PROCESS MONITORING FOR NONCONFORMING PROPORTION IN MANUFACURING

Ângelo Márcio Oliveira Sant'Anna
Industrial and Systems Engineering Graduate Program
Pontifical Catholic University of Parana
Imaculada Conceição Street, 1155, Curitiba, Parana, Brazil
angelo.santanna@pucpr.br

ABSTRACT

This paper presents the application of Statistical Process Control tool for the improvement of the manufacturing processes performance of defects proportion production processes. Successful applications of these traditional tools can be found in the literature. However, some companies do not manage to apply control charts successfully when the data do not Normal distribution. In this paper we apply Beta control chart for monitoring proportion data (p). The Beta Chart presents the control limits based on the Beta probability distribution. This control chart was used for monitoring proportion data in nitric acid manufaturing process and it was compared to the traditional control limits. The analysis showed that the charts proposed were more sensitive to the average run length (ARL). These results are useful in the study of the manufacturing of products that involve proportion data. The new approach is easy to implement and hence might be attractive and useful to practitioners.

KEYWORDS. statistical quality control, control charts, proportion data, Beta distribution.

(Statistics)
(OR in Industry)

1. Introdução

The literature reveals that the Statistical Process Control (SPC) has been used to monitor the quality and reliability of products in numerous industrial processes, due mainly to its operational simplicity. Traditionallly control charts have been used to monitor the quality characteristics of a process. Attribute control charts are important tools found in SPC to monitor processes with discrete data. The p-Charts and np-Charts are more popular for monitoring nonconforming items, developed by Shewhart in 1924. Estimates of mean and variance are calculated assuming a Binomial probability distribution with *n* and *p* parameters to the number of nonconforming items, and the control limits are calculated based on the Normal distribution approximation.

Many studies in different areas of knowledge assess variables with observations expressed in the interval [0,1], without adequately defining the terminology of the variable in question. The Figure 1 shows the classification of the terminologies used to define the term fraction data. The first category percentage - includes the ratio between two discrete numbers, i.e., the number of defective items divided by the total of items of the lot. The second category proportion - includes the ratio between two continuous numbers, i.e., the volume of crude oil converted into gasoline divided by the total volume.

Terms	Category	Form	Example	Distribution
Fraction Data	Percentage	$\frac{\text{discrete}}{\text{discrete}}$	$\frac{\text{n° of defective items}}{\text{total of items}}$	Binomial
	Proportion	$\frac{\text{continuous}}{\text{continuous}}$	$\frac{\text{researched volume}}{\text{total volume}}$	Beta

Figure 1 – Terminology of the term to fraction data

There are some rules that deal with the suppositions of symmetry and Normal distribution approximation. Fleiss et al. (2003) described an approximation as satisfactory, when the size of p is in the range ($0.3 \leq p \leq 0.7$) and n is extremely large for $np \geq 5$ and $n(1 - p) \geq n$, the variance $p(1 - p)$ remains constant, while Montgomery (2005) indicated that a Normal approximation to the Binomial distribution is satisfactory when $np \geq 10$ and p is in the range ($0.1 \leq p \leq 0.9$).

In many studies the p-Charts are used in situations where the parameter p is considered small (i.e. p = 0.001; 0.01; 0.05; 0.1; ...). In these cases a Binomial distribution is quite skewed and the approximation by a Normal distribution is not satisfactory, as it allows values: negative or greater than one. For example, Let Y be a random variable with Binomial distribution and parameters p = 0,001; 0,01; 0,05; 0,1 e n = 100, we can be seen several shape of the Binomial distribution to these data, see Figure 2.

Some studies introduce several enhancements of the p-Chart to monitor the quality characteristic of the products such as proportion data: Quesenberry (1991) proposed a Binomial Q Chart to monitor nonconforming fraction using a nonlinear transformation for the control limits and demonstrated that it approximates the Normal distribution closer to the Binomial. Heimann (1996) presented a modification of the p-Chart control limits for large sample sizes (n>10,000), noting that in this case the control limits are narrow, thus the false alarm rate increases. Schwertman and Ryan (1997) suggested modifications of the np-Chart control limits to fit on the Normal approximation when p<0.03. McColl and Motley (1998) applied a power transformation ($x^{0,2777}$) for small nonconforming fraction (5ppm) as a better approximation of the Normal distribution to the Binomial distribution.

