Estimation of mean and standard deviation using the stepwise robust Shewhart approaches

L. P. Nadeeka D. Premarathna

2017-04-03

Introduction

The stepwise robust chart procedure for Shewhart $\bar{X}$ chart by Nazir et al. (2014a) used both $\overline{IQR}_{10}$ and $\overline{TM}_{10}$. These two procedures were designed for equal subgroup sizes but unequal subgroup sizes in financial data are common due to market closures, public holidays etc. Hence the two robust estimators found in Nazir et al. (2014b) and Nazir et al. (2014a) are adopted for unequal subgroups using the appropriate unbiasing constants. These two procedures are used to estimate the unknown parameters μ and σ from limited Phase~I data which are required to find the Phase~II control limits.

The **QCCTS** package is developed to implement these two procedures.

Estimation of σ for Phase~II

The interquartile range of subgroup i is calculated as $IQR_i = Q_{i,3} - Q_{i,1}$, where $Q_{i,1} = x_{i,(a)}$ and $Q_{i,3} = x_{i,(b)}$, with $a = \lceil n_i/4 \rceil$, $b = n - a + 1$, and $x_{i,(m)}$ the m^{th}–order statistics of subgroup i. Each IQR_i is corrected by dividing its normalizing constant, $x_{\overline{IQR}_{10}}(n_i)$. These values are given in Table~1 of Nazir et al. (2014b). The 10% trimmed mean of the sample interquartile ranges is used as an unbiased estimate for σ,

$$\overline{IQR}_{10} = \frac{1}{k - 2(\lceil k/10 \rceil - 1)} \left[\sum_{m=\lceil k/10 \rceil}^{k-\lceil k/10 \rceil + 1} IQR_{(m)} \right]$$

This procedure involves two main steps namely screening of subgroups affected by parameter shifts and then the removal of individual outliers (signal points) in the remaining unaffected subgroups.

- Divide Phase I data of the differences $x_{ij}, i = 1, 2, ..., k$ and $j = 1, 2, .., n_i$ into k subgroups of size n_i.
- Calculate $\overline{IQR}_{10}$ as given in above. The variable control limits for screening the subgroup shifts are as follows: $\widehat{UCL}_1 = U_1(n_i)\frac{\overline{IQR}_{10}}{x_{\overline{IQR}_{10}}(n_i)}$ and $\widehat{LCL}_1 = L_1(n_i)\frac{\overline{IQR}_{10}}{x_{\overline{IQR}_{10}}(n_i)}$, and U_1, L_1 and $d_{\overline{IQR}_{10}}(n_i)$ are tabulated in Table~1 of Nazir et al. (2014b).
- Exclude all the subgroups whose IQR/d_{IQR} falls outside the control limits stated above.
- The residuals of each of the individual observation in the remaining subgroups are calculated by subtracting the *trimean* of corresponding subgroup. $res_{x_{ij}} = x_{ij} - TM_i$, where, $TM_i = (Q_{i,1} + 2Q_{i,2} + Q_{i,3})/4$ and $Q_{i,2}$ is the median of subgroup i. The residual values of each observation are then screened. The average IQR of remaining subgroups ($\overline{IQR}'$) is used to identify the outliers in residuals. The control limits are varied depending on the subgroup size: $UCL_{ind} = 3\overline{IQR}'/x_{IQR}(n_i), LCL_{ind} = -3\overline{IQR}'/x_{IQR}(n_i)$.
- The original individual values are removed in each subgroup, when the residuals breach the calculated limits in *Step 4*.
- After eliminating individual outliers, a new estimate of the standard deviation is obtained from the mean of standard deviations of remaining subgroups for Phase~II analysis. Each subgroup standard deviation is adjusted by dividing $c_4(n_i)$.

Estimation of μ for Phase~II

- Obtain the unbiased estimate for standard deviation ($\hat{\sigma_x}$) using the above procedure.
- For $\bar{x}$, instead of the traditional $\bar{x_i}$ of each subgroup, the trimeans of each subgroup is calculated; $TM_{(m)}$ denotes the m^{th} ordered value of the subgroup trimeans. The trimean of subgroup i is defined by, $TM_i = (Q_{i,1} + 2Q_{i,2} + Q_{i,3})/4$, where $Q_{i,2}$ is the median and $Q_{i,1} = x_{i,(a)}$ and $Q_{i,3} = x_{i,(a)}$ the first and third quartiles with $x_{i,(m)}$ the m^{th} order statistic in subgroup i and $a = \lceil n_i/4 \rceil, b = n_i - a + 1$.
- The 10% trimmed mean: TM_{10}, of sample trimeans is used to get the control limits for the $\bar{x}$ chart as follows:

$$\overline{TM}_{10} = \frac{1}{k - 2\lceil k/10 \rceil} \left[\sum_{m=\lceil k/10 \rceil + 1}^{k - \lceil k/10 \rceil} TM_{(m)} \right]$$

The upper and lower control limits are given by $\overline{TM}_{10} \pm 3\hat{\sigma_d}/\sqrt{n_i}$.

- Those subgroups whose TM_i fall outside the control limits are removed and the average of the trimeans ($\overline{TM}^{'}$) in the remaining subgroups is used to identify the individual outliers for the remaining unaffected subgroups.
- Remove individual observations falling outside the limits $\overline{TM}^{'} \pm 3\hat{\sigma_d}$.
- The mean of remaining subgroups (with remaining individual observations) is used for computing mean ($\bar{\bar{x}}$) and the associated Phase~II control limits.

Examples

The example given in Nazir et al. (2014b) and Nazir et al. (2014a) has choosen.

References

Nazir, Hafiz Z., Marit Schoonhoven, Muhammad Riaz, and Ronald J. M. M. Does. 2014a. "Quality Quandaries: A Stepwise Approach for Setting up a Robust Shewhart Location Control Chart." *Quality Engineering* 26 (2): 246–52. doi:10.1080/08982112.2013.874562.

———. 2014b. "Quality Quandaries: How to Set up a Robust Shewhart Control Chart for Dispersion?" *Quality Engineering* 26 (1): 130–36. doi:10.1080/08982112.2013.848367.