A C T A U N I V E R S I T A T I S L O D Z I E N S I S
FOLIA OECONOMICA 5(325) 2016
http://dx.doi.org/10.18778/0208-6018.325.11

*Dorota Pekasiewicz**

ORDER STATISTICS IN CONTROL CHARTS CONSTRUCTION

Abstract. Order statistics are applied in construction of control charts that are used in statistical process control. Control charts monitor the average process level, volatility or extreme values: maximum and minimum. In the paper selected control charts are presented. They were used to analyze the quality of component of a home appliance which are produced by a home appliance manufacturer.

Keywords: median, extreme statistics, control chart.

JEL: C13, C18, L23.

1. INTRODUCTION

First applications of control charts in statistical process control area appeared at the beginning of the 20th century, but their main development dates back to the second half of the 20th century. Classical methods of analysis of processes variability, aimed at assessing the stability of the process and determining when the process needs to be adjusted and when to leave it without changes, often refer to the assumption of normality of the distribution and independence of the sample elements. Control charts are deemed an effective quality management tool, which offer graphic illustration of the measurement, usually in the form of aggregated averages and other statistics. Popular control chart is the $\overline{X}$ chart, which presents moving average value of the process. The *R* control chart and the *S* control chart are used to control the volatility of the process and the *p* control card is dedicated to control share of defective products in the production process.

More rarely the „robust" control charts are used in practice i.e. charts that are constructed with the use of the median of the sample or the extreme values and their functions (see Thomas et al. 2005). These charts are analysed in the paper and their application to observation of a production process is presented.

* Chair of Statistical Methods, University of Łódź, pekasiewicz@uni.lodz.pl

2. APPLICATION OF THE MEDIAN IN CONTROL CHARTS CONSTRUCTION

The median is the order statistic which can be used to construct the control charts that monitor average process level. These cards are called „robust” control charts.

Let us assume that we have k random samples $X_1^i, X_2^i, ..., X_n^i$, $i = 1,\ldots,k$, where n is sample size. Let us denote by $\overline{X}_i$ the arithmetic mean for i-th sample. On the basis of the sample means, the median and the central line for this chart are calculated and they have the following form:

$$LC_{Me} = Me = \begin{cases} \frac{1}{2}\left(\overline{X}^{(k)}_{\left(\frac{k}{2}\right)} + \overline{X}^{(k)}_{\left(\frac{k}{2}+1\right)}\right), & \text{when} \quad k \text{ is even,} \\ \overline{X}^{(k)}_{\left(\frac{k+1}{2}\right)}, & \text{when} \quad k \text{ is odd,} \end{cases} \tag{1}$$

where $\overline{X}^{(k)}_{\left(\frac{k}{2}\right)}$, $\overline{X}^{(k)}_{\left(\frac{k}{2}+1\right)}$, $\overline{X}^{(k)}_{\left(\frac{k+1}{2}\right)}$ are order statistics with ranks, respectively, equal to $\frac{k}{2}$, $\frac{k}{2}+1$ and $\frac{k+1}{2}$.

Control limits are defined as:

$$DLK_{Me} = LC_{Me} - 3\frac{a(n)c(n)\tilde{s}}{\sqrt{n}}, \tag{2}$$

$$GLK_{Me} = LC_{Me} + 3\frac{a(n)c(n)\tilde{s}}{\sqrt{n}}, \tag{3}$$

where $\tilde{s}$ is the median of standard deviations s_i, for $i = 1,\ldots, k$, obtained from the n-element random sample and $a(n), c(n)$ are coefficients. For chosen sample sizes the values of $a(n), c(n)$ are presented in Table 1.

Table 1. Coefficients $a(n), c(n)$ for chosen sample sizes

n	3	4	5	6	7	8	9	10	15	20
$a(n)$	1.128	1.085	1.064	1.051	1.042	1.036	1.032	1.028	1.018	1.013
$c(n)$	1.065	1.037	1.026	1.0201	1.016	1.014	1.012	1.010	1.006	1.005

Source: Thompson et al. (2005).

