ARPN Journal of Engineering and Applied Sciences

ON GENERALIZED VARIANCE OF NORMAL-POISSON MODEL AND POISSON VARIANCE ESTIMATION UNDER GAUSSIANITY

Khoirin Nisa[1], Célestin C. Kokonendji[2], Asep Saefuddin[3], AjiHamim Wigena[3] and I WayanMangku[4]
[1]Department of Mathematics, University of Lampung, Bandar Lampung, Indonesia
[2]Laboratoire de Mathématiques de Besançon, Université Bourgogne Franche-Comté, France
[3]Department of Statistics, Bogor Agricultural University, Bogor, Indonesia
[4]Department of Mathematics, Bogor Agricultural University, Bogor, Indonesia
E-Mail: khoirin.nisa@fmipa.unila.ac.id

ABSTRACT

As an alternative to full Gaussianity, multivariate normal-Poisson model has been recently introduced. The model is composed by a univariate Poisson variable, and the remaining random variables given the Poisson one are real independent Gaussian variables with the same variance equal to the Poisson component. Under the statistical aspect of the generalized variance of normal-Poisson model, the parameter of the unobserved Poisson variable is estimated through a standardized generalized variance of the observations from the normal components. The proposed estimation is successfully evaluated through simulation study.

Keywords: covariance matrix, determinant, exponential family, generalized variance, infinitely divisible measure.

INTRODUCTION

Normal-Poisson model is a special case of normal stable Tweedie (NST) models which were introduced by Boubacar Maïnassara and Kokonendji [1] as the extension of normal gamma [2] and normal inverse Gaussian [3] models. The NST family is composed by distributions of random vector **X**= $(X_1, \ldots, X_k)^T$ where X_j is a univariate (non-negative) stable Tweedie variable and $(X_1, \ldots, X_{j-1}, X_{j+1},\ldots,X_k)^T =: \mathbf{X}_j^c$ given X_j are k-1 real independent Gaussian variables with variance X_i, for any fixed $j\in \{1, 2, \ldots, k\}$. Several particular cases have already appeared in different contexts; one can refer to [1] and references therein.

Normal-Poisson is the only NST model which has a discrete component and it is correlated to the continuous normal parts. Similar to all NST models, this model was introduced in [1] for a particular case of j that is j=1. For a normal-Poisson random vector X as described above, X_j is a univariate Poisson variable. In literatures, there is a model called "Poisson Gaussian" [4] [5] which is also composed by Poisson and normal distributions. However, normal-Poisson and Poisson Gaussian are two completely different models. Indeed, for any value of $j\in \{1, 2, \ldots, k\}$, a normal-Poisson$_j$ model has only one Poisson component and k-1 Gaussian components, while a Poisson-Gaussian$_j$ model has j Poisson components and k-j Gaussian components which are all independent. Normal-Poisson is also different from the purely discrete Poisson-normal model of Steyn [6] which can be defined as a multiple mixture of k independent Poisson distributions with parameters $m_1, m_2, \ldots, m_k$ and those parameters have a multivariate normal distribution. Hence, the multivariate Poisson-normal distribution is a multivariate version of the Hermite distribution [7].

Generalized variance (i.e. the determinant of covariance matrix expressed in term of mean vector) has important roles in statistical analysis of multivariate data. It was introduced by Wilks [8] as a scalar measure of multivariate dispersion and used for overall multivariate scatter. The uses of generalized variance have been discussed by several authors. In sampling theory, it can be used as a loss function on multiparametric sampling allocation [9]. In the theory of statistical hypothesis testing, generalized variance is used as a criterion for an unbiased critical region to have the maximum Gaussian curvature [10]. In the descriptive statistics, Goodman [11] proposed a classification of some groups according to their generalized variances. In the last two decades the generalized variance has been extended for non-normal distributions in particular for natural exponential families (NEFs) [12] [13].

Three generalize variance estimators of normal-Poisson models have been introduced (see [14]). Also, the characterization by variance function and by generalized variance of normal-Poisson have been successfully proven (see [15]). In this paper, a new statistical aspect of normal Poisson model is presented, i.e. the Poisson variance estimation under only observations of normal components leading to an extension of generalized variance term i.e. the "standardized generalized variance".

NORMAL POISSON MODELS

The family of multivariate normal-Poisson$_j$ models for all$j\in \{1, 2, \ldots, k\}$and fixed positive integer k>1 is defined as follows:

Definition 1. For$\mathbf{X} = (X_1, \ldots,X_k)^T$ a k-dimensional normal-Poisson random vector, it must hold that

1. X_j is a univariate Poisson random variable, and
2. $\mathbf{X}_j^c := (X_1, \ldots, X_{j-1}, X_{j+1},\ldots,X_k)^T$given $X_j$follows the$(k-1)$-variate normal$N_{k-1}(\mathbf{0}, X_j\mathbf{I}_{k-1})$ distribution, where$\mathbf{I}_{k-1}=\mathrm{diag}_{k-1}(1, \ldots, 1)$ denotes the $(k-1)\times(k-1)$unit matrix.

