

Simple models for macro-parasite distributions in hosts

Gonzalo Maximiliano LOPEZ^{1,3,4}, Juan Pablo APARICIO^{1,2}

¹ Instituto de Investigaciones en Energía no Convencional (INENCO),
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET),
Universidad Nacional de Salta, Av. Bolivia 5150, 4400 Salta, Argentina.

² Simon A. Levin Mathematical, Computational and Modeling Sciences Center,
Arizona State University, PO Box 871904 Tempe, AZ 85287-1904, USA

³ Departamento de Matemática,
Universidad Nacional de Salta, Av. Bolivia 5150, 4400 Salta, Argentina.

⁴ Corresponding author: gonzalo.maximiliano.lopez@gmail.com

Abstract

Negative binomial distribution is the most used distribution to model macro-parasite burden in hosts. However reliable maximum likelihood parameter estimation from data is far from trivial. No closed formula is available and numerical estimation requires sophisticated methods. Using data from the literature we show that simple alternatives to negative binomial, like zero-inflated geometric or hurdle geometric distributions, produce a good and even better fit to data than negative binomial distribution. We derived closed simple formulas for the maximum likelihood parameter estimation which constitutes a significant advantage of these distributions over negative binomial distribution.

Keywords: Hurdle geometric distribution; Macroparasite, Maximum likelihood estimation; Negative binomial distribution; Zero-inflated geometric distribution;

Contents

Abstract	1
1 Introduction	2
2 Zero-inflated(deflated) distributions and hurdle distributions	3
2.1 Zero-inflated(deflated) distributions	4
2.2 Hurdle distributions.	5
3 Parameter estimation and Maximum likelihood	5
3.1 Maximum likelihood estimation for zero-inflated(deflated) and hurdle distributions	6
3.1.1 Zero-inflated(deflated) distribution	6
3.1.2 Hurdle distributions	6
4 Maximum likelihood for the negative binomial distribution and two simple alternatives.	7
4.1 Zero-inflated geometric distribution	7
4.1.1 Maximum likelihood estimation	8
4.2 Hurdle geometric distribution	8
4.2.1 Maximum likelihood estimation	9
5 Some examples	10
5.1 Study of frequency distribution of <i>Ascaris lumbricoides</i> infection	10
5.2 Parasite distribution in crabs	10
6 Discussion and Conclusions	11
Bibliography	13
7 Tables	15

1 Introduction

Macroparasites usually present over-dispersed distributions where few hosts account for most of the parasites in the population (see, for example, [4][14]).

The most used distribution is the negative binomial distribution ([2][15]) which provide and accurate description of the observations.

However in many cases the negative binomial distribution (or other similar distributions) cannot account for the “excess” of zeros observed. A simple solution widely used is to consider zero-inflated distributions [9][6][7].

The negative binomial distribution is a two parameters distribution, which usually are the mean burden of parasites in the host population m and the inverse dispersion parameter k which is related with the degree of the over-dispersion [2]. Moreover, it can be shown that the limiting distribution of the NB(m, k) distribution, as $k \rightarrow \infty$, is a Poisson (m) distribution and if $k = 1$ the NB(m, k) distribution is a geometric ($\frac{k}{m+k}$) distribution.

However one problem with the negative binomial distribution is parameter estimation from the observations. The method of moments estimation is simple but not always precise [3]. Maximum likelihood estimation (MLE) provides one of the best parameters estimation ([11]) but for the negative binomial distribution there is no a closed formula for the parameter estimates in terms of the observations and should be obtained numerically which present some complexities respect to the parameters numerical estimation for other distributions [5][1].

In this article we show alternatives to the negative binomial distribution which describe the observations equally well (and in some cases provide a most precise description) but for which we can compute a formula for the maximum likelihood parameters estimation values.

2 Zero-inflated(deflated) distributions and hurdle distributions

Macroparasites usually show over-dispersed distribution where few individuals have most of the parasite in the host population. It is said that distributions follow the 20-80 rule, 20% percent of the individuals account for the 80% of the parasite burden [18]. Negative binomial distribution, among others, offers such over-dispersed distribution but in some cases an excess of zeros is observed for which the distributions fail to account. A simple and widely used solution is to consider zero-inflated (in some cases zero-deflated) distributions and hurdle distributions [16][10].

