**The effects of large round-off errors on the performance of control charts for the mean when the quality characteristic is normally distributed with a known variance**

**Background**

The Shewhart control charts are an important tool for monitoring processes in an ongoing manner. There is a natural variability in process variables in every production process due to random causes. A process where the variability is exclusively due to random causes is one that is statistically controlled. At times however, the variability not only stems from random causes but can also be attributed to problems with the machinery, a worker, the raw material, or other influencing factors. The manufacturer chooses the variables he would like to monitor, which are known as the quality characteristics. The primary role of control charts is to monitor parameters such as the mean and variance of the quality characteristics in an ongoing manner, the primary goals being to quickly identify deviations in the process, evaluate the process parameters, and determine its capacity, while minimizing the incidence of false alarms.

Deviations caused by random and systematic measurement errors are generally disregarded when using control charts. However, significant measurement errors are liable to effect the control chart and in some cases even lead to incorrect conclusions regarding the process control. This could be due to increased errors of a first type (the possibility of a false alarm, determining that the process has gone out of statistical control when there have been are no deviations), or increased errors of a second type (the possibility of not recognizing a deviation when one in fact exists).

The literature dealing with Shewhart’s classic control charts contains the assumption that the measured data are the true value of the monitored variable. Bennett (1954, XX) addressed the matter in his research, proposing that when the measurement error variance was smaller than the process variance it could be disregarded, as its effect on the strength of the control chart was minute. Abraham (1977, XX) proposed calculating the control charts without taking measurement errors into account and then adding a fixed value to represent the measurement error. Kanazuka (1986, XX) proposed increasing the sample size in order to solve the problem of the diminishing strength of control charts due to measurement errors. Walden (1990, XX) also suggested solving the problem by increasing the sample size. As another possible solution he proposed a process of repeated sampling, or a combination of increasing the sample size and performing repeated sampling. The results of their research along with others are mentioned in Linna and Woodall’s review article (2001).

Wheeler (2001a, XX) addresses control charts for features when products are incorrectly classified as intact or defective due to measurement errors. He proposes improving the quality of the production process rather than investing resources in a perfect measurement system, arguing that improving the measurement systems when sorting products so they match the specifications would increase the overhead, while improving the production process would lower costs.

The measurement process is influenced by many factors, such as environmental conditions, differences between measurers, rounding errors, calibration errors, errors in reading the results and measurement times. Therefore, when a certain variable is measured several times, there is likely to be a variance between the measurements. The different factors influencing this variance can be divided into two main categories: those stemming from the measurand itself, and those stemming from the measurement process and instruments, i.e. measurement errors (Gertsbakh 2003, XX).

Measurand Y can be described as the sum of two components:

(1)

X – the “true” value of the measurand

– the measurement error

The general description of the measurement can be expanded thusly:

(2)

– the mean of X

– the difference between the mean of X and its true value

– the noise of the true value, the random error

– the measurement error

The premise is that the two error components are non-dependent in relation to one another (ibid. 2003, XX).

In this paper we address measurement errors caused by rounding. Every measuring instrument displays measurements that are rounded-off, depending on the accuracy of the device. Round-off errors stem from two main causes. The first is the measurement device itself. Sometimes, due to technical, financial or other constraints, measurements are performed using cheap and fast measuring instruments with large measurement units. For example, if the measurement units of a scale are displayed in intervals of whole units (e.g. kilograms), the only results obtained will be whole numbers with no decimals, even though the value of the weighed variable could be a decimal (ibid. 2003, XX). The second cause is the accuracy of the systems receiving the measurement results. For example, the measurement device may send a 25-digit value but the system reading the data is only able to preserve 18 digits (Zhidong et al. 2009, XX).

One of the implications of rounding is that the final values obtained depend on the measurement units of the equipment and not the “true” value of the measured variable. This leads to rounded data: Y = with Y representing the rounded observations of X (Gertsbakh 2003, XX).

Rounded data affects statistical analysis because the rounding-off process itself leads to a data discretization process. In other words, the data goes from being treated as a continuous random variable to being treated as a discrete random variable, and this has a significant impact on statistical inference.

In many cases it is safe to disregard rounding; the parameters can be evaluated traditionally using rounded data and the evaluation is considered accurate enough for the purpose of statistical inference. However, in some cases disregarding the nature of the round-off process can lead to significant inaccuracies in the evaluation of the parameters, and as a result the use of theoretical statistical tools could lead to statistical inference errors (ibid. 2003, XX).

The degree of rounding is determined by the ratio between the standard deviation of the measured variable (σ) and the measuring unit of the measurement device (h). The ratio is denoted like so: δ ( ) with the measurement unit of the measurement device defined as the difference between two consecutive values on it, and it is generally presumed to be known.

The effect of a rounding error on the performance of the control chart depends on the size of δ. The smaller ratio δ is, the cruder the rounding and the greater its influence. Gertsbakh (ibid., XX) presented some rules of thumb for classification: when δ > 2 the measurement process is considered regular and allows for the use of traditional theoretical statistical tools. When δ < 0.5 the rounding is considered crude and the measurement process is considered special; in such cases the use of theoretical statistical tools may lead to incorrect statistical inference and therefore adaptations must be made. When 0.5 < δ < 2, the round-off level has no particular definition and this is considered an intermediate state (ibid., XX). The true value of the measurand, X, and the rounded results, Y, can be expressed thusly:

(3)

– the measurement error caused by crude rounding

We assume the normal measurement errors () and the random error () are negligible compared to the large rounding error, and therefore in Equation 3 replaces in Equation 1 and in Equation 2. The observations obtained are with .

