Statistical Process Control

L-28

Contents

Statistical Process Control

SPC involves the use of various methods to measure and analyze a process.

The overall objectives of SPC are to (i) improve the quality of the process output (ii) reduce process variability and achieve process stability (iii) Solve processing problems.

There are seven principal methods or tools used in SPC, they are called as magnificent seven.

(i) Control charts
(ii) Histograms
(iii) Pareto charts
(iv) Check sheets
(v) Defect concentration diagram
(vi) Scatter diagrams
(vii) Cause and effect diagrams

1) Process variability and process capability

a) Process Variations

(i) Random: result from intrinsic variability in the process, no matter how well designed or well controlled it is (ii) assignable: Something has occurred in the process that is not accounted for by random variations. Reasons for assignable variations include operator mistakes, defective raw materials, tool failures and equipment malfunctions

creased ($\sigma_2 > \sigma_0$). Finally, at sampling time t_3, both the mean and standard deviation of the process are observed to have increased ($\mu_3 > \mu_0$ and $\sigma_3 > \sigma_0$).

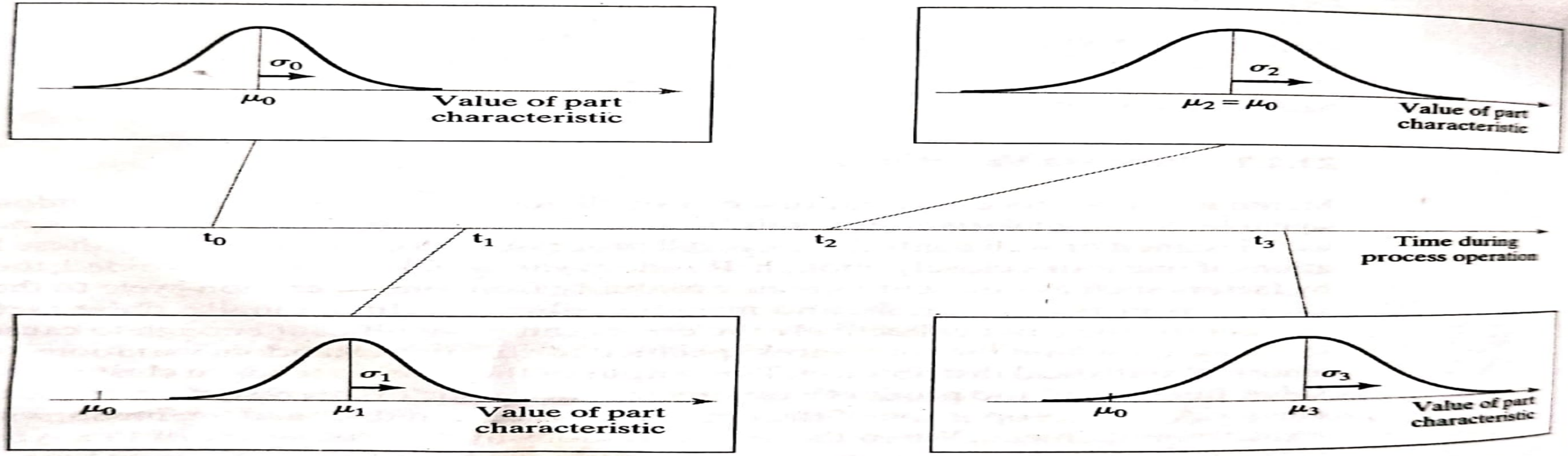

Figure 21.1 Distribution of values of a part characteristic of interest at four times during process operation: at t_0 process is in statistical control; at t_1 process mean has increased; at t_2 process standard deviation has increased; and at t_3 both process mean and standard deviation have increased.

b) Process capability and Tolerances

- Process capability relates to the normal variations inherent in the output when the process is in statistical control.

$PC = \mu \pm 3\sigma$

Where PC = process capability; μ = process mean ; σ = standard deviation of the process

- Assumptions underlying this definition are : (1) the output is normally distributed (2) Steady state has been achieved and the process is in statistical control. Under these assumptions 99.73% of the parts produced will have output values that fall within $\pm 3.0\sigma$

Sample Standard Deviation

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{21.2}$$

and the sample standard deviation s can be calculated from:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})}{n - 1}} \tag{21.3}$$

where x_i = measurement i of the part characteristic of interest; and n = the number of measurements in the sample, $i = 1, 2, \ldots, n$. Many hand-held calculators automatically compute these values based on input values of x_i. The values of $\bar{x}$ and s are then substituted for μ and σ in Eq. (21.1) to yield the following best estimate of process capability:

$$PC = \bar{x} \pm 3s \tag{21.4}$$

know the ratio of the specified tolerance range relative to the process capability, called *process capability index*, defined as:

$$PCI = \frac{UTL - LTL}{6\sigma} \quad (21.$$

where PCI = process capability index; UTL = upper tolerance limit of the tolerance rang LTL = lower tolerance limit; and 6σ = range of the natural tolerance limits. The under lying assumption in this definition is that the process mean is set equal to the nominal de

2) Control Charts

Control Chart is a graphical technique in which statistics computed from measured values of a certain process characteristic are plotted over time to determine if the process remains in statistical control