2.1 Zero-inflated(deflated) distributions

A discrete random variable Y follows a zero-inflated(deflated) distribution if its probability mass function P is given by

$$P(Y = y) = \begin{cases} \pi + (1 - \pi)p(0; \theta) & y = 0 \\ (1 - \pi)p(y; \theta) & y \neq 0 \end{cases} \quad (1)$$

where we denote by θ to the vector of parameters of the associated distribution p and then

$$\sum_{y=0}^{\infty} p(y; \theta) = 1.$$

When $\pi < 0$ we have a zero-deflated distribution.

If $G(z)$ is the probability generating function of the distribution p , then the probability generating function of the corresponding zero-inflated(deflated) distribution is given by

$$F(z) = \pi + (1 - \pi)G(z) \quad (2)$$

From the probability generating function we may obtain the mean and the variance for the distribution straightforwardly

$$\begin{aligned} E(Y) &= (1 - \pi)\mu \\ Var(Y) &= (1 - \pi)\sigma^2 + \pi(1 - \pi)\mu^2 \end{aligned} \quad (3)$$

where μ, σ^2 are the mean and variance of the distribution p . The coefficient of dispersion, or variance-to-mean ratio is therefore $D = \frac{Var(Y)}{E(Y)}$,

$$D = \frac{\sigma^2}{\mu} + \pi\mu \quad (4)$$

where σ^2/μ is the variance-to-mean ratio for the associated distribution p . As expected, zero-inflated distributions are more over-dispersed than the associated distribution.

2.2 Hurdle distributions.

Another common way to account for an excess of zeros are the hurdle distributions,

$$P(Y = y) = \begin{cases} \pi & y = 0 \\ (1 - \pi) \frac{p(y; \theta)}{1 - p(0; \theta)} & y \neq 0 \end{cases} \quad (5)$$

If $G(z)$ is the probability generating function for the associated distribution p , then the probability generating function for the hurdle distribution is

$$F(z) = \pi + (1 - \pi) \frac{G(z) - p(0; \theta)}{1 - p(0; \theta)} \quad (6)$$

from where we obtain the mean and the variance of the hurdle distribution as

$$\begin{aligned} E(Y) &= \alpha \mu \\ Var(Y) &= \alpha \sigma^2 + \alpha(1 - \alpha) \mu^2 \end{aligned} \quad (7)$$

where $\alpha = \frac{1 - \pi}{1 - p(0; \theta)}$, and μ, σ^2 are the mean and variance of the associated distribution p . Finally the variance-to-mean ratio for the hurdle distribution is given by

$$D = \frac{\sigma^2}{\mu} + (1 - \alpha) \mu \quad (8)$$

If $\pi > p(0; \theta)$ the hurdle distribution are more over-dispersed than the associated distribution.

3 Parameter estimation and Maximum likelihood

A simple, but in general inaccurate, way to fit the parameters of a distribution from a sample consist in the use of the sample moments. For example, the negative binomial distribution has two parameters which can be expressed in terms of the two first moments. While this method is quite simple, do not provide a reliable fit [3] [11]. Usually, the method of choice is maximum likelihood [11]. However, maximum likelihood estimation (MLE) not always produce a closed formula and parameters need to be estimated numerically [2][5][1]. In the following we pose the problem of maximum likelihood parameter estimation for zero-inflated and hurdle distributions.

3.1 Maximum likelihood estimation for zero-inflated(deflated) and hurdle distributions

The problem of parameter estimation by maximum likelihood for zero-inflated and hurdle distributions is presented in the following.

3.1.1 Zero-inflated(deflated) distribution

We denote by θ to the vector of parameters of the associated distribution p and therefore the set of parameters for the zero-inflated(deflated) distribution is (π, θ) . If y_1, \dots, y_N are N observations, N_0 is the number of observations with zero counts, then the log-likelihood function is given by

$$\ell(\pi, \theta) = N_0 \ln [\pi + (1 - \pi)p_0] + (N - N_0) \ln(1 - \pi) + \sum_{y_i \neq 0} \ln p_{y_i} \quad (9)$$

Maximizing ℓ for (π, θ) we obtain the following system,

$$\begin{aligned} \frac{\partial \ell}{\partial \pi} &= \frac{N_0(1 - p_0)}{\pi + (1 - \pi)p_0} - \frac{N - N_0}{1 - \pi} = 0 \\ \frac{\partial \ell}{\partial \theta_i} &= \frac{N_0(1 - \pi) \frac{\partial p_0}{\partial \theta_i}}{\pi + (1 - \pi)p_0} + \sum_{y_i \neq 0} \frac{\frac{\partial p_{y_i}}{\partial \theta_i}}{p_{y_i}} = 0 \end{aligned} \quad (10)$$

where we denote by $p_{y_i} = p(y_i; \theta)$.

