Open Journal of Applied Sciences
Supplement: 2012 world Congress on Engineering and Technology

On Efficient Monitoring of Process Dispersion using Interquartile Range

Shabbir Ahmad[1], Zhengyan Lin[1], Saddam Akber Abbasi[2], Muhammad Riaz[3,4]
[1] Department of Mathematics, Institute of Statistics, Zhejiang University, 310027, Hangzhou, China
[2]Department of Statistics, University of Auckland, New Zealand
[3] Department of Statistics, Quaid-i-Azam University, Islamabad Pakistan
[4] Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia
Email: [1]shabbirahmad786@yahoo.com

Abstract: The presence of dispersion/variability in any process is understood and its careful monitoring may furnish the performance of any process. The interquartile range (IQR) is one of the dispersion measures based on lower and upper quartiles. For efficient monitoring of process dispersion, we have proposed auxiliary information based Shewhart-type IQR control charts (namely IQR_r and IQR_p charts) based on ratio and product estimators of lower and upper quartiles under bivariate normally distributed process. We have developed the control structures of proposed charts and compared their performances with the usual IQR chart in terms of detection ability of shift in process dispersion. For the said purpose power curves are constructed to demonstrate the performance of the three IQR charts under discussion in this article. We have also provided an illustrative example to justify theory and finally closed with concluding remarks.

Keywords: Auxiliary Information, Bivariate Normal Distribution, Control Carts, Interquartile Range, Lower and Upper Quartiles, Power Curves.

1. Introduction

Statistical Process Control (SPC) is a collection of fundamental tools which are used to monitor process behavior. In the early 1920s Walter A. Shewhart developed control charting as a useful tool of Statistical Process Control (SPC) to monitor process parameters such as location, dispersion etc. The existence of variability is unavoidable in any process and its careful monitoring is necessary to improve the performance of any process. The variability in a process can be classified in two parts namely *natural* and *unnatural*. Natural/normal variation has a consistent pattern while unnatural/unusual variation has an unpredictable behavior over the time. The presence of natural variation in a process ensures that the process is in-control state, otherwise out-of-control. Control charts assist differentiating between natural and unnatural variations and hence declaring the process to be in-control or out-of-control.

To monitor process variability [1] proposed usual range and standard deviation charts (namely R and S charts). The efficiency of R chart is affected with the increment in sample size where as the performance of S chart becomes poor due to existence of outliers in data (cf. [2]). Later on different estimators of interquartile range (IQR) have been used to establish design structures of dispersion charts such as: [3] and [4] have used interquartile range by restricting the position of lower and upper quartiles as integer, which become cause of some uneven patterns in design structure of control chart. Rocke [5] proposed IQR based R_q chart which out performs the R chart for detecting shifts in process dispersion in outlier scenario. To avoid some irregularities of R_q chart, [2] proposed a new method of usual IQR chart based on the definition of [6]. Abbasi & Miller [7] compared the performances of different dispersion charts under normally and non-normally distributed environments and concluded that for small sample size the IQR chart exhibits reasonable performance while the performances of R and S charts are significantly influenced for highly skewed process environments. Much of the work related to dispersion control charts may be seen in the bibliographies of the above authors.

In this article we have proposed IQR control charts namely IQR_r and IQR_p charts to monitor the process dispersion in Shewhart setup. These charts are based on ratio and product estimators of lower and upper quartiles of study variable Y using one auxiliary variable X under bivariate normally distributed process. The rest of the article is organized as: Section II provides the design structure of IQR charts based on different quantile estimators considered here. In Section III the performance of IQR charts are investigated under the assumption of normality. An illustrative example is provided in Section IV to justify our proposal and finally the study is concluded with some recommendation in Section V.

2. Quantile Estimators and IQR Charting Structures

Let the quality characteristic of interest is Y (e.g. inner diameter of shaft) and X be an auxiliary characteristic (e.g.

outside diameter) associated with *Y*. Let $Q_y(\beta)$ & $Q_x(\beta)$ be the β -quantile of Y & X respectively and $f_y(Q_y(\beta))$ & $f_x(Q_x(\beta))$ be the values of density function at $Q_y(\beta)$ & $Q_x(\beta)$ respectively which can also be obtained by the kernel method or the k^{th} nearest neighbor (cf. [8]). Also ϕ_{yx} be the Cramer's coefficient defined as: $\phi_{yx} = (P_{11}(x,y) - \beta^2)/(\beta(1-\beta))$ where β lies between 0 and 1 depending upon the choice of quantile; $P_{11}(x,y) = P(X \le Q_x(\beta)$ & $Y \le Q_y(\beta))$ (cf. [9]).

