Current Applied Science and Technology Vol. 21 No. 2 (April-June 2021)

Designing of Optimal Required Sample Sizes for Double Acceptance Sampling Plans under the Zero-Inflated Defective Data

Pramote Charongrattanasakul[1] and Wimonmas Bamrungsetthapong[2]*

[1]Division of Mathematics, Faculty of Science and Technology, Rajamangala University of Technology Krungthep, Bangkok, Thailand
[2]Division of Applied Statistics, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Pathumthani, Thailand

Received: 9 May 2020, Revsied: 3 August 2020, Accepted: 16 September 2020

Abstract

This research proposes an optimal double acceptance sampling plan (DSP) for manufacturing that is affected by zero-inflated data. Suppose that the number of defective items for sample inspection is considered to be under Zero-inflated Poisson (ZIP) distribution. A multi-objective optimization using Genetic Algorithm is applied to calculate the optimal parameters $(n_1, n_2, c_1, c_2)^*$ of the proposed DSP, which is concerned with maximizing the probability of acceptance sampling plan (P_a) and minimizing the total cost of inspection (TC) and the average number of samples(ASN)simultaneously. The optimal solution was focused on the design of the required sample sizes (n_1, n_2) based on three different scenarios. The results showed that the first sample and the second sample should be equal $(n_1 = n_2)$. Moreover, it was found that the probability of extra zeros $(\emptyset)$ under the ZIP distribution affects the required sample sizes and the performance of the proposed DSP. Illustrations for selecting the optimal parameters of the proposed DSP are also provided. Real data with excess zero is used to illustrate the application of the proposed DSP.

Keywords: double acceptance sampling plan; zero-inflated poisson distribution; probability of extra zeros; probability of acceptance sampling plan; Genetic Algorithm; multi-objective optimization
DOI 10.14456/cast.2021.21

1. Introduction

One important tool in the product control technique is an acceptance sampling plan (ASP). An ASP is applied in many areas to inspect the quality of items such as raw materials, some partial products of the production process, and finished products. This technique helps consumers decide whether to accept or reject a product that is produced by manufacturers based on sampling results selected from a lot. Users can decide to choose the minimum sample size from a sampling plan to achieve the acceptance criteria or the rejection criteria for that lot. Dodge and Romig's tables are widely used to

*Corresponding author: Tel.: (+66) 2549-4137 Fax(+66) 2549-4138
E-mail: wimonmas_b@rmutt.ac.th

decide on a sampling plan in which users know the lot size, percent nonconformance of the lot, producer's risk, and consumer's risk for the production process [1]. In practice, if some necessary values are unknown, then they cannot choose the optimal ASP. This problem can affect the total cost of the inspection. Recently, many researchers have studied the determination of optimal ASP model using optimization techniques as follows. Duarte and Saraiva [2] proposed a method to find the optimal value of the ASP model. The objective function was used to find the lowest value of error for the probability of accepting the ASP model for single and double sampling plans that corresponded to the sample size and the acceptable number. Kaya [3] applied a Genetic Algorithm (GA) to determine the sample size of the attribute control chart for a multi-state process. The objective function used to find the minimum cost and the maximum probability of accepting the model was found. Kobilinsky and Bertheau [4] presented a cost function for the inspection process that depended on the number of inspection groups and sample sizes for single and double ASPs based on the manufacturer's risk and the consumer's risk. Cheng and Chen [5] applied the GA methods to design a DSP. These models increased the efficiency of the design ASP and reduced errors that impacted on the manufacturer's risk and consumer's risk. Moreover, GA methods were applied to find the best information more efficiently and more accurately. Sampath and Deepa [6] designed a DSP by applying the GA methods to determine the optimal sample size and acceptance number under the manufacturer's risk and the consumer's risk. Braimah *et al*. [7] evaluated the optimal value of a mathematical model to determine the sample size and the random range for the ASP. It was found that the condition of these models provided an acceptance number equal to zero.

In addition, some researchers have studied the economic aspects of various ASPs. Hsu and Hsu [8] studied the cost aspects of a single ASP in order to evaluate the minimum cost that was appropriate for both manufacturers and consumers. Results showed that the proposed cost model was used to inspect a group of defective items. Fallahnezhad and Aslam [9] designed an economic model of the ASPs that involved Bayesian inference in which a decision was taken depending on the proposed model. Fallahnezhad and Qazvini [10] designed a new economic model of the ASP in a two-stage approach based on the Maxima Nomination Sampling (MNS) technique. Fallahnezhad *et al*. [11] presented an ASP based on the MNS method with current inspection errors. An economical model was proposed in terms of inspection errors and clarified the impact of errors from an economic point of view.

Currently, most production processes have excellent quality control. It was found that when the production process was well inspected, zero defects were more often discovered in sample inspections. For this reason, some researchers presented the idea that a zero-inflated Poisson (ZIP) distribution was appropriate for the probability distribution of the number of defects. A ZIP distribution was presented by Lambert [12], McLachlan and Peel [13]. It was a special type of mixed distribution that degenerated at zero and yet was a Poisson distribution. The ZIP distribution was applied in many disciplines such as manufacturing, public health, epidemiology, medicine, etc. Recently, many researchers developed ASPs with ZIP distributions. Loganathan and Shalini [14] considered a single sampling plans (SSP) in which the number of defective items was a ZIP distribution. The optimal plan parameters were calculated using unity values. Uma and Ramya [15] presented a Quick Switching System (QSS) that fell under a ZIP distribution. Rao and Aslam [16] designed resubmitted lots plan in which the number of defects was a ZIP distribution. The parameters of the proposed sampling plan were considered based on nonlinear optimization solutions. A nonlinear optimization is used to determine the optimal plan parameters of the proposed sampling plan.Wang and Hailemariam [17] proposed a repetitive group sampling (RGS) and multiple dependent state (MDS) sampling under a ZIP distribution. Moreover, they developed DSP and sequential sampling plans for the ZIP distribution. The unity value approach was applied to determine the optimal plan parameters of the proposed sampling plans.

