ISSN 0972-0871

Reprinted from the

Far East Journal of
Mathematical Sciences (FJMS)
Volume 56, Number 2, 2011, pp 161-178

MODELLING DISPERSION USING FINITE MIXTURE OF POISSON

by

I. C. Kipchirchir

Pushpa Publishing House
Vijaya Niwas, 198 Mumfordganj
Allahabad 211002, INDIA
http://pphmj.com/journals/fjms.htm
fjms@pphmj.com & arun@pphmj.com

Information for Authors

Aims and Scope: The *Far East Journal of Mathematical Sciences (FJMS)* is devoted to publishing original research papers and critical survey articles in the field of *Pure and Applied Mathematics*, *Computer Applications* and *Applied Statistics. The FJMS is a fortnightly journal published in twelve volumes annually and each volume comprises of two issues.*

Indexing and Reviews: Mathematical Reviews, MathSciNet, IndexCopernicus Data, Zentralblatt für Mathematik, SCOPUS, Elsevier's bibliographic database, Ei databases index, EMBASE, EMCare.

Submission of Manuscripts: Authors may submit their papers for consideration in the Far East Journal of Mathematical Sciences (FJMS) by the following modes:

1. **Online submission (only .pdf and .doc files):** Please visit journal's homepage at http://www.pphmj.com/journals/fjms.htm
2. **Electronically (only .tex, .dvi, .pdf, .ps and .doc files):** At the e-mail address: fjms@pphmj.com or kkazad@pphmj.com
3. **Hard copies:** Papers in duplicate with a letter of submission at the address of the publisher.

The paper must be typed only on one side in double spacing with a generous margin all round. An effort is made to publish a paper duly recommended by a referee within a period of three months. One set of galley proofs of a paper will be sent to the author submitting the paper, unless requested otherwise, without the original manuscript, for corrections.

Abstract and References: Authors are requested to provide an abstract of not more than 250 words and latest Mathematics Subject Classification. Statements of Lemmas, Propositions and Theorems should be set in *italics* and references should be arranged in alphabetical order by the surname of the first author.

Page Charges and Reprints: Authors are requested to arrange page charges of their papers @ USD 40.00 per page for USA and Canada, and EUR 30.00 per page for rest of the world from their institutions/research grants, if any. However, for authors in India this charge is Rs. 500.00 per page. No extra charges for colour figures. Twenty-five reprints are provided to the corresponding author ex-gratis along with a copy of the issue in which the author's paper is appeared. Additional sets of reprints may be ordered at the time of proof correction.

Subscription Information for 2011

Institutional Price for all countries except India

Electronic Subscription	€ 840.00	US$ 1140.00
Print Subscription includes Online Access	€ 995.00	US$ 1445.00

For Institutions: On seeking a license for volume(s) of the Far East Journal of Mathematical Sciences (FJMS), the facility to download and print the articles will be available through the institutional 9 digits **IP address** to be provided by the appropriate authority. The facility to download will continue till the end of the next calendar year from the last issue of the volume subscribed. For having continued facility to keep the download of the same subscribed volume for another two calendar years may be had on a considerable discounted rate.

Free Online Access for Foreign Institutions with Print Subscription

Price in Indian Rs. (For Indian Institutions in India only)

Print Subscription Only	Rs. 12000.00

The subscription year runs from January 1, 2011 through December 31, 2011.

Information: The journals published by the "Pushpa Publishing House" are solely distributed by the "**Vijaya Books and Journals Distributors**".

Contact Person: Subscription Manager, Vijaya Books and Journals Distributors, Vijaya Niwas, 198 Mumfordganj, Allahabad 211002, India; sub@pphmj.com ; arun@pphmj.com

Far East Journal of Mathematical Sciences (FJMS)
Volume 56, Number 2, 2011, Pages 161-178
This paper is available online at http://pphmj.com/journals/fjms.htm

MODELLING DISPERSION USING FINITE MIXTURE OF POISSON

I. C. KIPCHIRCHIR

School of Mathematics
University of Nairobi
P. O. Box 30197-00100, Nairobi, Kenya
e-mail: kipchirchir@uonbi.ac.ke

Abstract

The negative binomial distribution is a versatile distribution in describing dispersion. The negative binomial parameter k is considered as a dispersion parameter. The aim of this paper is to demonstrate that finite mixture of Poisson can be used in modelling dispersion. The digamma function and the Sterling's expansion for the gamma function are used to construct a dispersion parameter in relation to the negative binomial parameter k. The construction is based on the hierarchical maximum likelihood estimation of the negative binomial parameter k. The method of moments estimates of the parameters of the finite mixture of Poisson is used in the analysis.

