Outlier identification and robust parameter estimation in a zero-inflated Poisson model

Article *in* Journal of Applied Statistics · February 2011

DOI: 10.1080/02664760903456426 · Source: RePEc

3 authors, including:

Jun Yang
Beihang University (BUAA)

Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Outlier identification and robust parameter estimation in a zero-inflated Poisson model

Jun Yang [a] , Min Xie [b] & Thong Ngee Goh [b]
[a] Department of System Engineering of Engineering Technology, BeiHang University, Beijing, 100191, People's Republic of China
[b] Department of Industrial and Systems Engineering, National University of Singapore, Singapore, 119260, Singapore

Available online: 27 Oct 2010

To cite this article: Jun Yang, Min Xie & Thong Ngee Goh (2011): Outlier identification and robust parameter estimation in a zero-inflated Poisson model, Journal of Applied Statistics, 38:2, 421-430

To link to this article: http://dx.doi.org/10.1080/02664760903456426

Journal of Applied Statistics
Vol. 38, No. 2, February 2011, 421–430

Outlier identification and robust parameter estimation in a zero-inflated Poisson model

Jun Yang[a]*, Min Xie[b] and Thong Ngee Goh[b]

[a] *Department of System Engineering of Engineering Technology, BeiHang University, Beijing 100191, People's Republic of China;* [b] *Department of Industrial and Systems Engineering, National University of Singapore, Singapore 119260, Singapore*

(*Received 28 May 2009; final version received 12 October 2009*)

The Zero-inflated Poisson distribution has been used in the modeling of count data in different contexts. This model tends to be influenced by outliers because of the excessive occurrence of zeroes, thus outlier identification and robust parameter estimation are important for such distribution. Some outlier identification methods are studied in this paper, and their applications and results are also presented with an example. To eliminate the effect of outliers, two robust parameter estimates are proposed based on the trimmed mean and the Winsorized mean. Simulation results show the robustness of our proposed parameter estimates.

Keywords: zero-inflated Poisson distribution; outlier identification; robust estimation; trimmed mean; Winsorized mean

1. Introduction

The zero-inflated Poisson (ZIP) model, as a popular modeling method for count data with excessive zeroes, has been studied extensively in the wide literature. Although the ZIP model without covariates has long been discussed [2,5,6], the first systematic analysis of this model was developed by Lambert [7]. Ever since then the ZIP model has been widely used in quality control and many other applications [1,9,12,14,15].

The research on the statistical aspects of the ZIP model has also been extensive. Van den Broek [11] proposed a score test for zero-inflation in a Poisson distribution. It is further extended to the more general situation by Jansakul and Hinde [4], where the zero probability is allowed to depend on covariates. Thas and Rayer [10] suggested three smooth goodness-of-fit tests for testing the ZIP distribution against general smooth alternatives in the sense of Neyman. Xiang *et al.* [13] proposed a score test for zero-inflation in correlated count data. Moghimbeigi *et al.* [8] suggested a score test for assessing the ZIP regression against Poisson regression in multilevel count data with excess zeroes. On the other hand, there has been very little research on outlier identification and robust parameter estimation, which is an important issue in using this type of models.

*Corresponding author. Email: tomyj2001@buaa.edu.cn

ISSN 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/02664760903456426
http://www.informaworld.com

As pointed by Davis and Adams [3], outliers are troublesome to practitioners. The ZIP distribution can be easily and seriously influenced by outliers as there are often a large number of zeroes; thus, outlier identification and robust parameter estimation are very important for the effective application of the ZIP model.

In this paper, we will discuss the problem of outlier identification and provide two robust parameter estimates. The ZIP model and its parameter estimation are first introduced in Section 2. In Section 3, some outlier identification methods are presented and their importance is shown by an example based on the read–write errors data used in Xie and Goh [14]. Two robust estimates of parameters are obtained from the trimmed mean and the Winsorized mean in Section 4. Some simulation studies are given in Section 5. Finally, some concluding remarks are given in Section 6.