The above-mentioned research showed that the production process was well inspected, the zero defects were more discovered in sample inspections. Three important objectives that the manufacturer expected to achieve from an optimal ASP: the lowest cost, the smallest ASN, and the highest probability of acceptance. From this idea, a multi-objective optimization is used to find the optimal DSP under the proposed process. Furthermore, GA is one of the most popular methods used to find the optimal value that provides the best answer to the problem and is flexible enough to solve complex problems, like those which are developed for genetic processes [18]. Many researchers indicate that GA can be used to resolve the problem of optimizing in ASP [3, 5, 6].

The aim of this research is to design the sample sizes required to achieve an optimal DSP with zero-inflated defective data. The optimal parameters for DSP under the ZIP distribution (DSP_{ZIP}) are calculated to maximize the P_a and minimize the TC and ASN simultaneously (multi-objective function). The MATLAB software (R2019b) [19] is used to provide a simulation study of the GA with multi-objective optimization. Additionally, we focus on studying the relationship between the required sample sizes and the economic models of the proposed ASP because, in practice, smaller value of required sample sizes or ASN are more satisfactory for designing an optimal ASP. Therefore, an economic model of DSP_{ZIP} under multi-objective optimization is also considered based on three different scenarios that compare the size of the first sample (n_1) and the second sample (n_2). In illustrations, the ratio of sample size, the different scenarios of the required sample sizes, and the optimal OC function of the DSP_{ZIP} are presented. Real data example were applied to determine the optimal plan parameters under proposed DSP_{ZIP}.

2. Materials and Methods

Currently, the number of defective items for many samples will be zero when most production processes have excellent quality control, and the production process is well inspected. In this situation, the proper probability distribution function of the number of defective items for sample inspection is the Zero-Inflated (ZI) distribution.

The ZI distribution is a mixture of a process that generates zeros and the other processes that are a counted distribution under non-negative integers. Suppose X is a random variable under the ZI distribution, then the probability mass function (pmf) of X is given by

$$P(X = x|\emptyset, \Theta) = \emptyset f(x) + (1-\emptyset)g(x;\, \Theta) \tag{1}$$

where

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

Let $\emptyset$ be a zero-inflation parameter, such that $0 < \emptyset < 1$, and $g(x;\, \Theta)$ is the pmf of X with a vector of the parameter, $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$.

Now, we consider a zero-inflated count model corresponding to the Poisson Binomial distribution, called zero-inflated Poisson (ZIP) distribution. Lambert [11], McLachlan and Peel [12] proposed the ZIP distribution, which is a special type of of mix of the Bernoulli distribution and Poisson distributions. From the pmf of Poisson distribution is given as $g(x;\, \lambda) = \frac{e^{-\lambda}\lambda^x}{x!};\ x = 0,1,2,\ldots,$ and substituting in eq.(1), then the pmf of ZIP distribution has the form

$$P(X = x|\emptyset, \lambda) = \begin{cases} \emptyset + (1-\emptyset)e^{-\lambda}, & x = 0 \\ (1-\emptyset)\frac{e^{-\lambda}\lambda^x}{x!}, & x = 1,2,\ldots, \end{cases} \tag{3}$$

where $\lambda = np, \lambda > 0$, $0 < \emptyset < 1$. Furthermore, the mean and the variance of the ZIP distribution are given by $\mu_{ZIP} = (1-\emptyset)\lambda$ and $\sigma^2_{ZIP} = \lambda(1-\emptyset)(1-\lambda\emptyset)$ respectively.

2.1 Design of the double acceptance sampling plan under ZIP distribution

The double acceptance sampling plan (DSP) requires the specification of four quantities which are known as its parameters. These parameters are n_1, c_1, n_2 and c_2. In a DSP, the decision of accepting or rejecting a lot is taken based on two samples.

1. The first sample: the lot is accepted if the number of defective units (d_1) in the first sample is less than the acceptance number c_1.

2. The second sample: the lot is accepted if the number of defective units $(d_1 + d_2)$ in both samples is greater than c_1 and less than or equal to the acceptance number c_2.