3.1.2 Hurdle distributions

The log-likelihood function for a hurdle distribution is given by

$$\ell(\pi, \theta) = N_0 \ln \pi + (N - N_0) \ln \left(\frac{1 - \pi}{1 - p_0} \right) + \sum_{y_i \neq 0} \ln p_{y_i} \quad (11)$$

and therefore the parameter values are obtained from the system

$$\begin{aligned} \frac{\partial \ell}{\partial \pi} &= \frac{N_0}{\pi} - \frac{N - N_0}{1 - \pi} = 0 \\ \frac{\partial \ell}{\partial \theta_i} &= \frac{(N - N_0) \frac{\partial p_0}{\partial \theta_i}}{1 - p_0} + \sum_{y_i \neq 0} \frac{\frac{\partial p_{y_i}}{\partial \theta_i}}{p_{y_i}} = 0 \end{aligned} \quad (12)$$

4 Maximum likelihood for the negative binomial distribution and two simple alternatives.

The probability mass function for the negative binomial distribution is given by

$$P(X = x) = \frac{\Gamma(x + k)}{\Gamma(x + 1)\Gamma(k)} p^k q^x \quad (13)$$

where $0 \leq p \leq 1$ y $k > 0$. Differentiating the log-likelihood function ℓ partially and setting them equal to zero yields the following system of likelihood equations

$$\begin{aligned} \frac{\partial \ell}{\partial p} &= \frac{Nk}{p} - \frac{mN}{1-p} = 0 \\ \frac{\partial \ell}{\partial k} &= \left[\sum_i \psi(y_i + k) \right] - N\psi(k) + N \ln p = 0 \end{aligned} \quad (14)$$

where $\psi(y) = \frac{\Gamma'(y)}{\Gamma(y)}$ is the digamma function. Substituting in the second equation the value of $p = \frac{k}{m+k}$ obtained from the first equation, gives:

$$\left[\sum_i \psi(y_i + k) \right] - N\psi(k) + N \ln \left(\frac{k}{m+k} \right) = 0 \quad (15)$$

This equation cannot be solved for k in a closed form and must be solved numerically. Iterative technique as Newton-Raphson method can be used, but this method may fail to find the MLE value. An analysis of the literature indicates that finding the MLE value is a challenge, since we could not obtain the root or obtain more than one for the equation (15) [3][11][5][17][13].

4.1 Zero-inflated geometric distribution

The geometric distribution is a special case of the negative binomial distribution for $k = 1$, and then, its probability mass function is given by

$$P(X = x) = pq^x \quad \text{con } q = 1 - p \quad (16)$$

where $x = 0, 1, 2, \dots$ and $0 \leq p \leq 1$. The mean is $\frac{1-p}{p}$ while the variance is $\frac{1-p}{p^2}$.

From (1) the corresponding zero-inflated(deflated) distribution is

$$P(Y = y) = \begin{cases} \pi + (1 - \pi)p & y = 0 \\ (1 - \pi)pq^y & y \neq 0 \end{cases} \quad (17)$$

Mean and variance is given by (3),

$$\begin{aligned} E(Y) &= (1 - \pi) \frac{(1 - p)}{p} \\ Var(Y) &= (1 - \pi) \frac{(1 - p)}{p^2} [1 + \pi(1 - p)]. \end{aligned} \quad (18)$$

The variance-to-mean ratio for the zero-inflated(deflated) geometric distribution is always greater than one, and therefore this distribution is always over-dispersed.

4.1.1 Maximum likelihood estimation

System of likelihood equations is given according to (10) by

$$\begin{aligned} \frac{\partial \ell}{\partial \pi} &= \frac{N_0(1 - p)}{\pi + (1 - \pi)p} - \frac{N - N_0}{1 - \pi} = 0 \\ \frac{\partial \ell}{\partial p} &= \frac{N_0(1 - \pi)}{\pi + (1 - \pi)p} + \frac{N - N_0}{p} - \frac{mN}{q} = 0 \end{aligned} \quad (19)$$

therefore, the best parameter estimations result

$$\begin{aligned} \hat{\pi} &= \frac{mN_0 - N + N_0}{mN - N + N_0} \\ \hat{p} &= \frac{N - N_0}{mN} \end{aligned} \quad (20)$$

which are in terms of the observed sample mean m , the sample size N and the number of zeros in the sample N_0 .

