

Graphical modeling for item difficulty in medical faculty exams

L Tomak, Y Bek, MA Cengiz¹

Department of Biostatistics and Medical Informatics, Faculty of Medicine, Ondokuz Mayıs University, ¹Department of Statistics, Faculty of Arts-Science, Ondokuz Mayıs University, Atakum, Samsun, Turkey

Abstract

Background: There are different indexes used in the evaluation of exam results. One important index is the difficulty level of the item that is also used in this study to obtain control charts. This article offers some suggestions for the improvement of multiple-choice tests using item analysis statistics.

Materials and Methods: The graphical modeling is important for the rapid and comparative evaluation of test results. The control chart is a tool that can be used to sharpen our teaching and testing skills by inspecting the weaknesses of measurements and producing reliable items. The research data for the application of control charts were obtained using the results of the fourth and fifth-grade student's exams at Ondokuz Mayıs University, Faculty of Medicine. I-chart or moving range chart (MR) is preferred for whole variable data.

Results: It is seen that all observations are within control limits for I-chart, but three points on MR-chart are settled on the LCL. Using \bar{X} -chart with subgroups, it was determined that control measurements were within the upper and lower limits in both charts. The difficulty levels of items were examined by obtaining different variable control charts. The difficulty level of the two items exceeded the upper control limit in R- and S-charts.

Conclusion: The control charts have the advantage for classifying items as acceptable or unacceptable based on item difficulty criteria.

Key words: Item difficulty, quality control, statistical process control, variable control charts

Date of Acceptance: 26-Jun-2015

Introduction

Statistical process control (SPC) is an important technique to assess a production process through control charts. Control charts are developed to monitor SPC and to determine variations in the production process.^[1]

Shewhart developed the first type of chart in 1920. Initially, control charts were used in order to solve industrial problems.^[2] After Shewhart, Levey and Jennings used control chart and SPC in clinical laboratory in 1950. The chart of Levey Jennings is widely preferred in clinical laboratory nowadays. This chart uses individual values, and it is plotted according to the reference value.^[3,4]

Control charts are classified according to whether they monitor attribute (discrete) data or variable (continuous) data. Attribute characteristics are used as binary data that are “conforming-not conforming,” “good-bad” or “defective-nondefective” in quality control.^[5,6]

Some control charts monitor variability of characteristics. The characteristics are measured as variable data for variable control charts.^[7] Control charts for variables include charts for individual measurements; charts for variability or dispersion of several measures, such as the range and the standard deviation, and charts for central tendency of several measures, such as averages, medians, and midranges.

Address for correspondence:

Dr. L Tomak,
Department of Biostatistics and Medical Informatics,
Faculty of Medicine, Ondokuz Mayıs University,
Atakum, Samsun, Turkey.
E-mail: lemantomak55@gmail.com

Access this article online

Quick Response Code:



Website: www.njcponline.com

DOI: 10.4103/1119-3077.173701

PMID: 26755220

Many applications include the companion use of a chart for variability and a chart for central tendency.^[8]

The principles of creating both variable control charts and attribute control charts are same. The aim is to determine the mean, the standard deviation, and the distance between the mean and the control limits based on the standard deviation. Control charts for variable characteristics are often used, and they are very effective in providing feedback about the performance of the process. These graphs can be classified in different ways for the individual measurement, for subgroups and for different combinations of subgroups.^[2]

The value of each measurement in control charts for individual observations is shown by a single point on the chart. There are three different charts for individual observations. They are individual observations (I) chart, moving range (MR) chart, and I-MR chart that are obtained using both of them.^[9]

Other control charts for variable characteristics are preferred if there are existing subgroups. These charts are respectively average (\bar{X}) chart, range (R) chart, standard deviation (S) chart, \bar{X} -R chart that consists of using a combination of \bar{X} and R and \bar{X} -S Chart that consists of using a combination of \bar{X} and S. In addition, there is I-MR-R/S-chart that is obtained with the combination of I, MR and R, or S.^[5]

The exam results were evaluated using the following control charts in this study. They are individual observations (I) chart, MR chart, and I-MR chart, average (\bar{X}) chart, R-chart, standard deviation (S) chart.

The production process to improve individual test items and to increase the quality of the test as a whole is evaluated through these control charts. When statistical process is an “in control” process, the test results are explained as a “good” item, but if statistical process is not an “in control” process, the test item identified need to be revised or thrown out; we may want to find out why that particular item had problems producing reliable measurements.

