*Journal of Institute of Science and Technology, 2014, **19(2):** 100-104, © Institute of Science and Technology, T.U.*

Bayesian Methods for Obtaining X-bar Chart: An Application on Food Quality Control

Ram Prasad Khatiwada
Central Department of Statistics
Tribhuvan University, Kathmandu, Nepal
E-mail: ramktd@gmail.com

ABSTRACT

This article is about the Bayesian methods of constructing control chart for process mean. The control chart for mean based on the use of normal prior distribution seems to be robust for normal observations. The concussion is derived from an experiment of a set of real dada of the weights of a food product filled in pouches. The results were also supported by the simulated data.

Keywords: control chart, uniform prior, posterior density, normal observations, predictive

INTRODUCTION

The main concern of quality control is realization of variability in the quality characteristics, between products, within the same process. This leads to make largest and smallest limits of the measurement, within which a product characteristic may vary, known as specification limits. In the beginning of the quality science the quality of product was measured with respect to specification limits. At that time a product within specification limits would be of acceptable quality level otherwise it would be inferior. This concept is still working in an acceptance sampling procedure for quality assurance.

A different approach of defining quality arose from the work of Shewhart (1931). His approach clearly based on the method of partitioning the variation within the product as due to *chance causes* or *assignable causes*. Deming (1986) renamed the causes of variation within process or product by *common causes* instead of *chance causes* and *special causes* instead of *assignable causes*. This approach of defining quality is that if the variability of the process or product is observed only due to the common causes and not affected by the special causes then the process is considered producing good-quality, otherwise it is assumed as producing not-good-quality. If the variation of the quality is observed only due to common causes then the state is considered to be *in-control* and if it is observed due to special causes then the state is considered as *out-of-control*.

Duncan (1974) described that the chance variations are due to time or possibly on some other basis and they behave in random manner and they may not show any cyclic behavior or runs. Nelson (1982) described the nature of statistical control as a state when changes in measure of variability and location from one sample period to the next are not greater than the statistical theory would predict. Taguchi (1976) argued that whatever be the source of variation, it is undesirable to be a deviation from the target. He further suggested needs of preplanning for reducing variability by the use of design of experiments. Shewhart (1931) developed the basis of control chart, which is a device used to decide whether the quality characteristic X of a product is in control or not. Girshick and Rubin (1952) introduced the perspectives of prior concept for process monitoring. Shiryaev (1963) and Robert (1959, 1966) used geometric distribution as the prior distribution for parameter p (proportion) for attribute control chart. de Finetti (1979) provided the mathematical concept of statistical control from exchangeable and subjective point of view.

The recent works on quality control are based on the application of Bayesian methods. This method is admired because its predictions are based on decision theoretic approach. Tsiamyrtzis (2000) urged that it usually incorporates prior distribution which is usually available in industrial settings. It is sensitive to detect a change in short runs where as the traditional approach of statistical process control require a bulky data gathering before charting and which is incapable to detect a change in short production run .

The application of Bayesian method in SPC is first appeared in Kihlstrom (1974), which contains good discussions on modeling uncertainty, product quality modeling and theory which explains consumer expenditures for product quality information. Chiu and Leung (1980) developed a theory for economic x-bar chart using reference prior distribution and derived the decision problem using the loss-cost function.

Woodward and Nayler (1993) presented the direct use of the Bayesian method in case of short runs. Ingleby and Lorenc (1993) presented and compared three different methods of quality control based on minimizing a Bayesian loss function.

Hautaniemi *et al* (2003) suggested a novel strategy for spot quality control by using Bayesian networks, which contain many appealing properties in the spot quality control in context of genetic research. Calle *et al* (2005) used Bayesian method for survival analysis for the model construction to sensory shelf life of foods. Marcellus (2008) presented a good application of Bayes' theorem to quality control as a rigid optimization model and to infer the values of structural parameters of the monitored process. Khatiwada and Sthapit (2008) described a Bayesian method to obtain the probabilistic control limits for the weights of a pouched product. Katiwada (2011) presented several ways of the use of the Bayesian methods to the food quality control problems.

