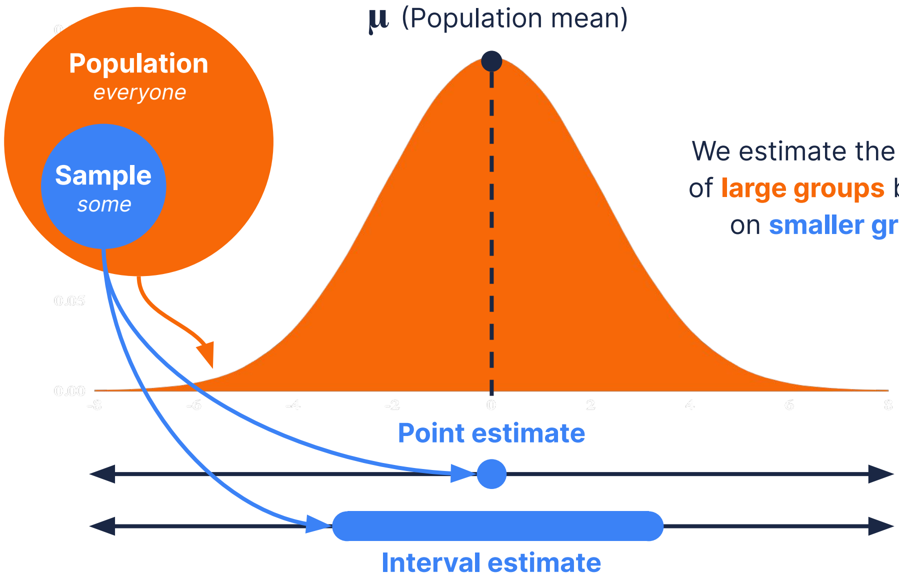
Errors and estimates

> Errors and estimates are important because they allow us to **properly interpret** the results of a **hypothesis test** and draw **accurate conclusions** about the **population**.

In hypothesis testing, we typically want to make **inferences** about a **population** based on a **sample** of data from that population.

Very often **we can't observe the entire population**, so we use the **sample to estimate** population parameters such as the mean or standard deviation.

When we use these estimates, there is a **degree of uncertainty** associated because our sample may **not perfectly represent the population**.



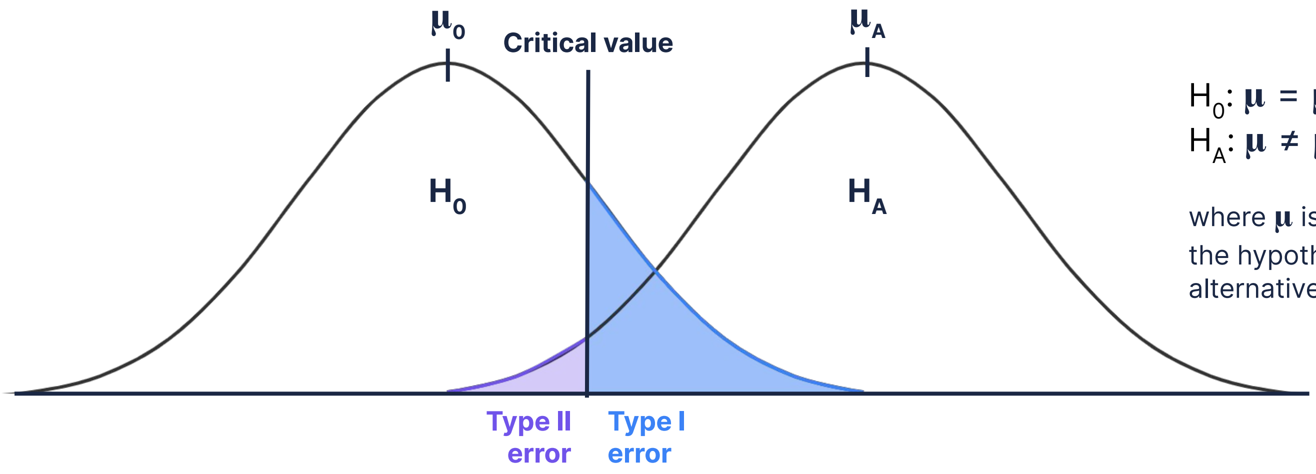
We estimate the value of **large groups** based on **smaller groups**.

