**Pak. j. eng. technol. sci.**
Volume 4, No 1, 2014, 1-12
ISSN: 2222-9930 print
ISSN: 2224-2333 online

PJETS

# On Shewhart Control Charts for Zero-Truncated Negative Binomial Distributions

**Anwer Khurshid*, Ashit B. Chakraborty****

Received on: Aug 1, 2013; Accepted on: Oct 2, 2013

**ABSTRACT**

The negative binomial distribution (NBD) is extensively used for the description of data too heterogeneous to be fitted by Poisson distribution. Observed samples, however may be truncated, in the sense that the number of individuals falling into zero class cannot be determined, or the observational apparatus becomes active when at least one event occurs. Chakraborty and Kakoty (1987) and Chakraborty and Singh (1990) have constructed CUSUM and Shewhart charts for zero-truncated Poisson distribution respectively. Recently, Chakraborty and Khurshid (2011 a, b) have constructed CUSUM charts for zero-truncated binomial distribution and doubly truncated binomial distribution respectively. Apparently, very little work has specifically addressed control charts for the NBD (see, for example, Kaminsky et al., 1992; Ma and Zhang, 1995; Hoffman, 2003; Schwertman. 2005).

The purpose of this paper is to construct Shewhart control charts for zero-truncated negative binomial distribution (ZTNBD). Formulae for the Average run length (ARL) of the charts are derived and studied for different values of the parameters of the distribution. OC curves are also drawn.

**Keywords:** control chart, Average Run Length (ARL), zero truncated negative binomial distribution (ZTNBD).

* Department of Mathematical and Physical Sciences, College of Arts and Sciences, University of Nizwa, P. O. Box 33, PC 616, Birkat Al Mouz, Oman
e-mail: anwer@unizwa.edu.om; anwer_khurshid@yahoo.com

** Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India
e-mail: abc_sac@rediffmail.com

## 1. INTRODUCTION

In industries a common monitoring tool is to construct control charts to observe whether a process is in control (Dou and Ping, 2002). A control chart is a statistical scheme devised for the purpose of checking and then monitoring the statistical stability of a process. The standard tool for this purpose is the Shewhart control chart, introduced by Walter A. Shewhart in 1924, elaborated upon it in his book entitled Economic Control of Quality Manufactured Product published in 1931. The advantage of Shewhart charts is its simplicity and is best used to detect large changes in the process, however, are not so sensitive to small changes.

Poisson distribution plays an important role in statistical quality control process through modeling random counts or control of defects per unit. Various types of processes can generate distributions of counts which can be modeled suitably by distributions other than Poisson distribution. Such processes include situations where counts tend to occur in clusters or where the intensity rate of the counts varies randomly over time. The negative binomial distribution (NBD) is a natural and more flexible extension of the Poisson distribution which allows for over-dispersion compared to the Poisson distribution (Hoffman, 2003). The application of NBD has been demonstrated in accident statistics, econometrics, quality control and biometrics. For detailed description, refer to Johnson et al. (2005), Khurshid et al. (2005) and Ryan (2011) among others.

The construction of Shewhart control chart for continuous distributions has been extensively studied in the literature (Mittag and Rinne, 1993; Wadsworth et al. 2002). However, compared with the continuous distributions, there are fewer investigations for discrete distributions. The most widely used for binomial random variable has been developed (see for example, Woodall, 1997; Ryan, 2011; Montgomery, 2013). Unfortunately, the literature on the control charts for the NBD is scanty (Kaminsky et al., 1992; Ma and Zhang, 1995; Xie and Goh, 1997; Hoffman, 2003; and Schwertman, 2005). For a comprehensive overview of Shewhart charts for numerous distributions, see Ryan (2011), Montgomery (2013).

In many cases, however, the entire distribution of counts is not observed. In particular, more often the zeros are not observed or sometimes a large number of zeros are contained in the data. In recent years, researchers have provided new complement models which are obtained by modifying the existing well known models.

These complement models are generally divided in to two categories; zero-truncated and zero-inflated.