Theoretically, Y can assume any value that is a multiple of h. However, in practice, when Y is measured with a measurement error stemming from crude rounding (δ < 0.5), it can assume no more than five different values with significant probability (greater than 0.00001). The other values obtained will have very little probability and can therefore be disregarded.

Under the assumption that the measurand prior to rounding, X, is distributed normally , the probability function of random variable Y is:

(4) 

(ibid., XX)

– the most frequent rounded value (the one with the highest probability).

Benson et al. (2013) studied the estimation of variance when the mean is known and the data is distributed normally with a measurement error stemming from crude rounding. Benson et al. (2014) then went on to study the estimation of variance using a confidence interval as well, for data normally distributed with a known mean that was collected with a measurement error stemming from crude rounding.

This paper aims to shed light on the effect of large rounding errors on the performance of the Shewhart control chart for the mean. The paper is structured as follows: in Section 2 we present the theoretical model for calculating the performances of Shewhart control charts for unrounded versus crudely rounded data. In Section 3 we compare the performances of control charts for rounded versus unrounded data. In Section 4 we present an analysis at the level of a single sample, using a simulation. Finally, in Section 5 we present our summary and conclusions.

**The theoretical model for calculating the performances of control chart for unrounded data and crudely rounded data**

In order to compare the performances of a control chart for unrounded data with its performances for rounded data, we performed several theoretical calculations for the following indexes:

Alpha (α) – an error of the first type – the probability that the sample mean will fall outside the control limits when in fact the process mean has not shifted.

Beta (β) – an error of the second type – the probability that the sample mean will fall within the control limits when in fact the process mean has shifted.

ARL0 – average run length; the expected number of samples set to run before the control chart signals an alarm, given that the process is under control.

ARL1 – the expected number of samples set to run before the control chart signals an alarm, given that the process is out of control (i.e. the mean has shifted).

We performed the calculations under the assumption that the observations were non-dependent and the process was normally distributed with a known standard deviation and a fixed sample size.

In the first stage the control limits were calculated for rounded and unrounded data, under the assumption that the process was controlled. When the data was unrounded, the indexes were calculated using formulas for calculating probabilities of a normal distribution, as is customarily done for Shewhart’s control charts (Montgomery 2013, XX).

(5)

The beta values were calculated under the assumption that the process mean had deviated by k standard deviations. Using the control limits that were calculated under the assumption that the process was controlled, we calculated the probability of the shift in the mean not being identified, that is, the probability of the sample mean falling within the control limits despite a shift in the process mean having occurred.

Formulaically, let k be the deviation of the mean from the original mean in standard deviation units, so that:

(6)

When the data is rounded it **is not** normally distributed, but discretely distributed according to the degree of rounding. Therefore, the indexes were calculated using a multinomial distribution. First we calculated the distribution of averages () that form the basis for all the calculations of rounded data. Then we calculated the general average () and the control limits according to which we measured the values of the four indexes: alpha (α), beta (β), ARL0 and ARL1,

The distribution of the averages ) was calculated in several stages: **In the first stage** the distribution of rounded data Y was built by finding the upper and lower limits of the *X* values [the original values (] for every possible Y (a half h from each direction). In addition, the probability of obtaining y was calculated by finding the difference between cumulative probabilities of the limits found using the normal cumulative distribution function.

(7)

We obtained a table with five values for Y in intervals of h in each direction from the most frequent value of Y, as well as the probability of each of these values (*P(Y*)). The other values had very little probability and could be disregarded. We represent the possible values for Y using a vector with five values (L):

with denoting the most frequent value (the one with the highest probability).

**In the second stage** sample size n was selected, and the mean was calculated for every combination of n possible values, under the assumption of the possibility of recurring values in the sample, and that the order in which the data was sampled was of no importance.

**In the third stage** we built a distribution table for the means (as several combinations could have the same mean)—the possible values and the probability of each of them occurring. This table served us later on in building the control chart.

The number of combinations (b) was determined by calculating the number of ways c out of d objects could be chosen, with repetitions and with no importance placed on the order. There are c+d-1 possible patterns, out of which c need to be chosen while minimizing the patterns that repeat themselves in a different order. In general:

Specifically, let us use LL to denote the number of possible values out of which values are chosen for a combination. In cases of crude rounding, there are no more than five values with significant probability, meaning d = LL = 5. The sample size is the number of objects selected (c = n), and we arrive at:

with

LL – as the length of vector L, LL = 5

n – as the sample size

b – as the number of combinations

For example, when the sample size is 7, the number of combinations is:

Let us mark:

i – as an index of a random variable in combination n…1

j – as an index of a combination out of total combinations b…1

Yi,j – as *the observation of* i *in combination* j

We calculated the average ( for every combination (j), and calculated the probability (P*j*) using the multinomial distribution formula:

with

*l* – as the index of values on vector L (1…5)

– as the number of times the value I of vector LL appears in combination j.

For example, if for n = 3 h = 2.5

The probability of each value on vector L

[0.001,0.1056,0.7887,0.1056,0.0001]

for combination (7.5, 10, 10)

In the next stage we calculated the probability of each possible outcome (the mean can be identical for several different combinations) by totaling the P*j*values of identical values. This stage resulted in a probabilities table for all the possible values. The values in the table appear in intervals of . In total the table has 4n + 1 values.

