ว.วิทย. มข. 42(1) 84-96 (2557) KKU Sci. J. 42(1) 84-96 (2014)

สมบัติบางประการของตัวประมาณค่าพารามิเตอร์สำหรับการแจกแจงแบบเจนเนอร์ไลด์ปัวส์ซง

Properties of Estimators for Generalized Poisson Distribution

Bunthom Suraporn[1]

บทคัดย่อ

สมบัติบางประการของตัวประมาณค่าพารามิเตอร์สำหรับการแจกแจงแบบเจนเนอร์ไลด์ปัวส์ซง ได้ถูกทำการ ศึกษามาบ้างแล้ว มีหลายท่านได้ศึกษาสมบัติโดยประมาณของตัวประมาณที่ได้มาโดยวิธีโมเมนต์ ตัวประมาณที่ได้มาโดยวิธีภาวะน่าจะเป็นสูงสุด และตัวประมาณที่ได้มาโดยวิธีเบย์เชียลภาวะน่าจะเป็นสูงสุด ในปี 1980 Kumar and Consul ได้หาค่าคาดหวังต่าง ๆ โดยใช้วิธีการประมาณ เขาได้หาค่าความแปรปรวนและความแปรปรวนร่วมโดยประมาณของตัวประมาณค่าที่ได้มาโดยวิธีโมเมนต์ ในปี 1984 Consul and Shoukri ได้หาค่าความแปรปรวนและความแปรปรวนร่วมโดยประมาณของตัวประมาณค่าที่ได้มาโดยวิธีภาวะน่าจะเป็นสูงสุด และในปี 2006 Suraporn, B. ได้หาค่าความแปรปรวนและความแปรปรวนร่วมโดยประมาณของตัวประมาณค่าที่ได้มาโดยวิธีเบย์เชียลภาวะน่าจะเป็นสูงสุด ซึ่งในบทความนี้ได้พิจารณาสมบัติบางประการของตัวประมาณ ความคงเส้นคงวา และพิจารณาประสิทธิภาพสัมพัทธ์ของตัวประมาณ

ABSTRACT

Some properties of estimators for generalized Poisson distribution were considered, they were derived the asymptotic properties of the method of moments estimators (MME), maximum likelihood estimators (MLE) and maximum Bayesian likelihood estimators (MBLE). Kumar and Consul (1980) have obtained their expectations up to the first order approximation. They derived asymptotic variances and the covariance of the method of moments estimators, $\hat{\lambda}^{MME}$ and $\hat{\theta}^{MME}$. Consul and Shoukri (1984) derived the asymptotic variance and the covariance of the maximum likelihood estimators, $\hat{\lambda}^{MLE}$, $\hat{\theta}^{MLE}$ and Suraporn, B. (2006) derived the asymptotic

[1]ภาควิชาคณิตศาสตร์ สถิติและคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

E-mail: scbuntsu@ubu.ac.th

variance and the covariance of the maximum Bayesian likelihood estimators, $\hat{\lambda}^{MBLE}$ and $\hat{\theta}^{MBLE}$. In this paper, some properties of existing estimators, the properties of estimators; consistency, bound and relative efficiency of estimators are considered.

คำสำคัญ: ฟังก์ชันความน่าจะเป็นแบบปัวส์ซง ฟังก์ชันความน่าจะเป็นแบบเจนเนอร์ไลด์ปัวส์ซง การกระจาย ตัวประมาณค่าที่ได้มาโดยวิธีโมเมนต์ ตัวประมาณค่าที่ได้มาโดยวิธีภาวะน่าจะเป็นสูงสุด ตัวประมาณค่าที่ได้มาโดยวิธีเบย์เชียลภาวะน่าจะเป็นสูงสุด

Keywords: Poisson probability distribution, Generalized Poisson distribution, Dispersion, The method of moments estimators, The maximum likelihood estimators, The maximum Bayesian likelihood estimators.

1. Introduction

1.1 Generalized Poisson Distribution (GPD)

Consul and Jain (1973a) proposed a new generalization of the discrete Poisson distribution which was later modified by Consul and Shoukri (1984). In 1989, Consul presented the GPD, a non-negative integer-valued distribution with two parameters λ and θ. A discrete random variable *X* has a generalized Poisson distribution, denoted by GPD (λ,θ), if and only if its probability mass function (pmf) $f(x;\lambda,\theta)$ is given by

$$f(x;\lambda,\theta) = \frac{\lambda(\lambda+\theta x)^{x-1}e^{-\lambda-\theta x}}{x!} \quad (1)$$

for *x* = 0, 1, 2, 3,..., where $\lambda > 0$, and $0 \leq \theta < 1$. Furthermore, all of its moments exist as long as $\theta < 1$. This GPD model accounts for many branching processes and queuing processes. Negative values of θ of this GPD model are considered else where, but in this study on only nonnegative values of θ will be considered.