The purpose of this study is to offer some suggestions for the improvement of multiple-choice tests using item analysis statistics and control charts that are mainly used to control a manufacturing process. The control charts will allow us to identify outliers among parts of items. Inspecting that particular part may give us some insights into the weaknesses in our measurements. As a result, the statistical processes of these examinations were evaluated by plotting different variable control charts.

Materials and Methods

The data for the application of control charts were obtained using the results of the fourth and fifth-grade student's exams at Ondokuz Mayıs University, Faculty of Medicine.

There are different indexes used in the evaluation of exams. The difficulty level was used to obtain control charts for this study. Item difficulty is a measure of the percentage of students answering a question correctly. Values for the difficulty index range from 0% (very difficult) to 100% (very easy). The difficulty index can be used to alert the instructor about potential problems. The level of difficulty is considered as between 0.2 and 0.8 for the acceptability of a problem.^[10]

We chose all of fourth-grade blocks which consist of 12 examination results; we also chose randomly one block among all blocks of fifth-grade class which consists of 20 questions results. Sensory block was recruited for the whole fifth class. The difficulty levels of 20 questions were evaluated for five different training groups. Because the number of training groups is five, the number of subgroups is accepted as 5.

We used control charts for individual observations of fourth-graders examination results of whole class, we illustrated the individual observation graphs for endocrine, gastroenterology, pediatrics, and pulmonary circulation blocks on the same graphics that belong to fourth grade class. We also used all of \bar{X} -chart, R-chart and S-chart for examination of sensory block of fifth grade class to see control charts of subgroups.

The analysis of examination was evaluated by control charts. MINITAB Statistical Software, Release 15 for Windows was used to get variable control charts.^[11]

Control charts for individual observations

Individual observation chart

The individual observation chart is the most basic variable control chart used for individual measurements of a variable type characteristic.^[7] For this chart, the measurable characteristics are selected. The frequencies of obtained measurements are determined.

The control limits of this chart are calculated according to the formula1, where “k” indicates the coefficient of the desired standard deviation. The arithmetic mean (\bar{X}) is generally used to estimate the mean of the process characteristic. The equality for k = 3 is as follows:^[8,12]

$$\begin{aligned} UCL &= \hat{\mu} + k\hat{\sigma} \Rightarrow \hat{\mu} + 3\hat{\sigma} \\ LCL &= \hat{\mu} - k\hat{\sigma} \Rightarrow \hat{\mu} - 3\hat{\sigma} \end{aligned} \quad (1)$$

The MR is used to estimate the standard deviation (σ). The MR is defined as the absolute value of the difference between the largest and smallest observations in consecutive subsets of “n” observations. So the MRs and the average MR are computed as:^[8,6]

$$MR = |X_i - X_{i-1}| \quad (2)$$

$$\overline{MR} = \frac{\sum_{i=2}^n |X_i - X_{i-1}|}{n-1} \tag{3}$$

When MR is used to estimate the standard deviation, derivation of control limits for the individual observations chart makes use of the relationship between standard deviation and ranges. The constant of proportionality is referred to as d_2 .^[6] The standard deviation is obtained as \overline{MR} / d_2 . Control limits are derived as follows:^[9]

$$\begin{aligned} UCL &= \bar{X} + \frac{3}{d_2} \times \overline{MR} \\ CL &= \bar{X} \\ LCL &= \bar{X} - \frac{3}{d_2} \times \overline{MR} \end{aligned} \tag{4}$$

Where d_2 , as a function of the subset size “n,” is tabulated in special table (Montgomery, 1996). For $n = 2$, d_2 is usually equal 1.128, and the coefficient is $3/1.128 = 2.66$. So control limits are simplified as follows:^[11,12]

$$\begin{aligned} UCL &= \bar{X} + 2.66 \times \overline{MR} \\ CL &= \bar{X} \\ LCL &= \bar{X} - 2.66 \times \overline{MR} \end{aligned} \tag{5}$$

If both μ and σ are known, they can be used directly. If only σ is known, \bar{X} is calculated but σ can be used directly. Control limits are determined accordingly:^[6,9]

$$\begin{aligned} UCL &= \mu + 3 \times \sigma & UCL &= \bar{X} + 3 \times \sigma \\ CL &= \mu & CL &= \bar{X} \\ LCL &= \mu - 3 \times \sigma & LCL &= \bar{X} - 3 \times \sigma \end{aligned} \tag{6}$$

Moving range chart

The MR is also used to estimate the standard deviation (σ) for MR-chart.^[11,5]

