Global Journal of Mathematical Sciences: Theory and Practical.
0974-3200 Volume 3, Number 5 (2011), pp. 493–506

http://www.irphouse.com/gjms.htm

On Approximate Confidence Intervals for a Damage Total Based on the Bivariate SUR Model

Nuri Eltabit Mohamed

Institute for Mathematical Stochastic,
Magdeburg-Germany,
E-mail: nurieltab@yahoo.ca

Abstract

When the bivariate SUR model errors are non-normal one might be interested in making inference about a damage total for a certain product based on the 'Bivariate Seemingly Unrelated Regression model'. The goal of this article is to construct approximate asymptotic confidence intervals for a damage total based on the 'Bivariate SUR' model of paired count data. The simulated coverage is evaluated using computer simulation and recommendations are provided for selecting the appropriate confidence intervals which there coverage closed to the nominal value.

AMS Subject Classification: 62D05.
Keywords: SUR model, Multivariate Central limit theorem, Slutsky's lemma, Asymptotic distribution, Approximate confidence intervals.

1. Introduction

The SUR model was introduced and formulated by Arnold Zellner. Zellner (1962) formulated k Seemingly Unrelated Regression models as correlated regression equations. The Seemingly Unrelated Regression (SUR) model is a technique for analyzing a system of k equations with correlated error terms cross equations. The errors are uncorrelated across individuals $i = 1, \dots, m$, but they are correlated across equations for the same individual i.

It may be important to determine a reliable estimates for minimum damage totals, which result as the estimate damage value of weighted sum of estimates damage percentage for two components.

Simon (1996), have been proposed two methods for constructing confidence intervals for the binomial proportion p to achieve a closeness of the coverage of these intervals to the corresponding nominal value, the universal confidence interval by the z-quantiles and confidence interval by the t-quantiles, he concluded that the t-quantile method achieves better coverage than the z-quantile method. Recently, Santner (1998) and Agresti and Coull (1998), Brown, Cai, and DasGupta (1999) show that the performance of the standard universal interval are miserable and erratic than is appreciated and is poor in this problem that can not be trusted. Santner (1998) have been recommended an alternative method, asymptotic confidence intervals by the z-quantiles produced intervals guarantee achieving at least the nominal value of the true coverage no matter the sample sizes and p, remarkably accurate even at modest sample sizes, further, it is widely recognized that the coverage of the standard universal interval is poor for p near 0 or 1.

In this paper, we will use the both quantiles to construct approximate confidence intervals for a damage total for a certain product based on the bivariate SUR model, the asymptotic (by the z-quantile) and the conservative confidence (by the t-quantile) with sample size mines one degrees of freedom.

In a simulation study and for approximating the true coverage of a damage total we simulate pairs of bivariate Poisson count observations at different sample sizes and model parameter values, we have considered two confidence intervals, the asymptotic (with the normal quantile) and the proposed conservative (with the adjusted t-quantile on the reduced sample sizes mines one) confidence intervals, as well as we will compare the performance of these confidence intervals among them.

The remainder of the paper is organized as follows: General view on the involved models will be given in Sect. 2. The bivariate SUR model will be given in Sect. 3. Section 4 highlight on the estimation of the bivariate SUR parameters. While Section 5 considers the derivation of the asymptotic normal distribution of the Ratio estimators with help of the Slutsky's lemma. Section 6 provides approximate confidence intervals for the damage total. In Sect. 7 we perform a simulation study to compare the coverage of the confidence intervals at a specific nominal value. Finally, Sect. 8 gives some conclusions.

2. The models

Suppose, there are m observations of 4 relevant components of data (taking in account any vertical dependencies between the components), i.e., $((Y_{11}, X_{11}), (Y_{12}, X_{12})), \ldots,$ $((Y_{m1}, X_{m1}), (Y_{m2}, X_{m2}))$ is a random sample of m observations drawn from an infinite population, such that each observational unit indexed by the subscript i associated with those random variables under the restrictions $0 \leq Y_{ij} \leq X_{ij}, \forall i = 1, \ldots, m, j = 1, 2,$

For illustration, the i^{th} individual represents count variables for example:
$X_{i1}, X_{i2} \equiv$ No.of units respectively for a product i.
$Y_{i1}, Y_{i2} \equiv$ No.of damage units respectively for the product i, $i = 1, \ldots, m$.

Poisson and Binomial Models

(Y_{i1}, Y_{i2}) and (X_{i1}, X_{i2}) have a bivariate Poisson model with parameters $(\lambda_0, \lambda_1, \lambda_2)$ and $(\lambda_0, \lambda_1 + \mu_1, \lambda_2 + \mu_2)$ respectively, where $Y_{ij} \sim \text{Poiss}(\lambda_0 + \lambda_j)$, $Z_{ij} \sim \text{Poiss}(\mu_j)$, $X_{ij} = Y_{ij} + Z_{ij} \sim \text{Poiss}(\lambda_0 + \lambda_j + \mu_j)$,
$\text{Cov}(Y_{i1}, Y_{i2}) = \text{Cov}(X_{i1}, X_{i2}) = \lambda_0$, as well as all $(Y_j, Z_j), (Z_j, \mu_j)$ and $(Y_j, Z_{j'})$,$j \neq j' = 1, 2$ are independent, for more details about the bivariate Poisson model see [6] and also [7].
Further, $P(Y_{ij} \mid X_{ij}) \sim \text{Bin}(X_{ij}, p_j)$, where $p_j = \dfrac{\lambda_0 + \lambda_j}{\lambda_0 + \lambda_j + \mu_j}$ is the binomial proportion, and a sample observation $X_{ij} > 0$, $\forall i = 1, \ldots, m, j = 1, 2$ ($\text{Bin}(n, p)$ denotes the Binomial distribution with sample size n and success probability p). Also denote $\hat{p}_j = \dfrac{\sum_{i=1}^m Y_{ij}}{\sum_{i=1}^m X_{ij}}$ as the ratio corresponding to the proportion p_j, and $E(\hat{p}_j) = p_j = \dfrac{E(Y_j)}{E(X_j)}$, since Y_{ij}, X_{ij} are $i.i.d$, $i = 1, \ldots, m,\ j = 1, 2$.

