



Plagiarism Checker X Originality Report

Similarity Found: 9%

Date: Selasa, Agustus 20, 2019

Statistics: 225 words Plagiarized / 2507 Total words

Remarks: Low Plagiarism Detected - Your Document needs Optional Improvement.

Far East Journal of Mathematical Sciences (FJMS) © 2017 Pushpa Publishing House, Allahabad, India <http://www.pphmj.com> <http://dx.doi.org/10.17654/MS102030645>
Volume 102, Number 3, 2017, Pages 645-654 ISSN: 0972-0871 Received: March 24, 2017; Accepted: June 5, 2017 2010 Mathematics Subject Classification: 62J12. **Keywords and phrases:** longitudinal data, outlier, regression model.

A **ROBUST PROCEDURE FOR GEE MODEL** Netti Herawati and Khoirin Nisa Department of Mathematics University of Lampung Indonesia Abstract In longitudinal studies, multiple measurements are taken on the same subject at different points in time. Thus, observations for the same subject are correlated. This paper proposes a robust procedure for estimating parameters of regression model when generalized estimating equation (GEE) applied to longitudinal data that contains outliers.

The procedure is a combination of the iteratively reweighted least square (IRLS) and least trimmed square (LTS) methods and is called iteratively reweighted least trimmed square (IRLTS). We conducted a simulation study for gamma model and Poisson model using the proposed method, the result shows that our approach can provide a better result than the classical GEE. 1.

Introduction In statistics, generalized estimating equation (GEE) [5] is used to estimate the parameters of a generalized linear model (GLM) [6] with a possible unknown correlation between outcomes. It is a general statistical approach to fit a marginal model for longitudinal data analysis, and it has been popularly applied into clinical trials and biomedical studies.

GEEs belong to a class of regression techniques that are referred to as Netti Herawati

and Khoirin Nisa 646 semiparametric because they rely on specification of **only the first two moments**. Under correct model specification and mild regularity conditions, parameter estimates from GEEs are consistent. The **generalized estimating equation** approach requires correct **specification of the first two moments** of a model.

However, **these moment assumptions can be distorted by contaminated or irregular measurements namely outliers**. As a result, the **generalized estimating equation** method **fails to** give consistent estimators, and more seriously this will lead to incorrect conclusions [1, 8]. In this situation, we need **a robust method that** can minimize the effect of outliers.

In recent years a few studies have considered robust methods **for longitudinal data** analysis, see e.g. [1, 2, 4, 8, 11]. **In this paper, we** combine the IRLS and LTS for obtaining a robust estimation of GEE when data contain outliers. We have shown the effectiveness of this procedure for normal model [7]. **In this paper** we apply the proposed procedure to gamma and Poisson models. 2.

Generalized Estimating Equation Let the vector of measurements on the i th subject be $Y_i = (Y_{i1}, \dots, Y_{ip})^T$ with corresponding vector of means $\mu_i = (\mu_{i1}, \dots, \mu_{ip})^T$ and $X_i = (X_{i1}, \dots, X_{in_i})^T$ be the $p \times n_i$ matrix of covariates. In general, the components of Y_i are correlated but Y_i and Y_k are independent for any $k \neq i$. To model the relation between the response and covariates, we can use a regression model similar to the generalized linear models: $g(\mu_i) = X_i \beta$ where $g(\cdot)$ is a specified link function, and $\beta = (\beta_1, \dots, \beta_p)^T$ is a vector of unknown regression coefficients to be estimated.

The GEE for estimating the $1 \times p$ vector **of regression parameter** β is given by:
$$D(\beta)^T \sum_{i=1}^S W_i (Y_i - \mu_i) = 0 \quad (1)$$
 where W_i be the covariance matrix of Y_i modeled as $W_i = A_i R_i A_i^T$ A_i is a diagonal matrix of variance functions $A_{ij} = V(\mu_{ij})$ and R_i is **the working correlation matrix** of Y_i indexed by a vector of parameters. Solutions to equation (1) are obtained by alternating between estimation of β and R_i .

There are several specific choices of the form of **working correlation matrix** R_i commonly used to model the correlation matrix of Y_i among them are exchangeable and autoregressive correlation matrices. Solving for β is done with iteratively reweighted least squares (IRLS). **The following is an algorithm for fitting the specified model using GEEs** as described in [3] and [8]: 1.

Compute **an initial estimate of** $\hat{\beta}$, GEE β for example **with an ordinary generalized linear model** assuming independence. 2. A current estimate GEE $\hat{\beta}$ is updated by regressing

the working response vector $(\mu y \beta \mu \beta X Z \hat{\alpha} - ? ? + = * \text{ on } X$. A new estimate new $\hat{\beta}$ is obtained by: $(\hat{\beta} - \hat{\beta}_{old}) = Z W X X W X \hat{\beta}_{old}$ where W is a block diagonal weight matrix whose i th block is the $n_i \times n_i$ matrix $(\hat{\beta} - \hat{\beta}_{old})$.

1111 ----* ? ? ? ? ? ? ? ? ? ? ? ? = $\beta \mu A a R A \beta \mu W$ iiiii 3. Use new $\hat{\beta}$ to update $\hat{\alpha} = H Z \hat{\beta}$ new where $(\hat{\alpha} - \hat{\alpha}_{old}) = W X X W X X H T T$ 4. Iterate until convergence. 3. Iterated **Reweighted Least Trimmed Square** Let us briefly recall that the robust **estimation of regression** parameters Netti Herawati and Khoirin Nisa 648 using LTS [9] method is given by: $\hat{\beta} = h \hat{\beta}_{LTS}$ where $h = \frac{1}{2} \min \arg \hat{\beta}$ where $222 \ 2 \ 2 \ 1 \ n h$ eeee ===== "" are the ordered squared residuals, from smallest to largest.

LTS is calculated by minimizing the h ordered squares residuals, where h can be chosen between the range $4 \leq h \leq 312$ with n being sample size and number of parameters, respectively. One can refer to e.g. [9, 10] for some details on LTS method. The IRLTS procedure is stated in the following short algorithm. To motivate this method, it is **convenient to write** the algorithm with involving the residuals. 1.

Compute **an initial estimate of** GEE $\hat{\beta}$ using IRLS, use the estimate to calculate fitted value: $(\hat{\beta} - \hat{\beta}_{old}) = 2$. Calculate residuals: $\hat{\beta} - \hat{\beta}_{old}$ Sort $\hat{\beta} - \hat{\beta}_{old}$ for $i = 1, \dots, n$ in ascending order: $1 \ 2 \ 1 \ 1 \ m$ eee = = = " 3. Choose h observations which have the lowest h -residuals, we denote as subset H . 4. Improve estimates of $\hat{\beta}$ new $\hat{\beta}$ based on subset H using IRLS. 5. Iterate until convergence.

4. Simulation Study We compare the performances of IRLTS and IRLS through simulation study. Two types of outcomes are considered, continuous and count responses. The models for data generation are as follows: **A Robust Procedure for GEE Model** 649, $1 \ 2 \ 2 \ 1 \ 1 \ 0 \ i j i j i j \ x x \beta + \beta + \beta = \mu$, $2 \ 2 \ 1 \ 1 \ 0 \ i j i j i j \ x x \log \beta + \beta + \beta = \mu$ where $k \ \beta$'s for $2, 1, 0 = k$ are randomly generated, $200 \dots, 2, 1 = i$ and $.5 \dots, 2, 1 = j$ The covariates $i j \ x_1$ are i.i.d.

from a uniform distribution $(0, 1)$ and $2 \ x$ is the measurement time variable, i.e., $.5, 4, 3, 2, 1 \ 2 = i \ x$ For each scenario, we generate the data **based on the** underlying true correlation structures as exchangeable and autoregressive with $.5 \ 0 = a = e$ Table 1. The expected values, standard errors and MSEs of $\hat{\beta}$ for gamma distributed model with exchangeable correlation matrix Method Classical GEE IRLTS $\hat{\beta} \ e \ () \ 0 \ \hat{\beta} \ E \ () \ 0 \ \hat{\beta} \ SE$ MSE $() \ 0 \ \hat{\beta} \ E \ () \ 0 \ \hat{\beta} \ SE$ MSE 5% 1.833421 0.417356 0.419098 1.292539 0.280084 0.080563 10% 1.963008 0.551804 0.694453 1.343505 0.345998 0.119739 20% 2.105067 0.522999 0.861099 1.302799 0.245013 0.061308 30% 2.158723 0.712847 1.180859 1.432421 0.581205 0.346613 Netti Herawati and Khoirin Nisa 650 $1 \ \hat{\beta} \ e \ () \ 1 \ \hat{\beta} \ E \ () \ 1 \ \hat{\beta} \ SE$ MSE $() \ 1 \ \hat{\beta} \ E \ () \ 1 \ \hat{\beta} \ SE$ MSE 5% 0.664892 0.297197 0.205406 0.957103 0.034595

0.003693 10% 0.587415 0.348410 0.297492 0.968876 0.075239 0.007119 20% 0.314225
 0.347369 0.600688 0.964959 0.039372 0.003323 30% 0.297561 0.338291 0.617832
 0.937923 0.088894 0.012682 $2 \hat{\beta}_e()$ $2 \hat{\beta}_E()$ $2 \hat{\beta}_{SE}()$ $2 \hat{\beta}_E()$ $2 \hat{\beta}_{SE}()$ MSE 5%
 0.743453 0.289166 0.174850 0.974127 0.039936 0.006689 10% 0.549449 0.423348
 0.425292 0.958451 0.018566 0.007922 20% 0.355655 0.349199 0.597828 0.975630
 0.032734 0.005954 30% 0.332191 0.336312 0.621917 0.969043 0.071512 0.010960 Table
 2.

The expected values, standard errors and MSEs of $\hat{\beta}$ for gamma distributed model with
 autoregressive correlation matrix Method Classical GEE IRLTS $0 \hat{\beta}_e()$ $0 \hat{\beta}_E()$ $0 \hat{\beta}_{SE}()$
 MSE $0 \hat{\beta}_E()$ $0 \hat{\beta}_{SE}()$ MSE 5% 1.713422 0.174403 0.127509 1.169724 0.219998
 0.102270 10% 2.051958 0.292341 0.476604 1.172668 0.169948 0.081395 20% 2.096749
 0.279746 0.561178 1.182726 0.281815 0.127424 30% 2.208355 0.570003 0.975395
 1.094088 0.105856 0.105908 $1 \hat{\beta}_e()$ $1 \hat{\beta}_E()$ $1 \hat{\beta}_{SE}()$ MSE $1 \hat{\beta}_E()$ $1 \hat{\beta}_{SE}()$ MSE 5%
 0.746269 0.169773 0.087549 0.977105 0.032371 0.001180 10% 0.493787 0.257591
 0.311197 0.974237 0.026179 0.000892 20% 0.290721 0.257720 0.553459 0.982542
 0.035247 0.001279 30% 0.228004 0.231963 0.595925 0.981851 0.023087 0.000579 $2 \hat{\beta}_e()$
 $2 \hat{\beta}_E()$ $2 \hat{\beta}_{SE}()$ MSE $2 \hat{\beta}_E()$ $2 \hat{\beta}_{SE}()$ MSE 5% 0.739142 0.224797 0.132593
 0.981539 0.021532 0.002405 10% 0.511909 0.302973 0.355672 0.981291 0.028264
 0.002762 20% 0.281145 0.270400 0.627330 0.983924 0.028529 0.002551 30% 0.225163
 0.377600 0.646722 1.003760 0.035796 0.001758 **A Robust Procedure for GEE Model** 651
 As shown in Table 1 and Table 2, our approach (IRLTS) performs better **than the classical
 GEE.**

The MSEs of IRLTS are smaller than the MSEs of classical GEE, the outliers influence the
 estimation of $10 \hat{\beta}$, $\hat{\beta}_\beta$ and $\hat{\beta}_2$. The parameter estimates of classical GEE are much
 more influenced than the parameter estimates of IRLTS. The more outliers contained in
 the data the larger the deviation of classical GEE estimates from the parameter value.

In Table 3 and Table 4, the behavior of MSEs of both methods is the same as the first
 case, here we can see that IRLTS performs better **than the classical GEE** because the
 MSEs of IRLTS are smaller than the MSEs of classical GEE. Table 3. The expected values,
 standard errors and MSEs of $\hat{\beta}$ for Poisson distributed model with exchangeable
 correlation matrix Method Classical GEE IRLTS $0 \hat{\beta}_e()$ $0 \hat{\beta}_E()$ $0 \hat{\beta}_{SE}()$ $0 \hat{\beta}_E()$ $0 \hat{\beta}_{SE}()$ MSE 5% 7.851928 0.578389 49.012922 2.242959 0.411831 1.041084 10% 9.299508
 0.776268 71.575965 2.084935 0.454712 1.670854 20% 10.710302 0.425014 96.915019
 2.351178 0.546663 1.478125 30% 11.413164 0.621123 111.439999 1.988002 0.486656
 1.475743 $1 \hat{\beta}_e()$ $1 \hat{\beta}_E()$ $1 \hat{\beta}_{SE}()$ MSE $1 \hat{\beta}_E()$ $1 \hat{\beta}_{SE}()$ MSE 5% 0.309635 0.108437
 0.418891 1.052717 0.059291 0.014543 10% 0.194983 0.044670 0.568586 1.011790
 0.074132 0.009602 20% 0.097441 0.077894 0.729016 0.918343 0.038557 0.002349 30%

0.092040 0.031461 0.733153 0.809874 0.101063 0.029211 $2^{\wedge} \beta e () 2^{\wedge} BE () 2^{\wedge} BSE MSE$
 $() 2^{\wedge} BE () 2^{\wedge} BSE MSE$ 5% 0.281547 0.064601 0.478304 0.947537 0.021808 0.000986
 10% 0.159566 0.060507 0.660655 0.967633 0.065401 0.004284 20% 0.087528 0.031481
 0.779957 0.900333 0.028748 0.005696 30% 0.067109 0.035320 0.816673 0.825752
 0.057643 0.024164 Netti Herawati and Khoirin Nisa 652 The result for Poisson model
 shows similar behavior to the result for gamma model.

For the result of Poisson model in Table 3 and Table 4, IRLTS also performs better than the classical GEE. The MSEs of IRLTS are smaller than the MSEs of classical GEE, the outliers influence the estimation of β_0 and β_1 . The parameter estimates of classical GEE are much more influenced than the parameter estimates of IRLTS.

The more outliers contained in the data the larger the deviation of classical GEE estimates from the parameter value. In Table 3 and Table 4, the behavior of MSEs of both methods is the same as the first case, here we can see that IRLTS performs better than the classical GEE. The estimation of IRLTS yields better results than classical GEE for both cases we considered here.

The MSEs of IRLTS is smaller than classical GEE, this means that IRLTS can reduce the influence of the high leverage points better than the classical GEE. Table 4. The expected values, standard errors and MSEs of β_i for Poisson distributed model with autoregressive correlation matrix Method Classical GEE IRLTS $0^{\wedge} \beta e () 0^{\wedge} BE () 0^{\wedge} BSE$
 $MSE () 0^{\wedge} BE () 0^{\wedge} BSE MSE$ 5% 8.493574 0.689463 40.055071 1.614313 0.331758
 0.455831 10% 9.788645 0.962589 58.478701 1.765065 0.204619 0.523307 20%
 10.525858 0.877176 70.050502 2.356166 0.470975 0.724548 30% 10.635832 0.570803
 71.449706 3.071576 0.974599 1.054251 $1^{\wedge} \beta e () 1^{\wedge} BE () 1^{\wedge} BSE MSE () 1^{\wedge} BE () 1^{\wedge} BSE$
 $MSE 5% 0.309635 0.108437 0.418891 1.052717 0.059291 0.014543 10% 0.194983$
 0.044670 0.568586 1.011790 0.074132 0.009602 20% 0.097441 0.077894 0.729016
 0.918343 0.038557 0.002349 30% 0.092040 0.031461 0.733153 0.809874 0.101063
 0.029211 A Robust Procedure for GEE Model 653 $2^{\wedge} \beta e () 2^{\wedge} BE () 2^{\wedge} BSE MSE () 2^{\wedge} BE$
 $() 2^{\wedge} BSE MSE 5% 0.281547 0.064601 0.478304 0.947537 0.021808 0.000986 10%$
 0.159566 0.060507 0.660655 0.967633 0.065401 0.004284 20% 0.087528 0.031481
 0.779957 0.900333 0.028748 0.005696 30% 0.067109 0.035320 0.816673 0.825752
 0.057643 0.024164 5.

Concluding Remark In this paper, we have shown that our proposed procedure can minimize the effect of outliers on parameter estimation; IRLTS can produce a relatively efficient and consistent estimator compared to the classical GEE (IRLS). Based on the MSE, IRLTS performs much better than the classical GEE for gamma and Poisson models.

Acknowledgement The authors would like to thank the anonymous referees for their valuable suggestions and comments that led to a considerably improved manuscript. References [1] A. Abebe, J. W. McKean, J. D. Kloeke and Y. Bilgic, Iterated reweighted rank-based estimates for GEE models, Technical Report, 2014. [2] P. S. Gill, A robust mixed linear model analysis for longitudinal data, Stat. Med. 19 (2000), 975-987. [3] G. Johnston and M.

Stokes, Repeated measures analysis with discrete data using the SAS system, SUGI Proceeding, SAS Institute Inc., Cary, NC, 1996. [4] S. H. Jung and Z. Ying, Rank-based regression with repeated measurements data, Biometrika 90 (2003), 732-740. [5] K. Y. Liang and S. L. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73 (1986), 13-22. [6] P. McCullagh and J. A.

Nelder, Generalized Linear Models, Chapman and Hall, London, 1989. Netti Herawati and Khoirin Nisa 654 [7] K. Nisa and N. Herawati, Robust estimation of generalized estimating equation when data contain outliers, INSIST 2 (2017), 1-5. [8] B. F. Qaqish and J. S. Preisser, Resistant fits for regression with correlated outcomes: an estimating equations approach, J. Statist. Plann. Inference 75(2) (1999), 415-431. [9] R. J.

Rousseeuw and A. M. Leroy, Robust Regression and Outlier Detection, Wiley, New York, 1987. [10] P. J. Rousseeuw and K. van Driessche, Computing LTS regression for large data sets, Data Mining and Knowledge Discovery 12 (2006), 29-45. [11] Y. G. Wang and M. Zhu, Rank-based regression for analysis of repeated measures, Biometrika 93 (2006), 459-464.

INTERNET SOURCES:

<1% - <http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/01/P24.V3s.pdf>
<1% - https://www.researchgate.net/profile/Netti_Herawati
1% -
https://www.researchgate.net/publication/5207255_Rank-based_regression_with_repeated_measurements_data
<1% -
https://www.researchgate.net/publication/251717629_Generalized_Estimating_Equations_for_Zero-Inflated_Spatial_Count_Data
<1% -
https://www.researchgate.net/publication/273311220_Robust_Weighted_Least_Squares_Estimation_of_Regression_Parameter_in_the_Presence_of_Outliers_and_Heteroscedastic

Errors

1% - <http://repository.lppm.unila.ac.id/2972/>

1% -

https://www.researchgate.net/publication/319087162_A_robust_procedure_for_gee_model

<1% - <http://ufdc.ufl.edu/UFE0018500/00001>

1% - <http://downloads.hindawi.com/archive/2014/303728.xml>

1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2888330/>

1% -

https://www.researchgate.net/publication/308391249_Iterated_Reweighted_Rank-Based_Estimates_for_GEE_Models

<1% -

https://www.researchgate.net/publication/260993329_Synergy_and_redundancy_in_the_Granger_causal_analysis_of_dynamical_networks

<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1472692/>

<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1839993/>

<1% - <https://www.science.gov/topicpages/r/regression+models+analyzed.html>

<1% -

https://www.researchgate.net/publication/308262708_A_fast_and_effective_method_for_a_Poisson_denoising_model_with_total_variation

<1% -

<https://epdf.pub/logistic-regression-using-the-sas-system-theory-and-application.html>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167668713001510>

1% - <https://www.lexjansen.com/nesug/nesug96/NESUG96107.pdf>

<1% - <http://pages.stern.nyu.edu/~wgreene/Text/revisions/Chapter14-Revised.doc>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167587709003110>

<1% -

https://www.researchgate.net/publication/5207463_Rank-based_regression_for_analysis_of_repeated_measures

<1% - <https://issuu.com/saimm/docs/saimm-201904-apr>

<1% - <https://link.springer.com/article/10.1186/s40488-014-0018-0>

1% - <https://www.sciencedirect.com/science/article/pii/S016926079390046N>

<1% - <https://dl.acm.org/citation.cfm?id=40031>

1% - <https://business.uq.edu.au/profile/1448/min-zhu>



Plagiarism Checker X Originality Report

Similarity Found: 16%

Date: Selasa, Agustus 20, 2019

Statistics: 564 words Plagiarized / 3571 Total words

Remarks: Low Plagiarism Detected - Your Document needs Optional Improvement.

Applied Mathematical Sciences, Vol. 10, 2016, no. 63, 3107 - 3118 HIKARI Ltd,
www.m-hikari.com <https://doi.org/10.12988/ams.2016.69238> Empirical Comparison of
ML and UML Estimators of the Generalized Variance for some Normal Stable Tweedie
Models: a Simulation Study Khoirin Nisa¹ Department of Mathematics, Lampung
University, Indonesia Célestin C.

Kokonendji Laboratoire de Mathématiques de Besançon Université Bourgogne
Franche-Comté, France Asep Saefuddin, Aji Hamim Wigena Department of Statistics,
Bogor Agricultural University, Indonesia I Wayan Mangku Department of Mathematics,
Bogor Agricultural University, Indonesia Copyright c 2016 Khoirin Nisa, Célestin C.
Kokonendji, Asep Saefuddin, Aji Hamim Wigena and I Wayan Mangku.

This article is distributed under the Creative Commons Attribution License, which
permits unrestricted use, distribution, and reproduction in any medium, provided the
original work is properly cited. Abstract This paper discuss a comparison of the
maximum likelihood (ML) estimator and the uniformly minimum variance unbiased
(UMVU) estimator of generalized variance for some normal stable Tweedie models
through simulation study. We describe the estimation of some particular cases of
multivariate NST models, i.e.

normal gamma, normal Poisson Also affiliated to Bogor Agricultural University,
Indonesia and Université Bourgogne Franche-Comté, France 3108 Khoirin Nisa et al.
and normal invers-Gaussian. The result shows that UMVU method produces better
estimations than ML method on small samples and they both produce similar
estimations on large samples.

Mathematics Subject Classification: 62H12 Keywords: Multivariate natural exponential family, variance function, maximum likelihood, uniformly minimum variance unbiased 1 Introduction Normal stable Tweedie (NST) models were introduced by Boubacar Maïnassara and Kokonendji [3] as the extension of normal gamma [5] and normal inverse Gaussian [4] models.

NST models are composed by a fixed univariate stable Tweedie variable having a positive value domain, and the remaining random variables given the fixed one are real independent Gaussian variables with the same variance equal to the fixed component. For a k -dimensional ($k \geq 2$) NST random vector $X = (X_1, \dots, X_k)^T$, the generating s-finite positive measure $\mu_{a,t}$ is given by $\mu_{a,t}(dx) = \mu_{a,t}(dx_1) \prod_{j=2}^k \gamma_{2,x_j}(dx_j)$, (1) where $\mu_{a,t}$ is the well-known probability measure of univariate positive s-stable distribution generating Lévy process $(X_a(t))_{t \geq 0}$ which was introduced by Feller [7] as follows $\mu_{a,t}(dx) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-xt} G(1 + a\gamma(r) \sin(-r\gamma(a)) r^{a-1} (1-a)x)^{a-1} dx$ $= \mu_{a,t}(x) dx$. (2) Here $a \in (0, 1)$ is the index parameter, $G(\cdot)$

is the classical gamma function, and I_A denotes the indicator function of any given event A that takes the value 1 if the event occurs and 0 otherwise. Parameter a can be extended into $a \in (-8, 2]$ [10]. For $a = 2$ in (2), we obtain the normal distribution with density $\gamma_{2,t}(dx) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx$. In multivariate analysis, including NST models, generalized variance has important roles in descriptive analysis and inferences.

In this paper we discuss the ML and UMVU generalized variance estimators of the following NST models: Comparison of ML and UMVU estimators ... 3109 1. Normal gamma (NG). For $a = 0$ in (1) one has the generating measure of normal gamma as follows: $\mu_{0,t}(dx) = t^{k-1} \prod_{j=1}^k \gamma_{2,x_j}(k-1)/2 \mu_{0,t}(t) \exp\left(-\frac{x_1^2}{2t}\right) \prod_{j=2}^k \gamma_{2,x_j}(t) dx_1 \dots dx_k$.

