11-1-2005

The Individuals Control Chart in Case of Non-Normality

Betül Kan
Anadolu University, kanbetul@yahoo.com

Berna Yazici
Anadolu University, Turkey, bernayazici@yahoo.com

The Individuals Control Chart in Case of Non-Normality

Betül Kan   Berna Yazıcı
Department of Statistics
Anadolu University

This article examines the effects of non-normality as measured by skewness and provides an alternative method of designing individuals control chart with non-normal distributions. A skewness correction method for constructing the individuals control chart is provided. An example of thickness of biscuit process is presented to illustrate the individuals control chart limits.

Key words: Quality control, shewhart control charts (S), skewness correction (SC).

Introduction

Statistical methods provide many useful applications in industrial process control. One of them is the control charts method, used to detect the occurrence of assignable causes in industry, presented by Shewhart in 1924. The Shewhart control charts are based on the assumption of the distribution of the quality characteristic is normal or approximately normal.

The setting of the control limits to utilize on a control chart assumes the assumption of normality. However, in many situations, this condition does not hold. There are numerous studies on the control charts when the underlying distribution is non-normal.

Burr (1967) studied the non-normality on the constants for the $\bar{x}$ and R charts. For skewed populations, Type I Risk probabilities grow larger as the skewness increases. For highly skewed populations, one possible solution is to increase sample size.

Betul Kan is Research Assistant in Department of Statistics. Her research areas are quality control, mathematical statistics and design of experiments. E-Mail: kanbetul@yahoo.com. Berna Yazıcı is Assistant Professor in the Department of Statistics. Her research interests include design of experiments, regression analysis and quality control. E-Mail: bernayazici@yahoo.com.

Control limits for median, range, scale, and location charts for the Weibull distribution have been developed, and they match Shewhart's assumption for samples of four or more; normal theory is satisfactory. There has been a suggestion to use a weighted variance method based on the semivariance approximation, and asymmetric control chart limits for the $\bar{x}$ and R charts are available. The geometric midrange and geometric range charts has also been studied for a lognormal population. An arcsin transformation to transform a binomial variable into a normal one has been suggested.

Also, Bai and Choi (1995) proposed a heuristic weighted variance method with no assumption on the form of the distribution. Another approach for non-normality is considered using some transformed Q statistics which are the standard normal or approximately normal. Others chose to transform the data in order to make them quasi-normal, or used systems of the distributions as a general tool for transforming the data to normality.

In this study, the application of generalized Burr and Weibull distributions are demonstrated. First, non-normality and the assumption of the normal are mentioned. Second, the skewness method formulas for an individual control chart, when the distribution of the data is non-normal, is proposed. Third, a data set is illustrated as an example at the application study. Finally, the results of the studies are presented.

Methodology

Non-Normality

The assumption of normality means that the probability density function of the quality characteristic X has to be normal or approximately normal. However, this assumption may not be tenable. In many situations, there may be reason to doubt the validity of the normality assumption. For instance, the distribution of measurements from chemical processes or observations of lifetimes are often skewed. If the measurements are really normally distributed, the statistic $\bar{x}$ is also normally distributed. If the measurements are asymmetrically distributed, the statistic $\bar{x}$ will be approximately normally distributed only when the sample size n is sufficiently large (the central limit theorem).

However, the sample size n is usually chosen according to the sampling cost. Therefore, if the sample size n is not sufficient and the measurements are not normally distributed, then using the Shewhart control charts for constructing the control limits may reduce the ability of detecting the assignable causes of the control chart. In order to compute the control chart constants, Weibull and Burr distributions are chosen, because they represent a wide variety of shapes as parameters change.

The Burr Distribution

Burr proposed new $\bar{x}$ charts based on non-normality by using the Burr distribution to modify the usual symmetrical control limits. He developed the tables for constructing the modified control charts. He tabulated the expected value, standard deviation, skewness coefficient, and kurtosis coefficient of the Burr distribution for various combinations of c and q. The user can make a standardized transformation between a Burr variate (Y) and another random variate (X) by these tables. The Burr system covers a wide and important range of the standardized third and fourth central moments and can be used to fit a wide variety of practical data distributions.