If $\theta = \delta\lambda$, equation (1) becomes

$$f(x;\lambda,\delta) = \frac{\lambda^{x}(1+\delta x)^{x-1}e^{-\lambda-\delta\lambda x}}{x!} \quad ; x = 0, 1, 2, 3,... \quad (2)$$

where $\lambda > 0$, and $0 \leq \delta < \theta^{-1}$. When $\theta = 0$ the GPD model reduces to the usual Poisson distribution. Consul and Jain (1973a) proved that

$$\sum_{x=0}^{\infty} \frac{\lambda(\lambda+\theta x)^{x-1}e^{-\lambda-\theta x}}{x!} = 1.$$

1.2 Estimation of Parameters

1.2.1 Method of Moments Estimators (MME)

Let $f(x;\lambda,\theta)$ be a pmf of a random variable X from the GPD model (1) with two unknown parameters λ and θ. As before let μ_r' denote the r^{th} moment about 0; that is, $\mu_r' = E(X^r)$. In general $\mu_r' = \mu_r'(\lambda,\theta)$ is a known function of the two parameters λ and θ. Let X_1, X_2, ..., X_n be a random sample from the density $f(x;\lambda,\theta)$, and, let M_j' be the j^{th} sample moment; that is

$$M_j' = \frac{1}{n}\sum_{i=1}^{n} X_i^j \text{ . Put } M_j' = \mu_j'(\lambda,\theta),\ j = 1, 2.$$

The estimators of the two variables λ and θ are obtained by replacing population moments by sample moments. We have

$$\mu_1' = E(X) = \frac{\lambda}{1-\theta} = M_1' = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

$$\mu_2' = E(X^2) = \frac{\lambda}{(1-\theta)^3} + \frac{\lambda^2}{(1-\theta)^2} = M_2' = \frac{1}{n}\sum_{i=1}^{n} X_i^2 .$$

So, $(1-\theta)^2 = \dfrac{\bar{X}}{S^2}$, for $S^2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$, hence, we have

$$\hat{\lambda}^{MME} = \sqrt{\frac{\bar{X}^3}{S^2}} \text{ and } \hat{\theta}^{MME} = 1 - \sqrt{\frac{\bar{X}}{S^2}}. \tag{3}$$

Shoukri (1980) computed the asymptotic variances of the moment estimators $\hat{\lambda}^{MME}$ and $\hat{\theta}^{MME}$ up to the second order of approximation. The computations are very messy and the expressions are very long. Herewith we give their values up to the first order of approximation.

$$V(\hat{\lambda}^{MME}) \approx \frac{\lambda}{2n}\left[\lambda + \frac{2-2\theta+3\theta^2}{1-\theta}\right],$$

$$V(\hat{\theta}^{MME}) \approx \frac{1-\theta}{2n\lambda}\left[\lambda - \lambda\theta + 2\theta + 3\lambda^2\right],$$

and

$$Cov(\hat{\lambda}^{MME}, \hat{\theta}^{MME}) \approx -\frac{1}{2n}\left[\lambda(1-\theta) + 3\theta^2\right]. \tag{4}$$

1.2.2 Maximum Likelihood Estimators (MLE)

The likelihood function of n random variables X_1, X_2, ..., X_n is defined to be the joint density of the n random variables, say $f_{X_1,X_2,...,X_n}(x_1,x_2,...,x_n;\lambda,\theta)$, which is considered to be a function of λ and θ. In particular, if X_1, X_2, ..., X_n are a random sample from the density GPD models (1), $f(x;\lambda,\theta)$, then the likelihood function is $f(x_1;\lambda,\theta)\cdot f(x_2;\lambda,\theta)\cdot\ldots\cdot f(x_n;\lambda,\theta)$, we shall use the notation $L(\lambda,\theta \mid x_1, x_2, ..., x_n) = L(\lambda,\theta \mid \underset{\sim}{x})$ for the likelihood function where (x_1, x_2, ..., x_n) are the values of a random sample observed (X_1, X_2, ..., X_n) of size n.