2. Zero-inflated Poisson model and parameter estimation problem

The ZIP distribution is a generalization of Poisson distribution, and its probability mass function is defined as follows:

$$f(x;p,\lambda)=\begin{cases}1-p+p\,\mathrm{e}^{-\lambda}, & \text{if } x=0,\\ p\left(\dfrac{\lambda^x}{x!}\right)\mathrm{e}^{-\lambda}, & \text{if } x>0.\end{cases} \tag{1}$$

For the ZIP distribution, we have $E(X)=p\lambda$ and $\mathrm{Var}(X)=p\lambda+p\lambda^2-p^2\lambda^2$.

Consider a set of observations $\{x_1,x_2,\ldots,x_n\}$ from the ZIP distribution with sample size n. The maximum-likelihood estimates (MLEs) of parameters p and λ are given as follows [15]:

$$\begin{aligned}\hat{p}&=\frac{1-n_0/n}{1-\exp(-\hat{\lambda})},\\ \hat{\lambda}&=\frac{\bar{x}}{\hat{p}},\end{aligned} \tag{2}$$

where n_0 is the number of zero values in the sample and $\bar{x}=(1/n)\sum_{i=1}^{n}x_i$.

Since n_0/n can be the estimate of the occurring probability of zero $P_0=1-p+p\,\mathrm{e}^{-\lambda}$, and $\bar{x}$ can be the estimate of the first moment $E(X)=p\lambda$, one kind of linear estimates of parameters p and λ can be the solution to the following equations:

$$\begin{aligned}1-\tilde{p}+\tilde{p}\,\mathrm{e}^{-\tilde{\lambda}}&=\frac{n_0}{n},\\ \tilde{p}\tilde{\lambda}&=\bar{x},\end{aligned} \tag{3}$$

and they are just the same as the MLEs of parameters p and λ.

3. Outlier identification

An outlier is an observation that is numerically distant from the majority of the sample. Outliers can play havoc with standard statistical methods since statistics derived from one sample with outliers may be misleading. For example, the sample mean $\bar{x}$ will be upset completely by a single outlier, if any data value $x_i\to\pm\infty$, then $\bar{x}\to\pm\infty$. The effect of outliers is obvious and serious for the ZIP distribution, since there are usually many zeroes.

There is no rigid mathematical definition for outliers, and determining an observation is whether or not an outlier is ultimately a subjective exercise. There are many outlier identification methods for normal distribution, such as Chauvenet's criterion, Grubbs' test and Pirce's criterion.

These methods are not suitable for the ZIP distribution since it is non-normal and asymmetrical. From the traditional exploratory data analysis, histogram and boxplot are applied to identify outliers in the following manner:

(1) Plot the histogram of the sample, where the vertical axis stands for the frequency number and the horizon axis represents the value of the random variable. If some data are judged to be distant from the rest of the sample, these data will be roughly recognized as outliers.
(2) Plot a boxplot of the positive part of the sample to determine ‘mild’ and ‘extreme’ outliers, since excessive zeroes may lead to $Q_1 = Q_3 = \text{IQR} = 0$ and the boxplot of the whole sample will have no effect, where Q_1 and Q_3 denote the lower quartile and the higher quartile, respectively, and IQR means the inter-quartile range and $\text{IQR} = Q_3 - Q_1$. Mild outliers are data points lower than Q_1 or higher than Q_3 from 1.5 IQR to 3 IQR, and extreme outliers are data points lower than the low quartile Q_1 or higher than the high quartile Q_3 by more than three IQR. Only extreme outliers are treated in this paper as the ZIP distribution is far from the normal distributions.
(3) The outlier percentage α of the data can be determined from the above two steps. For the ZIP distribution, the outlier can be identified from the histogram and boxplot only when it is larger than the high quartile Q_3 by more than three IQR. Then, we count the values with this characteristic in the data and set the proportion with this characteristic as $\alpha/2$, and the outlier percentage α is given.

Example 1 An example is analysed here to illustrate the outlier identification and the large influence of outliers. The data set in Table 1 shows the read–write errors discovered in a computer hard disk in a manufacturing process [14].

From the histogram of this sample illustrated in Figure 1, it is probably that 75, 75, 15, 11, 9, 9 are outliers, and from the boxplot of the positive part of this sample illustrated in Figure 2, 15 and two 75 are outliers. Thus 15 and two 75 are outliers in this sample.