(3) It is a member of simple quadratic natural exponential families (NEFs) [6] and was called as "gamma-Gaussian" which was characterized by Kokonendji and Masmoedi [8]. 2. Normal inverse Gaussian (NIG). For $a = 1/2$ in (1) we can write the normal inverse Gaussian generating measure as follows $\mu_{1/2,t}(dx) = t^{-(k+2)/2} \prod_{j=1}^k \gamma_{2,x_j}(k/2) \exp\left(-\frac{x_1^2}{2t}\right) \prod_{j=2}^k \gamma_{2,x_j}(t^2 + k \sum_{j=2}^k x_j^2) dx_1 \dots dx_k$.

(4) It was introduced as a variance-mean mixture of a univariate inverse Gaussian with multivariate Gaussian distribution [4] and has been used in finance (see e.g. [1, 2]). 3. Normal Poisson (NP). For the limit case $a = -8$ in (1) we have the so-called normal Poisson generating measure $\mu_{-8,t}(dx) = t^{k-1} \prod_{j=1}^k \gamma_{2,x_j}(k-1)/2 \exp\left(-\frac{x_1^2}{2t}\right) \prod_{j=2}^k \gamma_{2,x_j}(t) dx_1 \dots dx_k$.

(5) Since it is also possible to have $x_1 = 0$ in the Poisson part, the corresponding normal Poisson distribution is degenerated as d_0 . This model is recently characterized by Nisa et al. [9] 2 Generalized Variance of NST Models The cumulant function $K_{a,t}(\cdot)$ of NST models is given by $K_{a,t}(\cdot) = K_{a,t}(\cdot) + \sum_{j=2}^{\infty} \frac{1}{j!} \kappa_j X_j$ (6) where $K_{a,t} = \log R \kappa \exp(\cdot)_{a,t}(\cdot)$ is the cumulant function of the associated univariate stable Tweedie distribution a,t .

Then for each distribution we 3110 Khoirin Nisa et al. discuss here the corresponding cumulant function is given by $K_{a,t}(\cdot) = \sum_{j=2}^{\infty} \frac{1}{j!} \kappa_j X_j$ for NG - $t \sum_{j=2}^{\infty} \frac{1}{j!} \kappa_j X_j$, for NIG - $t \log \sum_{j=2}^{\infty} \frac{1}{j!} \kappa_j X_j$, for NP (7) (see [3, Section 2]).

The cumulant function is finite for \cdot in canonical domain $T(a,t) = \{\cdot \mid R \kappa; \sum_{j=2}^{\infty} \frac{1}{j!} \kappa_j P_j \in T(a,1)\}$ with $T(a,1) = (-8, 0)$ for NG $(-8, 0]$ for NIG R for NP. Let $G(a,t) = \int_0^\infty \int_{-\infty}^\infty \exp(-\cdot y) \cdot y^k dP(\cdot; a, y, t)$ be the set of probability distributions $P(\cdot; a, y, t)$ ($dx = \exp(-\cdot x) K_{a,t}(\cdot)_{a,t}(dx)$).

The variance function which is the variance-covariance matrix in term of mean parameterization; $P(\mu; G_{a,t}) := P(\cdot; \mu; a, t)$; is obtained through the second derivative of the cumulant function, i.e. $V_{G_{a,t}}(\mu) = K''_{a,t}(\mu)$ where $\mu = K'_{a,t}(\cdot)$. Then calculating the determinant of the variance function will give the generalized variance.

We summarize the variance function and the generalized variance of NG, NIG and NP models in Table 1. Table 1: Variance Function and Generalized Variance Model $V_{G_{a,t}}(\mu) = \det V_{G_{a,t}}(\mu)$ NG $(1/t)\mu\mu + \text{diag}(0, \mu_1, \mu_2, \dots, \mu_k)$ (1/t) $\mu_{k+1} \dots \mu_n$ NIG $(\mu_1/t^2)\mu\mu + \text{diag}(0, \mu_1, \mu_2, \dots, \mu_k)$ (1/t^2) $\mu_{k+2} \dots \mu_n$ NP $(1/\mu_1)\mu\mu + \text{diag}(0, \mu_1, \mu_2, \dots, \mu_k)$ μ_k 1 The ML and UMVU estimators of the generalized variance in Table 1 are stated in the following proposition. Proposition 1 Let X_1, \dots, X_n be random vectors with distribution $P(\cdot; a, y, t)$ $G(\cdot; p, t)$ in a given NST family. Denoting $X = (X_1, \dots, X_n)^T$ the sample mean with positive first component X_1 , the ML estimator of the generalized variance of NG, NP and NIG models is given by: $T_{n;k;t} = \det V_{G_p,t}(X) = \sum_{k=1}^n (1/t) X_{k+1}^2$, for NG X_{k+1} , for NP $(1/t^2) X_{k+2}^2$, for NIG Comparison of ML and UMVU estimators ...

3111 and the UMVU estimator is given by $U_{n;k,t} = \sum_{i=1}^n t G(nt) [G(nt + k + 1) - 1] S_{n,i} x_{k+1}(1i)$, for NG $n-k [S_{n,i=1} x(1i)] [S_{n,i=1} x(1i) - 1] \dots [S_{n,i=1} x(1i) - k + 1]$, $y_n = k$ for NP $t k^2 - 1 - k/2 [G(1 + k/2)] - 1 S_{n,i=1} x^3/2(1i) \exp(-nt) / [2 S_{n,i=1} x(1i)] \times R S_{n,i=1} x(1i) 0 y k/2 - 1 [S_{n,i=1} x(1i) - y_1] - 3/2 \times \exp(-y_1) - [(nt)^2 / 2 [S_{n,i=1} x(1i) - y_1]] dy_1$, for NIG (see Boubacar Maïnassara and Kokonendji, [3]) 3 Simulation Study In order to examine the behavior of ML and UMVU estimators empirically we carried out a

simulation study. We run Monte-Carlo simulations using R software.

We set several sample sizes (n) varied from 3 to 1000 and we generated 1000 samples for each n . We consider $k = 2, 4, 6$ to see the effects of k on generalized variance estimations. For simplicity we set $\mu_1 = 1$. Moreover, to see the effect of zero values proportion within X_1 in the case of normal Poisson, we also consider small mean values on the Poisson component i.e.

$\mu_1 = 0.5$ because $P(X_1 = 0) = \exp(-\mu_1)$. We report the numerical results of the generalized variance estimations for each model, i.e. the empirical expected value of the estimators with its standard errors (Se) and the empirical mean square error (MSE). We calculated the mean square error (MSE) of each method over 1000 data sets using the following formula: $MSE(\hat{b}) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{b}_i - b)^2$ where \hat{b} is the estimate of $\det V_{Ga,t}(\mu)$ using ML and UMVU estimators. 3.1

Normal gamma We generated normal gamma distribution samples using the generating sufficient positive measure of normal gamma in (1). Table 2 show the expected values of generalized variance estimates with their standard errors (in parentheses) and the means square error values of both ML and UMVU methods in case of normal gamma.

From the result in Table 2 we can observe different performances of ML estimator ($T_{n;k,t}$) and UMVU estimator ($U_{n;k,p,t}$) of the generalized variance. The expected values of $T_{n;k,t}$ converge while the values of $U_{n;k,t}$ do not, but $U_{n;k,t}$ is always closer to the parameter than $T_{n;k,t}$ for small sample sizes, i.e. 3112 Khoirin Nisa et al. for $n = 30$, this shows that UMVU is an unbiased estimator while ML is an asymptotically unbiased estimator.

For the two methods, the standard error of the estimates decreases when the sample size increase. Table 2: The expected values (with empirical standard errors) and MSE of $T_{n;k,t}$ and $U_{n;k,t}$ for normal-gamma with 1000 replications for given target value $\mu_{k+1} = 1$ with $k \in \{2, 4, 6\}$. Expected values and Standard errors

n	$T_{n;2,t}$	$U_{n;2,t}$	$T_{n;4,t}$	$U_{n;4,t}$	$T_{n;6,t}$	$U_{n;6,t}$
30	1.9805 (3.7192)	0.8912 (1.6736)	14.7935 (2.8128)	10.12878 (1.2875)	0.9756 (0.9754)	1.7405 (0.9520)
100	1.1648 (0.8236)	1.0085 (0.7131)	0.7054 (0.5085)	30.10998 (0.6031)	0.9978 (0.5471)	0.3736 (0.2994)
300	1.0380 (0.4115)	0.9881 (0.3917)	0.1708 (0.1536)	100.10231 (0.3152)	0.9931 (0.3060)	0.0999 (0.0937)
1000	1.0036 (0.1774)	0.9936 (0.1757)	0.0315 (0.0309)	500.10076 (0.1365)	1.0016 (0.1357)	0.0187 (0.0184)
4	4.2191 (13.3899)	0.8720 (2.7674)	189.6509 (7.6750)	2.3799 (5.0869)	0.9906 (2.1174)	27.7810 (4.4837)
20	1.6461 (2.0572)	1.0328 (1.2906)	4.6494 (1.6668)	30.13831 (1.3505)	1.0066 (0.9828)	1.9707 (0.9660)
60	1.1904 (0.8014)					

1.0117 (0.6811) 0.6784 0.4640 100 1.0869 (0.5706) 0.9849 (0.5171) 0.3332 0.2676 300
 1.0293 (0.2938) 0.9957 (0.2842) 0.0872 0.0808 500 1.0286 (0.2296) 1.0083 (0.2251)
 0.0535 0.0507 1000 1.0137 (0.1610) 1.0036 (0.1594) 0.0261 0.0254 6 7 13.7175
 (103.5833) 1.3062 (9.8634) 10891.2275 97.3811 10 6.6118 (36.8236) 1.1467 (6.3866)
 1387.4736 40.8103 20 2.2455 (4.3052) 0.8670 (1.6622) 20.0860 2.7806 30 1.9055 (3.4774)
 0.9905 (1.8076) 12.9123 3.2676 60 1.4151 (1.5070) 1.0092 (1.0748) 2.4434 1.1553 100
 1.2248 (0.8843) 0.9972 (0.7199) 0.8325 0.5183 300 1.0606 (0.4416) 0.9894 (0.4119)
 0.1986 0.1698 500 1.0182 (0.3160) 0.9765 (0.3030) 0.1002 0.0924 1000 1.0228 (0.2311)
 1.0016 (0.2263) 0.0539 0.0512 To examine the consistency of the estimators we have to
 look at their MSE.

The result shows that when n increases the MSE of the two methods become more similar and they both produced almost the same result for $n = 1000$. The MSE values for $n = 10$ in the table are presented graphically in Figure 1. In the figure we can easily see that all estimators become more similar when the sample size increase.

For small sample sizes, UMVU always has smaller MSE, in this situation UMVU is preferable than ML. The figure also shows that the difference between ML and UMVU for small sample sizes increases when the dimension increases. Comparison of ML and UMVU estimators ...

3113 (a) $k=2$ (b) $k=4$ (c) $k=6$ Figure 1: Bargraphs of the mean square errors of $T_{n;k,t}$ and $U_{n;k,t}$ for normal- gamma with $n \in \{10, 20, 30, 60, 100, 300, 500, 1000\}$ and $k \in \{2, 4, 6\}$. 3.2 Normal inverse-Gaussian The result for normal inverse-Gaussian is presented in Table 3. Similar with normal gamma, the result for normal inverse-Gaussian shows that UMVU method produced better estimates than ML method for small sample sizes.

From the result we can conclude that the two estimators are consistent. The bargraph of MSE values for $n = 10$ in Table 3 is presented in Figure 2. Notice that the result for this case is similar to the normal gamma case, i.e. for small sample sizes the difference between the MSEs of ML and UMVU estimators for normal inverse-Gaussian also increases when k increases.

(a) $k=2$ (b) $k=4$ (c) $k=6$ Figure 2: Bargraphs of the mean square errors of $T_{n;k,t}$ and $U_{n;k,t}$ for normal inverse Gaussian with $n \in \{10, 20, 30, 60, 100, 300, 500, 1000\}$ and $k \in \{2, 4, 6\}$. 3.3 Normal Poisson The simulation results for normal Poisson are presented in Table 4 and Table 5 for $\mu_1 = 1$ and $\mu_1 = 0.5$ respectively. In this simulation, the proportion of zero values in the samples increases when the mean of the Poisson component becomes smaller.

For normal-Poisson distribution with $\mu_j = 0.5$, we have many zero values in the samples. However, this situation does not affect the 3114 Khoirin Nisa et al. Table 3: The expected values (with standar errors) and MSE of $T_{n;k,t}$ and $U_{n;k,t}$ for normal inverse-Gaussian with 1000 replications for given target value $\mu_{k+2} = 1$ and $k \in \{2, 4, 6\}$. Expected values and Standard errors MSE $k \in T_n; k, t U_n; k, t T_n; k, t U_n; k, t$ 2 3 2.0068 (4.9227) 0.9135 (0.8235) 25.2469 0.6856 10 1.4249 (2.8513) 1.0316 (0.4388) 8.3103 0.1935 20 1.5936 (1.8951) 1.1340 (0.3718) 3.9439 0.1562 30 1.3677 (1.0155) 1.1641 (0.2668) 1.1664 0.0981 60 1.0846 (0.5341) 1.1104 (0.1856) 0.2924 0.0466 100 1.0819 (0.5166) 1.1102 (0.1675) 0.2735 0.0402 300 1.0006 (0.2570) 1.0843 (0.0919) 0.0660 0.0156 500 1.0356 (0.1890) 1.1374 (0.0727) 0.0370 0.0242 1000 1.0156 (0.1219) 1.0116 (0.0670) 0.0151 0.0115 4 5 9.3836 (30.0947) 1.3196 (1.1323) 975.9726 1.3843 10 4.6547 (13.8643) 1.2837 (0.8153) 205.5754 0.7452 20 2.7487 (5.1845) 1.2963 (0.6189) 29.9373 0.4709 30 1.4822 (2.1166) 1.1854 (0.4572) 4.7125 0.2434 60 1.3095 (1.1051) 1.2560 (0.3054) 1.3170 0.1588 100 1.1673 (0.8467) 1.2264 (0.2671) 0.7449 0.1226 300 1.0849 (0.4296) 1.2542 (0.1520) 0.1918 0.0877 500 1.0350 (0.2839) 1.0762 (0.0914) 0.0818 0.0416 1000 1.0107 (0.2080) 1.0102 (0.1137) 0.0434 0.0337 6 7 20.4865 (113.4633) 0.9423 (0.9984) 12056.9414 1.0001 10 12.1032 (55.7841) 1.0596 (0.8610) 2329.5787 0.7449 20 3.4498 (10.3056) 1.0054 (0.5933) 112.2060 0.3520 30 2.1422 (3.2262) 1.0246 (0.4970) 11.7130 0.2476 60 1.8236 (2.6064) 1.0587 (0.3744) 7.4717 0.1436 100 1.2468 (1.1599) 1.0129 (0.2643) 1.4062 0.1170 300 1.0781 (0.4953) 1.0568 (0.1596) 0.2514 0.0929 500 1.0815 (0.4065) 1.0230 (0.1110) 0.1719 0.0922 1000 1.0207 (0.2816) 1.0204 (0.0775) 0.0798 0.0760 generalized variance estimation as we can see that $T_{n;k,t}$ and $U_{n;k,t}$ have the same behavior for both values of μ_1 .

The MSE in Table 4 and 5 for $n = 10$ are displayed as bargraphs presented in Figure 3 and Figure 4. From those figures we see that UMVU is preferable than ML because it always has smaller MSE values when sample sizes are small ($n \leq 30$). 4 Conclusion In this paper we have discussed the generalized variance estimator of normal gamma, normal inverse-Gaussian and normal Poisson models using ML and UMVU methods. The simulation studies of the generalized variance estimators Comparison of ML and UMVU e stimators ...

3115 Table 4: The expected values (with standar errors) and MSE of $T_{n;k,t}$ and $U_{n;k,t}$ for normal Poisson with 1000 replications for given target value $\mu_k = 1$ and $k \in \{2, 4, 6\}$. Expected values and Standard errors MSE $k \in T_n; k, t U_n; k, t T_n; k, t U_n; k, t$ 2 3 1.3711 (1.4982) 1.0349 (1.3130) 2.3824 1.7252 10 1.0810 (0.6589) 0.9817 (0.6286) 0.4407 0.3955 20 1.0424 (0.4471) 0.9925 (0.4363) 0.2017 0.1904 30 1.0329 (0.3817) 0.9996 (0.3756) 0.1468 0.1411 60 1.0184 (0.2661) 1.0017 (0.2639) 0.0711 0.0697 100 1.0066 (0.2016) 0.9966 (0.2006) 0.0407 0.0403 300 1.0112 (0.1153) 1.0079 (0.1151) 0.0134 0.0133 500 0.9986 (0.0942) 0.9966 (0.0941) 0.0089 0.0089 1000 0.9998 (0.0641) 0.9988

(0.0641) 0.0041 0.0041 4 5 2.6283 (5.0058) 1.0721 (2.7753) 27.7093 7.7075 10 1.7362 (2.2949) 1.0422 (1.6267) 5.8085 2.6480 20 1.3276 (1.1713) 1.0073 (0.9588) 1.4793 0.9193 30 1.2274 (0.8892) 1.0167 (0.7750) 0.8424 0.6008 60 1.1111 (0.5643) 1.0085 (0.5250) 0.3308 0.2757 100 1.0647 (0.4448) 1.0038 (0.4260) 0.2021 0.1815 300 1.0245 (0.2389) 1.0043 (0.2354) 0.0577 0.0554 500 1.0092 (0.1889) 0.9972 (0.1872) 0.0358 0.0351 1000 1.0013 (0.1272) 0.9953 (0.1267) 0.0162 0.0161 6 7 4.5153 (12.8404) 0.9378 (4.0255) 177.2319 16.2084 10 3.6865 (8.1473) 1.1642 (3.3992) 73.5952 11.5816 20 1.9674 (2.9034) 1.0227 (1.7467) 9.3656 3.0514 30 1.5605 (1.8825) 0.9901 (1.3133) 3.8580 1.7250 60 1.2954 (1.0360) 1.0220 (0.8541) 1.1606 0.7300 100 1.2084 (0.7824) 1.0462 (0.6957) 0.6556 0.4861 300 1.0621 (0.3793) 1.0109 (0.3641) 0.1477 0.1327 500 1.0294 (0.2778) 0.9992 (0.2710) 0.0780 0.0734 1000 1.0185 (0.1939) 1.0034 (0.1915) 0.0379 0.0367 (a) k=2 (b) k=4 (c) k=6 Figure 3: Bargraphs of the mean square errors of $T_{n;k,t}$ and $U_{n;k,t}$ for normal Poisson with $\mu_1 = 1$, $n \in \{10, 20, 30, 60, 100, 300, 500, 1000\}$ and $k \in \{2, 4, 6\}$. 3116 Khoirin Nisa et al.

Table 5: The expected values (with standard errors) and MSE of $T_{n;k,t}$ and $U_{n;k,t}$ for normal Poisson with 1000 replications for given target value $\mu_k = 0.5k$ and $k \in \{2, 4, 6\}$. Expected values and Standard errors MSE $k \quad n \quad T_{n;k,t} \quad U_{n;k,t} \quad T_{n;k,t} \quad U_{n;k,t}$ 2 3 0.3930 (0.5426) 0.2320 (0.4223) 0.3148 0.1787 10 0.2868 (0.2421) 0.2378 (0.2212) 0.0600 0.0491 20 0.2652 (0.1660) 0.2407 (0.1583) 0.0278 0.0251 30 0.2642 (0.1374) 0.2476 (0.1332) 0.0191 0.0177 60 0.2598 (0.0903) 0.2514 (0.0888) 0.0083 0.0079 100 0.2534 (0.0712) 0.2484 (0.0705) 0.0051 0.0050 300 0.2495 (0.0418) 0.2478 (0.0417) 0.0017 0.0017 500 0.2491 (0.0313) 0.2482 (0.0313) 0.0010 0.0010 1000 0.2495 (0.0221) 0.2490 (0.0221) 0.0005 0.0005 4 5 0.2999 (0.8462) 0.0685 (0.3474) 0.7724 0.1207 10 0.1696 (0.3115) 0.0689 (0.1750) 0.1085 0.0306 20 0.1089 (0.1541) 0.0658 (0.1097) 0.0259 0.0120 30 0.0886 (0.0894) 0.0617 (0.0689) 0.0087 0.0048 60 0.0774 (0.0559) 0.0642 (0.0487) 0.0033 0.0024 100 0.0704 (0.0403) 0.0627 (0.0370) 0.0017 0.0014 300 0.0643 (0.0207) 0.0618 (0.0201) 0.0004 0.0004 500 0.0635 (0.0158) 0.0620 (0.0156) 0.0003 0.0002 1000 0.0631 (0.0115) 0.0624 (0.0114) 0.0001 0.0001 6 7 0.2792 (1.2521) 0.0268 (0.2274) 1.6371 0.0519 10 0.1212 (0.3918) 0.0165 (0.0858) 0.1646 0.0074 20 0.0427 (0.0883) 0.0124 (0.0345) 0.0085 0.0012 30 0.0356 (0.0539) 0.0151 (0.0271) 0.0033 0.0007 60 0.0236 (0.0281) 0.0149 (0.0196) 0.0009 0.0004 100 0.0211 (0.0183) 0.0159 (0.0145) 0.0004 0.0002 300 0.0173 (0.0089) 0.0157 (0.0082) 0.0001 0.0001 500 0.0166 (0.0068) 0.0157 (0.0064) 0.0000 0.0000 1000 0.0164 (0.0044) 0.0159 (0.0043) 0.0000 0.0000 for the three models show that UMVU produces better estimation than ML for small sample sizes.

However, the two methods are consistent and they become more similar when the sample size increases. References [1] K. K. Aase, Representative agent pricing of financial assets based on Lévy processes with normal inverse Gaussian marginals, Annals of Operations Research, 114 (2002), 15-31. <https://doi.org/10.1023/a:1021093615674>

Comparison of ML and UMVU estimators ...

3117 (a) $k=2$ (b) $k=4$ (c) $k=6$ Figure 4: Bargraphs of the mean square errors of $T_{n;k,t}$ and $U_{n;k,t}$ for normal Poisson with $\mu_1 = 0.5$, $n \in \{10, 20, 30, 60, 100, 300, 500, 1000\}$ and $k \in \{2, 4, 6\}$. [2] J. Andersson, On the normal inverse Gaussian stochastic volatility model, *Journal of Business & Economic Statistics*, 19 (2001), no. 1, 44-54. <https://doi.org/10.1198/07350010152472607> [3] Y. Boubacar Maïnassara and C. C.

Kokonendji, On normal stable Tweedie models and power-generalized variance functions of only one component, *TEST*, 23 (2014), 585-606. <https://doi.org/10.1007/s11749-014-0363-9> [4] O. E. Barndorff-Nielsen, J. Kent and M. Sørensen, Normal variance-mean mixtures and z distributions, *International Statistical Review*, 50 (1982), 145-159. <https://doi.org/10.2307/1402598> [5] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*, Wiley, New York, 1993. <https://doi.org/10.1002/9780470316870> [6] M.

Casalis, The $2d + 4$ simple quadratic natural exponential families on R^d , *The Annals of Statistics*, 24 (1996), 1828-1854. <https://doi.org/10.1214/aos/1032298298> [7] W. Feller, *An Introduction to Probability Theory and its Applications Vol. II*, Second edition, Wiley, New York, 1971. [8] C. C. Kokonendji and A. Masmoudi, On the Monge-Ampère equation for characterizing gamma-Gaussian model, *Statistics and Probability Letters*, 83 (2013), 1692-1698. <https://doi.org/10.1016/j.spl.2013.03.023> [9] K. Nisa, C.C. Kokonendji and A. Saefuddin, Characterizations of multi-variate normal Poisson model, *Journal of Iranian Statistical Society*, 14 (2015), 37-52.

[10] M. C. K. Tweedie, An index which distinguishes between some important exponential families, In: *Statistics: Applications and New Directions. Proceedings of the Indian Statistical Golden Jubilee International Conference*, 3118 Khoirin Nisa et al. Eds. J. K. Ghosh and J. Roy, Indian Statistical Institute, Calcutta, 1984, 579-604.

