



Exactly and almost compatible joint distributions for high-dimensional discrete conditional distributions



Kun-Lin Kuo^a, Chwan-Chin Song^b, Thomas J. Jiang^{b,*}

^a Institute of Statistics, National University of Kaohsiung, Kaohsiung, Taiwan

^b Department of Mathematical Sciences, National Chengchi University, Taipei, Taiwan

ARTICLE INFO

Article history:

Received 15 August 2016

Available online 24 March 2017

AMS subject classifications:

primary 62H05

secondary 62E17

Keywords:

Almost compatible joint distribution

Compatibility

Full conditional distributions

Incompatibility

Irreducible block diagonal matrix

Rank one positive extension matrix

Structural ratio matrix

ABSTRACT

A conditional model is a set of conditional distributions, which may be compatible or incompatible, depending on whether or not there exists a joint distribution whose conditionals match the given conditionals. In this paper, we propose a new mathematical tool called a “structural ratio matrix” (SRM) to develop a unified compatibility approach for discrete conditional models. With this approach, we can find all joint pdfs after confirming that the given model is compatible. In practice, it is most likely that the conditional models we encounter are incompatible. Therefore, it is important to investigate approximated joint distributions for them. We use the concept of SRM again to construct an almost compatible joint distribution, with consistency property, to represent the given incompatible conditional model.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

In probability modeling involving n random variables X_1, \dots, X_n , it is often simpler to specify the conditional distributions than to specify the entire joint distribution. Reproducing the associated joint distribution from the specified conditional distributions is one of the main problems we might encounter. However, there is no guarantee that such a joint distribution exists. We say that the specified conditional distributions are compatible if there exists a joint distribution whose conditional distributions match the specified conditional distributions. For practical circumstances when the compatibility problem can occur in statistical practice, one may refer to Arnold et al. [1] and Arnold and Press [5], among others.

Let $\mathbf{X} = (X_1, \dots, X_n)$ be a discrete random vector and let π denote its joint pdf. Let \mathbf{X}_a (or \mathbf{x}_a) be the a -component of \mathbf{X} (or \mathbf{x}), where a is a nonempty subset of $\{1, \dots, n\}$. For disjoint subsets a and b , the conditional pdf of \mathbf{X}_a at \mathbf{x}_a given $\mathbf{X}_b = \mathbf{x}_b$ is denoted by $\pi_{a|b}(\mathbf{x}_a|\mathbf{x}_b)$, and the pdf of \mathbf{X}_a is expressed as $\pi_a(\mathbf{x}_a)$. If $a \cup b = \{1, \dots, n\}$, we say that $\pi_{a|b}$ is a full conditional pdf of \mathbf{X} and $\pi_{a|b}(\mathbf{x}_a|\mathbf{x}_b)$ can also be denoted by $\pi_{a|b}(\mathbf{x})$. A full conditional model is a model that consists of only full conditional pdfs. We say that a full conditional model is compatible whenever its members are compatible.

Given a full conditional model, say $\{f_{a_i|\bar{a}_i} : 1 \leq i \leq k\}$ where $k \geq 2$, \bar{a}_i is the complement of a_i , and each $f_{a_i|\bar{a}_i}$ is considered as a putative conditional pdf of \mathbf{X}_{a_i} given $\mathbf{X}_{\bar{a}_i}$, it is natural to ask the following four fundamental questions in relation with compatibility issues: (Q1) How can we verify whether the model is compatible? (Q2) How can we check whether there is

* Corresponding author.

E-mail addresses: klkuo@nuk.edu.tw (K.-L. Kuo), ccsong@math.nccu.edu.tw (C.-C. Song), jiangt@nccu.edu.tw (T.J. Jiang).

a unique joint distribution if the model is compatible? (Q3) How can we find all the possible joint distributions from the given model if it is compatible? (Q4) How can we find an approximated joint distribution, with reasonable properties, to represent the given model if it is incompatible?

The issues (Q1)–(Q3) have been studied by many researchers, such as Arnold et al. [1–3], Arnold and Press [5], Berti et al. [6], Chen [7], Ip and Wang [12], Kuo and Wang [13], Ng [14], Slavkovic and Sullivant [15], Song et al. [16], Tian and Tan [17], Tian et al. [18], Wang [19], Wang and Kuo [20], Yao et al. [21], and Ghosh and Nadarajah [11]. For the most part, each of the above papers focuses on the case $a_i = \{i\}$ and $n = k$. Here, we address (Q1)–(Q3) without such restrictions. Both of the approaches given by Arnold et al. [3] and Arnold and Press [5] are based on arrays (multi-dimensional structures) using scalar labels. However, the computation of arrays may not be easy for users. Moreover, the complexity increases drastically as n and k increase.