Zero-truncated models are the ones when the number of individuals falling into zero class can not be determined, or the observational apparatus becomes active only when at least one event occurs. Chakraborty and Kakoty (1987) and Chakraborty and Bhattacharya (1989, 1991) have constructed CUSUM charts for zero-truncated Poisson distribution, doubly truncated geometric distribution and doubly truncated binomial distribution respectively. Chakraborty and Singh (1990) constructed Shewhart control charts for zero-truncated Poisson distribution where average length and operating characteristic function were obtained. Recently, Chakraborty and Khurshid (2011, 2012) have constructed CUSUM charts for zero-truncated binomial distribution and doubly truncated binomial distribution respectively. Another CUSUM control chart for zero truncated negative binomial distribution has been proposed by Khurshid and Chakraborty (2013).

Accordingly, distributions of negative binomial type often arise in practice where zero group is truncated. The main objective of this paper is to construct Shewhart control charts for zero-truncated negative binomial distribution (ZTNBD). Control charts based on these truncated distribution are studied and Average Run Length (ARL) computed accordingly alongside developing different expressions.

## 2. MATERIALS AND METHODS

### Zero-Truncated Negative Binomial Distribution (ZTNBD)

We consider a negative binomial distribution truncated at $x = 0$. The probability mass function of the ZTNBD is given by (Khurshid and Chakraborty, 2013)

$$f(x;k,p) = \frac{\binom{k+x-1}{x} p^k q^x}{1-p^k} \qquad \textbf{(2.1)}$$

where $x = 1, 2, ..., n.$ Here $f(x; k, p)$ denotes the probability that there are $x$ failures preceding the $k$ -th success in the $(x + k)$ trials. The last trial must be a success, the probability of which is $p$ and in the remaining $(k = x - 1)$ trials we must have $(k - 1)$ success the probability of which is given by binomial probability law by the expression $\frac{(x + k - 1)!}{(k - 1)!\ x!} p^{k-1} q^x$ Therefore, by compound probability theorem, is given by the product of two probabilities i. e.,

$$f(x; k, p) = \frac{(x + k - 1)!}{(k - 1)!\ x!} \frac{p^x}{(1 + p)^{k+x}} = \frac{(x + k - 1)!}{(k - 1)!\ x!} p^k q^x$$

for $x = 0, 1, 2, ..., n.$ More formally, assume a box contains $np$ non-defective items and $nq$ defective items. Items are drawn at random with replacement, the probability that exactly $(x + k)$ trials required to produce $k$ non-defective items is $\frac{(x + k - 1)!}{(k - 1)!\ x!} p^{k-1} q^x$. Thus $k$ and $p$ are the parameters of the negative binomial distribution, where the parameters satisfy $0 < p < 1$ and $k = 1, 2, 3, ....$

The statistical literature reveals that most probability distributions can be parameterized in many different ways; the NBD is no exception. An alternative widely used parameterization of the NBD can be obtained from the expansion of $(Q - P)^{-k}$ where $Q = 1 + P$, $k$ is positive real and $P > 0$ with $P$ not to be in $(0,1).$ Under this parameterization the probability mass function of ZTNBD will be (Zelterman, 2004 and Promislow, 2011)

$$f(x; k, p) = \binom{k + x - 1}{x} \left(1 - Q^{-k}\right)^{-1} \left(\frac{P}{Q}\right)^x \left(1 - \frac{P}{Q}\right)^k \quad \textbf{(2.2)}$$

where $x = 1, 2, ....$ ZTNBD has the mean and variance given by, respectively, (Johnson et al., 2005)

$$E(X) = \frac{kP}{1 - Q^{-k}} \quad \text{and} \quad V(X) = \frac{kPQ}{1 - Q^{-k}} \left[1 - \left(\frac{P}{Q}\right)\right] \left\{\left(1 - Q^{-k}\right)^{-1} - 1\right\}.$$

## 3. SHEWHART CONTROL CHART FOR ZTNBD

### 3.1 Shewhart Control Charts

Let us consider Shewhart $\overline{X}$ chart which contains center line that represents the average value of quality characteristics corresponding to in-control state. There are two horizontal lines namely Lower control limit ($LCL$) and Upper control limit ($UCL$). These control limits are selected so that if process is in control, nearly all sample points will fall between them. Let $W$ be a statistic that measures a quality characteristics of interest with mean $\mu_W$ and the standard deviation $\sigma_W$ then following configuration represents the Shewhart's general model for control chart:

$$LCL = \mu_W - A\sigma_W$$

$$CL = \mu_W$$

$$UCL = \mu_W + A\sigma_W$$

where $A$ is the distance of the control limits from the center line ($CL$), expressed in standard deviation units. It is customary to choose $A = 3$ (Montgomery, 2013) as it covers at least 99.73% of samples which is based on Shewhart's claim that control limits at 3 standard errors are the most economical (Wheeler and Chambers, 2010). Hence the control limits are known as $3\sigma$ limits.