Once all of MR values are obtained, the average of MR shows center line. Control limits are as follows:^[8]

$$\begin{aligned} UCL &= d_2 \times \hat{\sigma} + 3 \times d_3 \times \hat{\sigma} = d_2 \times \frac{\overline{MR}}{d_2} + 3 \times d_3 \times \frac{\overline{MR}}{d_2} \\ &= \overline{MR} + \frac{3 \times d_3}{d_2} \times \overline{MR} = \overline{MR} \times \left(1 + \frac{3 \times d_3}{d_2} \right) \\ CL &= d_2 \times \hat{\sigma} = \overline{MR} \\ LCL &= d_2 \times \hat{\sigma} - 3d_3 \times \hat{\sigma} \Rightarrow d_2 \times \frac{\overline{MR}}{d_2} - 3d_3 \times \frac{\overline{MR}}{d_2} \\ &= \overline{MR} - \frac{3 \times d_3}{d_2} \times \overline{MR} = \overline{MR} \times \left(1 - \frac{3 \times d_3}{d_2} \right) \end{aligned} \tag{7}$$

The following equations (D_3, D_4) are used to simplify the formulas of the control limits:^[11,13]

$$D_3 = \left(1 - 3 \frac{d_3}{d_2} \right) \text{ veyra } D_4 = \left(1 + 3 \frac{d_3}{d_2} \right) \tag{8}$$

Where both D_3 and D_4 , as a function of the subset size for “n,” are tabulated in specific table. So control limits are shown as follows:^[6,9]

$$\begin{aligned} UCL &= D_4 \times \overline{MR} \\ CL &= \overline{MR} \\ LCL &= D_3 \times \overline{MR} \end{aligned} \tag{9}$$

I-MR-chart

I-MR-chart is created using both I-chart and MR-chart.^[7]

Control charts for subgroups

Average (\bar{X}) chart

General principle for the creation of control limits of \bar{X} -chart is $\mu \pm 3\sigma$ for cases given standard and $\hat{\mu} \pm 3\hat{\sigma}_{\hat{\mu}}$ for cases with no standard or $\bar{X} \pm 3 \times S_{\bar{X}} = \bar{X} \pm 3 \times \frac{S}{\sqrt{n}}$.^[8]

When standard value is unknown, the mean average (\bar{X}) value is calculated. In order to determine \bar{X} , \bar{X} should be calculated for each subgroup (k). \bar{X} creates value of CL:^[12,13]

$$CL = \bar{X} = \frac{\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_k}{k} = \frac{\sum_{i=1}^k \bar{X}_i}{k} \tag{10}$$

The next step is to calculate UCL and LCL:^[8,9]

$$\begin{aligned} UCL &= \bar{X} + 3\hat{\sigma} \\ CL &= \bar{X} \\ LCL &= \bar{X} - 3\hat{\sigma} \end{aligned} \tag{11}$$

To obtain control limits, the value of standard deviation must be determined for the population. This can be determined in several ways. The first option is to use the standard error estimate σ/\sqrt{n} , and another option is to use the mean range.^[1]

Standard error-based \bar{X} -chart

Based on the central limit theorem, to obtain control limits, standard deviation of the process is divided by the square root of the sample’s size:^[5]

$$\begin{aligned} UCL &= \bar{X} + 3 \times \left(\frac{\hat{\sigma}}{\sqrt{n}} \right) \\ CL &= \bar{X} \\ LCL &= \bar{X} - 3 \times \left(\frac{\hat{\sigma}}{\sqrt{n}} \right) \end{aligned} \tag{12}$$

When the value of σ is estimated by S/c_4 for example, control limits are created as follows:^[7]

$$\bar{S} = \frac{(S_1 + S_2 + \dots + S_k)}{k} = \frac{\sum_{i=1}^k S_i}{k} \quad (13)$$

$$\begin{aligned} UCL &= \bar{X} + 3 \times \frac{(\bar{S}/c_4)}{\sqrt{n}} \\ CL &= \bar{X} \end{aligned} \quad (14)$$

$$LCL = \bar{X} - 3 \times \frac{(\bar{S}/c_4)}{\sqrt{n}}$$

Control limits can be simplified by the equation of A_3 :^[6,7]

$$\begin{aligned} UCL &= \bar{X} + A_3 \times \bar{S} \\ CL &= \bar{X} \\ LCL &= \bar{X} - A_3 \times \bar{S} \end{aligned} \quad (15)$$

Mean range-based \bar{X} -chart

For the data to have normal distribution, there is a special relationship between the range and the standard deviation as follows:^[8]

$$\sigma = \frac{R}{d_2} \text{ or } R = d_2 \times \sigma \quad (16)$$

R is called the relative range. The mean range is:^[11]

$$\bar{R} = \frac{R_1 + R_2 + \dots + R_k}{k} = \frac{\sum_{i=1}^k R_i}{k} \quad (17)$$