Control charts can be constructed on the basis of other median estimators. One proposition is application of the Bernstein estimator of the form (see Zieliński 1999):

$$Me^{Brs} = 0.5^{k-1}\sum_{i=1}^{k}\binom{k-1}{i-1}X_{(i)}^{(n)}. \tag{4}$$

The control lines are defined by the following formulas:

$$LC_{Me}^{Brs} = Me^{Brs}, \tag{5}$$

$$DLK_{Me}^{Brs} = LC_{Me}^{Brs} - 3\frac{a(n)c(n)\tilde{s}}{\sqrt{n}}, \tag{6}$$

$$GLK_{Me}^{Brs} = LC_{Me}^{Brs} + 3\frac{a(n)c(n)\tilde{s}}{\sqrt{n}}, \tag{7}$$

where $\tilde{s}$ is the Bernstein estimator of standard deviations s_i, so $\tilde{s} = 0.5^{k-1}\sum_{i=1}^{k}\binom{k-1}{i-1}s_{(i)}^{(k)}$.

The median is applied to construct the standard deviation chart, which is also a "robust" control chart and it is applied in monitoring the volatility of the process level. The control limits are defined by the following equations:

$$DLK_{\tilde{s}} = \max\left(0, c(n)\tilde{s}\left(1-3\left[\sqrt{a(n)^2-1}\right]\right)\right), \tag{8}$$

$$GLK_{\tilde{s}} = c(n)\tilde{s}\left(1+3\left[\sqrt{a(n)^2-1}\right]\right), \tag{9}$$

where $\tilde{s}$ is the median of standard deviations s_i, for i = 1,…,k.

3. EXTREME STATISTICS IN CONSTRUCTION OF CONTROL CHARTS

The extreme statistics are also applied in constructing control charts in the statistical process control. These statistics are used in the R control chart and the $X_{\max} - X_{\min}$ chart, and the $Me - R$ chart (see Kończak 2007).

The R control chart is applied when the production process is characterized by large volatility i.e. parameters have a wide range. The central line has the following form:

$$LC_R = \overline{R} = \frac{\sum_{i=1}^{k} R_i}{k}, \tag{10}$$

where $R_i = X_{(n),i}^{(n)} - X_{(1),i}^{(n)}$, for $i = 1, 2, ..., k$, and $X_{(n),i}^{(n)}, X_{(1),i}^{(n)}$ are maximum and minimum from sample $X_1^i, X_2^i, ..., X_n^i$.

Control limits are defined by formulas:

$$DLK_R = \max\left(0, \overline{R} - 3\frac{d_3(n)\overline{R}}{d_2(n)}\right), \tag{11}$$

$$GLK_R = \overline{R} + 3\frac{d_3(n)\overline{R}}{d_2(n)}, \tag{12}$$

and the values of coefficients for selected sample sizes are presented in Table 2.

Another control chart, which uses extreme statistics, is the $X_{\max} - X_{\min}$ chart. The central control line is defined by the *midrange*:

$$LC_{m,M} = M_r = \frac{\overline{X}_{(n)}^{(n)} + \overline{X}_{(1)}^{(n)}}{2}, \tag{13}$$

where $\overline{X}_{(n)}^{(n)} = \frac{1}{k}\sum_{i=1}^{k} X_{(n)i}^{(n)}$, $\overline{X}_{(1)}^{(n)} = \frac{1}{k}\sum_{i=1}^{k} X_{(1)i}^{(n)}$.

The lower and the upper control lines have the following forms:

$$DLK_{m,M} = M_r - \frac{3\overline{R}_{m,M}}{d_2(n)\sqrt{n}}, \tag{14}$$

$$GLK_{m,M} = M_r + \frac{3\overline{R}_{m,M}}{d_2(n)\sqrt{n}}, \tag{15}$$

where $\overline{R}_{m,M} = \overline{X}_{(n)}^{(n)} - \overline{X}_{(1)}^{(n)}$.

Table 2. Coefficients $a_2(n), d_2(n),\ d_3(n)$ for chosen sample sizes

n	3	4	5	6	7	8	9	10	15	20
$a_2(n)$	1.023	0.729	0.577	0.483	0.419	0.373	0.337	0.308	0.223	0.180
$d_2(n)$	4.358	4.698	4.918	5.078	5.204	5.306	5.393	5.469	5.571	5.921
$d_3(n)$	0.000	0.000	0.000	0.000	0.076	0.136	0.184	0.223	0.347	0.415

Source: www.zarz.agh.edu.pl/bsolins/karty_kontrolne.html.