Therefore, if P_a^1 and P_a^2 denote the probabilities of accepting a lot on the first sample and the second sample, as shown in eq.(4) and eq.(5), respectively, then the probability of accepting a lot (P_a) of p is given by equation (6).

$$P_a^1(p) = P(d_1 \leq c_1 : n_1) \tag{4}$$

$$P_a^2(p) = P(c_1 < d_1 \leq c_2 : n_1) \times P(d_1 + d_2 \leq c_2 : n_2) \tag{5}$$

$$P_a(p) = P_a^1(p) + P_a^2(p) \tag{6}$$

In this section, the optimal DSP under the ZIP distribution (DSP_{ZIP}) is described. From eq.(3) and eq.(6), the probability of accepting a lot for ZIP distribution is given in eq.(7).

$$\begin{aligned} P_a(p) = {} & \emptyset + (1-\emptyset)e^{-\lambda_1} \\ & + \sum_{d_1=1}^{c_1} (1-\emptyset)\frac{e^{-\lambda_1}{\lambda_1}^{d_1}}{d_1!} \\ & + \sum_{d_1=c_1+1}^{c_2} \left\{ \left[(1-\emptyset)\frac{e^{-\lambda_1}{\lambda_1}^{d_1}}{d_1!} \right] \times \left[\emptyset + (1-\emptyset)e^{-\lambda_2} + \sum_{d_2=1}^{c_2-d_1} (1-\emptyset)\frac{e^{-\lambda_2}{\lambda_2}^{d_2}}{d_2!} \right] \right\} \end{aligned} \tag{7}$$

Moreover, the average sample number function (ASN) of the DSP_{ZIP} is given in eq. (8).

$$\begin{aligned} ASN &= n_1 + n_2(1 - P_I) \\ &= n_1 + n_2 \textstyle\sum_{d_1=c_1+1}^{c_2} (1-\emptyset)\frac{e^{-\lambda_1}{\lambda_1}^{d_1}}{d_1!} \end{aligned} \tag{8}$$

where P_I is the probability of deciding on the acceptance or rejection of the lot on the first sample and is given by $P_I = P(d_1 \leq c_1 : n_1) + P(d_1 > c_2 : n_1)$.

2.2 Total cost function of DSP

In this section, the total cost function of the product inspection of a lot for the DSP plan under ZI distribution is discussed. In this research, three different types of costs are considered: cost of inspection per lot, cost of internal failure per lot, and cost of an outgoing defective per lot, as presented in eq. (9).

$$\begin{aligned} TC = {} & C_I\left(n_1 P_a^1(p) + (n_1+n_2)P_a^2(p) + N\big(1 - P_a(p)\big)\right) \\ & + C_F\left((n_1+n_2)p + \big(1 - P_a(p)\big)\big(N - (n_1+n_2)\big)p\right) \\ & + C_O\big(P_a(p)\big(N - (n_1+n_2)\big)p\big). \end{aligned} \tag{9}$$

where C_I is the cost of inspection per unit, C_F is the cost of the internal failure per unit, and C_O is the cost of an outgoing defective per unit. The component of the total cost function for the inspection of a lot for the DSP under the ZIP distribution can be expressed as follows:

First-term denotes the cost of inspection per lot, where $n_1P_a^1(p) + (n_1 + n_2)P_a^2(p) + N\big(1 - P_a(p)\big)$ represents the expected number of units inspected per lot.
Second- term denotes the cost of the internal failure per lot, where $(n_1 + n_2)p + \big(1 - P_a(p)\big)\big(N - (n_1 + n_2)\big)p$ represents the expected number of defective items detected per lot.
Third- term denotes the cost of an outgoing defective per lot, where $P_a(p)\big(N - (n_1 + n_2)\big)p$ represents the expected number of defective items not detected per lot.

3. Results and Discussion

In this section, the optimal parameters of the DSP_{ZIP} are calculated to achieve the minimum value of TC and ASN when the maximum probability of accepting a lot is received. The MATLAB software (R2019b) is used to perform a simulation study of the GA with multi-objective optimization. In the optimal solution of DSP_{ZIP}, the minimum of the TC and ASN is calculated by considering the optimal values of n_1, n_2, c_1 and c_2. The constraints of the producer's risk (α) and the consumer's risk (β) are satisfied immediately by the provision of the acceptable quality level (AQL) and the lot tolerance percent defective $(LTPD)$. In practice, the constraints of α and β are satisfied immediately when AQL and $LTPD$ are provided. For the effectiveness of the proposed sampling plan, two points $(AQL, 1 - \alpha)$ and $(LTPD, \beta)$ are considered for changes on the OC curve. A manufacturer intends that the probability of acceptance of a lot of items should be greater than $1 - \alpha$ at the quality level of AQL. In reality, a customer requests that the probability of the lot acceptance should be less than β at $LTPD$. In the optimization technique, the optimal solution is considered on three objective functions simultaneously.

Multi-objective function

$$\text{Minimize} \quad TC \text{ and } ASN \tag{10}$$

$$\text{Maximize} \quad P_a(p) \tag{11}$$

Subject to:

$$P_a(AQL) \geq 1 - \alpha \text{ and } P_a(LTPD) \leq \beta,$$
$$n_1 + n_2 \leq \delta N, n_1 > 0, n_2 > 0,$$
$$c_1 \geq 0, c_2 > 0 \text{ and } c_2 > c_1 \geq 0.$$

The following input parameters are used to design the proposed illustrations. Some input parameters are assigned the same values for the proposed illustrations, that is the lot size $N = 1{,}000$, the proportion of sample size from lot size $\delta = (0.10, 0.20)$, the producer's risk $\alpha = 0.05$, and the consumer's risk $\beta = 0.01$. Furthermore, the input parameters of the cost function are given [8]: $C_I = 1, C_F = 2$, and $C_O = 10$, respectively. Some input parameters are assigned the different values for each sample illustrations as shown in Table 1, that is, the proportion of defective (p), the zero-inflation parameter $(\emptyset)$, the acceptable quality level (AQL), and the lot tolerance percent defective $(LTPD)$. Most studies presented a comparison between different values of AQL and $LTPD$ used to find the optimal ASP [4-11, 15-17].