Consider the likelihood and the log-likelihood functions of GPD:

$$L(\lambda,\theta \mid x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i;\lambda,\theta)$$

$$= \prod_{i=1}^{n}\left[\frac{\lambda(\lambda+\theta x_i)^{x_i-1}}{x_i!}e^{-\lambda-\theta x_i}\right]$$

$$= \lambda^n e^{-n\lambda-\theta\sum_{i=1}^{n} x_i}\prod_{i=1}^{n}\left[\frac{(\lambda+\theta x_i)^{x_i-1}}{x_i!}\right] \quad (5)$$

and $\ln L(\lambda,\theta \mid x_1, x_2, ..., x_n) = n\ln(\lambda) - n\lambda - \theta\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}(x_i-1)\ln(\lambda+\theta x_i) - \sum_{i=1}^{n}\ln(x_i!)$

$$\frac{\partial}{\partial\lambda}\ln L(\lambda,\theta \mid x_1, x_2, ..., x_n) = \frac{n}{\lambda} - n + \sum_{i=1}^{n}\frac{(x_i-1)}{(\lambda+\theta x_i)} = 0$$

$$\sum_{i=1}^{n}\frac{(x_i-1)}{(\lambda+\theta x_i)} = n(1-\frac{1}{\lambda}) \quad (6)$$

$$\frac{\partial}{\partial\theta}\ln L(\lambda,\theta \mid x_1, x_2, ..., x_n) = -\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\frac{x_i(x_i-1)}{(\lambda+\theta x_i)} = 0, \text{ or}$$

$$\sum_{i=1}^{n}\frac{x_i(x_i-1)}{(\lambda+\theta x_i)} = n\bar{x} \quad (7)$$

From equation (6) and (7), we have $\hat{\lambda}^{MLE} = (1-\hat{\theta}^{MLE})\bar{X}$, and

$$H(\hat{\theta}^{MLE}) = \sum_{i=1}^{n}\frac{x_i(x_i-1)}{(\bar{X}-\hat{\theta}^{MLE}\bar{X}+\hat{\theta}^{MLE}x_i)} - n\bar{X} = 0$$

$$\sum_{i=1}^{n}\frac{x_i(x_i-1)}{\bar{X}+\hat{\theta}^{MLE}(x_i-\bar{X})} - n\bar{X} = 0. \quad (8)$$

The asymptotic variances and the covariances of the ML estimators, that is, $\hat{\lambda}^{MLE}, \hat{\theta}^{MLE}$ and $Cov(\hat{\lambda}^{MLE}, \hat{\theta}^{MLE})$ are

$$V(\hat{\lambda}^{MLE}) \approx \frac{\lambda(\lambda+2)}{2n},$$

$$V(\hat{\theta}^{MLE}) \approx \frac{(1-\theta)(\lambda+2\theta-\lambda\theta)}{2n\lambda},$$

$$Cov(\hat{\lambda}^{MLE}, \hat{\theta}^{MLE}) \approx -\frac{\lambda(1-\theta)}{2n}, \text{ (Consul and Shoukri, 1984).} \quad (9)$$

2. Main Results

2.1 Maximum Bayesian Likelihood Estimators (MBLE)

We have assumed that random sample comes from the same density $f(x;\lambda,\theta)$, where the function $f(x;\lambda,\theta)$ is assumed known. Moreover, we have assumed that λ and θ are fixed, though unknown. In some real situations where the density $f(x;\lambda,\theta)$ is attached with some prior information about λ and θ. Bayes-type estimator may be more appropriate.

While Bayes's posterior estimator is the mean of posterior distribution, the maximum Bayesian likelihood (MBL) estimators are defined as the posterior mode, which maximize $f(\lambda,\theta\,|\,\underset{\sim}{x})h(\lambda,\theta)$. This quantity bypasses the computation of the marginal distribution and can be also expressed as a penalized maximum likelihood estimator in the classical sense. If the sample size grows to infinity, maximum Bayesian likelihood estimators are asymptotically equivalent to the classical maximum likelihood estimators.