For distribution defined in (2.2) limits for ZTNBD control charts are given by

$$UCL = kP\left(1-Q^{-k}\right)^{-1} + 3\sqrt{kPQ\left(1-Q^{-k}\right)^{-1}\left[1-\left(\frac{P}{Q}\right)\left\{\left(1-Q^{-k}\right)^{-1}-1\right\}\right]}, \qquad \textbf{(3.1)}$$

$$LCL = kP\left(1-Q^{-k}\right)^{-1} - 3\sqrt{kPQ\left(1-Q^{-k}\right)^{-1}\left[1-\left(\frac{P}{Q}\right)\left\{\left(1-Q^{-k}\right)^{-1}-1\right\}\right]},$$

and $CL = kP\left(1-Q^{-k}\right)^{-1}$.

## 3.2 Average Run Length for Shewhart Chart

The most frequently used statistical characteristic of a control chart is its Average Run Length (ARL). ARL is the average number of points that must be plotted before a point indicates an out of control condition. For any Shewhart control chart, the ARL is $ARL = [p]^{-1}$ where $p$ is the probability that a single point exceeds the control limits. Now, if the mean shifts from in control value say $\theta_0$ to another value $\theta_1 = \theta_0 + k\sigma$, the probability of not detecting this shift on the first subsequent sample or the $\beta$ risk (Montgomery, 2013) is

$$\beta = P[X < UCL|\theta] - P[X \leq LCL|\theta].$$

Thus, for ZTNBD, we have

$$\beta = \sum_{x=1}^{UCL} \binom{k+x-1}{x} (1-Q^{-k})^{-1} \left(\frac{P}{Q}\right)^{x} \left(1-\frac{P}{Q}\right)^{k} - \sum_{x=1}^{LCL} \binom{k+x-1}{x} (1-Q^{-k})^{-1} \left(\frac{P}{Q}\right)^{x} \left(1-\frac{P}{Q}\right)^{k} \quad \textbf{(3.2)}$$

hence

$$ARL = \left[1 - \sum_{x=1}^{UCL} \binom{k+x-1}{x} (1-Q^{-k})^{-1} \left(\frac{P}{Q}\right)^{x} \left(1-\frac{P}{Q}\right)^{k} - \sum_{x=1}^{LCL} \binom{k+x-1}{x} (1-Q^{-k})^{-1} \left(\frac{P}{Q}\right)^{x} \left(1-\frac{P}{Q}\right)^{k}\right]^{-1}. \quad \textbf{(3.3)}$$

Operating Characteristic (OC) curve for the Shewhart control chart when the underlying distribution is ZTNBD can be constructed by plotting the $\beta$ risk against the magnitude of the shift of the process parameter that we wish to detect.

## 4. COMPUTATIONS AND CONCLUSIONS

The calculated values of $\beta$ risk (the probability of not detecting the shift on the first subsequent sample; Montgomery, 2013) and the corresponding values of ARL are shown in Tables 1.1 and 1.2. The effect of the parameter $Q$ on the control limits for different values of $k$ are shown in Tables 1.3 and 1.4.

It is evident from the Tables 1.1 and 1.2 that the values of ARL for fixed control limits $\mu \pm k\sigma; (k = 1,2,3)$ will go on decreasing as we go on increasing the values of $P$. But for fixed shifting of the parameter $P,$ the values of ARL increase as we increase the size of the control limits.

It has also been observed that as the values of $k$ increase (from $k = 2$ to $k = 3$), the values of ARL decrease for fixed control limits. The effects of $k$ and on control limits can be understood from the Tables 1.3 and 1.4. It has been observed from the tables that for fixed , the range of the control limits decrease as there is an increase in the values of , but for fixed and , the range of the control limits increase as we increase the value of in .
It has also been observed from the Tables 1.3 and 1.4, that the range of the control limits decrease as we increase in the values of (from to ) for fixed .