The median and the range are applied to control the process by the $Me - R$ control chart, for which the central line is:

$$LC_{Me-R} = \overline{Me}, \tag{16}$$

where $\overline{Me} = \frac{1}{k}\sum_{i}^{k} Me_i$ and Me_i is the sample median.

Control limits are defined as:

$$DLK_{Me-R} = \overline{Me} - a_2(n)\overline{R}, \tag{17}$$

$$GLK_{Me-R} = \overline{Me} + a_2(n)\overline{R}, \tag{18}$$

where $\overline{R} = \frac{\sum_{i=1}^{k} R_i}{k}$ and values $a_2(n)$ are presented in Table 2 for chosen sample sizes.

If the random variable distribution is known, it can be used to modify the control charts the R and the $Me - R$. In such case the median and the range distributions can be estimated and its fixed quantiles are used to determine the lower and the upper limit of the control chart.

In monitoring the process observation of the average level or the volatility level are not only a subject of interest, but they also serve to assess the maximum or the minimum value of the random variable. Extreme statistics and their distributions are applied in construction of control charts to detect a shift of the extreme values of the process (see Kończak 2013). For example when the maximum has the Gumbel distribution, the distribution parameters are estimated on the basis of k maximum values $X_{(n)}^{(n)}$ from k samples and the central line is defined by the equation:

$$LC_{ExV} = \hat{\lambda} + 0.36655\hat{\delta}, \quad (19)$$

where $\hat{\lambda}$, $\hat{\delta}$ are estimators of the Gumbel distribution parameters.

Upper and lower limit lines are defined by quantiles of fixed orders i.e. (1 – *p*)-quantile and *p*-quantile. In this case control limits are defined as:

$$DLK_{ExV} = \hat{\lambda} - \hat{\delta}\ln(-\ln p), \quad (20)$$

$$GLK_{ExV} = \hat{\lambda} - \hat{\delta}\ln(-\ln(1-p)). \quad (21)$$

In the analysis of the processes with a heavy-tailed random variable distribution, the construction of control charts should take into account the median and the Fréchet distribution quantile. In this case the control chart lines are defined as follows:

$$LC_{ExV} = \hat{\lambda} + \frac{\hat{\delta}}{\sqrt[\hat{\gamma}]{0.693}}, \quad (22)$$

$$DLK_{ExV} = \hat{\lambda} + \hat{\delta}(-\ln p)^{-\frac{1}{\hat{\gamma}}}, \quad (23)$$

$$GLK_{ExV} = \hat{\lambda} + \hat{\delta}(-\ln(1-p))^{-\frac{1}{\hat{\gamma}}}, \quad (24)$$

where $\hat{\lambda}$, $\hat{\delta}$, and $\hat{\gamma}$ are estimated values of the Fréchet distribution parameters.

4. APPLICATION OF CONTROL CHARTS IN A HOME APPLIANCE MANUFACTURER

Application of the proposed Bernstein median estimator in control chart construction requires using former of its properties, e.g., simulation studies. The simulation procedure consisted of the following steps:

1) control limits were determined based on 100 generated samples from distribution $F(x)$, where sample size $n = 9$,

2) 20 samples were generated, q from $G(x)$ distribution and 20-q from $F(x)$ distribution,

3) the number observations of exceeding control limits was calculated,

4) the mean of the number observations of exceeding control limits for repeated 10000 times was calculated.

The normal distribution $N(\mu,\sigma)$, lognormal distribution $LN(\mu,\sigma)$ and Weibull distribution $W(\lambda,k)$, where $\lambda > 0$, $k > 0$, are considered in the simulation study.

The selected results of the studies are presented in Table 3.