The MLEs of parameters p and λ obtained from the original data are $\hat{p} = 0.1346$ and $\hat{\lambda} = 8.6438$ [15], and the ZIP distribution is given by

$$f(x) = \begin{cases} 1 - 0.1346 + 0.1346\,\mathrm{e}^{-8.6438}, & \text{if } x = 0, \\ 0.1346 \dfrac{8.6413^x}{x!} \mathrm{e}^{-8.6438}, & \text{if } x > 0. \end{cases} \tag{4}$$

If the three outliers are deleted from the original data, the MLEs will be $\hat{p}_d = 0.1289$ and $\hat{\lambda}_d = 2.9140$.

Table 1. An actual set of data from a read–write error test for a certain computer hard disk.

0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	6	0	9
11	0	1	2	0	0	0	0	0	0	0	0	3	3	0	0	5	0	15	6
0	0	0	4	2	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
75	0	0	0	0	75	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0
0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
0	0	1	0	0	0	0	0												

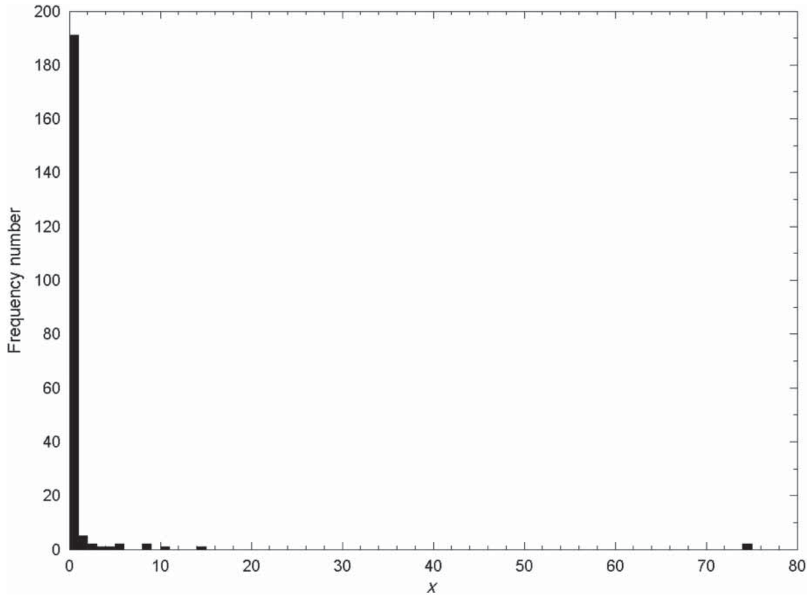

Figure 1. Histogram of the read–write error data.

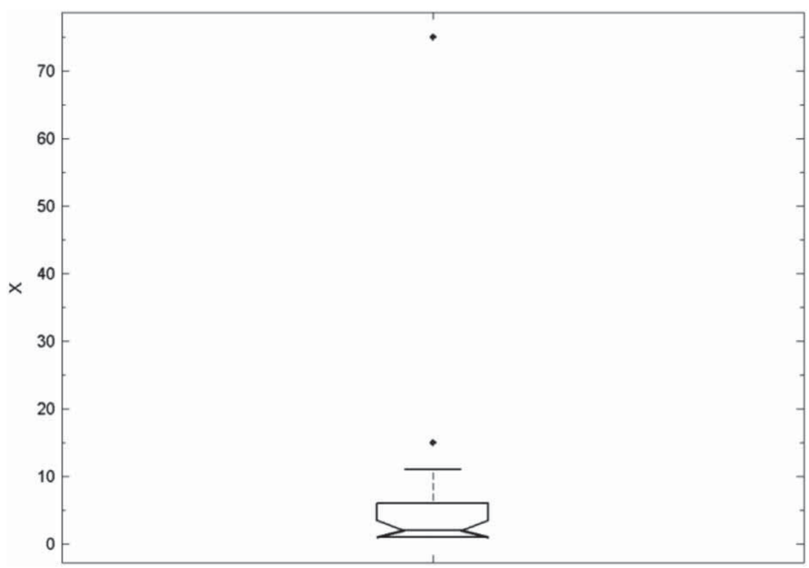

Figure 2. Boxplot of the positive part of the read–write error data.