In order to satisfy the second condition we need X_j>0. But in practice it is possible to have X_j=0 in the Poisson component. In this case, the corresponding normal components are degenerated as the Dirac mass $\delta_0$which makes their values become 0s. We have shown that zero values in X_j do not affect the estimation of the generalized variance of normal-Poisson [16].

From Definition 1, for a fixed power of convolution $t>0$ and given $j\in\{1, 2, \ldots, k\}$, denote $F_{t;j}=F(\nu_{t;j})$ the multivariate NEF of normal-Poisson with $\nu_t:=\nu^{t*}$, the NEF of a k-dimensional normal-Poisson random vector $\mathbf{X}$ is generated by

$$\nu_{t;j}(d\mathbf{x}) = \frac{t^{x_j}(x_j!)^{-1}}{(2\pi x_j)^{(k-1)/2}} \exp\left(-t - \frac{1}{2x_j}\sum_{\ell\neq j} x_\ell^2\right) \mathbb{1}_{x_j\in\mathbb{N}\setminus\{0\}}\delta_{x_j}(dx_j)\prod_{\ell\neq j} dx_\ell, \quad (1)$$

where 1_A is the indicator function of the set A. Since $t>0$ then $\nu_{t;j}$ is known to be an infinitely divisible measure; see, e.g., Sato [17].

The cumulant function of normal-Poisson is obtained from the logarithm of the Laplace transform of $\nu_{t;j}$, i.e. $K_{\nu_{t;j}}(\theta)=\log\int_{R^k}\exp(\theta^T x)\,\nu_{t;j}(d\mathbf{x})$ and the probability distribution of normal-Poisson$_j$ which is a member of NEF is given by

$$P(\theta;\nu_{t;j})(dx) = \exp\left\{\theta^T x - K_{\nu_{t;j}}(\theta)\right\}\nu_{t;j}(dx)$$

The mean vector and the covariance matrix of $F_{t;j}$ can be calculated using the first and the second derivatives of the cumulant function, i.e.:

$$\mu = K'_{\nu_{t;j}}(\theta)$$

and

$$V_{F_{t;j}}(\mu) = K''_{\nu_{t;j}}(\theta(\mu)).$$

For practical calculation we need to use the following mean parameterization:

$$P(\boldsymbol{\mu};F_{t;j}) := P(\boldsymbol{\theta}(\boldsymbol{\mu});\nu_{t;j}),$$

where $\theta(\mu)$ is the solution in $\boldsymbol{\theta}$ of the equation $\mu = K'_{\nu_{t;j}}(\theta)$. Then for a fixed $j\in\{1, 2, \ldots, k\}$, the variance function (i.e the variance-covariance matrix in term of mean parameterization) is given by

$$V_{F_{t;j}}(\mu) = \frac{1}{\mu_j}\mu\mu^T + \mathrm{diag}(\mu_j,\ldots,\mu_j,0_j,\mu_j,\ldots,\mu_j) \quad (2)$$

on its support

$$\mathbf{M}_{F_{t;j}} = \{\boldsymbol{\mu}\in R^k; \mu_j>0 \text{ and } \mu_\ell\in R \text{ for } \ell\neq j\}. \quad (3)$$

For $j = 1$, the covariance matrix of $\mathbf{X}$ can be expressed as follows:

$$V_{F_{t;j}}(\boldsymbol{\mu}) = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_j & \cdots & \mu_k \\ \mu_2 & \mu_1^{-1}\mu_2^2+\mu_1 & \cdots & \mu_1^{-1}\mu_2\mu_j & \cdots & \mu_1^{-1}\mu_k\mu_2 \\ \vdots & \vdots & & \vdots & & \vdots \\ \mu_j & \mu_1^{-1}\mu_j\mu_2 & \cdots & \mu_1^{-1}\mu_j^2+\mu_1 & \cdots & \mu_1^{-1}\mu_j\mu_k \\ \vdots & \vdots & & \vdots & & \vdots \\ \mu_k & \mu_1^{-1}\mu_k\mu_2 & \cdots & \mu_1^{-1}\mu_k\mu_j & \cdots & \mu_1^{-1}\mu_k^2+\mu_1 \end{bmatrix}$$

