

Effect of Inspection Error on CUSUM Chart for Truncated Negative Binomial Distribution

Abstract

In this paper, a mathematical investigation has been done for Cumulative Sum control Chart (CSCC) of Truncated Negative Binomial Distribution (TNBD) under inspection error. Average Run Length (ARL), d and ϕ is calculated for different values of parameter of TNBD.

Keywords: TNBD, CSCC, ARL, Inspection Error.

1. Introduction:

In statistics, truncated distributions are conditional distributions that effect from confining the area of additional probability distribution. Truncated distributions occur in realistic statistics in cases where the capability to record or even to know about, occurrences is restricted to values which lie within a specified range. The Poisson is often the standard distribution regarded as modeling random counts. As such, for analyzing control limits for processes with discrete responses, usual procedures are based on the Poisson distribution. Poisson distribution has the distinguished property that the mean and variance of the distribution is equal. Though, for several processes, the Poisson distribution offers an unsatisfactory model. Different kinds of procedures can make distributions of counts which are not effectively modeled by the Poisson distribution. Such processes comprise conditions where the intensity rates of the counts differ at random over time. Negative binomial distribution is a typical expansion of the Poisson distribution and authorizes for over-dispersion relative to the Poisson. The negative binomial distribution can be resultant of numerous models. The negative binomial is also resulting as a mixture of Poisson distributions however, usually resulting as a generalization of the geometric distribution. Applications of the negative binomial distribution are extensive. The negative binomial distribution is extensively used for the description of data excessively assorted to be fitted by a Poisson distribution. However, observed samples may be truncated, in the sense that the number of individuals falling into the zero class cannot be determined. Samford (1955) have illustrated the use of truncated negative binomial distribution over Poisson distribution by considering the method of moments. Hoffman (2003) considered the control limits of negative binomial distribution for count data with discrete responses based on Poisson distribution. Chakraborty et al. (2017) investigated the effect of measurement error on the power and ARL of control chart for ZTNBD based on standardized normal variates.

The CUSUM chart is exercises to examine the mean of a process based on samples taken from the process at given times. The measurements of the samples in a given time comprises a subgroup rather examining the mean of every subgroup independently, the CUSUM chart illustrates the accumulated information of existing and earlier samples. This is the reason why CUSUM chart is usually better than the \bar{X} –chart for detecting small shifts in the mean of a process. The CUSUM charts rely on the requirement of a target value and a known or consistent estimate of the standard deviation. This is the reason; the CUSUM chart is better used past the process control has been recognized. The CUSUM chart usually signals an out-of-control process by a growing or sliding drift of the cumulative sum until it crosses the boundary. An assignable cause is suspected whenever the CUSUM chart indicates an out-of-control process. Sankle et al. (2012) presented CSCC for truncated normal distribution considering the effect of inspection error. Sayyed and Singh (2015) considered CSCC for binomial parameters under the effect of inspection error where the underlying distribution is Poisson. Singh and Mishra (2017) have considered the effect of inspection error on singly truncated Binomial distribution.

In this paper, we have seen the effect of inspection error on CSCC for TNBD. ARL, d , and ϕ is calculated for different values of parameter of TNBD under inspection error.

2. Truncated Negative Binomial Distribution (TNBD) under Inspection Error:

The negative binomial distribution has the following probability density function (pdf):

$$f(x; r, p) = \begin{cases} \binom{x+r-1}{r-1} p^r q^x & ; x = 0, 1, 2, 3 \dots \\ 0, & otherwise, \end{cases} \quad (1)$$

where r and p are the parameters of Negative Binomial distribution which satisfy $0 < p < 1$ and $r = 0, 1, 2, 3 \dots$

If $p = \frac{1}{Q}$ and $q = \frac{P}{Q}$ then we have $Q - P = 1$, then the equation (1) will be written as:

$$f(x; r, p) = \begin{cases} \binom{x+r-1}{r-1} \left(\frac{P}{Q}\right)^x \left(1 - \frac{P}{Q}\right)^r & ; x = 0, 1, 2, 3 \dots \\ 0, & otherwise. \end{cases} \quad (2)$$

To obtain the corresponding probabilities for the truncated distribution the equation (2) must be divided by $(1 - f(x = 0; r, p))$ where $f(x = 0; r, p) = \left(1 - \frac{P}{Q}\right)^r$, it follows that:

$$f(x; r, p) = \begin{cases} \binom{x+r-1}{r-1} \left(\frac{1}{1-Q^{-r}}\right) \left(\frac{P}{Q}\right)^x \left(1 - \frac{P}{Q}\right)^r & ; x = 1, 2, 3 \dots \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The mean and variance of the above distribution are as follows:

$$\left. \begin{aligned} \text{Mean} &= E(X) = rP(1 - Q^{-r})^{-1} \\ \text{Variance} &= V(X) = \frac{rPQ}{(1-Q^{-r})} \left[1 - \frac{rP}{Q}((1 - Q^{-r}) - 1)\right] \end{aligned} \right\} . \quad (4)$$