For the Poisson distribution with parameter λ in the ZIP distribution, a c-chart can be used to monitor the process as in [16], and the control limits are as follows:

$$\mathrm{UCL} = \lambda + 3\sqrt{\lambda}, \tag{5}$$

$$\mathrm{CL} = \lambda, \tag{6}$$

$$\mathrm{LCL} = \lambda + 3\sqrt{\lambda}, \tag{7}$$

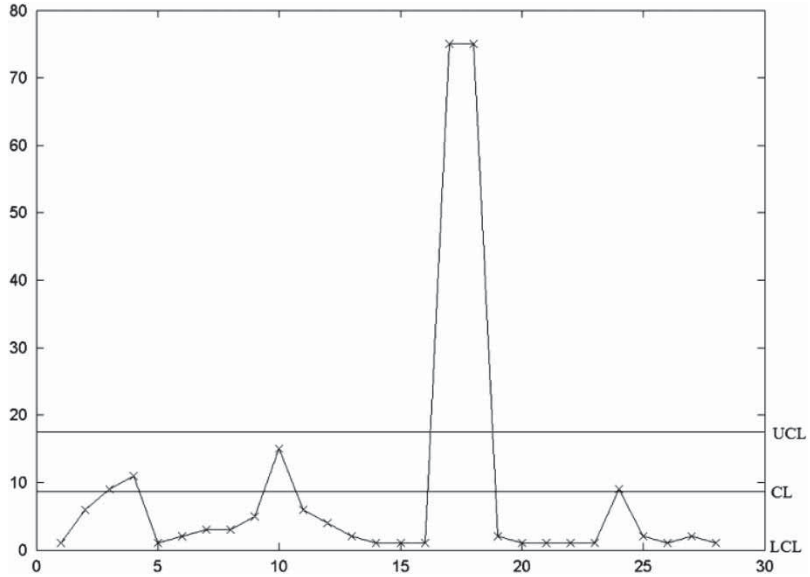

Figure 3. Control chart for the number of read–write error [16].

where CL, UCL, LCL denote the center line, the upper control limit and the lower control limit, respectively.

Then, if λ is set equal to 8.6438, the lower control limit and the upper control limit of the c-chart are 0 and 17.4639, respectively, and only two points with value 75 are out of control, which can be seen in Figure 3. Furthermore, most data points are below the CL, which is very abnormal. From the knowledge of the outliers, it can be seen that the basic reason of this phenomenon is that the LCL, CL and UCL have been distorted by those outliers.

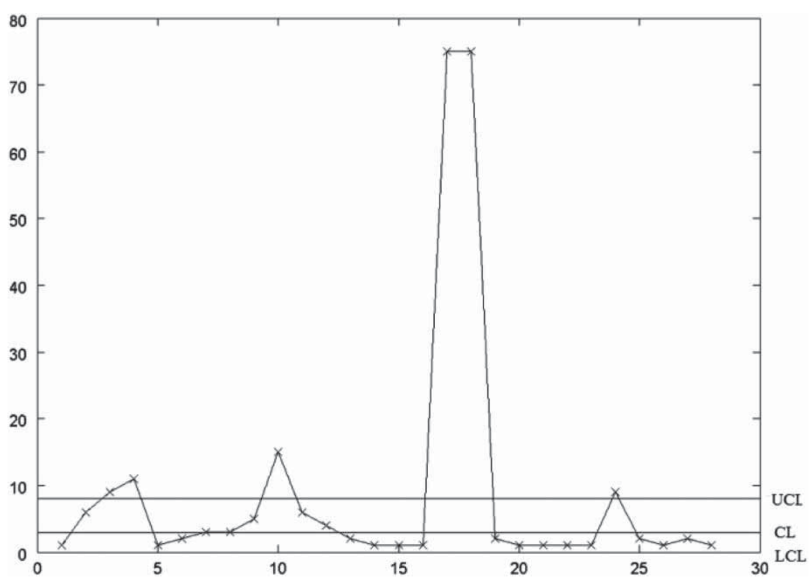

Figure 4. Control chart for the number of read–write error after discarding the outliers.

After discarding the three outliers, λ is set equal to 2.9140, whereas the lower control limit and the upper control limit of the c-chart are now 0 and 8.0351, respectively. It can be seen from Figure 4 that there are six out-of-control points, namely 9, 9, 11, 15, 75, 75, and 16 points are below the CL and 12 points are above the CL. It can be seen from Figures 3 and 4 that the influence of outliers on control chart can be large; thus, the outlier identification is very important for the ZIP model.