We calculated the theoretical mean for rounded numbers according to the formula put forth by Benson et al. (2013, XX):

The lower and upper control limits (*LCL, UCL*)were calculated according to Shewhart’s formulas for control charts for variables.

A = 3/sqrt(n)

The value of alpha was calculated by cumulative totaling of the probabilites of the values that were outside the control limits.

For every level of rounding, the beta values were calculated as dependent on the size of the shift of the mean in standard deviation units (k).

The method for calculating the distribution of means () given is identical to the method presented above. The value of beta was calculated by cumulatively totaling the probabilities of the values within the existing control limits.

**Comparing the performances of control charts for rounded versus unrounded data**

In this section we present several examples of theoretical calculations of performance indexes of control charts for unrounded and rounded data. The results presented are for unrounded values from a normal distribution with a mean of 10 and a standard deviation of 1, and for rounded data with rounding levels within the range of 0.3 < < 0.5, in intervals of 0.005, with a total of 41 rounding levels. For the purpose of illustration we chose three sample sizes, n = 7, 15 and 25. The sensitivity of the charts was tested with 12 deviation values (k), k = ± 0.1, 0.3. 0.5, 0.7, 1, 1.25.

Let us emphasize that only one value of the mean is presented, as the charts’ performances are influenced by the mean only through the value of their central line (CL). The value of the standard deviation also does not affect the values of the indexes. The size of the standard deviation bears no significance on alpha calculations, which are only influenced by the number of standard deviations according to which the control limits are determined. For beta calculations, the size of the deviation of the mean in relation to its size is important, and therefore we checked small, medium and large deviation values in relation to the size of the standard deviation we chose.

***Analysis of the results of the alpha and ARL0 values***

Whereas for unrounded data the value of alpha is determined by the user and is known in advance, for rounded data the distribution is not normal and the value of alpha is not set and known in advance. In the following charts we present comparisons of alpha values (Chart 1) and ARL0 values—the average run length before the alarm goes off (Chart 2), for rounded and unrounded data.

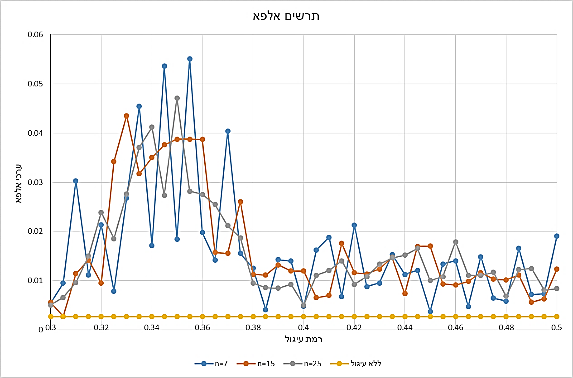


Chart 1. Alpha values for rounded and unrounded data as dependent on the rounding level and sample size.

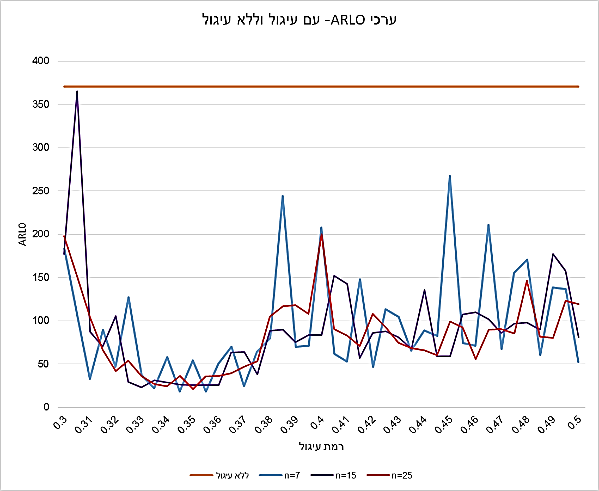


Chart 2. ARL0 values for rounded and unrounded data as dependent on the rounding level and sample size.

The charts above clearly demonstrate that the performances of the control chart are diminished when the data is rounded. The value of alpha is high for every rounding level within the studied range for each of the sample sizes tested, and accordingly the ARL0 values are significantly lower than the acceptable value. This results in a significantly higher number of false alarms that impede the continuity of work and the stability of the production line.

Two salient phenomena can be seen on these charts. Firstly, the spikes—the sharp ups and downs appearing on the chart. This phenomenon appears for every rounding level in all three sample sizes. Secondly, there is a difference between the alpha values when the rounding level is between 0.32–0.38 and when it is between the ranges of 0.3–0.32 and 0.38–0.5, with the damage to the performance being more prominent when the round-off level is in the range of 0.32–0.38.

In an attempt to understand the spiking in alpha values we repeated the experiment while **neutralizing the relative position** of the mean within the interval created as a result of the level of rounding. We set the mean precisely on the rounded value, in other words, for the mean of X to equal the most frequent value of Y. The most frequent value was set as and the other possible values were set around it.

L= [,, , , ]

As can be seen in Chart 3, we found that when the mean was set exactly on the most frequent values the chart obtained presents saw teeth behavior for each of the three sample sizes.

