1

Theory of Statistical Process Control

The most important tool in statistical process control is the control chart. Shewhart developed the first type of chart during the 1920s [1]. One of the most commonly used Shewhart charts is the $\overline{X}$ chart. In the following for illustration, we will use this chart. However, the principles reviewed may be broadened without much effort to include control charts in general.

Originally control charts were developed in order to solve industrial problems. We will start with pharmaceutical examples followed by applications within the health care sector. The analogies and differences between industrial and healthcare problems are also discussed.

1.1 STATISTICAL FOUNDATION OF CONTROL CHARTS

To characterise a process, products produced by the process may be sampled. A process variable is a random variable which value is obtained by observing or measuring a specified property of each product, produced by the process and reflecting its quality. A sample variable is also a random variable. However, its value is calculated as a function of the process variable values, measured in the sample.

Statistical Development of Quality in Medicine P. Winkel and N. F. Zhang

Sample values are used to construct a control chart. They are subsequently plotted on the chart to monitor the process.

1.1.1 Statistical Control

Statistical control is a concept fundamental to the theory of control charts. It is based on a distinction between two types of variation: one resulting from unavoidable causes, which one cannot identify (random variation), and one resulting from causes, which may be identified (assignable causes of variation). A process which sample values vary due to random causes alone is said to be in a state of statistical control. Additional variation caused by assignable causes may occur. If this is the case, the process is said to be out of statistical control. Since these causes may be identified, it is often possible to regulate and control them so that the process may be brought back into a state of statistical control.

Although the causes of variation of sample values from a process in statistical control cannot be identified, the type and extent of the variation may be described using large volumes of data. In other words, the values may be described approximately by a probability distribution. The parameters of this distribution characterise the state of the process. Information about this probability distribution may be obtained from random samples selected from the process while it is in statistical control.

1.1.2 Samples and Control Charts

Assume we are examining the production process for a pharmaceutical product (e.g., tablets) that is in statistical control. The machine producing the tablets has been adjusted to produce tablets with a weight that follows a Gaussian distribution with a mean of 63.000 mg and a standard deviation of 0.010 mg. Samples, each comprising one tablet, are selected from the production batch and their weights measured. We assume the error of measurement is negligible and that the machine is functioning as anticipated. Therefore, the results of the measurements follow a Gaussian distribution with mean 63.000 mg and standard deviation 0.010 mg. In the long run, we expect 99.73 % of the results to fall within an interval with its upper limit equal to the mean plus three standard deviations and the lower one equal to the mean minus three standard deviations, i.e., an interval between 62.970 mg and 63.030 mg (see Appendix A, Example A.11). The distribution with these limits entered is depicted in Figure 1.1 a.

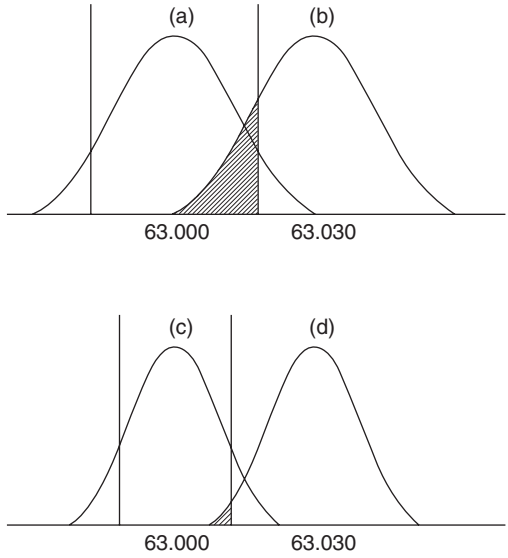

Figure 1.1 The distribution of the sample mean of the weight of tablets before ((a) and (c)) and after ((b) and (d)) the adjustment of a tablet-producing machine has been changed by 0.030 mg. Figures (a) and (b) show the distributions for sample size = 1 and figures (c) and (d) the distributions when the sample size = 2.

Had we selected samples comprising not one, but n ($n > 1$) tablets, it would have been natural to calculate the mean of the n measurement results if we wanted to follow the mean value of the process. The distribution of the weight is Gaussian. The mean of the results of n measurements generated by this distribution also follows a Gaussian distribution with the same mean, but with a standard deviation of $\frac{\sigma}{\sqrt{n}}$ (see Appendix A, Equation (A.23)). We may, therefore, calculate an interval within which 99.73 % of the sample means will fall in the long run. The lower limit of this interval is $\mu - \frac{3\sigma}{\sqrt{n}}$, and its upper limit is $\mu + \frac{3\sigma}{\sqrt{n}}$, i.e., $63.000 - \frac{3 \cdot 0.010}{\sqrt{n}}$ mg and $63.000 + \frac{3 \cdot 0.010}{\sqrt{n}}$ mg, respectively.

Figure 1.1 (a) depicts the distribution with these limits calculated for $n = 1$, and Figure 1.1 (c) depicts the distribution with the limits calculated for $n = 2$. Both of the intervals include 99.73 % of all values. However, the interval for the mean values ($n = 2$) is slimmer than that for the single values ($n = 1$) because it has a smaller standard deviation.

Now, assume that the machine is adjusted so that the mean value of the weight of tablets is increased by 0.030 mg. It will continue to produce tablets, the weights of which follow a Gaussian distribution with

standard deviation 0.010 mg. However, the mean has increased to 63.030 mg. The distribution of single values (see Figure 1.1 (b)) as well as the distribution of sample means (see Figure 1.1 (d)) will change. In both cases the distribution will be horizontally shifted towards the right so that the mean value now will be 63.030 mg instead of 63.000 mg. After the mean value of the process has changed, a large proportion of the sample values (single values for $n = 1$ and mean values for $n = 2$) will fall outside the upper limit in both cases. However, some of them will still fall within the two limits (the control limits). The proportion falling within the control limits will be larger when the sample size is 1 than when it is 2. This is so because the two distributions before and after the shift of the mean value of the process are slimmer and therefore better separated when the sample size is 2 than when it is 1.