Let y_i & x_i $(i = 1, 2, \ldots, n)$ be a sample of size n to get estimated values of $\beta-$quantile of Y & X as $\hat{Q}_y(\beta)$ & $Q_x(\beta)$ respectively. We consider three estimators of $Q_y(\beta)$, one usual and two based on an auxiliary characteristic *X* (using ratio and product patterns) defined as:

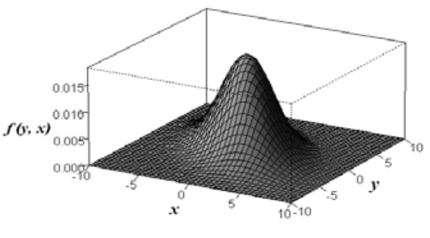
$$\left.\begin{array}{l} Usual:\ Q_u(\beta) = \hat{Q}_y(\beta) \\ Ratio:\ Q_r(\beta) = \hat{Q}_y(\beta)\left[Q_x(\beta)/Q_x(\beta)\right] \\ Product:\ Q_p(\beta) = \hat{Q}_y(\beta)\left[Q_x(\beta)/Q_x(\beta)\right] \end{array}\right\} \quad (1)$$

It is to be mentioned that we are taking $Q_x(\beta)$ to be a known quantile of auxiliary characteristic *X*. For the case of unknown $Q_x(\beta)$ we may estimate it by applying two phase sampling procedure (cf. [10] and [11]). The properties of the estimators, given in (1), can be easily obtained up to first order degree of approximation, following [10] and [12].

In our study we have considered normally distributed process environment under bivariate setup (Y, X) with density function given as:

$$f(y,x) = \frac{\exp[-0.5w/(1-\rho_{xy}^2)]}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}},\ \text{for}\begin{cases} -\infty < y,\ x < \infty & \sigma_y,\ \sigma_x > 0 \\ -\infty < \mu_y,\ \mu_x < \infty & -1 \le \rho_{yx} \le 1 \end{cases} \quad (2)$$

where,
$w = (y-\mu_y)^2/\sigma_y^2 + (x-\mu_x)^2/\sigma_x^2 - 2\rho_{xy}(y-\mu_x)(x-\mu_x)/(\sigma_y\sigma_x)$,
μ_y & μ_x are means of *Y* & *X* respectively, σ_y^2 & σ_x^2 are variances of *Y* & *X* respectively, σ_{yx} is covariance between *Y* & *X*, and $\rho_{xy} = \sigma_{xy}/(\sigma_x\sigma_y)$ be the correlation coefficient between Y & X. The bivariate normal density plot is given as:



Based on the estimators, given in (1) for $\beta = 0.25$ & 0.75, we define interquartile range statistic as:

$$IQR_i = Q_i(0.75) - Q_i(0.25)\ \ (\forall\ i = u,\ r\ \&\ p) \quad (3)$$

We may define the control charting structures based on *IQR$_i$* $(\forall\ i = u, r\ \&\ p)$ to monitor the dispersion parameter σ_y of quality characteristic *Y*. The probability limits of *IQR$_i$* based charting structures can be described as:

$$\text{Prob Limits:}\begin{cases} LPL = IQR_{il} \text{ with } F\left(IQR_i = IQR_{il}\right) \le \alpha_l \\ CL\ \ = IQR_{ic} \\ UPL = IQR_{iu} \text{ with } F\left(IQR_i = IQR_{iu}\right) \ge 1-\alpha_u \end{cases} \quad (4)$$

where *CL*, *LPL* and *UPL* refer to the Central Limit, Lower Probability Limit and Upper Probability Limit respectively of *IQR$_i$* charts and $\alpha = \alpha_l + \alpha_u$ is a pre-specified false alarm rate which is equally divided on both tails of the probability distribution of *IQR$_i$* to define the probability limits. It is to be mentioned that we may also define *K*-sigma limits of the structures based on *IQR$_i$* $(\forall\ i = u, r\ \&\ p)$ following [13].

