Comparison of Grand Median and Cumulative Sum Control Charts on Shuttlecock Weight Variable in CV Marjoko Kompas dan Domas

N Musdalifah¹, S S Handajani², E Zukhronah³
¹ Department of Mathematics, Faculty of Mathematics and Natural Sciences. Sebelas Maret University.
²,³ Department of Statistics, Faculty of Mathematics and Natural Sciences. Sebelas Maret University.

E-mail: nurulmdlfh@student.uns.ac.id, ssulistijovati@yahoo.com, etizkzukrnonah@yahoo.co.id

Abstract. Competition between the homoneous companies cause the company have to keep production quality. To cover this problem, the company controls the production with statistical quality control using control chart. Shewhart control chart is used to normal distributed data. The production data is often non-normal distribution and occured small process shift. Grand median control chart is a control chart for non-normal distributed data, while cumulative sum (cusum) control chart is a sensitive control chart to detect small process shift. The purpose of this research is to compare grand median and cusum control charts on shuttlecock weight variable in CV Marjoko Kompas dan Domas by generating data as the actual distribution. The generated data is used to simulate multiplier of standard deviation on grand median and cusum control charts. Simulation is done to get average run lenght (ARL) 370. Grand median control chart detects ten points that out of control, while cusum control chart detects a point out of control. It can be concluded that grand median control chart is better than cusum control chart.

Keywords: grand median control chart, cusum control chart, ARL.

1. Introduction
Free markets make the trade transactions among the regions in the world grow easy and fast. It caused the homogeneous companies that offer their products on the same area on a competition. Quality assurance of the production is a way to compete with the homogeneous companies. Efforts used to keep and to improve the quality of the result product can be done with quality control process statistics (Ramadhani [6]).

A tool used in statistical quality control of the production process is the control chart. Control chart is used to determine the process of controlled or uncontrolled through the upper control limit (UCL) and the lower control limit (LCL). A control chart may indicate an out of control condition when one or more points fall beyond the control limits or when the plotted points exhibit some nonrandom pattern of behavior (Montgomery [5]).

Shewhart control chart is a control chart that used to the normal distribution. Production quality characteristics is often non-normal distributed. To overcome this problem, we can use
Table 1. Data structure of samples

|   | 1   | 2   | ... | n   |
|---|-----|-----|-----|-----|
| 1 | x1,1| x1,2| ... | x1,n|
| 2 | x2,1| x2,2| ... | x2,n|
|   | ... |     | ... | ... |
| m | xm,1| xm,2| ... | xm,n|

Grand median control chart. According to (Altukife [1]), grand median control chart use the number of observation exceeding the grand median.

A production data often occur small process shift. Cusum control chart is used because it’s sensitive in detecting small process shift. The comparison between grand median and cusum control charts is to get the sensitive control chart. We obtained it by detecting small process shift on non-normal distribution. This research use shuttlecock weight variable in CV Marjoko Kompas dan Domas.

2. Grand Median Control Chart

Grand median is a statistic used to measure the center of m sample each of size n observations (Altukife [1]). Let \( x_{i1}, x_{i2}, x_{i3}, ..., x_{in} \) are a sample of independent observations of size \( n \) taken at period-i, where \( i = 1, 2, ..., m \). Data structure of samples can be seen in Table 1.

Let \( e \) is the grand median calculated from the combined total of \( N \) observations, \( N \) is the \( m \) samples each of size \( n \). Assume that \( O_{1i} \) is the number of observations that exceeds \( e \) in the \( i^{th} \) sample. Hence, \( O_{2i} \) is the number of observations less than or equal \( e \) in the \( i^{th} \) sample. The probability occurrence of \( O_{1i} \) and \( O_{2i} \) is 0.5. Therefore \( O_{1i} \) for \( i = 1, 2, ..., m \) is binomial(n,0.5). A mean and variance of random variable of binomial distribution are

\[
\mu = \frac{n}{2} \tag{1}
\]
\[
\sigma^2 = \frac{n}{4} \tag{2}
\]

