Bayesian Control Chart for the Fraction of Nonconforming Units

Lizanne Raubenheimer[1], Abrie J. van der Merwe[2]

Second Bayesian Young Statisticians Meeting (BAYSM 2014)
Vienna, September 18–19, 2014

[1] Department of Statistics, Rhodes University, Grahamstown, South Africa
L.Raubenheimer@ru.ac.za

[2] Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa

Abstract

Quality control is a process which is used to maintain the standards of products produced or services delivered. The binomial distribution is often used in quality control. The usual operation of the control chart for the fraction of nonconforming units, or the p – chart, will be extended by introducing a Bayesian approach. Control chart limits, average run lengths and false alarm rates will be determined by using a Bayesian method. A predictive distribution based on a Bayesian approach will be used to derive the rejection region. The proposed Bayesian method gives wider control limits than those obtained from the classical method, and gives larger values for the average run length and smaller values for the false alarm rate.

Keywords: Beta-binomial distribution; false alarm rate; p - chart; run length.

1 Introduction

In this paper the control chart for the proportion of nonconforming units, also known as the p - chart, will be studied. The binomial distribution is often used in quality control. The proportion, p, denotes the proportion of defective items in the population. Control chart limits, average run lengths and false alarm rates will be determined by using a Bayesian method. These results will be compared to the results obtained when using the classical method. [1] states that attributes control techniques, such as p - charts, plot statistics related to defective items and call for corrective action if the number of defectives becomes too large. [2] proposed a Bayesian approach to obtain control charts when there is parameter uncertainty, using a predictive distribution to derive the rejection region. [2] assumed that the prior information on p is a beta distribution, which means that the posterior distribution of p will also be a beta distribution. Let X_i, for $i = 1, 2, \ldots, m$, follow a binomial distribution with parameters n and p. The proportion of nonconforming items from sample i is defined as $\hat{p}_i = {X_i}/{n}$. Then $\overline{p}$ is calculated, where $\overline{p}$ is the average of the sample proportions and is defined as $\overline{p} = \sum_{i=1}^{m} {\hat{p}_i}/{m}$. From [3] the classical control chart is defined as:

$$\text{UCL} = \overline{p} + 3\sqrt{\tfrac{\overline{p}(1-\overline{p})}{n}} \quad \text{and} \quad \text{LCL} = \overline{p} - 3\sqrt{\frac{\overline{p}\,(1-\overline{p})}{n}}.$$

For the Bayesian method, the predictive density will be used to determine the control chart. From a Bayesian point, we have to decide on a prior for this unknown value of p.

2 Prior, Posterior and Predictive Density

The beta prior is a conjugate prior to the binomial distribution. Consider a beta prior, i.e. $p \sim Beta\,(a, b)$ for the unknown p

$$\pi\,(p) \quad \propto \quad p^{a-1}\,(1-p)^{b-1}\,. \tag{1}$$

For the p - chart the likelihood follows as

$$L\,(p\,|data\,) \quad \propto \quad p^{\sum\limits_{i=1}^{m} x_i}\,(1-p)^{mn-\sum\limits_{i=1}^{m} x_i}\,. \tag{2}$$

Combining Equations 1 and 2 it follows that the posterior distribution of p is a beta distribution, i.e.

$$\pi\,(p\,|data\,) \quad \propto \quad p^{\sum\limits_{i=1}^{m} x_i+a-1}\,(1-p)^{mn-\sum\limits_{i=1}^{m} x_i+b-1}\,. \tag{3}$$

If the process remains stable, the control chart limits for a future sample of n Bernoulli trials which results in T successes can be derived. Given n and p, the unconditional prediction distribution of T is

$$f\,(T\,|data\,) \quad = \quad \binom{n}{T} \frac{B\left(\sum\limits_{i=1}^{m} x_i + a + T, mn - \sum\limits_{i=1}^{m} x_i + b + n - T\right)}{B\left(\sum\limits_{i=1}^{m} x_i + a, mn - \sum\limits_{i=1}^{m} x_i + b\right)} \tag{4}$$

for $0 \leq T \leq n$, known as the beta-binomial distribution. It is assumed that the sample size is the same for the posterior distribution and the future sample. The predictive distribution in Equation 4 can be used to obtain the control chart limits, where the rejection region is defined as $\alpha = \sum_{R^{\star}(\alpha)} f\,(T\,|data\,)\,.$