The pairs of the univariate linear models will be considered first, as

$$Y_{i1} = X_{i1} p_1 + \epsilon_{i1},$$
$$Y_{i2} = X_{i2} p_2 + \epsilon_{i2},$$

with the following assumptions:
$E(\epsilon_{ij} \mid X_{ij}) = 0$, and with variances proportional to X_{ij}, i.e.,

$$\text{Var}(\epsilon_{ij} \mid X_{ij}) = \sigma_j^2 X_{ij},$$
$$\text{Cov}(\epsilon_{i1}, \epsilon_{i2} \mid X_{i1} X_{i2}) = \sigma_{12} \sqrt{X_{i1} X_{i2}},$$

and

$$\text{Cov}(\epsilon_{i1}, \epsilon_{i'2} \mid X_{i1} X_{i'2}) = 0, \quad \forall i \neq i',\ i, i' = 1, \ldots, m, j = 1, 2.$$

We merge these equations into a single linear model (for the i^{th} observation)

$$\mathbf{Y}_i = \mathbf{X}_i \mathbf{p} + \boldsymbol{\epsilon}_i,\ (i = 1, \ldots, m), \tag{2.1}$$

where, the response variable $\mathbf{Y}_i = (Y_{i1}, Y_{i2})^T$, the design matrix $\mathbf{X}_i = \begin{pmatrix} X_{i1} & 0 \\ 0 & X_{i2} \end{pmatrix}$, and the model coefficient $\mathbf{p} = \left(p_1, p_2\right)^T$, as well as the error component $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2})^T$. The error component $\boldsymbol{\epsilon}_i$ has the covariance matrix given by

$$\Sigma_i = \text{Cov}(\boldsymbol{\epsilon}_i) = \begin{pmatrix} \sigma_1^2 X_{i1} \sigma_{12} \sqrt{X_{i1} X_{i2}} \\ \sigma_{12} \sqrt{X_{i1} X_{i2}} \sigma_2^2 X_{i2} \end{pmatrix},$$

$\text{Cov}(\boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_{i'}) = \mathbf{0}$, when $i \neq i' = 1, \ldots, m$.

The linear model 2.1 will be standardized by the linear transformation $\mathbf{A}_i\mathbf{Y}_i = \mathbf{A}_i\mathbf{X}_i\mathbf{p} + \mathbf{A}_i\boldsymbol{\epsilon}_i$, to get the weighted linear model

$$\widetilde{\mathbf{Y}}_i = \widetilde{\mathbf{X}}_i\mathbf{p} + \widetilde{\boldsymbol{\epsilon}}_i, \tag{2.2}$$

where, the diagonal matrix

$$\mathbf{A}_i = \begin{pmatrix} \frac{1}{\sqrt{X_{i1}}} & 0 \\ 0 & \frac{1}{\sqrt{X_{i2}}} \end{pmatrix},$$

the weighed response variable $\widetilde{\mathbf{Y}}_i = \left(\widetilde{Y}_{i1}, \widetilde{Y}_{i2}\right)^T$, the weighted error component

$$\widetilde{\boldsymbol{\epsilon}}_i = \left(\widetilde{\epsilon}_{i1}, \widetilde{\epsilon}_{i2} \right)^T,$$

and the weighted design matrix

$$\widetilde{\mathbf{X}}_i = \begin{pmatrix} \widetilde{X}_{i1} & 0 \\ 0 & \widetilde{X}_{i2} \end{pmatrix}, \widetilde{X}_{ij} = \sqrt{X_{ij}},$$

and

$$\widetilde{Y}_{ij} = \frac{Y_{ij}}{\sqrt{X_{ij}}},$$
$$\widetilde{\epsilon}_{ij} = \frac{\epsilon_{ij}}{\sqrt{X_{ij}}},$$

given that $X_{ij} > 0$, $\forall i = 1, \ldots, m$, $j = 1, 2$. The covariance of the weighted error vector is given by

$$\widetilde{\Sigma} = \mathrm{Cov}(\widetilde{\boldsymbol{\epsilon}}_i) \quad = \mathbf{A}_i\mathrm{Cov}(\boldsymbol{\epsilon}_i)\mathbf{A}_i^T = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

and, $\mathrm{Cov}(\widetilde{\boldsymbol{\epsilon}}_i, \widetilde{\boldsymbol{\epsilon}}_{i'}) = \mathbf{0}$, $i \neq i' = 1, \ldots, m$, where

$$\sigma_j^2 = E\left(\widetilde{Y}_{ij} - \widetilde{X}_{ij}p_j\right)^2, \ \sigma_{12} = E\left((\widetilde{Y}_{i1} - \widetilde{X}_{i1}p_1)(\widetilde{Y}_{i2} - \widetilde{X}_{i2}p_2)\right), \quad j = 1, 2.$$