We will now construct a control chart. To do so we rotate Figure 1.1 (c) 90° counter clockwise and draw four horizontal lines passing through zero, the lower control limit, the process mean, and the upper control limit, respectively. The line passing through zero is used to indicate the time or the order of the samples. The result is depicted in Figure 1.2.

The fraction of a Gaussian distribution, with mean μ and standard deviation σ, that is delimited by the values $\mu \pm 3\sigma$, is 99.73 % and the remaining fraction located outside the interval is 100 % – 99.73 % = 0.27 %. Therefore, the probability that a sample mean falls outside the $(\mu \pm \frac{3\sigma}{\sqrt{n}})$ limits of a control chart is 0.27 %, as long as the process remains in statistical control. Each time we select a sample, we test the null hypothesis that the mean value of the process has not changed by checking if the sample mean falls within or outside the control limits given above. The level of significance of this test is 100 % – 99.73 % = 0.27 %. It follows that the control chart may be used repeatedly to test the hypothesis that the process is in statistical control. It is implicitly assumed that the value of σ never changes.

Example 1.1

A sample comprising five tablets is selected each day from a process producing tablets, and the weight of each tablet is measured. The mean value of the results of the measurements is calculated. The mean value and standard deviation of the process are known to be 63.000 mg and 0.010 mg, respectively. We want to construct a control chart with control limits equal to the mean ±3 standard deviations of the sample mean. Because the

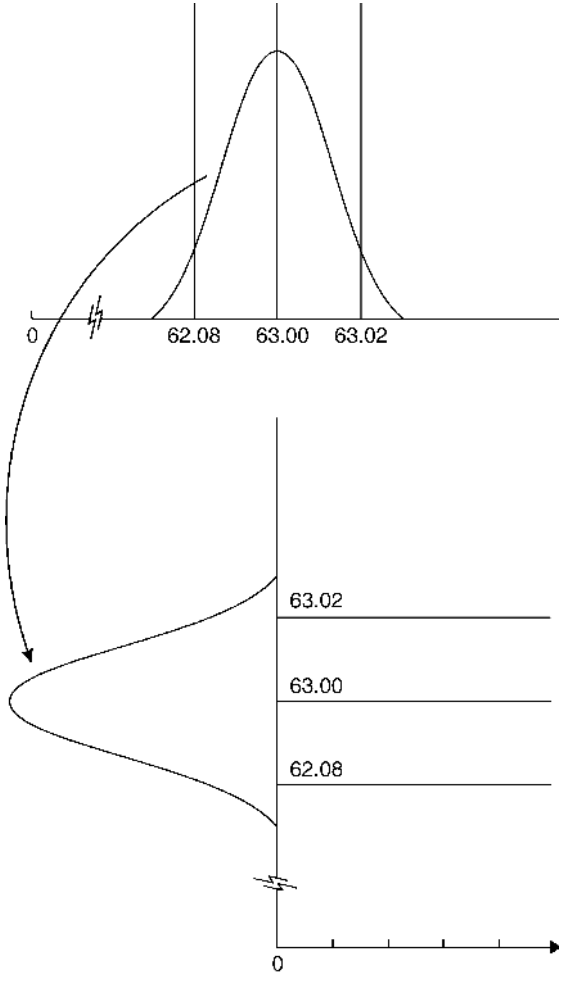

Figure 1.2 The construction of a control chart based on the distribution shown in Figure 1.1 (c). The distribution has been rotated 90 degrees counter clockwise. The line corresponding to the mean value is the centreline, and the lines corresponding to the limits of the 99.73 % confidence interval are the control limits of the chart.

sample size, n, is 5, the standard deviation of $\overline{x}$ (the sample mean) is: $\frac{0.010}{\sqrt{5}} = 0.0045$ mg (see Appendix A, Equation (A.23)). The mean of the distribution of sample means is the same as that of the process, i.e., 63.000 mg. The centreline of the chart, therefore, is at 63.000 mg. The upper and lower control limits are $63.000 + 3 \cdot 0.0045 = 63.035$ mg and $63.000 - 3 \cdot 0.0045 = 62.986$ mg, respectively.

In the above it has been assumed that the process has been so well described that its parameters may be considered known. As a rule this is not so, and we have to use estimates of the parameters when constructing a control chart.

1.2 USE OF CONTROL CHARTS

Initially, when one constructs a control chart, it is usually not known if the process is in statistical control. In the initial phase the goal is to reduce the variation of the process until it reaches a state of statistical control that is acceptable. To assess if a process is in statistical control, one often uses 20 to 25 samples, each comprising 4 to 5 observations. When the samples are collected, one should record those conditions that might possibly create variation in addition to the random variation. This could be, e.g., the temperature, the raw materials used, the identity of operators, etc. The average of the individual sample means ($\hat{\mu}$) is used as an estimate of the mean of the process variable. It defines the location of the centreline. An unbiased estimate of the process standard deviation ($\hat{\sigma}$) is calculated from the average of the standard deviations (s_i) of the individual samples ($\bar{s}$) divided by a factor (c_4), which depends of the sample size and is found using Table 1.1.

We have

$$\hat{\mu} = \frac{\sum_{i=1}^{k} \bar{x}_i}{k} \tag{1.1}$$

where k is the number of samples and

$$\hat{\sigma} = \frac{\bar{s}}{c_4} = \frac{\sum_{i=1}^{k} s_i}{k \cdot c_4} \tag{1.2}$$

The upper control limit (UCL) is calculated as

$$\text{UCL} = \hat{\mu} + 3\frac{\hat{\sigma}}{\sqrt{n}} \tag{1.3}$$

and the lower control limit (LCL) as

$$\text{LCL} = \hat{\mu} - 3\frac{\hat{\sigma}}{\sqrt{n}} \tag{1.4}$$

Table 1.1 Factors used for $\overline{X}$ charts and/or S charts.