With reference to equation (5), the likelihood function is

$$L(\lambda,\theta\,|\,\underset{\sim}{x}) = L(\lambda,\theta\,|\,x_1,\, x_2,\, ...,\, x_n) \;=\; \lambda^n e^{-n\lambda-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n}\left[\frac{(\lambda+\theta x_i)^{x_i-1}}{x_i\,!}\right].$$

Given prior distribution of λ and θ that λ is distributed as gamma (a, b), and θ is distributed as uniform (0,1). Assume that λ and θ are a prior independent, so the joint density of (λ,θ) is defined as follows:

$$h(\lambda,\theta) \;=\; \frac{1}{\Gamma(a)} b^a \lambda^{a-1} e^{-b\lambda} \quad \text{, where } \lambda > 0 \text{ and } 0 \le \theta < 1.$$

Then, the posterior density is

$$f(\lambda,\theta\,|\,\underset{\sim}{x}) = \frac{L(\lambda,\theta\,|\,\underset{\sim}{x})h(\lambda,\theta)}{\int\int L(\lambda,\theta\,|\,\underset{\sim}{x})h(\lambda,\theta)\,d\lambda\, d\theta}\,.$$

Consider $f(\lambda,\theta\,|\,\underset{\sim}{x})h(\lambda,\theta) = \dfrac{1}{\Gamma(a)} b^a \lambda^{a-1} e^{-b\lambda} \lambda^n e^{-n\lambda-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n}\left[\dfrac{(\lambda+\theta x_i)^{x_i-1}}{x_i\,!}\right]$

$$= \frac{1}{\Gamma(a)} b^a \lambda^{n+a-1} e^{-\lambda(n+b)-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n}\left[\frac{(\lambda+\theta x_i)^{x_i-1}}{x_i\,!}\right].$$

Hence, the posterior density is

$$f(\lambda,\theta \mid \underset{\sim}{x}) = \frac{\lambda^{n+a-1} e^{-\lambda(n+b)-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} (\lambda+\theta x_i)^{x_i-1}}{\int_0^1 \int_0^\infty \lambda^{n+a-1} e^{-\lambda(n+b)-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} (\lambda+\theta x_i)^{x_i-1} d\lambda d\theta}$$

Or $$f(\lambda,\theta \mid \underset{\sim}{x}) = \frac{\lambda^{n+a-1} e^{-\lambda(n+b)-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} (\lambda+\theta x_i)^{x_i-1}}{k(x)},$$

where $k(x) = \int_0^1 \int_0^\infty \lambda^{n+a-1} e^{-\lambda(n+b)-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} (\lambda+\theta x_i)^{x_i-1} d\lambda d\theta$ is the marginal density of *X*.

Consider a generalized Poisson distribution with probability mass function

$$f(x;\lambda,\theta) = \frac{\lambda(\lambda+\theta x)^{x-1} e^{-\lambda-\theta x}}{x!}, \lambda > 0,\ 0 \le \theta < 1.$$

The prior distribution of λ is gamma (*a, b*) and θ is uniform (0, 1), the maximum Bayesian likelihood estimators of λ and θ are as follows:

$$\hat{\lambda}^{MBLE} = \frac{n\bar{X}(1-\hat{\theta}^{MBLE}) + a - 1}{n+b}$$

where $\hat{\theta}^{MBLE}$ is such that

$$R(\hat{\theta}^{MBLE}) = (n+b)\sum_{i=1}^{n}\left[\frac{x_i(x_i-1)}{n\bar{X} + a - 1 + [(n+b)x_i - n\bar{X}]\hat{\theta}^{MBLE}}\right] - n\bar{X} = 0. \quad (10)$$

The asymptotic variances and the covariances of the MBL estimators, that is, $\hat{\lambda}^{MBLE}, \hat{\theta}^{MBLE}$ and $Cov(\hat{\lambda}^{MBLE}, \hat{\theta}^{MBLE})$ are

$$V(\hat{\lambda}^{MBLE}) \approx \frac{n\lambda(\lambda+2)}{2(n+b)^2},$$

$$V(\hat{\theta}^{MBLE}) \approx \left(\frac{n+b}{n}\right)^2 \frac{(1-\theta)(\lambda+2\theta-\lambda\theta)}{2n\lambda},$$

$$Cov(\hat{\lambda}^{MBLE}, \hat{\theta}^{MBLE}) \approx -\left(\frac{n+b}{n}\right)^2 \frac{\lambda(1-\theta)}{2n},$$ (Bunthom Suraporn.,2006). (11)

2.2 Some Properties of Estimators

2.2.1 Consistent Estimators

Lamma 1 $\frac{\hat{\lambda}}{(1-\hat{\theta})}$ is a consistence estimator of $\frac{\lambda}{1-\theta}$, when $\hat{\lambda}$ and $\hat{\theta}$ are MBLE of λ and θ, respectively.