The cumulative distribution function (CDF) of the Burr distribution is:

$$F(y) = 1 - (1 + y^c)^{-k} \quad y > 0 \quad (1)$$

where c and k are parameters ($c, k \geq 1$). Also, for large values of k, Burr distribution approaches the Weibull distribution.

The Weibull Distribution

The Weibull distribution is very flexible, and by appropriate selection of the parameters, the distribution can assume a wide variety of shapes. It is used extensively in reliability engineering as a model of time failure for electronic devices. The CDF of the Weibull's distribution is

$$F(y) = 1 - \exp(-[\lambda y]^\beta) \quad y > 0 \quad (2)$$

where λ and β are the scale and shape parameters, respectively.

The Individuals Control Chart

The individuals control chart, which may also be used to observe the magnitude of skewness, is used to find and remove the assignable causes quickly from the process if the sampling is possible, although $\bar{x}$-R or $\bar{x}$-s control charts can be constructed for the same situations.

The Process Distribution and the Parameters are Known

Let x_1, x_2, ...x_n be a sample from a process distribution whose mean is μ and standard deviation σ is known. The proposed the individuals control chart based on Skewness Correction (I_{sc} Chart) and the individuals control chart based on Shewhart (I_s Chart) methods are:

I_{sc} Chart:
$$UCL_{SC} = \mu + (3 + c_4^*)\sigma / \sqrt{n}$$
$$CL_{SC} = \mu$$
$$LCL_{SC} = \mu + (-3 + c_4^*)\sigma / \sqrt{n}$$

$$(3)$$

The constant c_4^* is the skewness correction.

$$c_4^* = \frac{\frac{4}{3}k_3(\overline{x})}{1 + 0.2k_3^2(\overline{x})}$$

(4)

$k_3(\overline{x})$ is the skewness (i.e., coefficient of skewness) of sample mean $\overline{x}$. When the process distribution is symmetric, $c_4^*=0$ and I_{SC} chart has the form of Shewhart chart.

The Process Distribution and the Parameters are Unknown

Let x_{i1}, $x_{i2, ... }$, x_{in}, i=1, 2,..., m, be m samples of size n from a process distribution with mean μ, standard deviation σ, and skewness k_3. When the process distribution is normal, Shewhart individuals control chart is

I_S Chart:
$$UCL_S = \overline{\overline{x}} + 3\overline{R}/d_2$$
$$CL_S = \overline{\overline{x}}$$
$$LCL_S = \overline{\overline{x}} - 3\overline{R}/d_2$$

(5)

Here,

$$\overline{\overline{x}} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}x_{ij},$$

$$\overline{R} = \frac{1}{m}\sum_{i=1}^{m}R_i,$$

R_i is the range of the i th sample. d_2 depends on the sample size n and calculated when the distribution is normal.

Let d_2^*, $d_2^* = \mu_R/\sigma_x$, be the constant for the given skewed process distribution corresponding to the role of the constant d_2 for the normal distribution. When $k_3=0$, the constant d_2^* is close to d_2 for the normal process distribution. Therefore, the skewness correction

method for individuals control chart is practically the Shewhart individuals control chart. The sample size n, is determined for the samples in the control process, the estimated value for the skewness of the sample averages

$$\hat{k}_{3\overline{x}} = \hat{k}_3 / \sqrt{n}$$

(6)

In many cases, the skewness k_3 needs to be estimated. It can be estimated by the sample skewness

$$\hat{k}_3 = \frac{1}{mn-3}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{x_{ij} - \overline{\overline{x}}}{\sqrt{\frac{1}{mn-1}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{\overline{x}})^2}}\right)^3$$

(7)

It is known that the estimator $\hat{k}_3$ of the k_3 is consistent. As well, it is known that the sample skewness $\hat{k}_3$ is essentially the third moment estimator. Based on the skewness correction method for the individuals control chart, the following is proposed