Table 1. Input parameters used to design the proposed illustrations

Input parameter	Illustration 1	Illustration 2	Illustration 3
p	0.01	0.05	0 to 0.20
$\emptyset$	0.001,0.01,0.05,0.10	0.01,0.05,0.10,0.20,0.30, 0.40, 0.50	0.01,0.05,0.10
AQL	0.01,0.05	0.05	0.05
$LTPD$	0.05,0.075,0.10	0.10	0.10

3.1 Illustration 1: The three conditions of the ratio of sample size

From eq.(8) and eq.(9), we can see that the values of TC and ASN of the DSP_{ZIP} depend on the required sample sizes (n_1, n_2). Therefore, in this illustration, the sensitivity analysis of the required sample sizes (n_1, n_2) is considered based on two constraints of the proportion of sample size from lot size (δ) by assigned that $\delta = 0.10\ (n_1 + n_2 \le 100)$, and $\delta = 0.20\ (n_1 + n_2 \le 200)$, respectively.

Moreover, the ratio of sample size, $r = \frac{n_1}{n_2}$, is used to measure of discrimination of the DSP_{ZIP} that is concerned with the required sample sizes. There are three conditions of the ratio of sample size: $r = 1$ or $n_1 = n_2$, $r > 1$ or $n_1 > n_2$, and $r < 1$ or $n_1 < n_2$.

Depending on the above, the following multi-objective optimization problem is solved to determine the optimal plan parameters $(n_1, n_2, c_1, c_2)^*$ of the DSP_{ZIP} as per eq.(10) and eq.(11). Referring to the numerical of this illustration from Table 1, suppose $p = 0.01$ and $P_a(0.01) = 0.99$ under the different combinations of $\emptyset = (0.001,0.01,0.05,0.10), AQL = (0.01,0.05)$, and $LTPD = (0.05,0.075,0.10)$.

From Table 2, the sensitivity analyses of (n_1, n_2, c_1, c_2) under the DSP_{ZIP} are shown by considering two constraints of the required sample sizes (δ) and three conditions of the ratio of sample size (r). The investigating values are given as follows.

1. Based on two constraints of the required sample sizes with $\emptyset = 0.01$ and $AQL = 0.01$, the results show that when $LTPD$ increases the value of TC and ASN decrease. Furthermore, r further approaches to 1 when $LTPD$ increases.

2. When $\emptyset$ increases under the same values of AQL and $LTPD$, the results are not different, that is, the more $\emptyset$ increases r also approachs 1.

3. Two constraints of the required sample sizes for the same values of $\emptyset$, AQL and $LTPD$ are considered. The results indicate that the first constraint $(n_1 + n_2 \le 100)$ is given a smaller value of TC and ASN than the second constraint $(n_1 + n_2 \le 200)$ with $P_a(0.01) = 0.99$. On the other hand, the second constraint is given a smaller r than the first constraint respectively. This means that TC and ASN further decrease when the value of r approaches 1.

For both constraints of the required sample sizes, it can interpret that when $AQL = 0.05$, $LTPD = 0.10$, and $\emptyset$=0.10, the value of r approaches 1. This means that the required sample sizes should be equal $(n_1 = n_2)$ which provides the maximum value of P_a, and the minimum value of TC and ASN.

3.2 Illustration 2: The optimal solution under three different scenarios of the required sample sizes

In the general sampling system, the user expects that the smaller value of required sample sizes (n_1, n_2) or ASN would be more satisfactory for designing the optimal ASP. So, this illustration aims

Table 2. Optimal parameters of the DSP_{ZIP} for the minimum value of TC and ASN under the constraints of the required sample sizes with $p = 0.01$ and suppose that $P_a(0.01) = 0.99$

AQL	$\emptyset$	$LTPD$	$n_1 + n_2 \leq 100$							$n_1 + n_2 \leq 200$						
			n_1	n_2	c_1	c_2	r	TC	ASN	n_1	n_2	c_1	c_2	r	TC	ASN
0.01	0.001	0.05	81	19	2	8	4.26	218	81.00	145	52	3	5	2.79	269	145.65
		0.075	71	22	2	3	3.23	186	71.18	90	67	3	8	1.34	198	90.01
		0.10	55	32	2	4	1.72	157	55.07	53	60	2	4	0.88	161	53.12
	0.01	0.05	82	16	2	3	5.13	220	82.20	148	52	3	9	2.85	277	148.00
		0.075	66	18	2	3	3.67	185	66.45	98	73	3	6	1.34	198	98.00
		0.10	55	44	2	4	1.25	151	55.0	55	85	2	4	0.65	161	55.18
	0.05	0.05	81	12	2	3	6.75	216	81.45	137	26	3	4	5.27	262	137.91
		0.075	65	35	2	3	1.86	182	65.80	83	87	2	4	0.95	201	83.71
		0.10	55	43	2	3	1.28	164	55.66	63	72	2	4	0.88	164	63.23
	0.10	0.05	81	19	2	3	4.26	212	80.63	143	56	3	5	2.55	263	143.61
		0.075	63	32	2	4	1.97	201	77.12	110	76	3	5	1.45	210	110.31
		0.10	52	47	3	4	1.10	179	80.14	51	85	1	3	0.60	195	52.01
0.05	0.001	0.05	90	10	2	5	9.00	230	90.02	158	42	4	8	3.76	261	158.01
		0.075	51	42	2	4	1.21	154	51.07	81	68	4	6	1.19	174	81.01
		0.10	44	54	3	5	0.81	137	44.01	41	103	2	4	0.40	139	41.08
	0.01	0.05	89	11	2	5	8.09	228	89.02	155	45	3	5	3.44	288	155.72
		0.075	55	18	2	4	3.06	162	55.04	66	68	3	7	0.97	162	66.01
		0.10	44	46	3	6	0.96	137	44.01	43	104	3	7	0.41	136	43.00
	0.05	0.05	51	49	1	3	1.04	199	51.60	116	84	2	4	1.38	265	117.89
		0.075	52	41	2	4	1.27	155	52.10	68	41	3	5	1.66	164	68.02
		0.10	36	48	2	5	0.75	133	36.01	58	92	4	6	0.63	142	58.01
	0.10	0.05	87	13	2	5	6.69	226	87.02	109	21	2	5	5.19	268	109.08
		0.075	52	41	2	4	1.27	155	52.07	66	42	3	5	1.57	161	66.02
		0.10	35	48	2	4	0.73	131	35.02	58	75	4	6	0.77	150	58.01

