APPLIED STATISTICS

Statistical process and product control

Renu Kaul and S. Roy Chowdhury
Reader, Department Of Statistics
Lady Shri Ram College For Women
Lajpat Nagar, New Delhi

4-Jan-2007 (Revised 20-Nov-2007)

CONTENTS

Keywords

Process control; Product control; Sub-grouping; Control charts; Process capability; CUSUM chart, 6-σ Capability.

Quality of a product

Quality now a days has become a major consumer decision factor. It is in the eye of the consumer and may be viewed from different angles. By quality we mean an attribute/characteristic of a product that determinates its fitness for use viz. chemical, physical, aesthetic etc.

Need for quality control

Quality control is a technique that controls the quality of the product or process from its design stage to end product. It covers all the aspects of process of production including the quality of raw materials, quality of manpower, quality of equipment and quality of management. Thus, statistical quality control means maintaining the quality of goods using statistical techniques in a continuous flow of production.

There is always a need to control the quality of product or process. This is because there are variations in products or processes due to presence of causes of variation in quality. There are mainly two major causes of variation in quality. One is chance cause (common cause) variation. It is the variation, which is inherent in the process due to the way it is managed. It is consistent over time and its effect is predictable. This type of variation is always present in the process. For example, the net content of a milk pouch varies slightly from pouch to pouch. The variation due to this cause is beyond the control of human hand and cannot be eliminated except under experimental conditions. The range of such variation is known as natural tolerance of the product or process. The second type of variation is assignable cause variation. The causes for this type of variation can easily be discovered but not predicted. The assignable causes may appear at any stage of the process starting from the raw form to final stage. This type of variation is due to defective/ substandard raw material, mishandling of machines, faulty equipment, unskilled technical staff, etc. The variation due to assignable causes can be controlled by controlling the cause. The main purpose of statistical quality control is to devise statistical methods to separate chance causes of variation from assignable causes of variation. The product or process is in statistical control if only common or chance causes of variation govern the whole system.

Basic concepts of process and product control

The main aim in any production process is to control and maintain the quality of the manufactured product so that it satisfies specified quality specifications. Its means the process should be controlled in such a manner so that the number of defective products produced by the process is minimum. This is called process control and can be obtained by developing and analyzing control charts. The other concept is the product control, which means controlling the quality of product and maintaining it at certain quality level accepted by consumers regardless of what quality level is being maintained by the producer. This can be obtained by Sampling Inspection Plans.

Process capability

Even if the process is under statistical control i.e. all assignable causes of variation are removed from the process, it gives different values under each observation due to the presence of common cause variation. The range of this variation is called natural tolerance (NT), which is the distance from -3σ to -3σ where σ is the process standard deviation.

This tolerance is also called process capability i.e. the capability of a process to meet the specifications when the process is in statistical control.

General theory of control charts

Control chart is a pictorial representation of process over a period of time. It shows the variation pattern of any process. A control chart tells if the process is in statistical control (i.e. only common cause variation is present) or is out of control (i.e. assignable causes variation is also present).
Control charts are based on a general model i.e. the control limits are $\pm 3\sigma$ from the average. A control chart consists of three horizontal lines representing:

(a) Upper control limit (UCL)
(b) Central limit indicating the average level of process (CL)
(c) Lower control limit (LCL)

and sample points.

Under controlled process most of the points (observations) should lie between UCL and LCL The general layout of a control chart is as shown in Fig. 1.

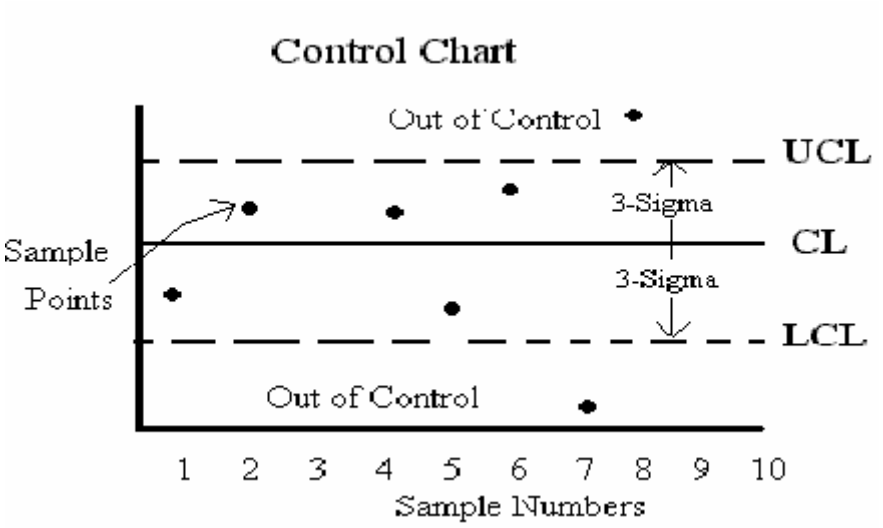


Fig. 1

The control limits are function of the underlying statistics 'τ' whose values depend on the sampling distribution of 'τ' defined as:

$$\begin{aligned} UCL &= E(\tau) + 3S.E.(\tau) \\ CL &= E(\tau) \\ LCL &= E(\tau) - 3S.E.(\tau) \end{aligned} \qquad (1)$$

Let $x_1, x_2 ..... x_n$ be sample observations and the statistics $\tau = f(x_1, x_2 ..... x_n)$ with $E(\tau) = \mu_\tau$ and $\text{Var}(\tau) = \sigma_\tau^2$. If the statistics τ follows normal distribution then we have:

$$\begin{aligned} & P[\mu_\tau - 3\sigma_\tau < \tau < \mu_\tau + 3\sigma_\tau] = 0.9973 \\ \Rightarrow\ & P[|t - \mu_\tau| < 3\sigma_\tau] = 0.9973 \\ \Rightarrow\ & P[|t - \mu_\tau| > 3\sigma_\tau] = 0.0027 \end{aligned} \qquad (2)$$

i.e. the probability that a random value of τ lies within 3σ limits i.e. $\mu_\tau \pm 3\sigma_\tau$ is 0.9973. In other words the probability that a random value of τ lies outside 3σ limits is 0.0027. When the process is in control the sample values should lie between $\mu_\tau - 3\sigma_\tau$ and $\mu_\tau + 3\sigma_\tau$ and process is said to be in statistical control. If the sample points fall out side the control limits then we must analyse the cause of process being not under control.

When the statistics τ does not follow the assumption of normality then for non-normal population we use Chebyechev's inequality from probability theory which is defined as: for any constant c > 0

$$P[|\tau - E(\tau)| < c] > 1 - \frac{Var(\tau)}{c^2}, \text{ take } c = 3\sigma_\tau$$

$$\Rightarrow P[|\tau - \mu_\tau| < 3\sigma_\tau] > 1 - \frac{\sigma_\tau^2}{9\sigma_\tau^2} = \frac{8}{9} \qquad (3)$$

$$\Rightarrow P[|\tau - \mu_\tau| < 3\sigma_\tau] > \frac{8}{9}$$

provided σ_τ^2 is known.

Dr. Walter Shewhart (1931) developed the concept of control charts while working at Bell Labs, the research arm of American Telephone and Telegraph Company. He developed two types of control charts, one for attributes and the other for variables.

Sub-grouping

An important idea introduced by Shewhart in the application of control charts is that of rational sub-groups. The basic principle behind this rationale is to choose the sub-groups in such a way that will give the maximum chance for the observations to be homogenous within a subgroup and the maximum chance for the subgroups to be different.

Of course, order of production is an obvious basis for forming rational subgroup. For example, we could take a sample of four units, then wait for sometime, take another sample of four units and continue. This method reduces the chance of assignable variation, if any, within the subgroup and maximizes the chance of assignable variation, if any, between subgroups. However, in this case it is quite likely that a process shift, which occurred and got corrected between samples is missed. The second method is that we could draw every third unit till four units have been selected, wait for some time and continue in the same manner. However, in this case it is quite likely that small process cycles are missed.

It is not always necessary to follow the order of production in forming rational subgroups. Suppose there are four machines in a factory. The final output from all the machines is blended. Now a sample from this mixture will not be able to detect which of the machines are not in control. In this case, one needs to apply the control chart technique separately to the output from each machine. The same concept may be applied to different operators, different suppliers, different shifts etc.

Control charts for attributes

In practical situations not all quality characteristics can be measured numerically. For example, number of defective items produced by a process or number of defects in the product. To study quality characteristics of such cases control charts for attributes were developed. Here we simply classify an item as conforming or non-conforming i.e. possessing or not possessing an attribute.