I_{SC} Chart:

$$UCL_{SC} = \overline{\overline{x}} + (3 + \frac{4k_3/(3\sqrt{n})}{1+0.2k_3^2/n})\frac{\overline{R}}{d_2^*} = \overline{\overline{x}} + A_{SC}^U\overline{R}$$

$$CL_{SC} = \overline{\overline{x}}$$

$$LCL_{SC} = \overline{\overline{x}} + (-3 + \frac{4k_3/(3\sqrt{n})}{1+0.2k_3^2/n})\frac{\overline{R}}{d_2^*} = \overline{\overline{x}} + A_{SC}^L\overline{R}$$

(8)

Table 1 gives the values of the constants A_{SC}^U and A_{SC}^L. If the skewness $k_3 < 0$, the A_{SC}^U is the same as the A_{SC}^L for $-k_3$ in Table 1, and vice versa. For instance, when n=2 and k_3=0,8, A_{SC}^U=3,35 and A_{SC}^L=2,08, with k_3= - 0,8, A_{SC}^U=2,08 and A_{SC}^L=3,35.

Table 1. The constants A_{SC}^U and A_{SC}^L for n=2,3,4,5,7,10

n	2		3		4		5		7		10	
k_3	A_{SC}^U	A_{SC}^L	A_{SC}^U	A_{SC}^L	A_{SC}^U	A_{SC}^L	A_{SC}^U	A_{SC}^L	A_{SC}^U	A_{SC}^L	A_{SC}^U	A_{SC}^L
0,0	2,66	2,66	1,78	1,78	1,46	1,46	1,30	1,30	1,11	1,11	0,98	0,98
0,4	3,03	2,36	1,96	1,59	1,64	1,38	1,41	1,19	1,19	1,03	1,04	0,92
0,8	3,35	2,08	2,17	1,45	1,74	1,22	1,52	1,12	1,27	0,98	1,11	0,89
1,2	3,69	1,87	2,37	1,33	1,90	1,14	1,65	1,03	1,38	0,93	1,17	0,82
1,6	4,00	1,73	2,58	1,25	2,06	1,08	1,77	0,98	1,48	0,87	1,23	0,79
2,0	4,27	1,63	2,77	1,18	2,20	1,02	1,90	0,94	1,56	0,85	1,33	0,79
2,4	4,51	1,58	2,93	1,13	2,36	0,98	2,03	0,89	1,67	0,79	1,39	0,73
2,8	4,70	1,60	3,08	1,11	2,48	0,94	2,12	0,87	1,75	0,77	1,45	0,70
3,2	4,88	1,64	3,22	1,11	2,58	0,94	2,24	0,85	1,83	0,77	1,52	0,66
3,6	4,98	1,70	3,33	1,13	2,68	0,94	2,33	0,83	1,90	0,74	1,58	0,66
4,0	5,08	3,36	3,41	1,14	2,78	0,94	2,39	0,83	1,98	0,71	1,61	0,38

If the k_3 value is not in this table, the user can take the nearest k_3 value or use interpolation. For instance, if a control chart for individuals n= 2 and k_3 = 1,35 is created, then one can take A_{SC}^U = 3,69 and A_{SC}^L = 1,87, or linearly interpolated values,

$$A_{SC}^U = \left(\frac{1,35-1,6}{1,2-1,6}\right)3,69 + \left(\frac{1,35-1,2}{1,6-1,2}\right)4,00$$
=0,625*3,69+0,375*4,00=3,806

$$A_{SC}^L = \left(\frac{1,35-1,6}{1,2-1,6}\right)1,87 + \left(\frac{1,35-1,2}{1,6-1,2}\right)1,73$$
0,625*1,87+0,375*1,73=1,817

Here, the user can also take the nearest k_3 value from Table 1, as k_3=1,2. Then, the values A_{SC}^U =3,69 and A_{SC}^L =1,87 are obtained.