The lines for the grand median control chart are given as follows

\[
LCL = \mu - D\sqrt{\sigma^2} \tag{3}
\]
\[
CL = \mu \tag{4}
\]
\[
UCL = \mu + D\sqrt{\sigma^2} \tag{5}
\]

where \( D \) is a multiplier of standard deviation and it is chosen to produce the desired type I error probability value (Altukife [1]). Equations (1) and (2) are substituted to equations (3) and (4) results control limit of grand median control chart

\[
LCL = \frac{n}{2} - D\sqrt{\frac{n}{4}} \tag{5}
\]
\[
CL = \frac{n}{2} \tag{6}
\]
\[
UCL = \frac{n}{2} + D\sqrt{\frac{n}{4}}. \tag{6}
\]
3. Cumulative Sum Control Chart

Cumulative sum (cusum) control chart collects all the information from the first sample until the last sample. Suppose that samples of sizes \( n > 1 \) are collected and \( \bar{x}_i \) is the average of the \( i \)th sample, then \( \mu_0 \) is the target of the process mean. However, if the mean shifts upward to some value \( \mu_1 > \mu_0 \), then an upward or positive drift will be developed in the cumulative sum \( C_i \). Conversely, if the mean shifts downward to some \( \mu_1 < \mu_0 \) then a downward or negative drift in \( C_i \) will be developed. The size of the shift is measured in terms of the process mean and standard deviation given by

\[
\delta = \frac{\mu_1 - \mu_0}{\sigma}.
\]

Observation of cusum control chart can be formed using tabular cusum defined by Koshti [4] as

\[
C_i^+ = \max[0, \bar{x}_i - (\mu_0 + K) + C_{i-1}^+] \quad (7)
\]

\[
C_i^- = \max[0, (\mu_0 - K) - \bar{x}_i + C_{i-1}^-] \quad (8)
\]

where the initial value \( C_i^+ \) and \( C_i^- \) is zero. Otherwise \( K \) is reference value defined by

\[
K = \frac{\delta}{2}\sigma = \frac{\mu_1 - \mu_0}{2}.
\]

where \( \delta \) is the size of the shift. Cusum control chart is a good control chart used to detect a small process shift. Usually the small process shift is approximately less than 1.5\( \sigma \) (Ferdinant [3]). Control limit of cusum control chart are defined as follows

\[
LCL = -h\sigma \quad (9)
\]

\[
UCL = h\sigma. \quad (10)
\]

4. Average Run Length

Average run length (ARL) is the average number of points that must be plotted before a point indicates an out of control condition. ARL is calculated from the average simulation run length. Run length was defined as the number of samples required to find samples out of the control limit. Calculating of the ARL is defined by

\[
ARL = \frac{1}{Pr(rejected H_0|H_0 \text{ is true})} = \frac{1}{\alpha}, \quad (11)
\]

where \( \alpha \) is probability of type I error or probability of production process is uncontrolled when the process is in control (Montgomery [5]).

Different technique that used to calculate the ARL is Siegmund approximation written as

\[
ARL = \begin{cases} 
\frac{\exp(-2b\Delta) + 2b\Delta - 1}{2\Delta^2}, & \text{for } \Delta \neq 0 \\
\frac{b^2}{\Delta}, & \text{for } \Delta = 0 
\end{cases} \quad (12)
\]

with \( \Delta = \delta - k \) for \( C_i^+ \) and \( \Delta = -\delta - k \) for \( C_i^- \),

\[
b = h + 1.166,
\]

\[
k = \frac{K}{\sigma}.
\]

To obtained ARL of the two sided cusum is written as

\[
\frac{1}{ARL} = \frac{1}{ARL^+} + \frac{1}{ARL^-}.
\]

Comparison between the two control chart can be done if the both of control chart have the same ARL value. ARL value is selected by setting alpha value. Simulation the multiplier of standard deviation (\( D \) and \( h \)) is to get the expected of ARL. Simulation is done by generating data as the actual distribution data.
5. **Data Distribution Test**

One of data distribution test is Anderson-Darling test statistic. According to Fallo et al [2], the steps of the hypothesis test are as follows.