In this context, this paper attempts to apply Bayesian methods of obtaining control limits of a control chart for mean by illustrating with the real as well as simulated data sets.

MATERIALS AND METHODS

Model

Let X be the measurement of the quality characteristic, x_{ij} denotes the measurement of j^{th} unit of i^{th} sample from a lot of size N; (i =1,2,……n) (j =1,2,……..k_i). The total number of samples observed is 'n'. The mean of the measurement (μ) of all units is the parameter of interest.

For an automated production system which produces batches of thousands items allows us to assume normality of the distribution of the characteristics being measured. Thus we take model assumption as data (X) follows Gaussian distribution with parameter μ and σ, $X \sim N\left(\mu,\sigma^2\right)$ where, μ and σ are the mean and standard deviation respectively.

The distribution of X for given μ is the *likelihood* of μ.

We assume, μ follows Gaussian *prior distribution* with mean μ_0 and standard deviation σ_0, $\mu \sim N\left(\mu_0,{\sigma_0}^2\right)$

Case I: When the variance of the process is set at a standard value (σ), process average is assumed normally distributed and unknown.

In such case the *posterior distribution* of parameter of interest (process average), given data will also be Gaussian with parameters μ_1 and σ_1; i.e.,

$$(\mu \mid X) \sim N\left(\mu_1,{\sigma_1}^2\right)$$

where, the posterior variance $({\sigma_1}^2)$ and posterior mean μ_1 are obtained (Carlin & Louis, 1996) as.

$${\sigma_1}^2 = \left(\frac{1}{{\sigma_0}^2}+\frac{n}{\sigma^2}\right)^{-1} \text{ and } \mu_1 = {\sigma_1}^2\left(\frac{\mu_0}{{\sigma_0}^2}+\frac{n\cdot\hat{\mu}}{\sigma^2}\right)$$

$\hat{\mu}$ is the estimated value of μ from data and n is the total number of sample.

The *predictive distribution* of the new sample after obtaining the posterior density of the first n samples is given by

$$\left(X_{n+1} \mid X_n\right) \sim N(\mu_p,{\sigma_p}^2)$$

where, $\mu_p = \mu_1$, and $\sigma_p{}^2 = \sigma^2 + {\sigma_1}^2$

Case II: When the process average value is assumed uniformly distributed unknown and process variance (σ^2) is known.

Let, the known process variance is σ^2, and we wish to obtain probability limits $(1-\alpha)100\%$ of the highest density region of the process mean θ, we set the model (for data) as

$$x_i \sim N(\mu,\ \sigma^2)$$

If we use uniform prior distribution on the whole real line for the parameter of interest

$$f(\mu) = C \text{ for } C > 0$$

It is easy to see that it combines with a normal density to give the standardized likelihood as posterior.

Then, the conditional posterior density of the parameter μ is

$$\mu \mid x \sim N(\bar{x},\ \sigma^2/n)$$

The data

The data were taken from an industry producing pouches of snacks. An experiment was designed to implement control chart procedure to maintain the weights of the product. The pouches were filled by an automated filling machine. The process was able to produce a large number of items so the weight of the pouches could be assumed to be normally distributed. The machine is so adjusted that the allowable standard deviation of the items produced was not more than 2 gm. 25 samples of each having 5 pouches drawn from the process revealed the information given in the Table 1.

Table 1. Sample mean and sample ranges of 25 samples of the product

Sample no.	1	2	3	4	5	6	7	8	9	10	11	12	13
Mean	98	97.6	97	98.4	100	100.8	99	99.4	96.8	95.6	96.6	97.2	99
Range	0.8	2.2	1.8	1.2	1.2	1	1.6	1	1.6	1	1	1.2	1.2
Sample no.	14	15	16	17	18	19	20	21	22	23	24	25	
Mean	100	97.6	100.8	100	100	94	99.8	95	94	98.8	98.8	100.8	
Range	1.8	2	0.8	0.8	2	0.8	1.2	1.6	1.6	2	1.2	1.2	

The estimated value of process mean $(\hat{\mu})$ and expected sample range $(\bar{R})$ were obtained as 98.20 and 1.35 respectively. The standard error for the estimated population mean computed as 0.8023 and the control limits for process mean in classical approach were obtained as 97.42 and 98.98.