Results

The measurement of the thickness of biscuit was obtained at a random time from a process line in a factory. ISO 9002 Quality

Assurance System is applied in this factory. The production of the dough is automatically materialized. The manufacturer uses an $\bar{x}$ chart to monitor the process. The data used in the study are independently taken in by equal time intervals. Based on an analysis of the skewness correction method, it is estimated that for n=5 with $\bar{\bar{x}}$=15,355 , $\bar{R}$=0,90 , k_3=0,051 from the Weibull distribution, by interpolation, the constants are A_{SC}^{U}= 1,30 and A_{SC}^{L}=1,29 from Table 1. It is noted that the constant is d_2=2,326 for n=5. Figure 1 shows the process for the Weibull distribution, n=5. Also the two different control limits are pointed out. The control limits of the individuals control chart based on the Shewhart method and skewness correction method are:

$$UCL_{SC} = \bar{\bar{x}} + A_{SC}^{U}\,\bar{R} = 16,525$$

I_{SC} Chart: $$CL_{SC} = \bar{\bar{x}} = 15,355$$

$$LCL_{SC} = \bar{\bar{x}} - A_{SC}^{L}\,\bar{R} = 14,194$$

$$UCL_S = \bar{\bar{x}} + 3\bar{R}/d_2 = 16,51$$

I_S Chart: $$CL_S = \bar{\bar{x}} = 15,355$$

$$LCL_S = \bar{\bar{x}} - 3\bar{R}/d_2 = 14,20$$

At the second stage, it is estimated that for n=7 with $\bar{\bar{x}}$=15,434, $\bar{R}$=1,065, k_3=0,056 for the Weibull distribution, by interpolation, the constants are A_{SC}^{U}=1, 11, and A_{SC}^{L}=1,11 from Table 1. It is noted that the constant is d_2=2,704 for n=7. Figure 2 shows the process for the Weibull distribution, n=7. Also, the two different control limits are pointed out. The control limits of the individuals control chart are based on the Shewhart method and skewness correction method which are:

$$UCL_{SC} = \bar{\bar{x}} + A_{SC}^{U}\,\bar{R} = 16,61$$

I_{SC} Chart: $$CL_{SC} = \bar{\bar{x}} = 15,434$$

$$LCL_{SC} = \bar{\bar{x}} - A_{SC}^{L}\,\bar{R} = 14,25$$

$$UCL_S = \bar{\bar{x}} + 3\bar{R}/d_2 = 16,61$$

I_S Chart: $$CL_S = \bar{\bar{x}} = 15,434$$
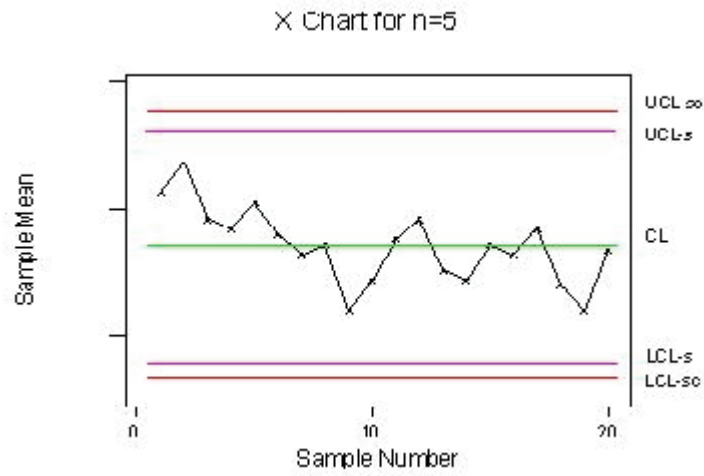
$$LCL_S = \bar{\bar{x}} - 3\bar{R}/d_2 = 14,25$$