It is to be noted that a variety of sensitizing rules are available in quality control literature which are used to differentiate between in-control and out-of-control states of process (cf. [14], [15], [16]). In our study we focus on first sensitizing rule to decide about process status for the control structures defined in (4). The first rule is defined as: Simulate bivariate random samples (y_i, x_i) of size n from the probability model (2) and compute the sample statistics *IQR$_i$s* for each sample. Plot the values of *IQR$_i$s* against the control limits defined in (4) or the appropriately defined K-sigma limits. By first sensitizing rule, any value of *IQR$_i$* falling outside the control limits indicates an out-of control signal for the dispersion parameter of quality characteristic *Y*.

3. Power Study of IQR Charts

To quantify the efficiency of a design structure, the discriminatory power is very famous performance measure in control charting setups. In this section we have evaluated the efficiency of *IQR$_i$* charts under investigation in terms of detection ability for shifts in process dispersion parameter and created power curves following [17], [7] and [18]. The in-control value of σ_y is considered as σ_y^0 while the out-of-control value is considered as σ_y^1 which can be defined in terms of λ and σ_y as $\sigma_y^1 = \lambda\sigma_y^0$, where λ is amount of shift in process dispersion σ_y. It is generally desired that for in-control state of process the false alarm rate should be low/close to pre-fixed value of α while for out-of-control state the power of charts should be high to detect the shifts in process parameters.

In order to investigate the performance of the IQR_i $(\forall\ i = u,\ r\ \&\ p)$ charts in terms of signaling probability, the power expression can be defined as:

$$\text{Power} = Pr[(IQR_i < LPL \text{ or } IQR_i > UPL) \mid \sigma_y^1 = \lambda \sigma_y^0] \quad (5)$$

A Monte Carlo simulation study with 100,000 replications is conducted under probability model (2) for different parameter values and different choices of n, α and ρ_{yx}. By varying the values of λ from $1.0-7.0$ we have evaluated (5) and provided the resulting power curves of IQR_i charts in Figures 1-3 for $\mu_y = \mu_x = 5,\ \sigma_y^2 = \sigma_x^2 = 1,\ \rho_{yx} = 0.50,\ 0.70\ \&\ 0.90,\ n = 10$ and $\alpha = 0.002$. In Figures 1-3, λ is plotted on horizontal axis and power values are plotted on vertical axis. The symbols IQR_u, IQR_r & IQR_p refer to the power curves of IQR_u, IQR_r & IQR_p charts respectively. The power curve analysis reveals the following points for the charting structures under dissection.

- The usual IQR_u chart performs better than ratio and product type IQR_r and IQR_p charts for low correlations (cf. Figure 1), while for moderate and high correlations IQR_r chart outperforms the IQR_u and IQR_p charts (cf. Figures 2 & 3).
- The performance of IQR_r chart keeps improving with the increase in the amount of correlation between *Y* and *X*, which is not the case with the IQR_p chart.
- The most inferior performance is exhibited by IQR_p chart. The reason behind this inferiority may be due to the fact that IQR_p chart is based on product estimator of $\beta-$quantile and according to [10], the product estimator is less efficient than usual estimator for $\rho_{yx} > 0$.
- The performance of all the charts has a direct relationship with the values of λ and n as expected.

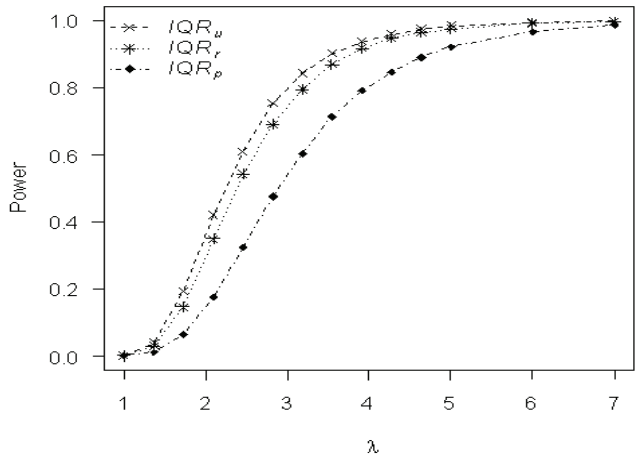


Figure 1. Power Curves of IQR_u, IQR_r and IQR_p charts under bivariate normal distribution for n=10, ρ_{yx} =0.50 & α =0.002

