

Discrimination between Some Over Dispersed Count Distributions

Yook-Ngor Phang^{1*}, Seng-Huat Ong² and Yeh-Ching Low^{3*}

¹*Universiti Teknologi Mara, Cawangan Melaka, Jalan Lendu, 78000 Alor Gajah, Melaka, Malaysia*

²*Department of Actuarial Science and Applied Statistics, UCSI University, 56000 Kuala Lumpur, Malaysia*

³*Department of Computing and Information Systems, Sunway University, 47500 Selangor, Malaysia*

The Poisson inverse Gaussian and generalized Poisson distributions are widely used in modelling overdispersed count data which are commonly found in healthcare, insurance, engineering, econometric and ecology. The inverse trinomial distribution is a relatively new count distribution arising from a one-dimensional random walk model (Shimizu & Yanagimoto, 1991). The Poisson inverse Gaussian distribution is a popular count model that has been proposed as an alternative to the negative binomial distribution. The inverse trinomial and generalized Poisson models possess a common characteristic of having a cubic variance function, while the Poisson inverse Gaussian has a quadratic variance function. The nature of the variance function seems to be an important property in modelling overdispersed count data. Hence it is of interest to be able to select among the three models in practical applications. This paper considers discrimination of three models based on the likelihood ratio statistic and computes via Monte Carlo simulation the probability of correct selection.

Keywords: Poisson inverse Gaussian distribution; generalized Poisson distribution; inverse trinomial distribution; overdispersed count data; Monte Carlo simulation; probability of correct selection; variance function

I. INTRODUCTION

The generalized Poisson and inverse trinomial distributions are cubic variance distributions belonging respectively to the Abel and Takac family of distributions (Letac & Mora, 1990). A distribution is said to have cubic variance function if the variance is a cubic function of its mean. Distributions with quadratic variance functions of the mean have been examined by Morris (1982). An example of a quadratic variance distribution is the negative binomial (NB) distribution. The generalized Poisson (GP) distribution (Consul, 1989) and the Poisson-inverse Gaussian distribution (PIG) (Holla, 1966; Sichel, 1971; Willmot, 1987; Ord & Whitmore, 1986; Ong, 1998) are well-known, and well-researched count models for overdispersion (variance greater than the mean) in data and they have been widely used in applications in diverse disciplines. Due to “its physical justification, and its abundance of convenient

mathematical properties”, Willmot (1987) proposed the PIG distribution as an alternative to the NB distribution. The inverse trinomial distribution (IT) was introduced by Shimizu and Yanagimoto (1991) as a random walk model. Aoyama *et al.*, (2008) generalized the IT distribution, and later Phang *et al.*, (2013) examined statistical analysis for the IT distribution. Although the IT distribution is a cubic variance function distribution, it is not adequately mentioned by Letac and Mora (1990). The NB distribution is a particular important case of the IT distribution. Since the IT and GP distributions both have cubic variance functions, it is of interest to be able to discriminate between the two models in practical applications. Phang *et al.*, (2013) showed that the IT distribution is a viable model for statistical analysis of overdispersed count data. Khang and Ong (2007) have shown the IT distribution to be a Poisson-stopped distribution. The GP, IT and PIG distributions are non-nested models and are rather flexible. In many instances,

*Corresponding author's e-mail: yehchingl@sunway.edu.my

they provide similar fits to data sets. Apart from the above-mentioned consideration about dispersion, they may differ in other aspects, such as tail probabilities which may affect the conducted statistical inferences. Therefore, it is necessary to consider discrimination between these three distributions by examining the correct selection probability (PCS).