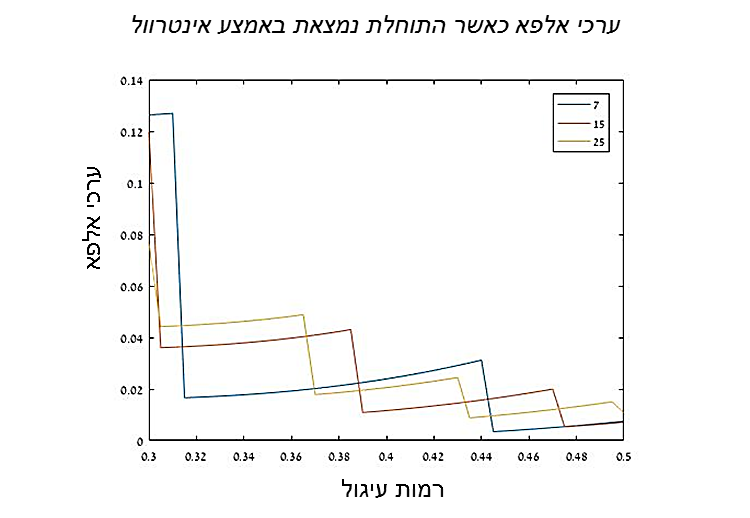


Chart 3. Alpha values as dependent on the level of round-off and sample size when the mean is mid-interval.

Alpha calculations depend on the rounding level, the sample size, and the control limits of the control chart for the mean.

Each rounding level δ has a fixed number of values (the average of a combination the size of n out of the possible values), for a specific scale size (h) and sample size (n). The size of the interval between values is (the difference between two adjacent values).

For example, for n = 7 and a rounding level of δ = 0.4 (i.e. h = 2.5), the interval size between values is . If we use D to denote the number of times the size appears within the control limits, we arrive at:

When the value of D is whole this indicates an interval, or in other words, the start of a new round of values in relation to the control limits. For example, for a sample size of 7, D receives a whole value of 5 when δ = 0.315. In the next round the size (δ = 0.4410) is entered once again on each side of the control chart so that D = 7. In this situation, the probability of being within the control limits significantly rises and therefore the alpha values drop.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **n=7** | **n=15** | **n=25** |
| **D** | **δ** | **δ** | **δ** |
| **3** | 0.1890 | 0.1291 | 0.1000 |
| **5** | 0.3150 | 0.2152 | 0.1667 |
| **7** | 0.4410 | 0.3012 | 0.2333 |
| **9** | 0.5669 | 0.3873 | 0.3000 |
| **11** | 0.6929 | 0.4734 | 0.3667 |
| **13** | 0.8189 | 0.5594 | 0.4333 |
| **15** | 0.9449 | 0.6455 | 0.5000 |
| **17** | 1.0709 | 0.7316 | 0.5667 |
| **19** | 1.1969 | 0.8176 | 0.6333 |
| **21** | 1.3229 | 0.9037 | 0.7000 |

Table 1. Rounding levels by sample sizes where there are complete rounds of values with the mean in the middle of the interval.

The table above shows the values of rounding levels that contain complete rounds of values within the range of the control limits and the central axis for the various sample sizes. It is possible to see that within the range of rounding levels studied, 0.3 < δ < 0.5, the number of times in which this situation occurs increases along with the sample size (when n = 7 it happens twice, as opposed to when n = 15, where it happens three times, and when n = 25, where it happens four times). In Chart 1 the two intervals of alpha values within the studied range for sample size 7 are at rounding levels 0.3150 and 0.4410, as noted in Table 1.

***Analysis of the beta and ARL1 values***

The following charts present beta values for unrounded data and rounded data at varioius round-off levels, according to sample and deviations sizes. Twelve sensitivity tests were conducted for six deviation sizes, k = ±0.1, 0.3, 0.5, 0.7, 1, 1.25

*Beta values for rounded and unrounded data*

Small sample size, n = 7

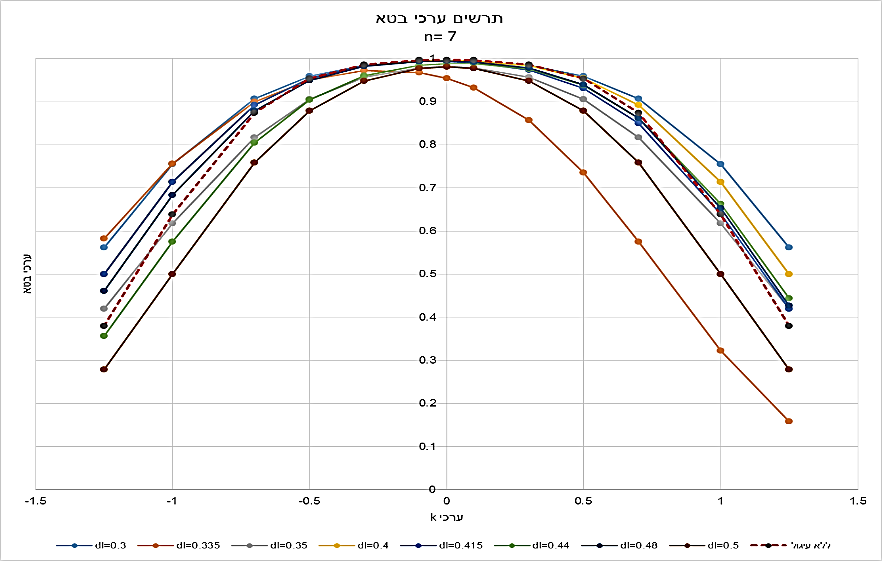


Chart 4. Beta values for unrounded and rounded data as dependent on the size of deviation and round-off level, n = 7.