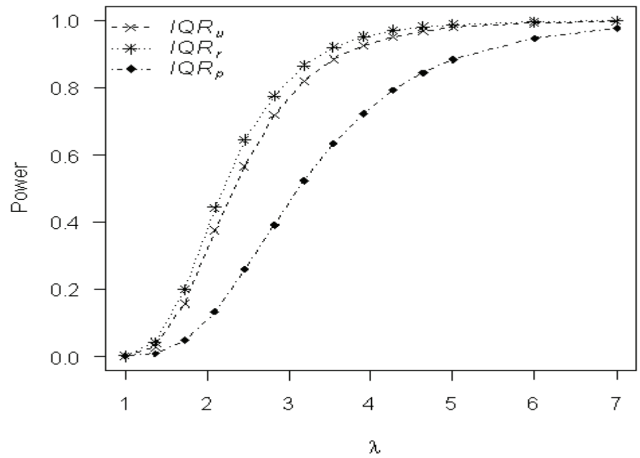


Figure 2. Power Curves of IQR_u, IQR_r and IQR_p charts under bivariate normal distribution for n=10, ρ_{yx} =0.70 & α =0.002

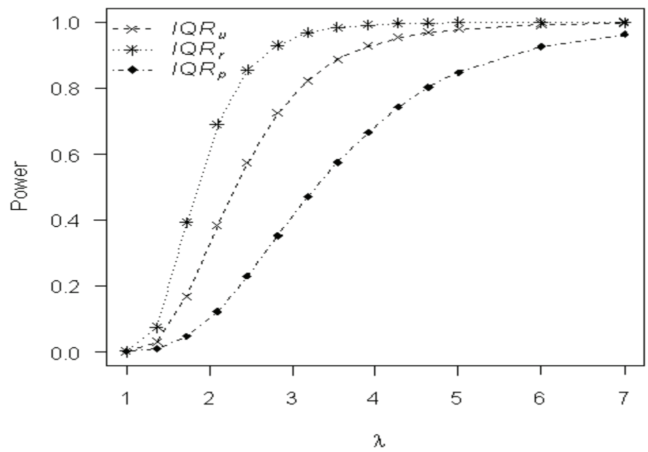


Figure 3. Power Curves of IQR_u, IQR_r and IQR_p charts under bivariate normal distribution for n=10, ρ_{yx} =0.90 & α =0.002

4. Illustrative Example

In order to justify our findings of power study in Section III, an example is provided to compare the performances of usual IQR_u chart and an auxiliary information based IQR_r chart. In real life examples the variables *Y* and *X* may refer as *i) Y:* the tensile strength in (psi) and *X*: the outside diameter in (mm) to monitor production of steel wire; *ii) Y:* production of pharmaceutical products in (units) and *X*: the temperate in (°C) in monitoring of pharmaceutical products etc.

For the said purpose we have simulated 30 samples each of size n=10 from probability model (2) with $\mu_y = \mu_x = 5$, $\sigma_y = \sigma_x = 1$ and $\rho_{yx} = 0.90$. The first $m_0 = 20$ samples are generated from in-control state i.e. $\lambda = 1$ whereas the remaining $m_1 = 20$ observations are generated from an out-of-control state with $\lambda = 2.5$ and computed the values of the charting statistics IQR_u and IQR_r. The resulting values are demonstrated in the form of control chart by plotting sample number on horizontal axis and values of IQR_u and IQR_r on vertical axis in Figure 4. The solid lines refer to the control limits and values of IQR_r chart while the dotted lines refer to the control limits and values of IQR_u chart.

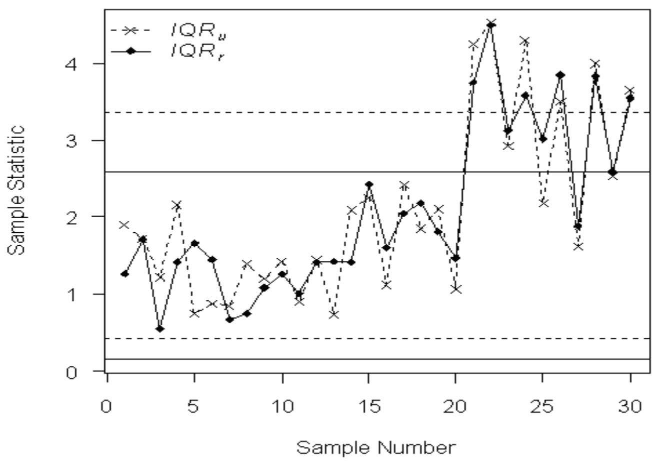


Figure 4. Control Chart Plots of IQR_u and IQR_r under bivariate normal distribution for n=10, ρ_{yx} =0.90, λ =2.5 & α =0.0020

It is obvious from Figure 4 that after 20th sample the IQR_u chart has detected 6 out-of-control points while the IQR_r chart has indicated 9 out-of-control signals. It mean that IQR_r chart has given 3 more out of control signals as compare to IQR_u chart which is in accordance with the finding of power study in Section III.