Sample size	Factors				
n	c_4	B_3	B_4	B_5	B_6
2	0.7979	0.000	3.267	0.000	2.606
3	0.8862	0.000	2.568	0.000	2.276
4	0.9213	0.000	2.266	0.000	2.088
5	0.9400	0.000	2.089	0.000	1.964
6	0.9515	0.030	1.970	0.029	1.874
7	0.9594	0.118	1.882	0.113	1.806
8	0.9650	0.185	1.815	0.179	1.751
9	0.9693	0.239	1.761	0.232	1.707
10	0.9727	0.284	1.716	0.276	1.669
11	0.9754	0.321	1.679	0.313	1.637
12	0.9776	0.354	1.646	0.346	1.610
13	0.9794	0.382	1.618	0.374	1.585
14	0.9810	0.406	1.594	0.399	1.563
15	0.9823	0.428	1.572	0.421	1.544
16	0.9835	0.448	1.552	0.440	1.526
17	0.9845	0.466	1.534	0.458	1.511
18	0.9854	0.482	1.518	0.475	1.496
19	0.9862	0.497	1.503	0.490	1.483
20	0.9869	0.510	1.490	0.504	1.470
21	0.9876	0.523	1.477	0.516	1.459
22	0.9882	0.534	1.466	0.528	1.448
23	0.9887	0.545	1.455	0.539	1.438
24	0.9892	0.555	1.445	0.549	1.429
25	0.9896	0.565	1.435	0.559	1.420

For $n > 25$, $c_4 \approx \dfrac{4(n-1)}{4n-3}$

where n is the sample size. Finally, the individual sample mean values are depicted on the chart. In the case where all points lie within the control limits, the data are consistent with the hypothesis that the process is in statistical control. If one or more points are located outside the limits, it is an indication that the process is not in statistical control, and the causes must be traced. In the cases where these causes are identified, the corresponding values are eliminated from the calculations, and a revised control chart is computed. It is now controlled if all of the remaining points fall within the revised control limits. Since the revised control limits are narrower than the original ones, data points that previously

fell within the original limits may now fall outside the revised limits. The cause, why a point fell outside the limits, may not necessarily be found. If this is the case for only one or few points, one may choose not to remove the values immediately, but wait and see how the control chart functions and eventually remove them later on. If a large number of points are falling outside the limits for unknown reasons, the pattern formed by the points should be inspected. By doing so, one may often be able to identify a cause common to all points. After a while, hopefully, the chart indicates that a state consistent with the hypothesis of statistical control has been reached. If the level of the process and the variation relative to this level are both acceptable, the chart specifies the objective of the process.

At this stage, it is vitally important that a protocol is written specifying how one should go about looking for special causes if a value falls outside the control limits, and how the report resulting from such a search should be made. The specifics of the protocol depend on the process. For a good clinical example see [2]. The chart may, then, be used to monitor regularly selected samples, the mean values of which are depicted on the chart. As long as these values are located within the control limits, one may assume that the process is in statistical control. The data cumulated in this way may be used to calculate relatively precise estimates of the parameters of the process. When reliable estimates are available, one may determine if the process actually meets the quality requirements. If this is not the case, it is advisable to revise the process, i.e., to improve it, until it meets the demands. In this phase statistical design of exploratory experiments is an important tool. A review of these techniques, however, is outside the scope of this book. The interested reader is referred to [3].

Example 1.2

At an outpatient clinic the management decided to take random samples comprising 30 ambulatory patients on each weekday for four weeks to study the patient waiting times and assess if the quality requirement for patient waiting times was met. The employees at the clinic were not aware of this investigation. The waiting time of each patient was recorded. A patient's waiting time is the period starting when the patient arrives at the clinic and ending when a technologist sees the patient. Thus, 20 samples each comprising 30 randomly selected waiting times were recorded.

Table 1.2 The mean and standard deviation of waiting times (minutes from patient's arrival at outpatient clinic until seen by a technologist) recorded on each of 20 weekdays.

Sample #	n	$\bar{x}$	$\bar{s}$
1	30	16.75	5.509
2	30	15.60	4.558
3	30	16.14	5.465
4	30	15.96	4.582
5	30	18.86	4.594
6	30	14.33	4.920
7	30	15.44	6.357
8	30	14.67	3.791
9	30	16.53	6.885
10	30	19.89	5.583
11	30	14.37	3.714
12	30	14.13	3.477
13	30	14.99	4.627
14	30	13.33	3.922
15	30	19.96	4.717
16	30	15.87	5.481
17	30	14.41	5.877
18	30	15.16	4.901
19	30	13.82	5.434
20	30	18.46	3.716
		$\hat{\mu} = 15.93$	$\bar{s} = 4.905$ $\hat{\sigma} = 4.95$

Table 1.2 shows the 20 sample mean values and standard deviations. The mean of the mean values ($\hat{\mu} = 15.93$ minute) estimating the process mean and the average of the standard deviations ($\bar{s} = 4.905$ minute) are also shown. An unbiased estimate of the process standard deviation ($\hat{\sigma}$) is calculated by dividing $\bar{s}$ by c_4. The latter quantity is calculated as $\frac{4(30-1)}{120-3} = 0.9915$ (see Table 1.1). Therefore, $\hat{\sigma}$ is $\frac{4.905}{0.9915} = 4.95$ minute. Using these data an $\overline{X}$ control chart may be constructed. The estimate of the standard deviation of the sample mean values is $\frac{4.95}{\sqrt{30}} = 0.90$ minute since the sample size is 30. The centreline of the $\overline{X}$ chart is at 15.9 minute ($\hat{\mu}$), the UCL is $15.9 + 3 \cdot 0.90 = 18.6$ minute, and the LCL is $15.9 - 3 \cdot 0.90 = 13.2$ minute.

Figure 1.3 (a) shows the $\overline{X}$ chart. The sample means are depicted on the chart. Since three of the values (samples # 5, # 10, and # 15) are located above the UCL, the process is not in statistical control.

