

## Limit of Detection; Type I (False Positive) and Type II (False Negative) errors

Undeniably the concept of limit of detection (LOD) has been and still is, one of the most controversial subject in analytical chemistry. It is vital in trace analysis where measurements are made at very low test analyte or property levels, such as in trace pesticide analysis. It is important to know what the lowest concentration of the analyte or property value that can be confidently detected by the analytical method is.

The LOD is often confused with the method sensitivity. In fact, the sensitivity of a test method is the capability of the method to discriminate small differences in concentration or mass of the analyte. For example, the slope of a calibration curve can show the sensitivity of the method.

The LOD of an analytical procedure is normally defined as the lowest amount of an analyte that can be *reliably* detected with a given analytical method, but not necessary quantitated as an exact value. Because of the word “*reliably*”, the LOD is itself a statistical estimate of concentration at which we can be fairly certain that the compound is present. Concentrations below this limit may not be detected.

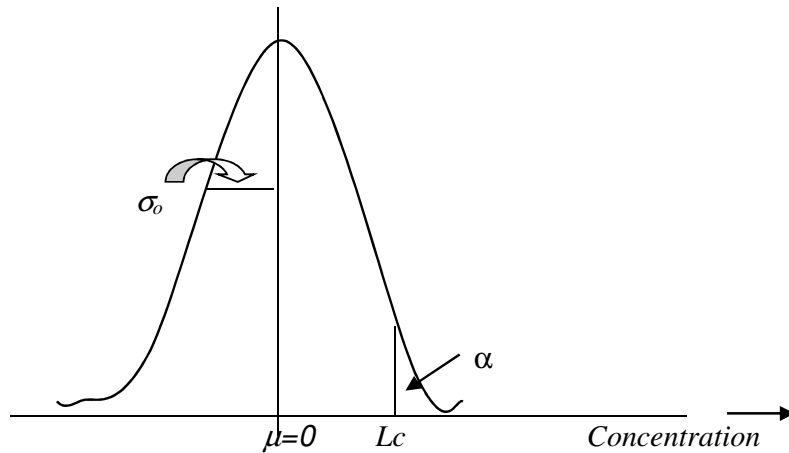
In other words, the LOD would be the lowest concentration obtained from the measurement of a sample (containing the component) that we would be able to discriminate from the concentration obtained from the measurement of a blank sample (a sample not containing the analyte).

So, when we state “Not Detected (ND)” for an analyte in our test report, it indicates that this analyte may be present below the detection limit. But, we do not wish to claim the presence of the analyte when it is actually absent; i.e. a false positive or Type I ( $\alpha$ ) decision error. Equally, we do not wish to report that the analyte is absent when it is truly present, i.e. a false negative or Type II ( $\beta$ ) decision error.

The meanings of these two types of error will be clearer by the following illustrations.

When there is no bias or systematic error in a test procedure, numerous repeats on the analysis of a blank sample, due to the presence of random error, would give concentration values distributed around zero with a given standard deviation  $\sigma_0$ . At some point in this zero-concentration value

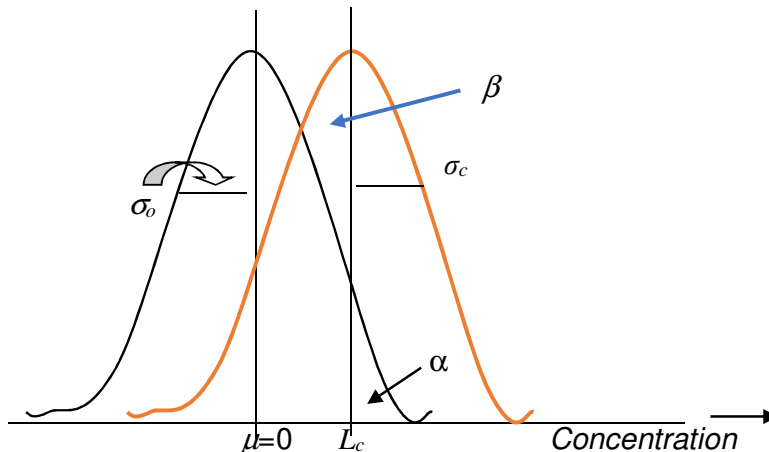
distribution, we can set a critical level  $L_c$  that allows us to make a decision whether the analyte is present or not. See the sketch below.



If the concentration obtained is higher than  $L_c$  hypothetically in this blank sample, we can state that the analyte is present in the sample but run a risk of falsely conclude that the analyte is actually present. This certain probability,  $\alpha$  is called type I error, indicating that we might probably commit a false positive.

