International Journal of Statistics and Probability; Vol. 11, No. 5; Setepmber 2022
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Parsimonious Bivariate T-distribution Type Symmetry Models for Square Contingency Tables

Kiyotaka Iki[1] & Sadao Tomizawa[2]

[1] Faculty of Economics, Nihon University, Chiyoda Ward, Tokyo, Japan

[2] Department of Information Science, Meisei University, Hino City, Tokyo, Japan

Correspondence: Kiyotaka Iki, Faculty of Economics, Nihon University, Chiyoda Ward, Tokyo, Japan.
E-mail: iki.kiyotaka@nihon-u.ac.jp

Received: July 23, 2022 Accepted: August 24, 2022 Online Published: September 12, 2022

doi:10.5539/ijsp.v11n5p44 URL: https://doi.org/10.5539/ijsp.v11n5p44

Abstract

For square contingency tables with ordered categories, Iki, Ishihara and Tomizawa (2013) considered the t-distribution type symmetry model and Iki, Okada and Tomizawa (2018) extended this model. These models are appropriate for a square contingency table if it is reasonable to assume an underlying bivariate t-distribution having any degrees of freedom. This study proposes three kinds of parsimonious models for these models. Additionally, this paper provides the decompositions of the parsimonious symmetry model using the proposed model. Some simulation studies based on bivariate t-distribution show the performances of the proposed models.

Keywords: bivariate t-distribution, square contingency table, symmetry, underlying distribution

1. Introduction

For analysis of contingency tables, we are interested in whether the two classificatory variables are independent of each another. When the independence does not hold, we may use Pearson's correlation coefficient to estimate the correlation between the two variables. Additionally, it is important to interpret the data, and propose models that fit the data well. Goodman (1979) considered the uniform association model, and Agresti (1983a) considered the linear-by-linear association model.

In particular, we consider tables with the same row and column classifications, which are known as square contingency tables. For square contingency tables, the independence between the row and column is unlikely to hold because many observations fall in the main diagonal cells, which indicates that the value of the row category is the same as the value of the column category. Therefore, for the analysis of square contingency tables, instead of independence, we are interested in whether or not the row variable is symmetric with the column variable. The symmetry (S) model (Bowker, 1948), the marginal homogeneity model (Stuart, 1955) and the quasi-symmetry model (Caussinus, 1965) have been proposed as models of symmetry. Moreover, for the research of the symmetry model, see Yoshimoto et al. (2019), Ando et al. (2021) and Shinoda et al. (2021).

We consider an $r \times r$ square contingency table with the same row and column ordinal classifications. Let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1, \ldots, r$; $j = 1, \ldots, r$). The S model is defined by

$$p_{ij} = p_{ji} \quad (i < j);$$

see Bishop et al. (1975, p.282). This model indicates a structure of symmetry of the probabilities with respect to the main diagonal of the table. Agresti (1983b) considered the linear diagonals-parameter symmetry (LDPS) model defined by

$$p_{ij} = \theta^{j-i} p_{ji} \quad (i < j).$$

This indicates that the probability of an observation falling in the (i, j)th cell, $i < j$, is θ^{j-i} times higher than the probability of it falling in the (j, i)th cell. A special case of the LDPS model obtained by putting $\theta = 1$ is the S model. Tomizawa (1991) proposed an extended linear diagonals-parameter symmetry (ELDPS) model defined by

$$p_{ij} = \theta_1^{j-i} \theta_2^{j^2-i^2} p_{ji} \quad (i < j).$$

This indicates that the probability of an observation falling in the (i, j)th cell, $i < j$, is $\theta_1^{j-i}\theta_2^{j^2-i^2}$ times higher than the probability of it falling in the (j, i)th cell. Agresti (1983; 1984, p.216) described the relationship between the LDPS model and the joint bivariate normal distribution as follows: the LDPS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances. Moreover, Tomizawa (1991) pointed out that the ELDPS model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with different marginal variances.

For any fixed constant m ($m > 2$), Iki et al. (2013) proposed the t-distribution type symmetry (TS(m)) model defined by

$$p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}} = \eta_m(j-i) \quad (i < j).$$

A special case of this model can be obtained by putting $\eta_m = 0$ in the S model. The TS(m) model indicates that the difference between the two symmetric probabilities raised to the power [$= -2/(m+2)$] is proportional to the distance from the main diagonal of the $r \times r$ table. The TS(m) model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with equal marginal variances having m degrees of freedom (see Iki et al., 2013). For any fixed constant m ($m > 2$), Iki et al. (2018) proposed the extended t-distribution type symmetry (ETS(m)) model defined by

$$p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}} = \gamma_m(j^2-i^2) + \eta_m(j-i) \quad (i < j).$$

A special case of this model can be obtained by putting $\gamma_m = 0$ in the TS(m) model. The ETS(m) model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with different marginal variances having m degrees of freedom (see Iki et al., 2018).