Issue (Q4) has been investigated by Arnold et al. [2], Arnold and Gokhale [4], Chen and Ip [8], Chen et al. [9] and Ghosh and Balakrishnan [10]. These papers discuss the most nearly compatible joint distribution, when the given full conditional model is incompatible, with various criteria to measure the incompatibility. Arnold and Gokhale [4] consider the Kullback–Leibler divergence as a measure of incompatibility and then provide an iterative algorithm to find the most nearly compatible joint distribution so that its conditional distributions have minimal Kullback–Leibler divergence to the given conditional distributions. Arnold et al. [2] propose the concept of ε -compatibility, based on either one of six different measuring criteria, to give the most nearly compatible joint distribution. Chen et al. [9] and Chen and Ip [8] propose the approach of Gibbs ensemble, based on Gibbs sampling, to search the most nearly compatible joint distribution. Ghosh and Balakrishnan [10] discuss several measurements of incompatibility for determining the most nearly compatible joint distribution. These papers discuss low-dimensional cases and do not give the explicit general results for any n and k .

To overcome the problems mentioned above, we provide a novel technique based on the Structural Ratio Matrix (SRM), which uses the (vector) values of conditioning variable(s) in the reference conditional pdf as column labels and uses the (vector) values of conditioning variable(s) in the non-reference conditional pdf(s) as row labels. This technique of using a reference conditional pdf to construct a 2-dimensional SRM is critical in solving complex high-dimensional problems and is described here for the first time. The SRM method we propose can be implemented to address the four compatibility issues for any high-dimensional full conditional models with/without any pattern of structural zeros. This method provides not only an efficient way to find the joint distribution(s) or an almost compatible joint distribution, but also a unified way that even a practitioner can easily apply.

2. Structural ratio matrix

Suppose that we have a full conditional model $\mathbb{F} = \{f_{a_i|\bar{a}_i} : 1 \leq i \leq k\}$, $k \geq 2$, for $\mathbf{X} = (X_1, \dots, X_n)$. We say that \mathbb{F} is compatible with a joint pdf π if $\pi_{a_i|\bar{a}_i} = f_{a_i|\bar{a}_i}$ for all $i \in \{1, \dots, k\}$. Let Ω_i , Ψ and Ψ_a denote the supports of $f_{a_i|\bar{a}_i}$, π and π_a , respectively. Specifically, $\Omega_i = \{\mathbf{x} : f_{a_i|\bar{a}_i}(\mathbf{x}) > 0\}$, $\Psi = \{\mathbf{x} : \pi(\mathbf{x}) > 0\}$, and $\Psi_a = \{\mathbf{x}_a : \pi_a(\mathbf{x}_a) > 0\}$.

When \mathbb{F} is compatible with π , it follows that $\Omega_1 = \dots = \Omega_k = \Psi$; see Arnold and Press [5, p. 152]. In addition, based on the definition of conditional pdf, we have

$$\forall \mathbf{x} \in \Psi \quad \frac{f_{a_i|\bar{a}_i}(\mathbf{x})}{f_{a_j|\bar{a}_j}(\mathbf{x})} = \frac{\pi_{a_i|\bar{a}_i}(\mathbf{x})}{\pi_{a_j|\bar{a}_j}(\mathbf{x})} = \frac{\pi(\mathbf{x})/\pi_{\bar{a}_i}(\mathbf{x}_{\bar{a}_i})}{\pi(\mathbf{x})/\pi_{\bar{a}_j}(\mathbf{x}_{\bar{a}_j})} = \frac{\pi_{\bar{a}_j}(\mathbf{x}_{\bar{a}_j})}{\pi_{\bar{a}_i}(\mathbf{x}_{\bar{a}_i})}.$$