Chen (1998) proposed an adjustment to the p-Chart control limits and compared them with the traditional p-Chart and the Binomial Q Chart using the false alarm rate, while Bourke (2008) compared the performance of four control charts by monitoring shifts of the nonconforming fraction in industrial processes. Thereby he noted similarities in the performance of the Synthetic control chart and np-Chart over a long time period of in-control process. With similar purposes, Hsieh *et al.*, (2007) applied fuzzy theory to monitor wafer defects Poisson distributed in manufacture process. Sim and Lim (2008) adapted the attribute control charts to monitor zero-inflated data and used the Blyth-Still interval with 3-sigma to calculate control limits assuming that this data follows a Binomial and Poisson distribution. Aebtarm and Bouguila (2011) compared the performance with eleven control charts for monitoring defects with Poisson distribution. Sant'Anna and Caten (2012) proposed control limits to monitor the nonconforming defects based on Beta distribution.

Using others distributions, Bourke (1991) presented a control chart to monitor the number of occurrences between successive nonconforming fractions and assumed the Geometric probability distribution, while Kaminsky *et al.*, (1992) proposed control charts fitting a modified Geometric distribution. Yang *et al.*, (2002) applied a Geometric distribution for monitoring processes with high quality ($p < 0.001$ and $n > 50{,}000$), although the false alarm rate is slightly affected when the sample size is small or there is a decrease of the nonconforming proportion, while Zhang *et al.*, (2007) used a Gamma distribution to monitor the time between the occurrences.

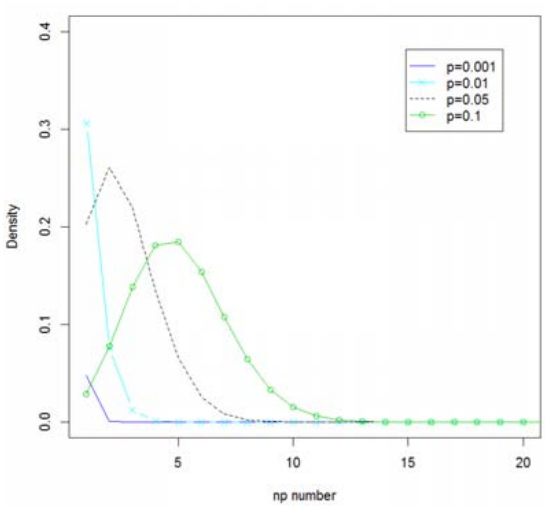

Figure 2 – Shape of the Binomial distribution to *n* and *p*

When the data distribution in industrial processes is asymmetric, the false alarm rate increases as the asymmetry because of the discrepancy between the shape of the data distribution and Normal distribution. Ferrell (1958) and Nelson (1979) suggested in these cases to assume that the data distribution is known and to construct control charts with exact limits, which provide desired false alarm rates.

This paper proposes to apply a Beta control chart for monitoring proportion data in nitric acid manufaturing process. This control chart assumes that the proportion data can be approximated by a Beta distribution and uses control limits based on this distribution. The results show that this scheme monitors well asymmetrically distributed data commonly found in industrial scenarios. In addition, sensitivity analyses demonstrate that this scheme is remarkably better than Shewhart charts in both, in-control and out-of-control process monitoring.

2. Probability Distributions

Let Y be a random variable that measures the number of nonconforming items (y_i) in a sample size of (n_i) independent items, i = 1, 2, ..., m. The probability of Y (P{$Y_i = y$}) is defined by the Binomial distribution,

$$P(Y_i = y) = \binom{n}{y} \pi^y (1-\pi)^{n-y} \quad (1)$$

If the percentage of nonconforming products is measured in a data set ($\pi_i = y_i / n_i$) which defines $Y_i \sim$ Bin(n_i ; π_i) and π_i as the nonconforming percentage. If the occurrence of this nonconforming percentage is independent and identically distributed, and based on central limit theorem, it can be assumed that the nonconforming percentage follows the Normal probability distribution for a sufficiently large n. This random variable Y Binomially distributed has the mean and variance given by, respectively, (Johnson *et al.*, 2005; Montgomery, 2005).

$$E(Y) = n\pi \quad and \quad Var(Y) = n\pi(1-\pi) \quad (2)$$

The approximation to the Binomial distribution by a Beta distribution may be more appropriate because the Beta density function can present a variety of form. Thus, it is assumed that the random variable (Y_i) follows the Binomial distribution and the percentage (π_i) obtained from the variable (Y_i) for each occurrence (i = 1, 2, ..., m) could follows a Beta probability distribution when indexed by the parameters (θ_1, θ_2), where θ_1, θ_2 > 0.

