Songklanakarin J. Sci. Technol.
44 (4), 1075–1085, Jul. – Aug. 2022

http://www.sjst.psu.ac.th

Original Article

Multi-objective optimal design of multiple dependent state sampling plan for over-dispersed data under the condition on a new zero-inflated distribution

Wimonmas Bamrungsetthapong, and Sirinapa Aryuyuen*

Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Pathum Thani, Khlong Luang, 12110 Thailand

Received: 19 February 2022; Revised: 5 May 2022; Accepted: 23 May 2022

Abstract

A sampling plan can help to determine the quality of products, monitor the goodness of materials, and validate whether the yields are free from defects. When the manufacturing process is precisely aligned, defects are minimized during sampling inspection. This study proposed a multiple dependent state (MDS) sampling plan under a zero-inflated Poisson quasi-Lindley (ZIPQL) distribution, denoted by MDS_{ZIPQL} to count zero-inflated data. A genetic algorithm with multi-objective optimization was used to estimate the optimal plan parameters to maximize the probability of accepting a lot (P_a) and minimize the total cost of inspection (TC) and the average sample number (ASN) simultaneously. A sensitivity analysis of the required sample size assessed the performance of the proposed MDS_{ZIPQL} as numerical examples compared to the MDS plan under a zero-inflated Poisson (MDS_{ZIP}) distribution. Simulation study results found that the required sample sizes and ASN of the MDS_{ZIPQL} plan were less than the MDS_{ZIP} plan, indicating that the MDS_{ZIPQL} plan performed better than the MDS_{ZIP} plan regarding the required sample size and ASN. Two real data sets were illustrated under the proposed MDS_{ZIPQL} plan and compared to the MDS_{ZIP} plan. Results showed that the MDS_{ZIPQL} plan had a smaller number of required sample sizes, ASN value and TC value than the MDS_{ZIP} plan (or maximum value of P_a). Therefore, the proposed MDS_{ZIPQL} plan was more efficient than the existing MDS_{ZIP} plan.

Keywords: multiple dependent state sampling plan, multi-objective optimization, zero-inflated distribution, zero-inflated Poisson quasi-Lindley distribution, over-dispersion

1. Introduction

Count data are often encountered in real-world applications. Discrete distributions play an important role in count data analysis in many research fields such as actuarial, environmental, actuarial, engineering, and economic sciences. Many discrete distributions have been developed for count data analysis (Li *et al.*, 2011). In practice, the Poisson distribution is usually used to fit count data. One important property of the Poisson distribution is that its mean and variance are equal (equidispersion). However, this property is often violated in real-life count data that is overdispersed when the variance is greater than the mean. For example, observed data on the number of claims often exhibit a variance that noticeably exceeds their mean, and the Poisson distribution is not appropriate in such a case. Copious researchers have developed distributions for counting data with many zeros. Mixed Poisson distributions have been applied to count data analyses with overdispersions, such as the Poisson-gamma distribution (Hilbe, 2011) and Poisson-Lindley (PL) distribution (Sankaran, 1970). In 2017, Grine and Zeghdoudi proposed a mixed Poisson distribution called the Poisson quasi-Lindley (PQL) distribution. The PQL model was fitted to several real data sets of the number of errors per copying group compared to the Poisson and PL distributions. Results showed that the PQL distribution gave a better fit than the Poisson and PL distributions. Therefore, the PQL distribution can be considered a good alternative for modeling

*Corresponding author
Email address: sirinapa_a@rmutt.ac.th

count data by providing more accurate estimates with a better fit (Grine, & Zeghdoudi, 2017).

Overdispersion in data can be caused by several factors depending on the situation, one of which is excess zeros in the data. The Poisson distribution cannot adequately accommodate the number of zeros in the sample. Zero-inflated distributions are introduced to analyze count data with excess zeros (Wagner, Robertson, & Harris, 2011). Some researchers tried to modify the Poisson distribution to combine excess zeros as the zero-inflated Poisson (ZIP) distribution (Lambert, 1992).

The ZIP distribution is often considered a probability distribution for overdispersed data consisting of many zeros (Lambert, 1992), and is applied in various fields including public health, economics, and production control. (Shalini & Abdullah, 2018). Ridout, Hinde, and Demétrio (2001) and Fang (2013) evaluated the performance between the ZIP model and the zero-inflated negative binomial (ZINB) model for count data with overdispersion. They concluded that the ZINB distribution was more flexible than the ZIP distribution.

A multiple dependent state (MDS) sampling plan decides to accept or reject the current lot based on the acceptance of the previous lot. This plan also requires continuous samples from both the current and previous lots. As a result, the sample size and cost of inspection are reduced. In 1976, Wortham and Baker were the first to introduce this plan (Wortham, & Baker, 1976). Later, Balamurali and Jun (2007) suggested an MDS sampling plan for variable data. They observed that the MDS sampling plan was more effective than the single sampling plan (SSP) because it provided the desired protection with a smaller required sample size. As a result, it was reasonable to conclude that the MDS plan gave a better performance. An MDS sampling plan for the ZIP distribution was created by Wang and Hailemariam (2018). The resubmitted SSP, repetitive group sampling and quick switching system have all been compared to the MDS sampling plan. The MDS sampling plan employs smaller required sample sizes and average sample numbers, and results show that it is more efficient than the other sampling plans. Many researchers proposed MDS sampling plans in different situations. For example, Aslam, Nazir, and Jun (2015) designed an attribute control chart based on the MDS sampling plan, while Nadi and Gildeh (2019) proposed a Weibull distribution-based group MDS sampling plan. They discovered that the actual mean lifetime of products was longer than the specified mean lifetime. Srinivasa, Rosaiah, and Naidu (2020) developed an MDS sampling plan for a time truncated life under exponentiated half logistic distribution, while Aslam *et al.*, (2021) developed a modified version of the MDS sampling plan. They claimed that this was more flexible and efficient than the existing MDS plan in terms of sample size and inspection cost over a time truncated life.