' k ' is the number of subgroups. The estimator of σ is $\hat{\sigma} = \frac{\bar{R}}{d_2}$.^[6]

The estimator of σ/\sqrt{n} is $\frac{\hat{\sigma}}{\sqrt{n}} = \frac{\bar{R}}{d_2 \sqrt{n}}$.^[9]

Control limits are:^[5]

$$\begin{aligned} UCL &= \bar{X} + \frac{3 \times \bar{R}}{d_2 \times \sqrt{n}} \\ CL &= \bar{X} \end{aligned} \quad (18)$$

$$LCL = \bar{X} - \frac{3 \times \bar{R}}{d_2 \times \sqrt{n}}$$

When the abbreviation of A_2 is used, control limits are:^[7,13]

$$\begin{aligned} UCL &= \bar{X} + A_2 \times \bar{R} \\ CL &= \bar{X} \\ LCL &= \bar{X} - A_2 \times \bar{R} \end{aligned} \quad (19)$$

If both μ_0 vey σ_0 are known, control limits will be as follows:^[5]

$$\begin{aligned} UCL &= \mu_0 + A \times \sigma_0 \\ CL &= \mu_0 \\ LCL &= \mu_0 - A \times \sigma_0 \end{aligned} \quad (20)$$

The value of A is found using a special table that shows the different values of A for different subgroup sizes.^[6]

R-Chart

Primarily, a measurable feature of the process is selected for the creation of R-chart. The range value for each subgroup is calculated. The mean range is detected using the range values for subgroups.^[5,6]

The center line of R-chart is the mean range that is shown by \bar{R} .^[1,12]

$$\bar{R} = \frac{R_1 + R_2 + \dots + R_k}{k} = \frac{\sum_{i=1}^k R_i}{k} \quad (21)$$

The equation of $s_R = d_3 \cdot \sigma$ is used to calculate the standard deviation of the mean range. Because, $\hat{\sigma} = \frac{R}{d_2}$ the equation of $\sigma_R = d_3 \cdot \sigma$ can be written as follows:^[11]

$$\sigma_R = \frac{d_3 \times \bar{R}}{d_2} \quad (22)$$

Both of the equations D_3 and D_4 are used to obtain control limits:^[5,13]

$$D_3 = \left(1 - 3 \times \frac{d_3}{d_2}\right) \text{ vey } D_4 = \left(1 + 3 \times \frac{d_3}{d_2}\right) \quad (23)$$

When both of the equations D_3 and D_4 are used, the control limits will be as follows:^[7]

$$\begin{aligned} UCL &= D_4 \times \bar{R} \\ CL &= \bar{R} \\ LCL &= D_3 \times \bar{R} \end{aligned} \quad (24)$$

If the value of σ is known, the control limits are as follows:^[5]

$$\begin{aligned} UCL &= D_2 \times \sigma_0 \\ CL &= R_0 \\ LCL &= D_1 \times \sigma_0 \end{aligned} \quad (25)$$

Both of the equations D_1 and D_2 are used according to the size of subgroups.^[6]

S-chart

A measurable feature of the process is used for S-chart, too. The value of standard deviation of every group is detected.^[8]

When neither μ or σ is known, standard deviation (S) is calculated using variance (S^2) of sample:^[2]

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \tag{26}$$

$$\bar{S} = \frac{1}{k} \sum_{i=1}^k S_i \tag{27}$$

\bar{S} is the average standard deviation for “k” groups. When S is used the estimate of σ , the biased results appear. To prevent this situation, it is recommended to use $c_4 \cdot \sigma$. Here, c_4 is a constant value associated with sample size. If the equation of $\bar{S} = c_4 \times \sigma$ is considered, σ will be equal to \bar{S}/c_4 . The expected value of S [E(S)] will be equal to $c_4 \cdot \sigma$, which is center line. The standard deviation of S is $\sigma \times \sqrt{1 - c_4^2}$.^[5,7]

When $c_4 \times \bar{S}$ is used as the estimate of σ , the control limits are:^[2]

$$\left. \begin{aligned} UCL &= \bar{S} + 3 \times \frac{\bar{S}}{c_4} \times \sqrt{1 - c_4^2} = \bar{S} \times \left(1 + \frac{3}{c_4} \times \sqrt{1 - c_4^2} \right) \\ CL &= \bar{S} \\ LCL &= \bar{S} - 3 \times \frac{\bar{S}}{c_4} \times \sqrt{1 - c_4^2} = \bar{S} \times \left(1 - \frac{3}{c_4} \times \sqrt{1 - c_4^2} \right) \end{aligned} \right\} \tag{28}$$