4.2 Hurdle geometric distribution

The hurdle distribution for the associated geometric distribution is given by

$$P(Y = y) = \begin{cases} \pi & y = 0 \\ (1 - \pi) \frac{pq^y}{1 - p} & y \neq 0 \end{cases} \quad (21)$$

Mean and variance are obtained in a straightforward way,

$$\begin{aligned} E(Y) &= \frac{(1 - \pi)}{p} \\ Var(Y) &= \frac{(1 - \pi)}{p^2} [1 + (\pi - p)] \end{aligned} \tag{22}$$

The variance-to-mean ratio for the hurdle geometric distribution is always greater than one, and therefore this distribution is always over-dispersed.

4.2.1 Maximum likelihood estimation

According to (12) the system of likelihood equations is given by

$$\begin{aligned} \frac{\partial \ell}{\partial \pi} &= \frac{N_0}{\pi} - \frac{N - N_0}{1 - \pi} = 0 \\ \frac{\partial \ell}{\partial p} &= (N - N_0) \left(\frac{1}{1 - p} + \frac{1}{p} \right) - \frac{mN}{1 - p} = 0 \end{aligned} \tag{23}$$

therefore, the best parameter estimations result

$$\begin{aligned} \hat{\pi} &= \frac{N_0}{N} \\ \hat{p} &= \frac{N - N_0}{Nm} \end{aligned} \tag{24}$$

which are in terms of the observed sample mean m , the sample size N and the number of zeros in the sample N_0 .

Note that the pair $\left(\frac{\hat{\pi} - \hat{p}}{1 - \hat{p}}, \hat{p} \right)$ is a maximum of log-likelihood function (9). By the change of variables $\pi_{\text{hg}} = \pi_{\text{zig}} + (1 - \pi_{\text{zig}})p$ the probability mass functions of the zero-inflated geometric (zig) and hurdle geometric (hg) models coincide and the corresponding Akaike's information criterion values will be the same. Due to this, the fit of the data by both models coincide. In what follows we will only carry out the analysis of the zero-inflated geometric model.

5 Some examples

5.1 Study of frequency distribution of *Ascaris lumbricoides* infection

The nematode *Ascaris lumbricoides* is one of the more common intestinal parasites of humans. Highly prevalent in tropical and temperate populations where poverty and lack of sanitation is common [12].

Burden of infestation is computed using the number of parasites in each host of the sample [14]. At present time this type of studies are not longer conducted and we will use the results from Seo [14]. He studied six rural populations in Korea where an endemic situation was observed.

The samples showed over-dispersed distributions which could be accurately fitted by a negative binomial distribution (see Figure 1).

In Figure 1 we show the observed (black) and expected values of the fitted models (negative binomial and zero-inflated geometric). In addition, Table 1 includes the maximum likelihood estimations, the chi-squared statistics and their corresponding p -values, and Akaike's information criterion (AIC). As we see, the zero-inflated geometric distribution fit the data as well as the negative binomial distribution in most cases. The AIC results show that the models negative binomial and zero-inflated geometric are similar. Indeed, the fit in Figure 1 improves the results obtained by negative binomial distribution in samples E and F. Using AIC, the model zero-inflated geometric showed the best performance in samples C, E and F. Hence, the $ZIG(\pi, p)$ is a suitable candidate model to fit such data.

5.2 Parasite distribution in crabs

Crofton [4] contributed significantly to the study of parasite distributions in hosts. In his works he observed that over-dispersion is one of the main characteristics of parasite host distributions. Analyzing data from Hynes and Nicholas [8] of parasite infestation of the crustacean *Gammarus pulex*, by the parasite acanthocephalan *Polymorphus minutus*. Crofton show that negative binomial distribution provides a good fit to the data.

In Figure 2 we compare the fit to the data obtained by Crofton using the negative binomial distribution and the zero-inflated(deflated) geometric distribution. In most cases our simple proposal provides a better fit than the negative binomial distribution.