to minimize the required sample sizes (n_1, n_2) under the optimal parameters of the DSP_{ZIP}. Based on this reason, the conditions of the required sample sizes (n_1, n_2) are considered as the constraints to find the optimal DSP_{ZIP} to achieve the maximum value of P_a and the minimum value of TC and ASN simultaneously.

In this illustration, a comparison between the size of the first sample (n_1) and the second sample (n_2) is considered. The smaller value of sample size (n_1, n_2) or ASN is always more satisfactory for designing the optimal ASP.

There are three different scenarios of the required sample sizes (n_1, n_2) used to find the optimal multi-objective function of a DSP_{ZIP} as follows.

Scenario1 (S1): $n_1 = n_2$ and $n_1 + n_2 \leq \delta N (r = 1)$

Scenario2 (S2): $n_1 > n_2$ and $n_1 + n_2 \leq \delta N (r > 1)$

Scenario3 (S3): $n_1 < n_2$ and $n_1 + n_2 \leq \delta N (r < 1)$

The maximum value of $P_a(p)$ and the minimum value of TC and ASN can be determined by solving eq.10 and eq.11, with a given input parameter from Table 1. Suppose $AQL = 0.05$, and $LTPD = 0.10$ based on the reasons from the Illustration 1, the optimal parameters $(n_1, n_2, c_1, c_2)^*$ of the DSP_{ZIP} are determined by satisfying two inequalities, $P_a(AQL) \geq 1 - \alpha$ and $P_a(LTPD) \leq \beta$, and three scenarios of the required sample sizes as mentioned above.

From Table 3, the main goal is achieved (the maximum P_a and the minimum TC and ASN) by obtaining the optimal values of $(n_1, n_2, c_1, c_2)^*$ under the DSP_{ZIP}. The investigating values are given as follows.

1. Based on three different scenarios of (n_1, n_2) with $\delta = 0.10(n_1 + n_2 \leq 100)$, and $\delta = 0.20(n_1 + n_2 \leq 200)$, when $\emptyset$ increases, the value of P_a tends to increase while TC tends to decrease.

2. The three different scenarios of (n_1, n_2) are compared under the same value of δ and $\emptyset$. The results show that S1, $(n_1, n_2, c_1, c_2)^* = (50{,}50{,}1{,}2)^*$, provides the lowest TC and highest P_a, whereas S3, $(n_1, n_2, c_1, c_2)^* = (40{,}60{,}0{,}2)^*$ provides the lowest ASN.

3. Considering under the same scenario, when δ increases under the same value of $\emptyset$, the value of P_a is increasing while the value of TC and ASN is decreasing.

It can interpret that the optimal required sample sizes is Scenario 1 $(n_1 = n_2)$ because S1 provides the lowest TC and highest P_a and these parameters are an important factor in the construction of the optimal DSP_{ZIP}. Moreover, the smaller of required sample sizes (lower δ) provides the optimal DSP_{ZIP} .

3.3 Illustration 3: The optimal OC function of the DSP_{ZIP}

Illustration 2 indicates that the required sample sizes of the first and second sampling should be the smallest and equal values $(n_1 = n_2)$. In this illustration, the input parameters are assigned from Table 1. The performances of the DSP_{ZIP} with different values of $\emptyset$ under the optimal scenario (S1) are presented in Figures 1-3. The OC curves of DSP_{ZIP} are shown in Figure 1 when $(n_1, n_2, c_1, c_2)^* = (50{,}50{,}1{,}2)^*$. Other than that, the OC curves of DSP_{ZIP} are compared with DSP_P . Suppose DSP_P is the DSP under traditional Poisson distribution.