Indeed, for the covariance matrix above one can use the Schur complement [18] of a matrix block to obtain the following representation of determinant

$$\det\begin{pmatrix}\gamma & \mathbf{a}^\top \\ \mathbf{a} & \mathbf{A}\end{pmatrix}^{-1} = \gamma \quad \det(\mathbf{A}-\gamma^{-1}\mathbf{a}\mathbf{a}^\top), \quad (4)$$

with the non-null scalar $\gamma = \mu_1$, the vector $\mathbf{a}^T=(\mu_2, \ldots,\mu_k)$ and the $(k-1)\times(k-1)$ matrix $\mathbf{A}=\gamma^{-1}\mathbf{a}\mathbf{a}^T+\mu_1\mathbf{I}_{k-1}$, where $\mathbf{I}_j= =\mathrm{diag}(1, \ldots, 1)$ is the $j\times j$ unit matrix. Consequently, the determinant of the covariance matrix for $j = 1$ is

$$\det\mathbf{V}_{F_{t;1}}(\boldsymbol{\mu}) = \mu_1^k \text{with } \boldsymbol{\mu}\in\mathbf{M}_{F_{t;1}}$$

Then, it is trivial to show that for $j\in\{1, 2, \ldots, k\}$ the generalized variance of normal-Poisson$_j$ model is given by

$$\det\mathbf{V}_{F_{t;1}}(\boldsymbol{\mu}) = \mu_j^k \text{with } \boldsymbol{\mu}\in\mathbf{M}_{F_{t;j}} \quad (5)$$

Equation (5) expresses the generalized variance of normal-Poisson model depends only on the mean of the Poisson component and the dimension space $k>1$.

CHARACTERIZATIONS AND GENERALIZED VARIANCE ESTIMATIONS

Among NST models, normal-Poisson and normal-gamma are the only models which are already characterized by generalized variance (see [19] for characterization of normal-gamma by generalized variance). In this section we present the characterizations of normal-Poisson by variance function and by generalized variance, then we present three estimations of generalized variance by maximum likelihood (ML), uniformly minimum variance unbiased (UMVU) and Bayesian methods.

Characterization

The characterizations of normal-Poisson models are stated in the following theorems without proof.

Theorem 1

Let $k\in\{2,3, \ldots\}$ and $t>0$. If an NEF $F_{t;j}$ satisfies (2) for a given $j\in\{1,2, \ldots,k\}$, then up to affinity, $F_{t;j}$ is a normal-Poisson$_j$ model.

Theorem 2

Let $F_{t;j}=F(\nu_{t;j})$ be an infinitely divisible NEF on R^k (k>1) such that

1) $\mathbf{\Theta}(\nu_{t;j}) = R^k$ and
2) $\det \mathbf{K}''_{\nu_{t;j}}(\boldsymbol{\theta}) = t \exp(k \times \boldsymbol{\theta}^{\mathrm{T}}\widetilde{\boldsymbol{\theta}}_j^c)$

for $\boldsymbol{\theta}=(\theta_1, \ldots \theta_k)^{\mathrm{T}}$ and $\widetilde{\boldsymbol{\theta}}_j^c=(\theta_1,\ldots, \theta_{j-1}, 1, \theta_{j+1}, \ldots, \theta_k)^{\mathrm{T}}$. Then $F_{t,j}$ is of normal-Poisson type.

All technical details of proofs can be seen in [15]. In fact, the proof of Theorem 1 is established by analytical calculations and using the well-known properties of NEFs described in Proposition 3 below.

Proposition 3

Let ν and $\tilde{\nu}$ be two σ-finite positive measures on R^k such that $F=F(\nu)$\$,$\tilde{F} = F(\tilde{\nu})$ and $\boldsymbol{\mu}\in M_F$.

(i) If there exists $(\mathrm{d},c)\in R^k \times R$ such that $\tilde{\nu}(dx) = \exp(\mathrm{d}^{\mathrm{T}}\mathrm{x}+c)\nu(dx)$, then $F = \tilde{F}$: $\Theta_{\tilde{\nu}} = \Theta_\nu - \mathrm{d}$ and $\mathrm{K}_{\tilde{\nu}}(\theta) = \mathrm{K}_\nu(\theta + \mathrm{d}) + c$; for $\tilde{\mu} = \mu\epsilon M_F$,
$\mathrm{V}_{\tilde{F}}(\tilde{\mu}) = \mathrm{V}_F(\mu)$ and $\det \mathrm{V}_{\tilde{F}}(\tilde{\mu}) = \det \mathrm{V}_F(\mu)$.

(ii) If $\tilde{\nu} = \varphi * \nu$ with $\varphi(\mathrm{x}) = \mathrm{Ax} + \mathrm{b}$, then: $\Theta(\tilde{\nu}) = \mathrm{A}^{\mathrm{T}}\Theta(\nu)$ and $\mathrm{K}_{\tilde{\nu}}(\theta) = \mathrm{K}_\nu(\mathrm{A}^{\mathrm{T}}\theta) + \mathrm{b}^{\mathrm{T}}\theta$; for $\tilde{\mu} = \mathrm{A}\mu + \mathrm{b} \,\epsilon\, \varphi M_F$,
$\mathrm{V}_{\tilde{F}}(\tilde{\mu}) = \mathrm{A}\mathrm{V}_F(\varphi^{-1}(\tilde{\mu}))\mathrm{A}^{\mathrm{T}}$
and
$\det \mathbf{V}_{\tilde{F}}(\tilde{\boldsymbol{\mu}}) = (\det \mathbf{A})^2 \det \mathbf{V}_F(\boldsymbol{\mu})$.