Now, we are interested in considering more parsimonious t-distribution type symmetry models, which can be described in terms of fewer parameters than the TS(m) (ETS(m)) models.

The purpose of this paper is to propose new models which may appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate t-distribution. The new models are different from the S, TS(m) and ETS(m) models. Section 2 proposes models and describes the properties of the new models. Section 3 includes the decompositions using the proposed models. Section 4 shows the maximum likelihood estimates of expected frequencies under the proposed models. Section 5 describes the relationships between the proposed models and t-distribution by the simulation study. Section 6 provides some concluding remarks.

2. Models

We consider random variables U and V having a joint bivariate t-distribution with m ($m > 2$) degrees of freedom, meaning $\mathrm{E}(U) = \mu_1$, $\mathrm{E}(V) = \mu_2$, variances $\mathrm{Var}(U) = m\sigma_1^2/(m-2)$, $\mathrm{Var}(V) = m\sigma_2^2/(m-2)$, and correlation coefficient $\mathrm{Corr}(U, V) = \rho$. The probability density function $f(u, v)$ is

$$f(u,v) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\left(1 + \frac{Q(u,v)}{m}\right)^{-\frac{m+2}{2}},$$

where,

$$Q(u,v) = \frac{1}{1-\rho^2}\left[\left(\frac{u-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho}{\sigma_1\sigma_2}(u-\mu_1)(v-\mu_2) + \left(\frac{v-\mu_2}{\sigma_2}\right)^2\right];$$

see Muirhead (2005, p.48). The probability density function is also expressed as

$$f(u,v) = c\left[1 + \frac{1}{m}(a_1u + b_1v + a_2u^2 + b_2v^2 + d(u,v))\right]^{-\frac{m+2}{2}}, \tag{1}$$

where

$$\begin{aligned}
c &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},\\
a_1 &= \frac{2}{\sigma_1(1-\rho^2)}\left(\frac{\rho\mu_2}{\sigma_2} - \frac{\mu_1}{\sigma_1}\right), \quad b_1 = \frac{2}{\sigma_2(1-\rho^2)}\left(\frac{\rho\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2}\right),\\
a_2 &= \frac{1}{\sigma_1^2(1-\rho^2)}, \quad b_2 = \frac{1}{\sigma_2^2(1-\rho^2)},\\
d(u,v) &= \frac{1}{1-\rho^2}\left(-\frac{2\rho}{\sigma_1\sigma_2}uv + \frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} - \frac{2\rho\mu_1\mu_2}{\sigma_1\sigma_2}\right),
\end{aligned}$$

and $d(u, v) = d(v, u)$. When $\mathrm{Var}(U) = \mathrm{Var}(V)$, that is, $\sigma_1^2 = \sigma_2^2\ (= \sigma^2)$, $f(u, v)$ is expressed as

$$f(u, v) = c\left[1 + \frac{1}{m}\left(a_1 u + b_1 v + t(u^2 + v^2) + d(u, v)\right)\right]^{-\frac{m+2}{2}}, \tag{2}$$

where

$$\begin{aligned} c &= \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}},\\ a_1 &= \frac{2}{\sigma^2(1-\rho^2)}\left(\rho\mu_2 - \mu_1\right), \quad b_1 = \frac{2}{\sigma^2(1-\rho^2)}\left(\rho\mu_1 - \mu_2\right),\\ t &= \frac{1}{\sigma^2(1-\rho^2)}\\ d(u, v) &= \frac{1}{\sigma^2(1-\rho^2)}\left(-2\rho uv + \mu_1^2 + \mu_2^2 - 2\rho\mu_1\mu_2\right), \end{aligned}$$

and $d(u, v) = d(v, u)$. Moreover, when $\mathrm{E}(U) = \mathrm{E}(V)$ and $\mathrm{Var}(U) = \mathrm{Var}(V)$, that is, $\mu_1 = \mu_2\ (= \mu)$ and $\sigma_1^2 = \sigma_2^2\ (= \sigma^2)$, $f(u, v)$ is expressed as

$$f(u, v) = c\left[1 + \frac{1}{m}\left(k(u + v) + t(u^2 + v^2) + d(u, v)\right)\right]^{-\frac{m+2}{2}}, \tag{3}$$

where

$$\begin{aligned} c &= \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}},\\ k &= -\frac{2\mu}{\sigma^2(1+\rho)}\\ t &= \frac{1}{\sigma^2(1-\rho^2)}\\ d(u, v) &= \frac{2}{\sigma^2(1-\rho^2)}\left(-\rho uv + \mu^2 - \rho\mu^2\right), \end{aligned}$$

and $d(u, v) = d(v, u)$.