The equations of both B_3 and B_4 are given below:^[7]

$$\left. \begin{aligned} B_3 &= 1 - \frac{3}{c_4} \times \sqrt{1 - c_4^2} \\ B_4 &= 1 + \frac{3}{c_4} \times \sqrt{1 - c_4^2} \end{aligned} \right\} \tag{29}$$

The control limits can be shown to use both B_3 and B_4 :^[2]

$$\left. \begin{aligned} UCL &= B_4 \times \bar{S} \\ CL &= \bar{S} \\ LCL &= B_3 \times \bar{S} \end{aligned} \right\} \tag{30}$$

If both μ and σ are known, the control limits are:^[6]

$$\left. \begin{aligned} UCL &= c_4 \sigma + 3 \times \sigma \times \sqrt{1 - c_4^2} = \sigma \times (c_4 + 3 \times \sqrt{1 - c_4^2}) \\ CL &= c_4 \sigma \\ LCL &= c_4 \sigma - 3 \times \sigma \times \sqrt{1 - c_4^2} = \sigma \times (c_4 - 3 \times \sqrt{1 - c_4^2}) \end{aligned} \right\} \tag{31}$$

The equations of both B_5 and B_6 are given as follows:^[5]

$$\left. \begin{aligned} B_6 &= c_4 + 3 \times \sqrt{1 - c_4^2} \\ B_5 &= c_4 - 3 \times \sqrt{1 - c_4^2} \end{aligned} \right\} \tag{32}$$

When these equations are used, control limits will be as follows:^[7]

$$\left. \begin{aligned} UCL &= \alpha_0 + 3 \times \sigma_0 = c_4 \times \sigma_0 + 3 \times \sqrt{1 - c_4^2} \times \sigma_0 = B_6 \times \sigma_0 \\ CL &= c_4 \times \sigma_0 \\ LCL &= \alpha_0 - 3 \times \sigma_0 = c_4 \times \sigma_0 - 3 \times \sqrt{1 - c_4^2} \times \sigma_0 = B_5 \times \sigma_0 \end{aligned} \right\} \tag{33}$$

Control charts for different combinations of subgroups
 \bar{X} -R-chart consists of a combination of \bar{X} and R. \bar{X} -S-chart consists of a combination of \bar{X} and S. I-MR-R/S-chart is the combination of I, MR and R or S.^[1,2]

Results

The variable control charts were created to evaluate item difficulty level derived from exam analysis.

The control charts for individual observations are as follows: I-chart is given in Figure 1 and MR-chart is given in Figure 2. There were four separate item groups and exam blocks. There are 12 points on these charts. The first four points refer to difficulty level of endocrine block items for every group (four groups). The second four points refer to difficulty level of gastrology block items for every group (four groups). The third-four points refer to difficulty level of pediatrics block items for every group (four groups). The last four points refer to difficulty level of pulmonary circulation block items for each group (four groups).

All of the data are within control limits on I-chart [Figure 1].

All the observations are within the control limits on MR-chart. However, the points 8, 12, and 13 are settled on the LCL [Figure 2].

Standard error-based \bar{X} -chart is given in Figure 3. All of the observations are within control limits and the variations