Received: September 16, 2016; Published: November 14, 2016

INTERNET SOURCES:

<1% - <http://www.m-hikari.com/ams/ams-2016/ams-61-64-2016/index.html>

10% -

<http://www.m-hikari.com/ams/ams-2016/ams-61-64-2016/p/nisaAMS61-64-2016.pdf>

1% - http://www.arpnjournals.com/jeas/volume_12_2017.htm

1% - <https://bio.biologists.org/content/biolopen/early/2016/08/03/bio.017673.full.pdf>

1% - <https://ceramics.onlinelibrary.wiley.com/doi/pdf/10.1111/jace.12732>

<1% -

<https://de.mathworks.com/matlabcentral/fileexchange/6050-normal-inverse-gaussian-distribution?tab=function>

1% - <https://www.cs.ubc.ca/~murphyk/Teaching/CS340-Fall07/reading/NG.pdf>

<1% -

<https://study.com/academy/answer/using-green-s-theorem-calculate-cf-dr-where-f-x-y-2xyi-y2-x2-j-x2-y2-2-and-c-is-any-positively-oriented-simple-closed-curve-that-encloses-the-origin-cf-dr.html>

1% -

https://www.researchgate.net/publication/4721160_On_the_Normal_Inverse_Gaussian_Stochastic_Volatility_Model

<1% - <https://www.sciencedirect.com/science/article/pii/S0370157397000161>

<1% - <https://nph.onlinelibrary.wiley.com/doi/10.1111/nph.15859>

<1% - <https://www.scribd.com/document/403961654/Evaluation-techniques>

<1% - <http://ufdc.ufl.edu/UF00001565/16633>

<1% -

https://mafiadoc.com/sufficient-sample-sizes-for-multilevel-modeling_5c3cc5ac097c4776678b45a5.html

<1% - <https://wenku.baidu.com/view/d7df2c1555270722192ef7d8.html>

<1% -

https://mafiadoc.com/xt-longitudinal-data-panel-data-stata_59fa11c71723dd6471623b49.html

<1% -

http://people.stat.sfu.ca/~lockhart/richard/830/13_3/lectures/nonparametric_basics/notes.pdf

<1% -

<https://epdf.pub/probability-and-random-processes-for-electrical-and-computer-engineers3392.html>

<1% - <https://www.scribd.com/document/106290277/Probdist-Ref>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167947313002806>

<1% - https://en.wikipedia.org/wiki/Standard_error

<1% - <http://people.stern.nyu.edu/wgreene/nonlinearfixedeffects.pdf>

<1% -

https://www.researchgate.net/publication/330010463_Effect_of_Sample_Size_on_Estimation_of_Fertility_using_Open_Birth_Interval_Data

<1% - <https://www.mdpi.com/2079-3197/4/1/3/htm>

<1% -

https://www.researchgate.net/publication/24116268_When_is_some_number_really_better_than_no_number_On_the_optimal_choice_between_non-market_valuation_methods

<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2829998/>

<1% -

https://www.researchgate.net/publication/321857310_Estimation_of_the_PMF_and_CDF_of_some_standard_discrete_distributions_useful_in_reliability_modelling

<1% -

https://www.researchgate.net/publication/266393627_Normal_variance-mean_mixtures_I_An_inequality_between_skewness_and_kurtosis

<1% -

https://www.researchgate.net/publication/220462437_Representative_Agent_Pricing_of_Financial_Assets_Based_on_Levy_Processes_with_Normal_Inverse_Gaussian_Marginals

<1% - <https://plato.stanford.edu/entries/logic-power-games/>

<1% - <https://www.sciencedirect.com/science/article/pii/S0377042716303077>

<1% - <http://www.umass.edu/wsp/resources/poisson/>

<1% -

https://www.researchgate.net/publication/11352301_Concurrent_Learning_of_Temporal_and_Spatial_Sequences

<1% -

https://www.researchgate.net/publication/232834424_Comparing_UMVU_and_ML_estimators_of_the_generalized_variance_for_natural_exponential_families

<1% -

https://www.researchgate.net/publication/316883157_Some_Estimation_Methods_for_the_Shape_Parameter_and_Reliability_Function_of_Burr_Type_XII_Distribution_Comparison_Study

<1% -

<https://www.scribd.com/document/354772208/D3665-Random-Sampling-of-Construction-Materials>

<1% - <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1467-9531.2009.01221.x>

<1% - <https://academic.oup.com/ije/article/44/3/1051/632956>

<1% - http://www.nucleide.org/ICRM_GSWG/Training/Efficiency.pdf

<1% - <https://amstat.tandfonline.com/doi/abs/10.1198/07350010152472607>

<1% - <https://link.springer.com/article/10.1007/s11749-014-0363-9>

<1% - <https://link.springer.com/article/10.1007/s11222-016-9649-y>

<1% - <https://epdf.pub/univariate-discrete-distributions-3rd-edition.html>

1% - <https://wenku.baidu.com/view/1ae57f1a10a6f524ccbf8514.html>

<1% - <https://rdr.io/cran/tweedie/man/tweedie-package.html>

<1% - https://en.wikipedia.org/wiki/Gwen_Lfill



Plagiarism Checker X Originality Report

Similarity Found: 20%

Date: Selasa, Agustus 20, 2019

Statistics: 1303 words Plagiarized / 6428 Total words

Remarks: Medium Plagiarism Detected - Your Document needs Selective Improvement.

JIRSS (2015) Vol. 14, No. 2, pp 37-52 DOI:10.7508 / jirss.2015.02.003 Characterizations of Multivariate Normal-Poisson Model Khoirin Nisa 1, 2, 3, Célestin C. Kokonendji 1, Asep Saefuddin 2 1 University of Franche-Comté, Besançon, France. 2 Bogor Agricultural University, Bogor, Indonesia. 3 Lampung University, Bandar Lampung, Indonesia. Abstract.

Multivariate normal-Poisson model has been recently introduced as a special case of normal stable Tweedie models. The model is composed of a univariate Poisson variable, and the remaining variables given the Poisson one are independent Gaussian variables with variance the value of the Poisson component.

Two characterizations of this model are shown, first by variance function and then by generalized variance function which is the determinant of the variance function. The latter provides an explicit solution of a particular Monge-Ampère equation. Keywords. Generalized variance, Infinitely divisible measure, Monge-Ampère equation, Multivariate exponential family, Variance function. MSC: 62H05; 60E07.

1 Introduction Motivated by normal gamma and normal inverse Gaussian models, Boubacar Maïnassara and Kokonendji (2014) introduced a new form of generalized variance functions which are generated by the so-called normal stable Tweedie (NST) models of k -variate Khoirin Nisa (khoirin.nisa@univ-fcomte.fr), Célestin C. Kokonendji (celestin.kokonendji@univ-fcomte.fr), Asep Saefuddin (asaefuddin@ipb.ac.id) 38 Nisa et al. distributions ($k > 1$).

The generating positive measure $\mu_{a,t}$ on \mathbb{R}^k of NST models is composed by the well-known probability measure $\nu_{a,t}$ of univariate positive s -stable distribution

generating function $\phi(t) = E[e^{t^T X}]$ for $t \in \mathbb{R}^k$ which was introduced by Feller (1971) as follows: $\phi(t) = \int_{\mathbb{R}^k} e^{t^T x} p(x) dx$, where $p(x)$ is the density function of X . For $a \in (0, 1)$ is the index parameter, $G(\cdot)$ is the classical gamma function, and I_A denotes the indicator function of any given event A that takes the value 1 if the event occurs and 0 otherwise. Parameter a can be extended into $a \in (-8, 2]$ (see Tweedie, 1984). For $a = 2$, we obtain the normal distribution with density $p(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$. For a k -dimensional NST random vector $X = (X_1, \dots, X_k)^T$, the generating function $\phi(t)$ is given by $\phi(t) = \prod_{j=1}^k \phi_j(t_j)$, (1.1) where X_1 is a univariate (non-negative) stable Tweedie variable and $(X_2, \dots, X_k)^T =: X_c$ given X_1 are $k-1$ real independent Gaussian variables with variance X_1 .

Normal-Poisson model is a special case of NST models; it is new and the only model which has a discrete component. Among NST models, normal-gamma is the only model which is also a member of simple quadratic natural exponential families (NEFs) of Casalis (1996); she called it "gamma-Gaussian" and it has been characterized by variance and generalized variance functions.

See Casalis (1996) or Kotz et al (2000, Chapter 54) for characterization by variance function, and Kokonendji and Masmoudi (2013) for characterization by generalized variance function which is the determinant of covariance matrix expressed in terms of the mean vector. In contrast to normal-gamma which is the same to gamma-Gaussian; normal-Poisson and Poisson-Gaussian (Kokonendji and Masmoudi, 2006; Koudou and Pommeret, 2002) are two completely different models. Indeed, for any value of $j \in \{1, \dots, k\}$, normal-Poisson 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. Poisson-Gaussian is a particular case of the simple quadratic NEFs with variance function $V_F(m) = \text{diag}(m_1, \dots, m_j, 1, \dots, 1)$. Characterizations of Multivariate Normal-Poisson Model 39 where $m = (m_1, \dots, m_k)^T$ is the mean vector, and its generalized variance function is $\det V_F(m) = m_1 \dots m_j$.

Some characterizations of Poisson-Gaussian j models have been done by several authors such as Letac (1989) for variance function, Kokonendji and Masmoudi (2006) for generalized variance function, and Koudou and Pommeret (2002) for convolution-stability. Also one can see Kokonendji and Seshadri (1996); Kokonendji and Pommeret (2007) for the generalized variance estimators of Poisson-Gaussian.

This normal-Poisson is also different from the purely discrete "Poisson-normal" model

of Steyn (1976), which can be defined as a multiple mixture of k independent Poisson distributions with parameters m_1, m_2, \dots, m_k and those parameters have a multivariate normal distribution. Three generalized variance estimators of normal Poisson model have been introduced (Kokonendji and Nisa, 2016).

In this paper we present the characterizations of multivariate normal-Poisson model by variance function and by generalized variance function which is connected to the Monge-Ampère equation (Gutiérrez, 2001). In Section 2 we present some properties of normal-Poisson model. We present the characterizations of normal-Poisson model by variance function in Section 3 and the characterization by generalized variance in Section 4.

2 Normal-Poisson model By introducing "power variance" parameter p defined by $(p-1)(1-a) = 1$ and equivalent to $p = p(a) = a - 2a - 1$ or $a = a(p) = p - 2p - 1$ (see Jorgensen, 1997, Chapter 4, for complete description of the power unit variance function of univariate stable Tweedie distributions), in the case of $a = -8$ or $p = p(-8) = 1$, Expression (1.1) will lead to k -variate normal-Poisson model.

Replacing $a(p)$ with $p(a)$ the generating measure of normal-Poisson model can be expressed as follows $\mu_t(dx) = \mu_1, t(dx) = \delta_1, t(dx_1) k_{j=2}^{\infty} 0, x_1(dx_j)$. (2.1) Then by (2.1), for a fixed power of convolution $t > 0$, denote $F_t = F(\mu_t)$ the multivariate NEF (Kotz et al, 2000, Chapter 54) of normal-Poisson with $\mu_t = \mu * t$, the NEF of a k -dimensional normal-Poisson random vector X is generated by $\mu_t(dx) = t x_1(x_1!) - 1 (2 p x_1) (k-1) / 2 \exp \{ \dots \} - t - 1 2 x_1 k_{j=2}^{\infty} x_2 j \dots 1 x_1 \{ N \setminus \{0\} \} dx_1(dx_1) k_{j=2}^{\infty} dx_j$. (2.2) 40 Nisa et al.

Since $t > 0$ then μ_t is known to be an infinitely divisible measure; see, e.g., Sato (1999). The cumulant function which is the logarithm of the Laplace transform of μ_t , i.e. $K_{\mu_t}(\theta) = \log \int \exp(\theta^T x) \mu_t(dx)$, is given by $K_{\mu_t}(\theta) = t \exp \{ \dots \} 1 + 1 2 k_{j=2}^{\infty} 2 j \dots \}$. (2.3) The function $K_{\mu_t}(\theta)$ is finite for all θ in the canonical domain $T(\mu_t) = \{ \theta \in \mathbb{R}^k; \theta^T \tilde{c}_1 := \theta^T 1 + 1 2 k_{j=2}^{\infty} 2 j < 0 \}$. (2.4) with $\theta = (\theta_1, \dots, \theta_k)^T$ and $\tilde{c}_1 := (1, \theta_2, \dots, \theta_k)^T$. (2.5) The probability distribution of normal-Poisson which is a member of NEF is given by $P(\theta; \mu_t)(dx) = \exp \{ \theta^T x - K_{\mu_t}(\theta) \} \mu_t(dx)$. From (2.3) we can calculate the first derivative of the cumulant function that produces a k -vector as the mean vector of F_t , and also its second derivative which is a $k \times k$ matrix that represents the covariance matrix. Using notations in (2.5) we obtain $K'_{\mu_t}(\theta) = K_{\mu_t}(\theta) \times \tilde{c}_1$ and $K''_{\mu_t}(\theta) = K_{\mu_t}(\theta) \times [\tilde{c}_1 \tilde{c}_1^T + I_{01k}]$, with $I_{01k} = \text{diag}_k(0, 1, \dots, 1)$.

The cumulant function presented in (2.3) and its derivatives are functions of the

We here state the first result as follows. Theorem 3.1. Let $k \in \{2, 3, \dots\}$ and $t > 0$. If an NEF F_t satisfies (2.6), then, up to a identity, F_t is normal-Poisson model. The proof is established by analytical calculations and using the well-known properties of NEFs described in Proposition 3.1 below. Proposition 3.1.

Let μ and $e\mu$ be two finite positive measures on \mathbb{R}^k such that $F = F(\mu)$, $eF = F(e\mu)$ and $m \in MF$. (i) If there exists $(d, c) \in \mathbb{R}^k \times \mathbb{R}$ such that $e\mu(dx) = \exp\{d \cdot x + c\} \mu(dx)$, then $F = eF: T e\mu = T \mu - d$ and $K e\mu(\cdot) = K \mu(\cdot + d) + c$; for $e m = m \in MF$, $V eF(e m) = V F(m)$ and $\det V eF(e m) = \det V F(m)$.

(ii) If $e\mu = f * \mu$ with $f(x) = Ax + b$, then: $T(e\mu) = A^T T(\mu)$ and $K e\mu(\cdot) = K \mu(A^T \cdot) + b^T \cdot$; for $e m = Am + b^T f(MF)$, $V eF(e m) = AV F(f^{-1}(e m))A^T$ and $\det V eF(e m) = (\det A)^2 \det V F(m)$. (iii) If $e\mu = \mu * t$ is the t -th convolution power of μ for $t > 0$, then, for $e m = t m \in t MF$, $V eF(e m) = t V F(t^{-1} e m)$ and $\det V eF(e m) = t^k \det V F(m)$. Proposition 3.1

shows that the generalized variance function of F , $\det V F(m)$, is invariant for any element of its generating measure (Part (i)) and for a linear transformation $f(x) = Ax + b$ such that $\det A = \pm 1$, in particular for a translation $x \mapsto x + b$ (Part (ii)). Sometimes we use terminology type to call an NEF F as a particular model up to a identity (Part (ii)) and convolution power (Part (iii)).

Characterizations of Multivariate Normal-Poisson Model 43 Proof. Without loss of generality, first we assume that $t = 1$ with a flashback to the identityability of Poisson component. Let $F = F(\mu)$ be an NEF satisfies (2.6) and (2.7) for $t = 1$. Using the blockwise inversion into $V F(m)$ in (2.6), one has: $[V F(m)]^{-1} = \begin{bmatrix} m_1^{-1} + 1 & m_1^{-1} m_3^T k^T \\ m_3^{-1} m_1^T + 1 & m_3^{-1} m_3^T k^T \end{bmatrix}$ (3.1) with $m_1 > 0$ and $m_3^T k^T := (m_2, \dots, m_k)^T \in \mathbb{R}^k$.

Since $m = K^T \mu(\cdot)$ and $V F(m) = K^T \mu(\cdot)$, then by writing μ in terms of m one gets $V F(m) = [\mu'(m)]^{-1}$ which implies $\mu'(m) = [V F(m)]^{-1} dm$. For $\mu \in T := \mu(MF)$ such that MF has the same elements as MF_t in (2.7), there exists a function $f: \mathbb{R}^k \rightarrow \mathbb{R}$ such that $\mu'(m) = [\mu^2 f(m)^T m^T]_{i,j=1,2,\dots,k}$. (3.2) Using (3.2) into (3.1) for getting the first information on Poisson component, we have $\mu^2 f(m)^T m^T = 1/m_1 + 1/m_3^T k^T = 2/m_2$ and then $\mu f(m)^T m^T = \log m_1 - 1/2 m_2^T k^T = 2/m_2 + f(m_2, \dots, m_k)$, (3.3) where $f: \mathbb{R}^{k-1} \rightarrow \mathbb{R}$ is an analytical function to be determined.

Note that since $m_1 > 0$ then $\log m_1$ and $1/(2 m_2^T k^T)$ in (3.3) are well-defined. Derivative of (3.3) with respect to m_j gives $\mu^2 f^T m^T m_j = -m_j m_2^T k^T + \mu f^T(m_2, \dots, m_k)^T m_j$. (3.4) 44 Nisa et al. Expression (3.4) is equal to the $(1, j)$ th element of

$[VF(m)] - 1$ in (3.1), that is $-m_j m^{2j-1} + f(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^j = -m_j m^{2j-1}$; therefore, $f(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^j = 0$ for all $j \in \{2, \tilde{y}, \tilde{y}, \tilde{y}, k\}$, this implies $f(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) = c_1$ (a real constant). Thus, (3.3) becomes $f(m^1) = \log m^1 - \frac{1}{2} m^{2j-1} k$ $j = 2, m^{2j} + c_1$ (3.5) and by integration with respect to m^1 , one gets $f(m) = m^1 \log m^1 - m^1 + \frac{1}{2} m^{1k} j = 2, m^{2j} + c_1 m^1 + h(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k)$, (3.6) where $h: R^{k-1} \rightarrow R$ is an analytical function to be determined. From now on, complete information of the model (i.e. normal components) begin to show itself.

The first and second derivatives of (3.6) with respect to m_j give, respectively, $f'(m) m^j = m^j m^1 + f'_j(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^j$, $f''_{jj}(m) m^{2j} = 1 m^1 + f''_{jj}(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^{2j}$, $f''_{jj}(m) m^{2j} = 1 m^1 + f''_{jj}(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^{2j}$. (3.7) and $f''_{jj}(m) m^{2j} = 1 m^1 + f''_{jj}(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^{2j}$. (3.8) Expression (3.8) is equal to the diagonal (j, j) th element of $[VF(m)] - 1$ in (3.1) for all $j \in \{2, \tilde{y}, \tilde{y}, \tilde{y}, k\}$, hence we have $1 m^1 + f''_{jj}(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^{2j} = 1 m^1$. Consequently, $f''_{jj}(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^{2j} = 0$ and $f'_j(m^2, \tilde{y}, \tilde{y}, \tilde{y}, m_k) m^j = c_j$ (a real constant) for all $j \in \{2, \tilde{y}, \tilde{y}, \tilde{y}, k\}$. Then, equation (3.7) becomes $f'(m) m^j = m^j m^1 + c_j j \in \{2, \tilde{y}, \tilde{y}, \tilde{y}, k\}$. (3.9) Characterizations of Multivariate Normal-Poisson Model 45 Using equation (3.5) and (3.9) one obtains $f(m) = \log m^1 - \frac{1}{2} m^{2j-1} k j = 2, m^{2j}, m^2 m^1, \tilde{y}, \tilde{y}, \tilde{y}, m_k m^1 + (c_1, \tilde{y}, \tilde{y}, \tilde{y}, c_k) m^1$ or $f(m) = \log m^1 - \frac{1}{2} m^{2j-1} k j = 2, m^{2j} + c_1 j = m^j m^1 + c_j, j = 2, \tilde{y}, \tilde{y}, \tilde{y}, k$. (3.10) From (3.10), each j belongs to R for $j \in \{1, 2, \tilde{y}, \tilde{y}, \tilde{y}, k\}$ because $m^1 > 0$ and $m^j \in R$ for $j \in \{2, \tilde{y}, \tilde{y}, \tilde{y}, k\}$.

Thus, one has $T(MF) = T \in R^k$ and also $m^1 = \exp \left(\frac{1}{2} (1 - c_1) + \frac{1}{2} k j = 2, (j - c_j)^2 \right)$, (3.11) $m^j = (j - c_j) \exp \left(\frac{1}{2} (1 - c_1) + \frac{1}{2} k j = 2, (j - c_j)^2 \right)$. (3.12) Since $m = K \mu(\cdot)$, then using (3.11) one can obtain $K \mu(\cdot)$ as follow: $K \mu(\cdot) = K' \mu(\cdot) \exp \left(\frac{1}{2} (1 - c_1) + \frac{1}{2} k j = 2, (j - c_j)^2 \right) + g(2, \tilde{y}, \tilde{y}, \tilde{y}, k)$, (3.13) where $g: R^{k-1} \rightarrow R$ is an analytical function to be determined. Again, derivative of (3.13) with respect to j produces $K \mu(\cdot) j = (j - c_j) \exp \left(\frac{1}{2} (1 - c_1) + \frac{1}{2} k j = 2, (j - c_j)^2 \right) + g_j(2, \tilde{y}, \tilde{y}, \tilde{y}, k) j$ which is equal to (3.12); then, one gets $g_j(2, \tilde{y}, \tilde{y}, \tilde{y}, k) = 0$ for all $j \in \{2, \tilde{y}, \tilde{y}, \tilde{y}, k\}$ implying $g(2, \tilde{y}, \tilde{y}, \tilde{y}, k) = C$ (a real constant).

Finally, it ensues from it that we have $K \mu(\cdot) = \exp \left(\frac{1}{2} (1 - c_1) + \frac{1}{2} k j = 2, (j - c_j)^2 \right) + C$. 46 Nisa et al. By Proposition 3.1 one can see that, up to a π -nity, this $K \mu$ is a normal-Poisson cumulant function as given in (2.3) with $t = 1$ on its corresponding support (2.4). Theorem 3.1 is therefore proven by using the analytical property of $K \mu$.

_4 Characterization by generalized variance function Before stating our next result, let

member of Equation (4.2) can be written as $\det K'' \mu(\cdot) = \det [S + \sum_{j=2}^k x_j \exp(\sum_{i=1}^{j-1} x_i) (dx)] = \det [S + L''(\cdot)]$. (4.3) For $S = \{i_1, i_2, \dots, i_j\}$, with $1 \leq i_1 < i_2 < \dots < i_j \leq k$, a non-empty subset of $\{1, 2, \dots, k\}$, and $tS: \mathbb{R}^k \rightarrow \mathbb{R}^j$ the map defined by $tS(x) = (x_{i_1}, x_{i_2}, \dots, x_{i_j})$, we define S the image measure of $H_j(dx_1, \dots, dx_j) = 1/j! (\det [tS(x_1), \dots, tS(x_j)])^2 (dx_1, \dots, dx_j)$ by $\pi_j: (\mathbb{R}^k)^j \rightarrow \mathbb{R}^k$, $(x_1, \dots, x_j) \mapsto x_1 + x_2 + \dots + x_j$. By Proposition 4.2 and Expression (4.3) the modified Lévy measure $\nu(\mu)$ in (2.1) can be expressed as $\nu(\mu) = (\det S)^{-1} d_0 + \sum_{S \neq \emptyset, S \subseteq \{1, 2, \dots, k\}} (\det S')^{-1} S$, (4.4) 48 Nisa et al.

where ν is a diagonal representation of S in an orthonormal basis $e = (e_i)_{i=1, \dots, k}$ (see Hassairi, 1999, page 384). Since S is the Brownian part, then it corresponds to the $k-1$ normal components from the right member of (4.2); that implies $r = \text{rank}(S) = k-1$ and $\det S = 0$. Therefore $\det \nu = 0$ with $\nu = \text{diag}(\nu_1, \nu_2, \dots, \nu_k)$ such that $\nu_1 = 0$ and $\nu_j > 0$ for all $j \in \{2, \dots, k\}$. For all non-empty subsets S of $\{1, 2, \dots, k\}$ there exist real numbers $a_S > 0$ such that $(\det \nu S')^{-1} S = \sum_{i=1}^k a_S \nu_i \delta_{e_i} < S \nu_i \delta_{e_i}$. $\sum_{i=1}^k a_S \nu_i \delta_{e_i} = a_S [\delta_{e_1} * \gamma_N(0, 1)(e_{c1})] * k$, (4.5) where $e_{c1} = (e_2, \dots, e_k)$ denotes the induced orthonormal basis of e without component e_1 ; i.e. $k-1$ is the dimension of e_{c1} .

With respect to Kokonendji and Masmoudi (2006, Lemma 7) for making precise the measure ν of (4.5), it is easy to see that $S_0 = \{1\}$ is a singleton (i.e. set with exactly one element) such that, for $x = x_1 e_1 + \dots + x_k e_k$, $x_2 \nu(dx) = \beta da_{e_1}$, with $\beta > 0$ and $a, 0$. Consequently, we have the following complementary set $S'_0 = \{1, 2, \dots, k\} \setminus \{1\}$. So, from (4.5) we have k th power of convolution of only one Poisson at the first component e_1 and $(k-1)$ -variate standard normal.

That means $K'' \mu(\cdot) = K \mu(\cdot) [\sum_{i=1}^k \nu_i \delta_{e_i} + I_{\{0\}}]$, with notations of (2.5). Let $B(\cdot) = \exp(\sum_{j=2}^k x_j = 2 \sum_{j=2}^k x_j)$ from (4.2). Since we check that $\sum_{i=1}^k (K \mu - B)(\cdot) \nu_i = 0$ for all $i = 1, \dots, k$, Proposition 4.3 allows that $(K \mu - B)(\cdot)$ is an affine function on \mathbb{R}^k and therefore $K \mu(\cdot) = \exp(\sum_{j=2}^k x_j = 2 \sum_{j=2}^k x_j) \sum_{i=1}^k \nu_i \delta_{e_i} + u \cdot + b$, for $(u, b) \in \mathbb{R}^k \times \mathbb{R}$. Hence $F = F(\mu)$ is of normal-Poisson type with $t = 1$. This completes the proof of the theorem. A reformulation of Theorem 4.1, by changing the canonical parameterization into the mean parameterization, is stated in the following theorem without proof. Theorem 4.2.

