

## 4.1 Introduction

Analytical chemistry is an applied measurement science in which quantitative studies predominate, and therefore one in which estimates of the inevitable errors are essential. In almost all applications of analysis the results obtained are supplied to a customer or user, and it is necessary that these users are satisfied as far as possible with the **quality** - the fitness for purpose - of the measurements. This has many important implications for analytical practice.

- Firstly, any assessment of the measurement errors must take into account the whole analytical process — including the sampling steps, which often contribute to the overall error very significantly.
- Secondly, the performance of the analyses undertaken in each laboratory must be checked internally on a regular basis, usually by applying them to standard or reference materials.
- Thirdly, in many application areas the results from different laboratories must be compared with each other, so that the users can be satisfied that the performance of the laboratories meets statutory, regulatory and other requirements.
- Finally, the analytical results must be supplied with a realistic estimate of their uncertainty, i.e. the range within which the true value of the quantity being measured lies.

The statistical methods used in such areas are often very simple in principle, many of them being based on techniques previously discussed in this module.

But their more frequent and improved application has been one of the major developments in analytical sciences in recent years, with a correspondingly large improvement in the quality and acceptability of many analytical results.

Moreover some of the methods discussed have broader applications. For example, the principles used to monitor the performance of a single analysis in a single laboratory over a period of time can also be applied to the monitoring of an industrial process.

## 4.2 Sampling

In most analyses we rely on chemical samples to give us information about a whole object. So unless the sampling stages of an analysis are considered carefully, the statistical methods discussed in this book may be invalidated, as the samples studied may not be properly representative of the whole object under study.

For example, it is not possible to analyse all the water in a stream for 2 toxic pollutant, and it is not possible to analyse all the milk in a tanker lorry to see if it contains a prohibited steroid hormone.

In other instances a small sample has to be used because the analytical method is destructive, and we wish to preserve the remainder of the material.

So in each case the sample studied must be taken in a way that ensures as far as possible that it is truly representative of the whole object.

To illustrate some aspects of sampling let us consider the situation in which we have a large batch of tablets and wish to obtain an estimate for the mean weight of a tablet.

Rather than weigh all the tablets, we take a few of them (say 10) and then weigh each one. In this example the batch of tablets forms the population and the 10 weighed tablets form a sample from this population

If the sample is to be used to deduce the properties of the population, it must be what is known statistically as a random sample, i.e. a sample taken in such a way that all the members of the population have an equal chance of inclusion.

### **Randomness**

Only then will statistics, such as the confidence limits of the mean, be valid. It must be appreciated that the term 'random' has, in the statistical sense, a different meaning from 'haphazard'. Although in practice an analyst might spread the tablets on his desk and attempt to pick a sample of 10 in a haphazard fashion: such a method could conceal an unconscious bias. The best way to obtain a random sample is by the use of a random number table.

Each member of the population is allocated a number in such a way that all the numbers have an equal number of digits e.g. 001, 002, 003, etc. Random numbers are then read off from a random number table (see Table A.8 of Book), starting at an arbitrary point to give,

for example, **964, 173**, etc., and the corresponding members of the population form the sample.

An alternative (and much simpler) method which is sometimes used is to select the population members at regular intervals, for example to take every hundredth tablet off a production line. This method is not entirely satisfactory since there might be a coinciding periodicity in the weight of the tablets: the importance of the randomness of the sample is evident.

(Recall sampling procedures from MA4603: systematic sampling, random sampling, quota sampling)

Again, if the last few tablets were taken and there had been a gradual decrease in weight during the production of the batch, then this sample would give an entirely erroneous value for the mean weight of the whole batch.

### **Bulk Sampling**

In the example above the population is made up of obvious discrete members that are nominally the same, viz. the tablets. Sampling from materials for which this is not true, such as rocks, powders, gases and liquids, is called **bulk sampling**.