(iii) If $\tilde{\nu} = \nu^*$ is the t-th convolution power of ν for t>0, then, for $\tilde{\boldsymbol{\mu}} = t\boldsymbol{\mu}\epsilon\, t\mathbf{M}_F$,

$\mathbf{V}_{\tilde{F}}(\tilde{\boldsymbol{\mu}}) = t\mathbf{V}_F(t^{-1}\tilde{\boldsymbol{\mu}})$ and $\det\mathbf{V}_{\tilde{F}}(\tilde{\boldsymbol{\mu}}) = t^k \det \mathbf{V}_F(\boldsymbol{\mu})$

The proof of Theorem 2 is obtained by using the infinite divisibility property of normal-Poisson, also applying two properties of determinant and affine polynomial. The infinite divisibility property used in theproof is provided in Proposition 4below.

Proposition 4

If ν is an infinitely divisible measure on R^k, then there exist a symmetric non-negative definite $d\times d$ matrix $\mathbf{\Sigma}$ with rank $r\le k$ and a positive measure ξ on R^k such that
$\mathbf{K}''_\nu(\mathbf{\theta}) = \Sigma + \int_{R^k} \mathbf{x}\mathbf{x}^{\mathrm{T}}\exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{x})\, \xi(d\mathbf{x})$.
See, e.g.[20, page 342].

The above expression of $\mathbf{K}''_\nu(\boldsymbol{\theta})$ is an equivalent of the Lévy-Khinchine formula [17]; thus, $\mathbf{\Sigma}$ comes from a Brownian part and the rest $\mathbf{L}''_\xi(\boldsymbol{\theta}) = \int_{R^k} \mathbf{x}\mathbf{x}^{\mathrm{T}}\exp(\boldsymbol{\theta}^{\mathrm{T}}\mathbf{x})\, \xi(d\mathbf{x})$ corresponds to jumps part of the associated Lévy process through the Lévy measure ξ.

Generalized variance estimators

Let $\mathbf{X}_1, \ldots,\mathbf{X}_n$ be random vectors i.i.d. with distribution $P(\boldsymbol{\mu};F_{t;j})$ in a normal-Poisson$_j$ model $F_{t;j}=F(\nu_{t;j})$ for fixed $j\in\{1,2, \ldots,k\}$. Denoting $\overline{\mathbf{X}} = \frac{(X_1+\cdots+X_n)}{n} = (\overline{\mathrm{X}}_1, \ldots, \overline{\mathrm{X}}_k)^{\mathrm{T}}$ the sample mean.

a) Maximum likelihood estimator

The ML generalized variance estimator of normal Poisson model $\det \mathbf{V}_{F_{t;j}}(\boldsymbol{\mu}) = \mu_j^k$ is given by

$$T_{n,t;j} = \det \mathbf{V}_{F_{t;j}}(\overline{\mathbf{X}}) = \bar{X}_j^k. \quad (6)$$

The ML estimator (6) is directly obtained from (5) by substituting μ_j with its ML estimator $\bar{X}_j$. For all $p\ge1$, $T_{n,t;j}$ is a biased estimator of $\det \mathbf{V}_{F_{t;j}}(\boldsymbol{\mu})$ with a given quadratic risk with tedious calculation of explicit expression or infinite.

b) Uniformly minimum variance unbiased estimator

The UMVU generalized variance estimator of normal Poisson models $\det \mathbf{V}_{F_{t;j}}(\boldsymbol{\mu}) = \mu_j^k$ is given by

$$\mathrm{U}_{n,t;j} = \mathrm{n}^{-k+1}\overline{\mathrm{X}}_j(\mathrm{n}\overline{\mathrm{X}}_j - 1) \ldots (\mathrm{n}\overline{\mathrm{X}}_j - \mathrm{k} + 1), \text{ if } n\bar{X}_j \ge k \quad (7)$$

The UMVU estimator of $\det \mathbf{V}_{F_{t;j}}(\boldsymbol{\mu})$ is deduced by using intrinsic moment formula of univarite Poisson distribution as follows:

$$E\left[Y_j(Y_j - 1) \ldots (Y_j - k + 1)\right] = \mu_j^k.$$

Indeed, letting $Y_j = n\bar{X}_j$ gives the result that (7) is the UMVU estimator of (5). Because, by the completness of NEF, the unbiased estimator is unique.

c) Bayesian estimator

Under assumption of prior gamma distribution of μ_j with parameter α>0 and β>0, the Bayesian estimator of $\det \mathbf{V}_{F_{t;j}}(\boldsymbol{\mu}) = \mu_j^k$ is given by

$$B_{n,t;j,\alpha,\beta} = \left(\frac{\alpha+n\bar{X}_j}{\beta+n}\right)^k. \quad (8)$$