Let $F_t = F(\mu_t)$ be an infinitely divisible NEF on \mathbb{R}^k such that 1. $M F_t = \{m \in \mathbb{R}^k; m_1 > 0 \text{ and } m_j \in \mathbb{R} \text{ with } j = 2, \dots, k\}$, and Characterizations of Multivariate Normal-Poisson Model 49 2. $\det V F_t(m) = m_k^{-1}$. Then F_t is of normal-Poisson type. Theorem 4.1 can be viewed as the solution to a particular Monge-Ampère equation (4.1). Whereas Theorem 4.2 is interesting for generalized variance estimation of the

model.

5 Conclusion In this paper we described some properties of normal-Poisson model. Then we showed that the characterization of normal-Poisson model by variance function was obtained through analytical calculations and using some properties of NEF. Also, the characterization of normal Poisson model by generalized variance which is the solution to a specific Monge-Ampère equation: $\det K''(\mu) = \exp(k \times \mu^c)$ on R^k can be solved using the infinite divisibility property of normal-Poisson.

Acknowledgment The authors would like to thank the Editor and Referees for their valuable and constructive comments that led to a considerably improved manuscript. **References** Bar-Lev, S., Bschouty, D., Enis, P., Letac, G., Lu, I. and Richard, D. (1994), The diagonal multivariate natural exponential families and their classification. *Journal of Theoretical Probability*, 7 (4), 883-929. Boubacar Maïnassara, Y. and Kokonendji, C. C.

(2014), Normal stable Tweedie models and power-generalized variance function of only one component, *TEST*, 23 (3), 585-606. Casalis, M. (1996), The $2d + 4$ simple quadratic natural exponential families on R^d . *The Annals of Statistics*, 24 (4), 1828-1854. Feller, W. (1971), *An Introduction to Probability Theory and its Applications Vol. II* Second edition. New York: Wiley. 50 Nisa et al. Gikhman, I. I. and Skorokhod, A.V.

(2004), *The Theory of Stochastic Processes II*. New York: Springer. Gutiérrez, C. E. (2001), *The Monge-Ampère Equation*. Boston: Birkhäuser. Hassairi, A. (1999), Generalized variance and exponential families. *The Annals of Statistics*, 27 (1), 374-385. Jorgensen, B. (1997). *The Theory of Dispersion Models*. London: Chapman & Hall. Kokonendji, C.C. and Masmoudi, A. (2006), A characterization of Poisson-Gaussian families by generalized variance. *Bernoulli*, 12 (2), 371-379.

Kokonendji, C.C. and Masmoudi, A. (2013), On the Monge-Ampère equation for characterizing gamma-Gaussian model. *Statistics and Probability Letters*, 83 (7), 1692-1698. Kokonendji, C. C. and Nisa, K. (2016). Generalized variance estimations of normal Poisson models. In *Forging Connections between Computational Mathematics and Computational Geometry, Springer Proceeding in Mathematics and Statistics* 124, Eds. K. Chen and A.

Ravindran. Switzerland: Springer, pp. 247-260. Kokonendji, C.C. and Pommeret, D. (2007), Comparing UMVU and ML estimators of the generalized variance for natural exponential families. *Statistics*, 41 (6), 547-558. Kokonendji, C.C. and Seshadri, V. (1996), On the determinant of the second derivative of a Laplace transform. *The Annals of Statistics*, 24 (4), 1813-1827. Kotz, S.,

Balakrishnan, N. and Johnson, N.L. (2000), Continuous Multivariate Distributions Vol.1: Models and Applications. 2nd edition . New York: Wiley. Koudou, A.E. and Pommeret, D. (2002), A characterization of Poisson-Gaussian families by convolution-stability. Journal of Multivariate Analysis , 81 , 120-127. Letac, G. (1989), Le problème de la classification des familles exponentielles naturelles sur \mathbb{R}^d ayant une fonction variance quadratique.

In Probability Measures on Groups IX - Lecture Notes in Mathematics . 1306 , Ed. H. Heyer. Berlin: Springer 194-215. Muir, T. (1960), A Treatise on the Theory of Determinants . New York: Dover. Sato, K. (1999), Lévy Processes and Infinitely Divisible Distributions . Cambridge: Cambridge University Press. Characterizations of Multivariate Normal-Poisson Model 51 Steyn, H.S. (1976), On the multivariate Poisson normal distribution.

Journal of the American Statistical Association , 71 (353), 233-236. Tweedie, M.C.K. (1984), An index which distinguishes between some important exponential families. In: Statistics: Applications New Directions, Proceedings of the Indian Statistical Golden Jubilee International Conference , Eds. J. K. Ghosh, J. Roy.

Calcutta: Indian Statistical Institute, 579-604.

INTERNET SOURCES:

<1% - <http://artikel.ubl.ac.id/index.php/LIT/article/view/1174>
 1% - <http://jirss.irstat.ir/article-1-307-en.html>
 <1% - http://personal.psu.edu/abs12/stat504/online/07_poisson/07_poisson_print.htm
 <1% -
https://www.researchgate.net/publication/2903743_Monotone_Runs_of_Uniformly_Distributed_Integer_Random_Variables_A_Probabilistic_Analysis
 <1% -
https://www.researchgate.net/publication/282399535_Simulations_of_full_multivariate_Tweedie_with_flexible_dependence_structure
 <1% -
https://mafiadoc.com/uncertainty-models-for-knowledge-based-systems_5c1bdb2c097c479a6b8b4783.html
 <1% - <https://www.ufdc.ufl.edu/AA00038040/00001>
 <1% - <http://www.math.unl.edu/~tmarley1/math314h/quizzes/hw1soln.pdf>
 <1% -
<http://www.m-hikari.com/ams/ams-2016/ams-61-64-2016/p/nisaAMS61-64-2016.pdf>
 <1% - http://www.maths.qmul.ac.uk/~bb/MS_Lectures_5and6.pdf

<1% - <http://www.columbia.edu/~ks20/4703-Sigman/4703-07-Notes-0.pdf>

<1% - https://www.researchgate.net/publication/272006326_On_normal_stable_Tweedie_models_and_power-generalized_variance_functions_of_only_one_component

<1% - [https://www.researchgate.net/publication/223218426_Orthogonality_of_the_Sheffer_system_associated_to_a_Levy process](https://www.researchgate.net/publication/223218426_Orthogonality_of_the_Sheffer_system_associated_to_a_Levy_process)

<1% - https://ceprofs.civil.tamu.edu/dlord/Papers/Zha_et_al_PIG.pdf

2% - https://www.researchgate.net/publication/38322468_A_characterization_of_Poisson-Gaussian_families_by_generalized_variance

<1% - <https://www.sciencedirect.com/topics/engineering/error-covariance-matrix>

<1% - <https://www.scribd.com/document/388664259/Hassani-3>

<1% - https://mafiadoc.com/mathematical-methods-for-physicists-a-concise-introduction_59bcccb51723dd7e3cb0999b.html

1% - https://www.researchgate.net/publication/38348721_On_the_determinant_of_the_second_derivative_of_a_Laplace_transform

<1% - <http://lmb.univ-fcomte.fr/IMG/pdf/arnpjeas2017nketal.pdf>

<1% - https://www.researchgate.net/publication/227726896_Occurrence_and_quantity_of_precipitation_can_be_modeled_simultaneously

<1% - https://mafiadoc.com/lists-decisions-and-graphs_59c08ce61723ddbda568399b.html

<1% - https://www.math.hmc.edu/calculus/tutorials/int_by_parts/int_by_parts.pdf

<1% - https://www.researchgate.net/publication/29641897_Generalized_variance_estimators_in_the_multivariate_gamma_models

<1% - <https://www.sciencedirect.com/science/article/pii/S0925231214002392>

<1% - https://www.researchgate.net/publication/318984339_Approximation_with_certain_genuine_hybrid_operators

<1% - <https://wenku.baidu.com/view/617426e80975f46527d3e1e3.html>

<1% - <http://www.naturalspublishing.com/files/published/a6w4k153vo7wt2.pdf>

<1% - https://www.math.uh.edu/~rohop/spring_13/Chapter2.pdf

<1% - <http://aleph.math.louisville.edu/teaching/2011SP-311/ps07-110425-solutions.pdf>

<1% - <https://recruit-mizukyo.jp/characteristic/early-program/pdf/shokiprg.pdf>

<1% - <https://www.ijert.org/research/a-modified-decomposition-covariance-matrix-estimation>

-for-undirected-gaussian-graphical-model-IJERTV3IS060820.pdf

<1% -

https://www.academia.edu/14726268/Estimating_the_variance_in_nonparametric_regression-what_is_a_reasonable_choice

<1% - https://www.casact.org/library/astin/vol24no2/265.pdf?origin=publication_detail

<1% -

https://www.researchgate.net/post/What_is_a_correct_practical_use_of_the_confidence_intervals

<1% -

https://mafiadoc.com/an-r-package-for-gaussian-poisson-and-binomial-random-arxiv_59dae30b1723dd378df4e1d2.html

1% - <https://www.sciencedirect.com/science/article/pii/S0167715213001065>

<1% -

https://archive.org/stream/in.ernet.dli.2015.141642/2015.141642.An-Introduction-To-Probability-Theory-And-Mathematical-Statistics_djvu.txt

<1% - https://mafiadoc.com/bayesian-network_5c14ab34097c4719208b45b8.html

<1% -

https://mafiadoc.com/a-tutorial-on-non-central-wishart-distributions_5b94687e097c47940a8b45e2.html

<1% -

https://research.cbs.dk/files/57133575/anders_r_nn_nielsen_excursion_sets_of_ininitely_divisible_random_fields_acceptedversion.pdf

<1% - <https://math.vanderbilt.edu/peters10/teaching/Fall2009/hwsolution9.pdf>

<1% - <https://arxiv.org/pdf/0902.0333.pdf>

<1% -

<http://appliedmaths.sun.ac.za/~htouchette/archive/talks/touchette-princeton2015.pdf>

<1% - <https://epdf.pub/modeling-infectious-diseases-in-humans-and-animals.html>

<1% - <https://ar.scribd.com/document/46978540/Ying>

<1% - <https://wenku.baidu.com/view/5d844e5377232f60ddcca1f5.html>

<1% -

<https://ipfs.io/ipfs/QmXoypizjW3WknFiJnKLwHCnL72vedxjQkDDP1mXWo6uco/wiki/Determinant.html>

<1% - <https://www.scribd.com/document/230717307/TOMACS-Paper-Revision>

<1% - <https://www.scribd.com/document/256113891/How-to-Build-a-Brain>

<1% -

<https://epdf.pub/principles-of-quantum-mechanics-as-applied-to-chemistry-and-chemical-physics.html>

<1% - <http://math.mit.edu/~sheffield/2019600/spring2019-600ProblemSet7.pdf>

<1% - https://issuu.com/urbanform/docs/volume_2_part_1_isuf_rome_2015

<1% - <http://www.columbia.edu/~kr2248/4109/chapter4.pdf>

<1% - http://math.stanford.edu/~ksound/Math171S10/Hw8Sol_171.pdf
<1% - <https://www.gmx.net/>
<1% - <https://www.fq.math.ca/Scanned/14-1/stringall.pdf>
<1% -
https://www.ifm.liu.se/edu/coursescms/tfyy68/lectures/Fo8_Rotation-av-fasta-kroppar-Vinkelacceleration-Energi-hos-roterande-kroppar-troghetsmoment-2017.pdf
<1% - http://web.stanford.edu/~tonyfeng/Zhu_all.pdf
<1% - <https://www.sciencedirect.com/science/article/pii/S0167715210000441>
<1% - <https://wenku.baidu.com/view/5b839439580216fc700afd03.html>
<1% -
https://www.researchgate.net/publication/292606455_The_Beta_Exponentiated_Gumbel_Distribution
<1% -
https://www.academia.edu/13973209/Blockwise_SVD_with_error_in_the_operator_and_application_to_blind_deconvolution
<1% -
<http://www.kfupm.edu.sa/sites/phys101/pdf/exams/exam2/Physics101-Second%20Major-152-Solution.pdf>
<1% - http://www.city.kashiwara.osaka.jp/_files/00101974/H27juurannzu05.pdf
<1% - <https://www.le.ac.uk/users/dsgp1/COURSES/LEISTATS/Lecture8.pdf>
<1% - <https://www.math.ucdavis.edu/~hunter/m125b/ch2.pdf>
<1% - <http://www.stat.ucla.edu/~rosario/classes/081/100a-2a/100aHW5Soln.pdf>
<1% -
https://mafiadoc.com/multispace-multistructure-neutrosophic-_5982205e1723ddef56dc042.html
<1% - <https://fam.tuwien.ac.at/~schmock/notes/Yamada-Watanabe.pdf>
<1% -
https://www.academia.edu/7347786/Boundary_element_method_analysis_for_the_transient_conduction_convection_in_2-D_with_spatially_variable_convective_velocity
<1% -
https://www.researchgate.net/publication/7798439_Regional_scale_evidence_for_improvements_in_surface_water_chemistry_1990-2001
<1% -
<https://ocw.mit.edu/courses/engineering-systems-division/esd-86-models-data-and-inference-for-socio-technical-systems-spring-2007/lecture-notes/lec13.pdf>
<1% - https://en.wikibooks.org/wiki/MATLAB_Programming/Print_Version
<1% -
https://www.academia.edu/11139023/Macroscopic_turbulence_modeling_for_incompressible_flow_through_undefromable_porous_media
<1% -

http://mpjmath.weebly.com/uploads/8/4/3/7/8437491/4th_six_weeks_study_guide-key-preap.pdf
<1% - https://issuu.com/hanifird/docs/zienkiewicz_o.c._taylor_r.l._vol._
<1% - https://www.academia.edu/1739173/groundwater_course
<1% -
https://mafiadoc.com/a-course-in-statistical-theory_5bb097d0097c4755428b4638.html
<1% - <https://docsplayer.net/22277132-2.html>
<1% -
<http://staff.katyisd.org/sites/thsanatomyphysiology/PublishingImages/Pages/documents/Part%201%20KEY.pdf>
<1% -
<http://www1.aucegypt.edu/faculty/mharafa/MENG%20475/Forced%20Vibration.pdf>
<1% -
http://people.math.aau.dk/~cornean/analyse2_F15/powerseries-are-analytic-16-02-2015.pdf
<1% - <http://dept.stat.lsa.umich.edu/~ionides/620/hw/hw1sol.pdf>
<1% - <https://epdf.pub/handbook-of-hydrocolloids.html>
<1% -
https://www.piwheels.org/simple/numpy/numpy-1.8.2-cp35-cp35m-linux_armv7l.whl
<1% - <https://epdf.pub/sensors-and-signal-conditioning-2nd-edition.html>
<1% - <https://fresh2refresh.com/c-programming/c-constants/>
<1% -
https://www.researchgate.net/publication/228988636_An_Inventory_of_Adaptation_to_climate_change_in_the_UK_challenges_and_findings
<1% - <https://www.sciencedirect.com/topics/mathematics/cayley-hamilton-theorem>
<1% - <https://arxiv.org/pdf/1905.00383>
<1% - <http://downloads.hindawi.com/journals/jfs/2015/620251.pdf>
<1% - <https://www.sciencedirect.com/science/article/pii/S0022039613004221>
<1% - <https://www.sciencedirect.com/topics/engineering/marangoni-force>
<1% - <https://www.iep.utm.edu/gettier/>
<1% - <https://arxiv.org/pdf/1903.09742>
<1% -
https://www.researchgate.net/publication/236604215_Dose-Response_Analysis_in_the_Joint_Action_of_Two_Effectors_A_New_Approach_to_Simulation_Identification_and_Modeling_of_Some_Basic_Interactions
<1% - <https://azdoc.pl/a-wesley-design-patterns-explained.html>
<1% - <https://wenku.baidu.com/view/e6da488b84868762caaed5a2.html>
<1% -
https://mafiadoc.com/control-and-optimization_59e033481723dd185692f4e8.html
<1% -

https://slidelegend.com/a-beginners-guide-to-dual-quaternions-wscg_59b48db61723dd6c7341f30b.html
<1% - <https://www.sciencedirect.com/science/article/pii/S0304414908001464>
<1% - <http://digitalhumanities.org/companion/index/lazy/default/blackwell/9781405148641/9781405148641.xml.lazy>
<1% - https://www.academia.edu/13655171/Algorithms_for_residues_and_Lojasiewicz_exponents
<1% - <https://www.sciencedirect.com/science/article/pii/S0022247X19306043>
<1% - <http://www.stat.cmu.edu/~aramdas/papers/AAAI-power-App.pdf>
<1% - https://mafiadoc.com/handbook-of-industrial-automation_598cb0471723ddcb690d826f.html
<1% - https://lasco-www.nrl.navy.mil/content/retrieve/polarize/2013_02/vig/c2/C2-PB-20130228_0257.fts
<1% - https://www.researchgate.net/publication/228858378_Controlled_Wild_Algebras
<1% - <https://epdf.pub/finite-element-methoda37abf9ed5ffdbf881c15c89c08c840657982.html>
<1% - https://www.academia.edu/645788/Consistency_of_the_Averaged_Cross-Periodogram_In_Long_Memory_Series
<1% - <http://oai.repec.org/?verb=ListRecords&set=RePEc:eee:jmvana&metadataPrefix=amf>
<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2858466/>
<1% - http://docshare.tips/journal-of-computational-and-applied-mathematics-vol-127-issues-1-2-s_5a551ee108bbc571379998fd.html
<1% - <https://www.scribd.com/document/393582161/8060-Vector-Analysis-5>
<1% - <https://epdf.pub/orbital-interaction-theory-of-organic-chemistry.html>
<1% - <https://yousource.it.jyu.fi/vrp-diss/vrp-diss/commit/8db446a79279d6f2dea94a6f1f3bfec050eb7e43?format=diff>
<1% - https://www.academia.edu/2992389/The_Nature_of_the_Hydrogen_Bond_in_DNA_Base_Pairs_The_Role_of_Charge_Transfer_and_Resonance_Assistance
<1% - https://mafiadoc.com/solutions-for-quantum-mechanics-textbook-problems_59bf81d01723dde301925982.html

<1% -

https://www.researchgate.net/publication/236858663_Sharp_regularity_of_linearization_for_C-1C-1_hyperbolic_diffeomorphisms

<1% -

https://www.researchgate.net/publication/2337099_Basis_Problem_For_Turbulent_Actions_II_c_0-equalities

<1% - <https://arxiv.org/pdf/quant-ph/0604004.pdf>

<1% - <https://cosmosis101.blogspot.com/feeds/posts/default>

<1% -

https://www.researchgate.net/publication/50401947_Composite_parameterization_and_Haar_measure_for_all_unitary_and_specialunitary_groups

<1% - <https://epdf.pub/random-and-vector-measures.html>

<1% -

https://static.springer.com/sgw/documents/1445494/application/vnd.ms-excel/Yellowsale_2014_US_titlelist.xls

<1% -

https://www.researchgate.net/publication/220056343_The_-Birnbaum-Saunders_distribution_An_improved_distribution_for_fatigue_life_modeling

<1% -

<https://www.docme.ru/doc/1342644/757.-springer-texts-in-statistics--christian-p.-robert---...>

<1% -

<https://www.cambridge.org/core/services/aop-cambridge-core/content/view/E6636AC79C657F70713AED844A696C9C/S0140525X01004149a.pdf/div-class-title-money-lies-and-replicability-on-the-need-for-empirically>

<1% - <https://arxiv.org/pdf/1906.09485>

<1% - <https://projecteuclid.org/euclid.ss/1219339107>

1% - <https://projecteuclid.org/euclid.bj/1145993979>

<1% - <https://www.facebook.com/london.chapman.9>

1% - <http://lmb.univ-fcomte.fr/Publications-310>

<1% -

https://archive.org/stream/springer_10.1007-10721064/10.1007-10721064_djvu.txt

<1% - https://link.springer.com/chapter/10.5176%2F2251-1911_CMCGS14.29_21

<1% -

https://www.researchgate.net/publication/1905089_Dispersion_Models_for_Extremes

1% - <https://www.statisticshowto.datasciencecentral.com/tweedie-distribution/>



Plagiarism Checker X Originality Report

Similarity Found: 17%

Date: Selasa, Agustus 20, 2019

Statistics: 986 words Plagiarized / 5643 Total words

Remarks: Low Plagiarism Detected - Your Document needs Optional Improvement.

Generalized Variance Estimations of Normal-Poisson Models Célestin C. Kokonendji and Khoirin Nisa Abstract This chapter presents three estimations of generalized variance (i.e., determinant of covariance matrix) of normal-Poisson models: maximum likelihood (ML) estimator, uniformly minimum variance unbiased (UMVU) estimator, and Bayesian estimator. First, the definition and some properties of normal-Poisson models are established.

Then ML, UMVU, and Bayesian estimators for generalized variance are derived. Finally, a simulation study is carried out to assess the performance of the estimators based on their mean square error (MSE). Keywords Covariance matrix • Determinant • Normal stable Tweedie • Maximum likelihood • UMVU • Bayesian estimator Introduction In multivariate analysis, generalized variance (i.e., determinant of covariance matrix) has important roles in the descriptive analysis and inferences.

It is the measure of dispersion within multivariate data which explains the variability and the spread of observations. Its estimation usually based on the determinant of the sample covariance matrix. Many studies related to the generalized variance estimation have been done by some researchers; see, e.g., [1–3] under normality and non-normality hypotheses.

A normal-Poisson model is composed by distributions of random vector $X = (X_1, X_2, \dots, X_k)^T$ with $k > 1$, where X_j is a univariate Poisson variable, and $(X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_k)$ given X_j are $k-1$ real independent Gaussian variables with variance X_j . It is a particular part of normal stable Tweedie (NST) models [4] with $p \geq 1$ where p is the power variance parameter of distributions within the Tweedie family.

This model was introduced in [4] for the particular case of normal-Poisson with $j \in \mathbb{N}$. Also, normal-Poisson is the only NST model which has a discrete component, and it is correlated to the continuous normal parts. C.C. Kokonendji (✉) • K. Nisa Laboratoire de Mathématiques de Besançon, University of Franche-Comté, Besançon, France e-mail: celestin.kokonendji@univ-fcomte.fr ; khoirin.nisa@univ-fcomte.fr © Springer International Publishing Switzerland 2016 K. Chen (ed.), Forging Connections between Computational Mathematics and Computational Geometry, Springer Proceedings in Mathematics & Statistics 124, DOI 10.5176/2251-1911_CMC GS14.29_21 247 celestin.kokonendji@univ-fcomte.fr 248 C.C. Kokonendji and K.

Nisa In literature, there is also a model known as Poisson-Gaussian [5–7] which is completely different from normal-Poisson. For any value of j , a normal-Poisson j model has only one Poisson component and $k-1$ normal (Gaussian) components, while a Poisson-Gaussian j model has j Poisson components and $k-j$ Gaussian components.

Poisson-Gaussian is also a particular case of simple quadratic natural exponential family (NEF) [5] with variance function $VF(\mu) = \text{Diag}(k(\mu_1, \dots, \mu_j, 1, \dots, 1))$, where $\mu \in (\mu_1, \dots, \mu_k)$ is the mean vector and its generalized variance function is $\det VF(\mu) = \mu_1, \dots, \mu_j$. The estimations of generalized variance of Poisson-Gaussian can be seen in [8, 9].

Motivated by generalized variance estimations of Poisson-Gaussian, we present our study on multivariate normal-Poisson models and the estimations of their generalized variance using ML, UMVU, and Bayesian estimators. Normal-Poisson Models In this section, we establish the definition of normal-Poisson j models as generalization of normal-Poisson 1 model which was introduced in [4], and then we give some properties. Definition 2.1

For a k -dimensional normal-Poisson random vector $X \in (X_1, X_2, \dots, X_k)^T$ with $k > 1$, it must hold that 1. X_j follows a univariate Poisson distribution. 2. $X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_k$ are independent normal variables with mean 0 and variance X_j , i.e., $X_c \mid X_j \sim N(0, X_j)$. 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 sample.

In this case, the corresponding normal components are degenerated as $\bullet 0$ which makes their values become 0s. The NEF $F_t \in \mathcal{F}(\bullet, t)$ of a k -dimensional normal-Poisson random vector X is generated by $\prod_{j=1}^k \frac{t^{x_j}}{D(t)^{x_j}} \frac{e^{-t}}{x_j!} \frac{1}{\Gamma(x_j)} \exp\left(-\frac{t}{2x_j}\right) \exp\left(-\frac{t}{2x_j}\right) \frac{1}{\Gamma(x_j)} \exp\left(-\frac{t}{2x_j}\right)$; for a fixed power of convolution $t > 0$, where I_A is the indicator function of the set A and δ_{x_j} is the Dirac measure at x_j . Since $t > 0$, then $\bullet t: D \bullet \ast t$ is an infinitely divisible measure.