Medium sample size, n = 15

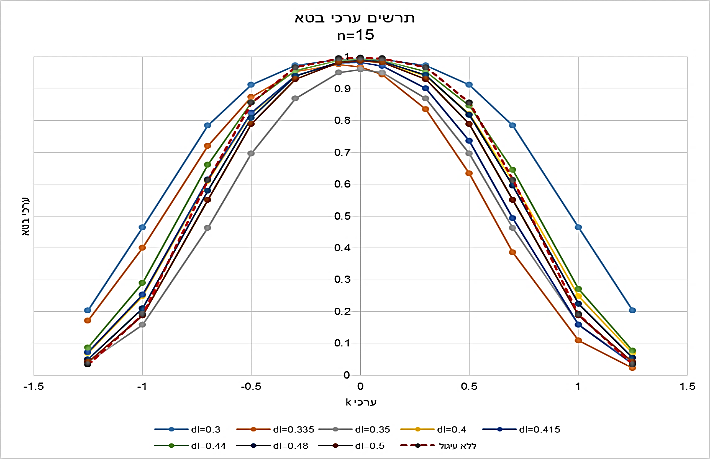


Chart 5. Beta values for unrounded and rounded data as dependent on the size of deviation and round-off level, n = 15.

Large sample size, n = 25

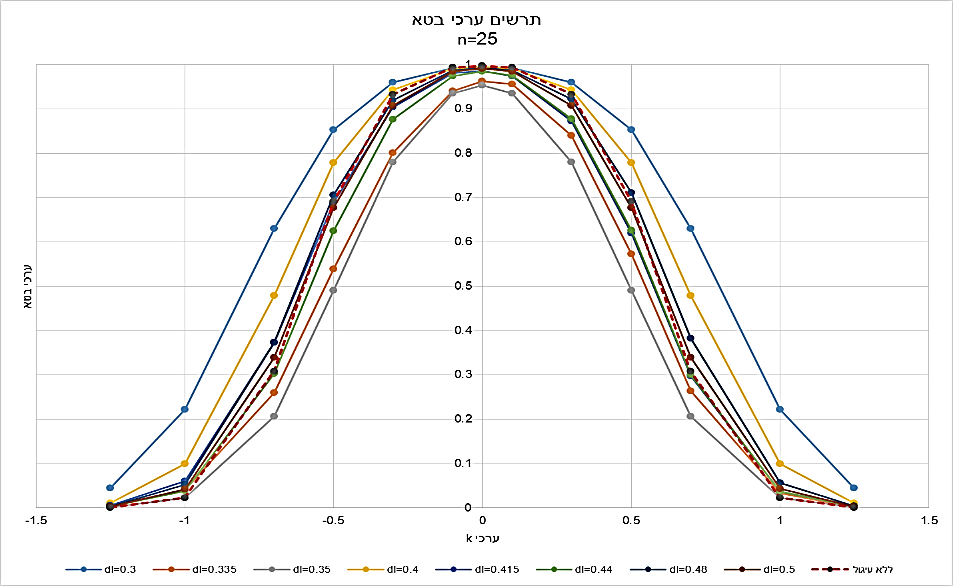


Chart 6. Beta values for unrounded and rounded data as dependent on the size of deviation and round-off level, n = 25.

These charts present beta values for eight different round-off levels, together with beta values for unrounded data (marked in red), for three sample sizes. The three charts demonstrate that in some cases the beta values of rounded data are smaller than the beta values for the unrounded data (values that are under the beta chart for unrounded data, represented by the dashed red line). This means that the control chart’s performances improved when the data was rounded. On the other hand, in other cases the beta values for the rounded data were higher than the beta values for the unrounded data (values that are above the beta chart for unrounded data, represented by the dashed red line), meaning that the chart’s performances were diminished when the data was rounded. It is important to note, and we will expound on this below, that the gain in terms of alpha values is significantly greater than the loss in terms of beta values.

When the data is not rounded there is complete symmetry in beta values for deviations from the mean at identical sizes and with opposite signs. When the data is rounded, the symmetry is maintained for some of the rounding levels while for others the symmetry is marred.

***The change in ARL0 versus the change in ARL1***

The following tables present the change in ARL0 as opposed to the change in ARL1 according to sample size, for the 12 sensitivity tests that were conducted. The change was calculated according to the difference between ARL values for unrounded data and ARL values for rounded data. For ARL0 a positive difference indicates diminished chart performances as there are less samples between false alarms when the data is rounded, whereas a negative difference indicates improved chart performances. For ARL1 on the other hand, a positive difference indicates improved chart performances, as it takes less samples to identify that the process is effectively no longer statistically controlled, compared to when the data is unrounded. Conversely, a negative difference in ARL1 values signifies a decline in the chart’s performances, as it takes more samples to recognize that the process is effectively no longer statistically controlled when the data is rounded.

|  |  |  |
| --- | --- | --- |
|  | ARL0 | ARL1 |
| Positive difference | Diminished chart performances | Improved chart performances |
| Negative difference | Improved chart performances | Diminished chart performances |

Table 2. The significance of the change in ARL0 and ARL1

The following table presents the results for round-off level 0.42 and three sample sizes:



Table 3. The change in ARL0 and the change in ARL1 for round-off level 0.42 and sample sizes 7, 15 and 25.