TABLE 21.1 Defect Rate as a Function of Process Capability Index (Tolerance Defined in Terms of Number of Standard Deviations of the Process), Given That the Process Is Operating in Statistical Control

Process Capability Index (PCI)	Tolerance = Number of Standard Deviations	Defect Rate (%)	Defective Parts per Million	Comments
0.333	±1.0	31.74	317,400	Sortation required.
0.667	±2.0	4.56	45,600	Sortation required.
1.000	±3.0	0.27	2,700	Tolerance = process capability.
1.333	±4.0	0.0063	63	Significant reduction in defects.
1.667	±5.0	0.000057	0.57	Rare occurrence of defects.
2.000	±6.0	0.0000002	0.002	Defects almost never occur.

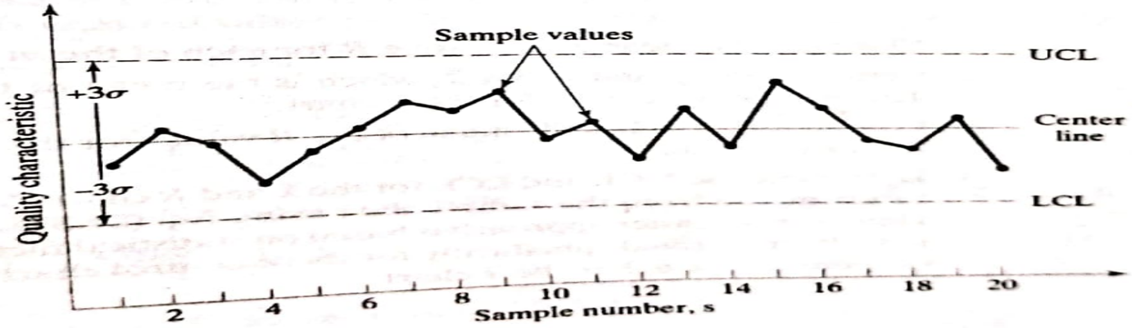

Figure 21.2 Control chart.

a) Control Charts for Variables

- A process that is out of statistical control manifests this condition in the form of significant changes in (i) Process mean (ii) Process variability.

Corresponding to these possibilities, there are two principal types of control charts for variables (1) X –bar chart (2) R chart

- The following procedure is used to construct the center, LCL and UCL for each chart

(i) Compute the mean x-bar and range R for each of the m samples

(ii) Compute the grand mean x-bar, which is the mean of the x-bar values for the m samples. This will be the center for the x-bar chart

(iii) Compute R-bar which is the mean of the R values for the m samples. This will be the center for the R chart

(iv) Determine the UCL and LCL for the X-bar and R charts.

Contd.

TABLE 21.2 Constants for the $\bar{x}$ and R Charts

Sample Size n	$\bar{x}$ Chart A_2	R Chart D_3	D_4
3	1.023	0	2.574
4	0.729	0	2.282
5	0.577	0	2.114
6	0.483	0	2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
9	0.337	0.184	1.816
10	0.308	0.223	1.777

$$\mathrm{LCL} = \bar{\bar{x}} - A_2\bar{R} \quad (21.6a)$$

$$\mathrm{UCL} = \bar{\bar{x}} + A_2\bar{R} \quad (21.6b)$$

And for the R chart:

$$\mathrm{LCL} = D_3\bar{R} \quad (21.7a)$$

$$\mathrm{UCL} = D_4\bar{R} \quad (21.7b)$$

EXAMPLE 21.1 $\bar{x}$ and R Charts

Although 20 or more samples are recommended, let us use a much smaller number here to illustrate the calculations. Suppose eight samples ($m = 8$) of size 5 ($n = 5$) have been collected from a manufacturing process that is in statistical control, and the dimension of interest has been measured for each part. It is desired to determine the values of the center, LCL, and UCL to construct the $\bar{x}$ and R charts. The calculated values of $\bar{x}$ and R for each sample are given below (measured values are in centimeters), which is step (1) in our procedure.

s	1	2	3	4	5	6	7	8
$\bar{x}$	2.008	1.998	1.993	2.002	2.001	1.995	2.004	1.999
R	0.027	0.011	0.017	0.009	0.014	0.020	0.024	0.018

Solution: In step (2), we compute the grand mean of the sample averages.

$$\bar{\bar{x}} = \frac{2.008 + 1.998 + 1.993 + 2.002 + 2.001 + 1.995 + 2.004 + 1.999}{8} = 2.000 \text{ cm}$$

In step (3), the mean value of R is computed.

$$\bar{R} = \frac{0.027 + 0.011 + 0.017 + 0.009 + 0.014 + 0.020 + 0.024 + 0.018}{8} = 0.0175 \text{ cm}$$

In step (4), the values of LCL and UCL are determined based on factors in Table 21.2. First, using Eq. (21.6) for the $\bar{x}$ chart,

$$\mathrm{LCL} = 2.000 - 0.577(0.0175) = 1.9899$$

$$\mathrm{UCL} = 2.000 + 0.577(0.0175) = 2.0101$$

And for the R chart using Eq (21.7),

$$\mathrm{LCL} = 0(0.0175) = 0$$

$$\mathrm{UCL} = 2.114(0.0175) = 0.0370$$

Contd.