The cumulant function which is the log of the Laplace transform of Π_t , i.e., $K_{-t}^{\text{TM}} / D \int \log Z R k \exp^{\text{TM}} T x_{-t} \cdot dx /$, is given by celestin.kokonendji@univ-fcomte.fr
 Generalized Variance Estimations of Normal-Poisson Models 249 $K_{-t}^{\text{TM}} / D t \exp 0_{-j}$
 $C 1 2 X \backslash \alpha j_{-2} \backslash 1$: (1) The function $K_{-t}^{\text{TM}} / \ln (1)$ is finite for all TM in the canonical domain: $,_{-t} / D 8^{\text{TM}} 2 R k I^{\text{TM}} T Q^{\text{TM}} c W D_{-j} C X \backslash \alpha j_{-2} \backslash = 2 < 0 9$ with $\text{TM} D$.

$_{-1};_{-};_{-} k / T$ and $Q^{\text{TM}} c W D_{-1};:::;_{-j}_{-1};_{-j} D 1;_{-j} C 1;:::;_{-k}_{-T}$: (2) The probability distribution of normal-Poisson j is $P_{-t}^{\text{TM}} I t / \cdot dx / D \exp^{\text{TM}} T x_{-t} K_{-t}^{\text{TM}} /_{-t} \cdot dx /$ which is a member of NEF $F_{-t} / D f P_{-t}^{\text{TM}} I t / I^{\text{TM}} 2,_{-t} / g$. From (1), we can calculate the first derivative of the cumulant function that produces a k -vector as the mean vector of F_{-t} and also its second derivative which is a $k_{-} k$ matrix that represents the covariance matrix. Using notations in (2), we obtain $K 0 t^{\text{TM}} / D K_{-t}^{\text{TM}} /_{-} Q^{\text{TM}} c$ and $K 00 t^{\text{TM}} / D K_{-t}^{\text{TM}}$

$/ j Q^{\text{TM}} c Q^{\text{TM}} c T C I 0 j k k$ with $I 0 j k D \text{Diag } k 1;:::; 1; 0 j; 1;:::; 1_{-}$. The cumulant function presented in (1) and its derivatives are functions of the canonical parameter TM . For practical calculation, we need to use the mean parameterization: $P_{-t} m I F t / W D P_{-t}^{\text{TM}} m / I_{-t} /$ with $\text{TM}(m)$ is the solution in TM of equation $m D K 0 t^{\text{TM}}$

$/$: The variance function of a normal-Poisson j model which is the variance-covariance matrix in term of mean parameterization is obtained through the second derivative of the cumulant function, i.e., $V F t_{-t} m / D K 00 t^{\text{TM}} m / \bullet$: Then we have $V F t_{-t} m / D 1 m j m m T C \text{Diag } k m j;:::; m j; 0 j; m j;:::; m j_{-}$ (3) with $m j > 0$ and $m \backslash 2 R; \backslash \alpha j$.

For $j D 1$, the covariance matrix of X can be expressed as below
 celestin.kokonendji@univ-fcomte.fr 250 C.C. Kokonendji and K. Nisa $V F t_{-t} m / D 2 m 1 ? m 2::: m j::: m k_{-}$?
 $m 2 ? m_{-1} 1 m 2 C m 1::: m_{-1} 1 m 2 m j::: m_{-1} 1 m 2 m k::: m j ? m_{-1} 1 m j m 2::: m_{-1} 1 m 2 C m 1::: m_{-1} 1 m j m k::: m k ? m_{-1} 1 m k m 2::: m_{-1} 1 m k m j::: m_{-1} 1 m 2 C m 1$
 3: Indeed, for the covariance matrix above, one can use the following particular Schur representation of the determinant $\det_{-} a T a A_{-} D \det A_{-} 1 a a T_{-}$ (4) with the non-null scalar $D m 1$, the vector $a D_{-} m 2;_{-}; m k / T$, and the $k_{-} 1 /_{-} k_{-} 1 /$ matrix $A D_{-} 1 a a T C m 1 I k_{-} 1$; where $I j D \text{Diag } j$.

$1;_{-};_{-}; 1 /$ is the $j_{-} j$ unit matrix. Consequently, the determinant of the covariance matrix $V F t_{-t} m /$ for $j D 1$ is $\det V F t_{-t} m / D m k$. Then, it is trivial to show that for $j \geq 1, :::, k g$, the generalized variance of normal-Poisson j model is given by $\det V F t_{-t} m / D m k$ (5) with $m j > 0; m \backslash 2 R; \backslash \alpha j$: (5) expresses that the generalized variance of normal-Poisson models depends mainly on the mean of the Poisson component (and the dimension space $k > 1$).

Among NST models, normal-gamma which is also known as gamma-Gaussian is the only model that has been characterized completely; see [5] or [10] for characterization by variance function and [11] for characterization by generalized variance function. For normal-Poisson models, here we give our result regarding to characterization by variance function and generalized variance.

We state the results in the following theorems without proof. Theorem 2.1 Let $k \geq 2, 3, \dots$ and $t > 0$. If an NEF F_t satisfies (3), then, up to affinity, F_t is of normal-Poisson model. Theorem 2.2 Let $F_t \in \mathcal{F}(\bullet, t)$ be an infinitely divisible NEF on \mathbb{R}^k such that 1. The canonical domain, $(\square) \in \mathbb{R}^k$ 2. $\det K(0) = 0$ 3. $\frac{1}{D(t)} \exp_{-k}^{-T} Q^T c_{-}$ celestin.kokonendji@univ-fcomte.fr Generalized Variance Estimations of Normal-Poisson Models 251 for t^m and $Q^T c_{-}$ given in (2). Then, up to affinity and power convolution, F_t is of normal-Poisson model.

All the technical details of proofs will be given in our article which is in preparation. In fact, the proof of Theorem 2.1 obtained by algebraic calculations and by using some properties of NEF is described in Proposition 2.1 below. An idea to proof Theorem 2.2 can be obtained using the infinite divisibility property of normal-Poisson for which this proof is the solution to the particular Monge–Ampère equation [12]: $\det K_{0,0}^{\text{TM}} / D_t \exp_{-k}^{\text{TM}} T Q^{\text{TM}} c_-$.

Gikhman and Skorokhod [13] showed that if μ is an infinitely divisible measure, then there exist a symmetric nonnegative definite $d \times d$ matrix Σ with rank $k-1$ and a positive measure ν on \mathbb{R}^k such that $K(0, t) = \Sigma + \int_0^t \int_{\mathbb{R}^k} x x^T \exp(-\langle x, z \rangle) \nu(dz) ds$. The Lévy–Khintchine formula of infinite divisibility distribution is also applied. Proposition 2.1 Let μ and ν be two finite positive measures on \mathbb{R}^k such that $\int \langle x, z \rangle^2 \mu(dz) < \infty$, $\int \langle x, z \rangle^2 \nu(dz) < \infty$, and $\mu(\{0\}) = 1$.

If there exists $(d, c) \in \mathbb{R}^{k \times r}$ such that $Q_{\cdot} d \in \text{Dexp}_{T_x E C c} \cdot d$; then $F D Q F W, \dots Q_{\cdot} / D, \dots / d$ and $K Q_{\cdot}^T / D K_{\cdot}^T C d / C c$; for $m \in \mathbb{M} F; V Q F \cdot m / D V F \cdot m$; and $\det V Q F \cdot m / D \det V F \cdot m / : 2$. If $Q_{\cdot} D \notin \cdot$ is the image measure of \cdot by the affine transformation $\cdot \mapsto A \cdot + b$; where A is a $k \times k$ nondegenerate matrix and $b \in \mathbb{R}^k$, then \cdot .

$Q_{-}/D_{AT}, \dots /$ and $KQ_{-}^{\text{TM}}/DK_{-AT}^{\text{TM}}CbT^{\text{TM}}I$ for $m \in \mathbb{D}AmCb2\mathbb{N}, MF/; VQF.$
 $m/D AVF\mathbb{N}_{-}1.m/_AT$; and $\det VQF.m/D.\det A/2 \det VF.m/ : 3.$ If $Q_{-}D_{-}t$
 is the t -th convolution power of \bullet for $t > 0$, then $\dots Q_{-}/D, \dots /$ and $KQ_{-}^{\text{TM}}/DtK_{-}^{\text{TM}}$
 $/I$ for $m \in \mathbb{D}tm2tMF; VQF.m/DtVF_{-}\mathbb{N}t_{-}1.m/$; and $\det VQF.m/Dtk\det VF$
 $(m).$ Proposition 2.1

shows that the generalized variance function $\det V F(m)$ of F is invariant for any element of its generating measure (Part 1) and for the affine transformation $y = Ax + b$ such that $\det A \neq 0$, particularly for a translation $y = x + b$ (Part 2). A reformulation of Theorem 2.2, by changing the canonical parameterization into mean parameterization, is stated in the following theorem. Theorem 2.3

Let $F_t \in \mathcal{F}(\cdot, t)$ be an infinitely divisible NEF on \mathbb{R}^k such that 1. $m_j > 0$ and $m \in \mathbb{R}^k$ with $\alpha_j \geq 0$. 2. $\det V F_t(m) \neq 0$: Then F_t is of normal-Poisson type.

celestin.kokonendji@univ-fcomte.fr 252 C.C. Kokonendji and K. Nisa Theorem 2.3 is equivalent to Theorem 2.2. The former is used for the estimation of generalized variance, and the latter is used for characterization by generalized variance.

Generalized Variance Estimations Here we present three methods for generalized variance estimations of normal-Poisson models $P(m; F_t) \in \mathcal{F}(\cdot, t) \in \mathcal{F}(\cdot, t)$, and then we report the result of our simulation study. Consider X_1, \dots, X_n be random vectors i.i.d. from $P(m; F_t)$ of normal-Poisson models, and we denote $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$.

$\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$ as the sample mean with positive j -th component X_j : The followings are ML, UMVU, and Bayesian generalized variance estimators. **Maximum Likelihood Estimator** Proposition 3.1 The ML estimator of $\det V F_t(m) / \det V F_t(m_0)$ is given by $T_n = \prod_{j=1}^k \frac{\Gamma(m_j)}{\Gamma(m_{0j})} \left(\frac{\bar{X}_j}{m_j} \right)^{m_j}$; (6) Proof The ML estimator above is easily obtained by replacing m_j in (5) with its ML estimator \bar{X}_j . **Uniformly Minimum Variance Unbiased Estimator** Proposition 3.2 The UMVU estimator of $\det V F_t(m) / \det V F_t(m_0)$

is given by $U_n = \prod_{j=1}^k \frac{\Gamma(m_j)}{\Gamma(m_{0j})} \left(\frac{\bar{X}_j}{m_j} \right)^{m_j}$; if $\bar{X}_j \leq m_j$; (7) Proof This UMVU estimator is obtained using intrinsic moment formula of univariate Poisson distribution as follows: $E \left(\frac{\bar{X}_j}{m_j} \right)^{m_j} = \frac{\Gamma(m_j)}{\Gamma(m_{0j})} \left(\frac{\bar{X}_j}{m_j} \right)^{m_j}$: Letting $Y \sim \text{Poisson}(\bar{X}_j)$ gives the result that (7) is the UMVU estimator of (5), because, by the completeness of NEFs, the unbiased estimation is unique. So, we deduced the desired result.

celestin.kokonendji@univ-fcomte.fr **Generalized Variance Estimations of Normal-Poisson Models** 253 A deep discussion about ML and UMVU methods on generalized variance estimations can be seen in [9] for NEF and [4] for NST models. **Bayesian Estimator** Proposition 3.3

Under assumption of prior gamma distribution of m_j with parameter $\alpha_j > 0$ and $\beta_j > 0$, the Bayesian estimator of $\det V F_t(m) / \det V F_t(m_0)$ is given by $B_n = \prod_{j=1}^k \frac{\Gamma(m_j)}{\Gamma(m_{0j})} \left(\frac{\bar{X}_j}{m_j} \right)^{m_j}$; (8) Proof Let X_1, \dots, X_n given m_j are $\text{Poisson}(m_j)$ with probability mass function $P(X_{ij} = x_{ij} | m_j) = \frac{e^{-m_j} m_j^{x_{ij}}}{x_{ij}!}$: Assuming that m_j follows

gamma(?, ?), then the prior probability distribution function of m_j is given by $f(m_j) = \frac{1}{\Gamma(k)} e^{-x} x^{k-1}$ for $m_j > 0$ and $k > 0$ where $\Gamma(\cdot)$ is the gamma function: $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Using the Bayes theorem, the posterior distribution of m_j given an observation sequence can be expressed as $f(m_j | x) = \frac{p(x | m_j) f(m_j)}{\int_0^\infty p(x | m_j) f(m_j) dm_j}$ which is a gamma density with parameters $\alpha = \sum_{i=1}^n x_{ij} + 1$ and $\beta = \frac{1}{C_j + 1}$.

Then with random sample X_{1j}, \dots, X_{nj} , the posterior will be gamma $\alpha = \sum_{i=1}^n X_{ij} + 1$; $\beta = \frac{1}{C_j + 1}$. The Bayesian estimator of m_j is given by the mean of the posterior distribution, i.e., $b_m = \frac{\alpha}{\beta}$, and then this concludes the proof. The choice of α and β depends on the information of m_j . Notice that for any positive value $c \in (0, 1)$; if $\alpha = c \sum_{i=1}^n X_{ij}$ and $\beta = \frac{1}{c(C_j + 1)}$, then the Bayesian estimator is the same as ML estimator.

In practice, the parameter of prior distribution of celestin.kokonendji@univ-fcomte.fr 254 C.C. Kokonendji and K. Nisa m_j must be known or can be assumed confidently before the generalized variance estimation. One can see, e.g., [14–16] for more details about Bayesian inference on m_j (univariate Poisson parameter).

Simulation Study In order to look at the performances of ML, UMVU, and Bayesian estimators of the generalized variance, we have done a Monte Carlo simulation using R software [17]. We have generated $k \in \{2, 4, 6, 8\}$ dimensional data from multivariate normal-Poisson distribution $F(\cdot, t)$ with $m_j \in \{1\}$. Fixing $j \in \{1\}$, we set several sample sizes n varied from 5 until 300, and we generated 1,000 samples for each sample size.

For calculating the Bayesian estimator, in this simulation we assume that the parameters of prior distribution depend on sample mean of Poisson component, X_j , and the dimension k . Then we set three different prior distributions: gamma X_j ; $k = 1$; gamma X_j ; $k = 2$; and gamma X_j ; $k = 3$. We report the results of the generalized variance estimations using the three methods in Table 1.

From these values, we calculated the mean square error (MSE) of each method over 1,000 data sets using the following formula $MSE = \frac{1}{1000} \sum_{i=1}^{1000} (G_{V_i} - G_{V_m})^2$ where G_{V_i} is the estimate of m_k using each method. From the values in Table 1, we can observe different performances of ML estimator ($T_{n,t}$), UMVU estimator ($U_{n,t}$), and Bayesian estimator ($B_{n,t}, ', ''$) of the generalized variance.

The values of $T_{n,t}$ and $B_{n,t}, ', ''$ converge, while the values of $U_{n,t}$ do not, but $U_{n,t}$ which

is the unbiased estimator always approximate the parameter ($m \leq 1$) and closer to the parameter than $T_{n,t}$ and $B_{n,t}$, for small sample sizes $n \leq 25$. For all methods, the standard error of the estimates decreases when the sample size increases.

The Bayesian estimator with gamma $X_j; k = 2$ prior distribution, i.e., $B_{n,t}; X_j; k = 2$, is exactly the same as $T_{n,t}$ for $k \leq 2$. This is because in this case, the Bayesian and ML estimators of m_1 are the same (i.e., $c \leq 1$). The goodness of Bayesian estimator depends on the parameter of prior distribution, γ and α .

From our simulation, the result shows that smaller parameter α gives greater standard error to the estimations in small sample sizes, and the accuracy of $B_{n,t}$, γ , α with respect to α varies with dimensions k . However, they are all asymptotically unbiased. There are more important performance characterizations for an estimator than just being unbiased. The MSE is perhaps the most important of them.

It captures the celestin.kokonendji@univ-fcomte.fr Generalized Variance Estimations of Normal-Poisson Models 255 Table 1 The expected values (with standard error) of $T_{n,t}$, $U_{n,t}$, and $B_{n,t}$, γ , α with $m_1 \leq 1$ and $k = 2, 4, 6, 8$ (target values $m \leq 1$) $k \leq 2$ n $T_{n,t}$ $U_{n,t}$ $B_{n,t}; X_j; k = 2$ $B_{n,t}; X_j; k = 3$ $k \leq 1$ 1.2790 (1.3826) 0.9533 (1.2050) 0.8186 (0.8849) 1.2790 (1.3826) 1.5221 (1.6454) $k \leq 5$ 1.1333 (0.8532) 0.9915 (0.8000) 0.8955 (0.6742) 1.1333 (0.8532) 1.2340 (0.9290) $k \leq 10$ 1.1121 (0.6295) 1.0276 (0.6056) 0.9589 (0.5428) 1.1121 (0.6295) 1.1714 (0.6631) 25 1.0357 (0.4256) 0.9959 (0.4175) 0.9604 (0.3946) 1.0357 (0.4256) 1.0628 (0.4367) 60 1.0090 (0.2526) 0.9924 (0.2505) 0.9767 (0.2445) 1.0090 (0.2526) 1.0201 (0.2553) 100 1.0086 (0.1988) 0.9986 (0.1979) 0.9890 (0.1950) 1.0086 (0.1988) 1.0153 (0.2002) 300 0.9995 (0.1141) 0.9962 (0.1140) 0.9929 (0.1134) 0.9995 (0.1141) 1.0017 (0.1144) $k \leq 4$ n $T_{n,t}$ $U_{n,t}$ $B_{n,t}; X_j; k = 2$ $B_{n,t}; X_j; k = 3$ $k \leq 1$ 2.3823 (4.6248) 0.9460 (2.5689) 0.4706 (0.9135) 1.2859 (2.4964) 1.9190 (3.7254) $k \leq 5$ 1.6824 (2.4576) 0.9531 (1.6995) 0.5890 (0.8605) 1.1491 (1.6786) 1.4756 (2.1555) $k \leq 10$ 1.4664 (1.6345) 1.0027 (1.2456) 0.7072 (0.7882) 1.1328 (1.2626) 1.3430 (1.4969) 25 1.2711 (1.0895) 1.0169 (0.9327) 0.8212 (0.7039) 1.0930 (0.9368) 1.2079 (1.0353) 60 1.0978 (0.5682) 0.9961 (0.5288) 0.9060 (0.4689) 1.0287 (0.5324) 1.0741 (0.5559) 100 1.0589 (0.4209) 0.9983 (0.4028) 0.9419 (0.3744) 1.0180 (0.4046) 1.0451 (0.4154) 300 1.0273 (0.2305) 1.0071 (0.2271) 0.9874 (0.2215) 1.0138 (0.2275) 1.0228 (0.2295) $k \leq 6$ n $T_{n,t}$ $U_{n,t}$ $B_{n,t}; X_j; k = 2$ $B_{n,t}; X_j; k = 3$ $k \leq 1$ 4.7738 (13.9827) 0.9995 (4.7073) 0.2593 (0.7594) 1.2514 (3.6655) 2.3548 (6.8972) $k \leq 5$ 2.9818 (6.2595) 0.9958 (2.7565) 0.3689 (0.7743) 1.1825 (2.4823) 1.8446 (3.8723) $k \leq 10$ 2.2232 (4.0454) 1.0124 (2.2131) 0.4733 (0.8612) 1.1406 (2.0756) 1.5778 (2.8709) 25 1.6399 (2.2478) 0.9555 (1.4833) 0.5708 (0.7824) 1.0513 (1.4410) 1.3076 (1.7923) 60 1.2479 (0.9978) 0.9827 (0.8226) 0.7778 (0.6220) 1.0283 (0.8222) 1.1319 (0.9051) 100 1.1830 (0.7646) 1.0235 (0.6800) 0.8853 (0.5722) 1.0517 (0.6798) 1.1151 (0.7207) 300 1.0530

(0.3758) 1.0022 (0.3608) 0.9539 (0.3404) 1.0119 (0.3612) 1.0322 (0.3684) $k D$ $8n$ $T_{n,t}$ $U_{n,t}$
 $B_{n,t}; X_{j;k}$ $B_{n,t}; X_{j;k} = 2$ $B_{n,t}; X_{j;k} = 3$ $k C 1$ 8.5935 (31.9230) 0.8677 (5.4574)
0.1232 (0.4576) 1.0535 (3.9134) 2.5038 (9.3010) $k C 5$ 4.7573 (12.5015) 0.8468 (3.0478)
0.1856 (0.4878) 1.0065 (2.6448) 1.9345 (5.0836) $k C 10$ 3.6816 (9.0892) 1.0394 (3.2258)
0.2994 (0.7392) 1.1394 (2.8130) 1.8789 (4.6387) 25 2.9055 (6.3150) 1.1341 (2.9623)
0.4314 (0.9377) 1.2129 (2.6362) 1.7675 (3.8416) 60 1.6201 (1.8804) 1.0511 (1.3062)
0.6794 (0.7885) 1.1035 (1.2807) 1.3059 (1.5156) 100 1.2890 (1.0907) 0.9850 (0.8667)
0.7541 (0.6381) 1.0199 (0.8630) 1.1308 (0.9569) 300 1.1056 (0.5378) 1.0086 (0.4968)
0.9199 (0.4474) 1.0213 (0.4967) 1.0578 (0.5145) bias and the variance of the estimator.

For this reason, we compare the quality of the estimators using MSE in Table 2 which are presented graphically in Figs. 1, 2, 3, and 4. From these figures, we conclude that all estimators become more similar when the sample size increases. For small sample sizes, $B_{n,t}; X_{j;k}$ always has the smallest MSE, while $T_{n,t}$ always has the greatest MSE (except for $k D 2$).

For $n \geq 25$, $U_{n,t}$ is preferable than $T_{n,t}$. In this situation, the difference between $U_{n,t}$ and $T_{n,t}$ increases when the dimension increases and also the difference between $T_{n,t}$ and $B_{n,t}$, ' , ". celestin.kokonendji@univ-fcomte.fr 256 C.C. Kokonendji and K. Nisa Table 2
The mean square error of $T_{n,t}$, $U_{n,t}$, and $B_{n,t}$, ?, ? of Table 1 $k D 2n$ MSE ($T_{n,t}$) MSE ($U_{n,t}$)
MSE ($B_{n,t}$)

$n ; t ; X_{j;k} / MSE . B_{n,t} ; X_{j;k} = 2 / MSE . B_{n,t} ; X_{j;k} = 3 / k C 1$ 1.9894 1.4542 0.8159
1.9894 2.9800 $k C 5$ 0.7458 0.6401 0.4654 0.7458 0.9179 $k C 10$ 0.4088 0.3675 0.2963
0.4088 0.4690 25 0.1824 0.1743 0.1573 0.1824 0.1947 60 0.0639 0.0628 0.0603 0.0639
0.0656 100 0.0396 0.0391 0.0381 0.0396 0.0403 300 0.0130 0.0130 0.0129 0.0130 0.0131
 $k D 4n$ MSE ($T_{n,t}$) MSE ($U_{n,t}$) MSE ($B_{n,t}$) $n ; t ; X_{j;k} / MSE . B_{n,t} ; X_{j;k} = 2 / MSE . B_{n,t} ; X_{j;k} = 3 / k C 1$ 23.2999 6.6019 1.1149 6.3136 14.7231 $k C 5$ 6.5055 2.8904 0.9093 2.8398
4.8724 $k C 10$ 2.8891 1.5514 0.7071 1.6118 2.3585 25 1.2604 0.8702 0.5274 0.8862
1.1151 60 0.3324 0.2797 0.2287 0.2843 0.3146 100 0.1806 0.1622 0.1435 0.1640 0.1746
300 0.0539 0.0516 0.0492 0.0519 0.0532 $k D 6n$ MSE ($T_{n,t}$) MSE ($U_{n,t}$) MSE ($B_{n,t}$) $n ; t ; X_{j;k} / MSE . B_{n,t} ; X_{j;k} = 2 / MSE . B_{n,t} ; X_{j;k} = 3 / k C 1$ 209.7568 22.1589 1.1254 13.4989 49.4073
 $k C 5$ 43.1085 7.5980 0.9979 6.1952 15.7078 $k C 10$ 17.8618 4.8981 1.0191 4.3278 8.5761
25 5.4622 2.2020 0.7964 2.0790 3.3071 60 1.0571 0.6769 0.4362 0.6769 0.8366 100
0.6181 0.4629 0.3406 0.4647 0.5327 300 0.1440 0.1302 0.1180 0.1306 0.1368 $k D 8n$ MSE
($T_{n,t}$) MSE ($U_{n,t}$) MSE ($B_{n,t}$) $n ; t ; X_{j;k} / MSE . B_{n,t} ; X_{j;k} = 2 / MSE . B_{n,t} ; X_{j;k} = 3 / k C 1$ 1,076.7380 29.8009 0.9782 15.3177 88.7698 $k C 5$ 170.4059 9.3124 0.9012 6.9951
26.7168 $k C 10$ 89.8046 10.4076 1.0373 7.9326 22.2895 25 43.5105 8.7931 1.2025 6.9949

$n ; t ; X_{j;k} = 2 / MSE . B_{n,t} ; X_{j;k} = 3 / k C 1$ 209.7568 22.1589 1.1254 13.4989 49.4073
 $k C 5$ 43.1085 7.5980 0.9979 6.1952 15.7078 $k C 10$ 17.8618 4.8981 1.0191 4.3278 8.5761
25 5.4622 2.2020 0.7964 2.0790 3.3071 60 1.0571 0.6769 0.4362 0.6769 0.8366 100
0.6181 0.4629 0.3406 0.4647 0.5327 300 0.1440 0.1302 0.1180 0.1306 0.1368 $k D 8n$ MSE
($T_{n,t}$) MSE ($U_{n,t}$) MSE ($B_{n,t}$) $n ; t ; X_{j;k} / MSE . B_{n,t} ; X_{j;k} = 2 / MSE . B_{n,t} ; X_{j;k} = 3 / k C 1$ 1,076.7380 29.8009 0.9782 15.3177 88.7698 $k C 5$ 170.4059 9.3124 0.9012 6.9951
26.7168 $k C 10$ 89.8046 10.4076 1.0373 7.9326 22.2895 25 43.5105 8.7931 1.2025 6.9949

15.3466 60 3.9204 1.7088 0.7246 1.6509 2.3907 100 1.2732 0.7515 0.4676 0.7452 0.9327
 300 0.3003 0.2469 0.2066 0.2472 0.2681 In this simulation, $B_{n,t}; X_{j;k}$ is the best estimator because of its smallest MSE, but in general we cannot say that Bayesian estimator is much better than ML and UMVU estimators since it depends on the prior distribution parameters. In fact, one would prefer $U_{n,t}$ as it is the unbiased estimator with the minimum variance.