Moreover, inspection of attributes are differentiated by two decision variables in which every item is analyzed and classified as good or faulty. Two types of error are possible, an item which is good but categorized as faulty, e_1 , or an item which is faulty but categorized as good, e_2 . If p is true fraction defective and p' is the apparent fraction defective, then we set:

$$p' = p(1 - e_2) + (1 - p)e_1, \quad (5)$$

with both e_1 and e_2 estimated.

3. CSCC for TNBD under Inspection Error

Let x_i ($i = 1, 2, 3, \dots$) are the independent random variables with the distribution given by equation (3). The sequential probability ratio test (SPRT) distinguishes between the hypotheses $H_0: P' = P'_0$ and $H_1: P' = P'_1 (> P'_0)$ has the value:

$$\frac{f(x_1, x_2, \dots, x_n | P'_1 r)}{f(x_1, x_2, \dots, x_n | P'_0 r)} = \prod_{i=1}^n \left(\frac{1 - (Q'_0)^{-r}}{1 - (Q'_1)^{-r}}\right)^{-n} (1 - (Q')^{-r})^{-1} \left(\frac{P'_1 Q'_0}{P'_0 Q'_1}\right)^{\sum x_i} \left(\frac{Q'_0(Q'_1 - P'_1)}{Q'_1(Q'_0 - P'_0)}\right)^{nr}. \quad (6)$$

The SPRT has the continuation region is then

$$\left(\frac{\beta}{1-\alpha}\right) < \prod_{i=1}^n \left(\frac{1 - (Q'_0)^{-r}}{1 - (Q'_1)^{-r}}\right)^{-n} (1 - (Q')^{-r})^{-1} \left(\frac{P'_1 Q'_0}{P'_0 Q'_1}\right)^{\sum x_i} \left(\frac{Q'_0(Q'_1 - P'_1)}{Q'_1(Q'_0 - P'_0)}\right)^{nr} < \left(\frac{1-\beta}{\alpha}\right), \quad (7)$$

that is,

$$\log\left(\frac{\beta}{1-\alpha}\right) < n \log\left(\frac{1 - (Q'_0)^{-r}}{1 - (Q'_1)^{-r}}\right) + r \log(1 - Q') + \sum_{i=1}^n x_i \log\left(\frac{P'_1 Q'_0}{P'_0 Q'_1}\right) + nr \log\left(\frac{Q'_0(Q'_1 - P'_1)}{Q'_1(Q'_0 - P'_0)}\right) < \log\left(\frac{1-\beta}{\alpha}\right). \quad (8)$$

Now consider the right hand side of the equation (8), we have:

$$n \log \left(\frac{1-(Q'_0)^{-r}}{1-(Q'_1)^{-r}} \right) + r \log(1-Q') + \sum_{i=1}^n x_i \log \left(\frac{P'_1 Q'_0}{P'_0 Q'_1} \right) + nr \log \left(\frac{Q'_0(Q'_1-P'_1)}{Q'_1(Q'_0-P'_0)} \right) < \log \left(\frac{1-\beta}{\alpha} \right), \quad (9)$$

$$\sum_{i=1}^n x_i \log \left(\frac{P'_1 Q'_0}{P'_0 Q'_1} \right) < \log \left(\frac{1-\beta}{\alpha} \right) - n \log \left(\frac{1-(Q'_0)^{-r}}{1-(Q'_1)^{-r}} \right) - nr \log \left(\frac{Q'_0(Q'_1-P'_1)}{Q'_1(Q'_0-P'_0)} \right) - r \log(1-Q'). \quad (10)$$

If we plot points $(n, \sum_{i=1}^n x_i)$ then the boundary line between the continuation region and the acceptance region for H_1 has the equation:

$$\sum_{i=1}^n x_i < \frac{\log \left(\frac{1-\beta}{\alpha} \right) - n \log \left(\frac{1-(Q'_0)^{-r}}{1-(Q'_1)^{-r}} \right) - nr \log \left(\frac{Q'_0(Q'_1-P'_1)}{Q'_1(Q'_0-P'_0)} \right) - r \log(1-Q')}{\log \left(\frac{P'_1 Q'_0}{P'_0 Q'_1} \right)}, \quad (11)$$

or using the approximation of $\beta = 0$, we have:

$$\sum_{i=1}^n x_i < \frac{-\log \alpha - n \log \left(\frac{1-(Q'_0)^{-r}}{1-(Q'_1)^{-r}} \right) - nr \log \left(\frac{Q'_0(Q'_1-P'_1)}{Q'_1(Q'_0-P'_0)} \right) - r \log(1-Q')}{\log(P'_1 Q'_0) - \log(P'_0 Q'_1)}. \quad (12)$$