1. Introduction

Dispersion is the description of the pattern of distribution of organisms in space (Southwood [6]) and often referred to as *spatial distribution*. It is a characteristic ecological property. Probability distributions are used to quantify and classify the dispersion of organisms. If the mean and variance are equal, then the spatial distribution is said to be *random* and the population pattern is said to be *random*. If

2010 Mathematics Subject Classification: 62E99.

Keywords and phrases: dispersion, negative binomial, Poisson, finite mixture, hierarchical estimation, measure of dispersion.

Received May 18, 2011

variance is greater than the mean, then the spatial distribution is said to be *contagious* and the population pattern is said to be *overdispersed* or *clumped* or *patchy* or *aggregated* or *clustered*. If variance is less than the mean, then spatial distribution is said to be *regular* and the population pattern is said to be *underdispersed* or *uniform*. Regular distribution is seldom observed unless during presence-absence sampling.

Many overdispersed pest populations that have been studied can adequately be described by the negative binomial distribution

$$p_x = \binom{k+x-1}{x}\left(\frac{p}{1+p}\right)^x\left(\frac{1}{1+p}\right)^k, \quad x = 0, 1, 2, \ldots; \; k > 0, \; p > 0 \tag{1}$$

so that $v = m(1+p) > m = kp$ implying it describes a contagious distribution.

However, for fixed $kp = \lambda$,

$$\lim_{k\to\infty} p_x = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots; \; \lambda > 0, \tag{2}$$

which is the Poisson distribution and describes a random distribution $(v = m = \lambda)$.

The positive exponent, k, is considered as a dispersion (or an aggregation) parameter (Anscombe [1]), so that the negative binomial distribution describes contagion $(k \to 0)$ and randomness $(k \to \infty)$.

Young and Young [7] reviewed measures of aggregation namely, variance to mean ratio, index of clumping, index of mean crowding and index of patchiness with respect to Poisson and negative binomial distributions. The four measures of aggregation revealed that decreasing values of k are associated with increasing measures of aggregation (departure from randomness).

Kipchirchir [4] demonstrated analytically that the negative binomial parameter k is a measure of dispersion by analyzing equicorrelation matrix in relation to coefficient of determination, partial correlation and principal components with respect to k. The analysis demonstrated that small values of k are associated with overdispersion whereas large values are associated with randomness.

Now, for given λ, let the number of individuals per unit X have a Poisson distribution with parameter λ, that is,

$$p_{x/\lambda} = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, \ldots; \; \lambda > 0 \tag{3}$$

so that

$$E(X/\lambda) = Var(X/\lambda) = \lambda \tag{4}$$

and the pattern would be random.

Suppose that some units provide more favourable environment than others (units are dissimilar). Then λ, the expected number of individuals in a unit, varies from unit to unit, that is, the environment is heterogeneous resulting in contagion. In particular, we assume that λ is a realization of a random variable Λ having a gamma density (Pearson Type III Distribution)

$$g(\lambda) = \frac{\lambda^{k-1}}{\Gamma(k)\,p^k} e^{-\lambda/p}, \quad \lambda,\ k,\ p > 0 \tag{5}$$

and the marginal distribution of X is

$$\begin{aligned} p_x &= \int_0^\infty p_{x/\lambda}\, g(\lambda)\, d\lambda \\ &= \frac{1}{\Gamma(k)\,x!\,p^k} \int_0^\infty \lambda^{k+x-1} e^{-(1+p)\lambda/p} d\lambda \\ &= \frac{\Gamma(k+x)}{\Gamma(k)\,x!} \left(\frac{p}{1+p}\right)^x \left(\frac{1}{1+p}\right)^k, \quad x = 0, 1, 2, \ldots;\ k,\ p > 0 \end{aligned} \tag{6}$$

which is the negative binomial distribution.

In the Bayesian context, the Poisson distribution is referred to as the *likelihood* and the gamma density as the *prior distribution of* Λ. The Bayes estimate of λ depends on the distribution function $G(\lambda)$ and an empirical Bayes estimate of λ depends on an empirical distribution function which is an estimate of $G(\lambda)$.

In our present context, the Poisson distribution is referred to as the *kernel*, the negative binomial distribution as a *continuous mixture of Poisson* and the gamma as the *mixing distribution*. Moreover, $G(\lambda)$ is identifiable in the continuous mixture of Poisson since the factorial moment generating function of the negative binomial distribution is the moment generating function of the Pearson Type III Distribution. In the sequel, we consider discrete counterpart of a continuous mixture of Poisson which is a finite mixture of Poisson.