RESULTS

Application of Bayesian Methods

Case I: using a normal prior from the process setup.

Considering the given process standard deviation 2, the variance of average weight obtained as $\sigma^2 / k = 4/5 = 0.8$; thus the prior distribution for average weight of the parameter of interest μ was assigned as $\mu \sim N(100,\ 0.8)$. The 3-sigma control limits for the process average were obtained 97.32 and 102.68 (Fig. 1). The distribution of sample average, $\bar{x}_i \mid \mu$ was $\bar{x}_i \mid \mu \sim N(98.2,\ 2^2)$. Clearly it gives a 3-sigma control limits within which sample statistics may fall are 92.2 to 104.2 (Fig. 2).

The posterior distribution of average weight, $\mu \mid \bar{x}_i \sim N(\mu_1, \sigma_1^2)$, was obtained as $\mu \mid \bar{x}_i \sim N(98.5, 0.133^2)$

where $\sigma_1^2 = \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right)^{-1} = \left(\frac{1}{0.8} + \frac{25}{4}\right)^{-1} = 0.133$ and

$$\mu_1 = \sigma_1^2\left(\frac{\mu_0}{\sigma_0^2} + \frac{n \cdot \hat{\mu}}{\sigma^2}\right) = 0.133\left(\frac{100}{0.8} + \frac{25 \times 98.2}{4}\right) = 98.5$$

The posterior 3-sigma control limits for process average were 97.40 and 99.60 (Fig. 3), which showed a sharper control range than the classical method (Fig. 4). Table 2 shows summaries of posterior for different number of additional sample observations. Results showed that the increase in the number of sample observations can give the same control limits.

Table 2. Posterior distribution of process average for different numbers of sample observations

Prior	Data distribution	Additional number of sample observations	Posterior distribution facts	3-sigma control limits for the process average weight
N(100, 0.8)	$N(\hat{\mu}, 4)$	5	N(98.5, 0.1333)	97.4-99.6
N(100, 0.8)	$N(\hat{\mu}, 4)$	100	N(98.5, 0.1333)	97.4-99.6
N(100, 0.8)	$N(\hat{\mu}, 4)$	1000	N(98.5, 0.1333)	97.4-99.6

In an analysis of predictive distribution, the 3-sigma range of the single sample drawn after implementing so obtained posterior control chart will be 92.4 to 104.6 (Fig. 5), since $(X_{n+1} \mid X_n) \sim N(98.5, 4.133)$. Table 3 shows summaries of predictive distribution of the average weight for different number of additional sample observations. Results showed that the increase in the number of sample observations could make a remarkable reduction on predictive variance and hence the size (range) of the control limits.

Table 3. Predictive distribution of average weight for different numbers of sample observations

Prior	Data variance	Additional number of sample observations	Posterior distribution facts	3-sigma control limits of the weights
N(100, 0.8)	$N(\hat{\mu}, 4)$	1	$N(98.5, 4.133)$	92.4-104.6
N(100, 0.8)	$N(\hat{\mu}, 4)$	5	$N(98.5, 0.933)$	95.6-101.4
N(100, 0.8)	$N(\hat{\mu}, 4)$	100	$N(98.5, 0.1733)$	97.25-99.75

Case II using a flat uniform prior.

The conditional posterior density of the parameter μ is

$$\mu \mid x \sim N(\bar{x},\ \sigma^2 / n) \quad = N(98.2\ ,\ 4/5)$$

Thus the posterior 3-sigma control limits for process average were 95.52 and 100.88. The control range so obtained was wider than that of normal prior selection, which happened due to a flat prior having a higher variance than that of the data (Fig. 6).