To show this, let $X_{j1},\ldots,X_{jn}$ given μ_j be Poisson distribution with mean μ_j, then the probability mass function is given by

$$p(\boldsymbol{x_{ji}}\,|\mu_j) = \frac{\mu_j^{x_{ji}}}{x_{ji}!}\exp(-\mu_j) \qquad \forall\, x_{ji}\epsilon\, N$$

Assuming that μ_j follows gamma(α,β), then the prior probability distribution function of μ_j is written as

$$f(\mu_j, \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\mu_j^{\alpha-1}\exp(-\beta\mu_j), \qquad \forall\mu_j > 0$$

with $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1}e^{-x}dx$ .Using the classical Bayes theorem, the posterior distribution of μ_j given an observation x_{ji} can be expressed as

$$f(\mu_j|x_{ji};\alpha,\beta) = \frac{p(x_{ji}|\mu_j)f(\mu_j,\alpha,\beta)}{\int_{\mu_j>0} p(x_{ji}|\mu_j)f(\mu_j,\alpha,\beta)d\mu_j}$$
$$= \frac{(\beta+1)^{\alpha+x_{ji}}}{\Gamma(\alpha+x_{ji})}\mu_j^{\alpha+x_{ji}-1}\exp\{-(\beta+1)\mu_j\}$$

which is the gamma density with parameters $\alpha'=\alpha+x_{ji}$, $\beta'=\beta+1$. Then with random sample $X_{j1}, \ldots, X_{jn}$ the posterior will be gamma$(\alpha+n\bar{X}_j, \beta+n)$. Since Bayesian estimator of μ_j is given by the expected value of the posterior distribution i.e. $\frac{\alpha+n\bar{X}_j}{\beta+n}$, then this will lead to (8).

MAIN RESULT

Poisson variance estimation under gaussianity

For a given random vector $\mathbf{X}= (X_1, \ldots, X_k)^{\mathrm{T}}$ on R^k of normal-Poisson$_j$, we now assume that only k-1 normal terms $\mathbf{X}_j^c$ of $\mathbf{X}$ are observed: $\mathbf{X}_{j1}^c,\ldots, \mathbf{X}_{jn}^c$ and, therefore, X_j is an unobserved Poisson random effect. Note that j is fixed in {1, 2,...,k}.

Assuming t=1 and following [1] with $\mathbf{X}$ having mean vector $\boldsymbol{\mu}=(\mu_1, \ldots, \mu_k)^{\mathrm{T}}\in\mathbf{M}_{F_{1;j}}$ and covariance matrix $\mathbf{V}=\mathbf{V}(\boldsymbol{\mu})$, then $\mathbf{X}_j^c$ follows a (k-1)-variate normal distribution, denoted by

$$\mathbf{X}_j^c \sim \mathcal{N}_{k-1}(\boldsymbol{\mu}_j^c, X_j\mathbf{V}_j^c), \quad (9)$$

with $\boldsymbol{\mu}_j^c = (\mu_1,\ldots,\mu_{j-1},\mu_{j+1},\ldots,\mu_k)^{\mathrm{T}}$. The ($k$-1)×($k$-1)-matrix $\mathbf{V}_j^c$ (which does not depend on $\boldsymbol{\mu}_j^c$) is symmetric and positive definite such that $\det\mathbf{V}_j^c = 1$ or $\mathbf{V}_j^c = \mathbf{I}_{k-1}$. Thus, without loss of generality, X_j in (9) can be a univariate Poisson variable with parameter $\mu_j>0$ which is known to be at the same time the mean and the variance. It follows that the unit generalized variance of $\mathbf{X} = (X_j, \mathbf{X}_j^{c\mathrm{T}})^{\mathrm{T}}$ is easily deduced as μ_j^k. Hence, the Poisson parameter μ_j of X_j can be estimated through generalized variance estimators of normal observations in the sense of ``standardized generalized variance" [21]:

$$\widehat{\mu}_j = \left(\det\left[\frac{1}{n-1}\sum_{i=1}^n \mathbf{X}_{ji}^c\mathbf{X}_{ji}^{c\,\mathrm{T}} - \overline{\mathbf{X}}_j^c\overline{\mathbf{X}}_j^{c\,\mathrm{T}}\right]\right)^{1/(k-1)} \quad \text{for } \det\mathbf{V}_j^c = 1$$

or

$$\widehat{\mu}_j = \left(\prod_{\ell\neq j}\left[\frac{1}{n-1}\sum_{i=1}^n X_{\ell i}^2 - \overline{X}_\ell^2\right]\right)^{1/(k-1)} \quad \text{for } \mathbf{V}_j^c = \mathbf{I}_{k-1}, \quad (10)$$

with $\overline{\mathbf{X}}_j^c = (\mathbf{X}_{j1}^c + \cdots + \mathbf{X}_{jn}^c)/n$ and $\bar{X}_\ell = (X_{\ell 1} + \cdots + X_{\ell n})/n$. This statistical aspect of normal-Poisson$_j$ models in (9) points out the flexibility of these models compared with the classical multivariate normal model $\mathcal{N}_{k-1}(\boldsymbol{\mu}_j^c, \boldsymbol{\Sigma})$, where the generalized variance $\det\boldsymbol{\Sigma}$ is replaced to the random effect $X_j\mathbf{V}_j^c$.