The two control charts are constructed in Figure 21.3 with the sample data plotted in the charts.

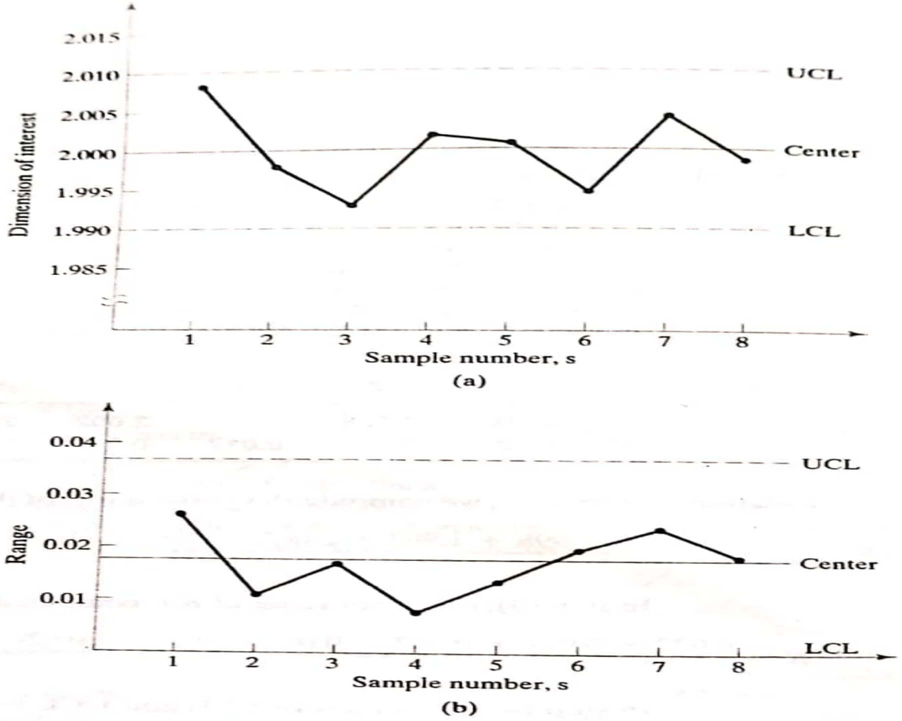

Figure 21.3 Control charts for Example 21.1: (a) $\bar{x}$ chart and (b) R chart.

If the mean and standard deviation for the process are known, an alternative way to calculate the center and UCL and LCL for the $\bar{x}$ chart is the following:

$$\text{LCL} = \mu - \frac{3\sigma}{\sqrt{n}} \tag{21.8a}$$

$$\text{UCL} = \mu + \frac{3\sigma}{\sqrt{n}} \tag{21.8b}$$

b) Control Charts for Attributes

- Control charts for attributes monitor the number of defects present in the sample or the fraction defect rate as the plotted statistic.
- Examples of these kinds of attributes include: number of defects per automobile, fraction of nonconforming parts in a sample, existence or absence of flash in plastic molding, and number of flaws in a roll of sheet steel.
- The two principal types of control charts for attributes are (1) the P-chart, which plots the fraction defect rate in successive samples (2) the C-chart, which plots the number of defects, flaws or other non-conformities per sample.

Contd.

p Chart. In the p chart, the quality characteristic of interest is the proportion (p for proportion) of nonconforming or defective units. For each sample, this proportion p_i is the ratio of the number of nonconforming or defective items d_i over the number of units in the sample n (assume samples are of equal size in constructing and using the control chart):

$$p_i = \frac{d_i}{n} \tag{21.10}$$

where i is used to identify the sample. If the p_i values for a sufficient number of samples are averaged, the mean value $\bar{p}$ is a reasonable estimate of the true value of p for the process. The p chart is based on the binomial distribution, where p is the probability of a nonconforming unit. The center in the p chart is the computed value of $\bar{p}$ for m samples of equal size n collected while the process is operating in statistical control.

$$\bar{p} = \frac{\sum_{i=1}^{m} p_i}{m} \tag{21.11}$$

The control limits are computed as three standard deviations on either side of the center. Thus,

$$\text{LCL} = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{21.12a}$$

$$\text{UCL} = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{21.12b}$$

where the standard deviation of $\bar{p}$ in the binomial distribution is given by

$$\sigma_p = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{21.13}$$

If the value of $\bar{p}$ is relatively low and the sample size n is small, then the LCL computed by the first of these equations is likely to be a negative value. In this case, let LCL = 0. (The fraction defect rate cannot be less than zero.)

EXAMPLE 21.2 *p* Chart

Ten samples ($m = 10$) of 20 parts each ($n = 20$) have been collected. In one sample there were no defects; in three samples there was one defect; in five samples there were two defects; and in one sample there were three defects. Determine the center, LCL, and UCL for the p chart.