There are four control charts for attributes:

1 p-chart or control chart for fraction defective
2 np-chart or control chart for number of defectives
3 c-chart or control chart for number of defects and
4 u-chart or control chart for average number of defects/units

p-Chart: Control chart for fraction defective

Let P be the population proportion and d be the number of defectives in a sample of size n then the sample proportion defective is p = d/n where d is a binomial variate with parameter n and P, thus E(d) = nP and Var(d) = nPQ, where Q = 1-P;

$$\Rightarrow \quad \boxed{\begin{aligned} E(p)&=E\left(\frac{d}{n}\right)=\frac{1}{n}E(d)=P \\ Var(p)&=Var\left(\frac{d}{n}\right)=\frac{1}{n^2}Var(d)=\frac{PQ}{n} \end{aligned}} \qquad (4)$$

thus for p-chart the 3-σ control limits are given by:

$$\boxed{\begin{aligned} &P\pm 3\sqrt{PQ/n} \\ &P\pm A\sqrt{PQ} \ , \text{ where } A=3/\sqrt{n} \end{aligned}} \qquad (5)$$

The value of A, computed for different values of n is shown in Annexure – 1
There are two cases to be considered while developing a p-chart:

(a) When standards are given

Let P_1 be the given value of P then the control limits are:

$$\boxed{\begin{aligned} UCL&=P_1+A\sqrt{P_1(1-P_1)} \\ CL&=P_1 \\ LCL&=P_1-A\sqrt{P_1(1-P_1)} \end{aligned}} \qquad (6)$$

(b)Standards are not given

Let for the i^{th} sample (i = 1, 2, …, m) of size n_i, d_i denote the number of defectives and p_i the fraction defective, then the control limits for p-chart are given by:

$$UCL=\bar{p}+A\sqrt{\bar{p}(1-\bar{p})}$$
$$CL=\bar{p}$$
$$LCL=\bar{p}-A\sqrt{\bar{p}(1-\bar{p})} \qquad (7)$$

where $$\bar{p}=\frac{\sum_{i=1}^{m} d_i}{\sum_{i=1}^{m} n_i}=\frac{\sum_{i=1}^{m} n_i p_i}{\sum_{i=1}^{m} n_i}$$

which is an unbiased estimate of P, since

$$E(\bar{p})=\frac{\sum_i E(d_i)}{\sum_i n_i}=\frac{\sum_i E(n_i P)}{\sum_i n_i}=P$$

Since p can't be negative so minimum value of LCL can be zero.

If the sample size n is different for all the samples then both UCL and LCL will change from sample to sample but the CL will remain same. However, If the sample size n is fixed for all the samples, then only one set of control limits needs to be computed and an estimate of the population proportion P is given by:

$$\hat{P}=\bar{p}=\frac{\sum_{i=1}^{m} np_i}{\sum_i n}=\frac{n\sum_{i=1}^{m} p_i}{nm}=\frac{1}{m}\sum_{i=1}^{m} p_i \qquad (8)$$

The process is said to be under statistical control if all the points are within control limits without following any pattern. When the points are lying outside UCL, there is an indication of deterioration in the quality and should be monitored regularly so that the cause can be identified and rectified by taking necessary steps. On other hand, when points are lying below LCL the process shows improvement in the product quality. However, before taking it for guaranteed, it should be properly investigated.

When a number of points fall outside the control limits, a revised estimate of p should be computed by eliminating all the points falling outside limits and new control limits should be computed to take a decision.

EXAMPLE - 1

A group working on printing of labels to be pasted on their product observed that defective labels produced every day affect the cost of printing. A label is said to be defective if it has incorrect price, incorrect code, incorrect batch number or incorrect product name. To reduce the cost by decreasing the percentage of defective labels the group decided to study the behavior of the process of producing labels by developing a p-chart. The group decided to pull randomly n labels per day for 20 days.

The number of defective labels produced every day is as follows:

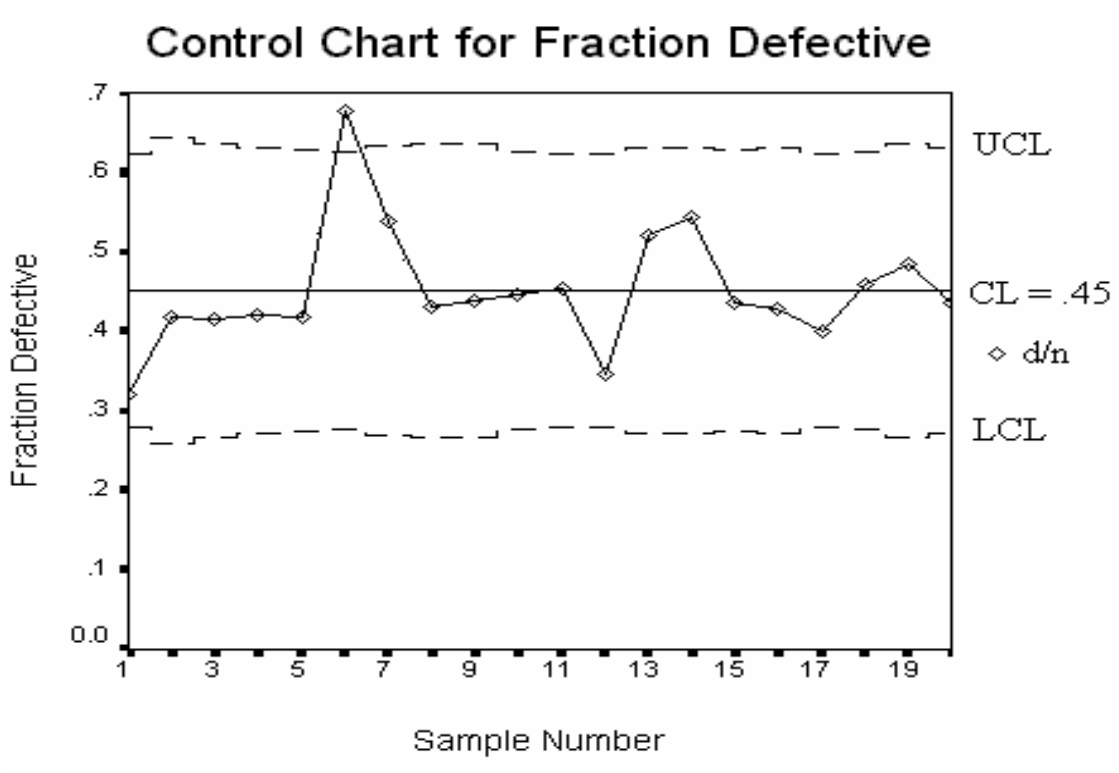
Day Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Number of labels	75	60	65	69	72	74	67	65	66	74	75	75	69	70	71	70	75	74	66	69
Number of defectives	24	25	27	29	30	50	36	28	29	33	34	26	36	38	31	30	30	34	32	30

SOLUTION

Since the subgroup size varies in this case, separate control limits have been calculated for each subgroup. However, the central limit for all of them remains the same. The following table gives the calculated control limits for the given observations:

Day Number	Number of Labels (n)	Number of Defectives (d)	Fraction Defective (d/n)	UCL	LCL
1	75	24	0.32	0.62	0.28
2	60	25	0.42	0.64	0.26
3	65	27	0.42	0.64	0.27
4	69	29	0.42	0.63	0.27
5	72	30	0.42	0.63	0.28
6	74	50	0.68	0.62	0.28
7	67	36	0.54	0.63	0.27
8	65	28	0.43	0.64	0.27
9	66	29	0.44	0.63	0.27
10	74	33	0.45	0.62	0.28
11	75	34	0.45	0.62	0.28
12	75	26	0.35	0.62	0.28
13	69	36	0.52	0.63	0.27
14	70	38	0.54	0.63	0.27
15	71	31	0.44	0.63	0.27
16	70	30	0.43	0.63	0.27
17	75	30	0.4	0.62	0.28
18	74	34	0.46	0.62	0.28
19	66	32	0.48	0.63	0.27
20	69	30	0.43	0.63	0.27

Control chart



Conclusion

Since one of the points is going out of control, showing the presence of assignable cause at only one point. Ignoring that sample the process is statistically under control and special attention has to be paid on the assignable cause present at that point.

np-Chart: Control chart for number of defective

When one is interested in making a decision on the basis of number of defectives in the sample instead of sample proportion defective, 3-σ control limits for np-chart need to be computed.

For np-chart, the 3-σ control limits are:

$$nP \pm 3\sqrt{nP(1-P)} \qquad (9)$$

(a) Control limits when standards are known:

Let P_1 be the given value of P then

$$\begin{aligned} UCL &= nP_1 + 3\sqrt{nP_1(1-P_1)} \\ CL &= nP_1 \\ LCL &= nP_1 - 3\sqrt{nP_1(1-P_1)} \end{aligned} \qquad (10)$$

(b) Control limits when standards are not known:

Let $\bar{p}$ be the estimate of P then

$$UCL=n\bar{p}+3\sqrt{n\bar{p}(1-\bar{p})}$$
$$CL=n\bar{p} \qquad (11)$$
$$LCL=n\bar{p}-3\sqrt{n\bar{p}(1-\bar{p})}$$

where, $\bar{p}$, can be computed in the same way as in the case of p-chart. Since again p can't be negative the minimum value of LCL can be zero.

Again, if the sample size n is same for all the samples, only one set of control limits needs to be developed whereas if the sample size is varying then all the three limits need to be computed afresh for all the samples. Hence it is more convenient to use p-chart than np-chart for variable sample size.