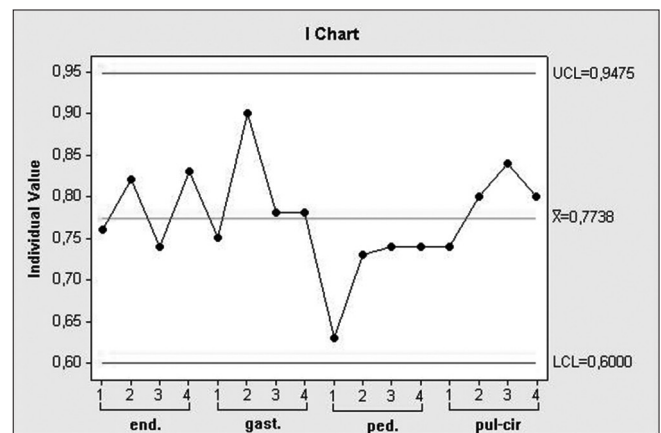


Figure 1: I-Chart of item difficulty level

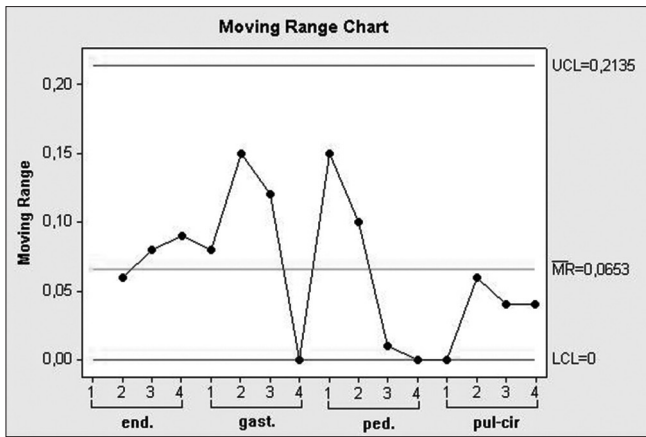


Figure 2: Moving range-chart of item difficulty level

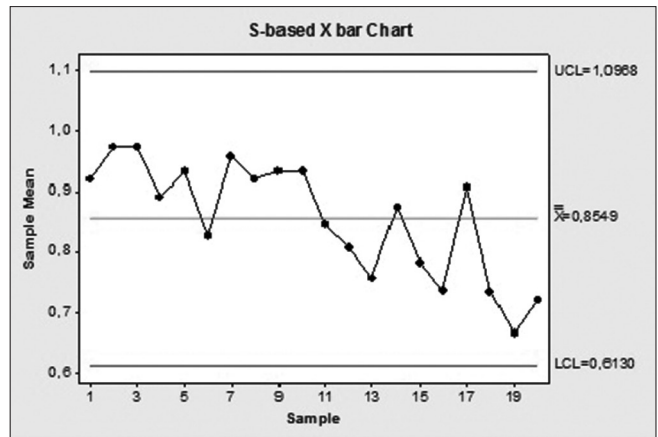


Figure 3: Standard error-based \bar{X} -chart of item difficulty level

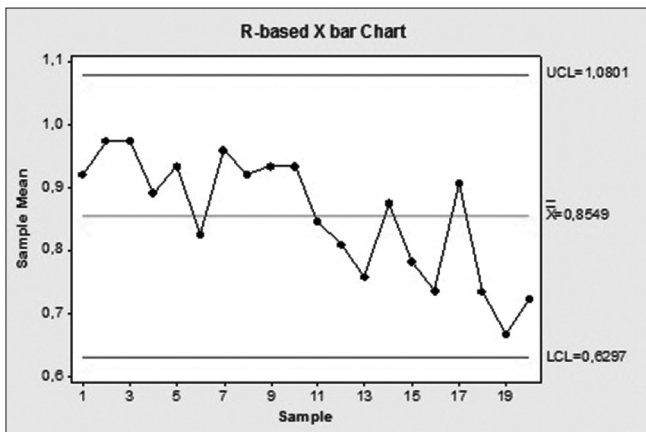


Figure 4: Mean range-based \bar{X} -chart of item difficulty level

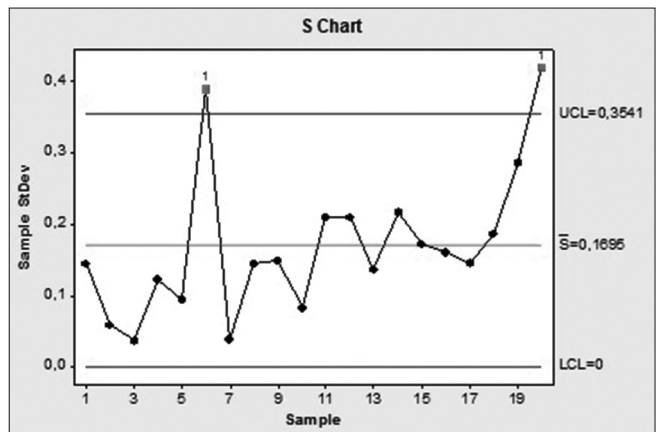


Figure 5: S-Chart of item difficulty level

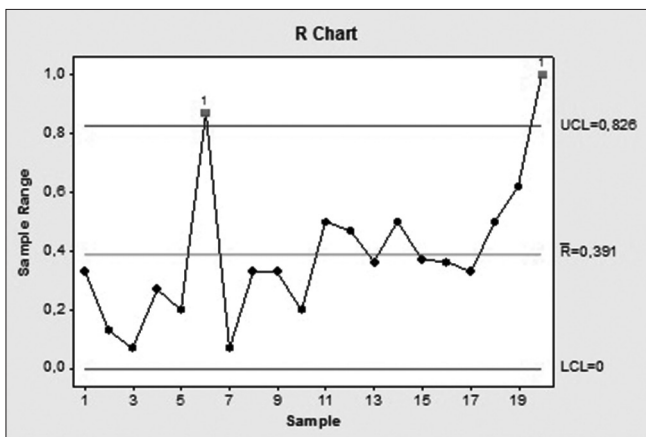


Figure 6: R-Chart of item difficulty level

exhibit a random pattern, so the process is stable and under control.