Likelihood ratio statistic is commonly used in model selection. The idea of using a logarithm of the ratio of the maximum likelihood function in selecting models is originated by Cox (1961, 1962). Bain and Englehardt (1980) use it in choosing between gamma and Weibull distributions. In addition, Gupta and Kundu (2003) employed this approach in choosing between gamma and generalized exponential (GE) distributions or between Weibull and GE distributions. Van der Hoeven (2005) discussed the probability to select the correct model using likelihood ratio based criteria in choosing between two nested models. Kundu *et al.*, (2005) considered discrimination between the lognormal and generalized exponential distributions. Recently, Qaffou and Zoglat (2017) used asymptotic results to distinguish the normal and Gumbel distributions. Virtually all the published work on discriminating between distributions are for continuous distributions. There seems to be a dearth of publications on discriminating discrete distributions. In this paper, we use the likelihood ratio statistic to discriminate the well-known PIG, GP and IT distributions. Since the exact distribution of this statistic is not known, the probabilities of correct selection are obtained via Monte Carlo simulations for different samples size and model parameters.

The structure of the paper is as follows. Section 2 presents a brief discussion of the IT, PIG and GP distributions. The PCS is given in section 3. A real-life data set is used as illustrations of the results in section 4. Section 5 provides some concluding remarks.

II. INVERSE TRINOMIAL, POISSON INVERSE GAUSSIAN AND GENERALIZED POISSON DISTRIBUTIONS

A. Inverse Trinomial Distribution

Inverse trinomial (IT) distribution is so named because its cumulant generating function is the inverse of the cumulant generating function for the trinomial distribution (Shimizu & Yanagimoto, 1991).

The probability mass function (pmf) of the IT distribution

is given by:

$$Pr(X = x) = \frac{\lambda p^\lambda q^x}{x + \lambda} \sum_{t=0}^{[x/2]} \binom{x + \lambda}{t, t + \lambda, x - 2t} \left(\frac{pr}{q^2}\right)^t \quad (1)$$

for $x = 0, 1, 2, \dots$, where $\lambda > 0, p \geq r, p + q + r = 1$, $[x]$ is the largest integer less than or equal to x and:

$$\binom{x + \lambda}{t, t + \lambda, x - 2t} = \frac{\Gamma(x + \lambda + 1)}{t! \Gamma(t + \lambda + 1) \Gamma(x - 2t + 1)}. \quad (2)$$

The probability generating function (pgf) of IT is given by:

$$G(u) = \left(\frac{2p}{[(1 - qu) + \sqrt{(1 - qu)^2 - 4pru^2}]} \right)^\lambda \quad (3)$$

$\lambda > 0, p + q + r = 1, 0 < 4pr/(1 - q)^2 < 1.$

If $r = 0$, this is the NB pgf.

The recurrence formula for pmf of IT is:

$$Pr(k + 1) = \frac{(q + 2\sqrt{pr})(\lambda + k)}{(k + 1)(2\lambda + 1 + k)} \left[\frac{(2\lambda + 1 + 2k)(1 - \frac{\phi}{2})}{(1 - \phi)(q + 2\sqrt{pr})(\lambda + k - 1)} Pr(k) - Pr(k - 1) \right] \quad (4)$$

where

$$\phi = \frac{4\sqrt{pr}}{(q + 2\sqrt{pr})} \quad (5)$$

with $Pr(0) = p^\lambda$ and $Pr(1) = p^\lambda \lambda q$. The mean and variance of IT are, respectively,

$$E[X] = \lambda \left\{ \frac{1 - (p - r)}{p - r} \right\} = \mu, \quad (6)$$

$$Var[X] = \frac{\lambda}{(p - r)^2} \left\{ 1 - (p - r) + \frac{2r}{p - r} \right\} = \sigma^2. \quad (7)$$

The variance σ^2 of the IT distribution can be expressed as a cubic function of the mean (Phang *et al.*, 2013):

$$\sigma^2 = \frac{2r}{\lambda^2} \mu^3 + \left(\frac{6r + 1}{\lambda} \right) \mu^2 + (6r + 1)\mu + 2r\lambda. \quad (8)$$

The Takács (1962) family of distributions examined by Letac

and Mora (1990, equation (4.37)) has a cubic variance function of the form:

$$V(m) = \left(\frac{\alpha+1}{\alpha p}\right)^2 m^3 + \frac{1}{p} \left(\frac{2\alpha+1}{\alpha}\right) m^2 + m, \alpha > 0, p > 0. \quad (9)$$

where m is the mean. Comparing this with equation (1) we see that they are of the same form apart from a constant. Refer also to Kokonendji (1995) for further discussion.