In fact, for $\mathbf{V}_j^c = \mathbf{I}_{k-1}$ in (9) with estimation $\hat{\mu}_j$ of (10) which corresponds to Part 2 of Definition 1, one has a kind of conditional homoscedasticity under the assumption of normality. However, we here have to handle the presence of zeros in the sample of X_j when the Poisson parameter μ_j is close to zero.

More precisely and without loss of generality, within the framework of one-way analysis of variance and keeping the previous notations, since there are at least two normal components to be tested, so the minimum value of k is 3 (or $k\geq 3$) for representing the number of levels k-1.

Simulation study

We present empirical analyses through simulation study to evaluate the consistency of $\hat{\mu}_j$. In order to apply this point of view, one can refer to [21] for a short numerical illustration; or in the context of multivariate random effect model, it can be used as the distribution of the random effects when they are assumed to have conditional homoscedasticity.

Using the standardized generalized variance estimation in (10) we assume that the Poisson component is unobservable and we want to estimate $\hat{\mu}_j$ based on observations of normal components. In this simulation, we fixed j=1 and we set some sample sizes n = 30, 50, 100, 300, 500, 1000. We consider k=3, 4, 6, 8 to see the effects of k on the standardized generalized variance estimations. Moreover, to see the effect of zero values proportion within X_j, we also consider small mean (variance) values on the Poisson component i.e. μ_j=0.5, 1, 5, because $P(X_j=0)=\exp(-\mu_j)$. We generated 1000 samples for each case. From the resulted $\hat{\mu}_j$ values of the generated samples we obtained the expected values and variance of $\hat{\mu}_j$ i.e. $E(\hat{\mu}_j)$ and $Var(\hat{\mu}_j)$ respectively. Then we calculated their MSE using the following formula

$$\mathrm{MSE}(\hat{\mu}_j) = [E(\hat{\mu}_j)-\mu_j]^2 + Var(\hat{\mu}_j),$$

where

$$E(\hat{\mu}_j) = \frac{1}{1000}\sum_{i=1}^{1000}\left[\hat{\mu}_j^{(i)}\right]$$

and

$$Var(\hat{\mu}_j) = \frac{1}{999}\sum_{i=1}^{1000}\left[\hat{\mu}_j^{(i)} - E(\hat{\mu}_j)\right]^2$$

We report the expected values and MSE of $\hat{\mu}_j$ in Table-1, Table-3.

VOL. 12, NO. 12, JUNE 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-1. The expected values and MSE of $\hat{\mu}_j$ with 1000 replications for $n\in\{30,50,100,300,500,1000\}$, $k\in\{3,4,6,8\}$, and μ_j=0.5.

k	*n*	$E(\hat{\mu}_j)$	MSE($\hat{\mu}_j$)
3	**30**	**0.473270**	**0.039251**
	50	0.487402	0.023864
	100	0.491117	0.010882
	300	0.495814	0.004058
	500	0.496612	0.002540
	1000	0.499035	0.001158
4	30	0.465915	0.031980
	50	0.488503	0.019574
	100	0.491804	0.009975
	300	0.494617	0.003457
	500	0.496200	0.002019
	1000	0.498271	0.000968
6	30	0.452953	0.026781
	50	0.478994	0.015763
	100	0.483284	0.007801
	300	0.495324	0.002713
	500	0.496771	0.001562
	1000	0.497542	0.000771
8	30	0.454636	0.023539
	50	0.468367	0.014280
	100	0.482915	0.007374
	300	0.495749	0.002395
	500	0.499078	0.001542
	1000	0.499199	0.000726

Table-2. The expected values and MSE of $\hat{\mu}_j$ with 1000 replications for $n\in\{30,50,100,300,500,1000\}$, $k\in\{3,4,6,8\}$, and μ_j=1.

k	*n*	$E(\hat{\mu}_j)$	MSE($\hat{\mu}_j$)
3	**30**	**0.962617**	**0.095854**
	50	0.983720	0.055901
	100	0.993564	0.029386
	300	0.994837	0.010214
	500	0.997781	0.005969
	1000	0.998467	0.003125
4	30	0.955849	0.078891
	50	0.973454	0.049405
	100	0.981452	0.023848
	300	0.992874	0.007467
	500	0.996215	0.004848
	1000	1.001149	0.002456
6	30	0.944165	0.058871
	50	0.972215	0.033577
	100	0.985437	0.017781
	300	0.992045	0.006229
	500	0.995822	0.003725
	1000	0.998113	0.001752
8	30	0.944031	0.052258
	50	0.973103	0.032476
	100	0.981169	0.015210
	300	0.992240	0.005135
	500	0.998451	0.002981
	1000	0.999042	0.001400