EXAMPLE – 2

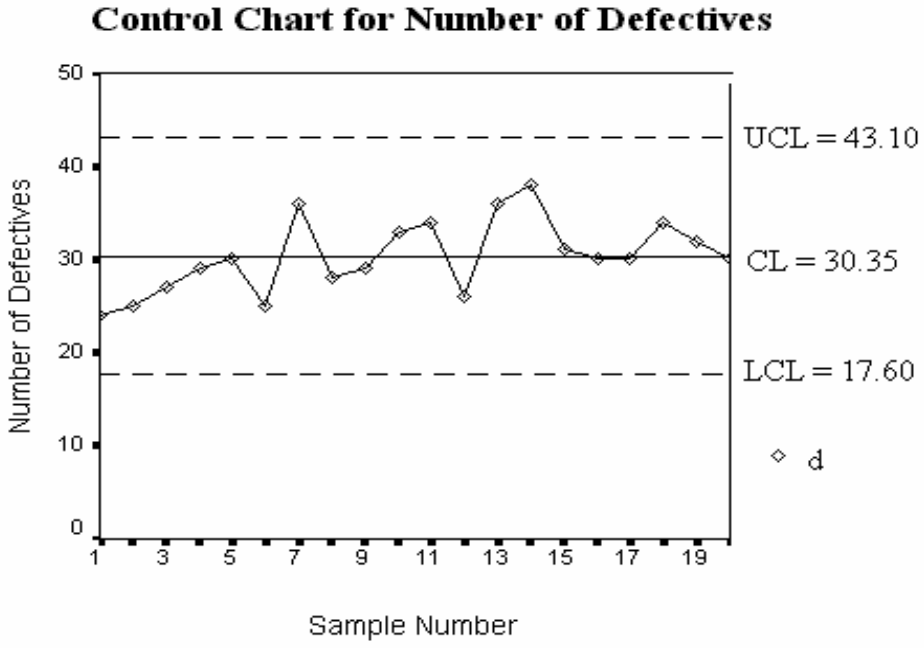
For the above example, another group working in the second printing unit decided to pull 75 labels every day. The data collected for 20 days is as follows. Construct a np- chart.

Day Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Number of Labels (n)	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75
Number of Defectives (np)	24	25	27	29	30	25	36	28	29	33	34	26	36	38	31	30	30	34	32	30

SOLUTION

The np-chart for the above observations is as follows:

CONTROL CHART



CONCLUSION

Since all the points are within control limits so the process is statistically under control. Its shows that the variation in the process is due to chance causes only and looking for assignable causes at this stage is not required.

c-Chart: Control chart for number of defects

In many situations, we need to study the number of defects in a product instead of just identifying it as a defective product. For example, a car consists of a number of parts. It's not necessary that all the parts are defective. If some of the parts are defective we can't say that the car is defective and we take decision on the basis of c-chart. Similarly, in a piece of cloth there may be some defects, which need to be counted and they vary from sample to sample even when the size of the cloth is same. To study such cases, one needs to compute control limits for the number of defects.

When sample size is same i.e. sample can be a single unit (for example, a car, a welding machine) or group of units or a unit of fixed length (for example, a piece of cloth), time, area (e.g. surface area of a base plate), one can study the process control by developing c-charts. But, when the sample size is varying (for example, the size of cloth varies from sample to sample or area of base plate varies) then instead of considering a c-chart one should consider a u (=c/n)-chart.

Let n be the size of the sample and let the probability p of occurrence of a defect over the sample space be very small such that np is finite. Then the number of defects (c) follows Poisson distribution with E(c) = λ and Var (c) = λ.

Thus for c-chart the 3-σ control limits are computed as:

$$\begin{aligned} UCL &= \lambda + 3\sqrt{\lambda} \\ CL &= \lambda \\ LCL &= \lambda - 3\sqrt{\lambda} \end{aligned} \qquad (12)$$

(a) When standards are given :

Let λ_1 be the given value of λ then

$$\begin{aligned} UCL &= \lambda_1 + 3\sqrt{\lambda_1} \\ CL &= \lambda_1 \\ LCL &= \lambda_1 - 3\sqrt{\lambda_1} \end{aligned} \qquad (13)$$

(b) **When standards are not given:**

If λ is unknown, let c_i be the number of defects observed on the i^{th} sample unit (i = 1, 2, …, m) then

$$E(\lambda) = \hat{\lambda} = \sum_{i=1}^{m} c_i / m = \bar{c}$$

then

$$UCL=\bar{c}+3\sqrt{\bar{c}}$$
$$CL=\bar{c}$$
$$LCL=\bar{c}-3\sqrt{\bar{c}} \qquad (14)$$

Again $\bar{c}$ can not be negative so minimum value of LCL should be taken as zero.

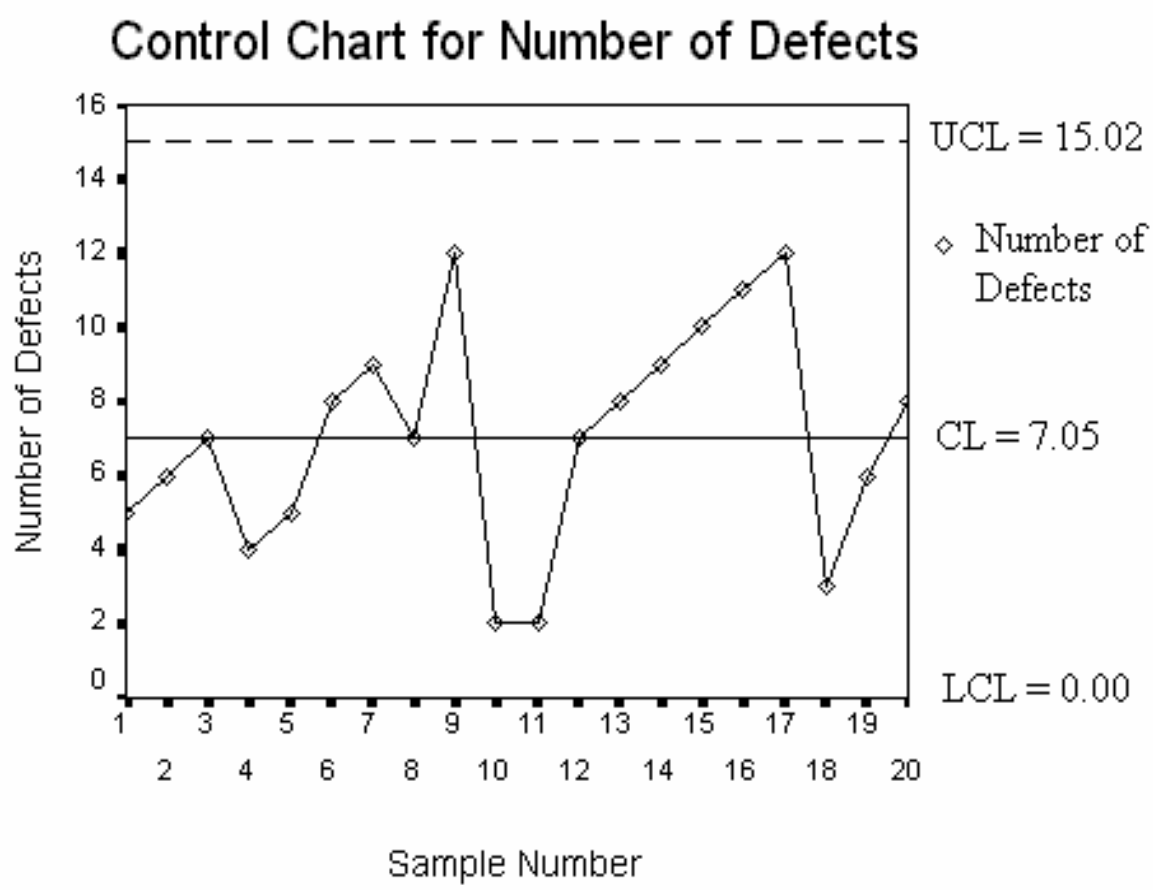
EXAMPLE 3

A shampoo manufacturing company wanted to reduce the number of surface defects on the bottles which make the bottle look less appealing. The specification requirements are no more than 15 surface defects per bottle. To study the process one bottle was randomly selected out of each 50 bottles manufactured per shift for 20 shifts. The data is as follows:

Sample Number	1	2	3	4	5	6	7	8	9	10
Number of defects	5	6	7	4	5	8	9	7	12	2

Sample Number	11	12	13	14	15	16	17	18	19	20
Number of defects	2	7	8	9	10	11	12	3	6	8

SOLUTION
CONTROL CHART



CONCLUSION

Since all the points are within control limits so the process is statistically under control. It shows that the variation in the process is only due to chance causes.

u-Chart: Control chart for average number of defects/unit

There are practical situations where sample size is not fixed. Let n_i be the sample size and c_i be the total number of defects observed in the i^{th} sample then $u_i = c_i/n_i$ be defined as the average number of defects per unit for the i^{th} sample (i = 1, 2, …, m)

Hence the standard error of the average number of defects per unit is given by:

$$S.E.(u) = \sqrt{\frac{\lambda}{n_i}} = \sqrt{\frac{\bar{u}}{n_i}}, \quad where \quad \hat{\lambda} = \bar{u} = \frac{\sum_{i}^{m} c_i}{\sum_{i}^{m} n_i} \tag{15}$$

Hence the 3-σ control limits for u-chart are:

$$\begin{aligned} UCL &= \bar{u} + 3\sqrt{\bar{u}/n_i} \\ CL &= \bar{u} \\ LCL &= \bar{u} - 3\sqrt{\bar{u}/n_i} \end{aligned} \tag{16}$$

It has been observed that CL will remain the same where as UCL and LCL will vary for each sample size. The interpretation of u-chart will be same as that of c-chart.