Solution: The center value of the control chart can be calculated by summing the total number of defects found in all samples and dividing by the total number of parts sampled:

$$\bar{p} = \frac{1(0) + 3(1) + 5(2) + 1(3)}{10(20)} = \frac{16}{200} = 0.08 = 8\%$$

The LCL is given by Eq. (21.12a):

$$\text{LCL} = 0.08 - 3\sqrt{\frac{0.08(1-0.08)}{20}} = 0.08 - 3(0.06066) = 0.08 - 0.182 \rightarrow 0$$

The upper control limit, by Eq. (21.12b):

$$\text{UCL} = 0.08 + 3\sqrt{\frac{0.08(1-0.08)}{20}} = 0.08 + 3(0.06066) = 0.08 + 0.182 = 0.262$$

Contd.

c Chart. In the c chart (c for count), the number of defects in the sample are plotted over time. The sample may be a single product such as an automobile, and c = number of quality defects found during final inspection. Or the sample may be a length of carpeting at the factory prior to cutting, and c = number of imperfections discovered per 100 m. The c chart is based on the Poisson distribution, where c = parameter represent-

ing the number of events occurring within a defined sample space (defects per car, imperfections per specified length of carpet). Our best estimate of the true value of c is the mean value over a large number of samples drawn while the process is in statistical control:

$$\bar{c} = \frac{\sum_{i=1}^{m} c_i}{m} \tag{21.14}$$

This value of $\bar{c}$ is used as the center for the control chart. In the Poisson distribution, the standard deviation is the square root of parameter c. Thus, the control limits are:

$$LCL = \bar{c} - 3\sqrt{\bar{c}} \tag{21.15a}$$

$$UCL = \bar{c} + 3\sqrt{\bar{c}} \tag{21.15b}$$

EXAMPLE 21.3 *c* Chart

A continuous plastic extrusion process is considered to be operating in statistical control, and it is desired to develop a c chart to monitor the process. Eight hundred meters of the extrudate have been examined and a total of 14 surface defects have been detected in that length. Develop the c chart for the process, using defects per hundred meters as the quality characteristic of interest.

Solution: The average value of the parameter c can be determined using Eq. (21.14):

$$\bar{c} = \frac{14}{8} = 1.75$$

This will be used as the center for the control chart. The LCL is given by Eq. (15a):

$$LCL = 1.75 - 3\sqrt{1.75} = 1.75 - 3(1.323) = 1.75 - 3.969 \rightarrow 0$$

And the UCL, using Eq. (21.15b):

$$UCL = 1.75 + 3\sqrt{1.75} = 1.75 + 3(1.323) = 1.75 + 3.969 = 5.719$$

C) Interpreting Control Charts

- Montgomery lists a set of specific indicators that a process is likely to be out of statistical control and that corrective action should be taken. These indicators are:

(i) One point that lies outside the UCL or LCL

(ii) Two out of three consecutive points that lie beyond $\pm 2\sigma$ on one side of the center line of the control chart

(iii) Four out of five consecutive points that lie beyond $\pm 1\sigma$ on one side of the center line of the control chart

(iv) Eight consecutive points that lie on one side of the center line

(v) Six consecutive points in which each point is always higher that its predecessor or six consecutive points in which each point is always lower than its predecessor.

Contd.