From the results in the tables we can see that when the sample size (n) increases, the expected values of $\hat{\mu}_j$ converge to the target value (μ_j) for all μ_j values we consider here. Also, the MSE of $\hat{\mu}_j$ decrease when sample size increase for all dimension k, this means that $\hat{\mu}_j$ is consistent. The simulation results with moderate sample sizes produce very good performances of $\hat{\mu}_j$. Note that the presence of zeros in the samples of the Poisson component does not affect the estimation of μ_j.

For a clear description of the performance of $\hat{\mu}_j$, we provide the bargraphs of MSE of $\hat{\mu}_j$ in Figure-1, Figure-3. The figures show that MSE value decrease when the sample size increase. From the result we conclude that $\hat{\mu}_j$ is a consistent estimator of μ_j. Notice that $\hat{\mu}_j$ produce smaller MSE for larger dimension.

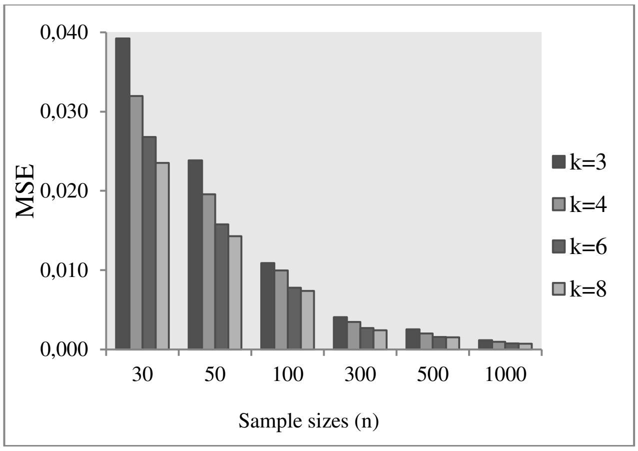

Figure-1. Bargraph of MSE($\hat{\mu}_j$) for μ_j= 0.5.

Table-3. The expected values and MSE of $\hat{\mu}_j$ with 1000 replications for $n \in$ {30,50,100,300,500,1000}, $k \in$ {3,4,6,8}, and μ_j=5.

k	***n***	$E(\hat{\mu}_j)$	MSE($\hat{\mu}_j$)
3	**30**	**4.886415**	**1.120641**
	50	4.942883	0.690184
	100	4.984949	0.356851
	300	4.987437	0.118193
	500	4.992459	0.064591
	1000	5.006692	0.035814
4	30	4.856583	0.928853
	50	4.921017	0.511915
	100	4.950201	0.269422
	300	4.983517	0.086223
	500	4.988398	0.050144
	1000	4.988551	0.025137
6	30	4.852608	0.589918
	50	4.926390	0.354075
	100	4.942147	0.175198
	300	4.974067	0.056670
	500	4.995231	0.033951
	1000	4.996774	0.016958
8	30	4.838751	0.457897
	50	4.910668	0.281625
	100	4.949142	0.135143
	300	4.985705	0.046987
	500	4.990750	0.027643
	1000	4.998134	0.013399

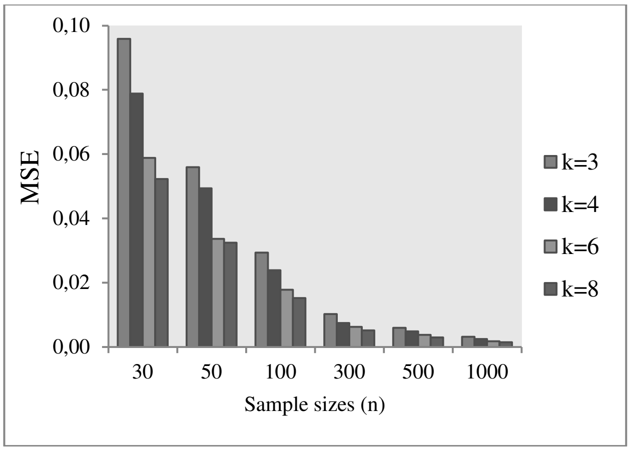

Figure 2. Bargraph of MSE($\hat{\mu}_j$) for μ_j= 1.