Figure 1.3 (b) shows a revised control chart after these three values have been eliminated from the calculations. Now the last value is outside the UCL. Figure 1.3 (c) shows the control chart calculated without using

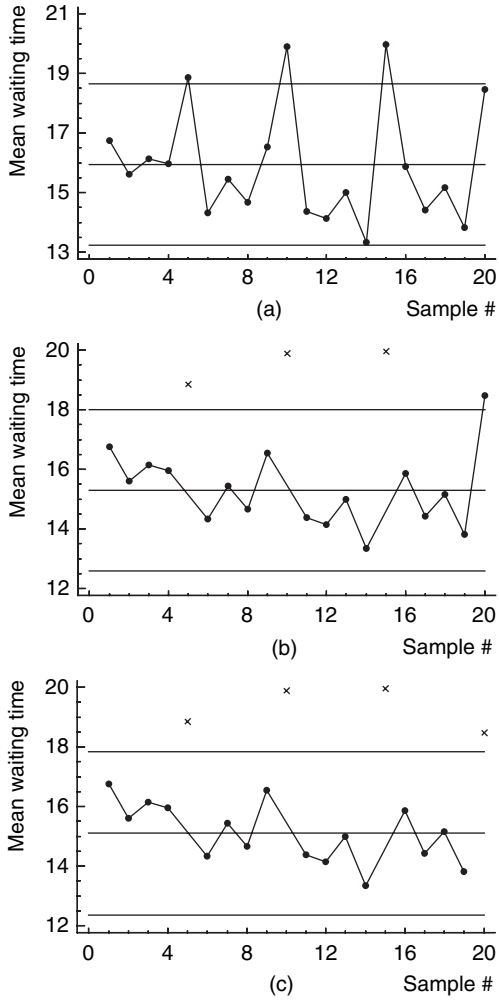

Figure 1.3 (a) An $\overline{X}$ chart showing the mean waiting times (minute) on weekdays. (b) The control chart shown in (a) after the results of samples # 5, # 10, and # 15 have been removed and the chart calculated without using these values. The latter values are depicted as crosses. (c) The control chart shown in (a), after the results of samples # 5, # 10, # 15, and # 20 have been removed and the chart calculated without using these values. The latter values are depicted as crosses.

this value. The chart is consistent with a process in statistical control. The management now had two jobs. First the special cause of the excess variation had to be found and removed and the process brought into a state of statistical control. Then the estimated process parameters had to be compared with the quality requirements for waiting times to assess the quality of the process.

Inspecting the pattern of values (# 5, # 10, # 15, and # 20) outside the UCL, one notes that all values were collected on a Friday. It turned out that on Fridays the patient mix differed from that of the other weekdays in that an unusually large number of patients from the cardiology department were scheduled for ECG recordings in addition to blood specimen collection. These patients required more time than those not scheduled for ECG recordings. A retrospective analysis revealed that overtime was much more common on Fridays than on any other weekday. To prepare the organisation for quality assessment and control on a routine basis, the management purchased a system for automatic recording of waiting times. The IDs of technologists doing venipunctures and recording ECGs were already automatically captured by the current clinical data processing system. We will return to this example in Chapter 9.

Table 1.3 shows the protocol they designed for the search for special causes to be used when the process was brought into a state of statistical control of a sufficiently high quality.

Table 1.3 Procedure for tracking special cause variation. Start at step 0 and proceed to the right.

Step	Actions	Questions and routing in table	
0	Control data and data processing.	Can error in data explain variation?	If yes, go to step 4. If no, go to step 1.
1	Compute ECG in number of venipuncture equivalents. Express production in number of venipunctures. Control staffing of ambulatory.	Can increased production or decreased staffing explain variation?	If no go to step 2. If yes, go to step 4.
2	Define productivity as venipuncture/technologist hour. Compute 1) average productivity, 2) average productivity per 30 minute period, 3) average productivity per technologist, and 4) average productivity per technologist per 30 minute period. Identify significantly outlying values.		Go to step 3.
3	Interview manager of ambulatory and technologists.		Go to step 4.
4	Write report and stop.		

1.3 DESIGN OF CONTROL CHARTS

When one constructs a control chart, it is necessary to decide which control limits, sample size, and sampling frequency one wants to use.

1.3.1 Control Limits

Clearly, the position of the control limits has a bearing on the function of a control chart. The further away from the centreline the limits are located, the fewer sample means will fall outside the limits. This implies that the probability that a type-1 error will be committed declines. A type-1 error is committed when a sample mean value falls outside the control limits even though the process is in statistical control. It is then assumed that the process is out of statistical control, and a search for the cause is initiated. The price of decreasing the expected number of type-1 errors, by widening the limits, is an increase in the expected number of type-2 errors. A type-2 error is committed if a sample mean value falls within the control limits and thereby prevents one from acknowledging that the process is no longer in statistical control. If one narrows the limits, the probability of committing a type-2 error declines, but at the same time that of committing a type-1 error increases.

The choice of control limits depends on a weighing of the pros and cons of the two types of errors. One approach is to decide initially how large a fraction (α) of the sample means one is willing to let fall outside the limits, while the process is in statistical control. The position of the control limits is then calculated so that this condition is fulfilled. If the fraction of values falling outside the limits is α, the fraction of values falling within the limits must be $(1-\alpha)$. In the case of the $\overline{X}$ chart, the problem may be stated as follows: We need to find a number, k, so that the probability that a sample mean will fall inside the control limits is

$$P\left(\mu - \frac{k\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + \frac{k\sigma}{\sqrt{n}}\right) = 1 - \alpha \tag{1.5}$$

where μ is the mean of the process, σ its standard deviation, $\overline{X}$ the sample mean, and n the sample size. It is not particularly difficult to find k in Equation (1.5) when the distribution of $\overline{X}$ is Gaussian and μ and σ are both known. When the value of k has been determined in this way, the UCL is set equal to $\mu + \frac{k\sigma}{\sqrt{n}}$, and the LCL is set equal to $\mu - \frac{k\sigma}{\sqrt{n}}$. Usually $k = 3$ is used. Inserting this value in Equation (1.5) the corresponding value of α may be calculated. One finds that $\alpha = 0.0027$. Therefore, in the long run, $(1 - 0.0027)100\,\% = 99.73\,\%$ of the sample mean values will

fall within the limits, as long as the process remains in statistical control. In the following we will use the factor 3, when calculating control limits. It is assumed that the sample mean follows a Gaussian distribution and the parameters are known. Due to the central limit theorem (see Appendix A, Section A.3.4.1), the assumption of a Gaussian distribution is not necessary when the sample size is large enough.