If a bulk material were perfectly homogeneous then only a small portion or **test increment** would be needed to determine the properties of the bulk. In practice bulk materials are non-homogeneous for a variety of reasons.

Materials such as ores and sediments consist of macroscopic particles with different compositions and these may not be uniformly distributed in the bulk. Fluids may be non-homogeneous on a molecular scale owing to concentration gradients. Such inhomogeneity can be detected only by taking a sample of test increments from different parts of the bulk.

If possible this should be done randomly by considering the bulk as a collection of cells of equal size and selecting a sample of cells by using random numbers as described above.

From the random sample, the mean and variance can be calculated.

- mean,  $\bar{x}$
- variance,  $s^2$

There are two contributions to variance:

- 1) The sampling variance- due to differences between the members of the sample, e.g. the tablets having different weights,
- 2) The measurement variance, e.g. random errors in weighing each tablet.

- **sampling variance,  $\sigma_1^2$**
- **measurement variance,  $\sigma_0^2$**

The next section describes how these two contributions can be separated and estimated by using ANOVA. For bulk materials the sampling variance is dependent on the size of the test increment relative to the scale of the inhomogeneities. As the test increment size increases, the inhomogeneities tend to be averaged out and so the sampling variance decreases.

## 4.5 Quality control methods – Introduction

If a laboratory is to produce analytical results of a quality that is acceptable to its clients, and allow it to perform well in proficiency tests or collaborative trials (see below), it is obviously essential that the results obtained in that laboratory should show excellent consistency from day to day. Checking for such consistency is complicated by the occurrence of random errors, so several statistical techniques have been developed to show whether or not time-dependent trends are occurring in the results, alongside these inevitable random errors. These are referred to as **quality control** methods.

Suppose that a laboratory uses a chromatographic method for determining the level of a pesticide in fruits. The results may be used to determine whether a large batch of fruit is acceptable or not, and their quality is thus of great importance.

The performance of the method will be checked at regular intervals by applying it, with a small number of replicate analyses, to a standard reference material (SRM), the pesticide level in which is certified by a regulatory authority.

Alternatively an internal quality control (IQC) standard of known composition and high stability can be used. The SRM or IQC standard will probably be inserted at random into the sequence of materials analysed by the laboratory, so that the IQC materials are not separately identified to the laboratory staff and are studied using exactly the same procedures as those used for the routine samples.

The known concentration of the pesticide in the SRM/IQC materials is the target value for the analysis,  $\mu_0$ . The laboratory needs to be able to stop and examine the analytical method if it seems to be giving erroneous results. On the other hand resources, time and materials will be wasted if the sequence of analyses is halted unnecessarily, so the quality control methods should allow its continued use as long as it is working satisfactorily.

If the values for the IQC samples do not show significant time-dependent trends, and if the random errors in the measurements are not too large, the analytical process is under control.

Quality control methods are also very widely used to monitor industrial processes. Again it is important to stop a process if its output falls outside certain limits, but it is equally important not to stop the process if it is working well. For example, the weights of pharmaceutical tablets coming off a production line can be monitored by taking small samples (see above) of tablets from time to time. The tablet weights are bound to fluctuate

around the target value  $\mu_0$  because of random errors, but if these random errors are not too large, and are not accompanied by time-dependent trends, the process is under control.

## 4.6 Shewhart charts for mean values

In MA4603 we learned how the mean of a sample of measurements could be used to provide an estimate of the population mean, and how the sample standard deviation,  $s$ , provided an estimate of the population standard deviation.

For a small sample size,  $n$ , the confidence limits of the mean can be computed (equation 2.9 in book) with the  $t$  value chosen according to the number of degrees of freedom ( $n - 1$ ) and the confidence level required.

The same principles can be applied to quality control work, but with one important difference. Over a long period, the population standard deviation of the pesticide level in the fruit (or, in the second example, of the tablet weights) will become known from experience.

In quality control work, the population standard deviation is given the title ***process capability***.

