Type I and Type II errors in significance tests

When we make a decision ruling on whether a measurement is in compliance with a specification, a certain threshold or a regulatory limit, we run into a risk of making either false accept or false reject situation. This is very similar to the conclusions drawn from a significance test.

As we know, the conclusion from a hypothesis test is drawn based on probabilities. This has the consequence that sometimes we draw the wrong conclusion.

A significance test at, for example, the $P \leq 0.05$ or 5% significance level involves that a null hypothesis, H_o, will be *rejected even though it is true*. This situation of *rejecting a true hypothesis* is known as a **Type I error** (α). In other words, the probability of committing a type I error is 5%.

For a general significance level α, the probability of committing a type I error is α. Hence, by adjusting the significance level we can change the probability of rejecting a true hypothesis.

That is, such error risk can be reduced by altering the significance level of the test to $P \leq 0.01$ or even $P \leq 0.001$. However, reducing this rate comes at the expense of increasing the rate of another type of error. This is because type I is not the only possible type of error.

It is also possible for us to *retain a null hypothesis even when it is false*. This is called a **Type II error** (β). Therefore, in order to calculate the probability of this type of error, it is necessary to postulate an alternative to the null hypothesis, known as an alternative hypothesis H_1.

In summary, a type I error is also called a *false negative*, i.e., we conclude that there is a difference or an association when in reality there is not. Likewise, a type II error is also called a *false positive* since we conclude that there is really no difference/association when in truth there is.

Controlling the probabilities of making errors

There are 4 possible scenarios as we carry out a hypothesis test: the null hypothesis Ho is either true or false, and it is either rejected or not rejected.

The conclusion is correct whenever we reject a false hypothesis or accept a true hypothesis. See table 1.

Table I: The 4 possible outcomes when testing an hypothesis

	Truth	
	Ho is true	Ho is false
Reject Ho	Type I error False positive	Correct conclusion True positive
Fail to reject Ho	Correct conclusion True negative	Type II error False negative

Example

Assume that the concentration of PSA (prostate specific antigen), a cancer detector in the blood is measured in order to detect prostate cancer. If the concentration is larger than 4 ng/L threshold, then the "alarm goes off" and the patient is sent for further investigation. But, how large should the threshold be? If it is large, then some patient will not be classified as sick although they are (so type II β error). On the other hand, if the threshold is set low, then patients will be classified as sick although they are not (type I α error).

(*We shall illustrate these two types of error graphically in the next article*.)