However, if in practice we know the information about prior distribution of m_j , we can get a better estimate (in the sense of having a lower MSE) than $U_{n,t}$ by using $B_{n,t}, ', ''$.
 celestin.kokonendji@univ-fcomte.fr Generalized Variance Estimations of Normal-Poisson Models 257 0.0 3 9 14 25 60 100 300 Sample Size 0.5 1.0 1.5 2.0 2.5 Mean Square Error ML -k BAYES-k/2 BAYES-k/3 Fig.

1 MSE plot of $T_{n,t}$, $U_{n,t}$, $B_{n,t}; x_{j;k}$; $B_{n,t}; x_{j;k} = 2$; and $B_{n,t}; x_{j;k} = 3$ for $k \in \mathbb{D} \subset \mathbb{C}$ on conclusion In this chapter, we have established the definition and properties of normal-Poisson models as a generalization of normal-Poisson 1 and showed that the generalized variance of normal-Poisson models depends mainly on the mean of the Poisson component.

The estimations of generalized variance using ML, UMVU, and Bayesian estimators show that UMVU produces a better estimation than ML estimator, while compared to Bayesian estimator, UMVU is worse for some choice of prior distribution parameters, but it can be much better for other cases. However, all methods are consistent estimators, and they become more similar when the sample size increases.
 celestin.kokonendji@univ-fcomte.fr 258 C.C.

Kokonendji and K. Nisa 0 5 9 14 25 60 100 300 Sample Size 5 10 15 20 Mean Square Error ML -k BAYES-k/2 BAYES-k/3 Fig. 2 MSE plot of $T_{n,t}$, $U_{n,t}$, $B_{n,t}; x_{j;k}$; $B_{n,t}; x_{j;k} = 2$; and $B_{n,t}; x_{j;k} = 3$ for $k \in \mathbb{D}$ 4 0 7 11 16 25 60 100 300 Sample Size 50 100 150 200 Mean Square Error ML -k BAYES-k/2 BAYES-k/3 Fig. 3 MSE plot of $T_{n,t}$, $U_{n,t}$, $B_{n,t}; x_{j;k}$; $B_{n,t}; x_{j;k} = 2$; and $B_{n,t}; x_{j;k}$ for $k \in \mathbb{D}$ 6 celestin.kokonendji@univ-fcomte.fr Generalized Variance Estimations of Normal-Poisson Models 259 0 13 9 18 25 60 100 300 Sample Size 200 400 600 800 1000 Mean Square Error ML -k BAYES-k/2 BAYES-k/3 Fig. 4 MSE plot of $T_{n,t}$, $U_{n,t}$, $B_{n,t}; x_{j;k}$; $B_{n,t}; x_{j;k} = 2$; and $B_{n,t}; x_{j;k} = 3$ for $k \in \mathbb{D}$ 8
 References 1. Hassairi, A.:

Generalized variance and exponential families. Ann. Stat. 27 (1), 374–385 (1999) 2.
 Kokonendji, C.C., Pommeret, D.: Estimateurs de la variance généralisée pour des familles exponentielles non gaussiennes. C. R. Acad. Sci. Ser. Math. 332 (4), 351–356 (2001) 3.
 Shorrock, R.W., Zidek, J.V.: An improved estimator of the generalized variance. Ann. Stat.

4(3), 629–638 (1976) 4. Boubacar Mainassara, Y., Kokonendji, C.C.:

On normal stable Tweedie models and power generalized variance function of only one component. TEST 23 (3), 585–606 (2014) 5. Casalis, M.: The 2d C 4 simple quadratic natural exponential families on \mathbb{R}^d . Ann. Stat. 24 (4), 1828–1854 (1996) 6. G. Letac, Le problème de la classification des familles exponentielles naturelles de \mathbb{R}^d ayant une fonction variance quadratique, in Probability Measures on Groups IX, H. Heyer, Ed. Springer, Berlin, 1989, pp. 192–216. 7. Kokonendji, C.C., Masmoudi, A.: A characterization of Poisson-Gaussian families by generalized variance.

Bernoulli 12 (2), 371–379 (2006) 8. Kokonendji, C.C., Seshadri, V.: On the determinant of the second derivative of a Laplace transform. Ann. Stat. 24 (4), 1813–1827 (1996) 9. Kokonendji, C.C., Pommeret, D.: Comparing UMVU and ML estimators of the generalized variance for natural exponential families. Statistics 41 (6), 547–558 (2007) 10. Kotz, S., Balakrishnan, N., Johnson, N.L.: Continuous Multivariate Distributions. Models and Application, vol. 1, 2nd edn.

Wiley, New York (2000) celestin.kokonendji@univ-fcomte.fr 260 C.C. Kokonendji and K. Nisa 11. Kokonendji, C.C., Masmoudi, A.: On the Monge–Ampère equation for characterizing gamma-Gaussian model. Stat. Probab. Lett. 83 (7), 1692–1698 (2013) 12. Gutiérrez, C.E.: The Monge–Ampère Equation. Birkhäuser, Boston (2001). Boston: Imprint: Birkhäuser 13. Gikhman, I.I., Skorokhod, A.V.: The Theory of Stochastic Processes 2. Springer, New York (2004) 14.

Berger, J.O.: Statistical Decision Theory and Bayesian Analysis, 2nd edn. Springer, New York (1985) 15. Sultan, R., Ahmad, S.P.: Posterior estimates of Poisson distribution using R software. J. Mod. Appl. Stat. Methods 11 (2), 530–535 (2012) 16. Hogg, R.V.: Introduction to Mathematical Statistics, 7th edn. Pearson, Boston (2013) 17. R Development Core Team: R: A Language and Environment for Statistical Computing.

R Foundation for Statistical Computing, Vienna (2009)
celestin.kokonendji@univ-fcomte.fr

INTERNET SOURCES:

4% - https://link.springer.com/chapter/10.5176%2F2251-1911_CMCGS14.29_21

<1% -

https://mafiadoc.com/catatan-kuliah-simulasi-dan-pemodelan-universitas-gunadarma_59d35b281723dd27b575daaf.html

1% -

<http://dl4.globalstf.org/?wpsc-product=generalized-variance-estimations-of-normal-poisson-models>

<1% -

https://www.researchgate.net/publication/220363689_An_improved_class_of_estimators_for_the_population_mean

<1% - <https://www.calameo.com/books/005388144e9a5f74f7dd0>

<1% - <https://quizlet.com/45664143/multivariate-data-analysis-flash-cards/>

<1% - <https://juniperpublishers.com/bboaj/pdf/BBOAJ.MS.ID.555612.pdf>

<1% -

https://www.researchgate.net/publication/38358071_An_Improved_Estimator_of_the_Generalized_Variance

<1% - https://www.academia.edu/36846641/Non-parametric_Tests_for_Complete_Data

<1% - <http://www.columbia.edu/~ks20/4703-Sigman/4703-07-Notes-0.pdf>

1% -

https://www.researchgate.net/publication/272006326_On_normal_stable_Tweedie_models_and_power-generalized_variance_functions_of_only_one_component

<1% - <https://singapore.kinokuniya.com/bw/9783319161389>

1% -

https://www.researchgate.net/publication/38322468_A_characterization_of_Poisson-Gaussian_families_by_generalized_variance

<1% - <https://www.sciencedirect.com/topics/computer-science/gaussian-component>

<1% -

https://www.academia.edu/28552332/Generalized_variance_estimators_in_the_multivariate_gamma_models

1% -

https://www.researchgate.net/publication/232834424_Comparing_UMVU_and_ML_estimators_of_the_generalized_variance_for_natural_exponential_families

<1% - <https://dergipark.org.tr/download/article-file/351094>

<1% -

https://mafiadoc.com/lecture-notes-on-thermodynamics-and-statistical-mechanics-a-_59d7047b1723ddc005fa4be2.html

<1% -

https://mafiadoc.com/applied-numerical-methods-using-matlabpdf_5b9a5c0b097c47562d8b46b7.html

<1% - <https://www.sciencedirect.com/science/article/pii/0013794494900817>

<1% - <https://arxiv.org/pdf/1905.00383>

<1% - <https://wenku.baidu.com/view/38eac0dbad51f01dc281f18a.html>

<1% -

https://www.academia.edu/4551730/Traffic_Collision_Analysis_Models_Review_And_Empirical_Evaluation

<1% - <https://www.sciencedirect.com/science/article/pii/S0002929707610301>
<1% - <https://www.sciencedirect.com/science/article/pii/S0167715213001065>
<1% - <https://manned.org/gromacs.1>
<1% - <https://www.sciencedirect.com/science/article/pii/S0165168419302622>
<1% - <https://pharmacy.ufl.edu/files/2013/01/Basic-Pharmacokinetics.pdf>
<1% - https://eprints.usq.edu.au/3888/1/Dunn_Smyth_Stats_and_Comp_v18n1.pdf
<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3152312/>
<1% -
http://support.sas.com/documentation/cdl/en/statug/67523/HTML/default/statug_gen_mod_details01.htm
<1% -
https://www.academia.edu/881943/FIR_system_identification_based_on_subspaces_of_a_higher_order_cumulant_matrix
<1% - <https://core.ac.uk/download/pdf/81838211.pdf>
<1% - <https://en.wikipedia.org/wiki/M.2>
<1% - https://www.youtube.com/watch?v=opwON-7J_wl
<1% - [https://www.wikiyy.com/en/Matrix_\(mathematics\)](https://www.wikiyy.com/en/Matrix_(mathematics))
<1% - <http://lmb.univ-fcomte.fr/IMG/pdf/arnjeas2017nketal.pdf>
<1% - https://slidelegend.com/gnu-octave_59c3d0641723dd5342dcefacc.html
<1% - <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018GC007585>
1% - [https://en.wikipedia.org/wiki/Dimension_\(vector_space\)](https://en.wikipedia.org/wiki/Dimension_(vector_space))
1% -
https://www.researchgate.net/publication/257013240_On_the_Monge-Ampere_equation_for_characterizing_gamma-Gaussian_model
<1% -
http://www.shsu.edu/~kws006/Precalculus/2.3_Zeroes_of_Polynomials_files/S%26Z%203.2.pdf
<1% - https://en.wikipedia.org/wiki/Talk:Cayley%E2%80%93Hamilton_theorem
<1% -
https://www.academia.edu/25064257/Pseudoeffective_and_nef_classes_on_abelian_varieties
<1% -
https://docs.wixstatic.com/ugd/a5a7e4_6edf196799b7438fb5ae15cf441a00c4.pdf?index=true
<1% -
<https://www.math.tugraz.at/~edson/Publications/Representations%20in%20Fibonacci%20Base.pdf>
<1% -
<https://www.sec.gov/Archives/edgar/data/866787/000086678702000052/0000866787-02-000052.txt>

<1% - <https://www.chordtela.com/2017/09/virgoun-bukti.html>
<1% - <https://www.bt.dk/>
<1% - <https://pytorch.org/docs/stable/torch.html>
<1% - <https://epdf.pub/kalman-filtering-theory-and-practice-using-matlab.html>
<1% -
https://slidelegend.com/linear-matrix-inequalities-in-control_5a96111c1723dd63ba0819a0.html
<1% -
https://www.researchgate.net/publication/50401947_Composite_parameterization_and_Haar_measure_for_all_unitary_and_specialunitary_groups
<1% - <https://www.sciencedirect.com/science/article/pii/S0003269705003222>
<1% - https://mafiadoc.com/evIEWS-8-users-guide-ii_5ca85b8c097c479d018b4631.html
<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3105782/>
<1% - <https://brainly.in/question/11681513>
<1% -
<https://www.scribd.com/doc/120449142/MATHEMATICAL-STATISTICS-Keith-Knight-pdf>
<1% - <https://civil29.blogspot.com/2012/09/upsc-syllabus-with-question-papers.html>
<1% - <https://gawhitaker.github.io/project.pdf>
<1% - <http://cfile209.uf.daum.net/attach/212C393D5743798C126386>
<1% -
http://docshare.tips/applications-of-statistics-and-probability-in-civil-engineering_577ec0bab6d87fdf4e8b4afb.html
<1% - https://issuu.com/vamsireddy22/docs/fe_reference_handbook
<1% - <https://www.sciencedirect.com/science/article/pii/S0377221709007905>
<1% -
<https://kraina-ua.com/up/files/infolist/????????????%20????%2019.08.2019-%2024.08.2019%20???pdf>
<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5031345/>
<1% - https://en.m.wikipedia.org/wiki/Log-normal_distribution
<1% -
https://www.researchgate.net/publication/319238246_Using_Bayesian_statistics_to_estimate_the_likelihood_a_new_trial_will_demonstrate_the_efficacy_of_a_new_treatment
<1% - <https://www.sciencedirect.com/science/article/pii/S0377221712002032>
<1% - <https://faculty.washington.edu/ezivot/econ583/mleLectures.pdf>
<1% -
https://www.researchgate.net/publication/271855086_Effects_of_associated_kernels_in_nonparametric_multiple_regressions
<1% - <https://academic.oup.com/biostatistics/article/11/2/317/268224>
<1% -
<http://oai.repec.org/?verb=ListRecords&set=RePEc:bot:rivsta&metadataPrefix=amf>

<1% - <http://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0162259>
 <1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3850654/>
 <1% - <https://www.mathworks.com/help/econ/what-is-bayesian-linear-regression.html>
 <1% - https://www.academia.edu/38063947/Marshall-Olkin_Generalized_Exponential_Distribution_Different_Methods_of_Estimations
 <1% - <https://www.mdpi.com/2220-9964/8/6/245/pdf>
 <1% - https://en.wikipedia.org/wiki/Student%27s_t-distribution
 <1% - http://nptel.ac.in/courses/112108149/pdf/M2/WE_M2.pdf
 <1% - <https://www.scribd.com/document/326871868/Bickel-Mathematical-Statistics-Basic-Ide-pdf>
 <1% - <https://www.scribd.com/document/346397522/Asymptotic-Statistics-Lecture-Notes>
 <1% - <https://open.library.ubc.ca/handle/2429/16692>
 <1% - <https://stats.stackexchange.com/questions/129885/why-does-increasing-the-sample-size-lower-the-sampling-variance>
 <1% - https://www.academia.edu/38341892/Bishop_nato_bayes
 <1% - https://link.springer.com/protocol/10.1007/978-1-4939-9074-0_13
 <1% - https://www.researchgate.net/publication/320244403_Parameter_Estimation_In_Weighted_Rayleigh_Distribution
 <1% - <https://www.sciencedirect.com/science/article/pii/S0304407694017166>
 <1% - <http://theanalysisofdata.com/notes/estimators1.pdf>
 <1% - <https://www.sciencedirect.com/science/article/pii/S0888613X17305686>
 <1% - https://www.researchgate.net/publication/229004125_A_Monte_Carlo_simulation_of_the_impact_of_sample_size_and_percentile_method_implementation_on_imagery_geolocation_accuracy_assessments
 <1% - https://mafiadoc.com/pdf-3185-kb_59d5e8bb1723dd68cfd31335.html
 <1% - <https://www.sciencedirect.com/science/article/pii/S0362546X78900135>
 <1% - <https://mcdougallbay.com/images/repository/%7B52381F8B-B90A-424C-B6A4-D250354423C0%7D/2019/48501%20-%20ISC%20Search.pdf>
 <1% - https://www.researchgate.net/publication/216300914_Design_and_Analysis_of_Simulation_Experiments
 <1% - <https://epdf.pub/sequential-estimation58073ef5b9fd5b0d5f3dcf3330c5c49819544.html>

<1% - <https://ufdc.ufl.edu/AA00062947/00048>

<1% -

http://www.archive.org/stream/historyandantiq00hollgoog/historyandantiq00hollgoog_djvu.txt

<1% - <https://en.wikipedia.org/wiki/Mean>

<1% -

https://www.researchgate.net/publication/282868157_On_the_Estimation_for_the_Weibull_Distribution

<1% - <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1467-9531.2009.01221.x>

<1% -

https://mafiadoc.com/2nd-maphysto-levy-conference_5a2c8a781723dd5d7993821f.html

<1% - <https://link.springer.com/article/10.3103%2FS1066530708010055>

<1% -

<https://www.cambridge.org/core/books/computation-and-modelling-in-insurance-and-finance/285CC67E533068A80FC119E9750A055D>

<1% - <https://www.loot.co.za/index/html/index4838.html>



Plagiarism Checker X Originality Report

Similarity Found: 20%

Date: Selasa, Agustus 20, 2019

Statistics: 679 words Plagiarized / 3384 Total words

Remarks: Medium Plagiarism Detected - Your Document needs Selective Improvement.

INSIST Vol. 2 No. 1, April 2017 (1 - 5) <http://insist.unila.ac.id/> DOI: 10.23960/ins.v2i1.23
Received: 20/08/2016 Accepted: 19/01/2017 Published online: 01/04/2017 eISSN: 2502-8588
Abstract —In this paper, a robust procedure for estimating parameters of regression model when generalized estimating equation (GEE) applied to longitudinal data that contains outliers is proposed.

The method is called 'iteratively reweighted least trimmed square' (IRLTS) which is a combination of the iteratively reweighted least square (IRLS) and least trimmed square (LTS) methods. To assess the proposed method a simulation study was conducted and the result shows that the method is robust against outliers. Keywords —GEE, IRLS, LTS, longitudinal data, regression model. I.

INTRODUCTION LONGITUDINAL studies are increasingly common in many scientific research areas, for example in the social, biomedical, and economical fields. In longitudinal studies, multiple measurements are taken on the same subject at different points in time. Thus, observations for the same subject are correlated. The analysis of data resulting from such studies often becomes complicated due to the within- subject correlation.

This correlation must be considered for any appropriate analysis method. Generalized linear models (GLM) as described by McCullagh and Nelder [1] is a standard method used to fit regression models for univariate data that are presumed to follow an exponential family distribution. The association between the response variable and the covariates is given by the link function.

GLM assume that the observations are independent and do not consider any correlation

between the outcome of the n observations. Liang and Zeger [2] introduced an approach to this correlation problem using GEE to extend GLM into a regression setting with correlated observations within subjects. The GEE method of Liang and Zeger gives consistent estimators of the regression parameter.

The parameter estimates are consistent even when the variance structure is misspecified under mild regularity conditions. However, problems can occur when data contain outliers. The method is not robust against outliers since it is based on score equations from the quasi likelihood method of estimation.

The working correlation matrix would be affected by the outliers and also the parameter estimates. In this situation, we need a robust method that can minimize the effect of outliers. In recent years, a few authors have considered robust methods for longitudinal data analysis.

For example, Qaqish and Preisser [3] proposed a resistant version of the GEE using M-type estimation by involving down-weighting influential 1Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Lampung, Jalan Prof. Soemantri Brojonegoro No. 1, Bandar Lampung, Indonesia. *Correspondence to Khoirin Nisa, email: khoirin.nisa@fmipa.unila.ac.id. Tel.: +62 721 701609; fax: +62 721 702767. data points.

Gill [4] proposed a robustified likelihood based on multivariate normal distribution. Jung and Ying [5] proposed an adaptation of the Wilcoxon-Mann-Whitney method of estimating linear regression parameters for use in longitudinal data analysis under the working independence model. And recently, Abebe et al. [6] proposed a robust GEE using iterated reweighted rank-based estimation.

In this paper, we adopt the LTS [7] method for robust linear regression in the sense of trimming the data for estimating the regression coefficients so that the observations with high residuals are not included in the parameter estimation. In Section 2 we present a brief review of GEE. In Section 3 we describe our proposed method IRLTS. In Section 4 we discuss some results from our simulation study. II.

GENERALIZED ESTIMATING EQUATION AND IRLS METHOD Let Y_{ij} , $j = 1, \dots, m_i$, $i = 1, \dots, n$ represent the j th measurement on the i th subject. There are m_i measurements on subject i and $N = \sum_{i=1}^n m_i$ total measurements. Assume that the marginal distribution of y_{ij} is of the exponential class of distributions and is given by: $\{ \} f f q q f q .()/(exp),,(y c a b y y f + - =$ where $a(\cdot)$, $b(\cdot)$, and $c(\cdot)$ are given, q is the canonical parameter, and f is the dispersion parameter.

Let the vector of measurements on the i th subject be $Y_i = [Y_{i1}, \dots, Y_{im_i}]^T$ with corresponding vector of means $\mu_i = [\mu_{i1}, \dots, \mu_{im_i}]^T$ and $X_i = [X_{i1}, \dots, X_{im_i}]^T$ be the $m_i \times p$ matrix of covariates. In general, the components of Y_i are correlated but Y_i and Y_k are independent for any $i \neq k$. To model the relation between the response and covariates, we can use a regression model similar to the generalized linear models: $g(\mu_i) = \eta_i = X_i \beta$ where $\mu_i = E(Y_i|X_i)$, g is a specified link function, and $\beta = [\beta_1, \dots, \beta_p]^T$ is a vector of unknown regression coefficients to be estimated.

The GEE for estimating the $p \times 1$ vector of regression parameters β is given by:
$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = - \frac{1}{n} \sum_{i=1}^n \frac{Y_i - \mu_i}{V_i} \frac{\partial \mu_i}{\partial \beta} \quad (1)$$
 where V_i be the covariance matrix of Y_i modeled as $V_i = \frac{1}{2} \frac{\partial^2 \eta_i}{\partial \beta \partial \beta^T} A_i R(a) A_i$, A_i is a diagonal matrix of variance function $V(\mu_{ij})$, and $R(a)$ is the working correlation matrix of Y_i indexed by a vector of parameters a . Solutions to (2) are obtained by alternating between estimation of β , a and q .

There are several specific choices of the form of working correlation matrix $R_i(a)$ commonly used to model Robust Estimation of Generalized Estimating Equation when Data Contain Outliers Khoirin Nisa 1,* , Netti Herawati 2 L 2 INSIST Vol. 2 No. 1, April 2017 (1 - 5) the correlation matrix of Y_i . A few of the choices are shown below, one can refer to [1] for additional choices.

The dimension of the vector a , which is treated as a nuisance parameter, and the form of the estimator of a are different for each choice. Some typical choices are: 1. $R_i(a) = R_0$, a fixed correlation matrix. For $R_0 = I$, the identity matrix, the GEE reduces to the independence estimating equation. 2. Exchangeable: $k_j Y Y \text{ Cor } i k j^{-1} = \rho$, $\rho(a)$. 3. Autoregressive-1: $\rho(k) = \rho^{k-j}$, $k_j Y Y \text{ Cor } i k j = \rho^{k-j}$. 4. Unstructured: $j k i k j Y Y \text{ Cor } a = \rho$, $\rho(a)$.

Solving for β is done with iteratively reweighted least squares (IRLS). The following is the algorithm for fitting the specified model using GEEs [3]: 1. Compute an initial estimate of GEE $\hat{\beta}$, for example with an ordinary generalized linear model assuming independence. 2. A current estimate GEE $\hat{\beta}$ is updated by regressing the working response vector $\eta_i(\hat{\mu}_i, \hat{\beta})$ on X_i .