(Remark: In practice the usual quantile values are rounded up: 1.96 is often rounded to 2 for 95% confidence limits and 2.97 is rounded to 3 for 99.7% confidence limits.)

$$\text{For 95\% confidence limits: } \mu = \bar{x} \pm \frac{2\sigma}{\sqrt{n}} \quad (4.2)$$

$$\text{For 99.7\% confidence limits: } \mu = \bar{x} \pm \frac{3\sigma}{\sqrt{n}} \quad (4.3)$$

These equations are used in the construction of the most common type of control chart: a Shewhart chart (Figure 4.1).

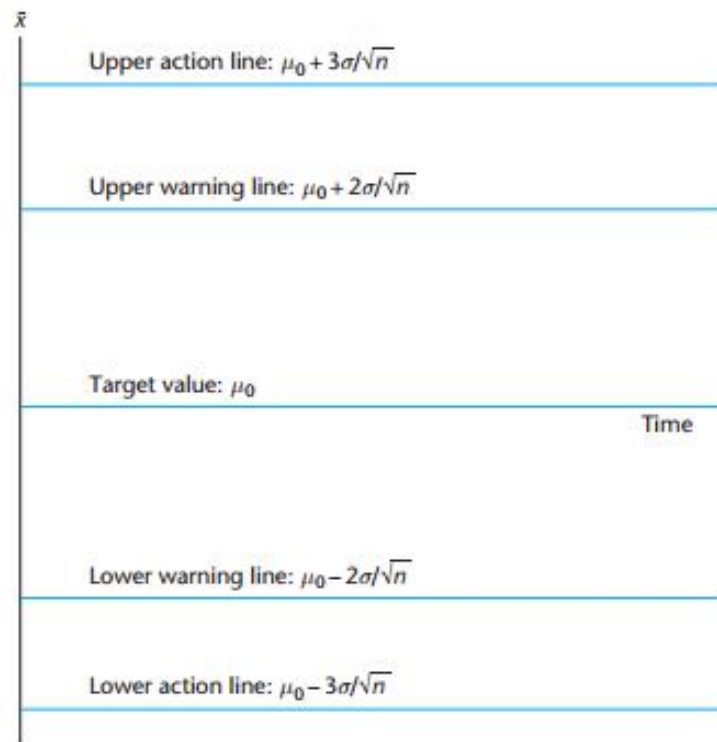


Figure 4.1 Shewhart chart for mean values.

The vertical axis of a Shewhart chart displays the process mean, ***X.bar***, of the measured values, e.g. of the pesticide concentration in the fruit, and the horizontal axis is a time axis, so that the variation of these X values with time can be plotted.

The target value,  $\mu_0$ , is marked by a horizontal line (in the middle). The chart also includes two further pairs of horizontal lines.

The horizontal lines defined using equations (4.2) and (4.3) and are called the **warning lines** and the **action lines** respectively. The purpose of these lines is indicated by their names.

Suppose a measured ***X.bar*** value falls outside the action lines. The probability of such an occurrence when the process is in control is known to be only 0.3%, i.e. 0.003, so in practice the process is usually stopped and examined if this occurs.

There is a probability of ca. 5% (0.05) of a single point falling outside either warning line (but within the action lines) while the process remains in control.

This alone would not cause the process to be stopped, but if two successive points fall outside the same warning line, the probability of such an occurrence (Prob=0.00125 in total for both warning lines) is again so low that the process is judged to be out of control.

These two criteria - one point outside the action lines, or two successive points outside the same warning line - are the ones most commonly applied in the interpretation of Shewhart charts.

Others are often used in addition: for example the probability of eight successive points lying on one specific side of the target value line is clearly low, i.e.  $0.5^8=0.0039$  ( i.e. 8 above average values in a row) and such an occurrence again suggests that the process is out of control.