Moitra (1990) states that the two parameter Beta distribution can model a large variety of variable since its probability density function (pdf) can assume several shapes. Johnson *et al*., (1995) corroborates that the Beta distribution can easily approximate other statistical distributions, while Teerapabolarn (2008) describes that the Beta distribution is bounded for modeling more specific cases.

The Beta distributions family comprises all probability distributions which present a random variable Y, the pdf depends of the parameters θ_1 and θ_2, and its pdf can be written as,

$$f(y;\theta_1,\theta_2) = \frac{\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)} y^{\theta_1-1}(1-y)^{\theta_2-1} \quad (3)$$

where $\Gamma(\theta)$ is a Gamma function assessed at point θ, i.e. with $\Gamma(\theta) = \int_0^{\propto} y^{\theta-1}e^{-y}$, $\theta > 0$.

Johnson *et al.* (1995) describes that a random variable Y with Beta distribution of two parameters has a mean and variance given by, respectively,

$$E(Y) = \frac{\theta_1}{\theta_1+\theta_2} \quad and \quad Var(Y) = \frac{\theta_1\theta_2}{(\theta_1+\theta_2)^2 \cdot (\theta_1+\theta_2+1)} \quad (4)$$

3. Control Charts

The control limits of the traditional Shewhart chart for monitoring the nonconforming data are determined by Eq. (5), assuming that the sample size (n) is too large that a Binomial distribution is approximately symmetrical on the mean ($\bar{p}$). This implies that this distribution can be approximated to a Normal distribution.

$$LCL = \bar{p} - w\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \quad ; \quad UCL = \bar{p} + w\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \quad (5)$$

where w is a constant that sets the width of the control limits corresponding to a control region (1 - α) and a desired average run length until a false alarm (ARL_0). Usually, the value w = 3 is used due to the Normal distribution approximation, corresponding a control region = 0.9973 and ARL_0 = 370.

The control limits of the Beta Chart presented in this paper for monitoring nonconforming data are determined by Eq. (6), assuming that usually it follow non-Normal and asymmetric distributions. This control chart has the advantages to monitor data with small values and in an asymmetrical shape.

$$LCL = \bar{p} - w_1\sqrt{s^2(\bar{p})} \qquad ; \qquad UCL = \bar{p} + w_2\sqrt{s^2(\bar{p})} \qquad (6)$$

where $\bar{p}$ and $s^2(\bar{p})$ represent the mean and variance of the fraction estimated through Eq. (4), w_1 and w_2 are constants that define the width of the control limits, corresponding a control region (1 - α) and a desired average run length (ARL_0). The), w_1 and w_2 represent the percentiles $\frac{\alpha}{2}$ e $\left[1-\frac{\alpha}{2}\right]$ of *cdf* of the random variable Y according to control region desired, for example, (1 - α) = 0.9973 and ARL_0 = 370.

4. Manufacturing Example

This section illustrates one example real data set of a manufacturing process of nitric acid by ammonia oxidation collected over 21 days, which will be analyzed using the control limits proposed by Shewhart and Beta charts. This data were published by Brownlee (1965, p.454). The variable monitored in this paper is proportion of ammonia unconverted (y_1). The variable y_1 is estimated by the ratio between the volume of nonconforming raw material and the total volume produced.

Figure 3 illustrates the histogram and density distribution of variables y_1 with Normal density overlain, when the asymmetric and bimodal distribution of the data can be seen. A statistical summary of the data investigated is showed too.

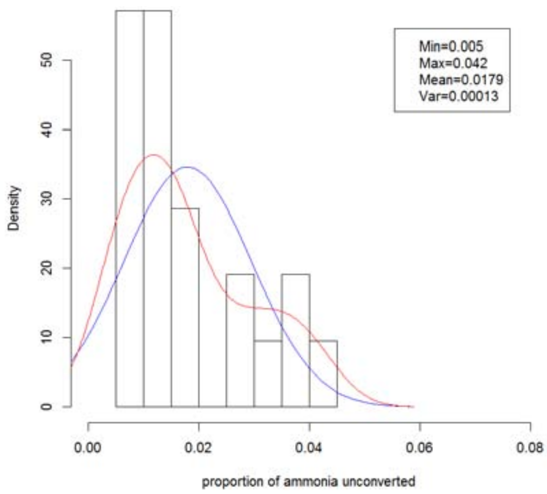

Figure 3 – Histogram and density distribution of the data with Normal density overlain

The chart control limits were calculated for the probability of false alarms (α = 0.0027) in the monitoring process based on the Normal distribution for the Shewhart and on the Beta distribution for the Beta Chart. The Eqs. (7, 8) show the control limits calculated using sample estimates depicted in Figure 3. The Shewhart chart extrapolate the lower limit for monitoring proportion of ammonia unconverted, introducing negative values, while the Beta Chart does not extrapolate this region, for considering in-control process [see Figure 4].