Increased technological developments showed that when the production process was well inspected, the number of defective items in the samples or products reduced to almost zero. Manufacturers can improve the acceptability of their sampling plans by decreasing the required sample size, as a larger required sample size increases inspection costs. Recently, many researchers have developed different types of acceptance sampling plans (ASPs) with ZIP distribution. For instance, Loganathan and Shalini (2014) constructed the SSP when the number of nonconforming items was under the condition of ZIP distribution. Uma and Ramya (2016) proposed a ZIP distribution with a quick switching system (QSS), while Rao and Aslam (2017) defined the number of nonconforming items under the ZIP distribution to construct the resubmitted sampling plan. Wang and Hailemariam (2018) suggested repetitive group sampling and MDS sampling under the ZIP distribution, while Charongrattanasakul and Bamrungsetthapong (2021) considered the double acceptance sampling plan under the ZIP distribution.

A comprehensive literature review indicated that when the production process is well developed zero defects are generally discovered in sample inspections. Three important manufacturing objectives from an optimal ASP are minimal cost, smallest average sample number (ASN) and highest probability of acceptance. In this study, multi-objective optimization was employed to determine the optimal ASP under the proposed process.

As mentioned, the PQL distribution was used to analyze count data with overdispersion. However, this may not be suitable for count data with many zeros. Therefore, in this paper proposed a zero inflated version of the PQL distribution, called the zero-inflated Poisson quasi-Lindley (ZIPQL) distribution. The MDS sampling plan was designed under a ZIPQL distribution represented by the MDS_{ZIPQL} plan. Optimal plan parameters for the MDS_{ZIPQL} plan were determined to obtain the minimum and maximum values of the multi-objective function. Sensitivity analysis of the MDS_{ZIPQL} plan was discussed using various parameter values, while the MDS_{ZIPQL} plan was presented with two real data sets to demonstrate its application. Comparative performances between the MDS_{ZIPQL} and MDS_{ZIP} plans were considered, and conclusions were included.

2. Methods

2.1 The PQL distribution

Let X be a random variable distributed as the PQL distribution with parameters θ and γ denoted by X~PQL(θ, γ), with a probability mass function (pmf) as (Grine and Zeghdoudi, 2017)

$$g(x)=\frac{\theta}{\gamma+1}\left[\frac{\gamma(\theta+1)+\theta(x+1)}{(\theta+1)^{x+2}}\right], \tag{1}$$

Where $x = 0, 1, 2, \ldots, \theta > 0, \gamma > -1$. The cumulative density function (cdf) is

$$G(x)=1-\frac{\gamma+2\theta+\gamma\theta+\theta x+1}{(\gamma+1)(\theta+1)^{x+2}}. \tag{2}$$

Its mean and variance are:

$$\mu_{\mathrm{PQL}}=\frac{2+\alpha}{\theta(1+\gamma)} \quad \text{and}$$

$$\sigma^2_{\mathrm{PQL}}=\frac{2+4\gamma+\gamma^2+\theta(2+\gamma)(1+\gamma)}{\theta(1+\gamma)^2}. \tag{3}$$

For $\gamma = \theta$ the PQL distribution reduces to the PL distribution (Sankaran, 1970). In 2019, re-parameterization of the PQL distribution was proposed by Altun (2019), by letting $\theta = \frac{2+\gamma}{(1+\gamma)\mu}$, and the pmf of the PQL distribution then becomes

$$g(x) = \frac{2+\alpha}{(1+\gamma)^2\mu}\left\{\frac{\gamma+(2+\gamma)(\mu+\gamma\mu)^{-1}(\mathrm{x}+\gamma+1)}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^{x+2}}\right\}, \tag{4}$$

where $x = 0, 1, 2, \ldots, \gamma > 0, \mu > 0.$ Its mean and variance are respectively

$$\mu_{\mathrm{PQL}} = \mu \text{ and } \sigma^2_{\mathrm{PQL}} = \mu + \frac{(2+4\gamma+\gamma^2)\mu^2}{(2+\gamma)^2}. \tag{5}$$

2.2 Zero inflated distributions

Zero inflated distribution mixes two zero generating processes as a counting distribution under non-negative integers, using the pmf $g(x;\boldsymbol{\xi})$ with vector parameters $\boldsymbol{\xi}$. The pmf of the zero inflated distribution (Lambert, 1992) is given by

$$f(x) = P(X = x \mid \phi, \boldsymbol{\xi}) = \phi\pi_0(x) + (1-\phi)g(x;\boldsymbol{\xi}), \tag{6}$$

where $x = 0, 1, 2, \ldots, 0 < \phi < 1$, and

$$\pi_0(x) = \begin{cases} 1 & , x = 0, \\ 0 & , x = 1, 2, 3, \ldots \end{cases}.$$

In 1992, Lambert proposed the zero inflated Poisson (ZIP) distribution with the pmf (Lambert, 1992) as

$$f(x;\phi,\lambda) = \begin{cases} \phi + (1-\phi)e^{-\lambda} & , x = 0, \\ (1-\phi)\dfrac{e^{-\lambda}\lambda^x}{x!} & , x = 1, 2, 3, \ldots, \end{cases} \tag{7}$$

where $f(x;\phi,\lambda)$ is the pmf of the ZIP with parameters ϕ and λ, denoted by X~ZIP (ϕ, λ). Its mean and variance are $(1 - \phi)\lambda$ and $\lambda(1 - \phi)(1 + \lambda\phi)$, respectively.

3. Results and Discussion

This section proposes a new zero-inflated distribution called the ZIPQL distribution. The MDS sampling plan was designed under the ZIPQL distribution and included some simulation and application studies.

3.1 A new zero inflated distribution

A new zero inflated distribution was obtained by mixing a process that generated zeros and the PQL distribution under non-negative integers, the so-called ZIPQL distribution.