We consider the $r \times r$ square contingency table with ordered categories. For any fixed constant m $(m > 2)$, we propose a model defined by

$$p_{ij} = \left[1 + \frac{1}{m}\left(\mu + \kappa(i + j) + \tau(i^2 + j^2) + \phi ij\right)\right]^{-\frac{m+2}{2}} \quad (i = 1, \ldots, r; j = 1, \ldots, r).$$

We shall refer to this model as a parsimonious symmetry (PaS(m)) model. From the form of equation (3), the PaS(m) model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with same marginal means and variances having m degrees of freedom. Under the PaS(m) model, we see that

$$p_{ij} = p_{ji} \quad (i < j).$$

Namely, the PaS(m) model implies the S model.

Next, for any fixed constant m $(m > 2)$, we propose a model defined by

$$p_{ij} = \left[1 + \frac{1}{m}\left(\mu + \alpha_1 i + \beta_1 j + \tau(i^2 + j^2) + \phi ij\right)\right]^{-\frac{m+2}{2}} \quad (i = 1, \ldots, r; j = 1, \ldots, r).$$

We shall refer to this model as a parsimonious t-distribution type symmetry (PaTS(m)) model. From the form of equation (2), the PaTS(m) model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with same marginal variances (and different marginal means) having m degrees of freedom. A special case of the PaTS(m) can be obtained by putting $\alpha_1 = \beta_1$ in the PaS(m) model. Under the PaTS(m) model,

$$p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}} = \frac{\beta_1 - \alpha_1}{m}(j - i) \quad (i < j).$$

Namely, the PaTS(m) model implies the TS(m) model. Additionally, under the PaTS(m) model, setting $\omega_{ij} = \mu + \alpha_1 i + \beta_1 j + \tau(i^2 + j^2) + \phi ij$, we see that

$$
\begin{aligned}
\lim_{m\to\infty} \frac{p_{ij}}{p_{ji}} &= \lim_{m\to\infty} \frac{(1+\frac{\omega_{ij}}{m})^{-\frac{m+2}{2}}}{(1+\frac{\omega_{ji}}{m})^{-\frac{m+2}{2}}} \\
&= \lim_{m\to\infty} \frac{\{(1+\frac{\omega_{ij}}{m})^{\frac{m}{\omega_{ij}}}\}^{-\frac{\omega_{ij}}{2}(1+\frac{2}{m})}}{\{(1+\frac{\omega_{ji}}{m})^{\frac{m}{\omega_{ji}}}\}^{-\frac{\omega_{ji}}{2}(1+\frac{2}{m})}} \\
&= \frac{\exp[-\frac{\omega_{ij}}{2}]}{\exp[-\frac{\omega_{ji}}{2}]} \\
&= \exp\left[\frac{1}{2}(\alpha_1 - \beta_1)(j-i)\right] \\
&= \theta^{j-i} \quad (i<j),
\end{aligned}
$$

where

$$
\theta = \exp\left[\frac{\alpha_1 - \beta_1}{2}\right].
$$

Namely, the PaTS(m) model approaches the LDPS model as m becomes larger.

Moreover, for any fixed constant m ($m > 2$), we propose a model defined by

$$
p_{ij} = \left[1 + \frac{1}{m}\left(\mu + \alpha_1 i + \beta_1 j + \alpha_2 i^2 + \beta_2 j^2 + \phi ij\right)\right]^{-\frac{m+2}{2}} \qquad (i = 1, \ldots, r; j = 1, \ldots, r).
$$

We shall refer to this model as a parsimonious t-distribution type symmetry (PaETS(m)) model. From the form of equation (1), the PaTS(m) model may be appropriate if it is reasonable to assume an underlying bivariate t-distribution with different marginal means and variances having m degrees of freedom. A special case of the PaETS(m) can be obtained by putting $\alpha_2 = \beta_2$ in the PaTS(m) model. Under the PaETS(m) model,

$$
p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}} = \frac{\beta_1 - \alpha_1}{m}(j-i) + \frac{\beta_2 - \alpha_2}{m}(j^2 - i^2) \quad (i<j).
$$