A new estimate new $\hat{\beta}$ is obtained by:
$$\hat{\beta}^{new} = (Z^T W X X^T W X \hat{\beta} + T^T T)^{-1} (Z^T W X \hat{\mu}_i + T^T T \hat{\beta}) \quad (2)$$
 where W is a block diagonal weight matrix whose i th block is the $m_i \times m_i$ matrix $W_i = \frac{1}{V_i} \frac{\partial \mu_i}{\partial \beta} \frac{\partial \mu_i}{\partial \beta^T}$. 3. Use new $\hat{\beta}$ to update $\eta_i = X_i \hat{\beta}$, where $\eta_i = X_i \hat{\beta}$. 4. Iterate until convergence. III.

ITERATIVELY REWEIGHTED LEAST TRIMMED SQUARE ALGORITHM First let us briefly recall that the robust estimation of regression parameters using LTS method is given by:

$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^h |e_i|$ which is based on the ordered absolute residuals $|e_1| \leq |e_2| \leq \dots \leq |e_n|$. LTS estimation is calculated by minimizing the h ordered squares residuals, where h can be chosen within $4 \leq h \leq n$, with n and p being sample size and number of parameters respectively. When $h = [n/2]$, LTS locates that half of the observations which has the smallest estimated variance.

In that case, the breakdown point is 50%. When h is set to the sample size, LTS and ordinary least square (OLS) coincide. In [7] Rousseeuw and Leroy shows $n^{1/2}$ consistency and asymptotic normality of LTS in the location-scale model. Věšek [8] extends this to the regression model with random regressors, the proof for fixed regressors is in later series of his papers: [9][10].

When n is very small, it is possible to generate all subsets of size h to determine which one minimizes the LTS criterion. Rousseeuw and Leroy computation of LTS based on subsets of size h requires $\binom{n}{h}$ subsets which is usually still too large for realistic applications. When n is small enough: 1. Select h . 2. Generate all possible subsets with h observations, and compute the regression coefficients, say $\beta_1, \dots, \beta_{\binom{n}{h}}$.

3.

Compute the residuals using all n observations, and from this the LTS criterion. 4. The LTS estimate corresponds to the β_{i^*} that minimizes the objective function (3). Rousseeuw and van Driessen [11] propose a fast algorithm for computing LTS. The trick is to iterate a few steps on a large number of starting values, and keep the 10 (say) most promising ones.

These are then used for full iteration, yielding the final estimate. The resulting algorithm makes LTS estimation faster. Our proposed procedure is a combination of IRLS and LTS methods. IRLTS estimator involves computing the hyperplane that minimizes the sum of the smallest h squared residuals and use the weighted least square estimation for β in each iteration. To motivate our estimator and following the fast LTS algorithm [11], it is convenient to write IRLTS algorithm with involving the residuals as follow.

Concentration-step: 1. Choose h observations. 2. Compute $\hat{\beta}$ based on h observations using IRLS method. 3. Use the estimate $\hat{\beta}$ to calculate residuals: $e_i = y_i - \hat{\beta}'x_i$ based on equation $\hat{\beta} = (X_h'X_h)^{-1}X_h'y_h$ of h observations. 4. Sort $|e_i|$ for $j = 1, \dots, n$, $i = 1, \dots, n$ in ascending order: $|e_{(1)}| \leq |e_{(2)}| \leq \dots \leq |e_{(n)}|$. 5.

Choose h observations which have the lowest h residuals, we denote the h observations as subset H . The repetitions of concentration-step will produce an iteration process.

IRLTS algorithm: 1. Choose h observations. 2. Compute $\hat{\beta}$ based on h observations using IRLS by (2), we obtain $\hat{\beta}^{(1)}$. 3. Calculate residuals: $e_i = y_i - \hat{\beta}^{(1)T} x_i$ of n observations. 4.

Sort $|e_i|$ in ascending order: $|e_{(1)}| \leq |e_{(2)}| \leq \dots \leq |e_{(n)}|$. 5. Choose h_1 observations which have the lowest h_1 residuals, we denote as subset H_1 . 6. Run concentration-step on H_1 twice, and we obtain H_1^* . 7. Repeat step 1- step 6 for h times 3 INSIST Vol. 2 No. 1, April 2017 (1 - 5) 8. From the h results, choose the best 10 subsets H_q , $q=1, \dots, 10$. 9. Run concentration-step on the best 10 subsets H_q until convergence. 10.

Choose the best subset H . IV. SIMULATION STUDY To look at the performance of the proposed method, we have done a simulation study by generating $N=1000$ observations from 200 subjects with 5 repeated measures. The model for data generation is as follows: $u_{ij} = \beta_0 + \beta_1 x_{ij}$ where $\beta_0 = \beta_1 = 1$, $i=1, 2, \dots, 200$ and $j=1, 2, \dots, 5$. The covariates x_{ij} are i.i.d.

from a uniform distribution $Unif(1,5)$. For this longitudinal data the normal distributed model is used. We generated data based on the underlying true correlation structures as exchangeable (EXCH) and autoregressive-1 (AR1) with $\alpha=0.3$ and 0.7 . We considered data without outliers ($e = 0\%$) as well as contaminated data ($e = 10\%, 20\%$ and 30%).

The contamination is generated from normal distribution $N(100,1)$, we set two cases for the contamination, i.e. randomly spread over the sample (case A) and randomly spread over the half upper x_{ij} values of the sample (case B). For each scenario 1000 Monte Carlo data sets were generated.

We evaluated the results using relative efficiency (RE) of IRLTS to IRLS and the mean square error (MSE) of $\hat{\beta}$ which we define as $\frac{MSE_{IRLTS}}{MSE_{IRLS}} = \dots$

!" and $\#\$ = " " " " ? ' _ ($
 $- _ * +$, with $0 = 0,1 " " " (2"$, where $Var(.)$ is the variance. We provide the expected values (E), and the relative efficiency resulted from our simulation in Table I - Table IV and the MSEs in Table V- Table VI.

The efficiency of IRLTS and IRLS for clean data (i.e. when $e = 0\%$) is almost equal since $RE \sim 1$ for each case, but IRLTS is more efficient than IRLS when data contain outliers. The parameter estimates of IRLS are much more influenced by the outliers than the parameter estimates of IRLTS.

From the expected values we can see that the more outliers contained in the data the larger the deviation of IRLS estimates from the parameter (i.e. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} b & b \\ b & b \end{pmatrix}$), while the parameter estimates of IRLTS are almost stable and close to the parameter. Table 1. Simulation Result for Longitudinal Data with Exchangeable Correlation Matrix with 3.0 = a Case Coeff.

$e) \wedge (IRLS \begin{pmatrix} b \\ b \end{pmatrix}) \wedge (IRLTS \begin{pmatrix} b \\ b \end{pmatrix}) RE (IRLTS/IRLS)$ Case A $0 \wedge b$ 0% 1.00500 1.00499 1.00178
 10% 10.08781 1.02474 0.00294 20% 17.55050 1.08504 0.00401 30% 23.73423 1.15940
 0.00657 $1 \wedge b$ 0% 0.99846 0.99848 1.00278 10% 0.99604 0.99948 0.00267 20% 0.98073
 0.99775 0.00393 30% 0.94544 0.99577 0.00649 Case B $0 \wedge b$ 0% 1.02179 1.01050 0.99466
 10% -5.20456 1.00195 0.00577 20% -10.30572 0.95625 0.01606 30% -14.27400 0.84177
 0.03562 $1 \wedge b$ 0% 0.99318 0.99693 0.99369 10% 5.95079 0.98675 0.00483 20% 9.99122
 0.99425 0.05797 30% 13.21288 1.04650 0.11281 Table 2. Simulation Result for Longitudinal Data with Exchangeable Correlation Matrix with 7.0 = a Case Coeff.

$e) \wedge (IRLS \begin{pmatrix} b \\ b \end{pmatrix}) \wedge (IRLTS \begin{pmatrix} b \\ b \end{pmatrix}) RE (IRLTS/IRLS)$ Case A $0 \wedge b$ 0% 1.01266 1.01187 1.01763
 10% 9.93676 1.01672 0.00433 20% 17.31940 1.05446 0.00505 30% 23.68107 1.16184
 0.00769 $1 \wedge b$ 0% 0.99600 0.99613 1.01882 10% 1.04154 1.00103 0.00403 20% 1.05223
 1.00661 0.00478 30% 0.95609 0.99443 0.00770 Case B $0 \wedge b$ 0% 1.05174 1.01922 0.99052
 10% -5.20125 1.00911 0.00837 20% -10.28303 0.98818 0.00528 30% -14.25209 0.83684
 0.04304 $1 \wedge b$ 0% 0.98159 0.99245 0.99948 10% 5.93523 0.98438 0.00722 20% 9.96316
 0.97828 0.00489 30% 13.21954 1.05368 0.12544 4 INSIST Vol. 2 No. 1, April 2017 (1 - 5)
 Table 3.

Simulation Result for Longitudinal Data with Autoregressive-1 Correlation Matrix with 3.0 = a Case Coeff. $e) \wedge (IRLS \begin{pmatrix} b \\ b \end{pmatrix}) \wedge (IRLTS \begin{pmatrix} b \\ b \end{pmatrix}) RE (IRLTS/IRLS)$ Case A $0 \wedge b$ 0% 0.99848
 0.99744 1.03077 10% 10.04370 1.02103 0.00247 20% 17.60751 1.07994 0.00359 30%
 23.63332 1.15166 0.00621 $1 \wedge b$ 0% 1.00043 1.00079 1.04004 10% 1.00954 1.00069
 0.00217 20% 0.96174 0.99879 0.00356 30% 0.97699 0.99897 0.00625 Case B $0 \wedge b$ 0%
 1.01800 1.00963 1.00481 10% -5.40516 0.99210 0.00113 20% -10.07597 0.94809 0.00393
 30% -14.18714 0.85628 0.05029 $1 \wedge b$ 0% 0.99341 0.99621 1.00108 10% 5.98712 0.99180
 0.00541 20% 9.89834 0.99916 0.01079 30% 13.19889 1.04016 0.16313 Table 4.
 Simulation Result for Longitudinal Data with Autoregressive-1 Correlation Matrix with 7.0 = a Case Coeff.

$e) \wedge (IRLS \begin{pmatrix} b \\ b \end{pmatrix}) \wedge (IRLTS \begin{pmatrix} b \\ b \end{pmatrix}) RE (IRLTS/IRLS)$ Case A $0 \wedge b$ 0% 1.00024 1.00079 1.07947
 10% 9.93814 1.00872 0.00412 20% 17.52048 1.07392 0.00486 30% 23.77262 1.16751
 0.00790 $1 \wedge b$ 0% 0.99960 0.99939 1.06915 10% 1.04489 1.00404 0.00379 20% 0.98595
 1.00065 0.00481 30% 0.94520 0.99331 0.00785 Case B $0 \wedge b$ 0% 1.03666 1.01330 0.99589
 10% -5.18276 1.01138 0.00618 20% -10.37403 0.97097 0.00484 30% -14.34357 0.80114

0.04723 1 \hat{b} 0% 0.98765 0.99545 0.99343 10% 5.92778 0.98511 0.00464 20% 10.00611 0.98523 0.00461 30% 13.25732 1.07041 0.14581 The consistency of the estimators is assessed through their MSEs (see Table V and Table VI).

When data contain outliers, the MSEs of IRLTS are relatively small compared to the MSEs of the classical GEE (IRLS). From the result we conclude that IRLTS is robust against outliers. Table 5. Mean Square Error of Parameter Estimates for Data with Exchangeable Correlation Matrix Case Coeff. e 3.0 = a 7.0 = a IRLS IRLTS IRLS IRLTS Case A 0 \hat{b} 0% 0.01664 0.01666 0.03141 0.03194 10% 90.44780 0.02372 88.48682 0.03765 20% 291.41893 0.07749 283.90429 0.09182 30% 544.84787 0.20944 540.15839 0.22414 1 \hat{b} 0% 0.00165 0.00165 0.00299 0.00305 10% 0.87157 0.00233 0.94041 0.00378 20% 1.91855 0.00754 1.92692 0.00924 30% 3.08263 0.02001 2.82052 0.02174 Case B 0 \hat{b} 0% 0.01726 0.01680 0.03467 0.03206 10% 41.61809 0.01801 41.89402 0.02886 20% 136.26362 0.13754 133.06421 0.03051 30% 240.63699 0.28655 240.56976 0.36852 1 \hat{b} 0% 0.00171 0.00167 0.00346 0.00318 10% 24.88252 0.00197 24.75140 0.00309 20% 82.04616 0.06984 80.97956 0.00361 30% 149.95039 0.09196 150.19042 0.11242 Table 6.

Mean Square Error of Parameter Estimates for Data with Autoregressive-1 Correlation Matrix Case Coeff. e 3.0 = a 7.0 = a IRLS IRLTS IRLS IRLTS Case A 0 \hat{b} 0% 0.01309 0.01349 0.02329 0.02515 10% 89.85427 0.02033 87.75463 0.03249 20% 292.90715 0.06774 289.43715 0.08572 30% 539.61490 0.19286 543.38625 0.22402 1 \hat{b} 0% 0.00131 0.00136 0.00231 0.00247 10% 0.88483 0.00192 0.86341 0.00328 20% 1.85020 0.00659 1.77790 0.00856 30% 2.98370 0.01864 2.71468 0.02134 Case B 0 \hat{b} 0% 0.01354 0.01337 0.02704 0.02577 10% 87.20910 0.05218 42.39676 0.02589 20% 164.20519 0.16609 135.05622 0.02836 30% 237.06421 0.34325 243.38623 0.41559 1 \hat{b} 0% 0.00135 0.00133 0.00273 0.00258 10% 27.51679 0.01437 24.80085 0.00262 20% 86.50801 0.07905 81.76371 0.00323 30% 149.53554 0.11950 151.10469 0.13076 31 INSIST Vol. 2 No. 1, April 2017 (1 - 5) V.

CONCLUSION Our proposed method have two different iterations in its procedure, one is the iteration for the estimation of regression parameter using IRLS method, and the other iteration is for selecting the best subset H for calculating the parameter estimate. We have shown that this procedure can minimize the effect of outliers on parameter estimation; IRLTS can produce a relatively efficient and consistent estimator compared to the classical GEE (IRLS). Base on the MSE, IRLTS performs much better than the classical GEE.

Hence, robust GEE using IRLTS is a good choice for longitudinal data analysis when data contains outliers. REFERENCES [1] P. McCullagh and J.A. Nelder, Generalized Linear Models . London : Chapman and Hall, 1989. [2] K.Y. Liang and S.L. Zeger, "Longitudinal

data analysis using generalized linear models", Biometrika , vol 73, pp. 13-22, 1986. [3] B.F. Qaqish and J.S.

Preisser, "Resistant fits for regression with correlated outcomes an estimating equations approach", Journal of Statistical Planning and Inference , vol. 75, pp. 15-431, 1999. [4] P.S. Gill, "A Robust Mixed Linear Model Analysis for Longitudinal Data", Statistics in Medicine , vol. 19, pp. 975-987, 2000. [5] S.H. Jung and Z. Ying, "Rank-Based Regression With Repeated Measurements Data", Biometrika , vol. 90, pp. 732-740, 2003. [6] Abebe, A., McKean, J. W. & Klope, J. D.,

Bilgic, Y. "Iterated Reweighted Rank-Based Estimates for GEE Models", submitted. [7] R.J. Rousseeuw and A.M. Leroy, Robust Regression and Outlier Detection . New York: Wiley, 1987. [8] J.A. Víšek, "The Least Trimmed Squares – random carriers". Bulletin of the Czech Econometric Society , vol. 6, pp. 1-30, 1999. [9] J.A. Víšek, The Least Trimmed Squares. Part II: v_{∞} -consistency. Kybernetika, vol. 42, pp. 181-202, 2006a. [10] J.A. Víšek, "The Least Trimmed Squares.

Part III: Asymptotic normality". Kybernetika, vol. 42, pp. 203-224, (2006b). [11] P.J. Rousseeuw, and K. van Driessen, " Computing LTS Regression for Large Data Sets". Data mining and Knowledge Discovery , vol. 12, pp. 29-45, 2006.

INTERNET SOURCES:

4% - <http://insist.unila.ac.id/index.php/ojs/article/view/23>

<1% -

<http://doczz.net/doc/610334/computational-and-financial-econometrics--cfe-10--comp>
uting

3% -

https://www.researchgate.net/publication/5207255_Rank-based_regression_with_repeated_measurements_data

<1% - <https://www.sciencedirect.com/science/article/pii/S1083879115018728>

1% -

https://www.researchgate.net/publication/2391452_Review_of_Software_to_Fit_Generalized_Estimating_Equation_Regression_Models

<1% -

https://www.researchgate.net/post/How_is_logistic_regression_used_What_conditions_and_types_of_variables_should_be_used

<1% -

<https://stats.idre.ucla.edu/sas/whatstat/what-statistical-analysis-should-i-usestatistical-analyses-using-sas/>

1% - <https://www.sciencedirect.com/science/article/pii/S0167947304002063>
1% -
https://www.researchgate.net/publication/308391249_Iterated_Reweighted_Rank-Based_Estimates_for_GEE_Models
<1% -
https://www.researchgate.net/publication/45870883_Tuning_parameter_selection_for_penalized_likelihood_estimation_of_inverse_covariance_matrix
<1% - <https://meloun.upce.cz/docs/publication/224-manus.pdf>
<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1226198/>
<1% -
https://www.researchgate.net/publication/271820538_Outlier_detection_and_robust_estimation_in_linear_regression_models_with_fixed_group_effects
<1% -
https://www.academia.edu/5366782/Effect_of_outliers_on_estimators_and_tests_in_covariance_structure_analysis
<1% - <https://www.sciencedirect.com/science/article/pii/S0096300314010509>
<1% -
<http://www.worldacademicunion.com/journal/1746-7659JIC/jicvol12no2paper04.pdf>
<1% - <https://www.sciencedirect.com/science/article/pii/S0167947317300178>
<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3252624/>
<1% - <https://www.scribd.com/document/328541591/Anton-Calculus-A-New-Horizon2>
<1% - https://archive.org/stream/gezhi/9_djvu.txt
<1% -
https://www.wyzant.com/resources/lessons/math/elementary_math/measurement
<1% -
https://mafiadoc.com/mechanics-in-phase-space-an-overview-with-_59c328211723ddd7d9bfc3e8.html
1% - <https://cran.r-project.org/web/packages/pscl/vignettes/countreg.pdf>
<1% - <https://m.open-open.com/pdf/a9513113c92944df85294ef748ad7faa.html>
<1% -
https://www.researchgate.net/publication/257338412_Functional_partially_linear_quantile_regression_model
<1% -
<https://epdf.pub/advances-on-theoretical-and-methodological-aspects-of-probability-and-statistics.html>
<1% - <https://www.sciencedirect.com/science/article/pii/0304393294011819>
<1% - <https://www.stern.nyu.edu/~wgreene/panel.doc>
<1% -
<https://nuit-blanche.blogspot.com/2012/09/compressive-sensing-and-advanced-matrix.html>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167947398000395>

<1% - <http://www.kaukauna.k12.wi.us/district/staff/2019-employee-benefits-guide.pdf>

1% - <https://www.lexjansen.com/nesug/nesug96/NESUG96107.pdf>

<1% - <http://ufdc.ufl.edu/AA00003585/00001>

<1% - <https://epdf.pub/generalized-linear-models-with-applications-in-engineering-and-the-sciences-wile.html>

<1% - <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD001293.pub3/full>

<1% - <http://counsel.kcue.or.kr/data/2017/175-1-1.pdf>

<1% - https://www.researchgate.net/publication/316709356_Audio_Data_Mining_for_Anthropogenic_Disaster_Identification_An_Automatic_Taxonomy_Approach

<1% - https://www.researchgate.net/publication/284946872_Comparison_of_Robust_Regression_Methods_in_Linear_Regression

<1% - <http://accounts.google.com/ServiceLogin?hl=it&service=mail>

<1% - https://www.academia.edu/32202817/Robust_methods_for_multivariate_data_analysis

<1% - https://www.researchgate.net/publication/273311220_Robust_Weighted_Least_Squares_Estimation_of_Regression_Parameter_in_the_Presence_of_Outliers_and_Heteroscedastic_Errors

1% - <http://iacmc.zu.edu.jo/ar/images/stories/IACMC2016/39.pdf>

<1% - https://www.academia.edu/15177828/Robust_estimators_in_semiparametric_partly_linear_regression_models

1% - <https://pdfs.semanticscholar.org/d2fd/55fc44935e2342250b8c21f13f3aeddeab0d.pdf>

1% - https://www.researchgate.net/publication/289667177_The_least_trimmed_squares_part_I_n-consistency

<1% - https://www.researchgate.net/publication/257564394_Active_learning_of_user's_preferences_estimation_towards_a_personalized_3D_navigation_of_geo-referenced_scenes

<1% - <https://www.science.gov/topicpages/l/laht+monte+carlo>

<1% - <http://mail.yahoo.com/>

<1% - https://www.researchgate.net/profile/Mia_Hubert/publication/257147083_Recent_developments_in_PROGRESS/links/0c960524951ac4cb86000000/Recent-developments-in-PROGRESS.pdf

<1% -

<https://epdf.pub/introduction-to-multivariate-statistical-analysis-in-chemometrics.html>

<1% - <https://www.science.gov/topicpages/c/curve+resolution+quantification.html>

<1% -

https://www2.cs.siu.edu/~qcheng/featureselection/Fisher_Markov_Final_minorRevision.pdf

<1% -

<http://metatoc.com/journals/1357-statistical-methods-in-medical-research-an-international-review-journal>

<1% - <https://www.sciencedirect.com/science/article/pii/S0378475409002882>

<1% -

https://www.academia.edu/23030159/Robust_Hypothesis_Testing_via_Lq-Likelihood

<1% - <https://quizlet.com/18153247/cph-exam-biostatistics-flash-cards/>

<1% - <http://support.sas.com/resources/papers/proceedings09/199-2009.pdf>

<1% -

https://www.researchgate.net/publication/228406284_Microfinance_A_Comprehensive_Review_of_the_Existing_Literature

<1% -

https://www.researchgate.net/publication/260758235_Shape_outlier_detection_and_visualization_for_functional_data_The_outliergram

<1% -

https://www.academia.edu/21510077/Marginal_Correlation_in_Longitudinal_Binary_Data_Based_on_Generalized_Linear_Mixed_Models

<1% -

<https://epdf.pub/time-series-data-analysis-using-eviews-statistics-in-practice.html>

<1% -

https://www.researchgate.net/publication/3481524_A_Sparsity-Based_Method_for_the_Estimation_of_Spectral_Lines_From_Irregularly_Sampled_Data

<1% - <http://www.ieomsociety.org/ieom2019/papers/538.pdf>

<1% -

https://www.researchgate.net/publication/225256092_Deletion_diagnostics_for_marginal_mean_and_correlation_model_parameters_in_estimating_equations

1% - https://link.springer.com/chapter/10.1007/978-3-319-39065-9_4

<1% - <https://www.sciencedirect.com/science/article/pii/S0378375814000573>

1% - <https://www.sciencedirect.com/science/article/pii/S0167947310001623>



Plagiarism Checker X Originality Report

Similarity Found: 25%

Date: Selasa, Agustus 27, 2019

Statistics: 1062 words Plagiarized / 4175 Total words

Remarks: Medium Plagiarism Detected - Your Document needs Selective Improvement.

www.arpnjournals.com ON GENERALIZED VARIANCE OF NORMAL-POISSON MODEL AND POISSON VARIANCE ESTIMATION UNDER GAUSSIANITY Khoirin Nisa^{1,*}, Célestin C. Kokonendji², Asep Saefuddin³, Aji Hamim Wigena³, I Wayan Mangku⁴ 1 Department of Mathematics, University of Lampung, Bandar Lampung, Indonesia 2Laboratoire 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 * Corresponding author: 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 $X = (X_1, \dots, X_k)^T$ where X_j is a univariate (non-negative) stable Tweedie variable and $(X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_k)^T =: X_j^\perp$ given X_j are $k-1$ real independent Gaussian variables with variance X_j , for any fixed $j \in \{1, 2, \dots, k\}$. 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_1 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 $j \in \{1, 2, \dots, k\}$, 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, \dots, 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 models for all $j \in \{1, 2, \dots, k\}$ and fixed positive integer $k > 1$ is defined as follows. **Definition 1.** For $X = (X_1, \dots, X_k)^T$ a k -dimensional normal-Poisson random vector, it must hold that X_j is a univariate Poisson random variable, and $X_j | X_{-j} := (X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_k)^T$ given X_j follows the $(k-1)$ -variate normal $N_{k-1}(0, X_j I_{k-1})$ distribution, where $I_{k-1} = \text{diag}(1, \dots, 1)$ 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 (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 $j \in \{1, 2, \dots, k\}$, denote $F_{t,j} = F((t)_j)$ the multivariate NEF of normal-Poisson with $(t) := (t^*)$, the NEF of a k -dimensional normal-Poisson random vector X is generated by $\int_A 1_A(x) dF_{t,j}(x)$ where 1_A is the indicator function of the set A . Since $t > 0$ then $(t)_j$ 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 $(t)_j$, i.e. $\psi_j(t) = \log \int \exp(t^T x) dF_{t,j}(x)$ and the probability distribution of normal-Poisson which is a member of NEF is given by $P_j(t) = \exp(\psi_j(t)) / \int \exp(t^T x) dF_{t,j}(x)$. 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.: $\mu_j = \psi_j'(t)$ and $\Sigma_j = \psi_j''(t)$.