Control charts serve as the feedback loop in SPC

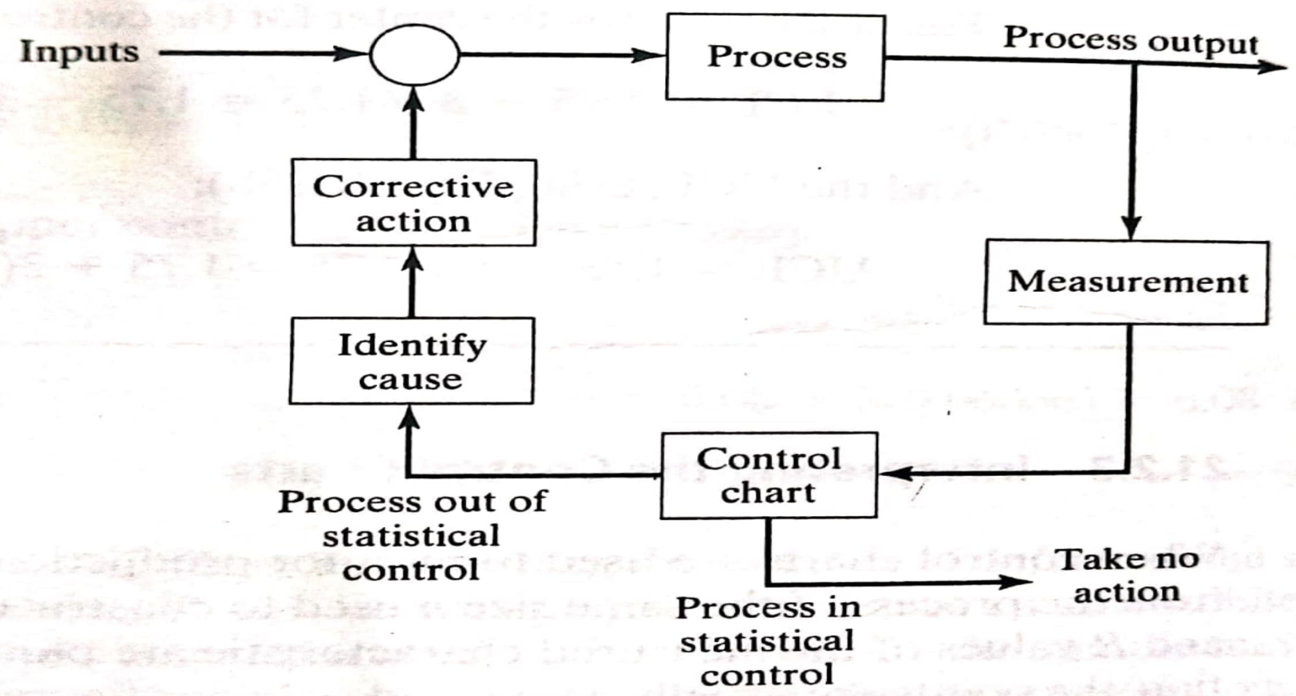

Figure 21.4 Control charts used as the feedback loop in SPC.

3) Other SPC Tools

a) Histogram

21.3.1 Histograms

The histogram is a basic graphical tool in statistics. After the control chart, it is probably the most important member of the SPC tool kit. A *histogram* is a statistical graph consisting of bars representing different values or ranges of values, in which the length of each bar is proportional to the frequency or relative frequency of the value or range, as shown in Figure 21.5. It is a graphical display of the *frequency distribution* of the numerical data. What makes the histogram such a useful statistical tool is that it enables the analyst to quickly visualize the features of a complete set of data. These features include: (1) the shape of the distribution, (2) any central tendency exhibited by the distribution, (3) approximations of the mean and mode of the distribution, and (4) the amount of scatter or spread in the data.

EXAMPLE 21.4 Frequency Distribution and Histogram

Part dimension data from the same process as in Example 21.1 are displayed in the frequency distribution of Table 21.3. The data are the dimensional values of individual parts taken from the process, while the process is in statistical control. Plot the data as a histogram and draw inferences from the graph.

TABLE 21.3 Frequency Distribution of Part Dimension Data

Range of Dimension	*Frequency*	*Relative Frequency*	*Cumulative Relative Frequency*
$1.975 \le x < 1.980$	1	0.01	0.01
$1.980 \le x < 1.985$	3	0.03	0.04
$1.985 \le x < 1.990$	5	0.05	0.09
$1.990 \le x < 1.995$	13	0.13	0.22
$1.995 \le x < 2.000$	29	0.29	0.51
$2.000 \le x < 2.005$	27	0.27	0.78
$2.005 \le x < 2.010$	15	0.15	0.93
$2.010 \le x < 2.015$	4	0.04	0.97
$2.015 \le x < 2.020$	2	0.02	0.99
$2.020 \le x < 2.025$	1	0.01	1.00

Solution: The frequency distribution in Table 21.3 is displayed graphically in the histogram of Figure 21.5. We can see that the distribution is normal (in all likelihood), and that the mean is around 2.00. We can approximate the standard deviation to be the range of the values (2.025 − 1.975) divided by 6, based on the fact that nearly the entire distribution (99.73%) is contained within $\pm 3\sigma$ of the mean value. This gives a σ value of around 0.008.

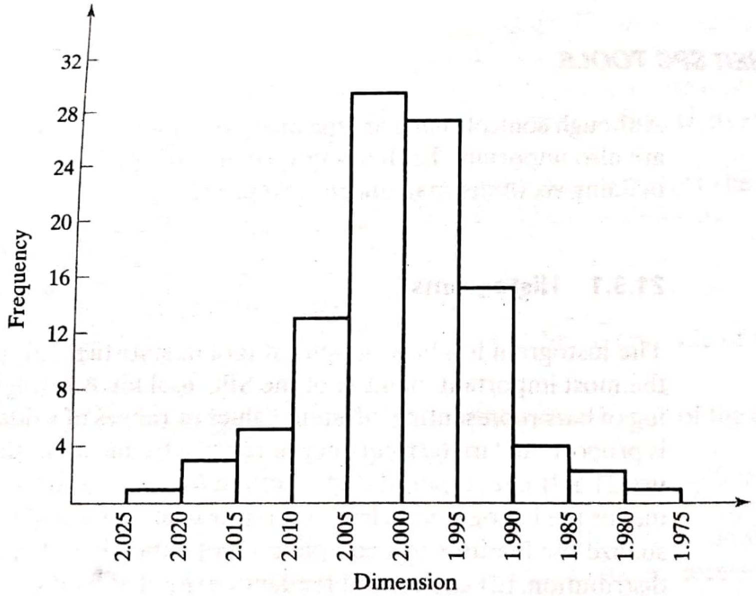

Figure 21.5 Histogram of the data in Table 21.3.

b)Pareto Charts

A *Pareto chart* is a special form of histogram, illustrated in Figure 21.6, in which attribute data are arranged according to some criterion such as cost or value. When appropriately used, it provides a graphical display of the tendency for a small proportion of a given population to be more valuable than the much larger majority. This tendency is sometimes referred to as *Pareto's Law*, which can be succinctly stated: "the vital few and the trivial many."[3] The "law" was identified by Vilfredo Pareto (1848–1923), an Italian economist and sociologist who studied the distribution of wealth in Italy and found that most of it was held by a small percentage of the population.