4. Two robust parameter estimates

Since the sample mean $\bar{x}$ is sensitive to outliers, if some robust statistics analogous to the sample mean are taken from the place of the sample mean in Equation (2), some robust estimates of parameters p and λ can be obtained. As is well known, trimmed estimators and Winsorized estimators are general methods to make statistics more robust; thus, robust parameter estimates of the ZIP model will be given by using the trimmed mean or the Winsorized mean instead of the sample mean in Equation (2).

4.1 *Using the α-trimmed mean*

The robust estimates of parameters p and λ are obtained by the MLEs and the α-trimmed mean will be the solution to the following equations:

$$\begin{aligned} 1 - \hat{p}_t + \hat{p}_t\, \mathrm{e}^{-\hat{\lambda}_t} &= \frac{n_0}{n}, \\ \hat{p}_t \hat{\lambda}_t &= \bar{x}_{t,\alpha}, \end{aligned} \tag{8}$$

where $\bar{x}_{t,\alpha}$ denotes the α-trimmed mean, that is, the mean of the central $1-\alpha$ part of the distribution and those $\alpha n/2$ observations are removed from each end. Obviously, the 0-trimmed mean gives the mean and the 1-trimmed mean gives the median. Since the breakdown point of the α-trimmed mean is $\alpha/2$, the α-trimmed mean will be more robust with the increase of α value, and the mean can be distorted by one single outlier as its breakdown point is 0.

4.2 *Using the Winsorized mean*

Although the trimmed mean is robust, simply discarding observation values will lose some information and may lead to some distortion. As is well known, Winsorized estimators are usually more robust to outliers than their unwinsorized counterparts, thus the Winsorized mean will be a good alternative for the sample mean. In the Winsorized mean, outliers will instead be replaced by certain percentiles (the trimmed minimum and maximum). The robust estimates of parameters p and λ obtained by the MLEs and the Winsorized mean will be the solution to the following equations:

$$\begin{aligned} 1 - \hat{p}_w + \hat{p}_w\, \mathrm{e}^{-\hat{\lambda}_w} &= \frac{n_0}{n}, \\ \hat{p}_w \hat{\lambda}_w &= \bar{x}_{w,\alpha}, \end{aligned} \tag{9}$$

where $\bar{x}_{w,\alpha}$ is the $(1-\alpha)$-Winsorized mean, for example, a 90% Winsorization would see all the data below the 5th percentile set to the 5th percentile, and data above the 95th percentile set to the 95th percentile.

Example 2 (Example 1 continued) The value α is set to be equal to 3% from outlier identification. In Example 1, we have $\bar{x}_{t,\alpha} = 0.3812$ and $\bar{x}_{w,\alpha} = 0.5288$.

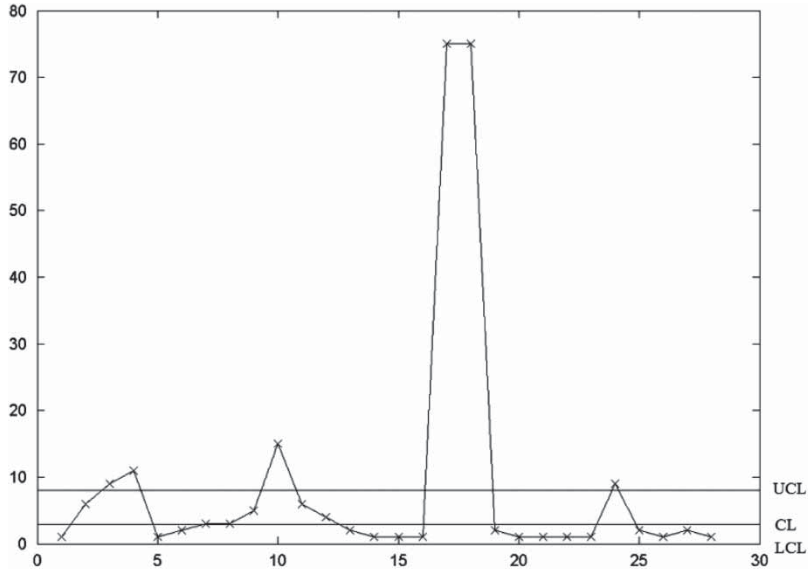

Figure 5. Control chart for the number of read–write error with the 3%-trimmed mean.