B. Poisson inverse Gaussian Distribution

The Poisson inverse Gaussian distribution is a mixed Poisson distribution derived from the Poisson distribution using the inverse Gaussian as a mixed distribution. It has received much attention in modelling overdispersed count data such as species abundance. Its theory and applications are discussed in Holla (1965), Sichel (1971) Willmot (1987), Ord and Whitmore (1986) and Ong (1998).

The pmf of the PIG is:

$$P(X = x) = \left(\frac{2\alpha}{\pi}\right)^{1/2} \frac{\mu^x e^{1/\lambda} K_{x-1/2}(\alpha)}{(\alpha\lambda)^x x!}, x = 0, 1, 2, \dots \quad (10)$$

where $\alpha = \sqrt{\lambda^{-2} + 2\mu/\lambda}$ and $K_y(\cdot)$ is the modified Bessel function of the third kind. The mean and variance are:

$$E(X) = \mu, \text{Var}(X) = \mu + \lambda\mu^2. \quad (11)$$

It is seen that the Poisson inverse Gaussian distribution has a quadratic variance function of the mean.

C. The Generalized Poisson Distribution

The generalized Poisson distribution is an extension of the Poisson distribution. For the properties and various applications of the generalized Poisson distribution, see Consul (1989).

The pmf of the GP is:

$$P(X = x) = \frac{(1+\gamma x)^{x-1} (\theta e^{-\gamma\theta})^x}{x! e^\theta}, x = 0, 1, 2, \dots \quad (12)$$

where the parameter space is as follows:

- i. $\theta > 0, \gamma \geq 0, 0 \leq \gamma\theta < 1$.
- ii. $\theta > 0, \gamma \leq 0, \max(-1, -\theta/m) \leq \gamma\theta < 0$, where m is the

largest positive integer such that $1 + \gamma m > 0$, see Consul (1989) and Johnson *et al.*, (1992, page 396).

The mean and variance are, respectively,

$$E(X) = \frac{\theta}{1-\gamma\theta} = \mu, \text{Var}(X) = \frac{\theta}{(1-\gamma\theta)^3} = \sigma^2. \quad (13)$$

The variance as a function of the mean μ is:

$$\sigma^2 = \frac{\theta}{(1-\gamma\theta)^3} = \frac{\mu^3}{\theta^2}. \quad (14)$$

This shows that the GP distribution is in the Abel class of distributions which have variance function (Letac & Mora, 1990, equation (4.35)):

$$V(m) = m \left(1 + \frac{m}{p}\right)^2, p > 0. \quad (15)$$

D. Likelihood Ratio Statistics

The likelihood and log-likelihood are, respectively, given by:

$$L = \prod_{x=0}^t \text{Pr}(x)^{F_x} \quad (16)$$

and

$$\ln L = \sum F_x \ln \text{Pr}(x) \quad (17)$$

where F_x is the observed frequency. The parameter estimates for the PIG, IT and GP distributions have been obtained by numerical maximization of the log-likelihood functions to determine the global maximum. For the IT distribution due to the constraint $p > r, p + q + r = 1$, the parameter r is fixed, and the maximization is with respect to λ and p .

III. PROBABILITY OF CORRECT SELECTION

Let L_M and L_N respectively be the likelihood functions for distributions M and N . The log-likelihood ratio statistic in selecting between two distributions M and N is given by:

$$T = \log(L_M/L_N). \quad (18)$$

If $T > 0$ we choose M , otherwise we choose N . The probability of correct selection (PCS) is given by:

$$Pr(T > 0 | \text{data from } M) \quad (19)$$

Monte Carlo simulation is employed to compute the probability of correct selection for different parameter values and sample sizes. For each model, several sets of parameters based on each model's respective constraints were chosen, and the model was simulated 5000 times for each set of parameter values.