Provision can also be made for stopping a process in cases where the plotted ***X.bar*** values show a trend (e.g. six increasing or decreasing points in succession, even if the points are within the warning lines), or where they seem to oscillate (e.g. 14 successive points, alternating up and down).

Users of control charts must establish clearly all the criteria to be used in declaring their process out of control.

#### 4.7 Shewhart charts for ranges

If a Shewhart chart for mean values suggests that a process is out of control, there are two possible explanations.

The most obvious is that the process mean has changed: the detection of such changes is the main reason for using control charts in which mean value ***X.bar*** values are plotted. An alternative explanation is that the process mean has remained unchanged but that the variation in the process has increased, i.e. that the action and warning lines are too close together, giving rise to indications that changes in ***X.bar*** have occurred when in fact they have not. Errors of the opposite kind are also possible, if the variability of the process has diminished (i.e. improved), then the action and warning lines will be too far apart, perhaps allowing real changes in ***X.bar*** to go undetected.

Therefore we must monitor the variability of the process as well as its mean value. This monitoring also has its own intrinsic value: the variability of a process or an analysis is one measure of its quality, and in the laboratory situation is directly linked to the repeatability (within—laboratory standard deviation) of the method.

The variability of a process can be displayed by plotting another Shewhart chart to display the **range, *R* (= highest value — lowest value)**, of each of the samples taken.



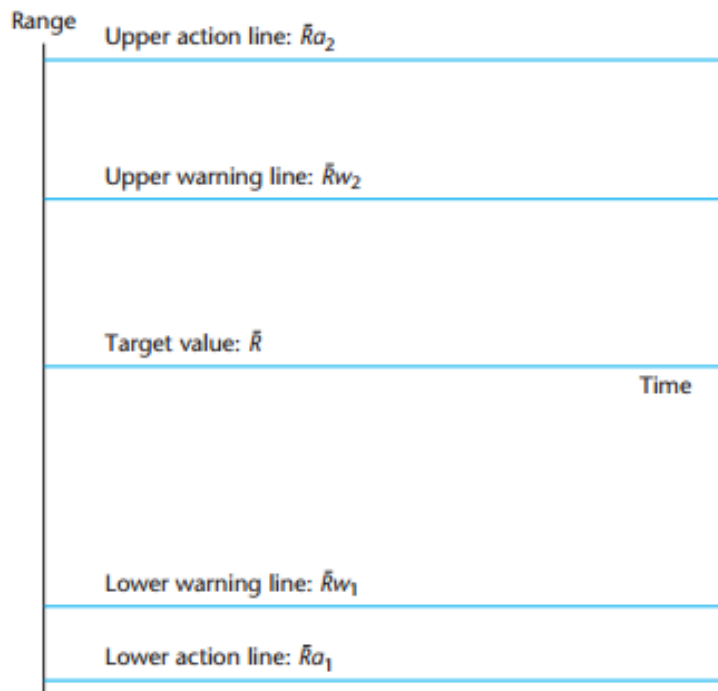


Figure 4.2 Shewhart chart for range.

A typical control chart for the range is shown in Figure 4.2. The general format of the chart is the same as is used in plotting mean values, with a line representing the target value, and also pairs of action and warning lines.

The most striking difference between the two charts is that these pairs of lines are not symmetrical with respect to the target value for the range, ***R.bar***. The value of ***R.bar*** can be calculated using the value of  $\sigma$  and the positions of the action and warning lines can be derived from ***R.bar***, using multiplying factors obtained from statistical tables.

These factors take values depending on the sample size,  $n$ .

The relevant equations are:

$$\bar{R} = \sigma d_1 \quad (4.4)$$

$$\text{Lower warning line} = \bar{R}w_1 \quad (4.5)$$

$$\text{Upper warning line} = \bar{R}w_2 \quad (4.6)$$

$$\text{Lower action line} = \bar{R}a_1 \quad (4.7)$$

$$\text{Upper action line} = \bar{R}a_2 \quad (4.8)$$

### Example 4.7.1

Determine the characteristics of the mean and range control charts for a process in which the target value is 57, the process capability is 5, and the sample size is 4.