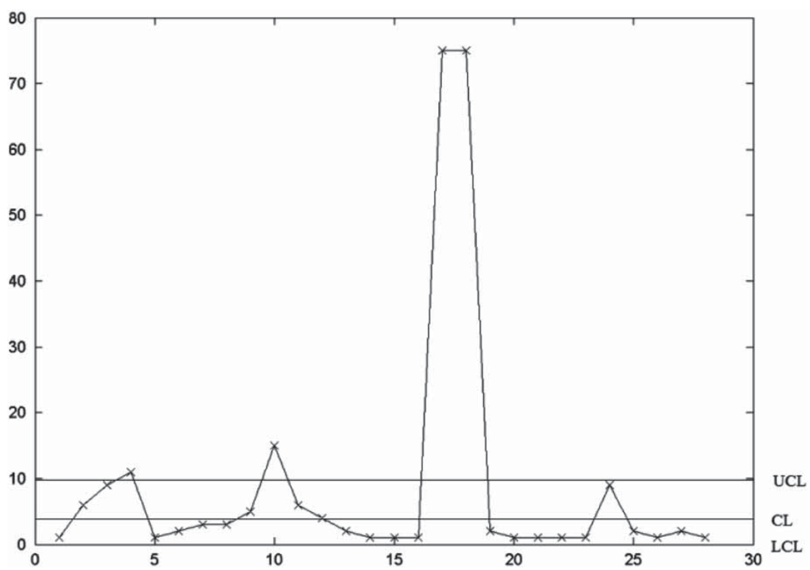

Figure 6. Control chart for the number of read–write error with the 3%-Winsorized mean.

Substitute $\bar{x}_{t,\alpha} = 0.3812$ into Equation (8) and solve it, then we have $\hat{p}_t = 0.1309$ and $\hat{\lambda}_t = 2.9121$. If we substitute $\hat{\lambda}_t$ into Equations (5) and (7) again, we have LCL $= 0$, UCL $= 8.0316$ and six out-of-control points, namely 9, 9, 11, 15, 75, 75, as illustrated in Figure 5.

Substitute $\bar{x}_{w,\alpha} = 0.5288$ into Equation (9) and solve it, then we have $\hat{p}_w = 0.1376$ and $\hat{\lambda}_w = 3.8430$. If we substitute $\hat{\lambda}_w$ into Equations (5) and (7) again, we have LCL $= 0$, UCL $= 9.7241$ and four out-of-control points, namely 11, 15, 75, 75, as illustrated in Figure 6.

5. Monte Carlo simulations

In this section, we investigate the robustness of the proposed two parameter estimates when the postulated model is a ZIP distribution but the data are contaminated by some outliers.

Suppose that the model of interest is a ZIP distribution with parameters (p, λ), but the observed data are contaminated by some data values from Poisson distribution $P(\Lambda)$ with probability α. Thus, the contaminated data can be seen as one generalized zero-inflated Poisson distribution (GZIP) as follows:

$$P(X = 0) = 1 - p_1 + p_2 + p_1\,\mathrm{e}^{-\lambda} + p_2\,\mathrm{e}^{-\Lambda},$$
$$P(X = x) = p_1\frac{\lambda^x}{x!}\mathrm{e}^{-\lambda} + p_2\frac{\Lambda^x}{x!}\mathrm{e}^{-\Lambda}, \quad x > 0, \tag{10}$$

where $p_2 = \alpha$, $p_1 = p - \alpha$.

In our simulation studies, we consider two scenarios. In both scenarios, we set the real underlying ZIP data with parameters $p = 0.8$ and $\lambda = 5$. In the first scenario, we set the contaminated data are sampled from GZIP$(0.8 - \alpha, \alpha, 5, 100)$ and $\alpha = 0.03, 0.04, 0.05, 0.06, 0.07, 0.08$. In the second scenario, we set the contaminated data are sampled from GZIP$(0.8 - 0.05, 0.05, 5, \Lambda)$ and $\Lambda = 20, 40, 60, 80, 100$.