Errors in hypothesis testing

The uncertainty resulting from the estimates we use means that there is a **chance of making an incorrect decision** in our hypothesis tests. These incorrect decisions are called **type I** and **type II** errors.

Truth

Decision	**H_0 is true**	**H_0 is false**
Fail to reject H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision



H_0: $\mu = \mu_0$
H_A: $\mu \neq \mu_0$

where μ is the population mean, μ_0 the hypothesised mean, and μ_A the alternative mean.

Type I error

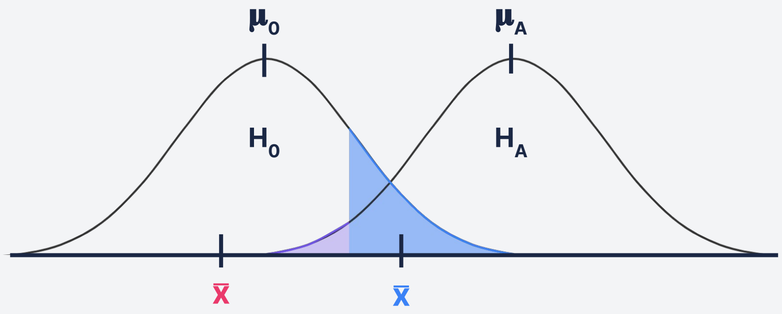
> In hypothesis testing, a **type I error** occurs when the null hypothesis is **rejected** when it is actually **true**. It is a false positive error.

A type I error occurs when **H_0 is rejected** even though it is **actually true**, i.e. we conclude that there is a significant effect or relationship where there is none, or falsely conclude there is an effect.

When H_0 ($\mu = \mu_0$) is **TRUE**, the distribution of the sample mean ($\bar{x}$) is centred around the hypothesised mean (μ_0).

When H_0 is **TRUE** and we *fail to reject* H_0 and $\bar{x} \approx \mu_0$ → **NO error, i.e. correct decision**

When H_0 is **TRUE** and we *reject* H_0 and $\bar{x} \approx \mu_0$ → **Type I error**



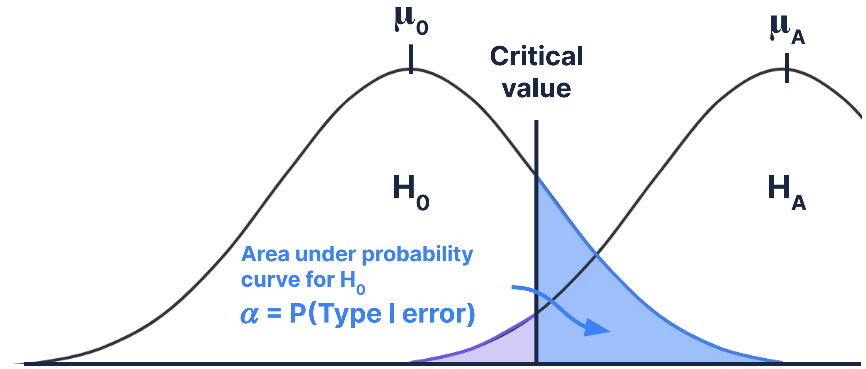
Type I error and alpha

> The **probability of a type I error** is denoted by α and relates to the **level of significance** (also denoted by α).

The **level of significance**, the predetermined threshold for deciding whether to reject or fail to reject the null hypothesis, is the **probability of making a type I error**, P(Type I error).

This means that when we set the level of significance, we actually set the **maximum probability of making a type I error** that we are willing to accept.

As such, if we set our level of significance to 5%, we were actually saying that we are willing to accept a 5% chance of making a type I error.



Thus, we can **reduce our chances of a type I error** by **decreasing the level of significance** rather than increasing our sample size.

Type II error

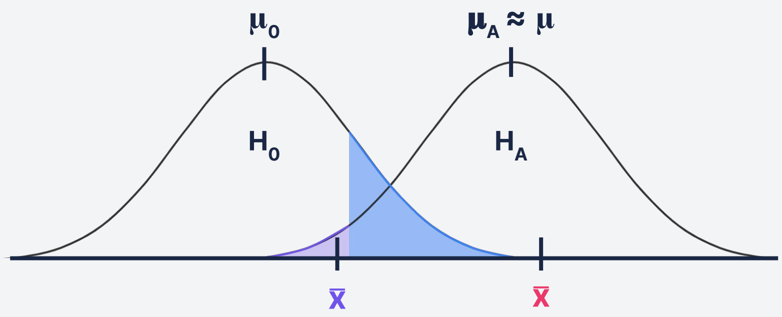
In hypothesis testing, a **type II error** occurs when the null hypothesis is **accepted** when it is actually **false**. It is a false negative error.