In addition to the control limits, two warning limits are sometimes used, one on each side of the centreline usually at a distance of two standard deviations from it. If a sample value falls between a warning limit and the corresponding control limit, then it is a warning that the process may be out of statistical control.

It is not always safe to assume that a sample variable follows a Gaussian distribution, as we have done previously. However, the consequences of erroneously making this assumption are limited. This appears from an improvement on Tchebichev's inequality [4], which may be phrased as follows: If the statistical variable X follows a unimodal distribution whose mode is equal to the mean, the probability that its value deviates from the distribution's mean value by more than k times its standard deviation is equal to or less than $\frac{1}{2.25k^2}$. A unimodal distribution is defined as a probability distribution which density function decreases monotonously to the left, as well as the right of its mode (see Appendix A, Section A.2.3). Based on the inequality above, for $k = 3$ the probability that a sample mean falls outside the control limits is $\frac{1}{2.25 \cdot 3^2} = 0.049$, or less, as long as the distribution of the sample mean is unimodal, and its mode and mean coincide.

1.3.2 Sample Size

If a process gets out of statistical control because its level is changing, the probability that a sample mean assumes a value outside the limits whereby the change will be acknowledged increases. The increase in probability depends on the sample size, as illustrated in Figure 1.1 in Section 1.1.2. The figure shows the distribution of the sample value before and after the mean has changed for sample size equal to 1 (upper frame) and for sample size equal to 2 (lower frame). When the sample size is 1, a much smaller fraction of the horizontally shifted distribution falls outside the control limits of the original distribution than in the case when the sample size is 2. Therefore, the probability (the area outside the control limits) that a specified change in the mean is acknowledged is smaller with sample size of 1 than with 2.

1.3.3 Sampling Frequency

The more frequently samples are selected, the sooner a change of the level of the process will be acknowledged; but also, the more frequently a sample mean value will fall outside the control limits while the process is in statistical control.

1.4 RATIONAL SAMPLES

A key issue to the construction of control charts is the formation of rational samples. Rational samples are composed so that assignable causes of variation may influence the variation between samples, but not the variation within samples. For instance, it will be inexpedient to mix the blood smears from two different technologists in the same sample because variation caused by differences between the two technologists will then not be acknowledged. The formation of rational samples is crucial and often requires considerable knowledge about the process in question.

If a process is in statistical control, all of the variation between the sample mean values can be explained by the variation within the samples because the process mean does not change. Therefore, the samples may be pooled and all be used to calculate a single estimate of the standard deviation. However, if the process is not in statistical control, the process mean may change between samples. In this case the standard deviation obtained by pooling the values will be larger than the standard deviation that one would obtain by calculating the average of the within-samples standard deviations. Since it is not known in advance whether a process is in statistical control, the average of the sample standard deviations should always be used. Otherwise the risk is that a lack of statistical control may be masked.

1.5 ANALYSING THE PROPERTIES OF A CONTROL CHART

Once a process is brought into a state of statistical control, a control chart may be used to monitor it. The purpose is to recognise quickly and in an objective way if the process gets out of statistical control. If a sample value falls outside the control limits, it is a very strong

indication that this has happened. However, the values may remain within the control limits, while the process is out of control. A state of statistical control is characterised by control values scattering at random around the centreline. Therefore, if systematic sequences of values begin to appear, it may be an indication that the process is out of control, even though all values stay within the limits. A sequence of values that all share the same quality is referred to as a run. Eight consecutive sample mean values, each being larger than its predecessor, or eight values all located on the same side of the centreline are examples of runs.

To better characterise the values entered on the chart, it is customary to enter two warning limits; a lower one and an upper one. As mentioned previously, each of them is located at a distance of two standard deviations from the centreline. The probability that a sample mean value falls outside the warning limits is approximately 0.05 if the process is in control. However, the probability that two mean values in a row fall outside the upper warning limit is quite low. Therefore, this indicates strongly that the mean value of the process has increased. The two warning limits may be supplemented by an additional pair of warning limits located on each side of the centreline, each at a distance of one standard deviation from the line. In this way, the region defined by the two control limits is divided into 6 zones. This makes it easier to recognise interesting runs, e.g., a run characterised by values located above the same inner warning limit.

1.5.1 Systematic Data Patterns

We may test statistically if it is improbable that a run is just a random phenomenon. Then, it may be concluded that the process is out of statistical control. However, if one applies several tests simultaneously, i.e., pays attention to many different types of runs, the combined probability of committing a type-1 error may be quite high. Therefore, it is not recommendable routinely to include tests based on various types of runs when assessing whether a process is in statistical control or not. Small changes of the mean level may certainly cause various types of runs to appear while all data points are still falling within the control limits. However, to identify small changes in the level, it is recommended instead to apply a time-weighted control chart. These control charts are reviewed in Chapter 3. This is not to say that one should not pay attention to extreme patterns and utilise the information thus gained. When a value

falls outside the control limits, statistically significant patterns of runs may be valuable clues in the search for the cause of the loss of control.

1.5.2 In-Control Average Run Length (ARL) and Out-of-Control ARL

In the following a process is considered to be in statistical control as long as the sample mean value stays within the control limits and out of control if a value falls outside the control limits. This usage of the chart implies that one is currently testing the hypothesis that the process is in control. Since the conclusions drawn on the basis of a statistical test result are inherently uncertain, one will occasionally commit an error. It is of practical interest to know how often one should expect a false alarm and thereby be led to commit a type-1 error. It is furthermore of interest to know the length of the period, from a loss of statistical control until the loss is acknowledged. The last question cannot be answered unequivocally unless it is specified how much out of control the process is in terms of the magnitude of the change of its parameter values.