EXAMPLE 4

A firm of architects was concerned about the number of mistakes being committed in their drawings requiring reworking. The firm decided to reduce the number of re-workings. To see whether the process was in control or not, the firm decided to use a u chart. The inspection unit was defined as a single drawing. The data that was collected for 20 weeks is as follows:

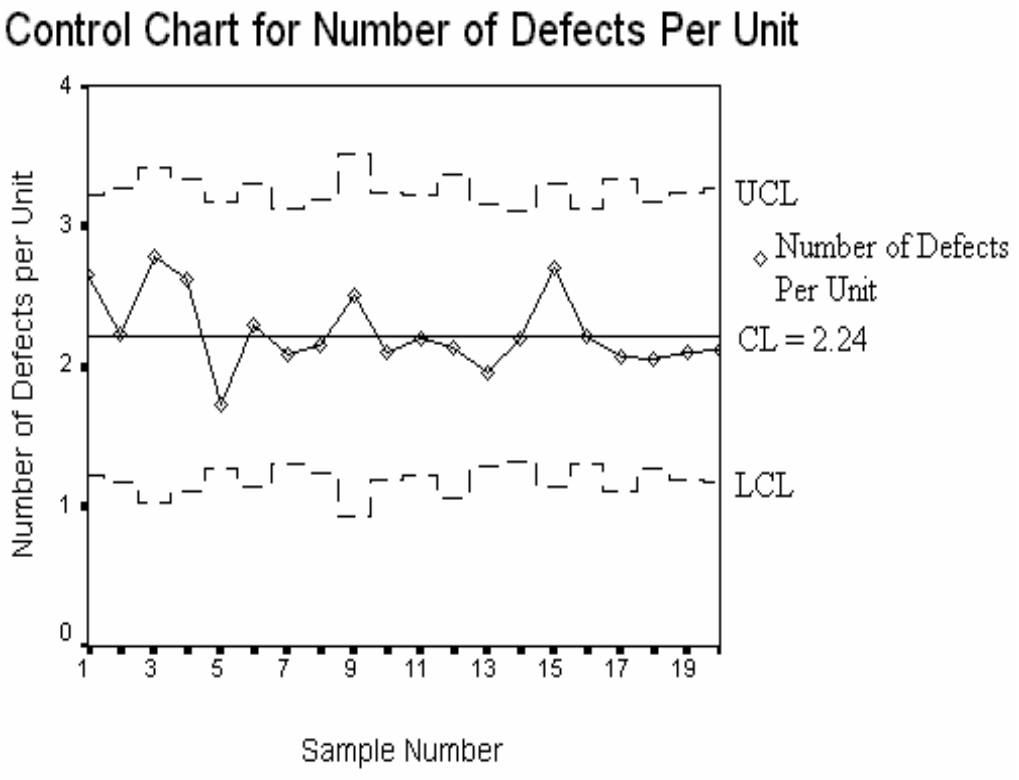
Week number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Number of drawings made	20	18	14	16	22	17	24	21	12	19	20	15	23	25	17	24	16	22	19	18
Number of mistakes (c)	53	40	39	42	38	39	50	45	30	40	44	32	45	55	46	53	33	45	40	38

SOLUTION

The following table shows the control limits for U-chart:

Week number	Number of drawings made	Number of mistakes (c)	Mistakes per unit	UCL	LCL
1	20	53	2.65	3.25	1.24
2	18	40	2.22	3.30	1.18
3	14	39	2.79	3.44	1.04
4	16	42	2.63	3.37	1.12
5	22	38	1.73	3.20	1.29
6	17	39	2.29	3.33	1.15
7	24	50	2.08	3.16	1.33
8	21	45	2.14	3.22	1.26
9	12	30	2.50	3.54	0.95
10	19	40	2.11	3.27	1.21
11	20	44	2.20	3.25	1.24
12	15	32	2.13	3.40	1.08
13	23	45	1.96	3.18	1.31
14	25	55	2.20	3.14	1.34
15	17	46	2.71	3.33	1.15
16	24	53	2.21	3.16	1.33
17	16	33	2.06	3.37	1.12
18	22	45	2.05	3.20	1.29
19	19	40	2.11	3.27	1.21
20	18	38	2.11	3.30	1.18

CONTROL CHART



CONCLUSION

Since all the points are within control limits the process is statistically under control.

Control charts for variables

These type of charts are developed when the quality characteristic of the product is measurable. For example- diameter of a wire, tensile strength, life of a bulb etc., the variable in these cases follows normal probability law. To study control charts for variables one can define two type of charts:

(i) Control charts for mean and range i.e. Control charts for $\overline{X}$ and R

(ii) Control charts for mean and standard deviation i.e. Control charts for $\overline{X}$ and σ

Control charts for $\overline{X}$ and R & $\overline{X}$ and σ

As discussed earlier, there are variations in the process due to both chance causes and assignable causes. The control limits in the $\overline{X}$ and R charts are developed in such a way that they reveal the presence or absence of assignable causes of the variation.

For developing control charts one needs to select the samples or sub-groups. These should be chosen in such a way so that the amount of variation within a sample is minimum. For details see discussion on sub-grouping. The size of the sample i.e. n and the time between two samples is important and it depends on type of process. Usually one may consider 25 samples of size 4 each or 20 samples of size 5 each. Generally, the time gap between two samples varies from 15 minutes to 30 minutes but it depends on the process and production cycle time.

Let X_{ij} (i=1,2,…m), (j=1,2,…n) be the measurement of the j^{th} observation in the i^{th} sample of size n, then the mean range and standard deviation for i^{th} sample can be computed as:

$$\overline{X_i} = \frac{1}{n}\sum_{j=1}^{n} X_{ij}$$
$$R_i = \max X_{ij} - \min X_{ij} \qquad (17)$$
$$s_i^2 = \frac{1}{n}\sum_{j=1}^{n}(X_{ij} - \overline{X_i})^2$$

Next we compute $\overline{\overline{X}}, \overline{R}$ *and* $\overline{s}$ as:

$$\overline{\overline{X}} = \frac{1}{m}\sum_{i=1}^{m} \overline{X}_i$$
$$\overline{R} = \frac{1}{m}\sum_{i=1}^{m} R_i \qquad (18)$$
$$\overline{s} = \frac{1}{m}\sum_{i=1}^{m} s_i$$

Control Limits for $\bar{X}$ Charts:

a) **When standards are known** i.e. when both μ and σ are known, then the 3-σ control limits for $\bar{X}$ chart are given by:

$$E(\bar{X}) \pm 3\ \text{S.E.}(\bar{X})$$

$$= \mu \pm 3\sigma/\sqrt{n} = \mu \pm A\sigma \qquad \text{(where } A=3/\sqrt{n}\text{)} \qquad (19)$$

(where A is a constant depending on the sample size n whose values are given in Annexure-I)

b) **When standards are unknown** i.e. the values of both μ and σ are unknown Then

$$\bar{\bar{X}} \pm 3\frac{\bar{R}}{d_2}\frac{1}{\sqrt{n}} = \bar{\bar{X}} \pm \left(\frac{3}{d_2\sqrt{n}}\right)\bar{R} = \bar{\bar{X}} \pm A_2\bar{R} \qquad (20)$$

where, $\hat{\mu} = \bar{\bar{x}}$, $\hat{\sigma} = \dfrac{\bar{R}}{d_2}$ (valid for a normal variable)

where d_2 and A_2 are constants depending on the value of n. For different value of n, values of d_2 and A_2 are given in Annexure-I

The 3-σ control limits for $\bar{X}$ chart are given by:

$$\begin{aligned} UCL &= \bar{\bar{X}} + A_2\bar{R} \\ CL &= \bar{\bar{X}} \\ LCL &= \bar{\bar{X}} - A_2\bar{R} \end{aligned} \qquad (21)$$

Now instead of $\bar{R}/d_2$ we estimate σ from the relation:

$$E(\bar{s}) = C_2\sigma \quad \text{(valid for a normal variable)}$$
$$\Rightarrow \hat{\sigma} = \bar{s}/C_2$$
$$\text{where } C_2 = \sqrt{\frac{2}{n}}\frac{\left(\frac{n-2}{2}\right)!}{\left(\frac{n-3}{2}\right)!}$$

Then the 3σ control limits for $\bar{x}$ chart are given by:

$$\begin{aligned} UCL &= \bar{\bar{X}} + \left(\frac{3}{\sqrt{n}C_2}\right)\bar{s} = \bar{\bar{X}} + A_1\bar{s} \\ CL &= \bar{\bar{X}} \qquad \text{where } A_1 = \frac{3}{\sqrt{n}C_2} \\ LCL &= \bar{\bar{X}} - \left(\frac{3}{\sqrt{n}C_2}\right)\bar{s} = \bar{\bar{X}} - A_1\bar{s} \end{aligned} \qquad (22)$$

Again A_1 is a constant depending on the sample size n whose values for different values of n are given in Annexure-I.