It is possible to see that the change in ARL0 is significant and only goes in one direction: there will be more false alarms in the monitoring process when the data is rounded. On the other hand, the change in ARL1 is smaller and goes in both directions. When the deviation from the mean is very small (k = ± 0.1, 0.3), there is an improvement in the chart’s performances, meaning the deviation will be discovered sooner. However, in these cases too, the change is small compared to the change in ARL0.

**Analysis at the level of a single sample using a simulation**

In the previous section we presented general comparisons we conducted. In the current section we present comparisons at the level of a single sample. The goal of the simulation is to compare the percent of samples in which the results of the monitoring process are identical for unrounded and rounded data, versus the percent of samples for which the results of the monitoring process for unrounded data would be the opposite of those for rounded data. In other words, our aim is to address cases where the sample mean is within the control limits when the data is unrounded, but falls outside the control limits when the data is rounded.

|  |  |  |
| --- | --- | --- |
|  | Unrounded – outside limits | Unrounded – inside limits |
| Rounded – outside limits | oo | io |
| Rounded – inside limits | oi | ii |

Table 4. Simulation states for estimating control chart performances at the level of a single sample.

States oo and ii represent cases where the round-off error has no impact (i.e. an identical result). State io represents cases where when the data is unrounded the sample mean is inside the control limits (i) and when the data is rounded the sample mean is outside the control limits (o). Under H0 this situation indicates an increase in alpha (diminished performances of the control chart) when the data is rounded. Under H1, this situation improves the performances of the control chart when the data is rounded, meaning it reduces beta.

The oi state represents cases where the sample mean is outside the control limits when the data is unrounded and inside the control limits when the data is rounded. Under H0, this situation indicates a decrease in alpha (i.e. improved performances of the control chart) when the data is rounded. Under H1, this situation indicates diminished performances of the control chart when the data is rounded, meaning it increases beta.

The value of alpha when the data is **unrounded** is the sum of both states in which the sample mean is outside the control limits when the data is unrounded, despite the process being under statistical control (the sum of the values in the left column in Table 4).

α = oo + oi

The value of alpha when the data is **rounded** is the sum of the two states in which the sample mean is outside the control limits when the data is rounded, despite the process being under statistical control (the sum of the values in the first row of Table 4).

αr = oo + io

The value of beta when the data are **unrounded** is the sum of the two states in which the sample mean is inside the control limits when the data is unrounded, despite the fact that the process is no longer under statistical control (the sum of the values in the right column in Table 4).

β = ii + io

The value of beta when the data are **rounded** is the sum of the two states in which the sample mean is inside the control limits when the data is rounded, despite the fact that the process is no longer under statistical control (the sum of the values in the second row of Table 4).

β = ii + oi

|  |  |  |
| --- | --- | --- |
|  | Large io | Large oi |
| H0 | Large alpha  Diminished control chart performances | Small alpha  Improved control chart performances |
| H1 | Small beta  Improved control chart performances | Large beta  Diminished control chart performances |

Table 5. Significance of states io and oi under H0 and H1.

*Simulation stages:*

*Under null hypothesis H0*

1. Drawing m sets (m = 100k) of various sizes of numbers from a normal distribution with a known mean and standard deviation:

.

Seven sample sizes were tested to determine the influence of sample size on chart performances:

n = [7, 10, 15, 20, 25, 30, 40]

1. Calculating the upper and lower control limits (LCL, UCL) on chart

For each of the seven databases (the different n sizes) we built control limits calculated according to three standard deviations:

1. Testing for every sample j out of the m samples whether or not was within the control limits and concentrating the binary results in a table with m rows (each row representing the sample number). If the value 1 is registered, otherwise the value 0 is registered.
2. Rounding the data set from stage a according various δ levels.
3. Estimating the mean for the rounded numbers based on a naïve estimate—averaging the n rounded numbers.

* Calculating new control limits (LCLr, UCLr) for the mean (calculated as the distance of three standard deviations from the estimation of the mean).

A = 3/sqrt(n)

Testing for every sample j out of the m samples whether or not was within the control limits and adding the binary results to the table from stage c (adding columns according to the number of rounding levels). If the value 1 is registered, otherwise the value 0 is registered.

1. Comparing between the results with and without rounding for each data set. We examined the rate of agreement between the two states (oo and ii), with and without rounding, and the rates of incidences of disagreement (states io and oi). The simulation was conducted while the process was under statistical control (under H0) and when it had gone out of statistical control (under H1).

*Under alternative hypothesis H1*. For every sample size eight new sets of m samples were drawn, each from a normal distribution with a mean deviated at orders of magnitude of a standard deviation.

Nine k values were tested, placed as positive and negative.

k = [0, 0.25, 0.3, 0.4, 0.5, 0.75, 0.9, 1, 1.25, 1.5]

We then repeated stages b–f for every data set.

Below is an analysis for rounding level δ= 0.355 and sample size n = 7

*Rounding level 0.355, sample size 7*



Table 6. Distribution of the states under H0 and H1 and calculation of alpha and beta values; rounding level 0.355, sample size 7.

The table above shows that the value of alpha for the rounded data is significantly greater than its value for the unrounded data (22 times greater), whereas the value of beta for the rounded data decreases less significantly compared to its value when the data is unrounded (a decrease of 20–50 percent).