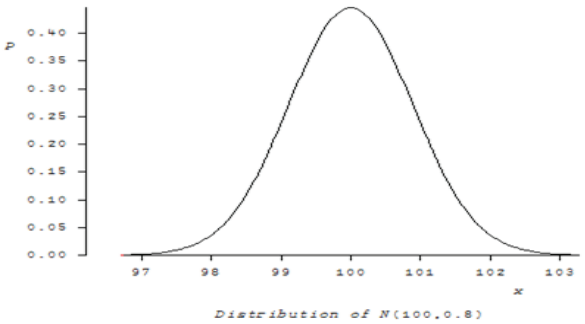

Fig. 1. Normal prior distribution of the process average

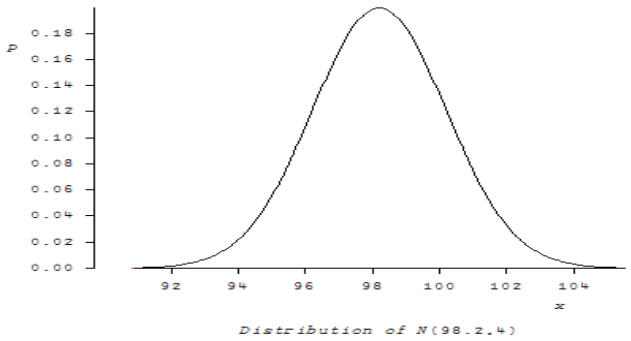

Fig. 2. Distribution of sample points (the likelihood of data)

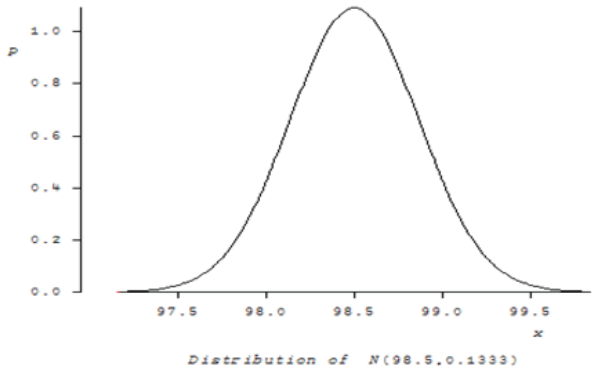

Fig. 3. Posterior distribution of process average, given data

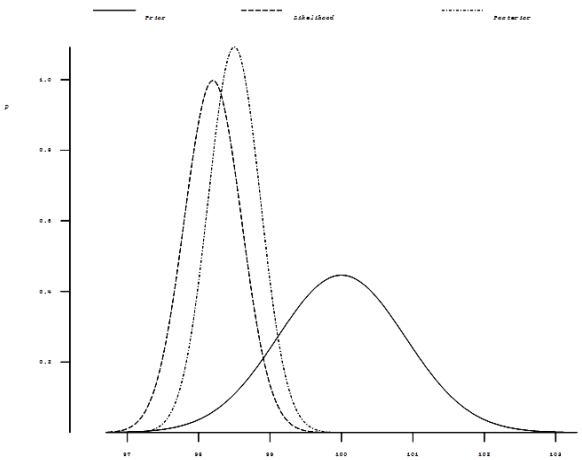

Fig. 4 Triplot of the prior to posterior for normal observation known variance model