3. The SUR model

And thus, the SUR model (see also [11]) will written by stacking the model (2.1) in row-wise as

$$\widetilde{\mathbf{Y}} \quad = \widetilde{\mathbf{X}}\mathbf{p} + \widetilde{\boldsymbol{\epsilon}}, \tag{3.1}$$

where, the $2m \times 1$ dimension weighted response vector

$$\widetilde{\mathbf{Y}} = \left(\widetilde{\mathbf{Y}}_1^T, \widetilde{\mathbf{Y}}_2^T, \cdots, \widetilde{\mathbf{Y}}_m^T\right)^T,$$

$2m \times 2$ dimension weighted design matrix

$$\widetilde{\mathbf{X}} = \left(\widetilde{\mathbf{X}}_1^T, \widetilde{\mathbf{X}}_2^T, , \cdots, \widetilde{\mathbf{X}}_m^T\right)^T,$$

2×1 dimension model parameter vector $\mathbf{p} = (p_1, p_2)^T$, $2m \times 1$ dimension model error vector

$$\widetilde{\boldsymbol{\epsilon}} = \left(\widetilde{\boldsymbol{\epsilon}}_1^T, \widetilde{\boldsymbol{\epsilon}}_2^T, \cdots, \widetilde{\boldsymbol{\epsilon}}_m^T\right)^T$$

with $2m \times 2m$ covariance matrix

$$\widetilde{\boldsymbol{\Sigma}} = \mathrm{Cov}(\widetilde{\boldsymbol{\epsilon}}) = \begin{pmatrix} \widetilde{\Sigma} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \widetilde{\Sigma} \end{pmatrix} = I_m \otimes \widetilde{\Sigma},$$

i.e,

$$\mathrm{Var}(\widetilde{\boldsymbol{\epsilon}}_i) = \widetilde{\Sigma},$$

and

$$\mathrm{Cov}(\widetilde{\boldsymbol{\epsilon}}_i, \widetilde{\boldsymbol{\epsilon}}_i{}') = \mathbf{0}(\forall i \neq i',\ i, i' = 1, \ldots, m),$$

where $\mathbf{0}$ is the 2×2 dimension matrix of zero's.

4. Estimation in SUR model

It well-known that the WLSE is the BLUE of the parameter **p**, however, the WLSE of the SUR model parameter vector will not results in the sample ratio estimator vector in question (which is the OLSE, where the equality WLSE $\equiv$ OLSE holds if the error covariance matrices are diagonal [11]), therefore the OLSE will be used although it is not the optimal estimator nevertheless produces the ratio estimators.

The OLS estimator of the model proportion **p** will now be obtained from (3.1) as following

$$\begin{aligned} \hat{\mathbf{p}} &= (\widetilde{\mathbf{X}}^T\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}^T\widetilde{\mathbf{Y}} = \left(\sum_{i=1}^m \widetilde{\mathbf{X}}_i^T\widetilde{\mathbf{X}}_i\right)^{-1} \sum_{i=1}^m \widetilde{\mathbf{X}}_i^T\widetilde{\mathbf{Y}}_i = \left(\sum_{i=1}^m \widetilde{\mathbf{X}}_i^2\right)^{-1} \sum_{i=1}^m \widetilde{\mathbf{X}}_i^T\widetilde{\mathbf{Y}}_i \\ &= \left(\sum_{i=1}^m \mathbf{X}_i\right)^{-1} \sum_{i=1}^m \widetilde{\mathbf{X}}_i^T\widetilde{\mathbf{Y}}_i = \begin{pmatrix} \dfrac{\sum_{i=1}^m \sqrt{X_{i1}}\widetilde{Y}_{i1}}{\sum_{i=1}^m X_{i1}} \\ \dfrac{\sum_{i=1}^m \sqrt{X_{i2}}\widetilde{Y}_{i2}}{\sum_{i=1}^m X_{i2}} \end{pmatrix} = \begin{pmatrix} \dfrac{\sum_{i=1}^m Y_{i1}}{\sum_{i=1}^m X_{i1}} \\ \dfrac{\sum_{i=1}^m Y_{i2}}{\sum_{i=1}^m X_{i2}} \end{pmatrix}, \end{aligned}$$