2. Finite Mixture of Poisson

et p_{x/λ_j}, $j = 1, 2, 3, \ldots, q$ be a family of Poisson probability distributions oisson kernels). Then

$$p_x = \sum_{j=1}^{q} p_j p_{x/\lambda_j}, \quad x = 0, 1, 2, \ldots; \; \lambda_j > 0, \tag{7}$$

$$p_j > 0 \; \forall j \; \text{ and } \sum_{j=1}^{q} p_j = 1 \tag{8}$$

finite mixture of Poisson probability distributions. The parameters p_j, , 2, 3, ..., q are called *mixing proportions* of the finite mixture. Direct ıation on p_x is supplied only by n observations on the random variable X.

inite mixtures can be used to describe some heterogeneous population which e regarded as being composed of a finite number of more homogeneous pulations. A useful result is obtained by considering kernel distributions such

$$E(X^r/\lambda) = \sum_{x=0}^{\infty} x^r p_{x/\lambda} = \sum_{i=0}^{r} a_i \lambda^i \tag{9}$$

is a polynomial of degree r in λ. Now, the rth raw moment of p_x is

$$\mu'_r = \sum_{x=0}^{\infty} x^r p_x \tag{10}$$

sing (7) we obtain

$$\mu'_r = \sum_{x=0}^{\infty} x^r \sum_{j=1}^{q} p_j p_{x/\lambda_j} = \sum_{j=1}^{q} p_j \sum_{x=0}^{\infty} x^r p_{x/\lambda_j}. \tag{11}$$

Ising (9) in (11), we obtain

$$\mu'_r = \sum_{j=1}^{q} p_j \left(\sum_{i=0}^{r} a_i \lambda_j^i \right) = \sum_{i=0}^{r} a_i \left(\sum_{j=1}^{q} p_j \lambda_j^i \right) = \sum_{i=0}^{r} a_i \alpha_i, \quad r = 1, 2, 3, \ldots, \tag{12}$$

where

$$\sum_{j=1}^{q} p_j \lambda_j^i = \alpha_i, \quad i = 0, 1, 2, ..., r \tag{13}$$

which is a linear system in powers of λ_j.

According to Everitt and Hand [2], to find the method of moments estimates of λ_j and p_j, $j = 1, 2, ..., q$, we require $2q$ equations given by (13), namely

$$\sum_{j=1}^{q} p_j \lambda_j^i = \alpha_i, \quad i = 0, 1, 2, ..., 2q - 1. \tag{14}$$

Now, suppose that we have found constants $\beta_1, \beta_2, ..., \beta_q$ such that $\lambda_1, \lambda_2, ..., \lambda_q$ are roots of

$$\lambda^q - \beta_1 \lambda^{q-1} - \beta_2 \lambda^{q-2} - \cdots - \beta_{q-1}\lambda - \beta_q = 0 \tag{15}$$

or equivalently

$$\sum_{i=0}^{q} \beta_{q-i} \lambda^i = 0 \text{ with } \beta_0 = -1, \tag{16}$$

then by multiplying the ith equation in (14) by β_{q-i} for $i = 0, 1, 2, ..., q-1$ and the qth by -1 and adding, we get

$$\sum_{i=0}^{q-1} \beta_{q-i} \left(\sum_{j=1}^{q} p_j \lambda_j^i \right) - \sum_{j=1}^{q} p_j \lambda_j^q = \sum_{i=0}^{q-1} \alpha_i \beta_{q-i} - \alpha_q,$$

that is,

$$\sum_{j=1}^{q} p_j \left(\sum_{i=0}^{q-1} \beta_{q-i} \lambda_j^i - \lambda_j^q \right) = \sum_{i=0}^{q-1} \alpha_i \beta_{q-i} - \alpha_q \tag{17}$$

which by virtue of (16) simplifies to

$$\sum_{i=0}^{q-1} \alpha_i \beta_{q-i} = \alpha_q. \tag{18}$$

Similarly, if we multiply $(i+1)$th equation in (14) by β_{q-i} for $i = 0, 1, 2, \ldots, q-1$ and the $(q+1)$th by -1 and adding, we get

$$\sum_{i=0}^{q-1} \alpha_{i+1}\beta_{q-i} = \alpha_{q+1}. \tag{19}$$

Constructing in this way, we can set up a system of q linear equations

$$\sum_{i=0}^{q-1} \alpha_{i+s}\beta_{q-i} = \alpha_{q+s}, \quad s = 0, 1, 2, \ldots, q-1 \tag{20}$$