Let F and G be the two models to be selected. The steps in obtaining the PCS are presented in the following algorithm.

1. Generate a sample from model F .
2. Estimate the parameters of model F .
3. Estimate the parameters of model G .
4. Determine the log-likelihood ratio T , whether it is positive or negative.
5. Calculate the probability of T positive, which is the probability of correct selection.

If we have to find the PCS between G and F , we repeat steps 1 to 5 by replacing model F with model G , and for step 5, we calculate the probability of T negative.

We use the procedures as mentioned above to compute the PCS for models IT and PIG ; IT and GP . We substitute the model F and model G accordingly.

1. Case 1: Between IT and PIG

Tables 1, 2 and 3 show the PCS between IT and PIG distributions when the data are generated from IT distribution with different λ values. The constraints for IT parameters are $p > r, p + r < 1$. Cells in the tables are left blank if these constraints are violated. The results show that overall they are easily distinguished between models IT and PIG especially when the λ increases from 0.5 to 2.5. When $\lambda = 0.5, p = 0.9$ and $r = 0.01$ and 0.05 , the PCS is very low. This is because most of the data are gathered at 0 and 1. There is a very high frequency of zeros for data generated from these two sets of parameters. When $\lambda = 0.5, p = 0.3$ and $r = 0.2$, the PCS is less than 0.5. This could be due to the high-over dispersion display by data generated from this set of parameters. The average dispersion index recorded is 6.23. The same reason applies to data generated with the

parameters $\lambda = 2.5, p = 0.3$ and $r = 0.2$. The PCS is 0.1334. The average dispersion index is 7.49. From Tables 1 and 2, we observed that when $p = 0.3$, the PCSs decrease as r increases. However, when $p = 0.5$ and $r \geq 0.1$ and when $p = 0.7$ the PCSs increase as r increases. The PCS for parameter sets when $\lambda = 2.5, p = 0.9$ is higher compared with $\lambda = 0.5, p = 0.9$. The PCS for different parameter sets appeared to be high when the sample size is large, $n = 1000$. The results given in Tables 4 and 5 show the computation of the PCS when the data come from PIG . All the PCSs are between 0.5-0.75 except when $v = 0.3$ and $\mu = 5$ which provide data with a very high overdispersion. The average dispersion index is 9.39. The results indicate that it is quite difficult to distinguish between PIG and IT for a certain range of parameter values. The PCS is higher when we increase the sample size to 1000 (Table 5). The PCS increases as v and μ increase.

Table 1. Probability of correct selection: IT and PIG distributions, data from IT distribution.

$\lambda = 0.5, n = 100$, subject to $p > r, p + r < 1$

p	0.3	0.5	0.7	0.9
r				
0.01	0.8106	0.7682	0.7006	0.0936
0.05	0.6812	0.6878	0.7072	0.4284
0.1	0.5412	0.6506	0.8638	-
0.2	0.4174	0.7726	0.9836	-
0.3	-	0.8538	-	-
0.4	-	0.9852	-	-

Table 2. Probability of correct selection: IT and PIG distributions, data from IT distribution.

$\lambda = 2.5, n = 100$, subject to $p > r, p + r < 1$

p	0.3	0.5	0.7	0.9
r				
0.01	0.7610	0.7606	0.7614	0.6334
0.05	0.6728	0.6846	0.7702	0.9844
0.1	0.5736	0.7146	0.9028	-
0.2	0.1334	0.7268	0.9972	-
0.3	-	0.8002	-	-
0.4	-	0.8062	-	-

Table 3. Probability of correct selection: IT and PIG distributions, data from IT distribution.
 $\lambda=2.5$, $n = 1000$, subject to $p > r, p + r < 1$

p	0.3	0.5	0.7	0.9
r				
0.01	0.9926	0.9224	0.8420	0.9738
0.05	0.9688	0.9010	0.9692	1.0000
0.1	0.9414	0.9710	1.0000	-
0.2	1.0000	0.9998	1.0000	-
0.4	-	1.0000	-	-