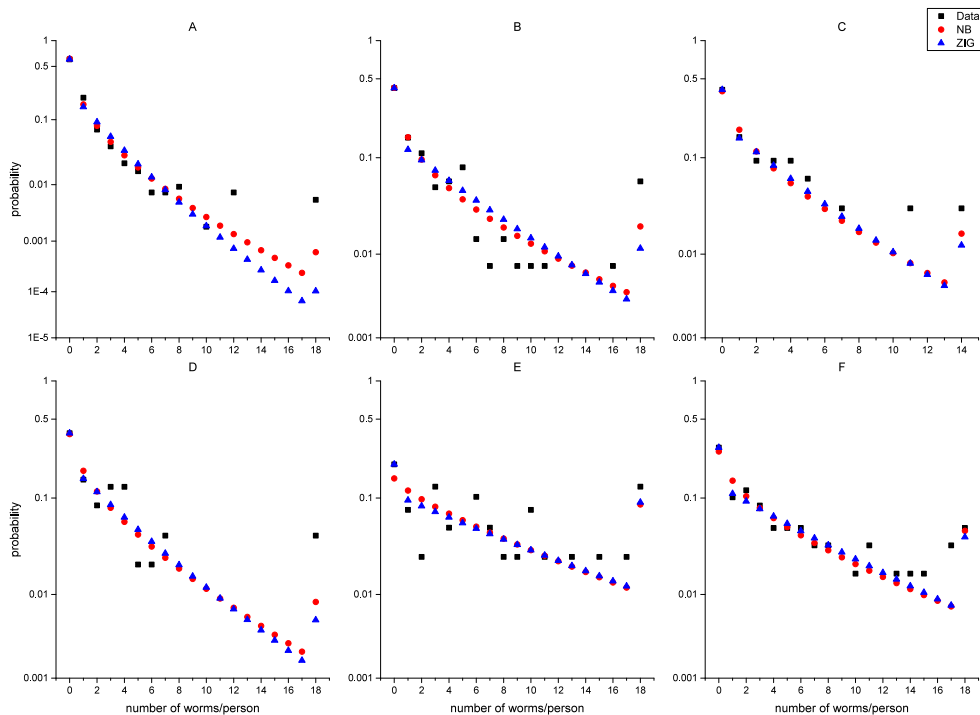


Figure 1: Fitting the parasite counts data (black) by NB (red) and ZIG (blue) distribution for Seo data set [14]. Except in the first case, the simple zero-inflated geometric distribution fit the data as well as the negative binomial distribution (see Table 1)

Based on the AIC and the chi-square goodness-of-fit test reported in Table 2 we conclude that the zero-inflated(deflated) geometric distribution provides a fit of the data as good as the negative binomial distribution.

6 Discussion and Conclusions

Negative binomial distribution is widely used to describe parasite burden in populations. It provides good fits of the observations which can be improved by the corresponding zero-inflated(deflated) distribution [4] [14].

However parameters estimation is far from trivial and maximum likelihood estimates must be found always numerically. Simple numerical methods (as the Newton method) are not easy to implement as there is not a closed