A type II error occurs when we **fail to reject H_0** even though it is **actually false**, i.e. we conclude that there is no significant effect or relationship where there is, or falsely conclude there is no effect.

When H_0 ($\mu = \mu_0$) is **FALSE**, the distribution of the sample mean ($\bar{x}$) is centred around the alternative mean (μ_A) which represents the true population mean under these conditions.

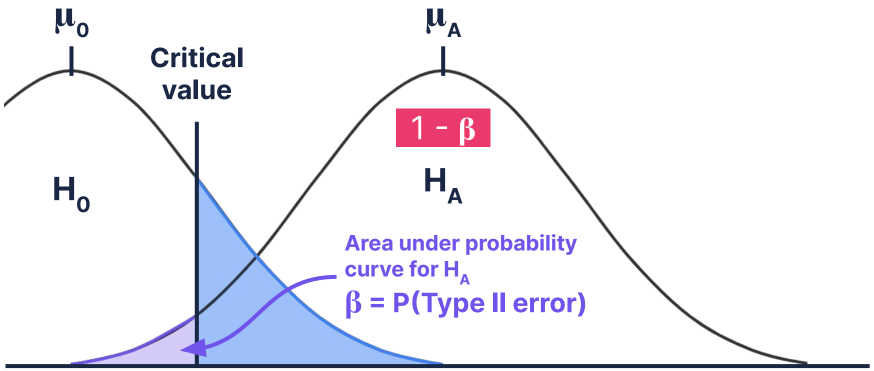
When H_0 is **FALSE** and we *reject* H_0 and $\bar{x} \approx \mu_A$ → **NO error, i.e. correct decision**

When H_0 is **FALSE** and we *fail to reject* H_0 and $\bar{x} \approx \mu_A$ → **Type II error**



Type II error and beta

The **probability of a type II error** is denoted by **β** and is used to determine the **power** of a test.



The **power** of a hypothesis test is the probability that the null hypothesis is rejected when it is false. In other words, it is the probability of identifying a real difference or relationship when one truly exists.

It is calculated from the probability of a type II error:

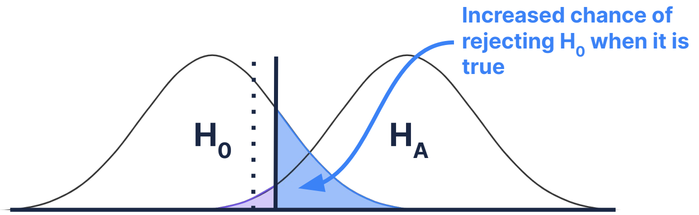
power = 1 - β = 1 - P(Type II error)

We can **reduce our chances of making a type II error** by **increasing the power** of our test, which means we need to **increase the sample size** or **increase the level of significance**.

Minimising the probability of errors

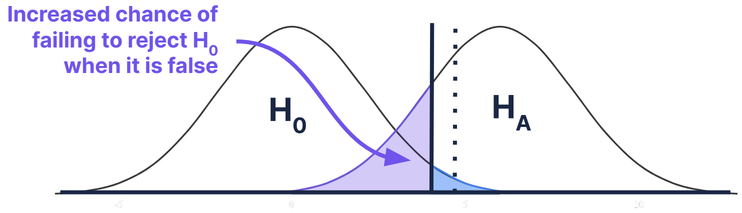
There is a **trade-off** between the **level of significance** and the **power of a test** because reducing the chances of one type of error could increase that of the other.

To **decrease the chances of a type II error**, we can either take a larger sample or we can **increase the power** by increasing the level of significance. However, if we do the second, we increase the probability of a type I error.



The probability of making a type I error is equal to the level of significance of the test.

To **decrease the probability of a type I error**, we need to decrease the level of significance, but changing the sample size has no effect on the probability of a type I error.



Increasing the sample size is often a practical solution in business settings, while changing the level of significance is not. The level of significance is often a fixed value, depending on the use case.