$\sqrt{X_{ij}}\widetilde{Y}_{ij} = Y_{ij},\ j = 1, 2.$

5. Asymptotic normality of the ratio estimator vector $\hat{\mathbf{p}}$

It will be assumed that the random error vectors $\widetilde{\boldsymbol{\epsilon}}_i$ are not normally distributed but $i.i.d$ random vectors, $i = 1, \ldots, m$, i.e $E\left(\widetilde{\boldsymbol{\epsilon}}_i\right) = \mathbf{0}$, and $\mathrm{Cov}(\widetilde{\boldsymbol{\epsilon}}_i) = \widetilde{\Sigma}$. Moreover, under certain conditions on the design matrix $\mathbf{X}_i$ one can show that in large sample size, $\hat{\mathbf{p}}$ has the multivariate asymptotic normal distribution. These conditions namely, the pairs $(\mathbf{X}_i, \mathbf{Y}_i)$ are $i.i.d \Rightarrow (\widetilde{\mathbf{X}}_i, \widetilde{\mathbf{Y}}_i)$ are also $i.i.d$, where $\widetilde{\mathbf{X}}_i = \begin{pmatrix} \widetilde{X}_{i1} & 0 \\ 0 & \widetilde{X}_{i2} \end{pmatrix}$, $\widetilde{X}_{ij} = \sqrt{X_{ij}}$, $X_{ij} > 0$, $\forall i = 1, \ldots, m$, $j = 1, 2$, as well as $E\left(\widetilde{\mathbf{X}}_i^T \widetilde{\mathbf{X}}_i\right) > 0$, $E(\widetilde{X}_{ij}\widetilde{X}_{ij'})$ exist, $(\forall\ i = 1, \ldots, m, j, j' = 1, 2)$.
The OLS estimation of the model (3.1) can be written as

$$\hat{\mathbf{p}} = \left(\widetilde{\mathbf{X}}^T \widetilde{\mathbf{X}}\right)^{-1} \widetilde{\mathbf{X}}^T \widetilde{\mathbf{Y}} = \left(\sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\mathbf{X}}_i\right)^{-1} \left(\sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\mathbf{X}}_i \mathbf{p} + \sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\boldsymbol{\epsilon}}_i\right),$$

and can be expression as

$$\begin{aligned} \sqrt{m}\left(\hat{\mathbf{p}} - \mathbf{p}\right) &= \left(\frac{1}{m}\sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\mathbf{X}}_i\right)^{-1} \left(\frac{1}{\sqrt{m}}\sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\boldsymbol{\epsilon}}_i\right) \\ &= \left(\frac{1}{m}\sum_{i=1}^{m} \mathbf{X}_i\right)^{-1} \left(\frac{1}{\sqrt{m}}\sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\boldsymbol{\epsilon}}_i\right), \end{aligned} \tag{5.1}$$

(the details are given in p. 612 and Appendix 16.3, p. 638 of [8]). Thus, in large m, it is straightforward to see that the denominator of Eq. 5.1 (by the Law of large numbers) is consistent

$$\left(\frac{1}{m}\sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\mathbf{X}}_i\right)^{-1} \longrightarrow^{\mathcal{P}} \left(E(\widetilde{\mathbf{X}}_i^T \widetilde{\mathbf{X}}_i)\right)^{-1},$$

provided that, the matrix $E(\widetilde{\mathbf{X}}_i^T \widetilde{\mathbf{X}}_i) > 0$, so

$$\left(\frac{1}{m}\sum_{i=1}^{m} \begin{pmatrix} X_{i1} & 0 \\ 0 & X_{i2} \end{pmatrix}\right)^{-1} \longrightarrow^{\mathcal{P}} \mathrm{diag}\left(E(X_1), E(X_2)\right)^{-1}.$$

The numerator of (5.1) obeys the Multivariate Central Limit Theorem

$$\frac{1}{\sqrt{m}}\sum_{i=1}^{m} \widetilde{\mathbf{X}}_i^T \widetilde{\boldsymbol{\epsilon}}_i \longrightarrow^{\mathcal{D}} N_k\left(\mathbf{0}, E\left(\widetilde{\mathbf{X}}_i^T \widetilde{\Sigma} \widetilde{\mathbf{X}}_i\right)\right),$$

thus,

$$\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\begin{pmatrix}\sqrt{X_{i1}}\widetilde{\epsilon}_{i1}\\ \sqrt{X_{i2}}\widetilde{\epsilon}_{i2}\end{pmatrix} \longrightarrow^{\mathcal{D}} N\left(\mathbf{0}, \begin{pmatrix}\sigma_1^2E(X_{i1}) & \sigma_{12}E\left(\sqrt{X_{i1}X_{i2}}\right)\\ \sigma_{12}E\left(\sqrt{X_{i1}X_{i2}}\right) & \sigma_2^2E(X_{i2})\end{pmatrix}\right)$$
$$\equiv N\left(\mathbf{0}, \begin{pmatrix}\sigma_1^2E(X_{1}) & \sigma_{12}E\left(\sqrt{X_{1}X_{2}}\right)\\ \sigma_{12}E\left(\sqrt{X_{1}X_{2}}\right) & \sigma_2^2E(X_{2})\end{pmatrix}\right),$$

where, the asymptotic covariance

$$\begin{aligned}\Sigma^* = \mathrm{Cov}\left(\widetilde{\mathbf{X}}_i^T\widetilde{\boldsymbol{\epsilon}}_i\right) &= E\left(\mathrm{Cov}\left(\widetilde{\mathbf{X}}_i^T\widetilde{\boldsymbol{\epsilon}}_i \mid \widetilde{\mathbf{X}}_i\right)\right) + \mathrm{Cov}\left(E\left(\widetilde{\mathbf{X}}_i^T\widetilde{\boldsymbol{\epsilon}}_i \mid \widetilde{\mathbf{X}}_i\right)\right)\\ &= E\left(\widetilde{\mathbf{X}}_i^T\mathrm{Cov}\left(\widetilde{\boldsymbol{\epsilon}}_i \mid \widetilde{\mathbf{X}}_i\right)\widetilde{\mathbf{X}}_i\right) = E\left(\widetilde{\mathbf{X}}_i^T\Sigma\widetilde{\mathbf{X}}_i\right)\\ &= \begin{pmatrix}\sigma_1^2E(X_{i1}) & \sigma_{12}E\left(\sqrt{X_{i1}X_{i2}}\right)\\ \sigma_{12}E\left(\sqrt{X_{i1}X_{i2}}\right) & \sigma_2^2E(X_{i2})\end{pmatrix}\\ &= \begin{pmatrix}\sigma_1^2E(X_{1}) & \sigma_{12}E\left(\sqrt{X_{1}X_{2}}\right)\\ \sigma_{12}E\left(\sqrt{X_{1}X_{2}}\right) & \sigma_2^2E(X_{2})\end{pmatrix},\end{aligned}$$