which can be expressed in matrix notation as

$$\begin{pmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_{q-1} \\ \alpha_1 & \alpha_2 & \cdots & \alpha_q \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{q-1} & \alpha_q & \cdots & \alpha_{2q-2} \end{pmatrix} \begin{pmatrix} \beta_q \\ \beta_{q-1} \\ \vdots \\ \beta_1 \end{pmatrix} = \begin{pmatrix} \alpha_q \\ \alpha_{q+1} \\ \vdots \\ \alpha_{2q-1} \end{pmatrix}, \tag{21}$$

that is,

$$A\underline{\beta} = \underline{\alpha}. \tag{22}$$

If the p_j's are non-zero and the λ_j's are distinct, A is non-singular and can be inverted to give

$$\hat{\underline{\beta}} = A^{-1}\underline{\alpha}. \tag{23}$$

The estimated β_i's are then substituted in (15) and solved for the λ_j, $j = 1, 2, \ldots, q$.

Next, consider the moments of the finite mixture (7), that is,

$$\mu'_r = \sum_{j=1}^{q} p_j \mu'_{rj}(\lambda_j), \tag{24}$$

where μ'_r is the rth raw moment of p_x which can be estimated by the sample raw moment m'_r and $\mu'_{rj}(\lambda_j)$ is the rth raw moment of the jth component of the finite mixture. In view of (8), we obtain estimates of the mixing proportions $p_1, \ldots, p_{q-1}$

empirical Bayes estimation, the mixing distribution being a step-function G_q, can be used as an approximation to $G(\lambda)$, moreover, G_q is identifiable in the finite mixture of Poisson.

We shall assume population interpretation of a prior distribution, that is, the prior distribution represents a population of possible parameter values, from which the λ of current interest has been drawn (Gelman et al. [3]). In other words we interpret the parameters λ_j's as 'observed' values of Λ_j's from a Pearson Type III distribution with hyperparameters k and p. Thus,

$$g(\lambda_j / k,\ p) = \frac{\lambda_j^{k-1}}{p^k \Gamma(k)} e^{-\lambda_j / p}, \quad k,\ p,\ \lambda_j > 0; \quad j = 1, 2, 3, \ldots, q \tag{31}$$

and assume the Λ_j's are conditionally independent given $(k,\ p)$ and we define the likelihood function

$$L(k,\ p) = \prod_{j=1}^{q} g(\lambda_j / k,\ p) = \prod_{j=1}^{q} \frac{\lambda_j^{k-1}}{p^k \Gamma(k)} e^{-\lambda_j / p} \tag{32}$$

for the hyperparameters k and p. A hierarchical model permits the interpretation of the Λ_j's as a random sample from a shared population distribution (Gelman et al. [3]). Hierarchically, X has information about λ which is a realization of Λ which has information about k.

Now, to determine the values of k and p which maximize the likelihood function, we consider the log-likelihood function

$$\ln L(k,\ p) = q(-k \ln p - \ln \Gamma(k)) + (k-1) \sum_{j=1}^{q} \ln \lambda_j - \frac{1}{p} \sum_{j=1}^{q} \lambda_j. \tag{33}$$

Differentiating (33) with respect to p, we obtain

$$\frac{\partial \ln L(k,\ p)}{\partial p} = -\frac{qk}{p} + \frac{1}{p^2} \sum_{j=1}^{q} \lambda_j = 0 \tag{34}$$

which simplifies to

$$kp = \frac{1}{q} \sum_{j=1}^{q} \lambda_j \tag{35}$$

and on taking log we obtain

$$\ln k + \ln p = \ln\left(\frac{1}{q}\sum_{j=1}^{q}\lambda_j\right). \tag{36}$$

Differentiating (33) with respect to k, we obtain

$$\frac{\partial \ln L(k, p)}{\partial k} = q\left(-\ln p - \frac{\partial \ln \Gamma(k)}{\partial k}\right) + \sum_{j=1}^{q} \ln \lambda_j = 0 \tag{37}$$

yielding

$$\frac{\partial \ln \Gamma(k)}{\partial k} + \ln p = \frac{1}{q}\sum_{j=1}^{q} \ln \lambda_{j}: \tag{38}$$

Next, we combine (36) and (38) and obtain the reduced maximum likelihood equation for determining estimate of k as

$$\ln k - \frac{\partial \ln \Gamma(k)}{\partial k} = \ln\left(\frac{1}{q}\sum_{j=1}^{q}\lambda_j\right) - \frac{1}{q}\sum_{j=1}^{q}\ln\lambda_j = \nu \tag{39}$$

which we generalize to

$$\ln k - \frac{\partial \ln \Gamma(k)}{\partial k} = \ln\left(\sum_{j=1}^{q} p_j\lambda_j\right) - \sum_{j=1}^{q} p_j \ln\lambda_j = \nu \tag{40}$$

so that if $p_j = 1/q$ for all j, then (40) reduces to (39).