Namely, the PaETS(m) model implies the ETS(m) model. Further, under the PaETS(m) model, we see that

$$
\begin{aligned}
\lim_{m\to\infty} \frac{p_{ij}}{p_{ji}} &= \exp\left[\frac{1}{2}(\alpha_1 - \beta_1)(j-i) + \frac{1}{2}(\alpha_2 - \beta_2)(j^2 - i^2)\right] \\
&= \theta_1^{j-i}\theta_2^{j^2-i^2} \quad (i<j),
\end{aligned}
$$

where

$$
\theta_1 = \exp\left[\frac{\alpha_1 - \beta_1}{2}\right],\ \theta_2 = \exp\left[\frac{\alpha_2 - \beta_2}{2}\right].
$$

Namely, the PaETS(m) model approaches the ELDPS model as m becomes larger.

$$
\begin{tikzcd}
PaS(m) \arrow[r] \arrow[d] & PaTS(m) \arrow[r] \arrow[d] & PaETS(m) \arrow[d] \\
S \arrow[r] & TS(m) \arrow[r] & ETS(m)
\end{tikzcd}
$$

Figure 1. Relationships among models

In Figure 1, we show the relationships among models. In Figure, $A \to B$ indicates that model A implies model B.

3. Decompositions of Models

Consider the $r \times r$ square contingency table. Let X and Y denote the row and column variables, respectively. We refer to the model of equality of marginal means, that is, $\mathrm{E}(X) = \mathrm{E}(Y)$, as the ME model. Additionally, we refer to model of equality of marginal means and variances, that is, $\mathrm{E}(X) = \mathrm{E}(Y)$ and $\mathrm{Var}(X) = \mathrm{Var}(Y)$, as the MVE model. Then, we obtain the following theorems.

Theorem 1 *The PaS(m) model holds, if and only if both the PaETS(m) and MVE models hold.*

Proof. If the PaS(m) model holds, then the PaETS(m) and MVE models hold. Assuming that the PaETS(m) and MVE models hold, then we shall show that the PaS(m) model holds. From the PaETS(m) model, we see

$$p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}} = \frac{1}{m}\left[(\alpha_1 - \beta_1)(i-j) + (\alpha_2 - \beta_2)\left(i^2 - j^2\right)\right] \quad (i < j).$$

Then, because the MVE model is given by to $\mathrm{E}(X) = \mathrm{E}(Y)$ and $\mathrm{E}(X^2) = \mathrm{E}(Y^2)$,

$$\begin{aligned}
&\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij}\left(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}}\right)\\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \frac{p_{ij}}{m}\left[(\alpha_1 - \beta_1)(i-j) + (\alpha_2 - \beta_2)\left(i^2 - j^2\right)\right]\\
&= \frac{\alpha_1 - \beta_1}{m}\sum_{i=1}^{r}\sum_{j=1}^{r}(i-j)\,p_{ij} + \frac{\alpha_2 - \beta_2}{m}\sum_{i=1}^{r}\sum_{j=1}^{r}\left(i^2 - j^2\right)p_{ij}\\
&= \frac{\alpha_1 - \beta_1}{m}\left(\mathrm{E}(X) - \mathrm{E}(Y)\right) + \frac{\alpha_2 - \beta_2}{m}\left(\mathrm{E}(X^2) - \mathrm{E}(Y^2)\right)\\
&= 0.
\end{aligned}$$

Additionally, we have

$$\begin{aligned}
&\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij}\left(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}}\right)\\
&= \sum\sum_{i<j} p_{ij}\left(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}}\right) + \sum\sum_{i>j} p_{ij}\left(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}}\right)\\
&= \sum\sum_{i<j}\left(p_{ij} - p_{ji}\right)\left(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}}\right).
\end{aligned}$$

For any $i < j$, if $p_{ij} \neq p_{ji}$, then $(p_{ij} - p_{ji})(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}}) < 0$, if $p_{ij} = p_{ji}$, then $(p_{ij} - p_{ji})(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}}) = 0$. Thus, when we assume that the PaETS(m) and MVE models hold, we can obtain $p_{ij} = p_{ji}$ for all $i < j$. Moreover, $p_{ij} - p_{ji} = 0$ for all $i < j$, that is,

$$(\alpha_1 - \beta_1)(i-j) + (\alpha_2 - \beta_2)(i^2 - j^2) = 0 \quad \text{for all } i < j.$$

Therefore we obtain $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$. Namely, the PaS(m) model holds. The proof is completed.

Theorem 2 *The PaS(m) model holds, if and only if both the PaTS(m) and ME models hold.*

The proof of Theorem 2 is omitted because that is obtained in a way similar to Theorem 1.