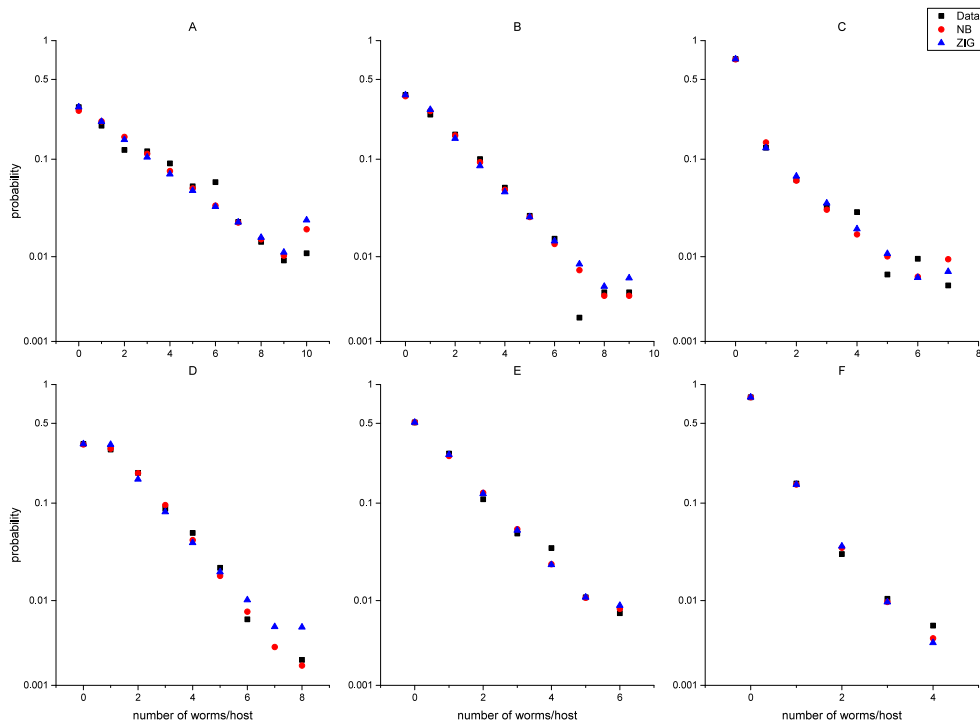


Figure 2: Fitting the parasite counts data (black) by NB (red) and ZIG (blue) distribution for data in parasite in crabs [4]. Zero-inflated(deflated) geometric distribution fit the data as well or better than the negative binomial distribution (see Table 2)

expression for the derivative of the Gamma function and fail if the starting value is not chosen appropriately [1].

Negative binomial distribution may also fail to fit the zero counts [4]. This may problem may be overcome using the zero-inflated negative binomial distribution but in this case parameters estimation by maximum likelihood is even more complex and also the AIC criteria penalize models with a larger number of parameters.

For zero-inflated geometric distribution or hurdle geometric distribution we found simple formulas for the maximum likelihood parameter estimates.

In the examples analyzed in this work zero-inflated geometric distribution or hurdle geometric distributions present in most cases a similar fit to the data than the negative binomial distribution. In the few cases this distribution significantly improve the fit. The AIC results show that the models are

similar.

However the major advantage of this distributions is not a little improvement in the data fitting but the fact that simple formula is provided for the maximum likelihood parameter estimates.

A simple formula for the distribution's parameters, avoiding the use of complex numerical methods for parameter estimation, may result of practical convenience for many researchers working in the area which are not familiar with programming and the numerical implementation of algorithms.

For all the models considered not good fit of the tail of the distribution is observed. However this fact does not necessarily indicates the need to consider other distributions. Because samples are small, rare events in the tail, when observed acts like outliers given too much weight to this rare observation.

Aknowledgements

This work was partially supported by grant CIUNSA 2018-2467. JPA is a member of the CONICET. GML is a doctoral fellow of CONICET.

Conflict of Interest

The authors have declared no conflict of interest.

Bibliography

- [1] Bandara, U., Gill, R., and Mitra, R. (2019). On computing maximum likelihood estimates for the negative binomial distribution. *Statistics & Probability Letters*, 148:54–58. 3, 5, 12
- [2] Bliss, C. and Fisher, R. (1953). Fitting the negative binomial distribution to biological data. *Biometrics*, 9(2):176–200. 2, 3, 5
- [3] Clark, S. and Perry, J. (1989). Estimation of the negative binomial parameter κ by maximum quasi-likelihood. *Biometrics*, 45(1):309–316. 3, 5, 7