Table 4. Probability of correct selection: IT and PIG distributions, data from PIG distribution.
 $n = 100$

v	0.3	0.6	1	3	5
μ					
0.3	0.6122	0.6560	0.6896	0.7302	0.7540
0.6	0.5826	0.5584	0.5634	0.6868	0.6930
1	0.6058	0.5900	0.5630	0.5724	0.6156
3	0.5844	0.5404	0.5928	0.6066	0.5856
5	0.2156	0.5860	0.6450	0.5702	0.5898

Table 5. Probability of correct selection: IT and PIG distributions, data from PIG distribution.
 $n = 1000$

v	0.3	0.6	1	3	5
μ					
0.3	0.5590	0.6136	0.6812	0.8206	0.7644
0.6	0.7554	0.7274	0.6524	0.5848	0.6672
1	0.9272	0.8014	0.7336	0.6404	0.5476
3	0.9694	0.9716	0.9588	0.9102	0.9368
5	0.9924	0.9908	0.9864	0.9656	0.9718

2. Case 2: Between IT and GP

Tables 6, 7 and 8 display the result of PCS calculated between IT and GP when the data were simulated from IT. For data sets which are highly dispersed, the PCSs are very low for example when $p = 0.3$ and $r = 0.1$ and 0.2 where the average dispersion indices are 3.63 and 6.23 respectively. Besides that, some data with certain parameter values are quite difficult to distinguish while others are easy to distinguish. We can

observe that IT and GP may possess the common characteristic such as skewness and dispersion under certain parameter values. Overall, Table 6 shows that it has the lowest PCS in comparison with IT and PIG. When λ increases to 2.5, the PCS's of p between 0.3 and 0.5 for IT and PIG are reduced. It is relatively more difficult to distinguish between IT and GP. Nevertheless, when the PCS's increase for $p > 0.5$, the PCS's are above 0.7. When the sample size increases to 1000, the PCSs are very high (close to 1). When data are coming from GP distribution under the constraint that $\gamma\theta < 1$, they are quite difficult to distinguish except some data sets with high dispersion such as when $\gamma = 0.7$, $\theta = 1$ and $\gamma = 1$, $\theta = 0.7$ which produce dispersion indices of 4.0 and 4.35 respectively. The PCSs for these data sets are 0.734 and 0.7322. Other PCSs are between 0.4306 and 0.6632 (Table 14). Contrary to other models, the PCSs decrease and most of them drop until less than 0.5 (Table 10) when the sample size is increased to 1000. This indicates that we cannot distinguish between IT and GP when the data come from GP, and the sample size is big. Cells in Tables 9, 10, 13 and 14 are left blank if the constraint $\gamma\theta < 1$ is violated.

Table 6. Probability of correct selection: IT and GP distributions, data from IT distribution.
 $\lambda = 0.5$, $n = 100$, subject to $p > r, p + r < 1$

p	0.3	0.5	0.7	0.9
r				
0.01	0.7196	0.7398	0.6864	0.0367
0.05	0.5078	0.6168	0.6690	0.2840
0.1	0.3240	0.5306	0.8306	-
0.2	0.1606	0.6196	0.9674	-
0.3	-	0.6460	-	-
0.4	-	0.8178	-	-

Table 7. Probability of correct selection: IT and GP distributions, data from IT distribution.
 $\lambda = 2.5$, $n = 100$, subject to $p > r, p + r < 1$

p	0.3	0.5	0.7	0.9
r				
0.01	0.6208	0.7074	0.7324	0.6208
0.05	0.4754	0.5542	0.8338	0.9746

0.1	0.3110	0.5338	0.8376	-
0.2	0.0190	0.4700	0.9904	-
0.3	-	0.4976	-	-
0.4	-	0.5110	-	-

Table 8. Probability of correct selection: IT and GP distributions, data from IT distribution.