Control Limits for R Chart

Let $E(R) = \overline{R}$ *and* σ_R be estimated by $\sigma_R = cE(R) = c\overline{R}$ where c is a constant depending on n, then the control limits for R-Chart are given by:

a) **When standards are given:**

$$\begin{aligned} UCL &= D_2\sigma \\ CL &= \overline{R} \\ LCL &= D_1\sigma \end{aligned} \qquad (23)$$

where D_1 and D_2 are constants and their values for different values of n are given in Annexure– I

b) **When standards are not given:**

$$\begin{aligned} UCL &= \overline{R} + 3c\overline{R} = (1+3c)\overline{R} = D_4\overline{R} \\ CL &= \overline{R} \\ LCL &= \overline{R} - 3c\overline{R} = (1-3c)\overline{R} = D_3\overline{R} \end{aligned} \qquad (24)$$

where D_3 and D_4 are constants and their values for different values of n are given in Annexure– I

Control Limits for σ Chart

Let a sample be drawn from a population with standard deviation σ then

$$\begin{aligned} & E(s^2) = \frac{n-1}{n}\sigma^2 \\ & \Rightarrow E(s) = C_2\sigma \\ & \text{where } C_2 = \sqrt{\frac{2}{n}}\frac{\left(\frac{n-2}{2}\right)!}{\left(\frac{n-3}{2}\right)!} \end{aligned} \qquad (25)$$

where, C_2, is a constant depending on the value of n. The different values of C_2 are given in Annexure – I

The first expression is true for samples drawn from any distribution whereas the second expression is true for normal distribution only.

When the samples are drawn from normal populations then

$$Var(s) = E(s^2) - \{E(s)\}^2 = \left[\frac{n-1}{n} - C_2^2\right]\sigma^2$$

$$S.E.(s) = C_3\sigma \text{ where } C_3 = \sqrt{\frac{n-1}{n} - C_2^2} \qquad (26)$$

a) **When standards are given:**

$$\text{UCL} = (C_2 + 3C_3)\sigma = B_2\sigma$$
$$\text{CL} = C_2\sigma$$
$$\text{LCL} = (C_2 - 3C_3)\sigma = B_1\sigma \qquad (27)$$

where, B_1 and B_2, are constants depending on the value of n and their values for different values of n are given in Annexure – I

b) When standards are not given:

$$UCL = \left(1 + \frac{3C_3}{C_2}\right)\bar{s} = B_4\bar{s}$$
$$CL = \bar{s} \qquad (28)$$
$$LCL = \left(1 - \frac{3C_3}{C_2}\right)\bar{s} = B_3\bar{s}$$

where, B_3 and B_4, are constants depending on the value of n and their values for different values of n are given in Annexure – I

Summary of out of control criteria

A process should be adjudged under control if points on both the charts i.e $\bar{X}$ and R or $\bar{X}$ and σ are distributed randomly within the 3σ limits. However, if some of the points go out of control on both the charts, it is best to start with R chart and remove the assignable causes of variation. In many cases this will automatically eliminate the nonrandom pattern on the $\bar{X}$ chart.

Sometimes, it may happen that though all points lie within the control limits, some of them follow specific patterns e.g. a run of seven or more points. By a run we mean consecutive points lying on one side of the average. The number of points in the run constitutes its length. These runs indicate an out of control situation. Also a run of points near 1σ or 2σ limits indicate an out of control situation. When the length of the run is too long, it implies that process average has changed. Also the number of runs on the control chart is a cause of concern whether too many or too few.

One may come across a continuous movement in one direction that is a trend on control chart. These are usually due to a gradual wearing out of a tool, operator fatigue or the presence of supervision or seasonal influences.

Cyclic patterns also sometimes appear on the control chart. Such a pattern on $\bar{X}$ chart may result from systematic changes in the temperature, operator fatigue, fluctuation in voltage or pressure, continuous rotation of operators or machines. Cycles on R charts are due to maintenance schedules, operator fatigue or tool wear.

Sometime points may cluster around the central line. This is generally due to incorrect calculation of control limits.

One may also come across a shift in the process level. These shifts are generally due to introduction of new workers, methods, raw materials or machines, a change in the inspection method or a change in the skill, attentiveness or motivation of the operator.

EXAMPLE 5

In the Petroleum production industry, the Management is interested in studying the API gravity of petrol as it found too much variation in its API values. The specifications for the API gravity are from 30 to 50. To determine whether the gravity specifications are met consistently, 20 groups of four samples each were collected at an interval of one hour. The following table gives the results:

Sub-Group number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Four samples Drawn at each hour	**I**	43	30	36	34	30	23	32	35	40	33	38	43	32	43	36	44	49	39	33	37
	II	36	32	35	37	37	33	33	34	38	35	39	35	33	44	38	45	46	37	35	36
	III	35	32	37	37	33	35	32	33	37	35	39	37	35	45	39	43	39	40	34	39
	IV	44	38	34	35	34	36	35	32	41	39	40	39	36	34	38	47	37	41	34	38

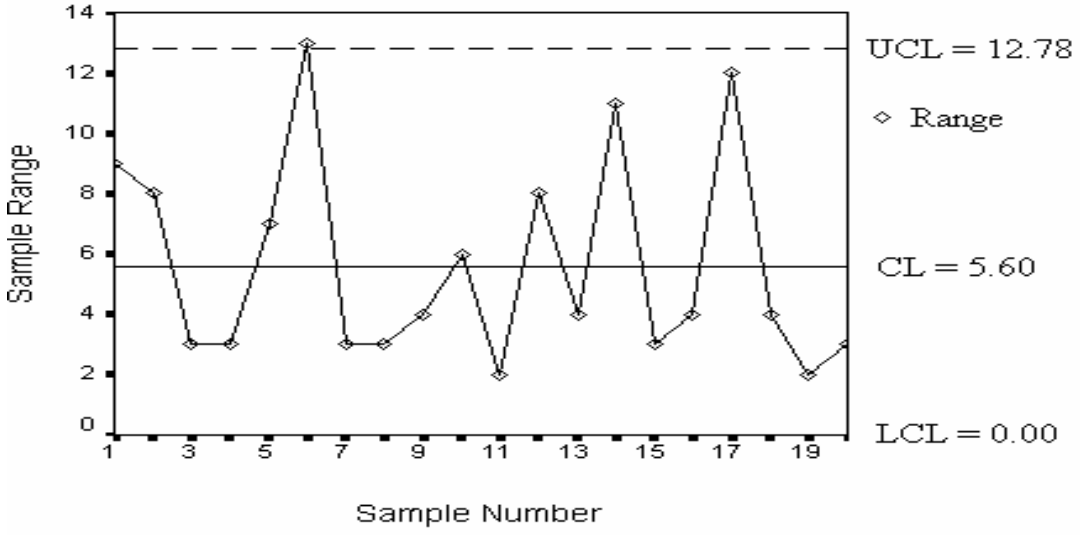
SOLUTION

The following table shows the $\bar{X}$ and R values for the given sample observations:

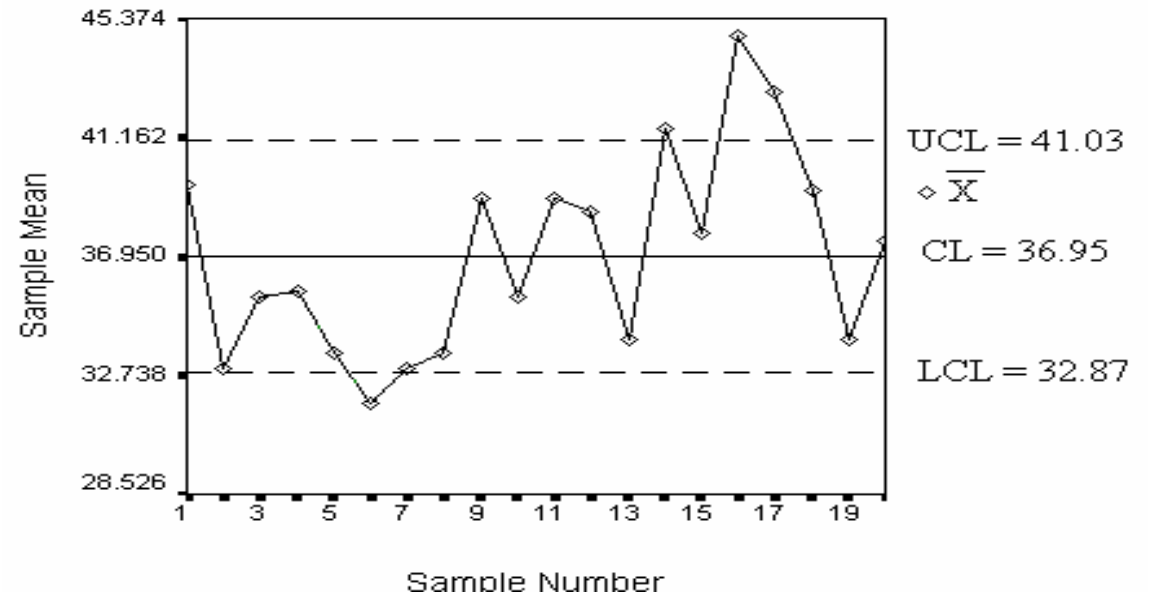
Sub-Group	***Four samples drawn* at each hour**					
Number	1	2	3	4	$\bar{X}$	R
1	43	36	35	44	39.50	9
2	30	32	32	38	33.00	8
3	36	35	37	34	35.50	3
4	34	37	37	35	35.75	3
5	30	37	33	34	33.50	7
6	23	33	35	36	31.75	13
7	32	33	32	35	33.00	3
8	35	34	33	32	33.50	3
9	40	38	37	41	39.00	4
10	33	35	35	39	35.50	6
11	38	39	39	40	39.00	2
12	43	35	37	39	38.50	8