proof

Necessary and sufficient conditions for consistency are

(a) $\lim_{n \to \infty} E\left[\frac{\hat{\lambda}}{(1-\hat{\theta})}\right] = \frac{\lambda}{1-\theta}$, and

(b) $\lim_{n \to \infty} V\left[\frac{\hat{\lambda}}{(1-\hat{\theta})}\right] = 0.$

These two conditions can be proved as follows:

Recall that $\hat{\lambda} = \frac{n\bar{X}(1-\hat{\theta})}{n+b}$, so $\frac{\hat{\lambda}}{(1-\hat{\theta})} = \frac{n\bar{X}}{n+b}$.

Therefore, $E\left[\frac{\hat{\lambda}}{(1-\hat{\theta})}\right] = E\left[\frac{n\bar{X}}{n+b}\right] = \frac{n}{n+b}E\left[\bar{X}\right]$

$$= \frac{n}{n+b}E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{1}{n+b}\sum_{i=1}^{n} E\left[X_i\right]$$

$$= \frac{1}{n+b}\sum_{i=1}^{n}\left[\frac{\lambda}{1-\theta}\right] = \frac{1}{n+b}\left(\frac{n\lambda}{1-\theta}\right) = \frac{1}{1+\frac{b}{n}}\left(\frac{\lambda}{1-\theta}\right).$$

Hence $\lim_{n \to \infty} E\left[\frac{\hat{\lambda}}{(1-\hat{\theta})}\right] = \lim_{n \to \infty}\left[\frac{1}{1+\frac{b}{n}}\left(\frac{\lambda}{1-\theta}\right)\right] = \frac{\lambda}{1-\theta}$,

i.e., equation (a) is true

$$V\left[\frac{\hat{\lambda}}{(1-\hat{\theta})}\right] = V\left[\frac{n\bar{X}}{n+b}\right] = \frac{n^2}{(n+b)^2}V\left[\bar{X}\right]$$

$$= \frac{n^2}{(n+b)^2}V\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{n^2}{(n+b)^2}\frac{1}{n^2}V\left[\sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{(n+b)^2}\sum_{i=1}^{n} V\left[X_i\right] = \frac{1}{(n+b)^2}\sum_{i=1}^{n}\left[\frac{\lambda}{(1-\theta)^3}\right]$$

$$= \frac{n}{(n+b)^2}\left(\frac{\lambda}{(1-\theta)^3}\right) = \frac{1}{(n+2b+\frac{b^2}{n})}\left(\frac{\lambda}{(1-\theta)^3}\right)$$

$$\lim_{n\to\infty} V\left[\frac{\hat{\lambda}}{1-\hat{\theta}}\right] = \lim_{n\to\infty}\left[\frac{1}{(n+2b+\frac{b^2}{n})}\left(\frac{\lambda}{(1-\theta)^3}\right)\right] = 0.$$

Therefore, equation (b) is proved.

2.2.2 Bound of Estimators

Recall that $\hat{\lambda} = \frac{n\bar{X}(1-\hat{\theta})}{n+b}$, $\sum_{i=1}^{n}\left[\frac{x_i(x_i-1)}{\hat{\lambda}+\hat{\theta}x_i}\right] - n\bar{X} = 0$, $\lambda > 0$, $0 \le \theta < 1$ and $x_i \geq 0$ for all i = 1, 2, ..., *n*.

So $$\sum_{i=1}^{n}\left[\frac{x_i(x_i-1)}{\hat{\lambda}+\hat{\theta}x_i}\right] - n\bar{X} = 0 = \sum_{i=1}^{n}\left[\frac{x_i(x_i-1)}{\frac{n\bar{X}(1-\hat{\theta})}{n+b}+\hat{\theta}x_i}\right] - n\bar{X}$$

$$n\bar{X} = \sum_{i=1}^{n}\left[\frac{x_i(x_i-1)}{\frac{n\bar{X}(1-\hat{\theta})}{n+b}+\hat{\theta}x_i}\right] \le \sum_{i=1}^{n}\frac{\left[x_i(x_i-1)\right]}{\frac{n\bar{X}(1-\hat{\theta})}{n+b}} = \frac{n+b}{n\bar{X}(1-\hat{\theta})}\sum_{i=1}^{n}\left[x_i(x_i-1)\right].$$