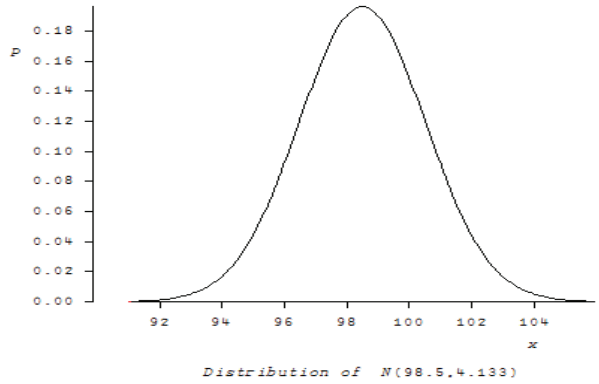

Fig. 5. Predictive distribution of single sample after posterior updating

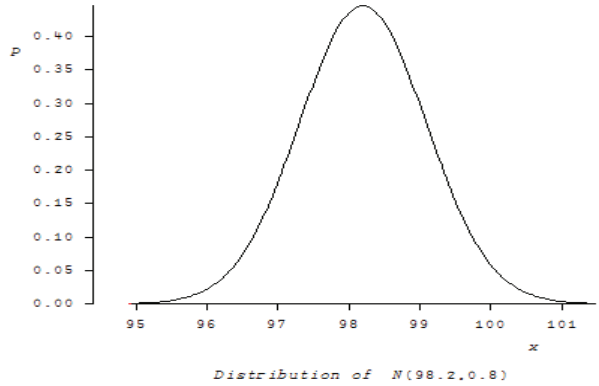

Fig. 6. Posterior distribution of process average with a uniform prior

CONCLUSION

This study clearly seems that Bayesian methods based on prior information provided by the previous experiments gives a good strength to obtain control chart for the current (running) process. The average value of the previous study can be taken as the mean of the normal prior and the variance as the known process variance. Information from the previous study can be used as the normal conjugate prior for the normal observations.

REFERENCES

Calle, M. L., Hough, G., Curia, A. and Gómez, G. 2006. Bayesian survival analysis modeling applied to sensory shelf life of foods, *Food Quality and Preference* **17** (3-4):307-312.

Carlin B. P. and Louis T. A. (1996). *Bayes and Empirical Bayes Methods for Data Analysis.* Chapman and Hall, New York.

Chiu W. K. and Leung M. P. 1980. A new Bayesian approach to quality control chart. *Metrika* **27**: 243-253.

de Finetti, B. 1979. Probability and exchangeability from a subjective point of view. *International Statistical Review **47:*** 129-135.

Deming, A. 1986. *Out of the Crisis.* MIT Press, Cambridge.

Duncan A. J. 1953. *Quality control and industrial statistics*. Richard D. Irwin, Homewood.

Girshick, M. A. and Rubin, H. 1952. A Bayes approach to a quality control model. *The Annals of Mathematical Statistics* **23:** 114-125.

Hautaniemi, H., Edgren, H., Vesanen, P., Wolf, M., Järvinen, A-K., Yali-Harja, O., Astola, J., Kallioniemi, O. and Monni, O. 2003. A novel strategy for microarray quality control using Bayesian networks. *Bioinformatics* **19** (16): 2031-2038.

Ingeby, N. B. and Lorence, A. C. 1993. Bayesian quality control using multivariate normal distributions. *Quarterly Journal of the Royal Meteorological Society* **119**: 195-1225.

Khatiwada, R. P. 2011. *Bayesian Modelling Approaches on some Issues of Agro-Food Production and Quality Control*. Ph. D. Thesis, Central Department of Statistics, Institute of Science and Technology, Tribhuvan University, Kathmandu, Nepal.

Khatiwada, R. P. and Sthapit, A. 2008. An application of Bayesian method in packaged food quality control. *Nepal Journal of Science and Technology* **9**: 41-48.

Kihlstrom, R. 1974. A Bayesian Model of Demand for Information about Product Quality. *International Economic Review* **15** (1): 99-118.

Marcellus R. L. 2008. Bayesian statistical process control. *Quality Engineering* **20** (1): 113 – 127.

Nelson, L. S. 1982. Control charts. In: *Encyclopedia of statistical sciences*, (ed.) Kotz and Johnson. Wiley, NewYork , 176-182 p.

Robert, S. W. 1959. Control chart based on geometric moving averages. *Technometrics* **1:** 239-250

Robert, S. W. 1966. A comparison of some control chart procedures. *Technometrics* **8:** 411-430.

Shewhart, W. A. 1931. *Economic control of quality of manufacturing product*. New York: Van Nostrand

Shiryaev, A. N. 1963. On optimum methods in quickest detection problems. *Theory of Probability and Applications* **8:** 22-46.

Taguchi, G. 1986. *Introduction to quality engineering-designing quality into products and process.* Asian Productivity Organization, Singapore.

Tsiamyrtzis, P. 2000. A Bayesian approach to quality control problems. Ph. D. Thesis, Faculty of Graduate School , University of Minnesota, U.S.

Woodward, P. W. and Naylor, J.C. 1993. An application of Bayesian methods in SPC. *The Statistician* **42**: 461-469.