Sub-Group	*Four samples drawn* at each hour					
Number	1	2	3	4	$\bar{X}$	R
13	32	33	35	36	34.00	4
14	43	44	45	34	41.50	11
15	36	38	39	38	37.75	3
16	44	45	43	47	44.75	4
17	49	46	39	37	42.75	12
18	39	37	40	41	39.25	4
19	33	35	34	34	34.00	2
20	37	36	39	38	37.50	3

The Range Chart



The $\bar{X}$ Chart



CONCLUSION

From the R-chart, it can be seen that point corresponding to only 6^{th} sample is going out of control, indicating slight deterioration in the process dispersion.

Next, to see whether process average is under control or not, we draw the $\bar{X}$-chart. From the chart we find that a number of sample points go out of control. Hence the process is not under control. Also sample number six needs special attention to know the cause of variation.

Cumulative sum control charts (cusum)

Till now we have studied Shewhart control charts that enable us to detect any variation in the process average or dispersion using information from the current group of observations. However, sometimes there are small shifts in the process average that are of concern, which may not be detected quickly by Shewhart control charts. The cumulative sum (CUSUM) control charts are quite useful in such cases. They help in detecting even small shifts in the process average from the target. They use the information given by the entire sequence of points by plotting cumulative sums of deviations of the observations from a target value. CUSUM means cumulative sum of the deviations from the target over the successive samples. If the sample observations are near the target value then the process is said to be operating on target and in this case subtractions are made from the cumulative sum. If the sample results are beyond the target value, the process is considered as not operating on target. In this case the difference is added to the cumulative sum. Again, if the cumulative sum exceeds the action limits the process is said to be out of control.

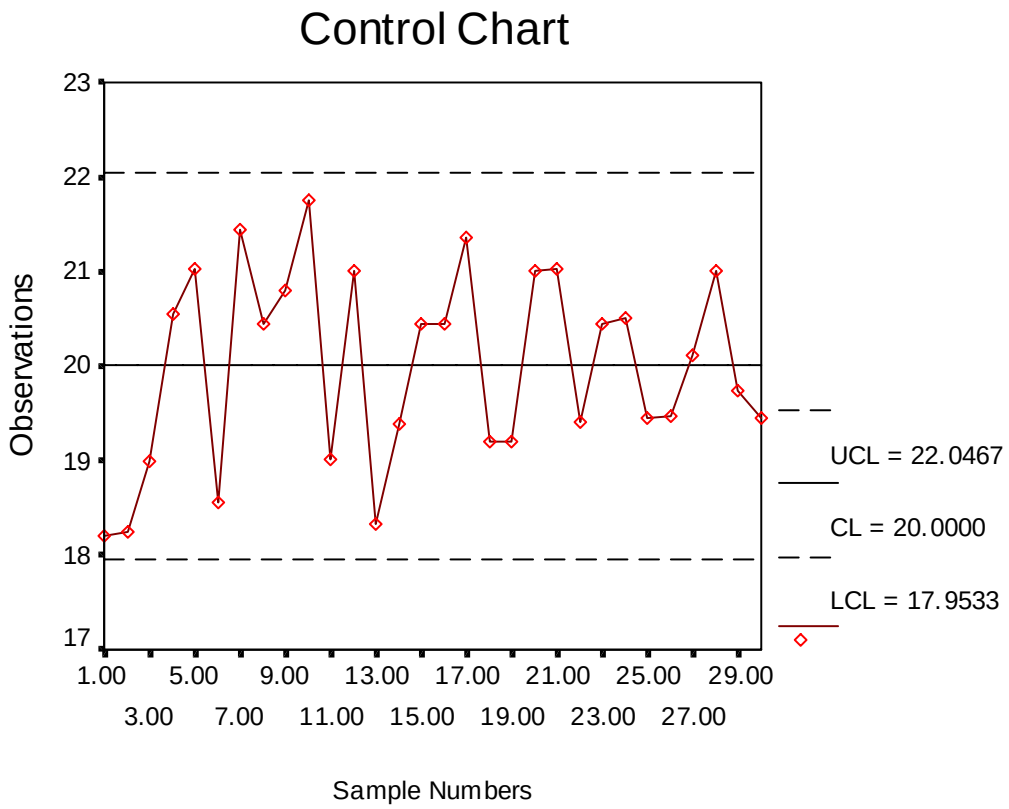
EXAMPLE

To observe the effect of shift in the average value of the process from target consider the following set of values:

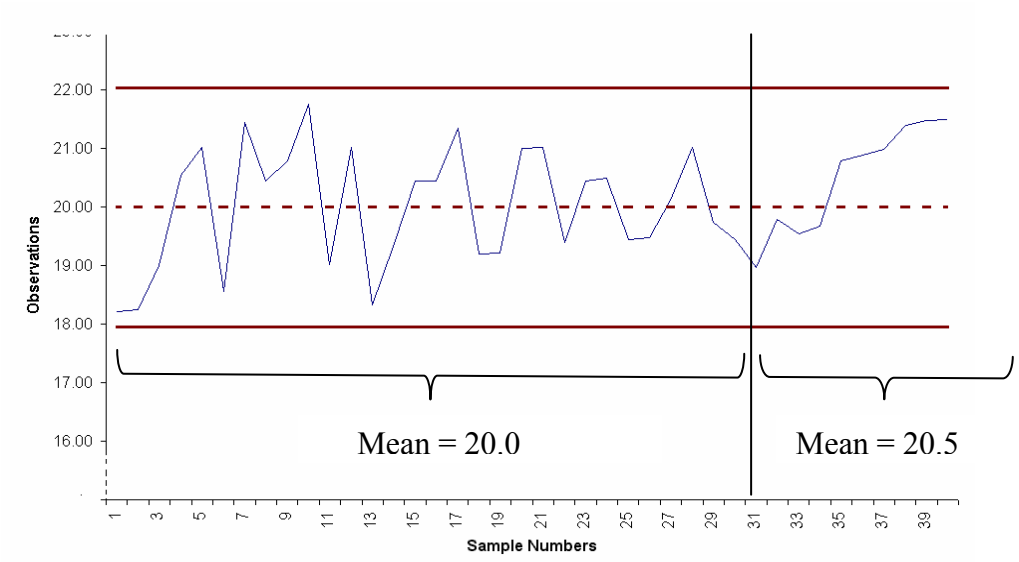
Sample No.	Observations	Sample No.	Observations
1	18.21	21	21.02
2	18.24	22	19.40
3	19.00	23	20.45
4	20.54	24	20.50
5	21.02	25	19.45
6	18.56	26	19.47
7	21.45	27	20.11
8	20.45	28	21.01
9	20.79	29	19.74
10	21.75	30	19.45
11	19.01	31	18.97
12	21.01	32	19.78
13	18.33	33	19.54
14	19.38	34	19.67
15	20.45	35	20.78

16	20.45	36	20.89
17	21.35	37	20.99
18	19.20	38	21.40
19	19.21	39	21.48
20	21.00	40	21.50

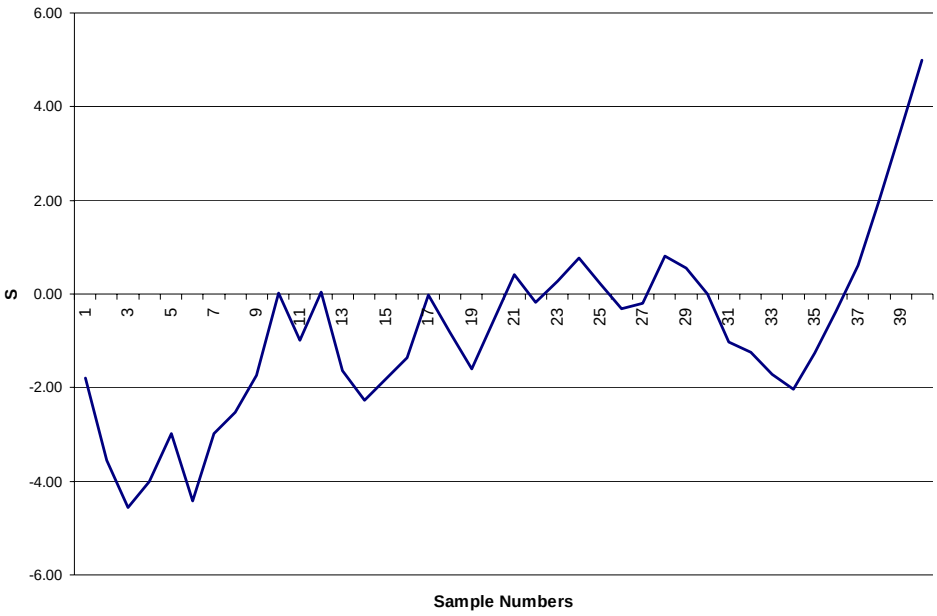
The first 30 samples were drawn from a normal population with mean 20.0 and s.d. 1. The control chart for the first 30 observations with LCL = 17.95, CL = 20.00 and UCL = 22.04 is as follows:



From the above chart it has been observed that the process is in control as all the observations are with in control limits. Further, the last 10 samples were drawn from a normal population with mean 20.5 and s.d. 1. When these 10 observations are plotted on the above control chart even when the target shifted from 20.0 to 20.5. The new control chart is shown below:



Next compute $S_i = (x_i - \mu_0) + S_{i-1}$ and plot a line graph corresponding to S_i. It has been observed an upward trend between the observations 31 to 40. The line graph is as follows:



Besides, the line graph representation one should also construct the CUSUM control chart. For the construction of CUSUM control charts three different points are considered which are:

1 deviation of sample from target value which is the difference between the sample value and target value
2 cumulative sum on the higher side (CSH)
3 cumulative sum on the lower side (CSL)

Besides the central line (Target), which is usually considered as zero, we consider four more lines denoted as:

A - the two lines defining the area above and below the target line where the process is assumed to be operating on target. Any deviations from the target lie in this area and result in subtractions from the cumulative sum of the deviations.

C -the two lines above and below A represent action lines or control lines. If the cumulative sum is greater than this value the process is said to be out of control.

If a deviation lies between A and C the process may be operating off target.

Construction of cusum charts for a given process

1 Determine the target value 'T'

2 If $\bar{X}_i$ denotes the average for the i^{th} sub-group (i = 1,….,m) then the process variation is estimated by computing

$$\hat{\sigma}_{\bar{x}} = \sqrt{\frac{\sum\left(\bar{X}_{i+1} - \bar{X}_i\right)^2}{2(m-1)}} \qquad (29)$$

3 Compute the value of allowable slack A as:

$$A = \frac{\hat{\sigma}_{\bar{X}}}{2} \qquad (30)$$

4. Compute the action limits C as:

$$C = 4.5\hat{\sigma}_{\bar{X}} \qquad (31)$$

5. Next for each subgroup, compute:

(i) the average $\bar{X}_i$, (i=1,…,m)

(ii) the deviation from target i.e. $D = (\bar{X}_i - T)$ (32)

(iii) the cumulative sum on the higher side as:

$$CSH_i = Max[0, D-A+CSH_{i-1}] \; ; \; CSH_0 = 0 \qquad (33)$$

(iv) the cumulative sum on the higher side as:

$$CSL_i = Max[0, -D-A+CSL_{i-1}] \; ; \; CSL_0 = 0 \qquad (34)$$

6. Draw the D, A, and C CSH_i and CSL_i lines on the control chart

7. Mark the values CSH_i and CSL_i computed in step 5 as points on the CUSUM chart.

8. If the process is out of control then the current process average can be computed as:

$$\text{Current Process Average} = \begin{cases} T + A + CSH/N_1 & \text{if } CSH > C \\ T - A + CSL/N_2 & \text{if } CSL > C \end{cases} \qquad (35)$$

where, N_1 is the number of consecutive subgroup for which CSH is greater then zero and N_2 is the number of consecutive subgroup for which CSL is greater then zero.

9. Eliminate the assignable causes if any and once the process is in control, repeat the chart for new control values by setting CSL = CSH = C/2

General form of cusum charts

The general form of CUSUM chart is shown in Fig. 2.

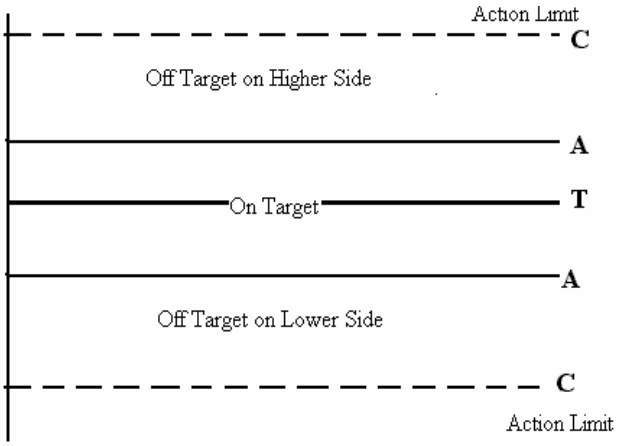
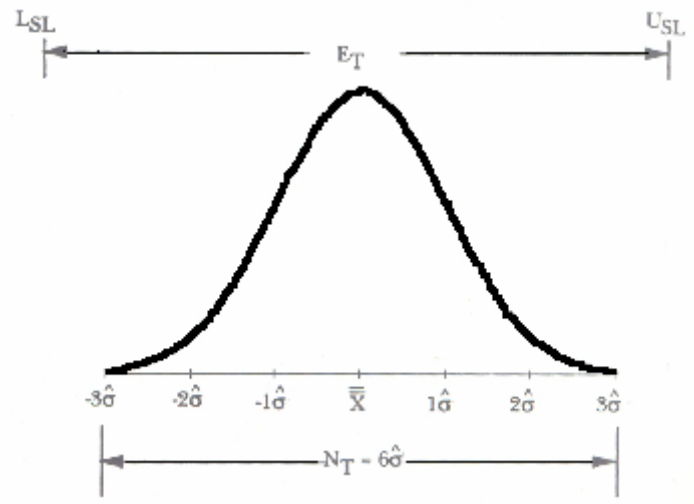


Fig. 2

Process capability measurements

If U_{SL} and L_{SL} be defined as the upper specification limit and the lower specification limit respectively, then the difference between the two specifications limits is called engineering tolerance (E_T). When the observed values are normally distributed one can denote N_T (natural tolerance) and E_T graphically as:



1. where $\bar{\bar{X}}$ is the estimate of population average μ and $\hat{\sigma}$ is the estimate of population standard deviation given by $\bar{R}/d_2$. If $E_T > N_T$, the process is capable of meeting the specifications.

Cp

There are various indices to measure process capability under the assumptions that the process is under statistical control, the observed values are normally distributed and the time interval over which data has been collected is the shortest. Cp is one of the methods to compute process capability, which is defined as the ratio of engineering tolerance to natural tolerance i.e.

$$Cp=\frac{E_T}{N_T}=\frac{U_{SL}-L_{SL}}{6\hat{\sigma}} \qquad (36)$$

In broader sense one may conclude:

If $Cp < 1.0$, the process is not capable of meeting the specifications.

If $Cp \geq 1.0$, the process is capable of meeting the specifications.

In practical situations the desired value of $Cp \geq 1.33$. Hence one may take a decision as follows:

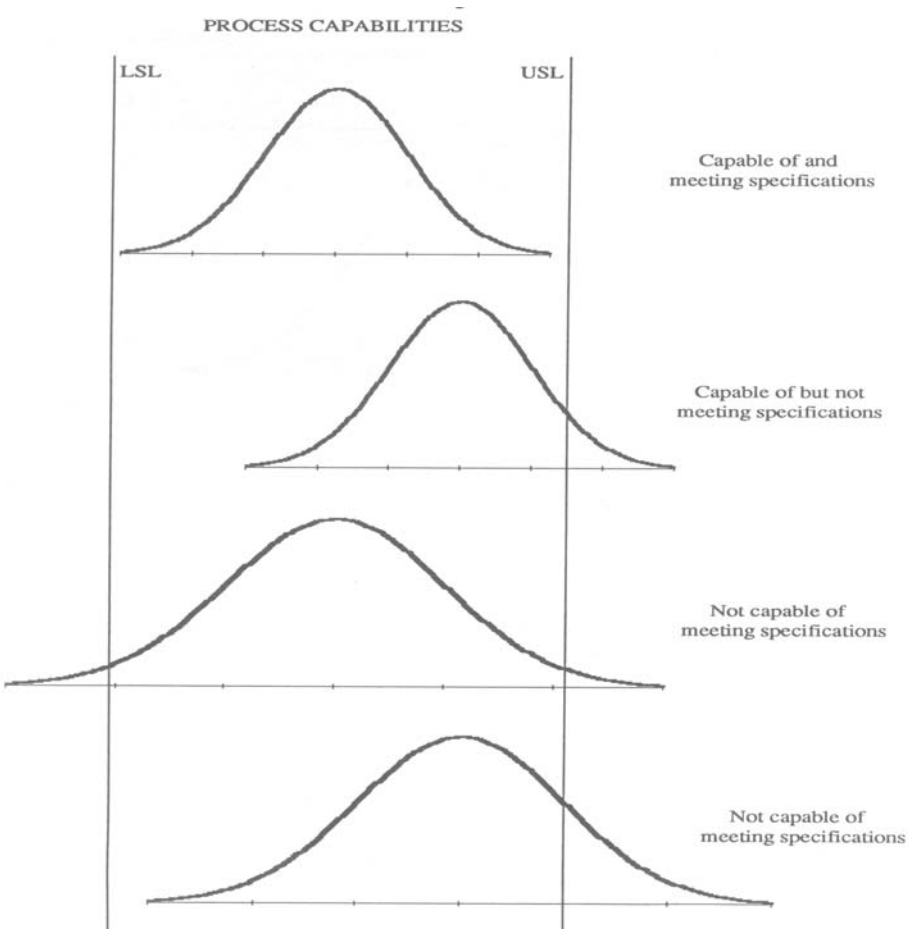
If $Cp \geq 1.33$, the process is capable of meeting the specifications.