The probability distribution of the sample means after the mean value of the process has changed is Gaussian with a mean equal to the new process mean value and a standard deviation equal to that characterising the distribution prior to the change. Using this information, one may calculate the fraction of the distribution delimited by the control limits (see Figure 1.1). This fraction is equal to the probability (β) that a sample mean will fall within the control limits and thereby prevent the change from being acknowledged. The probability that the mean value of the first sample selected – subsequent to a specified change of the process mean – falls outside the control limits is $1 - \beta$. This follows because β is the probability that it falls within the limits. The probability that the change will be acknowledged when the second sample is selected, is the probability (β) that it will not be acknowledged when the first sample is selected multiplied by the probability $(1 - \beta)$ that it will be when the second sample is selected, i.e., $\beta(1 - \beta)$. It is assumed that the sample values are statistically independent. The probability that the change will be acknowledged when the fifth sample is selected is $\beta^4(1 - \beta)$, etc. In general we have, that the probability that the change will be acknowledged at the kth trial is $\beta^{k-1}(1 - \beta)$. In the first example, it was necessary to obtain 2 samples before the change was acknowledged, and in the second example it was necessary to obtain 5 samples. To calculate the average number of samples necessary to select before a specified change

is acknowledged, each possible outcome should be weighted by its probability of taking place and the resulting products added. In principle, there are an infinite number of possible outcomes, and the sum, therefore, includes an infinite number of terms. It may be shown that the sum is $\frac{1}{1-\beta}$. We have

$$\sum_{k=1}^{\infty} k\beta^{k-1}(1-\beta) = \frac{1}{1-\beta} \tag{1.6}$$

In the special case when the process mean has not changed, $(1-\beta)$, the probability that a value falls outside the control limits is equal to α, the risk of committing a type-1 error. The average number of samples collected (ARL, the average run length) between type-1 errors is called the in-control ARL. According to Equation (1.6) it is calculated as

$$\text{ARL}_{\alpha} = \frac{1}{\alpha} \tag{1.7}$$

Example 1.3

Using the conventional control limits equal to the process mean ± 3 standard deviations implies that $\alpha = 0.0027$. Therefore, the $\text{ARL}_{0.0027}$ between type-1 errors is $\frac{1}{0.0027} = 370.37$, according to Equation (1.7).

When β is known, Equation (1.6) may be used to calculate the average number of samples selected subsequent to a specified change of the process mean and before the first value falls outside the limits and the change thereby is acknowledged. We have

$$\text{ARL}_{\beta} = \frac{1}{1-\beta} \tag{1.8}$$

1.6 CHECKLISTS AND PARETO CHARTS

Two helpful instruments may supplement a control chart: a checklist and a Pareto chart. A checklist is used to list in a chronological order the problems that one has come across so far while monitoring the process. It should include information about how often the various flaws and defects have been observed and who took care of them. A Pareto chart

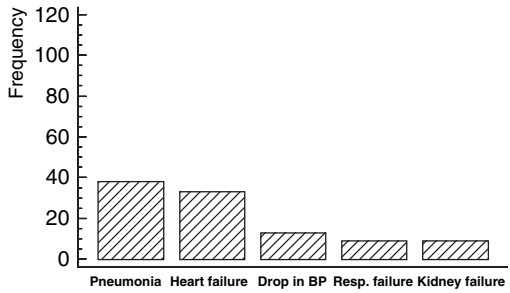

Figure 1.4 A Pareto chart showing the occurrence of postoperative medical complications in 102 surgical patients. BP stands for blood pressure, and Resp. stands for respiratory.

is in many ways similar to a histogram. The ordinate of a Pareto chart is the same, i.e., frequency. The abscissa, however, is qualitative instead of quantitative. It shows the type of problems one has come across. The chart gives a graphical representation of qualitative data, with the frequencies sorted according to size. The checklist may be used as input for the Pareto chart. The chart may be used to identify frequently occurring problems. However, it does not acknowledge the seriousness of the various problems. If some problems are serious and other problems are trivial ones, one might weigh the frequencies of the various problems according to their seriousness before drawing the Pareto chart.

Example 1.4

In each of 102 surgical patients the complications arising during the operation were noted. Figure 1.4 shows a Pareto chart of the frequencies of the various types of medical complications. It appears that heart failure and infection of the lungs are the predominant medical complications.

1.7 CLINICAL APPLICATIONS OF CONTROL CHARTS

The clinical applications of control charts are often less straightforward than the industrial ones. We will discuss the problems arising. Some of the issues will be further elaborated in Chapter 6.

1.7.1 Input/Output of Clinical Processes

In principle control charts may be used in clinical work in the same way as they are used in industrial work. The values of quality measures or quality indicators are measured as mentioned in the introduction and used to construct control charts. The quality may then be assessed and monitored using the charts.

The input to industrial processes may be controlled. However, this is not always the case for medical processes. In clinical medicine the patients vary considerably. Some patients may be so sick that they will not survive even if the clinical process, i.e., treatment and care, is optimal, while other patients whose diseases are less severe may survive even though the treatment and care they receive is of a poor quality. An industrial concern is able to standardise the input to its various processes. Therefore, one may safely assume that variation of the output (the products) mirrors the quality of the processes. By contrast, a hospital department or a practice cannot control the input (the number and types of patients received). Therefore, in this case, it is necessary to separate the variation of the output into two components: one that is caused by variation of the input (variation of severity of the patients' diseases, their co-morbidities, etc.) and one that is caused by the process (treatment and care). The problem may be dealt with in various ways as will be explained in Chapters 5, 6, and 7.

1.7.2 Samples

If possible, the samples formed should be rational. Therefore, assignable causes of variation should not be allowed to influence the within sample variation, and the selection of samples should be organised so that rational hypotheses of interest may be tested. Assume, e.g., that there are reasons to believe that the waiting-time between the arrival of a test request to a laboratory and the reporting of the corresponding result is not the same during working hours, as it is outside working hours. To assess this hypothesis, it is necessary to select 'waiting-time samples' so that a sample is either selected during working hours or outside working hours and not just at random times round the clock.