as

$$E\left(\widetilde{\mathbf{X}}_i^T\widetilde{\boldsymbol{\epsilon}}_i \mid \widetilde{\mathbf{X}}_i\right) = \widetilde{X}_i^T E\left(\widetilde{\boldsymbol{\epsilon}}_i \mid \widetilde{\mathbf{X}}_i\right) = \mathbf{0},\ \widetilde{\mathbf{X}}_i = (\widetilde{X}_{i1}, \widetilde{X}_{i2})^T,$$

consequently, yields

$$\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\widetilde{\mathbf{X}}_i^T\widetilde{\boldsymbol{\epsilon}}_i \longrightarrow^{\mathcal{D}} N_2\left(\mathbf{0}, \Sigma^*\right).$$

Therefore, with helps of the known Slutsky's lemma ([4], pp. 119-120), Eq. 5.1 can be written as

$$\sqrt{m}(\hat{\mathbf{p}} - \mathbf{p}) \longrightarrow^{\mathcal{D}} N_2\left(\mathbf{0}, \mathrm{diag}\left(E(X_1), E(X_2)\right)^{-1} \times \Sigma^* \times \mathrm{diag}\left(E(X_1), E(X_2)\right)^{-1}\right)$$
$$\equiv N_2\left(\mathbf{0}, \begin{pmatrix}\dfrac{\sigma_1^2}{E(X_1)} & \dfrac{\sigma_{12}E(\sqrt{X_1X_2})}{E(X_1)E(X_2)}\\ \dfrac{\sigma_{12}E(\sqrt{X_1X_2})}{E(X_1)E(X_2)} & \dfrac{\sigma_2^2}{E(X_2)}\end{pmatrix}\right), \quad (5.2)$$

which results in the multivariate asymptotic normality of the ratio estimator vector $\hat{\mathbf{p}}$ with the symmetric asymptotic covariance matrix

$$\Sigma^{**} = \mathrm{Cov}(\sqrt{m}\hat{\mathbf{p}}) = \begin{pmatrix}\dfrac{\sigma_1^2}{E(X_1)} & \dfrac{\sigma_{12}E(\sqrt{X_1X_2})}{E(X_1)E(X_2)}\\ \dfrac{\sigma_{12}E(\sqrt{X_1X_2})}{E(X_1)E(X_2)} & \dfrac{\sigma_2^2}{E(X_2)}\end{pmatrix}.$$

6. Approximate confidence intervals for the linear combination $\boldsymbol{\alpha}^T\mathbf{p}$

By applying the Cramer-Wold device ([4], p. 147), it shows that the convergence in (5.2) holds iff, for any $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^T \in \Re$, we have

$$\begin{aligned}\sqrt{m}\left(\boldsymbol{\alpha}^T\hat{\mathbf{p}} - \boldsymbol{\alpha}^T\mathbf{p}\right) &\longrightarrow^{\mathcal{D}} \quad N\left(0, \alpha^T\Sigma^{**}\alpha\right) \qquad (6.1)\\ &\equiv \quad N\left(0, \alpha_1^2\frac{\sigma_1^2}{E(X_1)} + \alpha_2^2\frac{\sigma_2^2}{E(X_2)} + 2\alpha_1\alpha_2\frac{\sigma_{12}E(\sqrt{X_1X_2})}{E(X_1)E(X_2)}\right).\end{aligned}$$

The estimators $\overline{X}_{.j}$, $\overline{X}_{.12}$, s_j^2, and s_{12} are the consistent estimators of the corresponding parameters $E(X_j)$, $E(\sqrt{X_1X_2})$, σ_j^2, and σ_{12} respectively, where

$$\overline{X}_{.j} = \frac{1}{m}\sum_{i=1}^{m} X_{ij}, \; \overline{X}_{.12} = \frac{1}{m}\sum_{i=1}^{m}\sqrt{X_{ij}X_{12}}, \; j = 1, 2,$$

as well as

$$\begin{aligned} s_j^2 &= \frac{1}{m-1}\sum_{i=1}^{m}\frac{\left(Y_{ij} - \hat{p}_j X_{ij}\right)^2}{X_{ij}} = \frac{1}{m-1}\sum_{i=1}^{m}\left(\widetilde{Y}_{ij} - \hat{p}_j\sqrt{X_{ij}}\right)^2 \\ s_{12} &= \frac{1}{m-1}\sum_{i=1}^{m}\frac{\left(Y_{i1} - \hat{p}_1 X_{i1}\right)\left(Y_{i2} - \hat{p}_2 X_{i2}\right)}{\sqrt{X_{i1}}\sqrt{X_{i2}}} \\ &= \frac{1}{m-1}\sum_{i=1}^{m}\left(\widetilde{Y}_{i1} - \hat{p}_1\sqrt{X_{i1}}\right)\left(\widetilde{Y}_{i2} - \hat{p}_2\sqrt{X_{i2}}\right),\end{aligned}$$

$\widetilde{Y}_{ij} = \dfrac{Y_{ij}}{\sqrt{X_{ij}}}$, $X_{ij} > 0 \; \forall j = 1, 2, i = 1, \ldots, m$.