For the control chart on which mean values will be plotted, the calculation is simple. The warning lines will be at  $57 \pm 2 \times 5/\sqrt{4}$ , i.e. at  $57 \pm 5$ ; and the action lines will be at  $57 \pm 3 \times 5/\sqrt{4}$ , i.e. at  $57 \pm 7.5$ . This chart is shown in Figure 4.3(a).

For the control chart on which ranges are plotted, we must first calculate  $\bar{R}$  using equation (4.4). This gives  $\bar{R} = 5 \times 2.059 = 10.29$ , where the  $d_1$  value of 2.059 is taken from statistical tables for  $n = 4$ . (See for example the table in the collection by Neave, the details of which are given in the Bibliography for Chapter 1.) The value of  $\bar{R}$  is then used to determine the lower and upper warning and action lines using equations (4.5)–(4.8). The values of  $w_1$ ,  $w_2$ ,  $a_1$  and  $a_2$  for  $n = 4$  are 0.29, 1.94, 0.10 and 2.58 respectively, giving on multiplication by 10.29 positions for the four lines of 2.98, 19.96, 1.03 and 26.55 respectively. These lines are shown in Figure 4.3(b).

It is not always the practice to plot the lower action and warning lines on a control chart for the range, as a reduction in the range is not normally a cause for concern.

However, as already noted, the variability of a process is one measure of its quality, and a reduction in ***R.bar*** represents an improvement in quality, the causes of which may be well worth investigating.

So plotting both sets of warning and action lines is recommended.

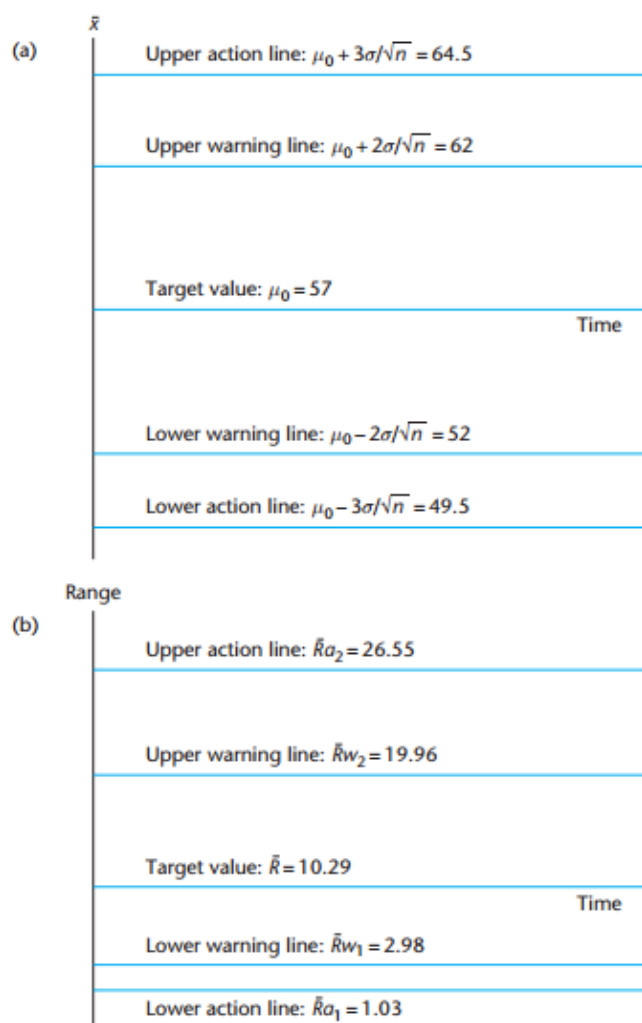


Figure 4.3 (a) Shewhart chart for mean values (example). (b) Shewhart chart for range (example).