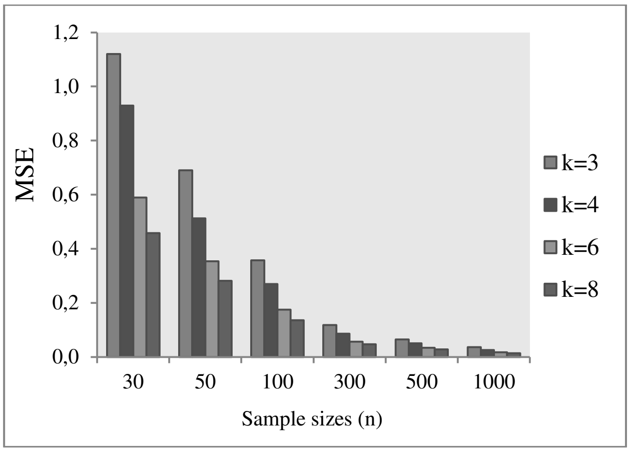

Figure-3. Bargraph of MSE($\hat{\mu}_j$) for μ_j= 5.

CONCLUSIONS

In this paper we discussed some properties of normal-Poisson model, its characterizations by variance function and by generalized variance, and also its generalized variance estimators. Then we showed that the variance (which is also the mean) of unobserved Poisson component can be estimated through the standardized generalized variance of the (k-1) normal components. The result from simulation study gives a conclusion that $\hat{\mu}_j$ is a consistent estimator of the Poisson variance.

REFERENCES

[1] Y. Boubacar Maïnassara and C.C. Kokonendji, "Normal stable Tweediemodels and power-generalized variance function of only one component", TEST, vol. 23, pp. 585-606, 2014.

[2] J. M. Bernardo and A.F.M. Smith, Bayesian Theory, New York: Wiley, 1993.

[3] O.E. Barndorff-Nielsen, J. Kent and M. Sørensen, "Normal variance-meanmixtures and z distributions", International Statistical Review, vol. 50, pp.145-159, 1982.

[4] C. C. Kokonendji and A. Masmoudi, "A characterization of Poisson Gaussian families by generalized variance", Bernoulli, vol. 12, pp. 371379, 2006.

[5] A. E. Koudou and D. Pommeret, "A characterization of Poisson-Gaussian families by convolution-stability", J. Multivariate Anal., vol. 81, pp. 120127, 2002.

[6] H. S. Steyn, "On the multivariate Poisson normal distribution", J. Am. Stat. Assoc., vol. 71, pp. 233-236, 1976.

ARPN Journal of Engineering and Applied Sciences

[7] C.D. Kemp and A.W. Kemp, "Some properties of the Hermite distribution", Biometrika, vol. 52, pp. 381-394, 1965.

[8] S. S. Wilks, Certain generalizations in the analysis of variance, Biometrika, vol. 24, pp. 471-494, 1932.

[9] L. G. Arvanitis and B. Afonja, "Use of the generalized variance and the gradient projection method in multivariate stratified sampling", Biometrics, vol. 27, pp. 119-127, 1971.

[10] S. L. Isaacson, "On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters", Ann. Math. Stat., vol. 22, pp. 217-234, 1951.

[11] M. Goodman, "A measure of overall variability in population", Biometrics, vol. 24, pp. 189-192, 1968.

[12] C. C. Kokonendji and V. Seshadri," On the determinant of the second derivative of a Laplace transform", Ann. Stat., vol. 24, pp. 1813-1827, 1996.

[13] C. C. Kokonendji and D. Pommeret, "Comparing UMVU and ML estimators of the generalized variance for natural exponential families", Statistics, vol. 41, pp. 547-558, 2007.

[14] C.C. Kokonendji and K. Nisa, "Generalized variance estimations of normal Poisson models", in Forging Connections between Computational Mathematics and Computational Geometry, Springer proceeding in Mathematics and Statistics, vol. 124, K. Chen and A. Ravindran, Eds., Switzerland: Springer, 2016, pp. 247-260.

[15] K. Nisa, C.C. Kokonendji and A. Saefuddin. Characterizations of multivariate normal Poisson models. Journal of Iranian Statistical Society, vol. 14, pp. 37-52, 2015.

[16] C.C. Kokonendji and K. Nisa, "Generalized variance in multivariate normal Poisson models and an analysis of variance", preprint of Laboratoirede mathematiques de Besancon, no. 2014/23, unpublished.

[17] K. Sato, Lévy Processes and Infinitely Divisible Distributions, Cambridge: Cambridge University Press, 1999.

[18] Y. Zi-Zhong, "Schur complement and determinant inequalities", Journal of Mathematical Inequalities, vol. 3, pp. 161-167, 2009.

[19] C. C. Kokonendji and A. Masmoudi, "On the Monge-Ampère equation for characterizing gamma-Gaussian model", Stat. Probabil. Lett. vol. 83,pp. 1692-1698, 2013.

[20] I. Gikhman and A. V. Skorokhod, the Theory of Stochastic Processes II, New York: Springer, 2004.

[21] A. Sen Gupta, "Tests for standardized generalized variances of multivariate normal populations of possibly different dimensions", Journal of Multivariate Analysis, vol. 23, pp. 209-219, 1987.