Table 3. Estimated q for the selected cases

Distribution $F(x)$	q	Distribution $G(x)$	DLK_{Me}^{Brs}	GLK_{Me}^{Brs}	$\bar{q}$
N(45,1)	2	N(47,1)	43.973	46.028	2.034
	2	N(47,1.5)			1.888
	6	N(47,1)			6.018
	6	N(47,1.5)			5.872
LN(3,0.5)	2	LN(5,1)	11.195	33.306	2.210
	2	LN(5,2)			2.204
	6	LN(5,1)			6.158
	6	LN(5,2)			6.156
W(50,2)	2	W(48,3)	1.930	2.028	2.094
	2	W(48,5)			2.102
	6	W(48,3)			6.080
	6	W(48,5)			6.074

Source: own calculation.

The results of the investigation allow to state that behavior of the control chart with Bernstein median estimator is satisfactory – estimated the number observations of exceeding control limits is near the fixed q.

The control chart with Bernstein median estimator and others chosen charts were applied in monitoring process in a home appliance manufacturer. The length of a detail of a home appliance was analized variable. Based on 100 samples, whose size is 9 elements (3 measurements derived from each of three shifts) arithmetic means and standard deviations were calculated. Next, control charts monitoring average levels were constructed.

The limits (2)–(3) for the first one and (6)–(7) for the second were as follows:

$$LC_{Me} = 45.4288 \text{ (mm.)},\ DLK_{Me} = 45.3296 \text{(mm.)},\ GLK_{Me} = 45.5349 \text{(mm.)},$$

$$LC_{Me}^{Brs} = 45.4328 \text{ (mm.)},\ DLK_{Me}^{Brs} = 45.3312 \text{(mm.)},\ GLK_{Me}^{Brs} = 45.5363 \text{ (mm.)}.$$

Control limits with Bernstein estimator significantly differ from control limits with a standard estimator of the median. Analysis of the process using two control charts gave identical results. Figure 1 presents the results of application of the control chart with Bernstein estimator. It can be seen that this process is not regular because of average level.

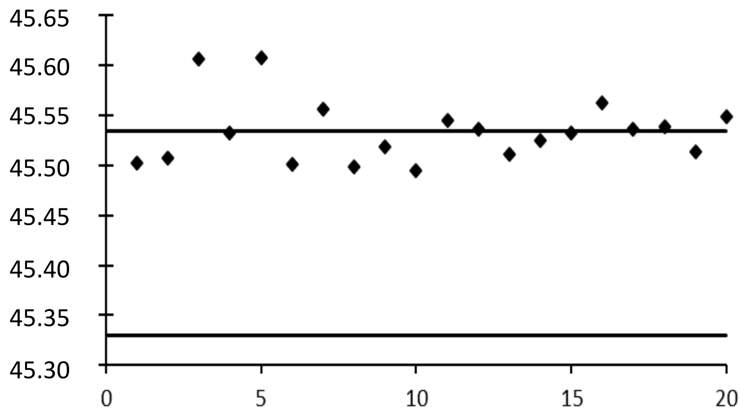

Figure 1. The results of application control chart for the average with limits (6) and (7)

Source: own elaboration.

Similar analyses were performed to study the volatility of constructing control chart with limits (8) and (9). The results are presented in Figure 2.

Another study concerned the stability of the process in terms of the maximum value.

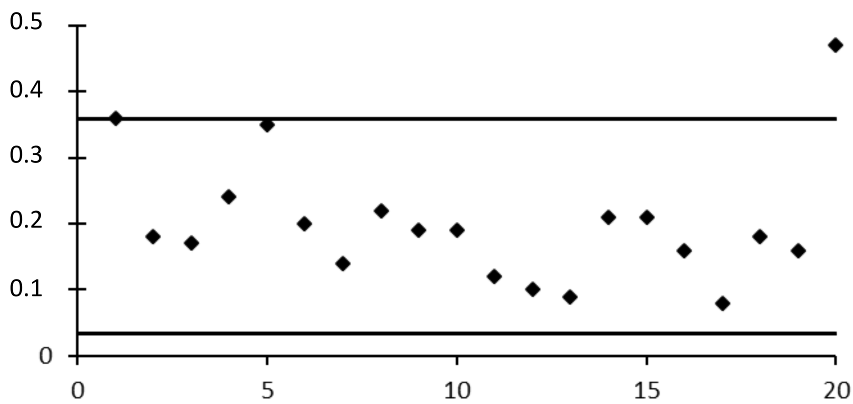

Figure 2. The results of application control chart for the volatility with limits (8) and (9)

Source: own elaboration.