$\lambda = 2.5, n = 1000$, subject to $p > r, p + r < 1$

p	0.3	0.5	0.7	0.9
r				
0.01	0.9662	0.9744	0.9284	0.9596
0.05	0.8126	0.9632	0.9114	1.0000
0.1	1.0000	0.8982	0.9980	-
0.2	1.0000	0.9524	1.0000	-
0.3	-	0.9918	-	-
0.4	-	0.9744	-	-

Table 9. Probability of correct selection: IT and GP distributions, data from GP distribution.

$n = 100$, subject to $\gamma\theta < 1$

γ	0.3	0.5	0.7	1	2
θ					
0.1		0.7070	0.6118	0.5334	0.4726
0.3	0.6052	0.4742	0.4556	0.5104	0.6632
0.5	0.5524	0.5022	0.4842	0.5978	-
0.7	0.6158	0.5186	0.6342	0.7322	-
1	0.4306	0.5704	0.7340	-	-
2	0.5928	0.9788	-	-	-

Table 10. Probability of correct selection: IT and GP distributions, data from GP distribution.

$n = 1000$, subject to $\gamma\theta < 1$

γ	0.3	0.5	0.7	1	2
θ					
0.1	0.6144	0.4990	0.4756	0.4702	0.0602
0.3	0.4054	0.4296	0.1300	0.0856	0.0352
0.5	0.1374	0.1310	0.1034	0.1220	-
0.7	0.3704	0.1256	0.1170	0.3110	-
1	0.1492	0.1490	0.3524	-	-
2	0.3952	-	-	-	-

3. Case 3: PIG and GP

Tables 11 and 12 showed the computed PCS for PIG and GP when data come from PIG. When $n = 100$, it is quite difficult to distinguish between PIG and GP especially for data that are highly dispersed with dispersion index more than 4 (see Table 11). The PCS are below 0.5. In addition, the PCS for other parameter values are between 0.5 and 0.63. Overall, the PCS increase as ν increases. Conversely, the PCS decrease as μ increases. Table 12 shows that when the sample size increases to 1000, the PCS also increase. It is clear that the two models can be discriminated easily. When the data come from GP, for sample size $n=100$, the two models are better discriminated in comparison to data coming from the PIG. The PCS are between 0.55 and 0.86 except when $\gamma = 0.5, \theta = 0.1$ with PCS equal to 0.4034. The dispersion index for data with these parameter values is 1.05. The PCS increase as γ and θ increase. Results are given in Table 13. When sample sizes increase (Table 14), the two models are well discriminated. The PCS increase as γ and θ increase and they are higher than when sample size $n = 100$.

Table 11. Probability of correct selection: PIG and GP distributions, data from PIG distribution.

$n = 100$

ν	0.3	0.6	1	3	5
μ					
0.3	0.5906	0.6058	0.6154	0.6118	0.6336
0.6	0.5612	0.5450	0.5530	0.3634	0.6220
1	0.5368	0.5656	0.5396	0.5676	0.4358
3	0.3910	0.4154	0.5062	0.5732	0.5694
5	0.0964	0.3504	0.4364	0.4960	0.5444

Table 12. Probability of correct selection: PIG and GP distributions, data from PIG distribution.

$n = 1000$

ν	0.3	0.6	1	3	5
μ					
0.3	0.5592	0.6140	0.6736	0.7148	0.3082
0.6	0.7568	0.7286	0.6538	0.5834	0.6324
1	0.9306	0.8052	0.7366	0.6406	0.5508
3	0.9626	0.9734	0.9616	0.9162	0.9398
5	0.9676	0.9786	0.9816	0.9706	0.9760

Table 13. Probability of correct selection: PIG and GP distributions, data from GP distribution.