To solve the likelihood equation, we shall use the Sterling's expansion for the gamma function, that is,

$$\ln \Gamma(k) \approx \left(k - \frac{1}{2}\right)\ln k - k + \frac{1}{2}\ln(2\pi) + \xi(k), \tag{41}$$

where

$$\xi(k) = \sum_{i=1}^{\infty} \frac{(-1)^{i-1} B_i}{2i(2i-1)k^{2i-1}} \tag{42}$$

with

$$B_1 = \frac{1}{6},\ B_2 = \frac{1}{30},\ B_3 = \frac{1}{42},\ B_4 = \frac{1}{30} \tag{43}$$

and is a convergent series (by alternating series test). The derivative of (41) is given by

$$\begin{aligned} \frac{\partial \ln \Gamma(k)}{\partial k} &\approx \ln k - \frac{1}{2k} + \xi'(k) \\ &= \ln k - \frac{1}{2k} + \sum_{i=1}^{\infty} \frac{(-1)^i B_i}{2ik^{2i}} \\ &= \ln k - \frac{1}{2k} - \frac{1}{12k^2} + \frac{1}{120k^4} - \frac{1}{252k^6} + \frac{1}{240k^8} - \cdots. \end{aligned} \tag{44}$$

An approximation of $\hat{k}$ can be obtained by ignoring terms of $O(k^{-1})$ in (44) so that

$$\frac{\partial \ln \Gamma(k)}{\partial k} \approx \ln k - \frac{1}{2k} \tag{45}$$

and the maximum likelihood equation (40) yields

$$k \approx \frac{1}{2\nu}. \tag{46}$$

A better approximation can be obtained by ignoring terms of $O(k^{-2})$ in (44) so that

$$\frac{\partial \ln \Gamma(k)}{\partial k} \approx \ln k - \frac{1}{2k} - \frac{1}{12k^2} \tag{47}$$

and the maximum likelihood equation (40) becomes

$$12\nu k^2 - 6k - 1 \approx 0 \tag{48}$$

yielding

$$k \approx \frac{1 + \sqrt{1 + \frac{4\nu}{3}}}{4\nu}. \tag{49}$$

An even better approximation can be obtained by using (44) and the maximum likelihood equation (40) becomes

$$\nu = \ln k - \frac{\partial \ln \Gamma(k)}{\partial k} \approx \frac{1}{2k} + \frac{1}{12k^2} - \frac{1}{120k^4} + \frac{1}{252k^6} - \frac{1}{240k^8} + \cdots \tag{50}$$

or equivalently the function

$$f(k) = 240\nu k^8 - 120k^7 - 20k^6 + 2k^4 - k^2 + 1 \approx 0. \tag{51}$$

An initial value k_1 can be obtained from (49) and better approximation can be generated by the Newton-Raphson iteration formula

$$k_{i+1} = k_i - \frac{f(k_i)}{f'(k_i)}, \quad i = 1, 2, 3, \cdots. \tag{52}$$

Table 1. Comparison of values of k for arbitrary values of ν

ν	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	γ	1	2	3
$k = \frac{1}{2\nu}$	2	1.5	1	0.866	0.5	0.25	0.167
$k = \frac{1+\sqrt{1+4\hat{\nu}/3}}{4\hat{\nu}}$	2.155	1.651	1.145	1.009	0.632	0.364	0.270
$k_{i+1} = k_i - \frac{f(k_i)}{f'(k_i)}$	2.152	1.647	1.136	0.993	0.558	0.364	0.270

$\gamma = 0.5772156$ is the Euler's constant.

We observe that values obtained by (49) and those obtained by (52) are more or less the same and hence for all practical purposes it suffices to use (49).