For each of the above set of parameter values in both scenarios, 2000 replicates with sample size $n = 100$ are generated and the parameters of the ZIP distribution are estimated by the MLE and the proposed two estimates, respectively. The performance comparison of these three kinds of parameter estimates with varied contaminated probability and varied contaminated Poisson distribution are listed in Tables 2 and 3, respectively, where the outlier percentage α in each replicates is determined by Step (3) in Section 3. For convenience, the MLEs modified by the trimmed mean and the Winsorized mean are referred to as Trim-MLE and Wins-MLE, respectively.

In Table 2, the mean of the MLE of parameter λ is far from the real parameter value of $\lambda = 5$ and its distance will become larger with the increasing of contaminated probability α. Compared with the mean of the MLE of λ, the means of the proposed two estimates of parameter λ are

Table 2. Comparison of the three kinds of estimates with varied contaminated probability.

Contaminated Poisson distribution	Contaminated probability α	Contaminated data distribution	Estimation method	Mean of $\hat{p}$	MSE of $\hat{p}$	Mean of $\hat{\lambda}$	MSE of $\hat{\lambda}$
$P(100)$	0.03	GZIP(0.77,0.03,5,100)	MLE	0.795	0.000074	8.603	13.085
			Trim-MLE	0.801	0.000052	4.952	0.0688
			Wins-MLE	0.799	0.000051	5.229	0.1251
$P(100)$	0.04	GZIP(0.76,0.04,5,100)	MLE	0.795	0.000077	9.823	23.384
			Trim-MLE	0.801	0.000054	4.952	0.0681
			Wins-MLE	0.799	0.000054	5.321	0.1790
$P(100)$	0.05	GZIP(0.75,0.05,5,100)	MLE	0.795	0.000068	10.99	36.110
			Trim-MLE	0.801	0.000051	4.939	0.0672
			Wins-MLE	0.799	0.000049	5.400	0.2362
$P(100)$	0.06	GZIP(0.74,0.06,5,100)	MLE	0.795	0.000071	12.20	51.953
			Trim-MLE	0.801	0.000054	4.909	0.0835
			Wins-MLE	0.799	0.000052	5.466	0.3092
$P(100)$	0.07	GZIP(0.73,0.07,5,100)	MLE	0.795	0.000068	13.41	70.978
			Trim-MLE	0.801	0.000048	4.901	0.0778
			Wins-MLE	0.798	0.000048	5.545	0.3843

Note: The real data follow ZIP(0.8, 5).

Table 3. Comparison of the three kinds of estimates with varied contaminated Poisson distribution.

Contaminated Poisson distribution	Contaminated probability α	Contaminated data distribution	Estimation method	Mean of $\hat{p}$	MSE of $\hat{p}$	Mean of $\hat{\lambda}$	MSE of $\hat{\lambda}$
$P(20)$	0.05	GZIP(0.75,0.05,5,20)	MLE	0.797	0.000055	5.950	0.9729
			Trim-MLE	0.801	0.000051	4.916	0.0725
			Wins-MLE	0.799	0.000050	5.371	0.2160
$P(40)$	0.05	GZIP(0.75,0.05,5,40)	MLE	0.796	0.000066	7.226	5.040
			Trim-MLE	0.801	0.000050	4.930	0.0688
			Wins-MLE	0.799	0.000050	5.392	0.2300
$P(60)$	0.05	GZIP(0.75,0.05,5,60)	MLE	0.796	0.000068	8.498	12.332
			Trim-MLE	0.801	0.000052	4.942	0.0676
			Wins-MLE	0.799	0.000050	5.408	0.2423
$P(80)$	0.05	GZIP(0.75,0.05,5,80)	MLE	0.795	0.000077	9.752	22.710
			Trim-MLE	0.801	0.000054	4.936	0.0710
			Wins-MLE	0.799	0.000054	5.397	0.2375
$P(100)$	0.05	GZIP(0.75,0.05,5,100)	MLE	0.795	0.000068	10.99	36.110
			Trim-MLE	0.801	0.000051	4.939	0.0672
			Wins-MLE	0.799	0.000049	5.400	0.2362

Note: The real data follow ZIP(0.8,5).

nearer to the real parameter value, and their MSEs are much smaller than that of the MLE, and furthermore, the MSEs of Trim-MLE are smaller than the Wins-MLE.