Operating Characteristic (OC) curve for the Shewhart control chart when the underlying distribution is ZTNBD can be constructed by plotting the risk (Tables 1.1 and 1.2) against the magnitude of the shift of the process parameter that we wish to detect.

A comparative study about the behavior of the OC curves under different control limits will be more apparent from Figures 1 and 2. It has been observed from the figures that OC curves deviate from the origin as the control limits decrease. It also shows value of has an effect on OC, ultimately, it will have an effect on producer’s and consumer’s risk.

## References

[1] Chakraborty, A. B. and Bhattacharya, S. K. (1989). CUSUM control charts for doubly truncated geometric and Poisson distributions. pp. 509-512. *Proceedings of Quality for Progress and Development, Asian Congress on Quality and Reliability*. Wiley Eastern Limited.

2] Chakraborty, A. B. and Bhattacharya, S. K. (1991). Cumulative sum control chart for a doubly truncated binomial distribution. *Egyptian Statistical Journal*, 35, 119-124.

[3] Chakraborty, A. B. and Kakoty, S. K. (1987). Control charts for ZTPD. *Journal of Indian Association for Productivity, Quality and Reliability*, 12, 17-25.

[4] Chakraborty, A. B. and Khurshid, A. (2011). One-sided cumulative sum (CUSUM) control charts for the zero-truncated binomial distribution. *Economic Quality Control*, 26, 41-51.

[5] Chakraborty, A. B. and Khurshid, A. (2012). Control charts for doubly-truncated binomial distributions. *Economic Quality Control*, 27, 187-194.

[6] Chakraborty, A. B. and Singh,. B. P. (1990). Shewhart control chart for ZTPD. pp. 18-24. *Proceedings of National Seminar on Quality and Reliability NIQR, Trivandrum, India.*

[7] Dou, Y. and Ping, S. (2002). One-sided control charts for the mean of positively skewed distributions. *Total Quality Management*, 13, 1021-1033.

[8] Hoffman, D. (2003). Negative binomial control limits for count data with extra-Poisson variation. *Pharmaceutical Statistics*, 2, 127-132.

[9] Johnson, N. L., Kemp, A. W. and Kotz, S. (2005). *Univariate Discrete Distributions*, Third Edition. John Wiley, New York.

[10] Kaminsky, F. C., Benneyan, J. C., Davis, R. D. and Burke, R. J. (1992). Statistical control charts based on a geometric distribution. *Journal of Quality Technology*, 24, 63-69.

[11] Khurshid, A. and Chakraborty, A. B. (2013). CUSUM control charts for zero-truncated negative binomial and geometric distributions. *Revista Investigación Operacional*, 34, 195-204.

[12] ]Khurshid, A., Ageel, M. I. and Lodhi, R. A. (2005). On confidence intervals for the negative binomial distribution. *Revista Investigación Operacional*, 26, 59-70.

[13] Ma, Y. and Zhang, Y. (1995). Q control charts for negative binomial distribution. *Computers and Industrial Engineering*, 31, 813-816.

[14] Montgomery, D. C. (2013). *Introduction to Statistical Quality Control*, Seventh Edition. John Wiley, New York.

[15] Mittag, H. J. and Rinne, H. (1993). *Statistical Methods of Quality Assurance*. Chapman and Hall, New York.

[16] Promislow, S. D. (2011). *Fundamentals of Actuarial Mathematics*, Second Edition. John Wiley, New York.

[17] Ryan, T. P. (2011). *Statistical Methods for Quality Improvement*, Third Edition. John Wiley, New York.

[18] Schwertman, N. C. (2005). Designing accurate control charts based on the geometric and negative binomial distribution. *Quality and Reliability Engineering International*, 21, 743-756.

[19] Xie, M. and Goh, T. N. (1997). The use of probability limits for process control based on geometric distribution. *International Journal of Quality and Reliability and Management*, 16, 64-73.

[20] Wadsworth, H. M., Stephens, K. S. and Godfrey, A. B. (2002). *Modern Methods for Quality Control and Improvement*, Second Edition. John Wiley, India Pvt. Ltd.