$n = 100$, subject to $\gamma\theta < 1$

γ	0.3	0.5	0.7	1	2
θ					
0.1		0.4034	0.6364	0.6076	0.5682
0.3	0.6296	0.6110	0.5912	0.5918	0.6840
0.5	0.5874	0.5908	0.6170	0.6472	-
0.7	0.5526	0.6138	0.6352	0.7128	-
1	0.6856	0.6424	0.6908	-	-
2	0.6424	0.8598	-	-	-

Table 14. Probability of correct selection: PIG and GP distributions, data from GP distribution.

$n = 1000$, subject to $\gamma\theta < 1$

γ	0.3	0.5	0.7	1	2
θ					
0.1	0.4718	0.5426	0.5562	0.5702	0.6508
0.3	0.5944	0.6106	0.6878	0.7668	0.8352
0.5	0.6384	0.7238	0.8076	0.9120	-
0.7	0.6630	0.8168	0.9052	0.9772	-
1	0.7496	0.8948	0.9560	-	-
2	0.9084	-	-	-	-

IV. APPLICATION TO DATA ANALYSIS

We apply the PCS to a real-life data set taken from Heilbron (1994). We apply IT, PIG, and GP distributions to the data set. Maximum likelihood estimates, χ^2 values and the log-likelihood values are presented in Table 15.

For the real-life data set (Table 15), when we compare IT and PIG distributions, the likelihood ratio $T_1 = -1014.75 - (-1011.49) = -3.26$ gives a negative value which implies that the PIG distribution is preferable than IT distribution. Referring to Table 4, the PCSs between IT and PIG distributions are approximately 80% which provide the same conclusion where PIG is more favourable than IT distribution.

Table 15. Sexual behaviour count data (Heilbron, 1994)

x	Observed frequency	Expected frequency		
		IT	P-IG	GP
0	466	472.13	465.73	471.62
1	182	179.40	192.94	179.58
2	103	78.78	76.44	78.75
3	29	37.79	34.42	37.89
4	15	9.27	17.36	19.38
5	3	10.26	9.48	10.34
6	5	5.65	5.48	5.70
7	0	3.19	3.29	3.21
8	2	1.84	2.04	1.85
9	0	1.07	1.29	1.08
10	2	0.64	0.84	0.64
11	0	0.38	0.55	0.38
12	0	0.23	0.36	0.23
13	0	0.14	0.24	0.14
14	1	0.09	0.16	0.08
15	1	0.05	0.11	0.05
16	1	0.03	0.08	0.03
17	1	0.04	0.18	0.05
Total	811			
-log-likelihood	1019.09	1014.75	1011.49	1014.95
χ^2		15.78	15.64	15.95

$$\begin{aligned} \hat{p}_{IT} &= 0.5527 & \hat{p}_{PIG} &= 0.4539 & \hat{\theta}_{GP} &= 0.5421 \\ \hat{r}_{IT} &= 0.0308 & \hat{\mu}_{PIG} &= 0.8390 & \hat{\gamma}_{GP} &= 0.6516 \\ \hat{\lambda}_{IT} &= 0.9125 \end{aligned}$$

V. CONCLUDING REMARKS

The χ^2 goodness-of-fit statistic and Akaike Information Criterion are popular methods for model selection. The probability of correct selection method serves as an alternative approach in model selection. It can be used for nested and non-nested models. Since the exact distribution of the likelihood ratio is complicated, the probability of correct selection is obtained through Monte Carlo simulation. The results indicate that under certain parameter values, it is quite difficult to discriminate the selected models. This could be due to the reason that under these parameter values, these models share the common characteristics such as they may have the same dispersion values, skewness or kurtosis.

Generally, the probability of correct selection is affected by the dispersion index and parameter values. From the simulation results, it is observed that the pair IT and GP and IT and PIG are more easier to discriminate when the sample size is large and data is from IT distribution. As an illustration

of application, a real-life data set (Heilbron, 1994) has been analyzed. The interesting problem of obtaining the asymptotic distribution of the likelihood ratio statistic will be considered elsewhere.