Thus $(1-\hat{\theta}) \le \frac{n+b}{(n\bar{X})^2}\sum_{i=1}^{n}\left[x_i(x_i-1)\right]$ or

$$\hat{\theta} \ge 1 - \frac{n+b}{(n\bar{X})^2}\sum_{i=1}^{n}\left[x_i(x_i-1)\right] \tag{12}$$

$$= 1 - \frac{n+b}{(n\bar{X})^2}\left[(n-1)S^2 + n\bar{X}(\bar{X}-1)\right]$$

$$= 1 - \left[\frac{(n+b)(n-1)S^2}{(n\bar{X})^2} + \frac{(n+b)(\bar{X}-1)}{n\bar{X}}\right]$$

So $$\hat{\theta} \ge \left[\frac{(n+b)n\bar{X} - nb\bar{X}^2 - (n+b)(n-1)S^2}{(n\bar{X})^2}\right],$$

Where $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ and $S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{n-1}$.

Similarly,

$$n\bar{X} = \sum_{i=1}^{n}\left[\frac{x_i(x_i-1)}{\frac{n\bar{X}(1-\hat{\theta})}{n+b}+\hat{\theta}x_i}\right] \le \sum_{i=1}^{n}\left[\frac{x_i(x_i-1)}{\hat{\theta}x_i}\right].$$

So $$\hat{\theta} \le \frac{\sum_{i=1}^{n}(x_i-1)}{n\bar{X}} = \frac{n\bar{X}-n}{n\bar{X}} = 1-\frac{1}{\bar{X}}. \quad (13)$$

From equation (12) and (13) , it is concluded that

$$1-\frac{n+b}{(n\bar{X})^2}\sum_{i=1}^{n}\left[x_i(x_i-1)\right] \le \hat{\theta} \le 1-\frac{1}{\bar{X}},$$

so

$$\frac{1}{\bar{X}} \le 1-\hat{\theta} \le \frac{n+b}{(n\bar{X})^2}\sum_{i=1}^{n}\left[x_i(x_i-1)\right] = \frac{1-\hat{\theta}}{\hat{\lambda}(n\bar{X})}\sum_{i=1}^{n}\left[x_i(x_i-1)\right]$$

$$1 \le \frac{1}{\hat{\lambda}(n\bar{X})}\sum_{i=1}^{n}\left[x_i(x_i-1)\right] \text{ or}$$

$$\hat{\lambda} \le \frac{1}{n\bar{X}}\sum_{i=1}^{n}\left[x_i(x_i-1)\right] = \frac{\sum_{i=1}^{n}x_i^2 - n\bar{X}}{n\bar{X}} = \frac{\sum_{i=1}^{n}x_i^2}{n\bar{X}} - 1$$

$$= \frac{(n-1)S^2 + n\bar{X}^2}{n\bar{X}} - 1 = \left(\frac{n-1}{n}\right)\frac{S^2}{\bar{X}} + \bar{X} - 1.$$

Consider the case where *a* and *b* are unknown.

Both values of *a* and *b* in prior distribution gamma (*a*, *b*) may be specified as follow. Let X_1, X_2, ..., X_n, be a random sample. If the mean and variance of prior distribution is the same as the mean and variance of parent distribution, we may define $\frac{a}{b} = \bar{X}$, and $\frac{a}{b^2} = S^2$.

Solving these two equations for *a* and *b* yield

$$a = b\bar{X} = \frac{\bar{X}^2}{S^2}, \text{ and } b = \frac{\bar{X}}{S^2}. \quad (14)$$

We suggest that the values of *a* and *b* in maximum Bayesian likelihood estimators should be the expressions above, so the maximum Bayesian likelihood estimators from equations (10), yield

$$\hat{\lambda}^{MBLE} = \frac{n\bar{X}(1-\hat{\theta}^{MBLE}) + \frac{\bar{X}^2}{S^2} - 1}{n + \frac{\bar{X}}{S^2}},$$

$$= \frac{n\bar{X}S^2(1-\hat{\theta}^{MBLE}) + \bar{X}^2 - S^2}{nS^2 + \bar{X}},$$

and $$(n+\frac{\bar{X}}{S^2})\sum_{i=1}^{n} \frac{x_i(x_i-1)}{n\bar{X} + \frac{\bar{X}^2}{S^2} - 1 + [(n+\frac{\bar{X}}{S^2})x_i - n\bar{X}]\hat{\theta}^{MBLE}} = n\bar{X} \qquad (15)$$

Note that these values of *a* and *b* will be used later in the simulation study as well.