From Figure 1, for the same value of p and a different value of $\emptyset$, the results indicate that a higher value of $\emptyset$ under DSP_{ZIP} provides the larger P_a. Based on the same condition of optimal $(n_1, n_2, c_1, c_2)^*$, DSP_{ZIP} gives a larger P_a than DSP_P for each value of p. In practice, an optimal sampling plan should give a small ASN. From Figure 2, it can be seen that $p = 0.04$ gives the maximum value of ASN of all the proposed sampling plans. Moreover, a larger of $\emptyset$ under DSP_{ZIP}

Table 3. The optimal solution of DSP_{ZIP} under three different scenarios of the required sample sizes

δ	$\emptyset$	S1 $(\mathbf{50,50,1,2})^*$			S2 $(\mathbf{60,40,1,2})^*$			S3 $(\mathbf{40,60,0,2})^*$		
		P_a	TC	ASN	P_a	TC	ASN	P_a	TC	ASN
0.10	0.01	0.2944	924	62.70	0.2071	977	68.87	0.1836	990	56.08
	0.05	0.3229	913	62.18	0.2392	963	68.51	0.2150	976	55.43
	0.10	0.3586	887	61.54	0.2792	942	68.07	0.2545	946	54.62
	0.20	0.4298	844	60.26	0.3593	890	67.17	0.3341	901	52.99
	0.30	0.5011	808	58.98	0.4394	847	66.27	0.4145	855	51.37
	0.40	0.5724	764	57.70	0.5195	796	65.38	0.4958	801	49.74
	0.50	***0.6436***	***721***	***56.41***	0.5996	752	64.48	0.5778	754	48.12
		$(\mathbf{100,100,2,4})^*$			$(\mathbf{120,80,2,4})^*$			$(\mathbf{80,120,1,3})^*$		
		P_a	TC	ASN	P_a	TC	ASN	P_a	TC	ASN
0.20	0.01	0.1380	1024	117.37	0.0778	1056	130.60	0.1028	1038	102.21
	0.05	0.1727	1001	116.67	0.1148	1039	130.17	0.1309	1020	102.27
	0.10	0.2160	978	115.79	0.1611	1007	129.64	0.1842	987	101.10
	0.20	0.3027	924	114.04	0.2538	957	128.57	0.2747	935	98.76
	0.30	0.3896	878	112.28	0.3466	907	127.50	0.3652	884	96.41
	0.40	0.4765	823	110.53	0.4395	857	126.42	0.4557	825	94.07
	0.50	***0.5635***	***777***	***108.77***	0.5326	799	125.35	0.5463	773	91.72

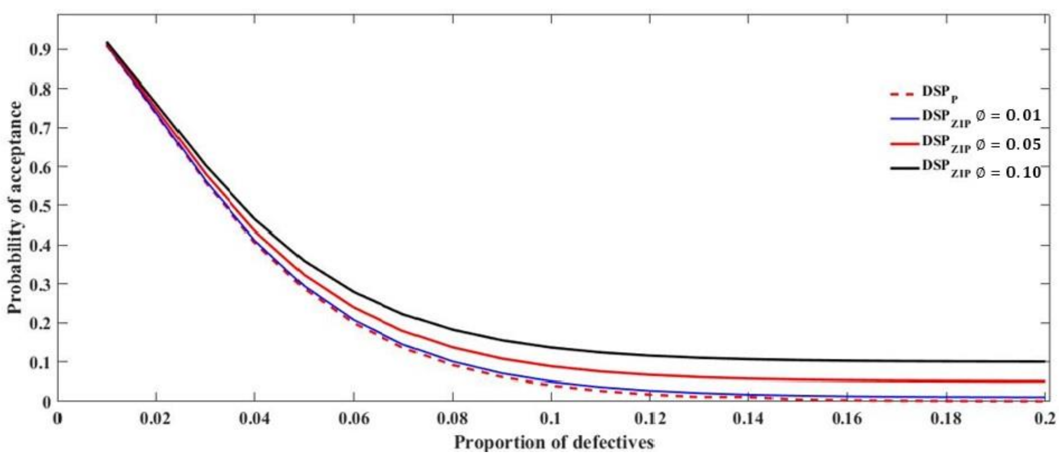

Figure 1. The optimal OC function of the DSP_{ZIP} under optimal plan $(50,50,1,2)^*$

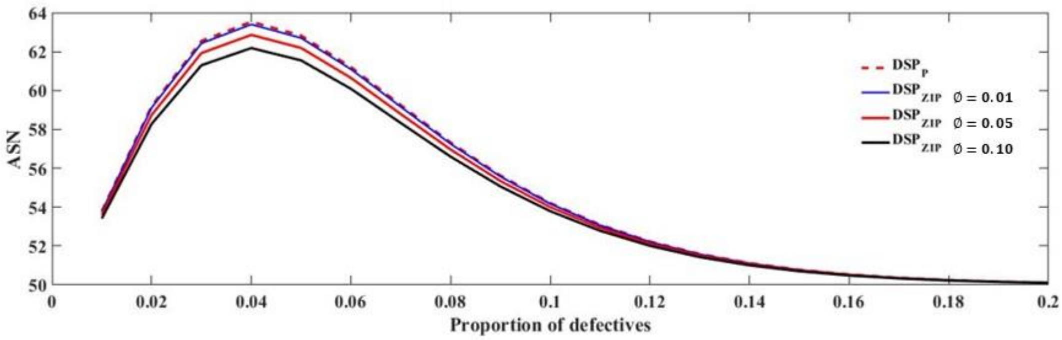

Figure 2. The optimal ASN curves of the DSP_{ZIP} under optimal plan $(50,50,1,2)^*$

