Quality Control for Multi-variable Problems

Submitted by: Chemex Laboratories

Analytical methods were many elements are determined simultaneously present special difficulties in quality control. The use of individual control charts for each variable (element) is generally not effective due to the combinative aspect of the variation. If there are n elements and the error on each is independent, the probability that all the results will fall within their respective 95% confidence limits is (0.95)n. If n = 30 (multi-element analytical packages typically have 25 to 40 elements) this probability is equal to (0.95)30 or 0.2146. In other words there will be one or more elements out of bounds in 78.53% of all batches, which on a simplistic interpretation suggests that about two-thirds of all batches will be rejected when nothing is actually wrong with the sample data.

In practice there will normally be both correlated error, (due to instrument variations that affect all elements or groups of elements) and independent error which is due to variations in the individual element measurement channels.

Some of the solutions that are being used or tested are listed below.

1. Widen Control Limits

   The general approach has been to widen the control limits for multi-element analyses to (3s. This has the effect of reducing the probability of false batch rejection for purely independent error to 1 - (0.9974)n which is 0.075 (7.5%).

   This disadvantages of this approach are a sacrifice in data quality for each individual element and as more and more elements are included in data packages (50 +) the probability of false batch rejection is relatively high - greater than 10% for 50 elements.

2. Multivariate Quality Control

   Principle component analysis has been examined. R.J. Howarth, M.H. Ramsey and B.J. Coles presented "The potential of multivariate quality control as a diagnostic tool in Geoanalysis" at Geoanalysis 97 (Vail, CO, June 97). Data from a 35 element multi-element analytical package was reduced to four principle components. They then suggested control be done by using the five control charts - one for each of the four components and a fifth for Hotelling's T - representing the multivariate distance.

   This is a sophisticated approach that requires large training sets and significant calculations for each test. Ideally the QC control can be done 'on-line' during the analytical sequence (the instrument software recognizes a control sample and subroutine testing for in / out of control is down before the next sample is analyzed - in less than a minute - so that time is not wasted doing out of specification analysis). The calculations may be too complex for on-line control.

   More problematic is the 'components' that are identified. Once a sample is identified as unacceptable the instrument operator needs to take corrective actions to resolve the problem before proceeding with further analyses. If the components are not closely related to physical properties (or are all combinations of a number of properties), identification of the 'fix' to the out of control problem may be difficult.

As well, multiple control charts, one for each component, are still being used.

3. Binomial Probabilities

   Tables of the binomial probability of M out of N points falling above the 95th (90th or 99th as well) can be constructed. Control limits can then set for 95/failing versus the total number determined be evaluated. If the probability of this event is less than a predetermined level (i.e. 0.5) the sample can be determined to be a failure.