2.3 Relative Efficiency of Estimators

The asymptotic relative efficiency (ARE) provides a reasonable basis for the comparison of the estimators. Previously, asymptotic variances of MME (method of moments estimators), MLE (maximum likelihood estimators) and MBLE (maximum Bayesian likelihood estimators) are given in equations (4), (9) and (11) respectively.

Therefore AREs of these estimators can be obtained as follows:

$$\frac{V(\hat{\lambda}^{MBLE})}{V(\hat{\lambda}^{MME})} \approx \frac{\frac{n\lambda(\lambda+2)}{2(n+b)^2}}{\frac{\lambda}{2n}\left(\lambda+2+\frac{3\theta^2}{1-\theta}\right)} = \frac{n\lambda(\lambda+2)}{2(n+b)^2}\left(\frac{2n}{\lambda}\right)\frac{1-\theta}{(1-\theta)(\lambda+2)+3\theta^2}$$

$$\approx \frac{n^2(\lambda+2)}{(n+b)^2}\frac{1-\theta}{(1-\theta)(\lambda+2)+3\theta^2} \le \frac{n^2(\lambda+2)}{(n+b)^2}\frac{1-\theta}{(1-\theta)(\lambda+2)}$$

$$= \left(\frac{n}{n+b}\right)^2 \le 1.$$

That is the ratio of $\frac{V(\hat{\lambda}^{MBLE})}{V(\hat{\lambda}^{MME})} \le 1$ or $V(\hat{\lambda}^{MBLE}) \le V(\hat{\lambda}^{MME})$. It means that the maximum Bayesian Likelihood Estimator is more efficient than the method of moments estimator. In addition when *n* approaches to infinity, the relative efficiency of the above equation goes to one as expected.

Consider

$$\frac{V(\hat{\lambda}^{MBLE})}{V(\hat{\lambda}^{MLE})} \approx \frac{\frac{n\lambda(\lambda+2)}{2(n+b)^2}}{\frac{\lambda}{2n}(\lambda+2)} = \frac{n\lambda(\lambda+2)}{2(n+b)^2}\left(\frac{2n}{\lambda(\lambda+2)}\right) = \left(\frac{n}{n+b}\right)^2.$$

Again, the MBLE is more asymptotic efficient than the MLE and the relative efficiency of the MBLE to MLE converges to one as $n \to \infty$.

2.4 Examples

According to the general properties of the GPD models described in session 1. The value of the first parameter λ is the average rate of the natural chance process for the occurrence of the events and, accordingly, it is an indicator of the intensity of the natural Poisson process. The second parameter θ is an indicator of overdispersion, underdispersion, or of no dispersion relative to the Poisson distribution. Thus the parameter θ may be regarded as a measure of departure from Poisson model.

The GPD model can be applied in many areas. The following 3 data set are selected from Consul (1989). They are : (1) Bortkiewicz' s data of deaths due to horse-kicks in the Prussian Army cited in Fisher (1954) ; (2) Data of Zaire (1976) on numbers of automobile accident injuries ; (3) Thorndike (1926)'s data on number of lost articles found in the Telephone and Telegraph Building, New York City.

In each data set, the expected frequencies are calculated under two underlying distribution namely usual Poisson distribution (UPD) and generalized Poisson distribution (GPD). For UPD, the parameter λ is estimated by the sample mean, $\hat{\lambda} = \overline{X}.$ For GPD, the two parameters, λ and θ, are estimated by method of moments (MM), maximum likelihood method (ML) and maximum Bayesian likelihood method (MBL). See equations (3), (8) and (10).

For each method of estimation, χ^2-test is applied for considering goodness of fit between the observed frequencies and the expected frequencies.

2.5 Discussion

For the first data set shown in Table 1, the sample mean and variance are 0.605 and 0.592, respectively. In this example the difference between the mean and the variance is negligible and hence usual Poisson distribution seems to fit the data very well. Estimates from all methods are slightly different. The other two examples are different situations where overdispersion exists. Table 2 exhibits on data of automobile accidents where the sample mean is 0.625 and variance is 1.002, while Table 3 exhibits the data of lost articles found with mean 1.033 and variance is 1.226. It is found that, in both cases, usual Poisson distribution does not fit to the data (χ^2-test statistics are respectively 10.694 and 9.346). The coefficient of variation (CV) of these two cases are 1.602 and 1.072 respectively. That is, overdispersion is concerned. Therefore, GPD is preferred to UPD.