Two principles may be applied to form rational samples. According to one principle a sample should only include products that are produced at the same time (or as close together in time as possible). This principle is applied when the primary purpose is to be able to acknowledge a

change in the process mean since the likelihood that such a change will affect the within-sample variation is very small. According to the other principle, the sampling period is extended. At the end of one sampling period, the next one is initiated, etc. Each sample consists of products representative of those produced since the last sample was selected. This principle is usually applied when a decision has to be made as to whether or not the products produced during the sampling period should be accepted. Supporters of the last principle often emphasise that a shift in the level, away from the control level and back again, in between sampling will not be acknowledged if the first principle is applied. However, if the process mean value fluctuates among different levels during the sampling period, the variation within the sample might be quite large. Therefore, it is possible to make any process look as if it were in control simply by lengthening the sampling period. In medicine all information about each patient is kept. Therefore, it makes sense to inspect all of the production. This implies that the second principle should be applied. Consequently, the samples should be as small as possible so that the conditions during the sampling are reasonably uniform. This also allows the search for the reason why a process is out of control to be conducted, while the trail is 'still hot'. Furthermore, actions necessary to remove the cause of a lack of statistical control will not be unduly delayed. For example we would want to detect an increase in the occurrence of Methicillin resistant *Staphylococcus aureus* infections as soon as possible to take the necessary precautions. On the other hand, the event one monitors (e.g., that a patient dies) may be a rare one. This requires the sample of patients to be made sufficiently large so that at least a few dead patients are included in each patient group. There are several considerations that one has to balance relative to each other before the sample size is decided.

Instead of using equally sized samples, it may be more practical to use equally sized sampling periods, e.g., a week, a month, or a quarter of a year. This implies that the sample size will vary. The above considerations then have to be balanced when the length of the sampling period is decided.

1.8 INAPPROPRIATE CHANGES OF A PROCESS

If the quality of a clinical process is not good enough for its purpose it has to be improved. To improve the quality of a process one has to change it. However, it is important to know when it is appropriate to change a process and when it is not. A process in statistical control may be

changed if its quality is deemed insufficient for its purpose on the basis of an assessment of its parameter values. However, it should not be changed on the basis of an assessment of a sample obtained from it. If the system is not in statistical control it should not be changed. The reason is that the results of a change cannot usually be interpreted. Instead, the process should be brought into a state of statistical control and then changed if deemed necessary.

We will present an example illustrating the effect of using sample values as a basis for changing a system that is in statistical control and examples showing the effect of changing a system that is not in statistical control.

1.8.1 Changing a Process in Statistical Control Guided by Samples

Example 1.5

Table 1.4 (column 2) shows a series of 15 numbers drawn at random from a Gaussian distribution with mean 10.00 and standard deviation 1.00.

Table 1.4 Simulation of a treatment where the dose of a drug is adjusted when the plasma concentration falls outside specified limits (8.50 to 11.50) even though the concentration without active adjustment would have remained in statistical control during the whole period.

Time (t)	Value without dose adjustment $X(t)$	Random change $R(t) = X(t) - X(t-1)$	Value with dose adjustment $Y(t) = Y(t-1) + R(t) + \text{ADJ-effect}(t)$ $(t > 1)$	Effect of dose adjustment ADJ–effect(t)
1	9.45		9.45	
2	7.99	−1.46	7.99	0.00
3	9.29	1.30	9.80	0.51
4	11.66	2.37	12.68	0.51
5	12.16	0.50	12.00	−1.18
6	10.18	−1.98	8.84	−1.18
7	8.04	−2.14	5.52	−1.18
8	11.46	3.42	11.92	2.98
9	9.20	−2.26	9.24	−0.42
10	10.34	1.14	9.96	−0.42
11	9.03	−1.31	8.23	−0.42
12	11.47	2.44	10.94	0.27
13	10.51	−0.96	10.25	0.27
14	9.40	−1.11	9.41	0.27
15	10.08	0.68	10.36	0.27

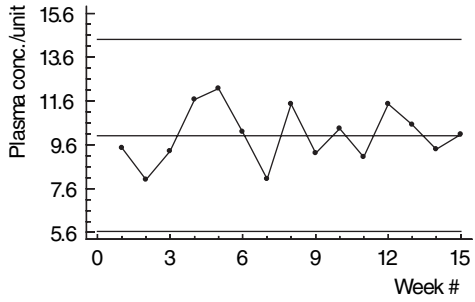

Figure 1.5 $\overline{X}$ chart, calculated from 15 random numbers generated by a Gaussian distribution with mean 10.00 and standard deviation 1.00, simulating the time course of the drug concentration measured in a patient receiving constant daily doses.

Figure 1.5 shows an $\overline{X}$ chart for $n = 1$, (see Chapter 2) constructed from these data and with the 15 values entered. As expected, the chart shows the picture one would expect when monitoring a process in statistical control the results of which follow a Gaussian distribution. The data points are all within the control limits.

Assume that the numbers represent the results of weekly measurements of the plasma concentration of a drug measured in the same patient who is taking this drug daily, without changing the dose (the basic dose). We want to calculate what would have happened if the dose, instead of having been kept constant, had been controlled by two limits, 11.50 and 8.50, as follows: when the concentration exceeds the upper limit of 11.50, the basic dose is reduced. The aim is to reduce the concentration by a quantity equal to the observed deviation from the 11.50 upper limit. In the same way, the basic dose is increased when the lower limit is exceeded.

When the dose is not adjusted, each new value is equal to the previous value plus the random biological variation that takes place between the measurements. When the dose is adjusted, the effect of the adjustment of the dose has to be added to the random biological variation. The random variation in the drug level, between measurements, is calculated as the difference between the measurements obtained when no adjustment is made. These random variations are shown in column 3 of Table 1.4. For example, the random variation from the first to the second value is $7.99 - 9.45 = -1.46$. Column 5 shows the change in the level, intended by adjustment of the basic dose according to the strategy. The values obtained when the strategy is applied and works as intended are shown in column 4. The initial value is 9.45. This is within the limits (8.50 and 11.50). So the dose is not adjusted. The second value is equal to the previous

value (9.45) plus the random variation that is −1.46 and the effect of adjustment that is 0. Therefore, the second value is 7.99. This value is 0.51 below the lower limit of 8.50. Consequently, the basic dose is adjusted to increase the concentration by 0.51. So it is assumed that everything else being equal the dose adjustment increases the concentration by 0.51 until the next time where the dose is changed relative to the basic dose. The third value is equal to 7.99 plus the random variation that is 1.30, plus the effect of the adjustment of the dose that is 0.51. Therefore, the third value is 9.80. Since the value is within the limits, the current dose is not changed. The fourth value becomes equal to $9.80 + 2.37$ (the random change) $+ 0.51 = 12.68$. Now the basic dose has to be adjusted again to achieve a change of $11.50 - 12.68 = -1.18$, etc.