Let X~PQL (μ, γ), then the pmf of X under the zero inflated distribution is

$$f(x;\boldsymbol{\Theta}) = \begin{cases} \phi + \dfrac{(1-\phi)(2+\gamma)}{(1+\gamma)^2\mu}\left\{\dfrac{\gamma+(2+\gamma)(1+\gamma)(\mu+\gamma\mu)^{-1}}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^2}\right\}; x = 0, \\ \dfrac{(1-\phi)(2+\gamma)}{(1+\gamma)^2\mu}\left\{\dfrac{\gamma+(2+\gamma)(\mu+\gamma\mu)^{-1}(\mathrm{x}+\gamma+1)}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^{x+2}}\right\}; x = 1, 2, 3, \ldots, \end{cases} \tag{8}$$

where a parameter vector $\boldsymbol{\Theta} = (\phi, \gamma, \mu)$ for $0 < \phi < 1$, $\gamma > 0$ and $\mu > 0$. Let X be a random variable distributed as the ZIPQL distribution with parameters ϕ, μ, *and* γ, denoted as X~ZIPQL (ϕ, μ, γ). Its mean and variance are respectively

$$\mu_{\mathrm{ZIPQL}} = (1-\phi)\mu \text{ and } \sigma^2_{\mathrm{ZIPQL}} = (1-\phi)\left[\mu + \frac{(2+4\gamma+\gamma^2)\mu^2}{(2+\gamma)^2} + \phi\mu^2\right]. \tag{9}$$

If ϕ, $\theta = (2+\gamma)/[(1+\gamma)\mu]$ and $\theta = \gamma$ then the ZIPQL distribution reduces to a zero-inflated Poisson Lindley (ZIPL) distribution (Xavier, Santos-Neto, Bourguignon, & Tomazella, 2017).

Some pmf plots of ZIPQL are shown in Figure 1. For fixed values of μ and γ, the proportion of zeros increases when the ϕ value increases; see Figure 1(a). For fixed values of ϕ and γ the proportion of zeros increases when the μ value decreases; see Figure 1(b). For fixed values of ϕ and μ the proportion of zeros increases when the γ value increases; see Figure 1(c).

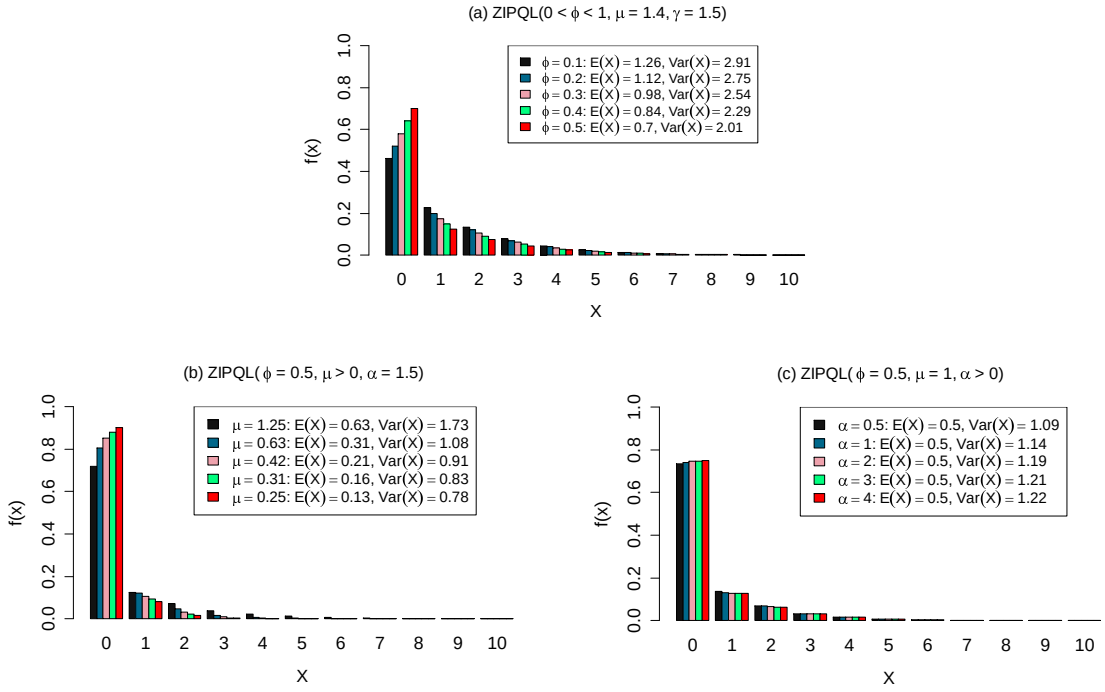

Figure 1. Plots of pmf for ZIPQL with specified parameters ϕ, μ and γ

3.2 Parameter estimation

Let X_i be the independent and identically distributed random variables of size n as $X_i \sim \text{ZIPQL}(\phi, \mu, \gamma)$. In this study, the maximum likelihood (ML) estimation was used to estimate the unknown parameters ϕ, μ and γ. From the pmf in (8), the log-likelihood function for the parameter vector of $\boldsymbol{\Theta} = (\phi, \mu, \gamma)$ can be written as

$$\log L(\boldsymbol{\Theta}) = \sum_{i=1}^{n} \left\{ I_{(x_i=0)} \left[\log\left(\phi + \frac{(1-\phi)(2+\gamma)}{\mu(1+\gamma)^2} \left(\frac{\gamma + (2+\gamma)(1+\gamma)(\mu+\gamma\mu)^{-1}}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^2} \right) \right) \right] \right.$$
$$\left. + I_{(x_i>0)} \left[\log\left(\frac{(1-\phi)(2+\gamma)}{\mu(1+\gamma)^2} \left(\frac{\gamma + (2+\gamma)(\mu+\gamma\mu)^{-1}(x_i+\gamma+1)}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^{x_i+2}} \right) \right) \right] \right\}, \quad (10)$$

To estimate the unknown parameters ϕ, μ and γ, the partial derivatives with respect to each parameter were equated to zero, i.e.,

$$\frac{\partial \log L(\boldsymbol{\Theta})}{\partial \phi} = 0, \frac{\partial \log L(\boldsymbol{\Theta})}{\partial \mu} = 0, \frac{\partial \log L(\boldsymbol{\Theta})}{\partial \gamma} = 0. \quad (11)$$

This equation cannot be derived in a closed form; therefore, solutions for the ML estimators of ϕ, μ and γ, were obtained using the *nlm* function on stats package in R program (R Core Team, 2020).

3.3 Design of MDS sampling plan

This section presents the MDS sampling plan, which defines the number of nonconforming items following the ZIPQL distribution, denoted by the $\text{MDS}_{\text{ZIPQL}}$ plan. Implementation steps of the $\text{MDS}_{\text{ZIPQL}}$ plan were described as follows (Wang & Hailemariam, 2018):

Step 1: Define two risk values (α and β) and two quality levels values (p_1 and p_2).

Step 2: Define the plan parameters: the sample size n, the acceptance number c_a, the rejection number c_r, and the number of previously accepted lots m.