Pareto's Law applies not only to the distribution of wealth but to many other distributions as well. The law is often identified as the 80%-20% rule (although exact percent-

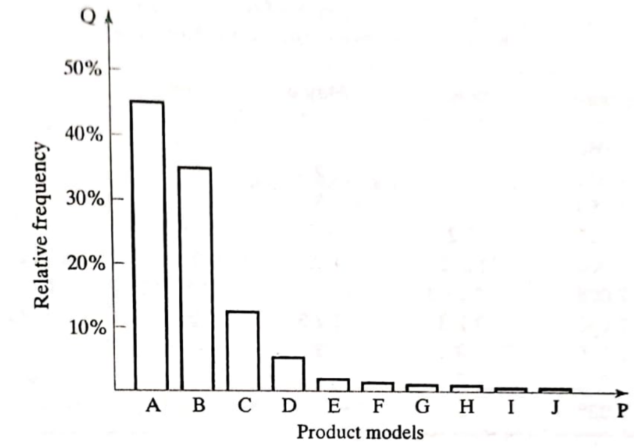

Figure 21.6 Typical (hypothetical) Pareto distribution of a factory's production output. Although there are ten models produced, two of the models account for 80% of the total units. This chart is sometimes referred to as a *P-Q* chart, where *P* = products and *Q* = quantity of production.

ages may differ from 80 and 20): 80% of the wealth of a nation is in the hands of 20% of its people; 80% of inventory value is accounted for by 20% of the items in inventory; 80% of sales revenues are generated by 20% of the customers; 80% of the quality savings can be obtained from 20% of the quality problems; and 80% of a factory's production output is concentrated in only 20% of its product models (as in Figure 21.6). What is suggested by Pareto's Law is that the most attention and effort in any study or project should be focused on the smaller proportion of the population that is seen to be the most important.

C) Check Sheets

The *check sheet* (not to be confused with "check list") is a data gathering tool generally used in the preliminary stages of the study of a quality problem. The operator running the process (e.g., the machine operator) is often given the responsibility for recording the data on the check sheet, and the data are often recorded in the form of simple check marks.

EXAMPLE 21.5 Check Sheet

For the dimensional data in the frequency distribution of Table 21.3, suppose we wanted to see if there were any differences between the three shifts that are responsible for making the parts. Design a check sheet for this purpose.

Solution: The check sheet is illustrated in Table 21.4. The data include the shift on which each dimensional value was produced (shifts are identified simply as 1, 2, and 3). The data in a check sheet are usually recorded as a function of time periods (days, weeks, months), as in our table.

TABLE 21.4 Check Sheet in Which Data from the Frequency Distribution of Table 21.3 Are Recorded According to Shift (1, 2, or 3) on Which the Parts Were Made

Range of Dimension	*May 5*	*May 6*	*May 7*	*May 8*	*May 9*	*Weekly Totals*
$1.975 \le x < 1.980$			3			1
$1.980 \le x < 1.985$		2		3	3	3
$1.985 \le x < 1.990$	1	3	3	1	3	5
$1.990 \le x < 1.995$	1 2	11 2 3	1 2	1 2	1 22	13
$1.995 \le x < 2.000$	11 222 3	11 22 3	111 222 3	11 222 3	11 22 3	29
$2.000 \le x < 2.005$	11 22 3	11 22 3	111 22 3	11 222 3	111 22	27
$2.005 \le x < 2.010$	1 2 3	1 2 3	22 3	1 33	1 2 3	15
$2.010 \le x < 2.015$	3	3	3		3	4
$2.015 \le x < 2.020$	3			3		2
$2.020 \le x < 2.025$	3					1
Total Parts/Day	20	20	21	20	19	100

It is clear from the data that the third shift is responsible for much of the variability in the data. Further analysis, shown in Table 21.5, substantiates this finding. This should lead to an investigation to determine the causes of the greater variability on the third shift, with appropriate corrective action to address the problem. The result of the corrective action might be to improve the process capability of the manufacturing operation making the parts.

TABLE 21.5 Summary of Data from Check Sheet of Table 21.4 Showing Frequency of Each Shift in Each of the Dimension Ranges

Range of Dimension	*Shift 1*	*Shift 2*	*Shift 3*	*Totals*
$1.975 \le x < 1.980$			1	1
$1.980 \le x < 1.985$		1	2	3
$1.985 \le x < 1.990$	2		3	5
$1.990 \le x < 1.995$	6	6	1	13
$1.995 \le x < 2.000$	11	13	5	29
$2.000 \le x < 2.005$	12	11	4	27
$2.005 \le x < 2.010$	4	5	6	15
$2.010 \le x < 2.015$			4	4
$2.015 \le x < 2.020$			2	2
$2.020 \le x < 2.025$			1	1
Weekly Total Parts/Shift	35	36	29	100
Average Daily Parts/Shift	7.0	7.2	5.8	

We also note from Table 21.5 that the average daily production rate for the third shift is somewhat below the daily rate for the other two shifts. The third shift seems to be a problem that demands management attention.