To construct the CUSUM chart and the V-Mask, the lead distance d and the angle ϕ of the mask is given by:

$$d = \frac{-\log \alpha}{\left[\log \left(\frac{1-(Q'_1)^{-r}}{1-(Q'_0)^{-r}} \right) + r \log \left(\frac{Q'_1(Q'_0-P'_0)}{Q'_0(Q'_1-P'_1)} \right) \right]}, \quad (13)$$

and

$$\phi = \tan^{-1} \left[\frac{\log \left(\frac{1-(Q'_1)^{-r}}{1-(Q'_0)^{-r}} \right) + r \log \left(\frac{Q'_1(Q'_0-P'_0)}{Q'_0(Q'_1-P'_1)} \right)}{\log(P'_1 Q'_0) - \log(P'_0 Q'_1)} \right]. \quad (14)$$

The approximate formula for the average run length (ARL) for detecting a shift for the parameter P from P'_0 to P'_1 for known r is given by,

$$ARL = \frac{-\log \alpha}{\left[\log \left(\frac{1-(Q'_0)^{-r}}{1-(Q'_1)^{-r}} \right) + \left(\frac{r P'_1}{1-(Q'_1)^{-r}} \right) \log \left(\frac{P'_1 Q'_0}{P'_0 Q'_1} \right) + r \log \left(\frac{Q'_0(Q'_1-P'_1)}{Q'_1(Q'_0-P'_0)} \right) \right]}. \quad (15)$$

Values of the d , ϕ and ARL of CUSUM chart are determined for a different combinations of the values of P , r and α for controlling the parameters P when r is known. These values are shown in Table 1.1 to Table 1.6.

4. Illustration and Conclusion:

For the purpose of numerical illustration, we have considered five cases: $(e_1, e_2) = (0,0)$, $(0.03,0.3)$, $(0.03,0.1)$, $(0.01, 0.05)$, $(0.005,0.02)$. Table- 1.1 shows that values of the lead distance d and ARL decreases as the value of α and k increases. In Table- 1.2 to Table- 1.5 we have taken different combinations of error rates. It is seen that the value of d and ARL decreases as the value of α and k increases. But if we compare error free case given in Table-1.1 to other error rates shown in Table 1.2 to Table- 1.5, it is seen that for $k=1$ and $\alpha = 0.05$, the value of d and ARL is 4.31 and 17.54 respectively when $(e_1, e_2) = (0,0)$ and for the same value of k and α , the values of $d= 4.41, 4.53, 4.64, 5.66$ and $ARL= 18.22, 19.10, 20.69, 26.55$ when $(e_1, e_2) = (0.005, 0.02), (0.03,0.3), (0.03,0.1), (0.01, 0.05)$ respectively. The above discussion shows that as the error rate increases the value of d and ARL increases, which is a good symptom to detect assignable causes less often in production process.

Table- 1.6 shows that the angle of mask, ϕ increases as the value of k increases. For comparison, if we take $(e_1, e_2) = (0,0), (0.005, 0.02), (0.03,0.3), (0.01, 0.05)$, the angle of mask is 67.50, 67.76, 68.24, and 72.96 respectively, which is the increased value of the mask, whereas $(e_1, e_2) = (0.03,0.1)$ the angle of mask is 67.45, lesser as compared to error free case.

Table- 1.1: Value of d and ARL when $(e_1, e_2) = (0, 0)$.