Step 3: Draw a random sample of n items from the current lot and count the number of nonconforming items d.

1) Accept the currently inspected lot if $d \le c_a$.

2) Reject the currently inspected lot if $d \ge c_r$.

3) Accept the currently inspected lot if $c_a < d < c_r$ and the number of nonconforming items from the m previous lots are less than or equal to c_a.

All the above steps are summarized in Figure 2. The probability of the acceptance function (P_a (p)) for the MDS plan from Wang and Hailemariam (2018) in (12) was considered.

$$P_a(p) = P(d \le c_a) + P(c_a < d < c_r) \cdot \left[P(d \le c_a)\right]^m \tag{12}$$

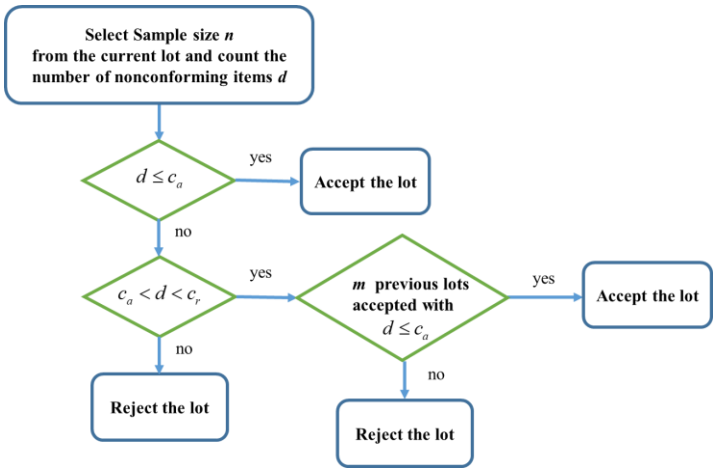

Figure 2. Operating procedure of the MDS sampling plan

Therefore, the P_a (p) of the MDS_{ZIPQL} plan under a ZIPQL (ϕ, μ, γ) is shown as (13).

$$\begin{aligned}
P_a(p) = {} & \phi + \frac{(1-\phi)(2+\gamma)}{(\gamma+1)^2\mu}\left\{\frac{\gamma+(2+\gamma)(1+\gamma)(\mu+\gamma\mu)^{-1}}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^2}\right\} \\
& + \frac{(1-\phi)(2+\gamma)}{(\gamma+1)^2\mu}\sum_{d=1}^{C_a}\left\{\frac{\gamma+(2+\gamma)(\mu+\gamma\mu)^{-1}(d+\alpha+1)}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^{d+2}}\right\} \\
& + \frac{(1-\phi)(2+\gamma)}{(\gamma+1)^2\mu}\sum_{d=c_a+1}^{C_r-1}\frac{\gamma+(2+\gamma)(\mu+\gamma\mu)^{-1}(d+\alpha+1)}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^{d+2}} \\
& \times\left\{\phi + \frac{(1-\phi)(2+\gamma)}{(\gamma+1)^2\mu}\left[\frac{\gamma+(2+\gamma)(1+\gamma)(\mu+\gamma\mu)^{-1}}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^2}\right]\right. \\
& \left. + \frac{(1-\phi)(2+\gamma)}{(\gamma+1)^2\mu}\sum_{d=1}^{C_a}\frac{\gamma+(2+\gamma)(\mu+\gamma\mu)^{-1}(d+\alpha+1)}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^{d+2}}\right\}^m,
\end{aligned} \tag{13}$$

where $\mu = np,\ \mu > 0,\ 0 < \phi < 1,\ \gamma > 0$ and p is the proportion of nonconforming items.

The ASN function for the MDS plan from Wang and Hailemariam (2018) was also applied to create the ASN function for the MDS_{ZIPQL} plan as:

$$\begin{aligned}\mathbf{ASN}(p) &= n + nP\left(c_a < d \le c_r\right) \\ &= n\left[1 + \sum_{d=c_a+1}^{c_r} \frac{(1-\phi)(2+\gamma)}{(\gamma+1)^2\,\mu}\left(\frac{\gamma+(2+\gamma)(\mu+\gamma\mu)^{-1}(d+\alpha+1)}{\left[1+(2+\gamma)(\mu+\gamma\mu)^{-1}\right]^{x+2}}\right)\right].\end{aligned} \tag{14}$$

3.4 Economic design of the MDS sampling plan

This section applies concepts from Charongrattanasakul and Bamrungsetthapong (2021) to design the total cost function for the inspection lot with the MDS_{ZIPQL} plan as shown in (15)-(18).

First: let C_I be the cost of inspection per lot as given below:

$$C_I = C_1\left(N - (N-n)\cdot P_a(p)\right) \tag{15}$$

where C_1 represents the inspection cost per unit.

Second: let C_F be the cost of the internal failure per lot as

$$C_F = C_2\left(N + (N-n)\cdot P_a(p)\right)p \tag{16}$$

where C_2 represents the internal failure cost per unit.

Third: let C_0 be the cost of outgoing nonconforming per lot as

$$C_0 = C_3 P_a(p)(N-n)\,p \tag{17}$$

where C_3 represents the cost of outgoing nonconforming per unit.

Therefore, the total cost of the MDS_{ZIPQL} plan can be expressed as:

$$\begin{aligned}TC &= C_I + C_F + C_0 \\ &= C_1\left(N-(N-n)P_a(p)\right) + C_2\left(N+(N-n)P_a(p)\right)p + C_3P_a(p)(N-n)\,p.\end{aligned} \tag{18}$$

The optimal plan parameter (n, c_a, c_r, m)* of the MDS_{ZIPQL} plan was calculated to obtain the minimum values of ASN, total cost (TC) of inspection and the maximum values of P_a (p) simultaneously. The genetic algorithm (GA) technique with multi-objective optimization was applied in a simulation study using the R program. The optimal solution of the proposed plan was obtained by substituting the optimal values of n, c_a, c_r, and m in the multi-objective function. The constraints of the producer's risk (α) and the consumer's risk (β) are immediately satisfied with the acceptable quality level (AQL or p_1) and the lot tolerance percent defective (LTPD or p_2). The MDS_{ZIPQL} plan is practical with two points (AQL, 1-α) and (LTPD, β), and considered for changes in the operating characteristic (OC) curve. The optimal solution was considered on three objective functions concurrently, as follows.