Contd.

Following types of check sheets can be distinguished

(i) Process distribution check sheet
(ii) Defective item check sheet
(iii) Defect location check sheet
(iv) Defect factor check sheet

d) Defect Concentration diagrams

This is a graphical method that has been found to be useful in analyzing the causes of product or part defects. The *defect concentration diagram* is a drawing of the product with all relevant views displayed, onto which have been sketched the various defect types at the locations where they each occurred. By analyzing the defect types and corresponding locations, the underlying causes of the defects can possibly be identified.

Montgomery [6] describes a case study involving the final assembly of refrigerators that were plagued by surface defects. A defect concentration diagram (Figure 21.7) was utilized to analyze the problem. The defects were clearly shown to be concentrated around the middle section of the refrigerator. On investigation, it was learned that a belt was wrapped around each unit for material handling purposes. It became evident that the defects were caused by the belt, and corrective action was taken to improve the handling method.

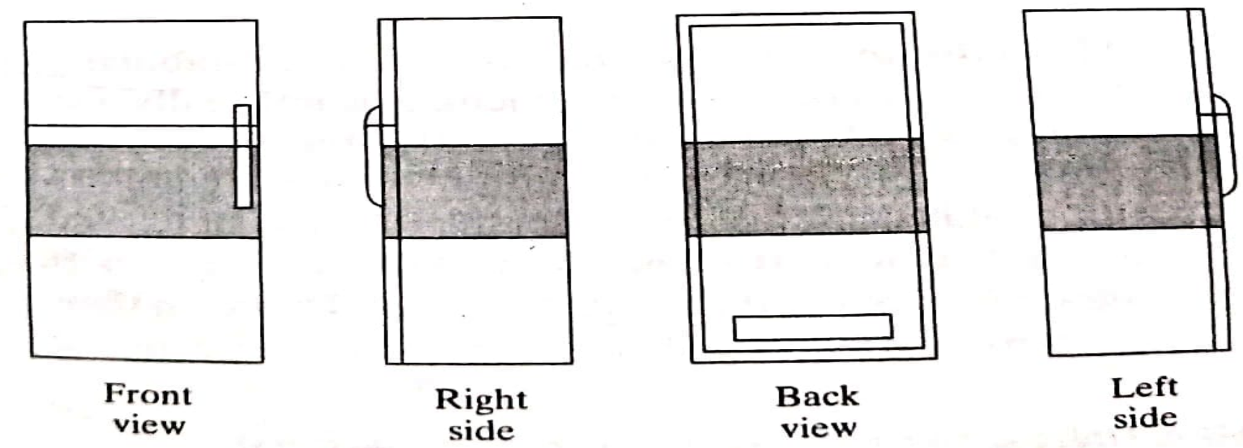

Figure 21.7 Defect concentration diagram showing four views of refrigerator with locations of surface defects indicated in cross-hatched areas.

e) Scatter Diagrams

In many industrial problems involving manufacturing operations, it is desirable to identify a possible relationship that exists between two process variables. The scatter diagram is useful in this regard. A *scatter diagram* is simply an x-y plot of the data taken of the two variables in question, as illustrated in Figure 21.8. The data are plotted as pairs; for each x_i value, there is a corresponding y_i value. The shape of the data points considered in aggregate often reveals a pattern or relationship between the two variables. For example, the scatter diagram in Figure 21.8 indicates that a negative correlation exists between cobalt content and wear resistance of a cemented carbide cutting tool. As cobalt content increases, wear resistance decreases. One must be circumspect in using scatter diagrams and in extrapolating the trends that might be indicated by the data. For instance, it might be inferred from our diagram that a cemented carbide tool with zero cobalt content would possess the highest wear resistance of all. However, cobalt serves as an essential binder in the pressing and sintering process used to fabricate cemented carbide tools, and a minimum level of cobalt is necessary to hold the tungsten carbide particles together in the final product. There are other reasons why caution is recommended in the use of the scatter diagram, since only two variables are plotted. There may be other variables in the process whose importance in determining the output is far greater than the two variables displayed.

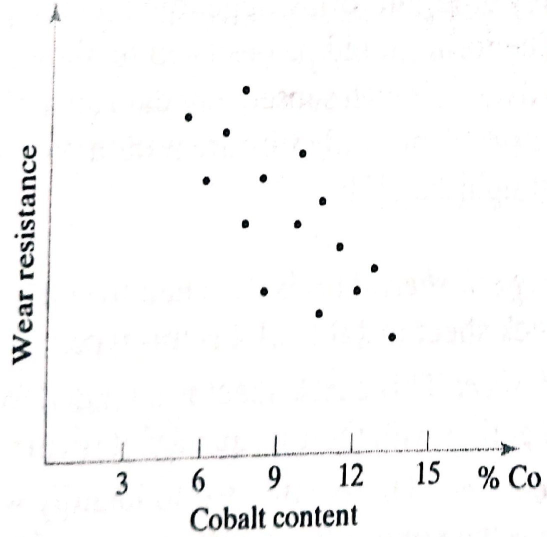

Figure 21.8 Scatter diagram showing the effect of cobalt binder content on wear resistance of a cemented carbide cutting tool insert.