5. Summary, Conclusion and Recommendations

For an imporved monitoring of process dispersion, we have investigated Shewhart-type interquatile range charts namely IQR_r and IQR_p charts. The design structures of these charts are based on ratio and product esitmators of lower and upper quartiles of quality charctersitc *Y* with the assitance of an auxiliary charcatersic *X*. For comparison purposes we have also included the usual interquatile range chart namely IQR_u chart. We have observed that the detection ablity of IQR_r chart is possitively releated with the correlation between *Y* & *X*. For low correlations the IQR_r chart offers lower detection ablity than the usual IQR_u chart but with the increament in ρ_{yx}, IQR_r chart outperforms IQR_u and IQR_p charts. The most inferior performance is exibited by IQR_p chart for any possitivly correlated process environmet because the product esitmator is more efficient than usual estimator in case of negative correlation between *Y* and *X*.

The scope of the study may be extended for different contaminated process scenarios under Shewhart, EWMA and CUSUM setups. Moreover, the multivariate versions of these control charts may be another direction to be explored.

6. Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 11171303) and the Specialized Research Fund for the Doctor Program of Higher Education (No. 20090101110020).

REFERENCES

[1] W.A. Shewhart, "Economic control of quality manufactured product. D. Van Nostrand", (reprinted by the American Society for Quality Control in 1980, Milwauker, WI), New York, 1931.

[2] M. Riaz, "A Dispersion Control Chart", Communications in Statistics - Simulation and Computation, 37(6), 1239-1261, 2008.

[3] D.M. Rocke, "Robust Control Charts", Technometrics, 31(2), 173-184, 1989.

[4] L.G. Tatum, "Robust Estimation of the Process Standard Deviation for Control Charts", Technometrics, 39(2), 127-141, 1997.

[5] D.M. Rocke, " $\bar{X}_Q$ and R_Q Charts: Robust Control Charts", Journal of the Royal Statistical Society. Series D (The Statistician), 41(1), 97-104, 1992.

[6] H.A. David, "Early sample measures of variability", Statistical Science, 13(4), 368-377, 1998.

[7] S.A. Abbasi, and A. Miller, "On proper choice of variability control chart for normal and non-normal processes", Quality and Reliability Engineering International, 28(3), 279-296, 2012.

[8] B.W. Silverman, "Density Estimation for Statistics and Data Analysis", Chapman & Hall, London, 1986.

[9] M.D. MartÃnez - Miranda, M. Rueda, and A. Arcos, "Looking for optimal auxiliary variables in sample survey quantile estimation", Statistics, 41(3), 241-252, 2007.

[10] P.V. Sukhatme, and B.V. Sukhatme, "Sampling Theory of Surveys with Application", Iowa Statistical University Press, New York, 1970.

[11] M.M. Rueda, A. Arcos, J.F. Muñoz, and S. Singh, "Quantile estimation in two-phase sampling", Computational Statistics & Data Analysis, 51(5), 2559-2572, 2007.

[12] M.M. Rueda, A. Arcos, and M.D. Martínez, "Difference Etimators of Quantiles in Finite Populations", TEST, 12(2), 481-496, 2003.
[13] D.C. Montgomery, "Introduction to Statistical Quality Control", 6, Wiley, New York, 2009.
[14] R.J.M.M. Does, and B.F. Schriever, "Variables control chart limits and tests for special causes", Statistica Neerlandica, 46(4), 229-245, 1992.
[15] M.B.C. Khoo, "Designs of runs rules schemes", Quality Engineering, 17, 27-43, 2004.
[16] M.V. Koutras, S. Bersimis, and P.E. Maravelakis, "Statistical process control using shewhart control chart with supplementary Runs Rules", Methodology and Computing in Applied Probability, 9, 207-224, 2007.
[17] M. Riaz, and A. Saghir, "A mean deviation-based approach to monitor process variability", Journal of Statistical Computation and Simulation, 79, 1173–1193, 2009.
[18] A. Saghir, Z. Lin, S.A. Abbasi, and S. Ahmad, "The Use of Probability Limits of COM–Poisson Charts and their Applications", Quality and Reliability Engineering International, 1-16, 2012.