Multi-objective function:

Minimize ASN and TC (19)

Maximize P_a (p) (20)

Subject to: $P_a\left(p_1 \mid \mu_1 = np_1\right) \ge 1-\alpha$, $P_a\left(p_2 \mid \mu_2 = np_2\right) \le \beta$, $n > 1$, $m \ge 1$, and $c_r > c_a \ge 0$.

Assume that the following sets of input parameters are given: N=1000, α=0.05, β=0.01, p_1=0.05, and p_2=0.1, while the fixed value of cost in each status is given by Hsu and Hsu (2012) as $C_I = 1$, $C_F = 2$, and $C_0 = 10$, respectively.

3.5 Simulation study

Sensitivity analyses of optimal (n, c_a, c_r, m)* under the MDS_{ZIPQL} plan by considering changes of the four variables: ϕ, γ, m and p are shown in Table 1. By solving equations (19) and (20) under the two inequalities $P_a(p_1 \mid \mu_1 = np_1) \ge 1-\alpha$ and $P_a(p_2 \mid \mu_2 = np_2) \le \beta$, this simulation study applied the GA technique with multi-objective optimization to determine the maximum value of P_a (p) and minimum values of TC and ASN.

Results in Table 1 give the optimal (n, c_a, c_r, m)* which can be considered as different zero-inflation parameter values (ϕ) for fixed values of p, γ and m. The results show that if ϕ increases, the required n decreases, causing ASN and TC to decrease

Table 1. Optimal plan parameters (n, c_a, c_r, m)* for the multi-objective functions of the MDS_{ZIPQL} plan under different values of γ, ϕ, m and p

p	γ	ϕ	m = 1						m = 2					
			n	c_a	c_r	ASN	TC	$P_a(p)$	n	c_a	c_r	ASN	TC	$P_a(p)$
0.01	1	0.0001	89	4	9	90.59	183	0.9972	76	4	10	76.84	171	0.9981
		0.001	85	4	8	86.29	180	0.9974	74	4	10	74.70	169	0.9989
		0.01	79	3	9	81.41	175	0.9974	74	4	8	74.74	169	0.9992
		0.10	77	3	9	79.15	173	0.9980	69	4	9	69.55	164	0.9998
		0.20	76	3	9	78.00	171	0.9984	64	3	8	65.11	160	0.9999
	5	0.0001	79	3	8	81.84	175	0.9950	64	3	8	65.38	161	0.9918
		0.001	70	2	9	74.77	169	0.9951	52	2	7	53.76	155	0.9922
		0.01	67	2	8	71.09	165	0.9956	51	2	8	52.70	153	0.9928
		0.10	63	2	9	66.48	162	0.9956	49	2	7	50.59	148	0.9952
		0.20	60	2	7	62.76	159	0.9958	48	2	8	49.37	147	0.9958
	10	0.0001	69	3	7	70.77	168	0.9946	65	3	8	66.49	162	0.9979
		0.001	67	2	8	71.12	167	0.9946	49	2	7	50.67	148	0.9982
		0.01	65	2	8	68.76	165	0.9960	42	2	9	43.06	140	0.9985
		0.10	64	2	9	67.58	162	0.9963	39	2	5	39.79	138	0.9986
		0.20	60	2	7	62.81	158	0.9966	36	2	6	36.59	134	0.9992
0.05	1	0.0001	75	3	8	97.41	685	0.7446	75	4	9	92.46	670	0.7750
		0.001	71	3	8	91.14	673	0.7555	61	3	9	77.94	660	0.7809
		0.01	70	3	9	89.82	680	0.7572	67	3	8	70.99	649	0.7971
		0.10	63	2	8	86.22	677	0.7630	56	3	8	69.28	637	0.8171
		0.20	60	2	8	78.89	665	0.7717	52	3	6	60.51	635	0.8178
	5	0.0001	69	3	8	85.14	675	0.7616	63	3	8	78.85	670	0.7657
		0.001	57	2	9	76.33	661	0.7662	61	3	8	75.84	664	0.7735
		0.01	53	2	7	67.55	666	0.7743	60	3	9	75.70	651	0.7963
		0.10	50	2	7	64.82	638	0.8116	45	2	8	57.58	637	0.8077
		0.20	48	2	6	59.43	631	0.8206	42	2	6	51.37	630	0.8168
	10	0.0001	59	2	9	79.28	665	0.7704	46	2	8	58.41	658	0.7717
		0.001	54	2	7	70.61	657	0.7809	43	2	6	53.58	652	0.7802
		0.01	48	2	6	60.60	646	0.7955	39	2	5	46.86	643	0.7922
		0.10	47	2	6	58.77	634	0.8144	29	1	6	37.14	643	0.7947
		0.20	43	2	5	51.18	628	0.8218	27	1	5	33.44	634	0.7978

while $P_a(p)$ increases. Different values of γ occurred under fixed values of p, ϕ and m. Results showed that if γ increases, the required n decreases, causing TC, ASN and $P_a(p)$ to decrease. Different values of m were considered under fixed values of p, ϕ, and γ. Results showed that if m increases, the required n decreases, causing TC, ASN and $P_a(p)$ to decrease. Different values of p were considered under given values of ϕ γ, and m. Results showed that if p increases, the required n decreases, causing TC to increase while ASN and $P_a(p)$ decrease. The optimal plan parameters under the MDS_{ZIPQL} plan provided maximum $P_a(p)$ and minimum values of ASN and TC.

Table 2 presents the effect of required sample sizes (n) on the efficiency of the MDS_{ZIPQL} plan with three different levels of γ, such as 1, 5, 20 and five different levels of ϕ such as 0.0001, 0.001, 0.01, 0.10 and 0.20 under given values of C_a = 0, C_r = 2, m = 2, and p = 0.01. These results suggested that larger values of γ provide smaller required n under the same value of ϕ, and ASN and TC decrease while $P_a(p)$ increases. Larger values of ϕ also provide smaller required n under a fixed value of γ and ASN, TC and $P_a(p)$ tend to decrease.