By plugging in these entries for the individual parameters in the asymptotic variance from (6.1), one can obtain the standard error of $\boldsymbol{\alpha}^T\hat{\mathbf{p}}$ (if defined) as

$$s.e(\boldsymbol{\alpha}^T\hat{\mathbf{p}}) = \sqrt{\frac{1}{m}\left(\alpha_1^2\frac{\sigma_1^2}{E(X_1)} + \alpha_2^2\frac{\sigma_2^2}{E(X_2)} + 2\alpha_1\alpha_2\frac{\sigma_{12}E(\sqrt{X_1X_2})}{E(X_1)E(X_2)}\right)}.$$

Construction of the confidence intervals for a linear combination of the proportions (i.e., confidence intervals for the damage total

$$Y = \boldsymbol{\alpha}^T\mathbf{p} = \sum_{j=1}^{2}\alpha_j p_j = \sum_{j=1}^{2} X_j p_j = \sum_{j=1}^{2} Y_j,$$

$\alpha_j = X_j$, where the estimated damage total $\widehat{Y} = \sum_{j=1}^{2} X_j\hat{p}_j$, $X_j = \sum_{i=1}^{m} X_{ij}$, $j = 1, 2$) needs the validity of the intervals, in the other words $(s.e(\widehat{Y}))^2 \geq 0$. Consequently, the asymptotic confidence interval for the damage total Y can be obtained by the normal quantile $z_{1-\frac{\alpha}{2}}$ as

$$\left[\widehat{Y} \pm z_{1-\frac{\alpha}{2}} s.e(\widehat{Y})\right],$$

as well as, the proposed conservative confidence interval by the t-quantile given by

$$\left[\widehat{Y} \pm t_{(m-1,1-\frac{\alpha}{2})} s.e(\widehat{Y})\right],$$

as, $\dfrac{\widehat{Y}-Y}{s.e(\widehat{Y})} \simeq t_{m-1}$, where, $t_{(m-1,1-\frac{\alpha}{2})}$ is a $\left(1-\dfrac{\alpha}{2}\right)$ quantile of the t-distribution with $(m-1)$ degrees of freedom.

7. Simulation study

The main goal of the simulations study, is to find reasonable and appropriate coverage of the suggested confidence intervals for the weighted proportions (or the confidence intervals of the damage total Y), a recommendation is based on the closeness of the achieved coverage of these intervals to the corresponding nominal value. The simulations was performed using R program. For approximating the true coverage of the damage total, we considered two confidence intervals, asymptotic (with the normal quantile) and conservative (with the adjusted t-quantile on the reduced sample sizes) confidence intervals.

The coverage of the runs are displayed graphically by the figures with the corresponding tables, also plotting horizontal lines at the nominal level. As a coverage of the confidence intervals demonstrates a similar pattern for all the confidence levels, we represent the results at the nominal level 0.95.

To simulate the true coverage for a linear combination $Y = \boldsymbol{\alpha}^T \mathbf{p} = \sum_{j=1}^{2} \alpha_j p_j$ and as our observations are count, we will simulate the count sample points $(Y_{ij}, X_{ij}), i = 1, \ldots, m, \ j = 1, 2$ at the sample sizes $m = 10, 15, 20, 50, 100$ from the bivariate Poisson distribution given in Sect. 2 with different values of the model parameter with 10000 number of replications, and $\alpha_1 = 1, \alpha_2 = 2, \alpha_j \geq 0$ (arbitrary). We will also use in this simulation the SUR sample variance s_j^2 and the covariance s_{12} given in Sec. 6.

Note that, small values of the model parameter will produce more noninformative observations (X_{ij} is a noninformative observation, if it associates with the event $A_i = \{X_{ij} = 0\}$, with probability of success $P(A_i) = e^{-\lambda}$, and $P(\overline{A_i}) = 1 - e^{-\lambda}$, where $\overline{A_i} = \{X_{ij} > 0\}, i = 1, \ldots, m, j = 1, 2$). The procedure is to exclude the invalid runs with non positive observations (intervals with invalid s.e's) which will not be counted. Further, one can calculate the percentage of these exclusions.

One need to take the following considerations:

- Taking only the informative observations will reduce the actual sample size m to the reduced random sample size m_j. Theoretically, one can calculate the random number of the noninformative observations in each sample, which equals $m - m_j$, where $m_j \leq m$, as well as, the percentage of the noninformative observations is $1 - \frac{m_j}{m}$.
- The valid runs are based on the positive variances (to ensure that, we will take only positive observations and on the positive summations of X_j.
- In this paper, we will adjust the degrees of freedom for the t-quantile on the reduced sample size $min(m_j)$, where m_j is the number of the informative observations of X_j, $j = 1, 2$. Since the damage total

$$Y = Y_1 + Y_2 = \alpha_1 p_1 + \alpha_2 p_2 = p_1 + 2p_2,$$

and

$$\widehat{Y} = \widehat{Y}_1 + \widehat{Y}_2 = \hat{p}_1 + 2\hat{p}_2.$$

So, the conservative confidence interval (safe boundaries) becomes

$$\left[\widehat{Y} \pm t_{(m1-1,1-\frac{\alpha}{2})} s.e(\widehat{Y})\right],$$

as well as, the approximate confidence interval is

$$\left[\widehat{Y} \pm z_{1-\frac{\alpha}{2}} s.e(\widehat{Y})\right],$$

where $s.e(\widehat{Y}) = \sqrt{\frac{1}{m} as\widehat{Var}(\widehat{Y})}$.