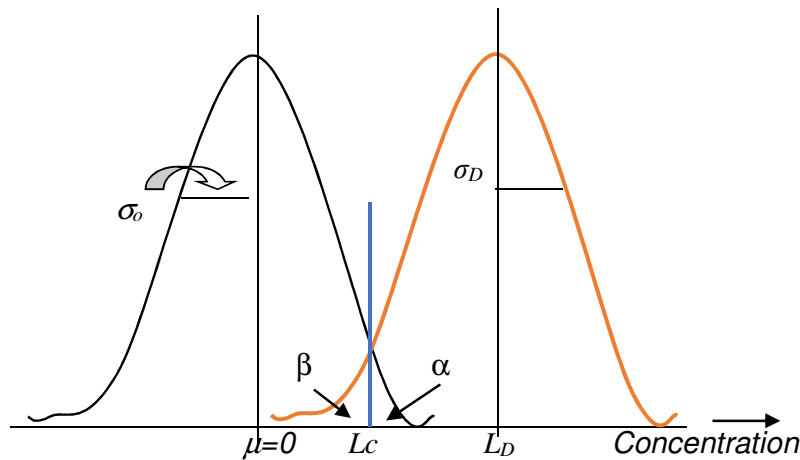
When we choose  $\alpha=0.50$ , we say on 50% occasions, a result on the blank sample would give a value greater than zero. But the risk of committing a false positive at this  $L_c$  is too high. More logically in most situations, we might want to fix  $\alpha = 0.05$ , indicating that the risk of declaring false positive is only 5%.

If we were to prepare a sample with a concentration  $L_c$  and repeat the analysis several times, we would obtain a distribution of values with a mean value of  $\mu = L_c$ , and a standard deviation of  $\sigma_c$ . As the concentration is very small, we can safely assume  $\sigma_0 = \sigma_c$ . See sketch below.



In this second round of analysis, we might declare in half of the cases that the analyte is not present in the sample with  $L_C$  concentration (i.e. it is not detected) because the concentrations are found below the critical level  $L_C$ . This is in contradiction to what actually happens when we analyse a sample with  $L_C$  concentration. Hence we are running a risk of falsely concluding that the analyte is not detected when in fact it is present. This probability  $\beta$  is called type II error, indicating that we might commit a false negative error.

Certainly it is not wise to put  $\beta = 0.50$  at a high 50% risk in committing this false negative error. To reduce such risk, it will be more logical to put  $\beta = 0.05$  by deciding on a concentration level  $L_D$  which gives  $\alpha=0.05$  in the sample analysis of  $L_C$  concentration. See the sketch below.



In here, we can safely assume that  $\sigma_o = \sigma_D$  because of small difference between concentration zero and  $L_D$ . It is now obvious that if the laboratory were to analyse samples containing  $L_D$  concentration, it would be running a much lower risk  $\beta$  by stating the analyte is not detected when in fact it is present. This  $L_D$  is known as limit of detection.

Mathematically, we can write the following equations to estimate  $L_D$ :

$$L_C = 0 + z_{(1-\alpha)}\sigma_o$$

$$L_D = L_C + z_{(1-\beta)}\sigma_D = z_{(1-\alpha)}\sigma_o + z_{(1-\beta)}\sigma_o = 2z_{(0.95)}\sigma_o$$

Where  $\alpha = \beta = 0.05$  and statistic  $z_{(0.95)} = 1.645$  under the one-sided normal (Gaussian) distribution

Hence, the limit of detection  $L_D = 0 + 2 \times 1.645 \sigma_o = 3.3 \sigma_o$ .

However, if the laboratory does not run sufficient number of replicates (say,  $n < 30$ ), one has to replace the population standard deviation  $\sigma$  with a sample standard deviation  $s$  and apply the  $t$ -statistic instead of using those parameters in the normal distribution. In that case, we use the following equation:

$$LC = t_{(1-\alpha, v=n-1)} S_o$$

$$LD = 2 \times t_{(1-\alpha, v=n-1)} S_o$$

where  $n$  is the number of repeats and  $v$  is the degree of freedom.

For some analytical instrumentation and by some learned professional organisations, the instrument's signal-to-noise (S/N) ratio is being used to estimate the LOD. Generally  $S/N \geq 3$  is taken as the LOD. In chromatography, this detection limit is the injected amount that results in a peak with a height at least three times as high as the baseline noise level.