Table 3 shows the performance of the MDS_{ZIPQL} plan in reducing the required sample sizes and ASN compared to the MDS_{ZIP} plan. The required sample sizes and ASN of the MDS_{ZIP} plan were obtained from Wang and Hailemariam (2018) and are shown in Table 3. The required sample sizes and ASN of the MDS_{ZIPQL} plan under different levels of ϕ = (0.001, 0.01, 0.04), p_1 = (0.01, 0.05), p_2 = [(0.05, 0.06, 0.09), (0.15, 0.17, 0.20)], α = 0.05, and β = 0.01 are given in Table 3. The required sample sizes and ASN of the MDS_{ZIPQL} plan were less than the MDS_{ZIP} plan. Results indicated that the MDS_{ZIPQL} plan performed better than the MDS_{ZIP} plan regarding the required sample size and ASN.

3.6 Application study

In this study, two real data sets were used to demonstrate the application of the ZIP and ZIPQL distributions, while the parameters in each distribution were estimated by the ML method. The Kolmogorov-Smirnov (KS) test, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used for the goodness of fit test of each distribution. Models with smaller values of AIC, BIC and KS showed the best fit for the data.

Example 1: Real data sets of numbers of read-write errors found in a computer hard disk in a production process (Xie *et al.*, 2001) were considered. This data set included 208 samples with a mean of 1.16 and a standard deviation of 1.20 as count data with overdispersion. Parameter estimates and the goodness of fit test for these data sets are given in Table 4. The ZIPQL distribution gave lower AIC, BIC and KS values than the ZIP distribution. Thus, the ZIPQL distribution was appropriate to fit this data.

Table 2. The effect of γ on optimal required sample sizes (n) of the MDS_{ZIPQL} plan under given values $C_a = 0$, $C_r = 2$, $m = 2$, and $p = 0.01$

ϕ	$\gamma = 1$				$\gamma = 5$				$\gamma = 20$			
	n	*ASN*	*TC*	$P_a(p)$	*n*	*ASN*	*TC*	$P_a(p)$	*n*	*ASN*	*TC*	$P_a(p)$
0.0001	25	29.86	209	0.9040	18	20.72	172	0.9389	16	18.16	159	0.9508
0.001	18	20.69	169	0.9421	17	17.71	150	0.9596	15	16.73	148	0.9621
0.01	15	15.68	148	0.9618	14	14.47	144	0.9653	13	13.52	143	0.9662
0.10	13	14.42	141	0.9677	12	12.97	135	0.9775	11	11.14	131	0.9739
0.20	11	12.04	132	0.9757	10	10.81	127	0.9809	9	9.66	123	0.9841

Table 3. Comparison of the required sample sizes and ASN of the MDS_{ZIPQL} plan and the MDS_{ZIP} plan under $m = 2$

ϕ	p_1	p_2	C_a	C_r	MDS_{ZIP*}		MDS_{ZIPQL}	
					n	*ASN*	*n*	*ASN*
0.001	0.01	0.05	1	3	79	90.2	58	63.67
	0.01	0.06	0	2	39	47.8	36	61.04
	0.01	0.09	0	2	26	31.2	26	30.66
0.01	0.01	0.05	1	3	81	92.5	58	63.62
	0.01	0.06	0	2	41	49.5	41	49.40
	0.01	0.09	0	2	27	32.5	25	29.31
0.04	0.01	0.05	1	3	90	99.7	78	85.59
	0.01	0.06	0	2	47	54.6	42	47.79
	0.01	0.09	0	2	32	37.0	29	32.83
0.001	0.05	0.15	4	6	54	58.3	40	42.55
	0.05	0.17	3	5	40	43.6	32	34.52
	0.05	0.20	2	4	27	29.9	21	22.81
0.01	0.05	0.15	4	6	55	59.5	36	37.91
	0.05	0.17	3	5	41	44.8	39	42.81
	0.05	0.20	2	4	28	31.0	28	30.80
0.04	0.05	0.15	5	7	68	71.8	62	66.18
	0.05	0.17	4	6	52	55.3	36	37.85
	0.05	0.20	2	4	31	33.4	31	33.01

* The result from Wang and Hailemariam (2018).

Table 4. Summary of goodness of fit test of ZIP and ZIPQL distributions for the number of errors

Number of errors	Observed frequency	Expected frequency of distributions	
		ZIP	ZIPQL
0	180	180.01	180.00
1	11	0.04	3.24
2	5	0.18	2.86
3	2	0.53	2.53
4	1	1.15	2.24
5	1	1.99	1.98
6	2	2.86	1.75
9	2	3.66	1.21
11	1	2.49	0.95
15	1	0.42	0.58
75	2	0.00	0.00
	ML estimators	$\hat{\phi} = 0.8654,\ \hat{\lambda} = 8.6413$	$\hat{\phi} = 0.8478,\ \hat{\mu} = 7.6429,\ \hat{\gamma} = 871.37$
	- log L	405.20	168.88
	AIC	814.40	343.76
	BIC	821.08	353.77
	KS test (p-value)	0.7745 (0.4430)	0.7740 (0.4490)

The number of read-write errors from real data was used to determine the optimal (n, c_a, c_r, m)* under the MDS_{ZIPQL} plan. The input parameters were assigned as follows: $N = 208$, ϕ=0.8478, γ=871.37, p_1=0.05, p_1=0.10, α = 0.05, β = 0.10, $C_I = 1$, $C_F = 2$, and $C_0 = 10$. The optimal (n, c_a, c_r, m)*, as shown in Table 6, was obtained by substituting input parameters in nonlinear multi-objective optimization.

Table 6 shows various optimal plan parameters based on values of p and m under given values ϕ =0.8478, γ =871.37, and $c_a = 0$. For example, if the producer knows that the proportion of read-write errors per lot is p = 0.001 and the number of previously accepted lots is m = 1, then the optimal plan parameter is (50, 0, 4, 1)*. Whereas if the producer needs to reduce ASN and TC, then the value of m should be increased, such as m = 4, and the optimal plan parameter is (30, 0, 3, 4)*.

Table 7 shows the performance comparison of our proposed MDS_{ZIPQL} plan and the MDS_{ZIP} plan proposed by Wang and Hailemariam (2018). Performances of both sampling plans were considered based on optimal required sample sizes n under four different levels of m and given values $c_a = 0$, and $c_r = 2$. At m = 1 the MDS_{ZIPQL} plan used only the required sample size of 26, which was smaller than the MDS_{ZIP} plan, with the required sample size of 35. As m increased, both sampling plans used equal required sample sizes n. Results showed that the MDS_{ZIPQL} plan gave lower ASN and TC and higher $P_a(p)$ than the MDS_{ZIP} plan.