$1 < Cp \geq 1.33$, the process is capable but needs close control and better centering as the value approaches to 1

$0.67 < Cp < 1$ the process is marginally capable of meeting the specifications.

$Cp \leq 0.67$ the process is not capable of meeting the specifications.

If the process is not centered relative to the specifications, the process can meet the specifications but is not meeting them. University of Tennesee (1986) has defined the various situations in the following figure:



To compute the percentage of product above the upper specification limit (or below the lower specification limit) one may use z table for the percentage of measurement for the value of z given by $z=(U_{SL}-\bar{\bar{X}})/\hat{\sigma}$ $(z=(\bar{\bar{X}}-L_{SL})/\hat{\sigma})$

where, $\bar{\bar{X}}$ is the overall process average and $\hat{\sigma}=\bar{R}/d_2$.

Cpk

The value of Cp $\geq$ 1.0 indicates that the process is capable but does not ensure that all the products produced is within the specification. This is because of the location of the process average which may be near to one of the specification limits. Thus does not give an idea whether the process is centered. Hence, the other index to measured process capability and ensure that the process is centered is Cpk. If Cpu and Cpl are the process capability based on U_{SL} and L_{SL} respectively then we define

$$\text{Cpk} = \text{minimum (Cpu,Cpl)}$$

Where, $\text{Cpu}=(U_{SL}-\bar{\bar{X}})/3\hat{\sigma}$, $\text{Cpl}=(\bar{\bar{X}}-L_{SL})/3\hat{\sigma}$ ($\bar{\bar{X}}$ is the overall process average) (37)

where, this takes into account where the process is centered. The desired value of Cpk is 1.0 and if

Cpk <1.0 the product is out of specification on any side.

Cpk may be calculated for either two-side or one-side specification limits whereas Cp can only be calculated for two-side specification limits.

Further, when the process average $\bar{\bar{X}}$ takes the value $(U_{SL}+L_{SL})/2$, the indices Cpk = Cp.

Practice problems

Question 1

An IC manufacturing company is interested in reducing the number of defects. Data on the number of defective IC's manufactured in each shift of output size 1000 are as given below. Develop an appropriate control chart.

Shift No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No. of Defectives	98	56	87	96	59	80	78	96	78	99	87	76	64	49	90	99	98	78	67	56

Question 2

In the above manufacturing unit suppose the number of IC's manufactured in each shift varies from 800 to 1000. Draw an appropriate control chart under variable sample size for the following data:

Shift No.	1	2	3	4	5	6	7	8	9	10
Shift Size	870	890	870	950	860	940	820	940	830	980
No. of Defectives	98	56	87	96	59	80	78	96	78	99

Shift No.	11	12	13	14	15	16	17	18	19	20
Shift Size	910	810	820	810	940	970	980	820	810	810
No. of Defectives	87	76	64	49	90	99	98	78	67	56

Question 3

A Packaging Company manufactures poly-film roles for making pouches for preserving vegetables. During manufacturing it was observed that the pores on the films are not exactly oval in shape and are considered as defects. The company is interested to know whether the process is in control. The quality assurance unit collected the following data. Draw an appropriate chart to take a decision on the process.

Length of Film Roll (in mts.)	10	12	14	16	12	17	13	15	17	18	19	17	20	18
Number of Defects	56	57	45	49	67	65	78	97	68	46	87	78	65	87

Also consider the case when the length of the film is constant i.e. 15 mts.

Question 4

A pen manufacturing company wanted to know whether the process of filling 15 ml. of ink in a refill is under control or not. To draw conclusion the quality control department drew 10 samples of size 5 each at an interval of one hour. The data collected is as follows:

S.N.	X1	X2	X3	X4	X5
1	14.89	14.87	13.98	14.99	14.67
2	13.89	13.56	14.23	14.38	13.78
3	13.89	13.45	13.89	13.97	14.98
4	14.65	14.98	15.02	15.21	13.67
5	14.87	15.23	15.11	14.34	13.99
6	15.00	15.32	14.23	13.78	13.99
7	13.78	14.76	14.45	15.00	15.12
8	14.99	15.02	15.00	14.99	13.89
9	14.23	15.00	15.11	13.98	14.67
10	15.00	15.01	15.12	14.99	14.98

Construct a control chart for mean and range and comments on the state of control.

References

1. Berenson, M.L. and Levine, D. M. (1992), Basic Business Statistics Concepts and Applications, Prentice Hall
2. Duncan A.J. (1967) Quality Control and Industrial Statistics. D.B. Taraporevala Sons & Co. Pvt. Ltd.
3. Gupta S.C. and Kapoor V.K. (1987), Fundamental of Applied Statistics, Sultan Chand & Sons.
4. Goon, A.M. Gupta, M.K. & Dasgupta, B. (1986), Fundamentals of Statistics, Vol. 2, The World Press Private Limited, Calcutta.
5. John, P.W.M. (1990) Statistical Methods in Engineering and Quality Assurance, John Wiley & Sons Inc.
6. Mcneese W.H. & Klein R.A. (1991) Statistical Methods for the Process Industries, ASQC Quality Press
7. Montgomery D.C. (1991), Introduction to Statistical Quality Control, John Wiley & Sons.

ANNEXURE

Factors for the Construction of Control Charts

	Mean Chart			Standard Deviation Chart					Range Chart				
Sample Size	Factors for control limits			Factors for central line	Factors for control limits				Factors for central line	Factors for control limits			
n	A	A_1	A_2	c_2	B_1	B_2	B_3	B_4	d_2	D_1	D_2	D_3	D_4
2	2.121	3.760	1.880	0.564	0	1.843	0	3.267	1.128	0	3.686	0	3.267
3	1.132	2.194	1.021	0.724	0	1.858	0	2.568	1.693	0	4.358	0	2.575
4	1.500	1.880	0.729	0.798	0	1.808	0	2.266	2.059	0	4.698	0	2.282
5	1.342	1.596	0.577	0.841	0	1.756	0	2.089	2.326	0	4.918	0	2.115
6	1.225	1.410	0.483	0.869	0.026	1.711	0.030	1.970	2.534	0	5.978	0	2.004
7	1.134	1.277	0.419	0.888	0.105	1.672	0.118	1.882	2.704	0.205	5.203	0.076	1.924
8	1.061	1.175	0.373	0.903	0.167	1.638	0.185	1.811	2.847	0.387	5.307	0.136	1.861
9	1.000	1.094	0.337	0.914	0.219	1.669	0.239	1.761	2.970	0.546	5.394	0.184	1.816
10	0.949	1.028	0.308	0.923	0.262	1.584	0.284	1.716	3.078	0.687	5.469	0.223	1.777
11	0.905	0.973	0.285	0.930	0.299	1.561	0.321	1.679	3.173	0.812	5.534	0.256	1.744
12	0.866	0.925	0.266	0.936	0.331	1.541	0.354	1.646	3.258	0.924	5.592	0.284	1.710
13	0.832	0.884	0.249	0.941	0.359	1.523	0.382	1.618	3.336	1.026	5.646	0.308	1.692
14	0.802	0.848	0.235	0.945	0.381	1.507	0.406	1.594	3.407	1.121	5.693	0.329	1.671
15	0.775	0.816	0.223	0.949	0.406	1.492	0.428	1.572	3.472	1.207	5.737	0.348	1.632
16	0.750	0.788	0.212	0.952	0.427	1.478	0.448	1.552	3.532	1.285	5.779	0.364	1.636
17	0.728	0.762	0.203	0.955	0.445	1.465	0.466	1.534	3.588	1.359	5.817	0.379	1.621
18	0.707	0.738	0.194	0.958	0.461	1.454	0.482	1.518	3.649	1.426	5.845	0.392	1.608
19	0.688	0.717	0.187	0.960	0.477	1.443	0.497	1.503	3.689	1.490	5.888	0.494	1.596
20	0.671	0.697	0.180	0.962	0.491	1.433	0.510	1.490	3.735	1.548	5.922	0.414	1.586
21	0.655	0.679	0.173	0.964	0.594	1.424	0.523	1.477	3.778	1.606	5.950	0.425	1.575
22	0.649	0.662	0.167	0.966	0.516	1.415	0.534	1.466	3.819	1.659	5.979	0.434	1.566
23	0.626	0.647	0.162	0.967	0.527	1.407	0.545	1.455	3.838	1.710	6.006	0.443	1.557
24	0.612	0.632	0.157	0.968	0.538	1.399	0.555	1.445	3.895	1.759	6.031	0.452	1.548
25	0.600	0.619	0.153	0.969	0.548	1.392	0.565	1.435	3.931	1.804	6.058	0.459	1.541