(i) **Hypothesis**
   
   \( H_0 \): Shuttlecock weight data followed certain distribution.
   
   \( H_1 \): Shuttlecock weight data unfollowed certain distribution.

(ii) **Significance Level**

   \( \alpha = 0.05 \).

(iii) **Critical area (CA)**

   \( CA = \{ A^* | A^* > A_\alpha \} \). \( H_0 \) is rejected if \( A^* \in CA \).

(iv) **Test Statistics**

   \[
   AD = -N - \sum_{i=1}^{N} \left( \frac{2i-1}{N} \right) \left[ \ln F(Z_t) + \ln(1 - F(Z_{N+1-t})) \right],
   \]

   where
   
   \( AD \): test statistics Anderson Darling,
   
   \( N \): number of observations,
   
   \( Z_t \): standardized variable, and
   
   \( F(Z_t) \): the probability of the standard normal cumulative distribution function in \( Z_t \).

Test for normal distribution data using a modification of the test statistic \( AD \) is written as

\[
A^* = AD(1 + 0.75 \frac{N}{N^2} + 0.25 \frac{N}{N^2}).
\]

Distribution test statistic having the smallest value of \( AD \) indicated that the distribution is a representative distribution (Wibawati et al. [7]).

6. **Research Method**

This research used primary data of shuttlecock weight variable that produced by CV Marjoko Kompas dan Domas in district Sukoharjo. Data were taken 50 samples, where each sample size 20. Before applying the both of control charts, it is determined a multiplier value of standard deviation \( D \) and \( h \). Multiplier value is selected to result in expected ARL. The following steps are to determine the value of \( D \).

(i) Generate data as the actual distribution data as much as 1000, then to calculate the grand median.

(ii) Set any limit value \( D \) and then to make the limit control chart.

(iii) Generate data as the actual data distribution with a sample size of 20, and then calculating the statistics \( O_{1i} \).

(iv) Statistics \( O_{1i} \) was plotted on limit control chart.

(v) If \( O_{1i} \) was inside of the control chart then back to plot the statistics \( O_{1i} \) on step 3 and if it was outside the control chart then defined as Run Length (RL).

(vi) Step 3 until step 5 was repeated 1.000 times.

(vii) Calculate ARL by sum the RL value and then it is divided in 1000.

The steps to determine the value of \( h \) cusum control chart are defined as follows.
(i) Determine the mean process shift
(ii) Determine the initial value of h and calculating ARL of cusum control chart
(iii) Repeat steps 1 and 2 to get the expected ARL.

Next applying both of controls chart on observation data. The steps are

(i) Apply the grand median control chart
   (a) Determining the grand median then search for observations in each sample whose value exceeds the grand median
   (b) The number of observations which value exceeds the grand median in each sample was plotted on a control limits of grand median control chart.
   (c) Count the points are outside the control limit.
(ii) Apply the cusum control chart
   (a) Determine the mean process shift and the multiplier of standard deviation.
   (b) Calculate value of tabular cusum $C_i^+$ and $C_i^-$ as an observation point of cusum control chart.
   (c) Count the points are outside the control limit
(iii) Compare the both of control chart.

7. Result

Test of data distribution to determine the observations follow the certain distribution. Steps of hypothesis test using the test statistic AD was written as follows

(i) Hypothesis
   $H_0$ : Shuttlecock weight data followed normal distribution.
   $H_1$ : Shuttlecock weight data unfollowed normal distribution.
(ii) Significance level
   $\alpha = 0.05$.
(iii) Critical area (CA)
   $CA = \{A^* | A^* > A_\alpha = A_{0.05} = 0.752\}$. $H_0$ is rejected if $A^* \in CA$.
(iv) Test statistics.

$$A^* = AD\left(1 + \frac{0.75}{N} + \frac{0.25}{N^2}\right)$$

Table 2 showed $A^* = 1.068$.

(v) Conclusion
   Because $A^*$ is in CA so $H_0$ is rejected, it can be stated that shuttlecock weight data followed non-normal distribution.