Calculating the mean and standard deviation of the two series, we get mean = 10.02 and standard deviation = 1.28 for the series without dose adjustment and mean = 9.77 and standard deviation = 1.79 for the series resulting from dose adjustment. Without adjustment, the mean is 0.2% away from the intended value of 10 and the standard deviation is 28% larger than 1. With active adjustment the values are 2.3% and 79% respectively. In other words, the quality of the treatment has declined considerably.

Clearly, the example is invented and rather simple-minded. However, it illustrates a phenomenon that is well known within the field of statistical process control, namely that the quality of a process that is in a state of statistical control deteriorates if one tries to adjust it on the basis of sample values. It is necessary to assess if the process is satisfactory or not, on the basis of its parameter values. If not, it must be adjusted. One then has to wait until a new state of statistical control has been reached. Then a decision has to be made if the quality of the revised process is satisfactory, etc.

1.8.2 Changing a Process That is Not in Statistical Control

RG Carey [5] reports some very interesting examples. Using somewhat modified data, but without changing the basic ideas, we present these examples. They illustrate that the interpretation of the effect of an adjustment of a process may be very difficult if the process is not in a state of statistical control.

Example 1.6

The annual death rate of coronary artery bypass graft operations at a hospital was 5 % in 1994. The protocol for the operation was changed in

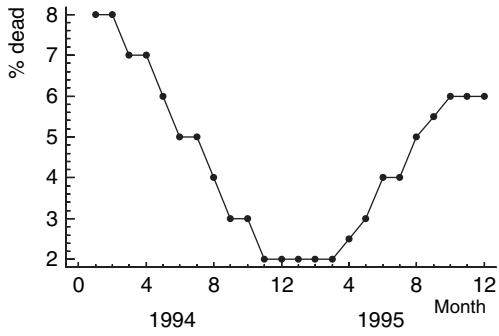

Figure 1.6 Percent of patients who died during coronary artery bypass graft operation as measured monthly during 1994 and 1995. The protocol for the operation was changed on January first 1995.

January 1995, and in January 1996 the annual death rate of 1995 was calculated and found to be 4 %. A statistical analysis comparing the annual death rates showed that this improvement was statistically significant.

Figure 1.6 shows the monthly death rates recorded during 1994 and 1995. It is obvious that for some unknown reason (perhaps improvement of the surgeons' skill to operate) the death rate has been declining throughout 1994, whereupon the trend has turned. In the beginning of 1995 the death rate dropped to 2 %, but at the end of the year it was as high as 6 %. Without examining whether the process is stable or not one may reach the conclusion that the change of protocol had a beneficial effect on the death rate. However, by examining the process, one realises that it is inappropriate to compare the annual rates because the process examined is not in a state of statistical control. In fact, inspection of the monthly rates leaves one with the impression that the change in protocol had a harmful effect.

Example 1.7

At two departments, A and B, the protocol for open-heart surgery was changed to reduce the transport time from the operating theatre to the intensive care department. The transport time is finished when the patient has stabilised and the monitoring of the patient begins. After the change in protocol had been instituted, the annual mean value of

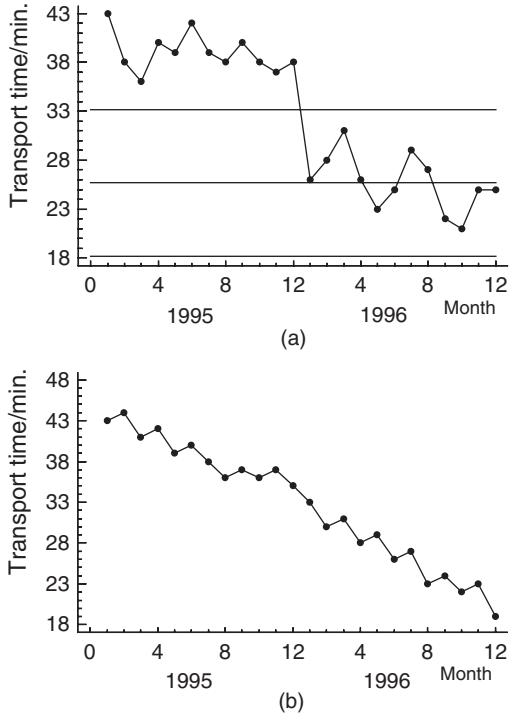

Figure 1.7 (a) The average transport time from the operating theatre to the intensive care unit as measured monthly during 1995 and 1996 in department A. The protocol for open-heart surgery was changed on January 1 1996. Only the control limits of the 1996 data are shown in the figure. (b) Data from Department B corresponding to those shown in Figure 1.7 (a) for Department A.

the transport times decreased from 39 minute to 26 minute in both departments.

Figure 1.7 (a) depicts the monthly average transport times in department A prior to and subsequent to the change in protocol. The transport time is in statistical control before as well as after the change has been introduced (only the control limits of the second period are shown in the figure). Further, the mean transport time has been reduced significantly as a result of the change. Looking at the corresponding figure for Department B, Figure 1.7 (b), one notes that the picture is completely different. Neither before nor after the introduction of the change is the process in statistical control. Furthermore, it is doubtful if the change in the protocol has had any appreciable effect on the steady decline of the monthly average transport time that started in 1995 and continued throughout 1996.

REFERENCES

[1] Shewhart WA. Quality control charts. AT&T Tech J 1926; 5:593–606.
[2] Beiles CB, Morton AP. Cumulative sum control charts for assessing performance in arterial surgery. ANZ J Surg 2004; 74:146–51.
[3] Cox DR. Planning of Experiments. John Wiley and Sons Inc, New York, 1958.
[4] Kotz S, Johnson S. Encyclopedia of Statistical Sciences 1. John Wiley and Sons Inc, New York, 1982 p351–2.
[5] Carey RG. Measuring health care quality. How do you know your care has improved? Eval Health Prof 2000; 23:43–57.