Example 2: Data on the number of defective LEDs within a lot (He *et al.*, 2012; Alevizakos & Koukouvinos, 2020) are presented in Table 7. This data set included 200 samples with a mean of 1.38 and a standard deviation of 3.57 as count data with overdispersion. Parameter estimates and the goodness of fit test for these data sets are illustrated in Table 5. The ZIPQL distribution gave lower AIC, BIC and KS values than the ZIP distribution. Therefore, the ZIPQL distribution was appropriate to fit this data.

Table 5. Summary of goodness of fit tests of ZIP and ZIPQL distributions for the number of defective LEDs

Number of defective LEDs	Observed frequency	Expected frequency of distributions	
		ZIP	ZIPQL
0	162	162.01	162.01
1	1	0.19	5.23
2	6	0.71	4.51
3	2	1.71	3.89
4	3	3.10	3.35
5	4	4.50	2.89
6	5	5.44	2.49
7	2	5.64	2.15
8	4	5.11	1.85
9	3	4.13	1.60
10	1	2.99	1.38
12	2	1.20	1.03
16	2	0.08	0.57
18	1	0.01	0.42
0.3619	2	0.00	0.36
	ML estimators	$\hat{\phi} = 0.8099,\ \hat{\lambda} = 7.2580$	$\hat{\phi} = 0.7797,\ \hat{\mu} = 6.2631,\ \hat{\gamma} = 2346.1106$
	- log L	224.66	207.85
	AIC	453.32	421.70
	BIC	459.92	431.60
	KS test (p-value)	0.7440 (0.2144)	0.7434 (0.3793)

Table 6. Optimal plan parameters $(n,c_a,c_r,m)^*$ of the MDS_{ZIPQL} plan to inspect the example data sets

m	Data of example 1						Data of example 2					
	p	n	c_r	ASN	TC	$P_a(p)$	p	n	c_r	ASN	TC	$P_a(p)$
1	0.001	50	4	51.38	52.68	0.9999	0.001	57	5	57.68	58	0.9999
	0.005	51	4	52.57	59.54	0.9988	0.005	59	5	61.99	67	0.9973
	0.01	51	4	53.59	68.37	0.9955	0.01	59	5	63.88	75	0.9919
	0.05	52	4	56.16	139.78	0.9511	0.05	61	5	68.85	143	0.9265
2	0.001	35	3	35.79	44.25	0.9985	0.001	31	3	31.21	33	0.9999
	0.005	46	4	46.31	47.73	0.9999	0.005	32	3	32.98	41	0.9977
	0.01	48	4	50.34	65.88	0.9937	0.01	34	3	34.80	52	0.9913
	0.05	40	4	54.02	139.10	0.9462	0.05	40	4	44.85	131	0.9268
3	0.001	32	3	32.67	41.40	0.9983	0.001	28	3	28.17	30	0.9999
	0.005	41	4	42.06	50.09	0.9979	0.005	32	3	32.98	41	0.9968
	0.01	42	4	44.79	79.51	0.9812	0.01	35	4	37.02	54	0.9898
	0.05	45	4	48.65	136.37	0.9450	0.05	41	4	45.97	133	0.9137
4	0.001	30	3	30.13	31.85	0.9980	0.001	25	3	25.13	27	0.9999
	0.005	35	4	35.79	44.34	0.9979	0.005	25	3	25.62	34	0.9974
	0.01	35	3	36.36	54.29	0.9919	0.01	32	3	33.71	51	0.9907
	0.05	37	4	40.01	131.48	0.9485	0.05	35	4	39.22	129	0.9147

Table 7. Optimal required sample sizes (n) with $(n,0,c_r,m)^*$ under MDS_{ZIPQL} and MDS_{ZIP} plans to inspect the example data sets

Data sets	MDS_{ZIPQL} plan					MDS_{ZIP} plan			
	m	n	ASN	TC	$P_a(p)$	n	ASN	TC	$P_a(p)$
Example 1	1	26	26.10	28	0.9999	35	36.37	54	0.9921
$(n,0,2,m)^*$	2	25	25.09	27	0.9999	25	25.73	45	0.9949
	3	23	23.08	25	0.9999	23	23.63	43	0.9950
	4	22	22.07	24	0.9999	22	22.58	42	0.9948
Example 2	1	36	38.10	54	0.9929	38	40.64	56	0.9917
$(n,0,3,m)^*$	2	31	32.62	50	0.9922	34	36.15	53	0.9912
	3	28	29.35	47	0.9914	29	30.61	48	0.9907
	4	25	26.11	44	0.9912	25	26.22	44	0.9908

The number of defective LEDs was used to determine the optimal $(n, c_a, c_r, m)^*$ under the MDS_{ZIPQL} plan. The input parameters were N = 200, ϕ =0.7797, γ =2,346.1106, p_1 = 0.05, p_2 = 0.10, α = 0.05 and β = 0.10. The various optimal plan parameters $(n, c_a, c_r, m)^*$ based on p and m under fixed values ϕ = 0.7797, γ = 2,346.1106 and c_a = 0 are reported in Table 6. For example, if the proportion of defective LEDs per lot is p =0.05 and the number of previously accepted lots is m = 3, and the optimal plan parameter is (41, 0, 4, 3)*.

Performance comparison of the proposed MDS_{ZIPQL} plan and the MDS_{ZIP} plan under the data set on number of defective LEDs is shown in Table 7. Performances of both sampling plans were considered based on optimal values of n and m for fixed values c_a = 0, and c_r = 3. Results showed that at m = 1 the MDS_{ZIPQL} plan used only the required sample size of 36, which was smaller than the MDS_{ZIP} plan with the required sample size of 38. As m increased, the MDS_{ZIPQL} plans used required sample sizes less than the MDS_{ZIP} plan. Therefore, the MDS_{ZIPQL} plan gave lower ASN and TC and higher $P_a(p)$ than the MDS_{ZIP} plan.