(P'_0, P'_1)		α														
		0.05			0.025			0.01			0.005			0.001		
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
(1, 2)	d →	4.31	3.04	2.27	5.31	3.75	2.80	6.62	4.68	3.50	7.62	5.38	4.02	9.94	7.03	5.25
	ARL →	17.54	9.49	6.20	21.60	11.69	7.64	26.96	14.60	9.53	31.03	16.80	10.97	40.45	21.90	14.30
(1, 3)	d →	2.72	1.86	1.36	3.35	2.29	1.67	4.19	2.86	2.09	4.82	3.29	2.41	6.28	4.29	3.14
	ARL →	5.72	3.04	1.98	7.05	3.74	2.44	10.12	4.67	3.05	10.13	5.37	3.50	13.20	7.01	4.57
(1, 4)	d →	2.16	1.44	1.04	2.66	1.77	1.28	3.32	2.21	1.60	3.82	2.54	1.84	4.98	3.32	2.40
	ARL →	3.11	1.63	1.06	3.83	2.01	1.35	4.77	2.50	1.63	5.49	2.88	1.88	7.16	3.76	2.46
(1,5)	d →	1.85	1.21	0.87	2.28	1.49	1.07	2.85	1.86	1.33	3.28	2.14	1.54	4.28	2.80	2.01
	ARL →	2.04	1.06	0.69	2.51	1.31	0.85	3.13	1.63	1.06	3.60	1.87	1.23	4.70	2.45	1.60

Table- 1.2: Value of d and ARL when $(e_1, e_2) = (0.01, 0.05)$

(P'_0, P'_1)		α														
		0.05			0.025			0.01			0.005			0.001		
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
(1, 2)	d →	4.53	3.17	2.35	5.59	3.91	2.90	6.97	4.87	3.62	8.02	5.61	4.17	10.47	7.32	5.43
	ARL →	19.10	10.30	6.72	23.52	12.68	8.28	29.36	15.83	10.33	3.78	18.22	4.88	44.04	23.75	15.50
(1, 3)	d →	2.86	1.93	1.41	3.52	2.38	1.74	4.40	2.98	2.17	5.06	3.42	2.49	6.61	4.47	3.25
	ARL →	6.28	3.32	2.17	7.74	4.09	2.67	9.66	5.11	3.34	11.11	5.88	3.84	14.49	7.67	5.01
(1, 4)	d →	2.27	1.50	1.08	2.79	1.85	1.33	3.49	2.31	1.66	4.01	2.65	1.91	5.24	3.46	2.49
	ARL →	3.44	1.80	1.17	4.24	2.22	1.45	5.29	2.77	1.81	6.09	3.19	2.08	7.94	4.16	2.72
(1,5)	d →	1.95	1.27	0.90	2.40	1.56	1.12	3.00	1.95	1.39	3.45	2.24	1.61	4.50	2.93	2.09
	ARL →	8.28	1.18	0.77	2.81	1.46	0.95	3.55	1.82	1.19	4.04	2.10	1.37	5.26	2.73	1.79

Table- 1.3: Value of d and ARL when $(e_1, e_2) = (0.03, 0.1)$

(P'_0, P'_1)		α														
		0.05			0.025			0.01			0.005			0.001		
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
(1, 2)	$d \rightarrow$	4.64	3.28	2.45	5.72	4.04	3.02	7.14	5.05	3.77	8.21	5.81	4.34	10.71	7.57	5.66
	ARL \rightarrow	20.69	11.22	7.33	25.48	13.81	9.02	31.82	17.24	11.26	36.60	19.84	12.96	47.73	25.87	16.90
(1, 3)	$d \rightarrow$	2.95	2.02	1.48	3.64	2.49	1.83	4.54	3.11	2.28	5.23	3.58	2.63	6.81	4.67	3.42
	ARL \rightarrow	6.97	3.71	2.42	8.58	4.57	2.98	10.71	5.70	33.72	12.33	6.56	4.28	16.07	8.55	5.58
(1, 4)	$d \rightarrow$	2.36	1.58	1.14	2.90	1.94	1.41	3.62	2.43	1.76	4.17	2.80	2.03	5.44	3.65	2.64
	ARL \rightarrow	3.90	2.05	1.34	4.80	2.53	1.65	6.00	3.15	2.06	6.90	3.63	2.37	9.00	4.73	3.09
(1,5)	$d \rightarrow$	2.04	1.34	0.97	2.51	1.66	1.19	3.13	2.07	1.49	3.61	2.38	1.71	4.70	3.11	2.23
	ARL \rightarrow	2.64	1.38	0.90	3.25	1.70	1.11	4.06	2.12	1.38	4.67	2.44	1.59	6.09	3.18	2.08