Finally, the Shewhart chart with approximation to the Normal distribution showed negative, while the Beta chart with control limits based on the Beta distribution presented satisfactory estimates within the [0,1]-interval. The Beta control limits used are more stable for monitoring nonconforming proportion.

$$\text{Shewhart Chart:}\quad \begin{aligned} LCL &= 0.0179 - (3)(0.0289) = -0.0689 \\ CL &= 0.0179 \\ UCL &= 0.0179 + (3)(0.0289) = 0.1047 \end{aligned} \tag{7}$$

$$\text{Beta Chart:}\quad \begin{aligned} LCL &= 0.0179 - (1.49)(0.0115) = 0.0007 \\ CL &= 0.0179 \\ UCL &= 0.0179 + (4.65)(0.0115) = 0.0715 \end{aligned} \tag{8}$$

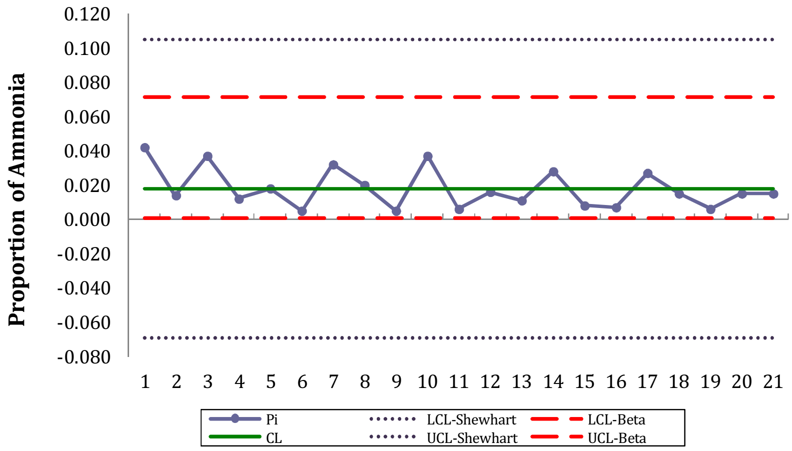

Figure 4 – Control charts based on Shewhart and Beta control limits by proportion data

We now perform a sensitivity study to compare Shewhart and the Beta charts in two scenarios: in-control and out-of-control; calculating the average run length (ARL_0) for values of interval *p*. For the in-control process, we evaluate the ARL_0 until a false special cause and the higher the ARL_0 is better. It can be written as function of the type I error probability (α). For the out-of-control process, we calculated the average run length until a true special cause; the lower the ARL_1 is better. The ARL_1 can be written as a function of the type II error probability (β), which is the probability of the control chart not detecting a change in the out-of-control process, [see Eq. 9] (Montgomery, 2005).

$$ARL_0 = \frac{1}{\alpha} \qquad ; \qquad ARL_1 = \frac{1}{[1-\beta]} \tag{9}$$

The sensitivity of the control charts investigated for p = 0.0179 and n = 21 presents for a process in-control of ARL_0 = 167 and ARL_0 = 370 for Shewhart and Beta charts, respectively. As the Binomial distribution is discrete, the 3-sigma lower control limits are truncated at the value zero, the nominal value of α is not equal to 0.0027.

Thus, the control charts with approximation by Normal distribution present narrow limits, generating ARL_0 less than 370 samples and false alarms more frequently. While using the approximation by Beta distribution it is possible to define an exact nominal value [see Table 1].

Table 1 – Sensitivity analysis using ARL from the manufacture data

	p	*LCL*	*UCL*	α	*β*	*ARL*
Shewhart	0.0179	-0.0689	0.1047	0.0060	-	167
	0.020	-0.0717	0.1117	-	0.992	123
Beta	0.0179	0.00074	0.0715	0.0027	-	370
	0.020	0.00104	0.0716	-	0.988	83

7. Conclusion

The aim of this paper was to apply a Beta control chart for monitoring proportion data from nitric acid manufaturing process. This control chart assumes the Beta probability distribution for the proportion has values restricted to the interval [0,1] and it have a symmetric density in several industrial processes.

The Beta chart led to more precise results than the Shewhart chart used for proportion monitoring. The control limits used in the Beta Chart present the in-control average run length (ARL_0 = 370) of 0.9973 for the probability (α = 0.0027) and a faster detection of the changes for all the out-of-control scenarios with lower values for ARL_1.