Table 1. Number of death due to horse-kicks in Prussian Army with expected frequencies obtained from different distributions assumed and method of parameter estimation.

Number of Deaths	Observed Frequency	Expected Frequencies			
		UPD assumed	GPD assumed		
			MME	MLE	MBLE
0	109	109.215	108.492	108.309	109.023
1	65	66.075	67.090	67.350	66.894
2	23	19.988	19.999	20.000	19.766
3	2	4.031	3.828	3.775	3.746
4 or more	1	0.691	0.529	0.509	0.512
Total	200	200.000	200.000	200.000	200.000
estimates		$\hat{\lambda}=\bar{X}=0.605$ $S^2=0.592$	$\hat{\lambda}=0.612$ $\hat{\theta}=-0.011$	$\hat{\lambda}=0.613$ $\hat{\theta}=-0.014$	$\hat{\lambda}=0.607$ $\hat{\theta}=-0.011$
χ^2	Statistics	**1.633**	**1.674**	**1.704**	**1.719**

Source: Consul, 1989.

Table 2. Number of automobile accident injury with expected frequencies obtained from different distribution assumed and methods of parameter estimation.

Number of Accidents	Observed Frequency	Expected Frequencies			
		UPD assumed	GPD assumed		
			MME	MLE	MBLE
0	36	29.975	34.185	35.155	35.859
1	10	18.734	13.672	12.683	12.297
2	6	5.854	5.064	4.794	4.589
3	3	1.219	1.892	1.922	1.843
4 or more	1	0.218	1.187	1.446	1.412
Total	56	56.000	56.000	56.000	56.000
Estimates		$\hat{\lambda}=\bar{X}=0.625$ $S^2=1.002$	$\hat{\lambda}=0.493$ $\hat{\theta}=0.210$	$\hat{\lambda}=0.466$ $\hat{\theta}=0.255$	$\hat{\lambda}=0.446$ $\hat{\theta}=0.262$
χ^2	Statistics	**10.694***	**1.934**	**1.633**	**1.710**

Source: Consul, 1989. **Remark**: * represent significance at 5% level.

Table 3. Number of lost articles found in a day in the Telephone and Telegraph Building, New York City with expected frequencies obtained from different distributions assumed and methods of parameter estimation.

Number of Lost articles	Observed Frequency	Expected Frequencies			
		UPD assumed	GPD assumed		
			MME	MLE	MBLE
0	169	150.547	163.887	164.861	165.399
1	134	155.529	143.137	142.269	142.088
2	74	80.338	73.342	72.842	72.598
3	32	27.666	28.891	28.927	28.826
4	11	7.145	9.704	9.865	9.844
5 or more	3	1.775	4.039	4.236	4.245
Total	423	423.000	423.000	423.000	423.000
estimates		$\hat{\lambda}=\bar{X}=1.033$ $S^2=1.226$	$\hat{\lambda}=0.948$ $\hat{\theta}=0.082$	$\hat{\lambda}=0.942$ $\hat{\theta}=0.088$	$\hat{\lambda}=0.939$ $\hat{\theta}=0.091$
χ^2	Statistics	**9.346***	**1.524**	**1.421**	**1.416**

3. References

Cramer, H. (1946). Mathematical methods of statistics. Princeton, N.J. Princeton University Press.

Consul, P.C. and Jain, G.C. (1973a). A generalization of the Poisson distribution. Technometrics. 15(4): 791–799.

Consul, P.C. and Shoukri, M.M. (1984). Maximum likelihood estimation for the generalized Poisson distribution. Comm. Statist. A-Theory Methods 13(2): 1533–1547.

Consul, P.C. (1989). Generalized Poisson Distributions, properties and applied. New York : Marcel Dekker.

Kumar, A. and Consul, P.C. (1979). Negative moments of a modified power series distribution and bias of the maximum likelihood estimator. Comm. Statist. A-Theory Methods 8(2): 151–166.

Kumar, A. and Consul, P.C. (1980). Minimum variance unbiased estimation for modified power series distribution, Comm. Statist. A-Theory Methods. 9(12): 1261–1275.

Shoukri, M.M. (1980). Estimation problems for some generalized discrete probability distributions. Doctoral dissertation, University of Calgary.

Suraporn, B. (2006). The generalized Poisson distribution :some properties and estimation. Doctoral dissertation, The National Institute Development Administration.

❑❑❑❑❑