4. Measure of Dispersion with Respect to Finite Mixture

In Table 1, we observe that k decreases as ν increases (k increases as ν decreases). Generally, we find the limit of ν as $k \to \infty$ by considering the digamma function

$$\frac{d \ln \Gamma(k)}{dk} = \frac{\Gamma'(k)}{\Gamma(k)} = -\gamma + \sum_{y=1}^{\infty}\left(\frac{1}{y} - \frac{1}{k+y-1}\right), \tag{53}$$

where γ is the Euler's constant. This infinite series is convergent by integral test. Expanding the digamma function, we have

$$\frac{\Gamma'(k)}{\Gamma(k)} = -\gamma + \left(\frac{1}{1} - \frac{1}{k}\right) + \left(\frac{1}{2} - \frac{1}{k+1}\right) + \left(\frac{1}{3} - \frac{1}{k+2}\right) + \cdots \tag{54}$$

which can be expressed as

$$\frac{\Gamma'(k)}{\Gamma(k)} = -\gamma + \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k} + \frac{1}{k+1} + \frac{1}{k+2} + \cdots\right)$$

$$- \left(\frac{1}{k} + \frac{1}{k+1} + \frac{1}{k+2} + \cdots\right)$$

$$= -\gamma - \frac{1}{k} + \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k}\right) \tag{55}$$

so that from (40)

$$\nu = \ln k - \frac{\Gamma'(k)}{\Gamma(k)} = \gamma + \frac{1}{k} - \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k} - \ln k\right). \tag{56}$$

In particular, $\nu = \gamma$ when $k = 1$ which corresponds to the geometric distribution.

Since

$$\lim_{k \to \infty}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k} - \ln k\right) = \gamma, \tag{57}$$

on taking limit of (56), we obtain

$$\lim_{k \to \infty} \nu = \lim_{k \to \infty}\left(\gamma + \frac{1}{k} - \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k} - \ln k\right)\right) = 0. \tag{58}$$

Thus, as $k \to \infty$ (randomness), $\nu \to 0$, but from (40),

$$\nu \to 0 \Leftrightarrow \lambda_j \to \lambda \ \forall j \tag{59}$$

and consequently the finite mixture

$$p_x = \sum_{j=1}^{q} p_j p_{x/\lambda_j} \to p_{x/\lambda} \sum_{j=1}^{q} p_j = p_{x/\lambda}, \tag{60}$$

a single Poisson which describes randomness.

On the other hand as $k \to 0$ (overdispersion), then $\nu \to \infty$ and we conclude from (40) that λ_j, $j = 1, 2, 3, \ldots, q$ are distinct and hence the population is heterogeneous having a finite number of more homogeneous subpopulations. In other words, there is clustering (overdispersion).

Thus, ν as in (40) can be used as a dispersion parameter when using finite mixture of Poisson in describing dispersion.

5. Illustration

We shall use the data in Table 2 generated by a mixture of Poisson and gamma distributions where gamma distribution is the Pearson Type III with parameters $k = 10$ and $p = \frac{1}{2}$. The mixture distribution is the negative binomial with parameters $k = 10$ and $p = \frac{1}{2}$.

Table 2. Mixture of Poisson and Pearson Type III Data

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
n_x	0	3	8	10	2	11	4	4	0	1	2	4	0	1	0	0	0

Source: Maritz and Lwin [5]

As per the following seven categories:

$$x = \overbrace{0, 1, 2}^{1}, \overbrace{3}^{2}, \overbrace{4}^{3}, \overbrace{5}^{4}, \overbrace{6}^{5}, \overbrace{7, 8}^{6}, \overbrace{9, 10, 11, 12, 13, 14, 15, 16}^{7};$$

these data fit negative binomial distribution with parameters $k = 10$ and $p = \frac{1}{2}$ at 5% level of significance. For the data in Table 2, $n = 50$, $\bar{x} = 5$, and the method of moments estimate of negative binomial parameter k is 6.068.

The first four raw moments of Poisson kernel are

$$E(X/\lambda) = \lambda, \quad E(X^2/\lambda) = \lambda^2 + \lambda,$$

$$E(X^3/\lambda) = \lambda^3 + 3\lambda^2 + \lambda, \quad E(X^4/\lambda) = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda. \tag{61}$$

5.1. Estimation of k using G_3

From (21) and (24) with $\alpha_r = \mu'_r$, which are estimated by m'_r, we have

$$\begin{pmatrix} 1 & 5 & 34.12 \\ 5 & 34.12 & 286.52 \\ 34.12 & 286.52 & 2736.04 \end{pmatrix} \begin{pmatrix} \beta_3 \\ \beta_2 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 286.52 \\ 2736.04 \\ 28239.80 \end{pmatrix}$$

yielding

$$\begin{pmatrix} \hat{\beta}_3 \\ \hat{\beta}_2 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} 138.78 \\ -101.88 \\ 19.26 \end{pmatrix}$$

and from (15), we solve

$$\lambda^3 - \hat{\beta}_1\lambda^2 - \hat{\beta}_2\lambda - \hat{\beta}_3 = 0$$

yielding the roots

$$\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3 \end{pmatrix} = \begin{pmatrix} 2.115 \\ 5.768 \\ 11.377 \end{pmatrix}.$$