The results of the statistic test AD on some distributions are shown in Table 2. Distribution test having the smallest value AD was 3-parameter Weibull distribution. In order that, hypothesis test using $H_0$ 3-parameter Weibull distribution data got statistic test $A^* = 0.648 < A_{0.05} = 2.492$. Because $A \notin CA$, $H_0$ was not rejected which it stated data followed 3-parameter Weibull distribution

Data distribution test was used to generate data as the actual distribution. The generated data used to the simulation of multiplier standard deviation was to get ARL 370. The multiplier of standard deviation of $D$ value gave amount 2.71 by equations (11) and $h$ value amount 4.87 that calculated from equation (12). Observations of grand median control chart are defined by $O_{1i}$. Equation (5) and (6) gave LCL amount 3.9043 and UCL amount 16.0597.
Table 2. The result of statistics test on some distributions

| Distribusi               | $AD$  |
|--------------------------|-------|
| Normal                   | 1.067 |
| Lognormal                | 0.762 |
| Exponential              | 442.195 |
| 2-p Exponential          | 176.951 |
| Weibull                  | 17.243 |
| 3-p Weibull              | 0.648 |
| Smallest Extreme Value   | 18.927 |
| Largest Extreme Value    | 5.705 |
| Gamma                    | 0.857 |
| Logistic                 | 1.384 |
| Loglogistic              | 1.237 |

Observations of cusum control chart are defined from tabular cusum value using equation (7) and (8). Equation (9) and (10) gave LCL amount $-0.2922$ and UCL amount $0.2922$ of cusum control chart. Grand median and cusum control charts can be seen in Figure 1 and Figure 2.

Figure 1. Grand median control chart

Figure 2. Cusum control chart

Figure 1 showed the control chart with LCL of grand median at 3.94 and UCL at 16.06 detect the 10 points that out of control. Figure 2 showed the cusum control chart with LCL at 0.2922 and UCL at 0.2922 detect a point that out of control. Observation points that out of control limit LCL and UCL on the grand median control chart were the sample 11, 24, 27, 29, 30, 31, 38, 39, 41, and 45, while on the cusum control chart was the sample to 18.

8. Conclusion

Simulation multiplier of standard deviation with the ARL 370 results $D$ value amount 2.71 and $h$ value amount 4.87. It’s to get the LCL and UCL the both of control charts. From the number out of control points it can be concluded that the grand median was better than cusum control chart on weight shuttlecock data that 3-parameter Weibull distribution. It means that the company produces the shuttlecock that qualified the specification, so the consumers are satisfied. The satisfaction of consumers can cause the increasing number of sale production. So the income of the company will be increased.
References
[1] Altukife, F.S., *A New Nonparametric Control Chart Based on Observations Exceeding The Grand Median*, Pakistan Journal of Statistic **19** (2003), no.3, 343-351.
[2] Fallo, J.O., Setiawan, and Susanto, B., *Perbandingan Uji Berdasarkan Anderson-Darling, Cramer, dan Lilliefors*, Prosiding Seminar Nasional Mathematika, UNNES, Semarang, (2013), 152-158.
[3] Ferdinant, P.F., *Pengendalian Kualitas Proses dengan Peta Kendali CUSUM dan EWMA*, Prosiding Seminar Nasional Industrial Services (SNIS) III, UNTIRTA, Banten, (2013), 147-152.
[4] Koshti, V.V., *Cumulative Sum Control Chart*, International Journal of Physics and Mathematical Sciences **1(1)** (2011), 28-32.
[5] Montgomery, D.C., *Introduction to Statistical Quality Control*, 6th ed., John Willey and Sons, Washington, 2009.
[6] Ramadhani, G.S., *Analisis Pengendalian Kualitas Menggunakan Diagram Kendali Demerit*, Skripsi, FMIPA Universitas Diponegoro, Semarang, 2014.
[7] Wibawati, Mashuri M and A.K. Gupta, *Diagram Kontrol Nonparametrik Mann Whitney untuk Target*, Skripsi, FMIPA Universitas Teknologi Surabaya, Surabaya, 2007.