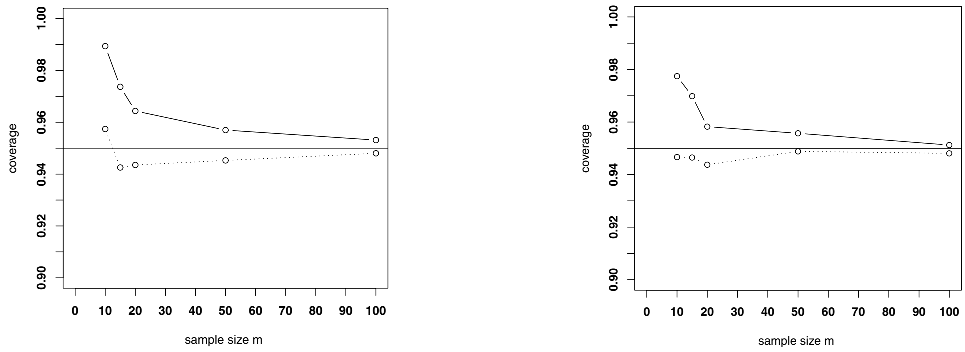
Finally, due to space constraints we will present only the more interesting results.

a) For very small values of the Poisson parameter the procedure results in many noninformative observations with high probability.

b) For small Poisson parameter values (< 1).

The coverage of confidence intervals by the z-quantile more closed to the nominal level especially at $m \geq 20$, while the confidence intervals by the adjusted t-quantile falls upper the nominal level for all sample sizes, Fig. 1 (left and right panel).

Tables 1, 2 show the percentage of the excluded runs decreases as sample size become larger by the both quantiles. and very high at small m and λ (because the simulation procedure results in many noninformative observations, or non valid intervals been excluded),

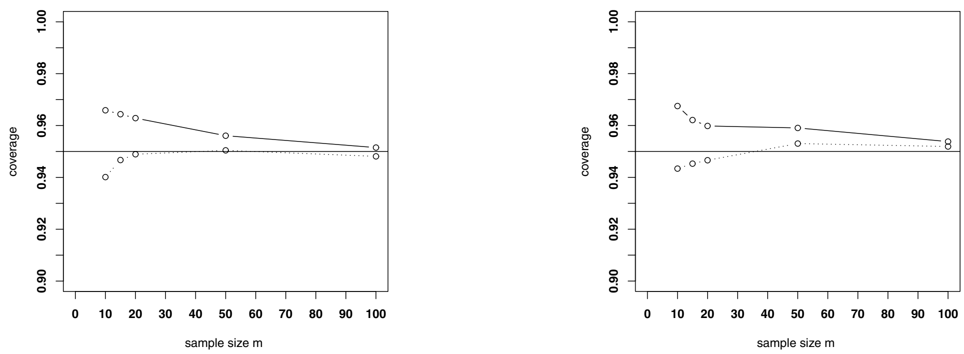


(**left panel**) (**right panel**)

Figure 1: Coverage of confidence intervals based on normal quantiles (dashed lines) and adjusted t-quantiles (solid lines) for $\lambda_0 = 0.2$, $\lambda_1 = 0.3$, $\lambda_2 = 0.3$, $\mu_1 = 0.4$, $\mu_2 = 0.3$ (**left panel**). And for $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = 0.5$ (**right panel**), $\alpha_1 = 1, \alpha_2 = 2$.

	sample size m				
	10	15	20	50	100
The percentage of the excluded runs	0.9249	0.3090	0.2651	0.1376	0.0592

Table 1: The percentage of the excluded runs of the confidence intervals for $\lambda_0 = 0.2, \lambda_1 = 0.3, \lambda_2 = 0.3, \mu_1 = 0.4, \mu_2 = 0.3, \alpha_1 = 1, \alpha_2 = 2$.



(**left panel**) (**right panel**)

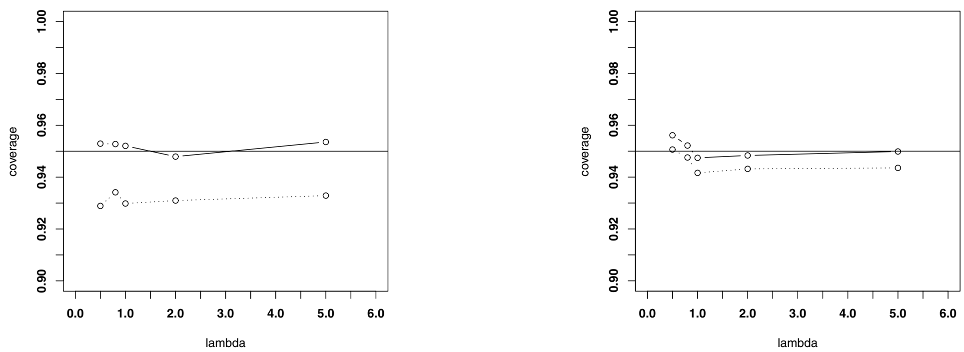
Figure 2: Coverage of confidence intervals based on normal quantiles (dashed lines) and adjusted t-quantiles (solid lines) for $\lambda_0 = 2$, $\lambda_1 = 3$, $\lambda_2 = 3$, $\mu_1 = 4$, $\mu_2 = 3$ (**left panel**). And for $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = 5, \alpha_1 = 1, \alpha_2 = 2$ (**right panel**).