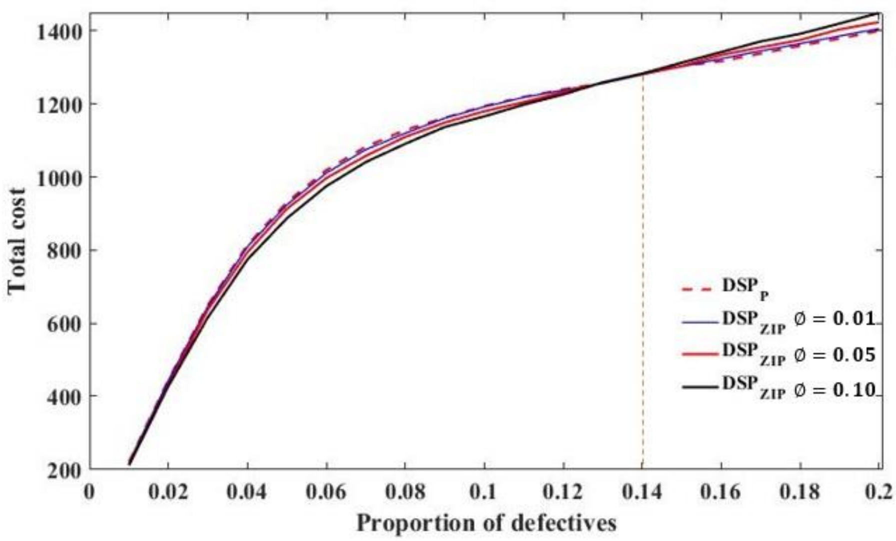

Figure 3. The optimal total cost of the DSP_{ZIP} under optimal plan $(50,50,1,2)^*$

provides smaller ASN. From Figure 3, the results show that larger values of $\emptyset$ under DSP_{ZIP} provides the lower TC for $0 < p \leq 0.13$. Furthermore, the OC function under the optimal plans $(50,50,1,2)^*$ and $(100,100,2,4)^*$ are compared in Figure 4. The result indicates that for $p = 0.01$, the optimal plan $(100,100,2,4)^*$ has a slightly larger P_a than $(50,50,1,2)^*$. On the other hand, when p increase, the optimal plans $(50,50,1,2)^*$ has a slightly larger P_a than $(100,100,2,4)^*$.

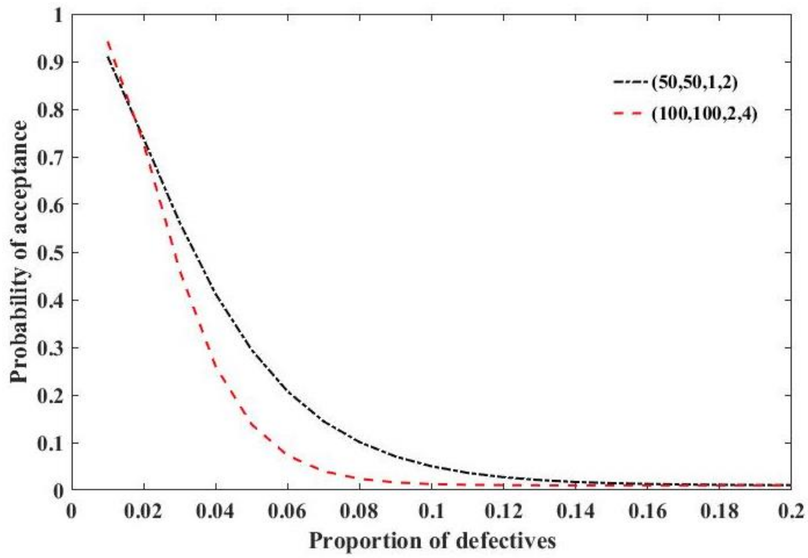

Figure 4. A comparison of the OC function under the optimal plans$(50,50,1,2)^*$and $(100,100,2,4)^*$

3.4 Real data application

In this section, a real dataset with excess zero counts is the number of read-write errors discovered in a computer hard disk in a manufacturing process (see Xie *et al*. [20]) as shown in Figure 5. This data set includes a total of 208 samples with a mean of 1.16 and a standard deviation of 1.20. From this dataset, the input parameters are calculated as follows:

$$\emptyset = \frac{Number\ of\ zero\ value\ samples}{Total\ number\ samples} = 0.87,\ p = 0.006,\ \text{and}\ \ N = 208.$$

Some of the input parameters are assigned by the user:

$\delta = (0.10{,}0.20), \alpha = 0.05,\ \beta = 0.01, AQL = 0.05, LTPD = 0.10, C_I = 1, C_F = 2,$ and $C_O = 10.$

Substituting input parameters in multi-objective optimization using GA methods, we then obtained the optimal plan parameters $(n_1, n_2, c_1, c_2)^*$.The result shows that at $\delta = 0.10$ $(n_1 + n_2 \leq 20)$, the optimal plan parameters are $(10{,}10{,}0{,}2)$ which provide the optimal solution $P_a = 0.9917$, $TC = 37$, and $ASN = 10.27$. However, for $\delta = 0.20$ $(n_1 + n_2 \leq 40)$, the optimal plan parameters are $(20{,}20{,}0{,}1)$, which provide an optimal solution $P_a = 0.9911, TC = 47$, and $ASN = 20.95$.