- [4] Crofton, H. (1971). A quantitative approach to parasitism. *Parasitology*, 62(2):179–193. 2, 10, 11, 12, 16
- [5] Dai, H., Bao, Y., and Bao, M. (2013). Maximum likelihood estimate for the dispersion parameter of the negative binomial distribution. *Statistics & Probability Letters*, 83(1):21–27. 3, 5, 7
- [6] Greene, W. (1994). Accounting for excess zeros and sample selection in poisson and negative binomial regression models. *NYU Working Paper No. EC-94-10*. 3
- [7] Hall, D. (2000). Zero-inflated poisson and binomial regression with random effects: a case study. *Biometrics*, 56(4):1030–1039. 3
- [8] Hynes, H. and Nicholas, W. (1963). The importance of the acanthocephalan polymorphus minutus as a parasite of domestic ducks in the united kingdom. *Journal of Helminthology*, 37(3):185–198. 10
- [9] Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14. 3
- [10] Min, Y. and Agresti, A. (2005). Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling*, 5(1):1–19. 3
- [11] Piegorsch, W. (1990). Maximum likelihood estimation for the negative binomial dispersion parameter. *Biometrics*, 46(3):863–867. 3, 5, 7
- [12] Pullan, R. and Brooker, S. (2012). The global limits and population at risk of soil-transmitted helminth infections in 2010. *Parasites & vectors*, 5(81):1–14. 10
- [13] Saha, K. and Paul, S. (2005). Bias-corrected maximum likelihood estimator of the negative binomial dispersion parameter. *Biometrics*, 61(1):179–185. 7
- [14] Seo, B., Cho, S., and Chai, J. (1979). Frequency distribution of ascaris lumbricoides in rural koreans with special reference on the effect of changing endemicity. *Korean J Parasitol*, 17(2):105–113. 2, 10, 11, 15
- [15] Shaw, D., Grenfell, B., and Dobson, A. (1998). Patterns of macroparasite aggregation in wildlife host populations. *Parasitology*, 117(6):597–610. 2

- [16] Welsh, A., Cunningham, R., Donnelly, C., and Lindenmayer, D. (1996). Modelling the abundance of rare species: statistical models for counts with extra zeros. *Ecological Modelling*, 88(1):297–308. 3
- [17] Willson, L., Folks, J., and Young, J. (1984). Multistage estimation compared with fixed-sample-size estimation of the negative binomial parameter k . *Biometrics*, 40(1):109–117. 7
- [18] Woolhouse, M., Dye, C., Etard, J., Smith, T., Charlwood, J., Garnett, G., Hagan, P., Hii, J., Ndhlovu, P., Quinnell, R., et al. (1997). Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proceedings of the National Academy of Sciences*, 94(1):338–342. 3

7 Tables

Table 1: Parameters of NB and ZIG distributions calculated from observed Seo data [14] and results of chi-squared test and AIC

Theoretical distribution	Calculated parameters	Samples					
		A ($n = 540$)	B ($n = 136$)	C ($n = 32$)	D ($n = 47$)	E ($n = 39$)	F ($n = 59$)
NB	m	1.0167	2.8235	2.3125	2.5106	6.6410	4.6102
	k	0.3546	0.4240	0.5761	0.5893	0.8726	0.6193
	chi-squared statistic	50.0660	25.8883	6.8092	17.4885	15.6914	11.2927
	p -value	< 0.0001	0.0556	0.8699	0.3547	0.4747	0.7911
	AIC	1438.0235	574.5621	131.7761	198.1410	235.0388	310.0239
ZIG	π	0.3653	0.2687	0.2011	0.1819	0.0971	0.1581
	p	0.3843	0.2057	0.2568	0.2458	0.1197	0.1544
	chi-squared statistic	213.1743	43.2591	6.8197	22.2336	14.4552	10.6473
	p -value	< 0.0001	0.0003	0.8693	0.1358	0.5648	0.8307
	AIC	1458.0335	579.2921	131.5360	198.1585	233.3277	308.9315
	df	16	16	12	16	16	16

Table 2: Parameters of NB and ZIG distributions calculated from observed of parasite in crabs data [4] and results of chi-squared test and AIC.

Theoretical distribution	Calculated parameters	Samples					
		A ($n = 549$)	B ($n = 509$)	C ($n = 633$)	D ($n = 486$)	E ($n = 276$)	F ($n = 191$)
NB	m	2.2732	1.4165	0.6003	1.3189	0.8913	0.2670
	k	1.2564	1.5837	0.2974	3.0544	1.2679	0.6069
	chi-squared statistic	20.6558	3.1086	10.5075	2.9993	2.3843	0.2776
	p -value	0.0081	0.8748	0.0621	0.8089	0.6655	0.8704
	AIC	2211.4460	1662.1623	1279.5742	1506.5751	724.9286	252.4139
ZIG	π	-0.0256	-0.1304	0.4875	-0.3313	-0.1020	0.2195
	p	0.3109	0.4438	0.4605	0.5023	0.5528	0.7451
	chi-squared statistic	23.4185	6.0467	6.5825	8.6706	2.1793	0.4542
	p -value	0.0029	0.5343	0.2536	0.1930	0.7028	0.7969
	AIC	2215.4026	1665.9254	1274.8293	1514.0280	724.8346	252.4989
	df	8	7	5	6	4	2