4. Conclusions

This research proposed MDS sampling plans to inspect serially submitted lots in production. Most production processes have excellent quality control and inspection. With excellent quality control, the manufacturer can operate the production process with no or almost no defects. Excellent inspection results occur when a product is of good quality and there are often zero defects in the inspection process. Our proposed MDS sampling plan was designed under a new zero-inflated distribution, called the ZIPQL distribution, denoted by the MDS_{ZIPQL} plan. The optimal plan parameter under the MDS_{ZIPQL} plan was calculated to obtain the maximum $P_a(p)$ and minimum ASN and TC simultaneously. Sensitivity analysis of the MDS_{ZIPQL} plan indicated that different values of ϕ, γ, m, and p, affected the required sample size n, ASN, TC and $P_a(p)$. The MDS_{ZIPQL} plan performed well when ϕ, γ, and m were large values, while p was close to zero. Finally, two real data sets were considered under the ZIPQL distribution which fitted this data. The optimal plan parameter under the MDS_{ZIPQL} plan was determined. These results were also compared with the MDS_{ZIP} plan under the same conditions. The MDS_{ZIPQL} plan showed better performance than the MDS_{ZIP} plan. In practice, the MDS_{ZIPQL} plan gave a good performance when p was small while ϕ was large. This allows producers to use a small sample size and save on costs using the MDS_{ZIPQL} plan. Inaccurate data in the production process leads to indecision about product quality. Neutrosophic statistics might be applied to the MDS_{ZIPQL} plan in future studies.

Acknowledgements

The author would like to thank the anonymous reviewers for comments and suggestions.

References

Alevizakos, V., & Koukouvinos, C. (2020). Monitoring of zero-inflated Poisson processes with EWMA and DEWMA control charts. *Quality and Reliability Engineering International, 36*(1), 88-111.

Altun, E. (2019). A new model for over-dispersed count data: Poisson quasi-Lindley regression model. *Mathematical Sciences, 13*(3), 241-247.

Aslam, M., Nazir, A., & Jun, C. H. (2015). A new attribute control chart using multiple dependent state sampling. *Transactions of the Institute of Measurement and Control, 37*(4), 569-576.

Aslam, M., Jeyadurga, P., Balamurali, S., Azam, M., & Al-Marshadi, A. (2021). Economic determination of modified multiple dependent state sampling plan under some lifetime distributions. *Journal of Mathematics*, Article ID 7470196. doi:10.1155/2021/7470196

Balamurali, S., & Jun, C. H. (2007). Multiple dependent state sampling plans for lot acceptance based on measurement data. *European Journal of Operational Research*, *180*(3), 1221-1230.

Charongrattanasakul, P., & Bamrungsetthapong, W. (2021). Designing of optimal required sample sizes for double acceptance sampling plans under the zero-inflated defective data. *Current Applied Science and Technology, 21*(2), 227-239.

Fang, R. (2013). *Zero-inflated negative binomial (ZINB) regression model for over-dispersed count data with excess zeros and repeated measures, an application to human microbiota sequence data* (Doctoral dissertation, University of Colorado at Denver, Anschutz Medical Campus). Retrieved from http://hdl.handle.net/10968/244.

Grine, R., & Zeghdoudi, H. (2017). On Poisson quasi-Lindley distribution and its applications. *Journal of Modern Applied Statistical Methods, 16*(2), 403-417.

He, S., Huang, W., & Woodall, W. H. (2012). CUSUM charts for monitoring a zero-inflated Poisson process. *Quality and Reliability Engineering International, 28*(2), 181-192.

Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge, England: Cambridge University Press.

Hsu, L. F., & Hsu, J. T. (2012). Economic design of acceptance sampling plans in a two-stage supply chain. *Advances in Decision Sciences Article, ID 359082*.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Techno metrics, 34*(1), 1-14.

Li, S., Yang, F., Famoye, F., Lee, C., & Black, D. (2011). Quasi-negative binomial distribution: properties and applications. *Computational Statistics and Data Analysis, 55*(7), 2363-2371.

Loganathan, A., & Shalini, K. (2014). Determination of single sampling plans by attributes under the conditions of zero-inflated Poisson distribution. *Communications in Statistics-Simulation and Computation, 43*(3), 538-548.

Nadi, A. A., & Gildeh, B. S. (2019). A group multiple dependent state sampling plan using truncated life test for the Weibull distribution. *Quality Engineering*, 31(4), 1–11.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Rao, G. S., & Aslam, M. (2017). Resubmitted lots with single sampling plans by attributes under the conditions of zero-inflated Poisson distribution. *Communication in Statistics-Simulation and Computation, 46*(2), 1814-1824.

Ridout, M., Hinde, J., & Demétrio, C. G. (2001). A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics, 57*(1), 219-223.

Sankaran, M. (1970). The discrete Poisson-Lindley distribution. *Biometrics, 26*(1), 145-149.

Shalini, K., & Abdullah, A. S. (2018). Designing double sampling plans under the conditions of zero-inflated Poisson distribution. *Journal of Emerging Technologies and Innovative Research, 5*(12), 529-534.

Srinivasa Rao, G., Rosaiah, K., & Naidu, C. R. (2020). Design of multiple-deferred state sampling plans for exponentiated half logistic distribution. *Cogent Mathematics and Statistics, 7*(1), 1857915.

Uma, G., & Ramya, K. (2016). Determination of quick switching system by attributes under the conditions of zero-inflated Poisson distribution. *International Journal of Statistics and Systems*, 11, 157-165.

Wagner, B. D., Robertson, C. E., & Harris, J. K. (2011). Application of two-part statistics for comparison of sequence variant counts. *PloS One, 6* (5), e20296.

Wang, F. K., & Hailemariam, S. S. (2018). Sampling plans for the zero-inflated Poisson distribution in the food industry. *Food Control, 85*, 359-368.

Wortham, A. W., & Baker, R. C. (1976). Multiple deferred state sampling inspection. *The International Journal of Production Research, 14*(6), 719-731.

Xavier, D., Santos-Neto, M., Bourguignon, M., & Tomazella, V. (2017). Zero-Modified Poisson-Lindley distribution with applications in zero-inflated and zero-deflated count data. *arXiv preprint arXiv:1712. 04088*. Retrieved from https://arxiv.org/abs/1712. 04088.

Xie, M., He, B., & Goh, T. N. (2001). Zero-inflated Poisson model in statistical process control. *Computational Statistics and Data Analysis, 38*(2), 191-201.