Table- 1.4: Value of d and ARL when $(e_1, e_2) = (0.03, 0.30)$

(P'_0, P'_1)		α														
		0.05			0.025			0.01			0.005			0.001		
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
(1, 2)	$d \rightarrow$	5.66	3.70	2.64	6.98	4.56	3.25	8.71	5.69	4.06	10.02	6.54	4.68	13.07	8.53	6.10
	ARL \rightarrow	26.55	13.96	9.11	32.69	17.20	11.22	40.81	21.47	14.01	46.95	24.70	16.12	61.22	32.21	21.01
(1, 3)	$d \rightarrow$	3.53	2.24	1.58	4.34	2.77	1.95	5.42	3.45	2.44	6.24	3.97	2.81	8.14	5.18	3.66
	ARL \rightarrow	8.88	4.62	3.02	10.94	5.69	3.72	13.66	7.10	4.64	15.71	8.17	5.34	20.49	10.65	6.96
(1, 4)	$d \rightarrow$	2.78	1.74	1.22	3.43	2.15	1.50	4.28	2.68	1.88	4.92	3.08	2.16	6.42	4.02	2.82
	ARL \rightarrow	4.97	2.57	1.68	6.12	3.16	2.07	7.65	3.95	2.58	8.80	4.54	2.97	11.47	5.92	3.88
(1,5)	$d \rightarrow$	2.39	1.48	1.03	2.94	1.82	1.27	3.67	2.27	1.59	4.23	2.62	1.82	5.51	3.41	2.38
	ARL \rightarrow	3.38	1.73	1.14	4.16	2.13	1.40	5.19	2.67	1.75	5.97	3.07	2.01	7.79	4.00	2.62

Table- 1.5: Value of d and ARL when $(e_1, e_2) = (0.005, 0.02)$

(P'_0, P'_1)	$d \rightarrow$ ARL \rightarrow	α														
		0.05			0.025			0.01			0.005			0.001		
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
(1, 2)		4.41	3.10	2.31	5.43	3.82	2.85	6.78	4.77	3.56	7.80	5.49	4.09	10.16	7.15	5.33
		18.22	9.85	6.43	22.44	12.13	7.92	28.01	15.15	9.89	32.23	17.43	11.38	42.02	22.72	14.84
(1, 3)		2.79	1.90	1.39	3.43	2.33	1.71	4.28	2.91	2.13	4.93	3.35	2.45	6.42	4.37	3.20
		5.97	3.17	2.07	7.35	3.90	2.55	9.18	4.87	3.18	10.56	5.60	3.66	13.77	7.31	4.77
(1, 4)		2.21	1.47	1.06	2.72	1.81	1.31	3.40	2.26	1.63	3.91	2.60	1.88	5.09	3.39	2.45
		3.26	1.71	1.12	4.01	2.10	1.38	5.01	2.63	1.72	5.76	3.02	1.97	7.51	3.94	2.57
(1,5)		1.90	1.24	0.89	2.34	1.53	1.09	2.92	1.91	1.37	3.36	2.19	1.57	4.38	2.86	2.05
		2.15	1.12	0.73	2.65	1.38	0.90	3.31	1.72	1.13	3.80	1.98	1.30	4.96	2.58	1.69

Table- 1.6: Values of ϕ for different values of (e_1, e_2) :

(P'_0, P'_1) \rightarrow (e_1, e_2) \downarrow	(1, 2)			(1, 3)			(1, 4)			(1, 5)		
	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
(0, 0)	67.50	73.70	77.68	69.78	75.90	79.58	71.31	77.30	80.75	72.45	78.31	81.57
(0.01, 0.05)	68.24	74.42	78.31	70.40	76.47	80.06	71.85	77.77	81.14	72.92	78.71	81.89
(0.03, 0.1)	67.45	73.64	77.63	69.54	75.67	79.39	70.92	76.95	80.46	71.94	77.86	81.21
(0.03, 0.30)	72.96	78.70	81.88	74.61	80.07	82.96	75.69	80.94	83.63	76.48	81.55	84.09
(0.005, 0.02)	67.76	73.96	77.91	69.99	76.10	79.75	71.49	77.46	80.88	72.60	78.44	81.67

References

Chakraborty, A. B., Khurshid, A. and Acharjee, R. (2017). Measurement error effect on the power of control chart for Zero-Truncated Negative Binomial Distribution, *Yugoslav Journal of Operations Research*, DOI: 10.2298/YJOR161028002C.

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

Sankle, R., Singh, J.R. and Mangal, I.K. (2012). Cumulative sum control charts for truncated normal distribution under measurement error. *Statistics in Transition (New Series)*, 13, 95-106.

Samford, M. R. (1955). The truncated negative binomial distribution. *Biometrika*, 42, 58-69.

Mujahida Sayyed and Singh, J. R. (2015). CSCC Chart for Binomial Parameters under Inspection Error, *Revista Investigacion Operacional*, 36 (2), 127-132.

Singh, J. R. and Mishra, U. (2017). Power of Control Chart for Singly Truncated Binomial Distribution under Inspection Error, *Global and Stochastic Analysis*, 25th International Conference of Forum for Interdisciplinary Mathematics.