Furthermore, a sensitivity analysis corroborated the superior performance of the Beta chart compared to the Shewhart. This analysis allowed the assessment that the approximation of the Binomial distribution by the Beta distribution was more appropriate, as it provided values restricted to the interval [0,1].

Referências

Aebtarm, S.; Bouguila, N. An empirical evaluation of attribute control charts for monitoring defects, Expert Systems with Applications, v. 38, p. 7869–7880, 2011.

Bourke, P.D. Detecting a shift in fraction nonconforming using Run-Length control charts with 100% Inspection, Journal of Quality Technology, Milwaukee, v. 23, n. 3, p. 225–238, 1991.

Bourke, P.D. Performance comparisons for the Synthetic control chart for detecting increases in fraction nonconforming, Journal of Quality Technology, Milwaukee, v. 40, n. 4, p. 461–475, 2008.

Brownlee, K.A. Statistical theory and methodology in science and engineering, New York: John Wiley & Sons, 1965, 590p.

Chen, G. An improved p Chart through simple adjustments, Journal of Quality Technology, Milwaukee, v. 30, n. 2, p. 142–151, 1998.

Ferrell, E. B. Control charts for log-normal universe, Industrial Quality Control, Québec, v. 15, n. 1, p. 4–6, 1958.

Fleiss, J.L., Levin, B.; Paik, M.C. Statistical methods for Rates and Proportions, 3ª ed., New Jersey: John Wiley & Sons, 2003, 760p.

Heimann, P.A. Attribute control charts with large sample sizes, Journal of Quality Technology, Milwaukee, v. 28, n. 4, p. 451–459, 1996.

Hsieh, K.L., Tong, L.I., Wang, M.C. The application of control chart for defects and defect clustering in IC manufacturing based on fuzzy theory, Expert Systems with Applications, v. 32, p. 765–776, 2007.

Johnson, N.L.; Kemp, A.W.; Kotz, S. Univariate Discrete Distributions. 3ª ed., New York: John Wiley & Sons, 2005, 646p.

Johnson, N.L.; Kotz, S.; Balakrishnan, N. Continuous Univariate Distributions, v. 2, 2ª ed., New York: John Wiley & Sons, 1995, 752p.

Kaminsky, F.C.; Benneyan, J.C. Davis, R.D.; Burke, R. Statistical control charts based on a Geometric distribution, Journal of Quality Technology, Milwaukee, v. 24, n. 2, p. 63–69, 1992.

McColl, J.I.; Motley, T.J. Control charts applicable when the fraction nonconforming is small, Journal of Quality Technology, Milwaukee, v. 30, n. 3, p. 240–247, 1998.

Moitra, S.D. Skewness and the Beta distribution, Journal of the Operational Research Society, Houndmills, v. 41, n. 10, p. 953–961, 1990.

Montgomery, D.C. Introduction to Statistical Quality Control, 5ª ed., New York: John Wiley & Sons, 2005, 732p.

Nelson, P. R. Control charts for Weibull processes with standards given, IEEE Transactions on Reliability, Knoxville, v. 28, n. 3, p. 383-387, 1979.

Quesenberry, C.P. SPC Q charts for a Binomial parameter p: short and long runs, Journal of Quality Technology, Milwaukee, v. 23, n. 3, p. 239–246, 1991.

Sant'Anna, A.M.O.; Caten, C.S. Beta control charts for monitoring fraction data, Expert Systems with Applications, v. 39, p. 10236–10243, 2012.

Schwertman, N.C.; Ryan, T.P. Optimal limits for attributes control charts, Journal of Quality Technology, Milwaukee, v. 29, n. 1, p. 86–98, 1997.

Sim, C.H.; Lim, M.H. Attribute charts for zero-inflated process, Communications in Statistics – Simulations and Computation, Hamilton, v. 37, n. 7, p. 1440–1452, 2008.

Teerapabolarn, K. A Bound on the Binomial Approximation to the Beta Binomial Distribution, International Mathematical Forum, Sofia, v. 3, n. 28, p. 1355–1358, 2008.

Yang, S-F.; Yeh, J-T. Using cause selecting control charts to monitor dependent process stages with attributes data, Expert Systems with Applications, v. 38, p. 667–672, 2011.

Zhang, C.W.; Xie, M.; Goh, T.N.; Liu, J.Y. A control chart for the Gamma distribution as a model of time between events, International Journal of Production Research, London, v. 45, n. 23, p. 5649–5666, 2007.