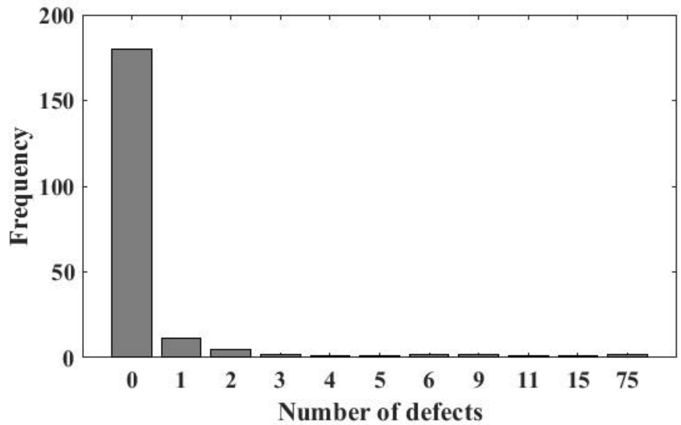

Figure 5. Number of read-write errors discovered in a computer hard disk

4. Conclusions

Nowadays, It can be observed that when the production processes are well inspected, zero defects are often found in sample inspections. There are many ways to achieve the optimal DSP that is affected by zero-inflated data. In this research, the proposed method was modified to make an optimal decision for the manufacturer. The optimal parameters were proposed to maximize the $P_a(p)$ and minimize the TC and ASN simultaneously. The proposed method was designed based on three different scenarios of the required sample sizes using multi-objective optimization with GA methods.

In conclusion, according to the proposed method, the result indicates that the values of $P_a(p)$, TC and ASN under the DSP_{ZIP} depend on the required sample sizes. Based on the same value of AQL and $LTPD$, when $\emptyset$ increases, the smaller value of required sample sizes provides the optimal DSP_{ZIP} to achieve maximum value of the $P_a(p)$ and minimum value of the TC and ASN. Furthermore, under 3 different scenarios of the required sample sizes, we found that the first sample (n_1) and the second sample (n_2) should be equal $(n_1 = n_2)$. The performance of the DSP_{ZIP} is considered by the OC function with a different value of $\emptyset$ based on $n_1 = n_2$ scenario. The results indicate that the DSP_{ZIP} provides more performance when $\emptyset$ is increased. Furthermore, under the same condition, DSP_{ZIP} provides more performance than the DSP_P in each value of p.

To apply the proposed methods, the manufacturer should know some necessary values of input parameters such as lot size, the proportion of defect per lot, cost per unit, etc. In future work, the proposed method can be applied to optimize multiple acceptance sampling plans. Moreover, the proposed method can be extended under other zero-inflated distributions such as the zero-inflated Negative binomial distribution.

5. Acknowledgements

The first author is deeply thankful to the Division of Mathematics, Faculty of Science and Technology, RMUTK. The second author thanks the Division of Applied Statistics, Faculty of Science and Technology, RMUTT for their support.

References

[1] Montgomery, D.C., 2005. *Introduction to Statistical Quality Control*. New York: John Wiley and Sons.

[2] Duarte,B.P.M. and Saraiva, P.M., 2008. An optimization-based approach for designing attribute acceptance sampling plans. *International Journal of Quality and Reliability Management*, 25, 824-841.

[3] Kaya, I., 2009. A genetic algorithm approach to determine the sample size for attribute control charts. *Information Sciences*,179, 1552-1566.

[4] Kobilinsky, A. and Bertheau, Y., 2005. Minimum cost acceptance sampling plans for grain control, with application to GMO detection. *Chemometrics and Intelligent Laboratory Systems*, 75(2), 189-200.

[5] Cheng, T.M. and Chen, Y.L., 2006. A GA mechanism for optimizing the design of attribute-double-sampling-plan. *Automation in Construction*, 16(3), 345-353.

[6] Sampath, S. and Deepa, S.P., 2012. Determination of optimal double sampling plan using Genetic Algorithm. *Pakistan Journal of Statistics and Operation Research*, 8(2), 195-203.

[7] Braimah, O.J., Saheed, Y.K., Owonipa, R.O. and Adegbite, I.O., 2015. Economic reliability acceptance sampling plan design with zero acceptance. *African Journal of Computing and ICT*, 8(3), 69-84.

[8] Hsu, L.F. and Hsu, J.T., 2012. Economic design of acceptance sampling plans in a two-stage supply chain. *Advances in Decision Sciences*, 2012, https://doi.org/10.1155/2012/359082

[9] Fallahnezhad, M.S. and Aslam, M., 2013. A new economical design of acceptance sampling models using Bayesian inference. *Accreditation and Quality Assurance*,18(3), 187-195.

[10] Fallahnezhad, M.S. and Qazvini, E., 2017. A new economical scheme of acceptance sampling plan in a two-stage approach based on the maxima nomination sampling technique. Published online in *Transactions of the Institute of Measurement and Control*, 39(7), https://doi.org/ 10.1177/0142331216629203

[11] Fallahnezhad, M.S., Qazvini, E. and Abessi, M., 2018. Designing an economical acceptance sampling plan inthe presence of inspection errors based on maxima nomination sampling method. *Scientia Iranica*, 25(3), 1701-1711.

[12] Lambert, D., 1992. Zero-inflated Poisson regression, with application to defects in manufacturing. *Technometrics*, 34, 1-14.

[13] McLachlan, G. and Peel, D., 2000. *Finite Mixture Models*. New York: John Wiley and Sons.

[14] Loganathan, A. and Shalini, K., 2014. Determination of single sampling plans by attributes under the conditions of zero-inflated Poisson distribution. *Communication in Statistics-Simulation and Computation*, 43, 538-548.

[15] Uma, G. and 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.

[16] Rao, G.S. and 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, 1814-1824.

[17] Wang, F.K. and Hailemariam, S.S., 2018. Sampling plans for the zero-inflated Poisson distribution in the food industry. *Food Control*, 85, 359-368.

[18] Holland, J.H., 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

[19] The Math Works^{TM}, 9.7.0 (R2019b), License Number 40791029.

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