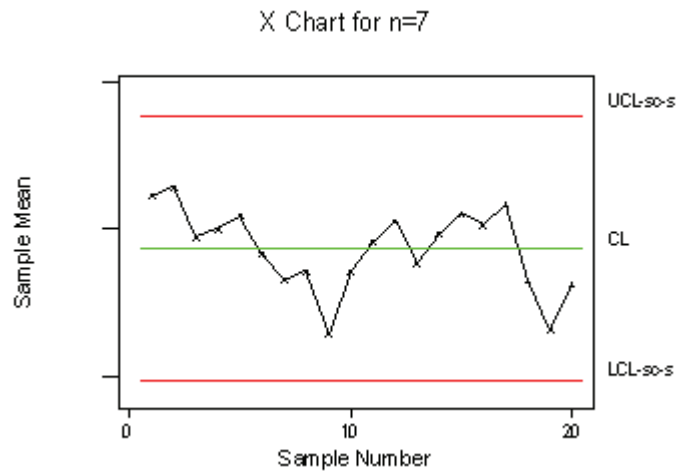Figure 1.Individuals chart for Weibull Distribution, n=5

Figure 2. Individuals chart for the Weibull Distribution, n=7

If a comparison is made between the control limits constructed by two methods, it may be said that for n=5, the range of control limits have expanded but for n=7 the range between the control limit have not changed.

On the other hand, based on an analysis of skewness correction method, it is estimated that for n=5 with $\bar{\bar{x}}$=14,599, $\bar{R}$=0, 74 from the Burr distribution, by interpolation, the constants are A_{SC}^{U}=1, 48 and A_{SC}^{L}=1, 14 from Table 1. For the set of data, the sample skewness =0, 67 and kurtosis=0, 64 coefficients are estimated, the tables given in Burr (1942) may be used to obtain the mean and the standard deviation of the corresponding Burr distribution. Then the data set may be described by the Burr distribution with the parameters c=3 and k=4. Figure 3 shows the process for the Weibull distribution, n=5. Also, the two different control limits are pointed out. The control limits of the individuals control chart based on the Shewhart method and skewness correction method are

I_{SC} Chart:

$$UCL_{SC} = \bar{\bar{x}} + A_{SC}^{U}\bar{R} = 15,69$$

$$CL_{SC} = \bar{\bar{x}} = 14,59$$

$$LCL_{SC} = \bar{\bar{x}} - A_{SC}^{L}\bar{R} = 13,75$$

I_{S} Chart:

$$UCL_{S} = \bar{\bar{x}} + 3\bar{R}/d_{2} = 15,51$$

$$CL_{S} = \bar{\bar{x}} = 14,59$$

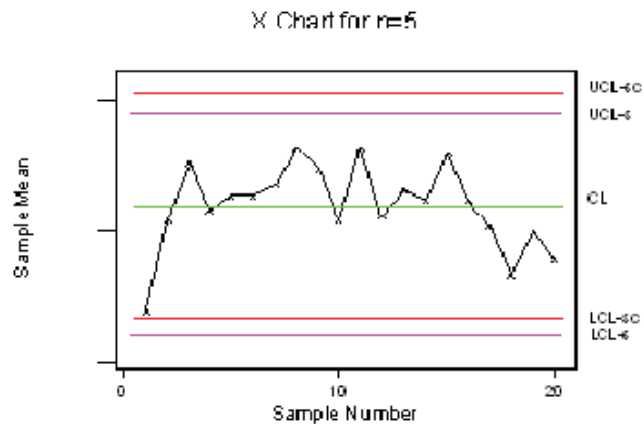$$LCL_{S} = \bar{\bar{x}} - 3\bar{R}/d_{2} = 13,64$$

Figure 3. Individuals chart for the Burr Distribution, n=5

At the second stage, it is estimated that for n=7 with $\overline{\overline{x}}$=14,646, $\overline{R}$=0,855 from the Burr distribution, by interpolation, the constants are A_{SC}^{U} = 1, 21 and A_{SC}^{L} =1, 02 from Table 1. For the data set, the sample skewness = 0,484 and kurtosis = 0,384 coefficients are estimated, the tables given in Burr (1942) may be used to obtain the mean and the standard deviation of the corresponding Burr distribution. Then, the data set may be described by the Burr distribution with the parameters c=3 and k=6. Figure 4 shows the process for the Weibull distribution, n=7. Also, the two different control

limits are pointed out. The control limits of the individuals control chart based on the Shewhart method and the skewness correction method are:

I_{SC} Chart:

$$UCL_{SC} = \overline{\overline{x}} + A_{SC}^{U}\,\overline{R} = 15,68$$

$$CL_{SC} = \overline{\overline{x}} = 14,64$$

$$LCL_{SC} = \overline{\overline{x}} - A_{SC}^{L}\,\overline{R} = 13,77$$

I_{S} Chart:

$$UCL_{S} = \overline{\overline{x}} + 3\overline{R}/d_{2} = 15,59$$

$$CL_{S} = \overline{\overline{x}} = 14,64$$

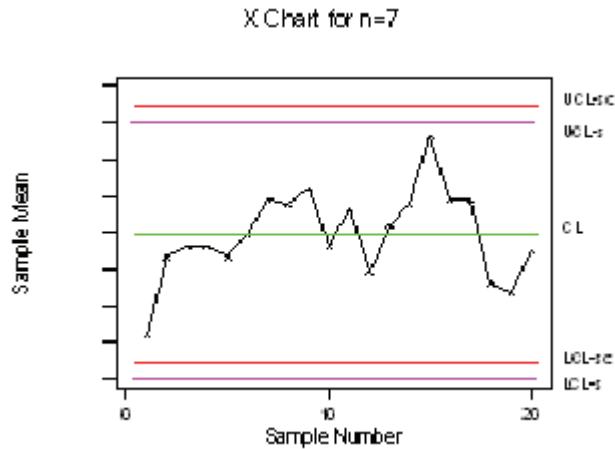$$LCL_{S} = \overline{\overline{x}} - 3\overline{R}/d_{2} = 13,70$$

Figure 4. Individuals chart for the Burr Distribution, n=7

If a comparison is made between the control limits constructed by two methods, it may be stated that for n=5 and n=7, the range of control limits have expanded. Differing from the Weibull distribution's results for n=7, the Burr distribution's results are not satisfactory.

Conclusion

The development of the skewness correction method for the individuals control charts when the distribution of the data is non-normal has been illustrated. The function used here depends on Chan and Cui's (2003) model. The Burr and the Weibull distributions are applied to derive the control limits. An example of the biscuit thickness process is presented to illustrate the individuals control chart. For n = 5, the individuals control chart limits obtained by the skewness correction method are larger than the limits of Shewhart method. However, for n = 7, when the data set is Weibull distributed, there is no difference between the two limit's values. There is a difference when the distribution of the data set is Burr distributed. In other words, for the Burr distributed data for n = 7, the limits in question are expanded for both methods.

In this study, according to the control charts which are compared, when the distribution of the data set is Weibull, the individuals control chart limits calculated by the skewness correction method are close to individuals control chart limits calculated by the Shewart method for the larger values of sample size n. Hence, if the data set is Weibull distributed, researchers may use individuals control chart with skewness correction instead of Shewhart individuals control chart.

References

Bai, D. S. & Choi, I. S. (1995). $\bar{x}$ and R control charts for skewed populations. *Journal Of Quality Technology, 27*(2).

Burr, W. I. (1967). The effect of non-normality on constants for $\bar{x}$ and R charts. *Industrial Quality Control*, 563.

Burr, W. I. (1973). Parameters for a general system of distributions to match a grid of α_3 and α_4. *Communications In Statistics, 2*(1), 1-21.

Burr, W. I, (1942). Cumulative frequency distribution. *Annals Of Mathematical Statistics, 13,* 215-232.

Chan, L. K., & Cui, H. J. (2003). *Skewness correction $\bar{x}$ and R charts for skewed distributions. Naval Research Logistics*. Wiley Periodicals. Inc, *50.*

Chou, C., Chen, C., & Liu, H. (2000). Economic-statistical design of $\bar{x}$ charts for non-normal data by considering quality loss. *Journal of Applied Statistics, 27*, 8.