Mean range-based \bar{X} -chart is shown in Figure 4. Because all of the data are settled within control limits, we can conclude that the process is under control.

S-chart is given in Figure 5. As two points, 6 and 20 are outside UCL, statistical process is out of control.

R-chart is shown in Figure 6. Two points, 6 and 20, are outside UCL on this chart, so the process is unstable.

Discussion

Six different variable control charts are used for continuous data, and the quality of item analysis process is evaluated. Here, difficulty levels of items were examined by obtaining different variable control charts.

The graphical modeling is important for the rapid and comparative evaluation of test results. The control chart is just one more tool we can use to sharpen our teaching and testing skills. The control charts will allow us to identify outliers among parts of items. Inspecting that particular unstable part may give us some insights into the weaknesses in our measurements.

The control charts have the advantage for quick summaries of various aspects of the quality of the items of exams.

Simply classify items as acceptable or unacceptable based on item difficulty criteria. Also, this type of charts tend to be more easily understood by lecturer or other related people unfamiliar with quality control procedure or other statistical quality control index expectations.

The item difficulty data has normal distribution. So it is assumed that 68% of the values fall in the interval $\pm 1.S$; 95% of the values fall in the interval $\pm 2.S$; 99.7% of the values fall in the interval $\pm 3.S$.^[14,15] If a value is distant $\pm 3.S$ from the average value, this is an unexpected observation (0.3% randomly).^[16] So this usually indicates a problem with the process. A measurement value that is distant $\pm 2.S$ from the average value may occur at least 5% randomly.^[17-19]

When there are not any subgroups, I-chart or MR-chart is preferred for variable data. Both I-chart and MR-chart were created to analyze item difficulty of fourth-grade exam. When these are evaluated, it is seen that all observations are within control limits, but three points on MR-chart are settled on the LCL. However, the processes are assumed to be in control for both charts.

In I-chart, the mean difficulty index was found to be 0.7738. This indicates that mean difficulty level was high, and it is a proof that the questions were easy. However, it was determined that values of all the blocks were within the desired limits. While the distribution for most of the observations was close to the mean difficulty level, it was determined that the questions of gastrology items block of the second group were quite easy (0.90) and the questions of pediatric items block of the first group were more difficult (0.63).

MR-chart is a chart that demonstrates the variability. Although it is within the control limits, there is a variation on this chart originating from the first items group. This difference is specifically remarkable in pediatric and pulmonary circulation training groups (internships).

I-chart is the simplest chart among all of control charts. This chart is preferred to get a result in a short time. It is particularly preferred when using \bar{X} , R and S control charts is not practical. It is much more sensitive in detecting significant variations (variances) in cases where the sample size to be determined by the binormal distribution is very large.^[7,20] MR is generally used to evaluate the mean variability of the difficulty of item. If we do not take the subgroups into consideration, one or both of these charts can be preferred.^[2-21]

Other variable charts used in the presence of subgroups were obtained in this study. First of them is \bar{X} -chart. In cases with unknown standard values, for the estimation of σ , this chart was generated in two different ways, using standard

error and mean range, and it was determined that control measurements were within the upper and lower limits in both charts. So we can say that the process is "in-control." The mean value was detected to be 0.86 in both charts. This indicates that the questions were easy.

\bar{X} -chart is the most commonly used control chart for the central tendency since it is highly sensitive in detecting the variations related to the distribution of the mean.^[12,22] Standard Error-based \bar{X} -chart is not very practical. Instead of standard error, using the mean range as the measurement of variation, an alternative mean range-based \bar{X} -chart can be created more easily. When sample sizes are relatively small ($n \leq 10$), the mean range-based \bar{X} -chart can replace standard error-based \bar{X} -chart. Since the number of the subgroups in this study is five, it would be better to use the mean range-based \bar{X} -chart.^[8,23]

In the charts obtained in this study, it was found out that difficulty level values of the two questions exceeded the upper control limit in R- and S-charts. Variability in the process was outside the determined limit at two points. R-chart is one of the most widely used control charts in variation measurement. The chart is popular because it is easy to calculate the mean range. Its effectiveness in estimating the variation of the process is better when the number of subgroups is smaller. S-chart is used to determine if the level of variation in the process is significant or not. Thus, control limits of the standard deviations related to the samples taken at regular intervals are created in this chart. When there is a strong variation within the established limits, the process is considered to be unstable.^[1,2,24] Since the number of subgroups in this study were five, R-chart was preferred to S-chart.