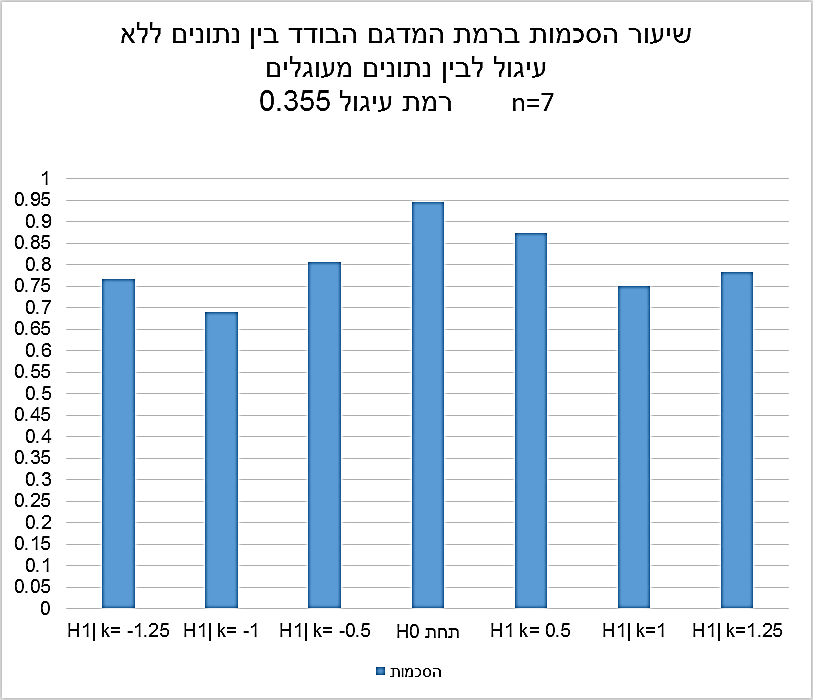


Chart 7. The rate of agreement between unrounded and rounded data; rounding level 0.355, sample size 7.

For a small sample and a rounding level of 0.355 it is not possible to determine the trend of the rate of agreement in relation to the level of deviation from the mean (in absolute values). Under H0 the rate of agreement is high (95%), whereas under H1 it goes down and then comes up, shifting between 70%–90%.

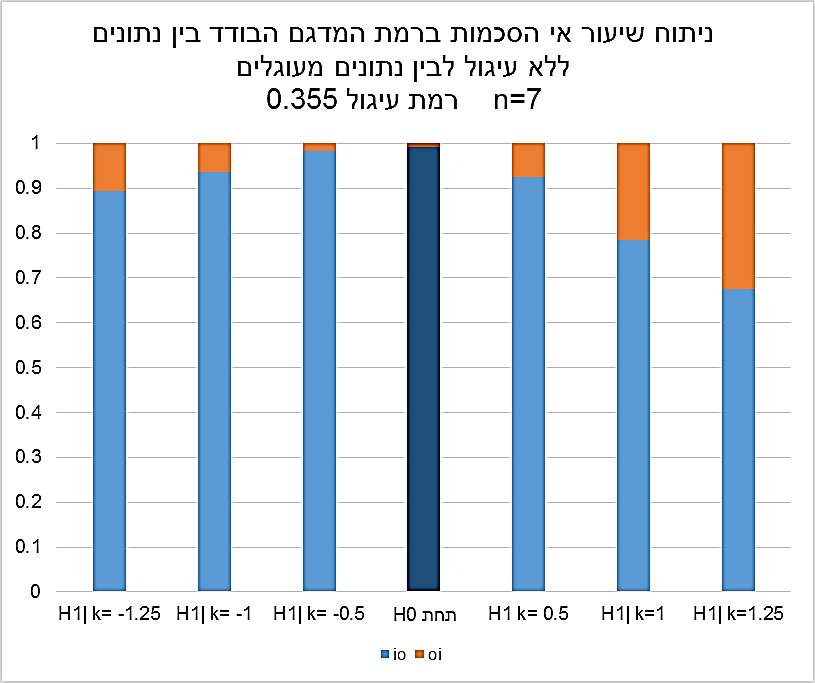


Chart 8. The rate of disagreement between unrounded and rounded data; rounding level 0.355, sample size 7.

The chart above demonstrates that for a small sample and a rounding level of 0.355 under H0 the performances of the charts are diminished. In almost 100% of the disagreements regarding the sample’s position in relation to the limits, the samples were found to be outside the control limits when the data were rounded. In other words, the value of alpha increased when the data was rounded (state io).

When there was a change in the mean (under H1), we found that in most cases where the results with and without rounding were not in agreement, the sample was outside the control limits when the data was rounded, meaning the value of beta decreased (the io state) and the chart performances improved.

It appears that when the deviation from the mean is positive there are less states in which the samples are found outside the control limits when the data is rounded. However here too, in most cases it is possible to see that the value of beta is smaller.

It is possible to determine that for a small sample size and a rounding level of 0.355 the performances of the control chart are diminished for alpha but improved for beta.