3 Simulation Study

In this simulation study the average run lengths and false alarm rates will be compared using the classical and proposed Bayesian method. The run length of a control procedure is the number of samples required before an out-of-control signal is given. A good control procedure has a suitably large average run length when the process is in-control and a small average run length otherwise, from [4]. We will consider a number of different samples sizes, n, and number of samples, m. For the classical method it is assumed that $p = 0.5$ when determining the average run length and the false alarm rate. For the Bayesian method, the value of p is of course unknown and the prior distribution given in Equation 1 is used. Four different priors were considered. For the simulation procedure, we randomly generated m binomial random variables. The 3-sigma control chart limits were calculated for the classical method. Followed by calculating the false alarm rate and the run length, using the binomial distribution. For the Bayesian method, the control chart limits were calculated using the predictive density with $\alpha = 0.0027$. Followed by calculating the the false alarm rate and the run length, using the beta-binomial distribution. This was repeated 100 000 times, and then the average of the false alarm rates was calculated, and also the average of the run lengths. From Table 1 we can see that in every single case, the false alarm rate is lower for one of the Bayesian methods. For the majority of the cases, the Bayesian method yielded a larger average run length. Typically, one wants a smaller false alarm rate and a larger average run length. For the Bayesian method, the false alarm rate is generally closer to the nominal level of 0.0027.

Table 1: (a) Average run lengths and (b) average false alarm rates for different values of m and n.

		Classical	**Bayes** $B(0.5, 0.5)$	**Bayes** $B(1, 1)$	**Bayes** $B(3, 3)$	**Bayes** $B(10, 3)$
$m = 2$ &	(a)	188.8749	338.0909	289.3581	309.4339	327.6048
$n = 25$	(b)	0.01832	0.00326	0.00364	0.00339	0.00335
$m = 4$ &	(a)	259.1217	327.80139	317.3218	288.4798	323.5345
$n = 25$	(b)	0.00778	0.00328	0.00349	0.00377	0.00331
$m = 4$ &	(a)	259.3864	341.1857	352.6781	339.7289	347.0327
$n = 50$	(b)	0.00708	0.00303	0.00295	0.00303	0.00295
$m = 10$ &	(a)	408.9574	303.4763	303.9957	305.9098	321.7337
$n = 20$	(b)	0.00374	0.00388	0.00388	0.00385	0.00365
$m = 2$ &	(a)	187.3209	358.9326	359.4457	363.4456	367.0650
$n = 250$	(b)	0.01424	0.00281	0.00280	0.00277	0.00274
$m = 10$ &	(a)	328.6722	333.4760	331.1047	332.1746	339.1109
$n = 50$	(b)	0.00399	0.00309	0.00313	0.00312	0.00304

4 Conclusion

The usual operation of the p - chart was extended by introducing a Bayesian approach. We considered four different beta priors. We conclude that the proposed Bayesian method gives wider control limits than those obtained from the classical method. The Bayesian method generally gives larger values for the average run length and smaller values for the false alarm rate.

References

[1] Calabrese, J. M. (1995). "Bayesian process control for attributes." *Management Science*, **41**(4), 637–645.

[2] Menzefricke, U. (2002). "On the evaluation of control chart limits based on predictive distributions." *Communications in Statistics - Theory and Methods*, **31**(8), 1423–1440.

[3] Montgomery, D. C. (1996). *Introduction to Statistical Quality Control.* John Wiley & Sons, New York, 3rd edition.

[4] Woodall, W. H. (1985). "The statistical design of quality control charts." *The Statistician*, **34**, 155 – 160.