Conclusion

Control charts are effective instruments in detecting the specific causes of the variability of the difficulty of items in exams. If some points are outside the control limits, it indicates that the item is out of control; it could be either easy or difficult. The balance of the construction of items of exam has been deteriorated, and corrective operations should be performed. As soon as deterioration occurs in the process, the factor leaving the present observation outside the limits (the factor pushing the present observation outside the limits) should be detected. It does help identify questions that need to be revised or thrown out. The tests can be improved by maintaining and developing a pool of "good" items from which future tests will be drawn in part or whole.

References

1. Bass I. Six Sigma Statistics with Excel and Minitab. USA: McGraw-Hill Company; 2007.

2. Winkel P, Zhang NF. *Statistical Development of Quality in Medicine*. England: John Wiley and Sons Press; 2007.
3. Westgard JO. *Basic QC Practices*. 2nd ed. Madison: Westgard QC Inc.; 2002.
4. Westgard JO, Groth T. Design and evaluation of statistical control procedures: Applications of a computer "quality control simulator" program. *Clin Chem* 1981;27:1536-45.
5. Montgomery DC. *Introduction to the Statistical Quality Control*. 3rd ed. New York: John Wiley and Sons Inc.; 1996.
6. Oakland JS. *Statistical Process Control*. 5th ed. London: MPG Books Limited; 2003.
7. Chandra MJ. *Statistical Quality Control*. USA: CRC Press; 2001.
8. Wadsworth HM, Stephens K, Godfrey AB. *Modern Methods for Quality Control and Improvement*. New York: John Wiley and Sons Inc.; 1986.
9. Ryan TP. *Statistical Methods for Quality Improvement*. New York: John Wiley and Sons Inc.; 1986.
10. Backhoff E, Larrazolo N, Rosas M. The level of difficulty and discrimination power of the basic knowledge and skills examination (EXHCOBA). *Revista Electrónica de Investigación Educativa* 2000;2:1-16.
11. Minitab Inc. *MINITAB Statistical Software, Release 15 for Windows*. State College, Pennsylvania; 2009.
12. Kume H. *Statistical Methods for Quality Improvement*. 10th ed. Tokyo: 3A Corporation; 1992.
13. Ishikawa K. *Guide to Quality Control*. Tokyo: Asian Productivity Organization; 1986.
14. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem* 1981;27:493-501.
15. Gray JJ, Wreghitt TG, McKee TA, McIntyre P, Roth CE, Smith DJ, *et al.* Internal quality assurance in a clinical virology laboratory. I. Internal quality assessment. *J Clin Pathol* 1995;48:168-73.
16. Westgard JO, Klee GG. Quality management. In: Burtis CA, Ashwood ER, Bruns DE, editors. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 4th ed. New York: Elsevier Press; 2006. p. 485-523.
17. Cembrowski GS, Martindale RA. Quality control and statistics. In: Bishop ML, Fody EP, Schoeff LE, editors. *Clinical Chemistry: Principles, Procedures, Correlations*. 4th ed. Philadelphia: Lippincott Williams and Wilkins; 2004. p. 48-89.
18. Goris N, De Clercq K. Quality assurance/quality control of foot and mouth disease solid phase competition enzyme-linked immunosorbent assay – Part II. Quality control: Comparison of two charting methods to monitor assay performance. *Rev Sci Tech* 2005;24:1005-16.
19. Taghizadegan S. *Essentials of Lean Six Sigma*. USA: Elsevier Press; 2006.
20. Lenz HC, Wilrich TH. *Frontiers in Statistical Quality Control*. Germany: German Copyright Law; 2006.
21. Koch DD, Oryall JJ, Quam EF, Feldbruegge DH, Dowd DE, Barry PL, *et al.* Selection of medically useful quality-control procedures for individual tests done in a multitest analytical system. *Clin Chem* 1990;36:230-3.
22. Thompson JR, Koronacki J. *Statistical Process Control*. 2nd ed. USA: Chapman and Hall Press; 2002.
23. Juran JM, Godfrey AB. *Juran's Quality Handbook*. 5th ed. New York, USA: McGraw-Hill; 1999.
24. Revere L, Black K. Integrating six sigma with total quality management: A case example for measuring medication errors. *J Healthc Manag* 2003;48:377-91.

How to cite this article: Tomak L, Bek Y, Cengiz MA. Graphical modeling for item difficulty in medical faculty exams. *Niger J Clin Pract* 2016;19:58-65.

Source of Support: Nil, **Conflict of Interest:** None declared.