4. Goodness-of-fit Test

For an $r \times r$ contingency table, let n_{ij} denote the observed frequency in the ith row and jth column of the table, where $n = \sum\sum n_{ij}$ and let m_{ij} denote the corresponding expected frequency ($i = 1, \ldots, r$; $j = 1, \ldots, r$). Assume that the observed frequencies have a multinomial distribution. Let $G^2(M)$ denote the likelihood ratio chi-squared statistic, defined by

$$G^2(M) = \sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij}\log\left(\frac{n_{ij}}{\hat{m}_{ij}}\right),$$

where $\hat{m}_{ij}$ is the maximum likelihood estimate of expected frequency m_{ij} under model M. Under model M, these statistics have a asymptotically central chi-squared distribution with the corresponding degrees of freedom. For the PaS(m) model, $\{p_{ij}\}$ are determined by μ, κ, τ and ϕ. Therefore, the numbers of degrees of freedom for the PaS(m) model are $r^2 - 4$.

Similarly, the numbers of degrees of freedom for the PaTS(m) and PaETS(m) models are $r^2 - 5$ and $r^2 - 6$, respectively. We consider the maximum likelihood estimates of expected frequencies $\{m_{ij}\}$ under the PaS(m), PaTS(m) and PaETS(m) models in the log-likelihood equation. For the PaS(m) model, we must maximize the Lagrangian

$$\begin{aligned} L = &\sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij}\log p_{ij} - \lambda\left(\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij} - 1\right) - \sum\sum_{i<j} \psi_{ij}(p_{ij} - p_{ji}) \\ &- \sum\sum_{(i,j)\in D} \lambda_{ij}\left\{p_{ij}^{-\frac{2}{m+2}} - \left(\mu + \kappa(i+j) + \tau(i^2+j^2) + \phi ij\right)\right\}, \end{aligned}$$

where

$$\begin{aligned} \mu &= \frac{1}{2}\left(11p_{11}^{-\frac{2}{m+2}} - 13p_{12}^{-\frac{2}{m+2}} + 3p_{13}^{-\frac{2}{m+2}} + p_{23}^{-\frac{2}{m+2}}\right), \\ \kappa &= \frac{1}{2}\left(-6p_{11}^{-\frac{2}{m+2}} + 9p_{12}^{-\frac{2}{m+2}} - 2p_{13}^{-\frac{2}{m+2}} - p_{23}^{-\frac{2}{m+2}}\right), \\ \tau &= \frac{1}{2}\left(p_{11}^{-\frac{2}{m+2}} - 2p_{12}^{-\frac{2}{m+2}} + p_{13}^{-\frac{2}{m+2}}\right), \\ \phi &= \frac{1}{2}\left(p_{11}^{-\frac{2}{m+2}} - p_{12}^{-\frac{2}{m+2}} - p_{13}^{-\frac{2}{m+2}} + p_{23}^{-\frac{2}{m+2}}\right), \\ D &= \{(i,j) | i<j, (i,j) \neq (1,1),(1,2),(1,3),(2,3)\}, \end{aligned}$$

with respect to $\{p_{ij}\}$, λ, $\{\psi_{ij}\}$ and $\{\lambda_{ij}\}$. For the PaTS(m) model, we must maximize the Lagrangian

$$\begin{aligned} L = &\sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij}\log p_{ij} - \lambda\left(\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij} - 1\right) \\ &- \sum\sum_{(i,j)\in E_1} \psi_{ij}\left(p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}} - (j-i)p_{12}^{-\frac{2}{m+2}} + (j-i)p_{21}^{-\frac{2}{m+2}}\right) \\ &- \sum\sum_{(i,j)\in E_2} \lambda_{ij}\left\{p_{ij}^{-\frac{2}{m+2}} - \left(\mu + \alpha i + \beta j + \tau(i^2+j^2) + \phi ij\right)\right\}, \end{aligned}$$

where

$$\begin{aligned} \mu &= \frac{1}{2}\left(11p_{11}^{-\frac{2}{m+2}} - 10p_{12}^{-\frac{2}{m+2}} + 3p_{13}^{-\frac{2}{m+2}} - 3p_{21}^{-\frac{2}{m+2}} + p_{23}^{-\frac{2}{m+2}}\right), \\ \alpha &= \frac{1}{2}\left(-6p_{11}^{-\frac{2}{m+2}} + 6p_{12}^{-\frac{2}{m+2}} - 2p_{13}^{-\frac{2}{m+2}} + 3p_{21}^{-\frac{2}{m+2}} - p_{23}^{-\frac{2}{m+2}}\right), \\ \beta &= \frac{1}{2}\left(-6p_{11}^{-\frac{2}{m+2}} + 8p_{12}^{-\frac{2}{m+2}} - 2p_{13}^{-\frac{2}{m+2}} + p_{21}^{-\frac{2}{m+2}} - p_{23}^{-\frac{2}{m+2}}\right), \\ \tau &= \frac{1}{2}\left(p_{11}^{-\frac{2}{m+2}} - 2p_{12}^{-\frac{2}{m+2}} + p_{13}^{-\frac{2}{m+2}}\right), \\ \phi &= \frac{1}{2}\left(p_{11}^{-\frac{2}{m+2}} - p_{13}^{-\frac{2}{m+2}} - p_{21}^{-\frac{2}{m+2}} + p_{23}^{-\frac{2}{m+2}}\right), \\ E_1 &= \{(i,j) | i<j, (i,j) \neq (1,2)\}, \\ E_2 &= \{(i,j) | i<j, (i,j) \neq (1,1),(1,2),(1,3),(2,3)\}, \end{aligned}$$