**Summary of the findings based on the simulation at the level of a single sample**

* The agreement rate regarding the sample’s position in relation to the control limits when comparing unrounded and rounded data is very high, especially under H0. However, there is still a percent of disagreements that affects the performances of the control chart.
* For the cases that were tested we found that when the size of the sample was small the rate of agreement decreased the more the deviation from the mean increased, or that there was a decrease followed by an increase (in most cases a relatively small increase) in the rate of agreements when the deviation increased. When the sample size was large we found that the rate of agreement increased the more the deviation from the mean increased.
* In cases of disagreement regarding the sample’s position in relation to the control limits, and no change in the mean (under H0), in all the incidents we tested the vast majority of samples were found to be outside the control limits when the data was rounded. In other words, the value of alpha had increased significantly and the performances of the control chart had been diminished.
* In cases of disagreement regarding the sample’s positon in relation to the control limits, it was not possible to decisively determine the state of beta when the data was rounded. We saw cases were most of the samples were within the control limits (the value of beta increased, the performances of the control chart decreased) and cases were most of the samples were outside the control limits (the value of beta decreased, the performances of the control chart improved). We also saw cases of near equilibrium. In all cases the change for the better or for the worse in the value of beta was relatively small compared to the change in the value of alpha.
* We found that for small deviations there was generally an improvement in the performances of the chart. In other words, the deviation is discovered sooner when the data is rounded, whereas for medium and large deviations the performances of the charts were diminished in most cases. The quality controller expects to receive a signal that the process has gone out of statistical control when there is a deviation from the process. When the deviations are large he expects to get the signal sooner compared to when they are small. When the data is rounded, the opposite process takes place: it takes longer to identify that the system has gone out of control when the deviations are large compared to when they are small.

The findings of the simulation at the level of the sample are in alignment with the findings from the analysis of the control charts and the alpha and beta indexes.

**Summary and conclusions, including follow-up research**

For the alpha and ARL0 indexes there is a significant decrease in the performances of the control chart when the data is rounded and the control limits are calculated using Shewhart’s tools. The direct implication is that false alarms occur more frequently, disrupting the work routine on the production line. Additional damage that may occur due to the large amount of false alarms is that it could lead to quality controllers on the production line ignoring signals they receive during the control process.

In regards to the beta and ARL1 indexes, there is no conclusive effect. Some of the rounding levels diminished the performance of the control charts, meaning it would have taken less samples to identify loss of statistical control had the data been unrounded. However other rounding levels actually improved the performances of the control chart, requiring less samples to identify the deviation compared to when the data was unrounded. Based on the analyses we conducted, we found that when the deviation from the mean was small there was an improvement, meaning the deviation was discovered sooner. Surprisingly, when the deviation was medium-sized or large, precisely when we would have expected to be informed of losing statistical control relatively quickly, the performances of the control charts were diminished, taking more samples to recognize the deviation.

Our primary finding is that when the data is rounded and the control limits are calculated according to Shewhart’s classic theory, performances of the control charts under H0 are significantly diminished for alpha and ARL0, in relation to their values when the data is unrounded. Under H1, the performances of the charts are diminished or improved based on the size of the deviation from the mean. However, in all cases the change in beta and ARL! values is relatively small compared to the change under H0.

Another important finding is that there was no conclusive direction of the change in beta values when the data was rounded. The analysis at the level of the single sample showed inconclusive results for all the samples. Where beta increased, we found substantial rates of samples behaving in the opposite manner. This finding indicates that standard control charts are inappropriate for crudely rounded data.

Bibliography

Abraham, B. 1977. “Control Charts and Measurement Error.” *Annual Technical Conference of the American Society for Quality Control* 31, 370–374.

Bennett, C A. 1954. “Effect of Measurement Error on Chemical Process Control.” *Industrial Quality Control* 10, 17–20.

Benson, D, E Dvir-Harcabi, I Regev, and E Schechtman. 2013. “Estimation of a Normal Process Variance from Measurements with Large Round-Off Errors.” *IET Science, Measurement and Technology* 7. no. 3, 180–189.

Benson, D and E Schechtman. 2015. “Using Measurements with Large Round-Off Errors Interval Estimation of Normal Process Variance.” *IET Science, Measurement and Technology*, 1–7.

Gertsbakh, I. 2003. *Measurement Theory for Engineers*. Berlin: Springer Verlag.

Kanazuka, T. 1986. “The Effects of Measurement Error on the Power of X̄ - R Charts.” *Journal of Quality Technology* 18, 91–95.

Linna, K, and W Woodall. 2001. “Effect of Measurement Error on Shewhart Control Charts.” *Journal of Quality Technology*, 213­–222.

Montgomery, D C. 2013. *Statistical Quality Control: A Modern Introduction.* 7th ed. International Student Version. NewYork: Wiley.

Vardeman, S B. 2005. “Sheppard’s Correction for Variances and the Quantization Noise Model.” *IEEE Transactions on Instrumentation and Measurement* 54, 2117–2119.

Walden, C T. 1990. “An Analysis of Variables Control Charts in the Presence of Measurement Error.” Unpublished master's thesis, Department of Industrial Engineering, Mississippi State University, UMI no. 1343291.

Wheeler, D. 2011. “100% Inspection and Measurement Error.” Quality Digest, May 3, 2011.

<https://www.qualitydigest.com/inside/quality-insider-column/100-inspection-and-measurement-error.html>.

Zhidong B, Z Shurong Z, Z Baoxue, and H Guorong. 2009. “Statistical Analysis for Rounded Data.” *Journal of Statistical Planning and Inference*139, 2526–2542.