[21] Wheeler, D. J. and Chambers, D. S. (2010). *Understanding Statistical Process Control*, Third Edition. SPC Press, Knoxville, TN.

[22] Woodall, W. H. (1997). Control Charting Based on Attribute Data: Bibliography and Review. *Journal of Quality Technology*, 29, 172-183.

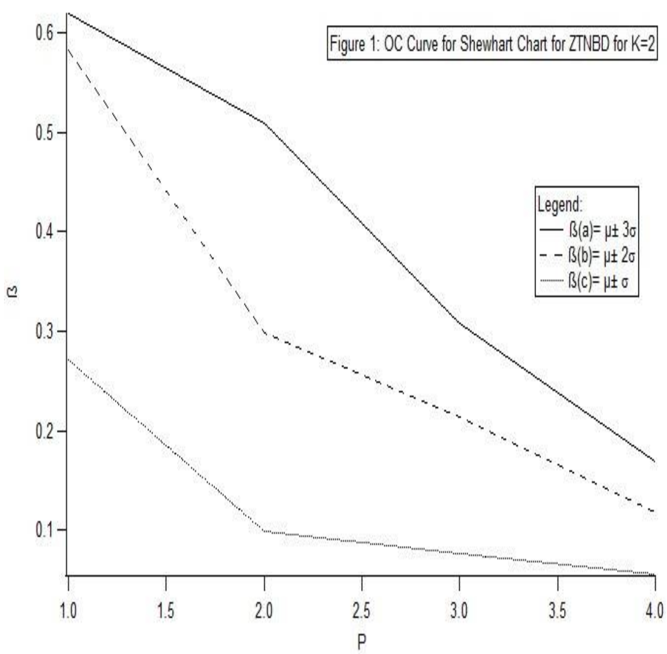
[23] Zelterman, D. (2004). *Discrete Distributions: Applications in the Health Sciences*. John Wiley, New York.

**Table 1.1: Some numerical values of** $\beta$ and $ARL$ for selected values of $P$ when $k = 2$

**Table 1.2: Some numerical values of** $\beta$ and $ARL$ for selected values of $P$ when $k = 3$

**Table 1.3: Some numerical values of** $\beta$ and $ARL$ for selected values of $Q$ when $k = 2$

**Table 1.4: Some numerical values of $\beta$ and $ARL$ for selected values of $Q$ when $k = 3$**



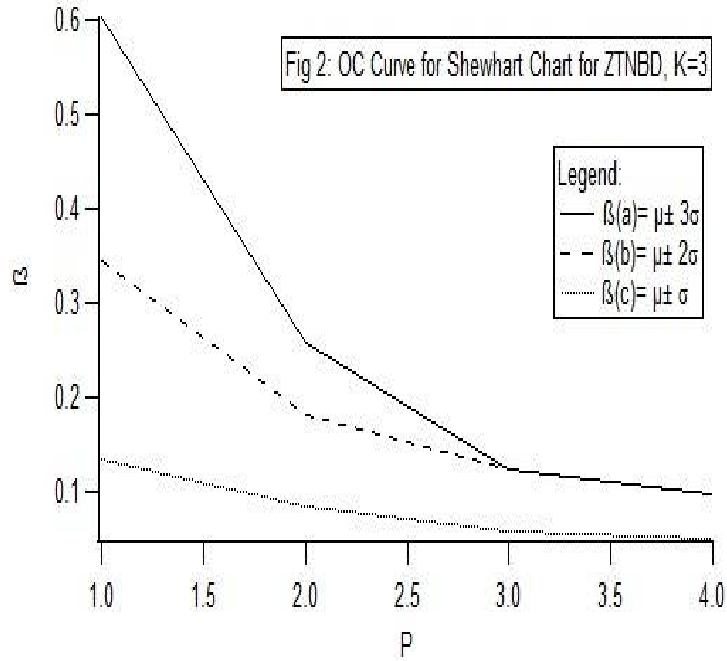


Fig 2: OC Curve for Shewhart Chart for ZTNBD, K=3
Legend:
ß(a)= μ± 3σ
ß(b)= μ± 2σ
ß(c)= μ± σ
ß
0.6
0.5
0.4
0.3
0.2
0.1
1.0
1.5
2.0
2.5
3.0
3.5
4.0
P