The stability of the process of the maximum value was also analysed.

Gumbel distribution parameters were estimated based on 100 observations using the percentile method. The 0.25-quantile and 0.75-quantile were selected in this method, because the mean squared errors of Gumbel parameter estimators are the smallest for p and $1-p$ quantiles when p is near to 0.2–0.3 (see Pekasiewicz, (2015)). The 0.01 and 0.99 quantiles of Gumbel distribution were control limits (formulae (20), (21)). The results for 20 data are presented in Figure 3. This process is regular because of volatility level and maximum level.

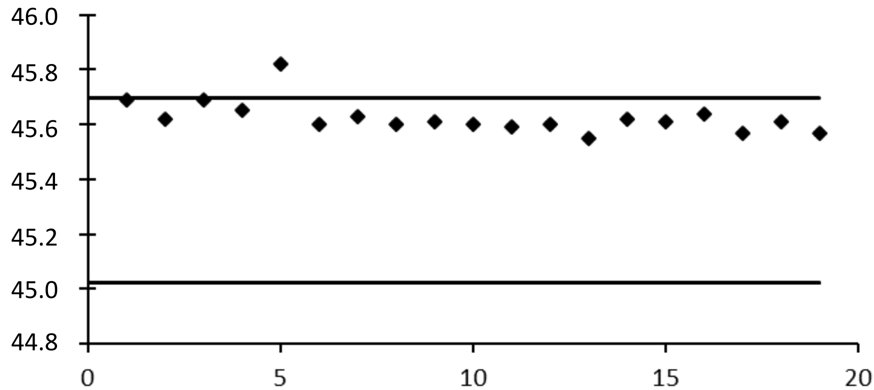

Figure 3. The results of control chart application for the maximum with limits (20) and (21)

Source: own elaboration.

5. CONCLUSIONS

The order statistics are applied in construction of robust control charts which monitor the production process in terms of average level, volatility level and maximum or minimum values. The chosen charts are presented in this paper. The results of simulation studies show that the control chart with Bernstein median estimator gives a good performance. It indicates that it can be applied in investigation of the production quality of home appliance component. Besides this chart, a few other presented charts were used to analyse this process. These control charts are a useful tool in the statistical process control for the diagnostic variables characterized by non-normal distribution.

REFERENCES

Kończak G. (2007), *Metody statystyczne w sterowaniu jakością produkcji*, Wydawnictwo Akademii Ekonomicznej im. K. Adamieckiego, Katowice.

Kończak G. (2013), *On the Use of the Extreme Value Distribution in Monitoring Production Processes*, "Acta Universitatis Lodziensis. Folia Oeconomica", 286, 307–315.

Pekasiewicz D. (2015), *Statystyki pozycyjne w procedurach estymacji i ich zastosowania w badaniach ekonomicznych*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Thompson J.R., Koronacki J., Nieckuła J. (2005), *Techniki zarządzania jakością od Shewhardta do metody „Six Sigma"*, Akademicka Oficyna Wydawnicza Exit, Warszawa.

Zieliński R. (1999), *Best Equivalent Nonparametric Estimator of Quantile*, "Probability & Statistic Letter", 45, 79–84.

Dorota Pekasiewicz

STATYSTYKI POZYCYJNE W KONSTRUKCJI KART KONTROLNYCH

Streszczenie. Statystyki pozycyjne mogą być stosowane do konstrukcji kart kontrolnych, za pomocą których analizuje się przeciętny poziom wielkości produkowanego elementu, czy też poziom jej zróżnicowania lub maksymalnej lub minimalnej jej wartości. W pracy zaprezentowane zostały wybrane karty kontrolne, które zastosowane zostały do analizy jakości elementu urządzenia gospodarstwa domowego produkowanego przez pewną firmę.

Słowa kluczowe: mediana, statystyk ekstremalne, karta kontrolna.

JEL: C13, C18, L23.