For practical calculation we need to use the following mean parameterization: $P_j(t) = \exp(\psi_j(t)) / \int \exp(t^T x) dF_{t,j}(x)$, where μ_j is the solution in \mathbb{R}^k of the equation $\psi_j'(\mu_j) = t$. Then for a fixed $j \in \{1, 2, \dots, k\}$, the variance function (i.e. the variance-covariance matrix in term of mean parameterization) is given by $\Sigma_j(\mu_j) = \psi_j''(\mu_j) = I_{k-1} + \text{diag}(\mu_j^{-1}, \dots, \mu_j^{-1})$ (2) on its support $\mathcal{S}_j = \{\mu \in \mathbb{R}^k; \mu_j > 0 \text{ and } \mu_l \in \mathbb{R} \text{ for } l \neq j\}$.

$I_j\}$.

(3) For $j = 1$, the covariance matrix of X can be expressed as follows: / Indeed, for the covariance matrix above one can use the Schur complement [18] of a matrix block to obtain the following representation of determinant / with the non-null scalar $(= (1,$ the vector $aT=((2, ..., (k)$ and the $(k-1) \times (k-1)$ matrix $A = (-1aaT + (1lk-1,$ where $lj= =diag(1, ..., 1)$ is the $j \times j$ unit matrix.

Consequently, the determinant of the covariance matrix for $j = 1$ is $\det ?? ?? ??;1 ?? = ?? 1 ??$ with $???? ?? ??;1$ Then, it is trivial to show that for $j (\{1, 2, ..., k\}$ the generalized variance of normal-Poisson j model is given by $\det ?? ?? ??;1 ?? = ?? ?? ??$ with $???? ?? ??;??$ (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 $k (\{2,3, ... \}$ and $t>0$. If an NEF $Ft;j$ satisfies (2) for a given $j (\{1,2, ...,k\}$, then up to affinity, $Ft;j$ is a normal-Poisson j model. **Theorem 2** Let $Ft;j=F((???;??)$ be an infinitely divisible NEF on $R_k (k>1)$ such that $?? (???;?? = ?? ??$ and $\det ?? " (???;?? ?? = ?? \exp(?? \times ?? T ?? ?? ??)$ for $?? =((1, ..., (k)T$ and $?? ?? ?? =((1,..., (j-1, 1, (j+1, ..., (k)T$. Then $Ft;j$ 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 $($ and $($ be two (-finite positive measures on R_k such that $F=F(()$, $?? =??(()$ and $\mu(MF)$. (i) If there exists $(d,c)(?? ?? \times ??$ such that $((????) = \exp (dT_x+c)((dx)$, then $??= ?? : ?? ?? = ?? ?? -??$ and $?? ?? ?? = ?? ?? ??+?? +??$; for $?? =?? ? ?? ?? , ?? ?? ?? = ?? ?? (??)$ and $\det ?? ?? ?? = \det ?? ?? ??$.$

(ii) If $?? =??*??$ with $?? ?? =????+??$, then: $?? ?? = ?? T ?? ??$ and $?? ?? ?? = ?? ?? ?? T ?? + ?? T ??$; for $?? =????+?? ? ?? ?? ?? , ?? ?? (??)=?? ?? ?? ?? -1 ?? ?? T$ and $\det ?? ?? ?? = \det ?? ?? \det ?? ?? (??)$. (iii) If $?? = ?? *$ is the t -th convolution power of $??$ for $t>0$, then, for

$\mu_j = \sum_{k=0}^{\infty} \frac{t_k}{k!} \mu_j^{(k)}$, $\mu_j^{(k)} = \frac{1}{k!} \frac{d^k}{dt^k} \mu_j(t)$ and $\det \mu_j^{(k)} = \det \mu_j^{(k)}$. 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 the proof is provided in Proposition 4 below.
Proposition 4 If μ is an infinitely divisible measure on \mathbb{R}^k , then there exist a symmetric non-negative definite $d \times d$ matrix Σ with rank $r(\Sigma)$ and a positive measure ν on \mathbb{R}^k such that $\mu(t) = \exp\left(-\frac{1}{2} t^T \Sigma t + \int_{\mathbb{R}^k} (e^{it \cdot} - 1) \nu(dt)\right)$. See, e.g. [20, page 342].

The above expression of $\mu(t)$ is an equivalent of the Lévy-Khinchine formula [17]; thus, Σ comes from a Brownian part and the rest ν corresponds to jumps part of the associated Lévy process through the Lévy measure ν . **Generalized Variance Estimators** Let X_1, \dots, X_n be random vectors i.i.d.

with distribution $P(\cdot; F_{t,j})$ in a normal-Poisson $_j$ model $F_{t,j} = F((t,j))$ for fixed $j \in \{1, 2, \dots, k\}$. Denoting $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = (X_1, \dots, X_k)^T$ the sample mean. **Maximum Likelihood Estimator** The ML generalized variance estimator of normal Poisson model $\det \mu_j^{(k)}(\bar{X})$ is given by $\hat{\mu}_j^{(k)}(\bar{X}) = \det \mu_j^{(k)}(\bar{X})$. (6) The ML estimator (6) is directly obtained from (5) by substituting μ_j with its ML estimator $\hat{\mu}_j$.

For all $p=1, \dots, n$, t_j is a biased estimator of $\det \mu_j^{(k)}$ with a given quadratic risk with tedious calculation of explicit expression or infinite. **Uniformly Minimum Variance Unbiased Estimator** The UMVU generalized variance estimator of normal Poisson models $\det \mu_j^{(k)}$ is given by $\hat{\mu}_j^{(k)}(\bar{X}) = \det \mu_j^{(k)}(\bar{X})$ if $\mu_j^{(k)}(\bar{X}) = \det \mu_j^{(k)}(\bar{X})$. (7) The UMVU estimator of $\det \mu_j^{(k)}$ is deduced by using intrinsic moment formula of univariate Poisson distribution as follows $\det \mu_j^{(k)}(\bar{X}) = \det \mu_j^{(k)}(\bar{X})$. Indeed, letting $\mu_j^{(k)}(\bar{X}) = \det \mu_j^{(k)}(\bar{X})$ gives the result that (7) is the UMVU estimator of (5).

Because, by the completeness of NEF, the unbiased estimator is unique. **Bayesian Estimator** Under assumption of prior gamma distribution of μ_j with parameter $(\alpha > 0$ and $\beta > 0$, the Bayesian estimator of $\det \mu_j^{(k)}$ is given by $\hat{\mu}_j^{(k)}(\bar{X}) = \det \mu_j^{(k)}(\bar{X})$.

(8) To show this, let X_1, \dots, X_n given μ_j be Poisson distribution with mean μ_j , then the probability mass function is given by $P(X_i = x_i | \mu_j) = \frac{\mu_j^{x_i}}{x_i!} \exp(-\mu_j)$. Assuming that μ_j follows gamma (α, β) , then the prior probability distribution function of μ_j is written as $\pi(\mu_j) = \frac{\beta^\alpha}{\Gamma(\alpha)} \mu_j^{\alpha-1} \exp(-\beta \mu_j)$, $\alpha > 0$ with $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} \exp(-t) dt$.

Using the classical Bayes theorem, the posterior distribution of μ_j given an observation x_{ji} can be expressed as $\pi(\mu_j | x_{ji}) \propto \pi(\mu_j) \pi(x_{ji} | \mu_j)$, $\pi(\mu_j) > 0$, $\pi(\mu_j) \propto \mu_j^{\alpha-1} \exp(-\beta \mu_j)$, $\pi(x_{ji} | \mu_j) = \frac{\beta^{\mu_j}}{\Gamma(\mu_j)} \mu_j^{\mu_j-1} \exp(-\beta x_{ji})$, which is the gamma density with parameters $\alpha = \beta + x_{ji}$, $\beta = \beta + 1$. Then with random sample X_{j1}, \dots, X_{jn} the posterior will be $\text{gamma}(\beta + \sum_{i=1}^n x_{ji}, \beta + n)$.

Since Bayesian estimator of μ_j is given by the expected value of the posterior distribution i.e. $E(\mu_j | x_{ji})$, then this will lead to (8). **MAIN RESULT** **Poisson Variance Estimation under Gaussianity** For a given random vector $X = (X_1, \dots, X_k)^T$ on R_k of normal-Poissonj, we now assume that only $k-1$ normal terms of X are observed: X_1, \dots, X_{k-1} and, therefore, X_j is an unobserved Poisson random effect. Note that j is fixed in $\{1, 2, \dots, k\}$.

Assuming $t=1$ and following [1] with X having mean vector $\mu = (\mu_1, \dots, \mu_k)^T$ ($\mu_j > 0$) and covariance matrix $V = V(\mu)$, then X follows a $(k-1)$ -variate normal distribution, denoted by $X \sim N_{k-1}(\mu, V)$, (9) with $\mu = (\mu_1, \dots, \mu_{k-1}, \mu_k)^T$. The $(k-1) \times (k-1)$ -matrix V (which does not depend on μ) is symmetric and positive definite such that $\det V = 1$ or $\det V = \sigma^2$.

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 $X = (\mu_1, \dots, \mu_k)^T$ is easily deduced as $\det V$. 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]: $/$ or $/$ with $\mu = (\mu_1, \dots, \mu_k)^T$ and $I = (I_1 + \dots + I_k)/k$.

This statistical aspect of normal-Poissonj models in (9) points out the flexibility of these models compared with the classical multivariate normal model $N_{k-1}(\mu, V)$, where the generalized variance $\det V$ is replaced to the random effect μ_j . In fact, for $\mu_k = \mu_{k-1}$ in (9) with estimation μ_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 $k=3$) 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 μ_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. $\mu_j = 0.5, 1, 5$, 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 $MSE(\hat{\mu}_j) = [E(\hat{\mu}_j) - \mu_j]^2 + Var(\hat{\mu}_j)$, where $E(\hat{\mu}_j) = \frac{1}{1000} \sum_{i=1}^{1000} \hat{\mu}_j^{(i)}$ and $Var(\hat{\mu}_j) = \frac{1}{999} \sum_{i=1}^{1000} (\hat{\mu}_j^{(i)} - E(\hat{\mu}_j))^2$. We report the expected values and MSE of $\hat{\mu}_j$ in Table 1 - Table 3. 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 $\mu_j = 0.5$. $k \quad n \quad E(\hat{\mu}_j) \quad MSE(\hat{\mu}_j)$

3	30	0.473270	0.039251
3	50	0.487402	0.023864
3	100	0.491117	0.010882
3	300	0.495814	0.004058
3	500	0.496612	0.002540
3	1000	0.499035	0.001158
4	30	0.465915	0.031980
4	50	0.488503	0.019574
4	100	0.491804	0.009975
4	300	0.494617	0.003457
4	500	0.496200	0.002019
4	1000	0.498271	0.000968
6	30	0.452953	0.026781
6	50	0.478994	0.015763
6	100	0.483284	0.007801
6	300	0.495324	0.002713
6	500	0.496771	0.001562
6	1000	0.497542	0.000771
8	30	0.454636	0.023539
8	50	0.468367	0.014280
8	100	0.482915	0.007374
8	300	0.495749	0.002395
8	500	0.499078	0.001542
8	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 $\mu_j = 1$. $k \quad n \quad E(\hat{\mu}_j) \quad MSE(\hat{\mu}_j)$

3	30	0.962617	0.095854
3	50	0.983720	0.055901
3	100	0.993564	0.029386
3	300	0.994837	0.010214
3	500	0.997781	0.005969
3	1000	0.998467	0.003125
4	30	0.955849	0.078891
4	50	0.973454	0.049405
4	100	0.981452	0.023848
4	300	0.992874	0.007467
4	500	0.996215	0.004848
4	1000	1.001149	0.002456

k \ n	30	50	100	300	500	1000
3	0.944165	0.972215	0.985437	0.992045	0.995822	0.998113
4	0.944031	0.973103	0.981169	0.992240	0.998451	0.999042
6	0.944165	0.972215	0.985437	0.992045	0.995822	0.998113
8	0.944031	0.973103	0.981169	0.992240	0.998451	0.999042

From the results in the tables we can see that when the sample size (n) increases, the expected values of μ_j converge to the target value (μ_j) for all μ_j values we consider here. Also, the MSE of μ_j decrease when sample size increase for all dimension k, this means that μ_j is consistent.

The simulation results with moderate sample sizes produce very good performances of μ_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 μ_j , we provide the bargraphs of MSE of μ_j in Figure 1 - Figure 3.

The figures show that MSE value decrease when the sample size increase. From the result we conclude that μ_j is a consistent estimator of μ_j . Notice that μ_j produce smaller MSE for larger dimension. / Figure 1. Bargraph of MSE(μ_j) for $\mu_j = 0.5$ Table 3. The Expected Values and MSE of μ_j with 1000 replications for n ({30,50,100,300,500,1000}, k ({3,4,6,8} , and $\mu_j = 5$.

k \ n	30	50	100	300	500	1000
3	4.886415	4.942883	4.984949	4.987437	4.992459	4.998113
4	4.856583	4.921017	4.950201	4.983517	4.988398	4.998113
6	4.852608	4.926390	4.942147	4.974067	4.995231	4.998113
8	4.838751	4.910668	4.949142	4.985705	4.990750	4.998113

/ Figure 2. Bargraph of MSE(μ_j) for $\mu_j = 1$ / Figure 3.

Bargraph of MSE(μ_j) for $\mu_j = 5$ CONCLUSION 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 μ_j is a consistent estimator of the Poisson variance. REFERENCES Y.

Boubacar Mainassara and C.C. Kokonendji, "Normal stable Tweedie models and

power-generalized variance function of only one component", TEST, vol. 23, pp. 585-606, 2014. J. M. Bernardo and A.F.M. Smith, Bayesian Theory, New York: Wiley, 1993. O.E. Barndorff-Nielsen, J. Kent and M. Sørensen, "Normal variance-mean mixtures and z distributions", International Statistical Review, vol. 50, pp. 145-159, 1982. C. C. Kokonendji and A.

Masmoudi, "A characterization of Poisson Gaussian families by generalized variance", Bernoulli, vol. 12, pp. 371379, 2006. A. E. Koudou and D. Pommeret, "A characterization of Poisson-Gaussian families by convolution-stability", J. Multivariate Anal., vol. 81, pp. 120127, 2002. H. S. Steyn, "On the multivariate Poisson normal distribution", J. Am. Stat. Assoc., vol. 71, pp. 233-236, 1976. C. D. Kemp and A. W.

Kemp, "Some properties of the Hermite distribution", Biometrika, vol. 52, pp. 381-394, 1965. S. S. Wilks, Certain generalizations in the analysis of variance, Biometrika, vol. 24, pp. 471-494, 1932. 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. 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. M. Goodman, "A measure of overall variability in population", Biometrics, vol. 24, pp. 189-192, 1968. 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. 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. 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. 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. C.C. Kokonendji and K. Nisa, "Generalized variance in multivariate normal Poisson models and an analysis of variance", preprint of Laboratoire de mathématiques de Besançon, no. 2014/23, unpublished. K. Sato, Lévy Processes and Infinitely Divisible Distributions, Cambridge: Cambridge University Press, 1999. Y.

Zi-Zhong, "Schur complement and determinant inequalities", Journal of Mathematical Inequalities, vol. 3, pp. 161-167, 2009. 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. I. I. Gikhman and A. V. Skorokhod, The Theory of Stochastic Processes II, New York: Springer, 2004. A.

SenGupta, "Tests for standardized generalized variances of multivariate normal populations of possibly different dimensions", Journal of Multivariate Analysis, vol. 23, pp. 209-219, 1987.

INTERNET SOURCES:

<1% -
https://www.ut.ee/sites/default/files/ut_files/7d1699c4142a0b137759549e6548a633.vnd.ms-powerpoint

2% - http://www.arpnjournals.org/jeas/research_papers/rp_2017/jeas_0617_6137.pdf

3% -
https://www.researchgate.net/publication/272006326_On_normal_stable_Tweedie_models_and_power-generalized_variance_functions_of_only_one_component

8% - <http://lmb.univ-fcomte.fr/IMG/pdf/arpnjeas2017nketal.pdf>

<1% -
https://www.researchgate.net/publication/29641897_Generalized_variance_estimators_in_the_multivariate_gamma_models

<1% -
https://www.academia.edu/2919042/Computational_field_theory_and_pattern_formation

<1% - <https://www.sciencedirect.com/science/article/pii/S0377221719306241>

<1% -
https://www.academia.edu/24601963/DasGupta_A._Probability_for_statistics_and_machine_learning

1% -
https://www.researchgate.net/publication/23644434_A_Characterization_of_Poisson-Gaussian_Families_by_Convolution-Stability

<1% -
https://www.researchgate.net/publication/265856434_On_the_number_of_components_in_a_Gaussian_mixture_model

1% -
https://www.researchgate.net/publication/38348721_On_the_determinant_of_the_second_derivative_of_a_Laplace_transform

<1% - <http://ufdc.ufl.edu/AA00003537/00001>

<1% -
https://www.researchgate.net/publication/23644545_Descriptive_measures_of_multivariate_scatter_and_linear_dependence

<1% -

https://www.researchgate.net/profile/Fahim_Ashkar/publication/227159365_The_generalized_method_of_moments_as_applied_to_the_generalized_gamma_distribution/links/552fa34b0cf2acd38cbc249c.pdf?inViewer=0&pdfJsDownload=0&origin=publication_detail

<1% - <https://www.sciencedirect.com/science/article/pii/S0005109813000368>

<1% - https://en.m.wikipedia.org/wiki/Normal_distribution

1% -

https://www.researchgate.net/publication/232834424_Comparing_UMVU_and_ML_estimators_of_the_generalized_variance_for_natural_exponential_families

<1% -

https://www.researchgate.net/publication/269238659_Generalized_Variance_Estimations_of_Normal_Poisson_Models

1% -

https://answers.microsoft.com/en-us/windows/forum/windows8_1-update/definition-update-for-windows-defender-kb2267602/f023b4a3-bbba-472f-8d0b-3cad328de7ab

<1% - <https://studylib.net/doc/14809954/preface>

<1% - <https://arxiv.org/pdf/1510.05593.pdf>

<1% - <https://open.library.ubc.ca/handle/2429/44647>

<1% - https://github.com/obspy/obspy/blob/master/obspy/signal/array_analysis.py

<1% - <https://psychology.wikia.org/wiki/Cumulant>

<1% - https://en.wikipedia.org/wiki/Parametric_equation

<1% - <https://stats.stackexchange.com/q/155424>

<1% - <http://iopscience.iop.org/article/10.1088/1367-2630/10/8/083030>

<1% - <https://wenku.baidu.com/view/b0361709bb68a98271fefacb.html>

<1% -

https://www.academia.edu/29798232/A_Bayesian_Poisson_specification_with_a_conditionally_autoregressive_prior_and_a_residual_Moran_s_Coefficient_minimization_criterion_for_estimating_leptokurtic_distributions_in_regression-based_Multi-Drug_Resistant_Tuberculosis_treatment_protocols

<1% - https://link.springer.com/chapter/10.5176%2F2251-1911_CMCGS14.29_21

<1% - <https://www.sciencedirect.com/science/article/pii/S0167715218303092>

<1% -

https://www.researchgate.net/publication/38322468_A_characterization_of_Poisson-Gaussian_families_by_generalized_variance

<1% - <https://wenku.baidu.com/view/b2f1191fc281e53a5802ffa6.html>

<1% - <https://www.sciencedirect.com/science/article/pii/0022123672900882>

<1% - <https://www.math.upenn.edu/~ekorman/teaching/det.pdf>

<1% - https://archive.org/stream/arxiv-1202.3670/1202.3670_djvu.txt

<1% -

https://mafiadoc.com/fundamentals-of-linear-algebra_59c6485d1723dd830e59ffd1.html

<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3855520/>
<1% - <https://www.sciencedirect.com/science/article/pii/S0378437116002429>
<1% -
https://www.academia.edu/1771185/Robustification_Process_on_Bayes_Estimators
<1% - <https://www.sciencedirect.com/topics/mathematics/likelihood-equation>
<1% - <https://www.sciencedirect.com/science/article/pii/S147466701734750X>
<1% - <https://www.sciencedirect.com/science/article/pii/S0377042718302243>
<1% -
https://mafiadoc.com/introduction-to-time-series-and-forecasting_5bb64f3c097c47744a8b465b.html
<1% - <http://math.arizona.edu/~jwatkins/o-mle.pdf>
<1% - https://www.math.uh.edu/~jiwenhe/Math1432/lectures/lecture04_handout.pdf
<1% - <https://www.sciencedirect.com/science/article/pii/S1573441283010090>
<1% - <https://www.scribd.com/document/362571583/Astro-Stat-Book-of-Notes>
<1% -
http://www.dbs.ifi.lmu.de/Lehre/MaschLernen/SS2019/lecture/07_FrequentistBayesian.pdf
<1% - https://www.academia.edu/22874793/Introductory_Econometrics
<1% - https://en.wikipedia.org/wiki/Central_limit_theorem
<1% - https://en.wikipedia.org/wiki/Gamma_distribution
<1% - <https://www.sciencedirect.com/topics/engineering/deformation-gradient-tensor>
<1% -
https://www.researchgate.net/publication/26619988_A_simple_robust_control_chart_based_on_MAD
<1% - <https://www.sciencedirect.com/topics/engineering/conditional-distribution>
<1% -
https://www.academia.edu/1317892/Minimum_Hellinger_distance_estimation_for_Poisson_mixtures
<1% -
<https://math.stackexchange.com/questions/1577443/estimator-variance-for-l1-norm-and-l2-norm>
<1% - <https://www.eurasip.org/Proceedings/Ext/NSIP99/Nsip99/papers/18.pdf>
<1% -
<https://stats.stackexchange.com/questions/4700/what-is-the-difference-between-fixed-effect-random-effect-and-mixed-effect-mode>
<1% - https://issuu.com/rcopeland42/docs/probability_and_statistics_for_engineers
<1% - <https://www.sciencedirect.com/science/article/pii/S0169743906001419>
<1% - <https://itl.nist.gov/div898/handbook/eda/section3/eda355.htm>
<1% - <https://bmjopen.bmj.com/content/6/8/e010983>
<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3885826/>

<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2971698/>

<1% - http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapters1-12.pdf

<1% - <https://www.sciencedirect.com/science/article/pii/S0927539816300433>

<1% - <https://www.sciencedirect.com/science/article/pii/S037837580000207X>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167947307002903>

<1% - http://dept.stat.lsa.umich.edu/~kshedden/Courses/Stat406/Notes/Mean_Estimators.pdf

<1% - <https://www.sciencedirect.com/science/article/pii/S0304407601000732>

<1% - <https://quizlet.com/12855033/probability-statistics-flash-cards/>

<1% - <https://www.onlinelibrary.wiley.com/doi/10.1002/sim.6767>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167947316000232>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167668714001279>

<1% - https://www.researchgate.net/publication/270661878_Differential_effects_of_ThetaBeta_and_SMR_neurofeedback_in_ADHD_on_sleep_onset_latency

<1% - https://www.researchgate.net/publication/330983326_Estimation_of_entropy_for_inverse_Weibull_distribution_under_multiple_censored_data

<1% - http://docshare.tips/re-1989-11_59d8ab6408bbc52579a0db05.html

<1% - <https://www.bls.gov/osmr/research-papers/2002/pdf/st020120.pdf>

<1% - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2844672/>

<1% - <https://link.springer.com/article/10.1007/s10986-018-9406-3>

<1% - https://www.researchgate.net/publication/222216108_EM-estimation_and_modeling_of_heavy-tailed_processes_with_the_multivariate_normal_inverse_Gaussian_distribution

<1% - https://link.springer.com/chapter/10.1007/978-3-319-13881-7_39

<1% - <https://www.tandfonline.com/doi/full/10.1080/03610926.2016.1146770>

<1% - https://link.springer.com/chapter/10.1007/978-1-4614-1412-4_17

<1% - <https://link.springer.com/article/10.3103%2FS1066530708010055>

<1% - <https://www.sciencedirect.com/science/article/pii/S0167715210000441>

<1% - <http://www.mscs.mu.edu/~jsta/issues/11%284%29/JSTA11%284%29p3.pdf>

<1% - <http://lmb.univ-fcomte.fr/Chapitres-d-ouvrages>

<1% - <http://lmb.univ-fcomte.fr/Publications-310>

<1% - <https://www.sciencedirect.com/science/article/pii/S0377042715002356>

<1% - <https://link.springer.com/content/pdf/bfm%3A978-3-642-61943-4%2F1.pdf>

<1% - <https://www.sciencedirect.com/science/article/pii/0047259X87901539>

<1% - <https://wol.jw.org/en/wol/s/r1/lp-e?q=disaster%20prepared>