This implies that

$$\frac{f_{a_i|\bar{a}_i}(\mathbf{x})}{f_{a_j|\bar{a}_j}(\mathbf{x})} = \frac{\pi_{\bar{a}_j}(\mathbf{x}_{\bar{a}_j})}{\pi_{\bar{a}_i}(\mathbf{x}_{\bar{a}_i})} = \frac{\pi_{\bar{a}_j}(\mathbf{y}_{\bar{a}_j})}{\pi_{\bar{a}_i}(\mathbf{y}_{\bar{a}_i})} = \frac{f_{a_i|\bar{a}_i}(\mathbf{y})}{f_{a_j|\bar{a}_j}(\mathbf{y})},$$

for any $\mathbf{x}, \mathbf{y} \in \Psi$ provided that $\mathbf{x}_{\bar{a}_i} = \mathbf{y}_{\bar{a}_i}$ (i.e., the \bar{a}_i -components of \mathbf{x} and \mathbf{y} are equal) and $\mathbf{x}_{\bar{a}_j} = \mathbf{y}_{\bar{a}_j}$. Thus, in addressing the compatibility issues, we need to assume the following conditions for any given \mathbb{F} .

(C1) $\Omega_1 = \dots = \Omega_k$ which can be used to define the set Ψ .

(C2) Ratios $f_{a_i|\bar{a}_i}/f_{a_j|\bar{a}_j}$ at $\mathbf{x} \in \Psi$ and at $\mathbf{y} \in \Psi$ are equal whenever \mathbf{x} and \mathbf{y} have the same \bar{a}_i - and \bar{a}_j -components.

In other words, if either condition (C1) or (C2) does not hold, then \mathbb{F} is automatically incompatible.

Given $f_{a_i|\bar{a}_i}$ and $f_{a_j|\bar{a}_j}$ in \mathbb{F} , we define the *basic* SRM (BSRM), denoted by $R^{[i:j]} = [r_{\mathbf{x}_{\bar{a}_j}, \mathbf{x}_{\bar{a}_i}}^{[i:j]}]$, of $f_{a_i|\bar{a}_i}$ over $f_{a_j|\bar{a}_j}$ as follows:

$$r_{\mathbf{x}_{\bar{a}_j}, \mathbf{x}_{\bar{a}_i}}^{[i:j]} = \begin{cases} f_{a_i|\bar{a}_i}(\mathbf{x})/f_{a_j|\bar{a}_j}(\mathbf{x}) & \text{if } \mathbf{x} \in \Psi, \\ * & \text{otherwise,} \end{cases}$$

where $r_{\mathbf{x}_{\bar{a}_j}, \mathbf{x}_{\bar{a}_i}}^{[i:j]}$ is the entry with row label $\mathbf{x}_{\bar{a}_j}$ and column label $\mathbf{x}_{\bar{a}_i}$ in $R^{[i:j]}$, and $*$ refers to an undefined entry. The BSRM $R^{[i:j]}$ has size $|\Psi_{\bar{a}_j}| \times |\Psi_{\bar{a}_i}|$ and uses the elements of $\Psi_{\bar{a}_j}$ and of $\Psi_{\bar{a}_i}$ as respective row labels and column labels, where $|\Psi_{\bar{a}_j}|$ is the number of elements in the set $\Psi_{\bar{a}_j}$.

To construct a SRM for \mathbb{F} , we use one of k conditional pdfs, say $f_{a_i|\bar{a}_i}$, as the reference conditional pdf and then build $k - 1$ BSRMs $R^{[i:j]}$ of $f_{a_i|\bar{a}_i}$ over $f_{a_j|\bar{a}_j}$ for $j \neq i$ as defined above. For convenience, we use $f_{a_i|\bar{a}_i}$ as the reference conditional pdf. As all these BSRMs have the same column size, we can construct a 2-dimensional matrix by arranging them in a cascade, as can be seen in the following definition for SRM.

Definition 1. Given \mathbb{F} , the SRM of $f_{a_1|\bar{a}_1}$ over $\{f_{a_2|\bar{a}_2}, \dots, f_{a_k|\bar{a}_k}\}$ is defined as

$$R^{[1;2,\dots,k]} = \begin{pmatrix} R^{[1;2]} \\ \vdots \\ R^{[1;k]} \end{pmatrix},$$

where $R^{[1;j]}$ is the BSRM of $f_{a_1|\bar{a}_1}$ over $f_{a_j|\bar{a}_j}$.

Notice that our SRM is structurally different from the ratio arrays given by Arnold et al. [3] and Arnold and Press [5] when $k > 2$, although they are the same when $n = k = 2$. When $k > 2$, the conditional pdf ratios considered by Arnold et al. [3] and Arnold and Press [5] are expressed in several n -dimensional arrays. However, ours are arranged in a specially designed matrix.

