Outliers in analytical data – the purpose and action on detecting

If we do a visual inspection of a series of analytical data using dot plots, box plots, etc., we might see one or two seemingly outlying values in this data set. It is appropriate to confirm them by carrying out some outlier tests.

The purpose of outlier testing

The purpose of outlier tests is to show whether such values could reasonably arise from chance variation or are so extreme as to indicate some other cause such as an unusual test sample, a human error or an instrument fault.

This outlier testing both protects the analyst from taking unnecessary action based on such chance or probability variation and directs attention towards likely analytical problems that require inspection and possible correction. The most important role of outlier testing is therefore to provide objective criteria for taking appropriate investigative or corrective action.

Outlier tests are also used in some circumstances to provide a degree of robustness. Although outright rejection of an extreme value on statistical grounds alone is not generally recommended, if there is good reason to believe that errors are likely, rejection of extreme outliers can, like robust statistics, prevent errors from unduly influencing results.

Action on outlier detection

Before considering individual tests, it is useful to see what action should be taken on the basis of outlier tests. It must be stressed that statistical outlier is only *unlikely* to arise by chance. A positive outcome from an outlier test is best considered as a signal to investigate the cause; usually, outlier should not be removed from the data set easily solely because of the result of a statistical test without other considerations.

This is because experience suggests that human or other error is among the most common causes of extreme outliers. This experience has given rise to fairly widely used guidelines for acting on outlier tests on analytical data, based on the outlier testing and inspection

procedure included in ISO 5725 Part 2 for processing inter-laboratory data. The main features are:

1. Test at the 95% and the 99% confidence level
2. All outliers should be investigated and any errors corrected
3. Outliers significant at the 99% level may be rejected unless there is a technical reason to retain them
4. Outliers significant only at the 95% level (often termed stragglers) should be rejected only if there is an additional technical reason to do so
5. Successive testing and rejection is permissible, but not to the extent of rejecting a large proportion of the data

This procedure leads to results which are not seriously biased by rejection of chance extreme values, but are relatively insensitive to outliers at the frequency commonly encountered in measurement work.

Do remember that an outlier is only 'outlying' in relation to some prior expectation. The outlier tests assume underlying normality. If the data were Poisson distributed, for example, many valid high values might be incorrectly rejected because they appear inconsistent with a normal distribution. For example, total bacterial count test is Poisson distributed with wide range of replicated results expected. It is therefore crucial to consider whether outlying values might represent genuine features of the population.

In another case, testing relatively inhomogeneous or granular materials will often show apparent outliers where different particles have very different analyte present. A good example is the determination of aflatoxins in nuts, where a small portion – even one kernel in a laboratory sample – can contain hundreds of times more toxin than the majority. Rejecting the observation due to that kernel on statistical grounds would be entirely incorrect; the 'outlier' is genuinely representative of the population and it must be included in calculating mean values to obtain a correct decision.

Similarly, an outlying observation in a process control environment is an important signal of a process problem; if all the outlying values were rejected, its process control would be no longer effective. Hence, it follows that outlier testing needs careful consideration where the population characteristics are unknown or, worse, known to be non-normal.