	sample size m				
	10	15	20	50	100
The percentage of the excluded runs	0.3570	0.2935	0.2552	0.1283	0.0504

Table 2: The percentage of the excluded runs of the confidence intervals for $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = 0.5, \alpha_1 = 1, \alpha_2 = 2$.

	sample sizes m				
	10	15	20	50	100
The percentage of the excluded runs	0.3287	0.2814	0.2462	0.1348	0.0557

Table 3: The percentage of the excluded runs of the confidence intervals for $\lambda_0 = 2, \lambda_1 = 3, \lambda_2 = 3, \mu_1 = 4, \mu_2 = 3$ and $\alpha_1 = 1, \alpha_2 = 2$.



(**left panel**) (**right panel**)

Figure 3: Coverage at $m = 15$ for $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = \lambda$ (**left panel**). And at m=50 for $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = \lambda$ (**right panel**).

c) For larger parameter values (≥ 1)

The coverage with t-quantiles are larger than nominal for all sample sizes, while with z-quantile falls below the nominal at $m < 20$ for both chosen parameter values as seen in Fig. 2 left and right panels. The corresponding table 3 shows the percentage of excluded intervals which decrease as sample sizes increase also for large parameter values.

On the other side, the coverage at sample sizes 15, and 50 for the parameter values $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = \lambda$, where $\lambda = 0.5, 0.8, 1, 2, 5$, shown in Fig. 3 (left and right panels): the suggested t-quantile method gives reasonable coverage (recommended) at both sample sizes, while by the z-quantile method

	λ				
	0.5	0.8	1	2	5
The percentage of the excluded runs	0.2884	0.2802	0.2803	0.2742	0.2700

Table 4: The percentage of the excluded runs at sample size $m = 15$ for $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = \lambda$ and $\alpha_1 = 1, \alpha_2 = 2$.

	λ				
	0.5	0.8	1	2	5
The percentage of the excluded runs	0.1290	0.1343	0.1382	0.1308	0.1302

Table 5: The percentage of the excluded runs at sample size $m = 50$ for $\lambda_0 = \lambda_1 = \lambda_2 = \mu_1 = \mu_2 = \lambda$ and $\alpha_1 = 1, \alpha_2 = 2$.

produced dramatically smaller coverage specially at small sample size for all λ (not recommended).

From the corresponding tables 4 and 5 respectively, one can observe that: for small sample size the percentage of the excluded runs is more than that at larger sample size and the exclusions decrease as lambda or sample size become larger.

8. Conclusions

- For small parameter values the procedure results in more noninformative observations with high probability. Further, as sample size m or λ increases, the percentage of the excluded runs decrease.
- The proposed confidence intervals by the t-quantile method is highly conservative specially at small sample sizes and small λ (not recommended), however, there coverage exceed always than that by the z-quantile. While by the z-quantile is recommended when λ is larger at enough sample sizes.
- When the proportions nearly 0.5, one expect to attain better coverage and more conservative if p near 0 or 1.
- With findings of these simulations, one may recommending that the conservative confidence intervals should not be often used at small parameter values or small sample sizes, while the asymptotic is better and needs only enough sample sizes.

References

[1] Aitken, A.C., A further note on multivariate selection, *Proceedings of the Edinburgh Mathematical Society*, 5:37–40, 1936.

[2] Agresti A., and Coull, B.A., Approximate is better than exact for interval estimation of binomial proportions, *The American Statistician*, 52:119–126, 1998.

[3] Brown, L.D., Cai, T., and DasGupta, A., Interval estimation for a binomial proportion, Thechnical Report, Department of Statistics, Purdue University, 1999.

[4] Knight, K.C., Mathematical Statistics, Chapman & Hall/CRC, *Texts in Statistical Science*, pp. 120–149, 2000.

[5] Montgomery, D.C, Peck, E.A., and Vining, G.G., Introduction to linear regression analysis, Third Edition, Wiley, New York, p. 586, 2001.

[6] Johnson, N., Kotz, S., and Balakrishnan, N., Discrete Multivariate Distributions, Wiley, New York, pp. 124–138, 1997.

[7] Karlis, D., and Tsiamyrtzis, P., Exact Bayesian modeling for bivariate Poisson data and extensions, *Stat. Comput.*, 18:27–40, 2007.

[8] Stock, H., and Waston, W.C., Introduction to econometrics, Pearson Education, Inc, pp. 588–591, 2003.

[9] Simon, Gary, z or t for Binomial Confidence Intervals Teaching Statistics, 18(3):90–91, 1996.

[10] Santner, T. J., A note on teaching binomial confidence intervals, *Teaching Statistics*, 20:20–23, 1998.

[11] Timm, N.H., Applied Multivariate analysis, Springer verlag, Inc, New York, 2002.

[12] Zellner, A., An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *J. Amer. Statist. Ass.*, 57:348–368, 1962.