with respect to $\{p_{ij}\}$, λ, $\{\psi_{ij}\}$ and $\{\lambda_{ij}\}$. For the PaETS(m) model, we must maximize the Lagrangian

$$\begin{aligned} L = &\sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij}\log p_{ij} - \lambda\left(\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij} - 1\right) \\ &- \sum\sum_{(i,j)\in F_1} \psi_{ij}\left[p_{ij}^{-\frac{2}{m+2}} - p_{ji}^{-\frac{2}{m+2}} + \frac{(j-i)}{2}\left\{(2i+2j-8)(p_{12}^{-\frac{2}{m+2}} - p_{21}^{-\frac{2}{m+2}})\right.\right. \\ &\qquad\qquad\qquad\qquad\qquad\qquad \left.\left. -(i+j-3)(p_{13}^{-\frac{2}{m+2}} - p_{31}^{-\frac{2}{m+2}})\right\}\right] \\ &- \sum\sum_{(i,j)\in F_2} \lambda_{ij}\left\{p_{ij}^{-\frac{2}{m+2}} - \left(\mu + \alpha_1 i + \beta_1 j + \alpha_2 i^2 + \beta_2 j^2 + \phi ij\right)\right\}, \end{aligned}$$

where

$$
\begin{aligned}
\mu &= \frac{1}{2}\left(11p_{11}^{-\frac{2}{m+2}} - 6p_{12}^{-\frac{2}{m+2}} + p_{13}^{-\frac{2}{m+2}} - 7p_{21}^{-\frac{2}{m+2}} + p_{23}^{-\frac{2}{m+2}} + 2p_{31}^{-\frac{2}{m+2}}\right),\\
\alpha_1 &= \frac{1}{2}\left(-6p_{11}^{-\frac{2}{m+2}} + p_{13}^{-\frac{2}{m+2}} + 9p_{21}^{-\frac{2}{m+2}} - p_{23}^{-\frac{2}{m+2}} - 3p_{31}^{-\frac{2}{m+2}}\right),\\
\beta_1 &= \frac{1}{2}\left(-6p_{11}^{-\frac{2}{m+2}} + 8p_{12}^{-\frac{2}{m+2}} - 2p_{13}^{-\frac{2}{m+2}} + p_{21}^{-\frac{2}{m+2}} - p_{23}^{-\frac{2}{m+2}}\right),\\
\alpha_2 &= \frac{1}{2}\left(p_{11}^{-\frac{2}{m+2}} - 2p_{21}^{-\frac{2}{m+2}} + p_{31}^{-\frac{2}{m+2}}\right),\\
\beta_2 &= \frac{1}{2}\left(p_{11}^{-\frac{2}{m+2}} - 2p_{12}^{-\frac{2}{m+2}} + p_{13}^{-\frac{2}{m+2}}\right),\\
\phi &= \frac{1}{2}\left(p_{11}^{-\frac{2}{m+2}} - p_{13}^{-\frac{2}{m+2}} - p_{21}^{-\frac{2}{m+2}} + p_{23}^{-\frac{2}{m+2}}\right),\\
F_1 &= \{(i,j) | i < j, (i,j) \neq (1,2), (1,3)\},\\
F_2 &= \{(i,j) | i < j, (i,j) \neq (1,1), (1,2), (1,3), (2,3)\},
\end{aligned}
$$

with respect to $\{p_{ij}\}$, λ, $\{\psi_{ij}\}$ and $\{\lambda_{ij}\}$. Setting the partial derivations of L equal to zero using the Newton-Raphson method, we can obtain the maximum likelihood estimates of $\{m_{ij}\}$ under the PaS(m), PaTS(m) and PaETS(m) models.