In Table 3, the mean of the MLE of parameter λ is also far from the real value of $\lambda = 5$ and their distance becomes larger with the increasing of the mean Λ of the contaminated Poisson distribution. The MSEs of the MLE of parameter λ are much larger than that of the other two ones, and the MSEs of Trim-MLE are still smaller than the Wins-MLE. Even for parameter p whose MLE is robust, the MSEs of the MLE are also larger than the other two ones in Tables 2 and 3. In general, the proposed two estimates perform much better than the MLE, especially when Λ or α is relatively large, which show that the proposed two parameter estimates are more robust than the traditional MLEs. In additional, it seems that the trim-MLE is better than the Wins-MLE.

6. Conclusions

ZIP distribution is often influenced by outliers, thus a necessary step for data analysis is to reduce the outliers influence. Histogram and boxplot can be used for outlier identification, and boxplot is only applied to the positive part of the sample, since the probability of zero is usually high in the ZIP distribution.

Two robust parameter estimates of the ZIP model are given by the modification of the MLEs through the use of the trimmed mean and the Winsorized mean instead of the sample mean; the αvalue of these two means can be obtained from the earlier outlier identification. Simulation results show that the proposed two parameter estimates are more robust than the traditional MLE.

Acknowledgements

This research is supported by NUS research grant No. R-266-000-044-112 and the work was done while the first author was with the National University of Singapore as a research fellow. The authors would like to thank the three anonymous referees for their comments on an earlier version of this paper.

References

[1] A.K. Bohara and R.G. Krieg, *A zero-inflated Poisson model of migration frequency*, Int. Reg. Sci. Rev. 19(3) (1996), pp. 211–222.

[2] A.C. Cohen, *Estimation in Mixtures of Discrete Distributions*, Proceedings of the International Symposium on Discrete Distributions, Montreal, Pergamon Press, New York, 1963, pp. 373–378.

[3] C.M. Davis and B.M. Adams, *Robust monitoring of contaminated data*, J. Qual. Technol. 37(2) (2005), pp. 163–174.

[4] N. Jansakul and J.P. Hinde, *Score tests for zero-inflated Poisson models*, Comput. Stat. Data Anal. 40(1) (2002), pp. 75–96.

[5] N.L. Johnson and S. Kotz, *Distributions in Statistics: Discrete Distributions*, Houghton Mifflin, Boston, 1969.

[6] N. Johnson, S. Kotz, and A.W. Kemp, *Univariate Discrete Distributions*, 2nd ed., Wiley, New York, 1992.

[7] D. Lambert, *Zero-inflated Poisson regression with an application to defects in manufacturing*, Technometrics 34(1) (1992), pp. 1–14.

[8] A. Moghimbeigi, M.R. Eshraghian, K. Mohammad, and B. Mcardle, *Multilevel zero-inflated negative binomial regression modeling for over-dispersed count data with extra zeros*, J. Appl. Stat. 35(10) (2008), pp. 1193–1202.

[9] C.H. Sim and M.H. Lim, *Attribute charts for zero-inflated processes*, Commun. Stat. Simul. Comput. 37 (2008), pp. 1440–1452.

[10] O. Thas and J.C.W. Rayner, *Smooth tests for the zero-inflated Poisson distribution*, Biometrics 61(3) (2005), pp. 808–815.

[11] J. Van den Broek, *A score test for zero-inflation in a Poisson distribution*, Biometrics 51(2) (1996), pp. 738–743.

[12] A.H. Welsh, R.B. Cunningham, C.F. Donnelly, and D.B. Lindenmayer, *Modeling the abundance of rare species: Statistical models for counts with extra zeroes*, Ecol. Model. 88 (1996), pp. 297–308.

[13] L.M. Xiang, A.H. Lee, K.K.W. Yau, and G.J. Mclachlan, *A score test for zero-inflation in correlated count data*, Stat. Med. 25(10) (2006), pp. 1660–1671.

[14] M. Xie and T.N. Goh, *SPC of a near zero-defected process subject to random shock*, Qual. Reliab. Eng. Int. 9 (1993), pp. 89–93.

[15] M. Xie, B. He, and T.N. Goh, *Zero-inflated Poisson model in statistical process control*, Comput. Stat. Data Anal. 38 (2001), pp. 191–201.

[16] W. Xie, M. Xie, and T.N. Goh, *Control charts for processes subject to random shocks*, Qual. Reliab. Eng. Int. 11 (1995), pp. 355–360.