VI. REFERENCES

- Aoyama K, Shimizu, K & Ong, SH 2008, 'A first-passage time random walk distribution with five transition probabilities: a generalization of the shifted inverse trinomial', *Annals of the Institute of Statistical Mathematics*, vol. 60, pp. 1-20.
- Bain, LJ & Englehardt M 1980, 'Probability of correct selection of Weibull versus gamma based on likelihood ratio', *Communication in Statistics - Series A*, vol. 9, pp. 375-381.
- Consul, PC 1989, *Generalized Poisson Distribution*, Dekker, New York, US.
- Cox, DR 1961, 'Tests of separate families of hypotheses', *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 96.
- Cox, DR 1962, 'Further results on tests of separate families of hypotheses', *Journal of Royal Statistical Society*, vol. B24, pp. 406-424.
- Gupta, RD & Kundu, D 2003, 'Discriminating between Weibull and generalized exponential distributions', *Computational Statistics & Data Analysis*, vol. 43, pp. 179-196.
- Heilbron, D 1994, 'Zero-altered and other regression models for count data with added zeros', *Biometrical Journal*, vol. 36, pp. 531-547.
- Holla, MS 1966, 'On a Poisson-inverse Gaussian distribution', *Metrika*, vol. 11, pp. 115-121.
- Johnson, NC, Kotz, S & Kemp, AW 1992, *Univariate Discrete Distributions*, 2 edn, John Wiley & Sons, New York, US.
- Khang, TF & Ong SH 2007, 'A new generalization of the logarithmic distribution arising from the inverse trinomial distribution', *Communications in Statistics - Theory and Methods*, vol. 36, pp. 3-21.
- Kokonendji, CC 1995, 'Sur les familles exponentielles naturelles réelles de grand-Babel', *Annales de la Faculté des Sciences de Toulouse*, vol. 4, no. 4, pp. 763-800.
- Kundu, D, Gupta, RD & Manglick, A 2005, 'Discriminating between the lognormal and generalized exponential distributions', *Journal of Statistical Planning and Inference*, vol. 127, pp. 213-227.
- Letac, G & Mora, M 1990, 'Natural real exponential families with cubic variance functions', *Annals of Statistics*, vol. 18, pp. 1-37.
- Morris, C 1982, 'Natural exponential families with quadratic variance functions', *Annals of Statistics*, vol. 10, no. 1, pp. 65-80.
- Ong, SH 1998, 'A note on the mixed Poisson formulation of the Poisson-inverse Gaussian distribution', *Communications in Statistics-Simulations*, vol. 27, no. 1, pp. 67-78.
- Ord, JK & Whitmore, GA 1986, 'The Poisson-inverse Gaussian distribution as a model for species abundance', *Communications in Statistics - Simulation and Computation*, vol. 27, pp. 67-78.
- Phang, YN, Sim, SZ & Ong, SH 2013, 'Statistical Analysis for the Inverse Trinomial Distribution', *Communications in Statistics - Simulation and Computation*, vol. 42, no. 9, pp. 2073-2085.
- Qaffou, A & Zoglat, A 2016, 'Discriminating between normal and Gumbel distributions', *Revstat - Statistical Journal*, vol. 15, pp. 523-536.
- Shimizu, K & Yanagimoto, T 1991, 'The inverse trinomial distribution', *Japanese Journal of Applied Statistics*, vol. 20, no. 2, pp. 89-96.
- Sichel, HS 1971, 'On a family of discrete distributions particularly suited to represent long-tailed frequency data', ed Laubscher NF, in *Proceedings of the third Symposium on Mathematical Statistics*, Pretoria, CSIR, pp. 51-97.
- Van der Hoeven, N 2005, 'The probability to select the correct model using likelihood-ratio based criteria in choosing between two nested models of which the more extended one is true', *Journal of Statistical Planning and Inference*, vol. 135, no. 2, pp. 477-486.
- Willmot, GE 1987, 'The Poisson-inverse Gaussian distribution as an alternative to the negative binomial', *Scandinavian Actuarial Journal*, vol. 2, pp. 113-127.