5. Simulation Study

As described in Section 2, the PaS(m), PaTS(m) and PaETS(m) models may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate t-distribution having m degrees of freedom. We shall consider the relationships between the proposed models and bivariate t-distribution in terms of simulation studies, and the comparison between the proposed models and S, TS(m) and ETS(m) models.

Consider random variables U and V having a bivariate t-distribution with m degrees of freedom, meaning $\mathrm{E}(U) = 0$, $\mathrm{E}(V) = \mu_2$, variances $\mathrm{Var}(U) = m/(m-2)$, $\mathrm{Var}(V) = m\sigma_2^2/(m-2)$, and correlation coefficient $\mathrm{Corr}(U, V) = \rho$. Suppose that there are some conditions; $m = 30, 100$, $\mu_2 = 0, 0.2$, $\sigma_2^2 = 1, 1.2$, $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$, a 4×4 table of sample size 5000 is formed using cut points for each variable at $-0.7, 0, 0.7$.

We count the frequencies of acceptance (at the 0.05 significance level) based on the likelihood ratio chi-squared statistic for testing the hypothesis that the models with the corresponding m degrees of freedom hold per 10000 times for 4×4 tables on each condition.

From Tables 1 and 2, we see that the ETS(m) model is a good fit for all conditions. Further the TS(m) model is a good fit when $\sigma_2^2 = 1$, and the S model gives good fit on when $\mu_2 = 0$ and $\sigma_2^2 = 1$. In contrast, the PaS(m), PaTS(m) and PaETS(m) models show a similar trend when ρ is close to 0. Thus, from the result of this simulation, we obtain that if it is reasonable to assume an underlying bivariate t-distribution with a low correlation coefficient, the parsimonious models would fit the data well.

6. Concluding Remarks

Each of the S, TS(m) and ETS(m) models is saturated on the main diagonal cells of the table, but the PaS(m), PaTS(m) and PaETS(m) models are unsaturated on them. Thus, under the PaS(m), PaTS(m) and PaETS(m) models, the estimated expected frequencies on the main diagonal are always not equal to the observed frequencies on the main diagonal. The PaS(m), PaTS(m) and PaETS(m) models may be useful when we want to utilize the information on the main diagonal.

From Section 5, when observations are not so concentrated in the main diagonal cells, that is, a correlation coefficient between row and column variables is close to 0, the proposed models (PaS(m), PaTS(m) and PaETS(m)) may be better for application to a square table than the S, TS(m) and ETS(m) models.

Acknowledgements

The authors would like to thank the referee and the editorial team for many comments and suggestions.

Table 1. The frequencies of acceptance (at the 0.05 significance level) per 10000 times for 4×4 tables based on the likelihood ratio chi-squared statistic for testing the hypothesis that the S, TS(30), ETS(30), PaS(30), PaTS(30) or PaETS(30) model hold

μ_2	σ_2^2	ρ	S	TS(30)	ETS(30)	PaS(30)	PaTS(30)	PaETS(30)
0	1	0.1	9501	9520	9494	9156	9150	9116
0.2	1	0.1	0	9273	9241	0	8888	8837
0	1.2	0.1	1535	1308	9490	2403	2234	9082
0.2	1.2	0.1	0	1107	9289	0	1806	8927
0	1	0.2	9474	9460	9483	8589	8549	8489
0.2	1	0.2	0	9265	9246	0	8336	8304
0	1.2	0.2	1575	1317	9508	1936	1779	8462
0.2	1.2	0.2	0	1055	9294	0	1418	8258
0	1	0.3	9490	9493	9472	7417	7316	7171
0.2	1	0.3	0	9168	9137	0	7064	6935
0	1.2	0.3	1437	1212	9489	1237	1118	7124
0.2	1.2	0.3	0	983	9117	0	922	6765
0	1	0.4	9459	9505	9497	5159	4981	4751
0.2	1	0.4	0	9133	9058	0	4720	4565
0	1.2	0.4	1338	1111	9478	567	486	4599
0.2	1.2	0.4	0	901	9086	0	400	4390
0	1	0.5	9504	9491	9504	2349	2211	2025
0.2	1	0.5	0	9068	9003	0	2081	1866
0	1.2	0.5	1214	1017	9480	145	126	1888
0.2	1.2	0.5	0	765	8990	0	92	1755
0	1	0.6	9499	9487	9518	533	449	384
0.2	1	0.6	0	9028	8993	0	424	350
0	1.2	0.6	897	737	9455	7	7	315
0.2	1.2	0.6	0	636	8921	0	8	263

Table 2. The frequencies of acceptance (at the 0.05 significance level) per 10000 times for 4×4 tables based on the likelihood ratio chi-squared statistic for testing the hypothesis that the S, TS(100), ETS(100), PaS(100), PaTS(100) or PaETS(100) model hold