To obtain the estimates of the mixing proportions we have from (27), (28) and (61)

$$\begin{pmatrix} 9.261 & 5.610 \\ 134.220 & 101.788 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 6.377 \\ 106.693 \end{pmatrix}$$

yielding

$$\begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \end{pmatrix} = \begin{pmatrix} 0.2665 \\ 0.6968 \end{pmatrix}$$

and finally from (8)

$$\hat{p}_3 = 1 - \hat{p}_1 - \hat{p}_2 = 0.0367.$$

From (40), an estimate of ν is

$$\hat{\nu} = \ln\left(\sum_{j=1}^{3} \hat{p}_j \hat{\lambda}_j\right) - \sum_{j=1}^{3} \hat{p}_j \ln \hat{\lambda}_j = 0.0995498$$

and Table 3 gives estimates of k with respect to linear, quadratic and iteration formulae.

Table 3. Estimates of k

Formula	$k = \frac{1}{2\nu}$	$k = \frac{1 + \sqrt{1 + 4\nu/3}}{4\nu}$	$k_{i+1} = k_i - \frac{f(k_i)}{f'(k_i)}$
$\hat{k}$	5.02	5.18	5.16

5.2. Estimation of k using G_4

From (21) and (24) with $\alpha_r = \mu'_r$, which are estimated by m'_r, we have

$$\begin{pmatrix} 1 & 5 & 34.12 & 286.52 \\ 5 & 34.12 & 286.52 & 2736.04 \\ 34.12 & 286.52 & 2736.04 & 28239.80 \\ 286.52 & 2736.04 & 28239.80 & 305791.72 \end{pmatrix} \begin{pmatrix} \beta_4 \\ \beta_3 \\ \beta_2 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 2736.04 \\ 28239.80 \\ 305791.72 \\ 3416182.52 \end{pmatrix}$$

yielding

$$\begin{pmatrix} \hat{\beta}_4 \\ \hat{\beta}_3 \\ \hat{\beta}_2 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} -834.06596 \\ 831.17729 \\ -245.40271 \\ 27.17911 \end{pmatrix}$$

and from (15), we solve

$$\lambda^4 - \hat{\beta}_1\lambda^3 - \hat{\beta}_2\lambda^2 - \hat{\beta}_3\lambda - \hat{\beta}_4 = 0$$

yielding the roots

$$\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3 \\ \hat{\lambda}_4 \end{pmatrix} = \begin{pmatrix} 1.724307 \\ 4.538120 \\ 8.788438 \\ 12.128250 \end{pmatrix}.$$

To obtain the estimates of the mixing proportions we have from (27), (28) and (61)

$$\begin{pmatrix} 10.403943 & 7.59013 & 3.339812 \\ 154.52516 & 134.09004 & 73.197618 \\ 2221.6391 & 2077.6277 & 1318.122 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 7.12825 \\ 125.1027 \\ 1950.8898 \end{pmatrix}$$

yielding

$$\begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \\ \hat{p}_3 \end{pmatrix} = \begin{pmatrix} 0.1123815 \\ 0.7087769 \\ 0.1734621 \end{pmatrix}$$

and finally from (8)

$$\hat{p}_4 = 1 - \hat{p}_1 - \hat{p}_2 - \hat{p}_3 = 0.0053795.$$

From (40), an estimate of ν is

$$\hat{\nu} = \ln\left(\sum_{j=1}^{4} \hat{p}_j \hat{\lambda}_j \right) - \sum_{j=1}^{4} \hat{p}_j \ln \hat{\lambda}_j = 0.0857416$$

and Table 4 gives estimates of k with respect to linear, quadratic and iteration formulae

Table 4. Estimates of k

Formula	$k := \frac{1}{2\nu}$	$k = \frac{1 + \sqrt{1 + 4\nu/3}}{4\nu}$	$k_{i+1} = k_i - \frac{f(k_i)}{f'(k_i)}$
$\hat{k}$	5.831	5.994	5.993

5.3. Discussion on optimal G_q

The case for $q = 5$ fails to estimate $G(\lambda)$ since $\hat{p}_1 < 0$. The case for $q = 6$ fails to estimate $G(\lambda)$ since the matrix A in (23) for estimating β_j's is singular which means either the λ_j's are not distinct or $p_j = 0$ for some j. Thus, G_4 is the optimal estimate of $G(\lambda)$. Geometrically, a smoothed G_4 is closer to $G(\lambda)$ and hence is a good estimate of $G(\lambda)$.

