Errors in hypothesis testing

From previous notes, we have learnt that, in any testing procedure, we always start by assuming that the null hypothesis H_0 is true before carrying out any relevant investigation. However, we have to bear in mind that sampling errors often lead us to the wrong conclusion and decision as regards the acceptance or rejection of H_0.

Sometimes, it may happen that a null hypothesis is, in fact, correct but that our *sample statistic* (point estimate of the *parameter* under investigation) is relatively far from the proposed value. But it may also happen that the null hypothesis is false and yet the value of the sample statistic is relatively close to the proposed value! Whatever be the case, we are making a wrong decision.

The following table displays all the possible states-of-nature and decisions in a decision-making process:

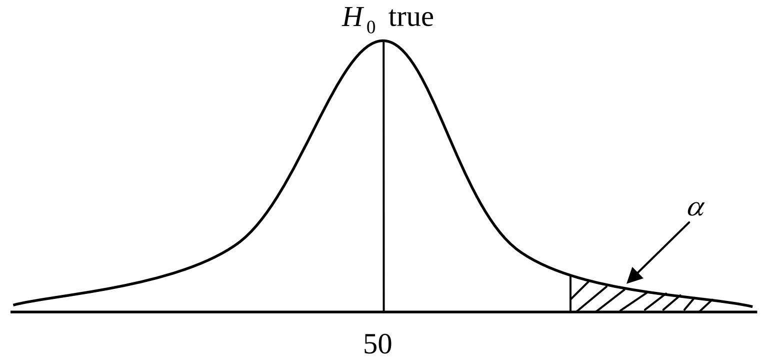
		Decision	
		Accept H_0	Reject H_0
States of nature	H_0 true	Correct	**Type I error**
	H_0 false	**Type II error**	Correct

Type I error

From definition, a Type I error is made when we reject a true null hypothesis whereas a Type II error is made when we accept a false null hypothesis. It is definitely easier to understand the meaning of a Type I error since we always assume that a null hypothesis is true before testing. We will refer to the Normal distribution in our discussion, on the basis that our sample is large so that we may quote the Central Limit Theorem.

Consider an example where we are testing the null hypothesis that the population mean is 50 kg against the alternative hypothesis that the population mean is greater than 50 kg at a 5% level of significance. Assuming that H_0 is true, we would draw a Normal distribution centred at 50 with the critical region (region of rejection of H_0) lying on the extreme right-hand side (shaded in the diagram below).

$H_0: \mu = 50$
$H_1: \mu > 50$



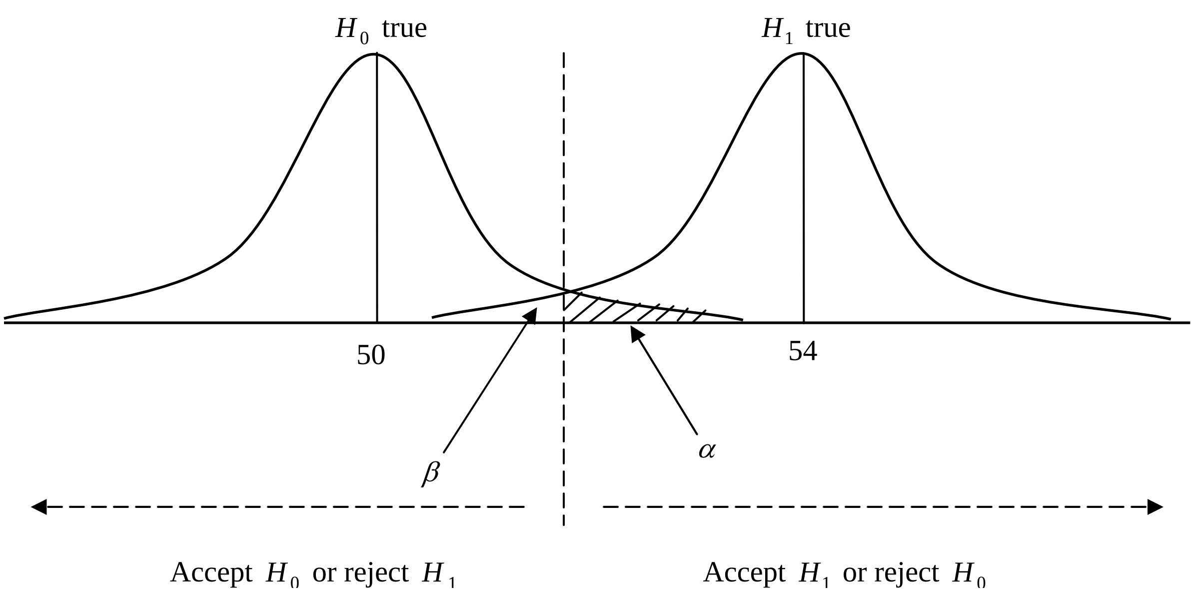
It is clear in the above diagram that the region where we reject H_0 is the shaded region, that is, the significance level (critical region). Therefore, the probability of rejecting a true null hypothesis is just the probability of making a Type I error, that is, α. Note, however, that α is just a nominal value which is given before testing the hypothesis. In general, P[Type I error] is not necessarily equal to α. It all depends on two factors:

(1) The normality of the population and
(2) The randomness of the sample.

If these two conditions are satisfied, we can safely say that P[Type I error] = α.

Type II error

The probability of a Type II error is slightly harder to calculate since it is often confusing to locate the region of acceptance of a false null hypothesis. To accept H_0 when H_0 is false would be somewhat equivalent to reject H_1 when H_1 is true. However, we would need the distribution of H_1 true, that is, the centre or mean of H_1 true before calculating the probability of making a Type II error. Let us consider the above example with the true value of the mean being given as 54 (remember we are looking at the distribution of H_1 true). The diagram would be as follows:



If we consider the distribution of H_1 true, we will see that its region of rejection is the little area labelled as β, which is exactly the probability of making a Type II error (accept H_0 when H_0 is false). Note that α and β are inversely proportional in size since they lie next to each other, being divided by the line of demarcation established by the nominal significance level. The P[Type I error] can be minimized by choosing a large sample, thus decreasing the standard error of the estimate.

The *power* of a test is defined as the probability of rejecting a false null hypothesis (rejecting H_0 when H_0 is false, which is indeed a correct decision). It should be clear that we are considering the quantity 1 – β, which has to be maximized in any testing procedure. This would mean minimizing the probability of a Type II error and thus maximizing the probability of a Type I error.