μ_2	σ_2^2	ρ	S	TS(100)	ETS(100)	PaS(100)	PaTS(100)	PaETS(100)
0	1	0.1	9486	9485	9495	9350	9350	9352
0.2	1	0.1	0	9271	9259	0	9117	9089
0	1.2	0.1	1525	1283	9502	2583	2412	9371
0.2	1.2	0.1	0	1062	9225	0	1840	9108
0	1	0.2	9480	9490	9486	8970	8951	8880
0.2	1	0.2	0	9242	9175	0	8647	8610
0	1.2	0.2	1469	1222	9489	2060	1897	8854
0.2	1.2	0.2	0	981	9247	0	1516	8585
0	1	0.3	9497	9483	9496	7942	7842	7739
0.2	1	0.3	0	9174	9154	0	7560	7445
0	1.2	0.3	1407	1168	9480	1397	1272	7700
0.2	1.2	0.3	0	921	9102	0	973	7252
0	1	0.4	9501	9521	9492	6038	5857	5672
0.2	1	0.4	0	9087	9049	0	5324	5151
0	1.2	0.4	1358	1145	9485	694	597	5407
0.2	1.2	0.4	0	785	8994	0	419	4833
0	1	0.5	9468	9470	9481	2880	2728	2508
0.2	1	0.5	0	8942	8847	0	2531	2337
0	1.2	0.5	1122	917	9465	165	141	2269
0.2	1.2	0.5	0	651	8881	0	110	1959
0	1	0.6	9529	9517	9512	669	609	541
0.2	1	0.6	0	8907	8844	0	525	467
0	1.2	0.6	838	651	9451	14	11	396
0.2	1.2	0.6	0	534	8764	0	6	326

References

Agresti, A. (1983a). A survey of strategies for modeling cross-classifications having ordinal variables. *Journal of the American Statistical Association, 78*, 184-198. https://doi.org/10.1080/01621459.1983.10477950

Agresti, A. (1983b). A simple diagonals-parameter symmetry and quasi-symmetry model. *Statistics and Probability Letters, 1*, 313-316. https://doi.org/10.1016/0167-7152(83)90051-2

Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. New York: John Wiley & Sons.

Ando, S., Tahata, K., & Tomizawa, S. (2021). Extended diagonal uniform association symmetry model for square contingency tables with ordinal categories. *Statistics and Probability Letters, 169*, 1-4. https://doi.org/10.1016/j.spl.2020.108973

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice.* Cambridge: The MIT Press.

Bowker, A. H. (1948). A test for symmetry in contingency tables. *Journal of the American Statistical Association, 43*, 572-574. https://doi.org/10.1080/01621459.1948.10483284

Caussinus, H. (1965). Contribution à l'analyse statistique des tableaux de corrélation. *Annales de la Faculté des Sciences de l'Université de Toulouse, 29*, 77-182. https://doi.org/10.5802/afst.519

Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association, 74*, 537-552. https://doi.org/10.1080/01621459.1979.10481650

Iki, K., Ishihara, T., & Tomizawa, S. (2013). Bivariate t-distribution type symmetry model for square contingency tables with ordered categories. *Model Assisted Statistics and Applications: An International Journal, 8*, 315-319. https://doi.org/10.3233/MAS-130269

Iki, K., Okada, M., & Tomizawa, S. (2018). An extended bivariate t-distribution type symmetry model for square contingency tables. *Open Journal of Statistics, 8*, 249-257. https://doi.org/10.4236/ojs.2018.82015

Muirhead, R. J. (2005). *Aspects of Multivariate Statistical Theory*. New Jersey: John Wiley & Sons.

Shinoda, S., Tahata, K., Yamamoto, K., & Tomizawa, S. (2021). Marginal continuation odds ratio model and decomposition of marginal homogeneity model for multi-way contingency tables. *Sankhya: The Indian Journal of Statistics, Ser. B, 83*, 304-324. https://doi.org/10.1007/s13571-020-00228-9

Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika, 42*, 412-416. http://dx.doi.org/10.1093/biomet/42.3-4.412

Tomizawa, S. (1991). An extended linear diagonals-parameter symmetry model for square contingency tables with ordered categories. *Metron: International Journal of Statistics, 49*, 401-409.

Yoshimoto, T., Tahata, K., Iki, K., & Tomizawa, S. (2019). Moment symmetry models and decompositions of symmetry for multi-way contingency tables. *Calcutta Statistical Association Bulletin, 71*, 83-98. https://doi.org/10.1177/0008068319880443

Copyrights