6. Conclusion

The average number of individuals per unit within the 50 units sampled as in Table 1, is as follows:

$$x = \overbrace{1, 2}^{1.73}, \overbrace{3, 4, 5, 6}^{4.33}, \overbrace{7, 9, 10}^{8.14}, \overbrace{11, 13}^{11.4}$$

which mirrors the parameter estimates of the Poisson kernels, namely, $\hat{\lambda}_1 = 1.72$, $\hat{\lambda}_2 = 4.53$, $\hat{\lambda}_3 = 8.79$, $\hat{\lambda}_4 = 12.13$. Thus, the finite mixture describes a heterogeneous population which is composed of four homogeneous subpopulations. This is reminiscent of overdispersion.

In particular, the estimate of the four component finite mixture dispersion parameter is $\hat{\nu} = 0.0857416$ and the corresponding estimate of the negative binomial dispersion parameter estimated using G_4 is $\hat{k} = 5.993$. In fact, G_4 yielded more or less the same estimate as the negative binomial method of moments estimate (6.068).

Thus, finite mixture of Poisson can be used to describe dispersion with ν as a measure of dispersion.

Acknowledgement

I acknowledge EAUMP-ISP for their partial support.

References

[1] F. J. Anscombe, The statistical analysis of insect counts based on the negative binomial distribution, Biometrics 5 (1949), 165-174.

[2] B. S. Everitt and D. J. Hand, Finite Mixture Distributions, Chapman and Hall, 1981.

[3] A. Gelman, J. B. Carlin, H. S. Stern and D. B. Rubin, Bayesian Data Analysis, Chapman and Hall CRC, 1995.

[4] I. C. Kipchirchir, The negative binomial parameter k as a measure of dispersion, ICASTOR J. Math. Sci. 4(2) (2010), 197-207.

[5] J. S. Maritz and T. Lwin, Empirical Bayes Methods, Chapman and Hall, 1989.

[6] T. R. E. Southwood, Ecological Methods, Methuen and Co. Ltd., 1966.

[7] L. J. Young and J. H. Young, A spatial view of the negative binomial parameter k when describing insect populations, Proceedings of the Kansas State University Conference on Applied Statistics in Agriculture (1990), 13-20.

Our Publications

1. Advances and Applications in Discrete Mathematics (ISSN: 0974-1658)
2. Advances and Applications in Fluid Mechanics (ISSN: 0973-4686)
3. Advances and Applications in Statistics (ISSN: 0972-3617)
4. Advances in Computer Science and Engineering (ISSN: 0973-6999)
5. Advances in Differential Equations and Control Processes (ISSN: 0974-3243)
6. Advances in Fuzzy Sets and Systems (ISSN: 0973-421X)
7. Current Development in Oceanography (ISSN: 0976-6960)
8. Current Development in Theory and Applications of Wavelets (ISSN: 0973-5607)
9. Far East Journal of Applied Mathematics (ISSN: 0972-0960)
10. Far East Journal of Dynamical Systems (ISSN: 0972-1118)
11. Far East Journal of Electronics and Communications (ISSN: 0973-7006)
12. Far East Journal of Mathematical Education (ISSN: 0973-5631)
13. Far East Journal of Mathematical Sciences (FJMS) (ISSN: 0972-0871)
14. Far East Journal of Theoretical Statistics (ISSN: 0972-0863)
15. International Journal of Functional Analysis, Operator Theory and Applications (ISSN: 0975-2919)
16. International Journal of Information Science and Computer Mathematics (ISSN: 1829-4969)
17. International Journal of Materials Engineering and Technology (ISSN: 0975-0444)
18. International Journal of Numerical Methods and Applications (ISSN: 0975-0452)
19. JP Journal of Algebra, Number Theory and Applications (ISSN: 0972-5555)
20. JP Journal of Biostatistics (ISSN: 0973-5143)
21. JP Journal of Fixed Point Theory and Applications (ISSN: 0973-4228)
22. JP Journal of Geometry and Topology (ISSN: 0972-415X)
23. JP Journal of Heat and Mass Transfer (ISSN: 0973-5763)
24. JP Journal of Solids and Structures (ISSN: 0973-5615)

Pushpa Publishing House, Vijaya Niwas, 198, Mumfordganj, Allahabad 211002, India

arun@pphmj.com **www.pphmj.com**