f) Cause and effect diagram

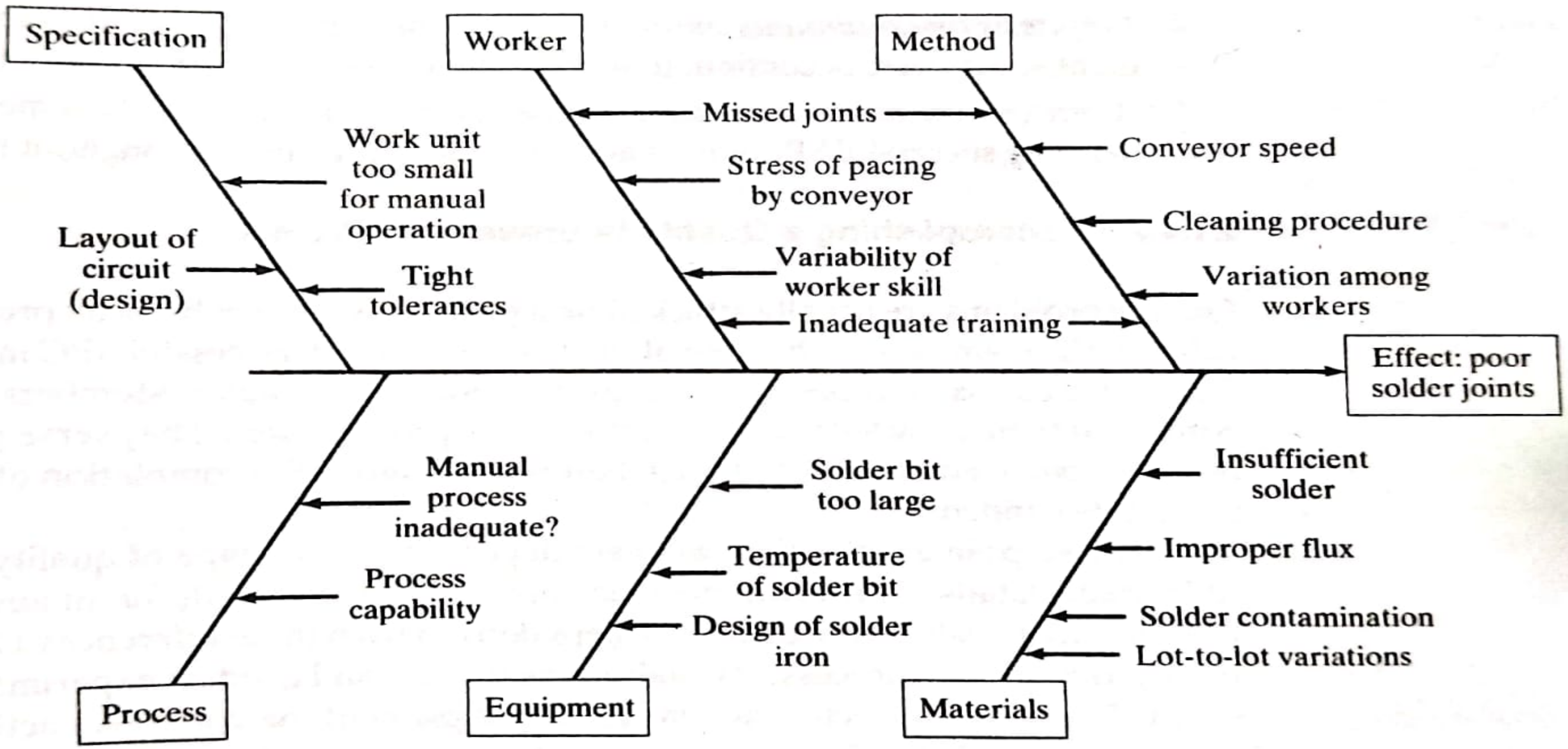

Figure 21.9 Cause and effect diagram for a manual soldering operation. The diagram indicates the effect (the problem is poor solder joints) at the end of the arrow, and the possible causes are listed on the branches leading toward the effect.

4) Implementing Statistical Process Control

a) Elements of a successful SPC Program

(i) Management commitment and leadership

(ii) Team approach to problem solving

(iii) SPC training for all employees

(iv) Emphasis on continuous improvement

(v) A recognition and communication system

b)Accomplishing a Quality Improvement Project

TABLE 21.6 Applications of the Seven SPC Tools in a Quality Improvement Project

Quality improvement Project Step	*SPC Tool*	*Other Techniques and Actions*
1. Select the project	Control charts Pareto chart	Pareto priority index
2. Observe the process	Check sheet	Check list Propose theories and hypotheses
3. Analyze the project	Histogram Pareto chart Defect concentration diagram Scatter diagram Cause and effect diagram	Conduct experiments Computer simulations Evolutionary operations on actual process Literature review
4. Formulate corrective action	Scatter diagram Cause and effect diagram	Make recommendations Management approval and authorization
5. Implement corrective action		Revise procedures Manage change Project assessment (audit) Disband team