We say that a SRM, say $R = [r_{a,b}]$, has a rank 1 positive extension (ROPE) matrix $E = [e_{a,b}]$ if E has the same size as R , and satisfies the following three conditions: (i) all entries of E are positive; (ii) $e_{a,b} = r_{a,b}$ for all $r_{a,b} \neq *$; and (iii) E is of rank 1. Note that E may not be unique when it exists.

3. Checking compatibility and identifying associated joint distribution(s)

In this section, we will show how to use SRM to check the compatibility and to find the associated joint distribution(s) for a given \mathbb{F} . First, we call a positive vector $\mathbf{u} = (u_1, \dots, u_m)^\top$ a probability vector if $u_1 + \dots + u_m = 1$, and an *inverted* probability vector if $1/u_1 + \dots + 1/u_m = 1$. Then, we give the following lemma to prove our main result, [Theorem 3](#), which gives a necessary and sufficient condition for checking compatibility.

Lemma 2. Suppose that $R^{[1;2,\dots,k]}$ has a ROPE matrix E . Then there exist an inverted probability vector \mathbf{v} (with $|\Psi_{\bar{a}_1}|$ components) and probability vectors $\mathbf{u}^{(2)}, \dots, \mathbf{u}^{(k)}$ (with $|\Psi_{\bar{a}_2}|, \dots, |\Psi_{\bar{a}_k}|$ components, respectively) such that

$$E = \begin{pmatrix} \mathbf{u}^{(2)} \\ \mathbf{u}^{(3)} \\ \vdots \\ \mathbf{u}^{(k)} \end{pmatrix} \mathbf{v}^\top. \tag{1}$$

Moreover, $\mathbf{u}^{(j)} \mathbf{v}^\top$ is a ROPE matrix of $R^{[1;j]}$.

Proof. Since E is a ROPE matrix, there exists positive vectors $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{u}}^{(2)}, \dots, \tilde{\mathbf{u}}^{(k)}$ such that $E = (\tilde{\mathbf{u}}^{(2)\top}, \dots, \tilde{\mathbf{u}}^{(k)\top})^\top \tilde{\mathbf{v}}^\top$. For convenience, we use the column and row labels of $R^{[1;j]}$ to represent the subscripts of components of $\tilde{\mathbf{v}}$ and components of $\tilde{\mathbf{u}}^{(j)}$, respectively. That is, $\tilde{\mathbf{v}} = (\tilde{v}_{\mathbf{x}_{\bar{a}_1} : \mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1}})^\top$ and $\tilde{\mathbf{u}}^{(j)} = (\tilde{u}_{\mathbf{x}_{\bar{a}_j} : \mathbf{x}_{\bar{a}_j} \in \Psi_{\bar{a}_j}})^\top$. For $\mathbf{x} \in \Psi$ with \bar{a}_1 - and \bar{a}_j -components $\mathbf{x}_{\bar{a}_1}$ and $\mathbf{x}_{\bar{a}_j}$, respectively, we have

$$\forall_{j \in \{2, \dots, k\}} \tilde{u}_{\mathbf{x}_{\bar{a}_j}}^{(j)} f_{a_j|\bar{a}_j}(\mathbf{x}) = \frac{f_{a_1|\bar{a}_1}(\mathbf{x})}{\tilde{v}_{\mathbf{x}_{\bar{a}_1}}}.$$

We then obtain

$$\begin{aligned} \sum_{\mathbf{x}_{\bar{a}_j} \in \Psi_{\bar{a}_j}} \tilde{u}_{\mathbf{x}_{\bar{a}_j}}^{(j)} &= \sum_{\mathbf{x}_{\bar{a}_j} \in \Psi_{\bar{a}_j}} \tilde{u}_{\mathbf{x}_{\bar{a}_j}}^{(j)} \sum_{\substack{\mathbf{y} \in \Psi \\ \mathbf{y}_{\bar{a}_j} = \mathbf{x}_{\bar{a}_j}}} f_{a_j|\bar{a}_j}(\mathbf{y}) = \sum_{\mathbf{x} \in \Psi} \tilde{u}_{\mathbf{x}_{\bar{a}_j}}^{(j)} f_{a_j|\bar{a}_j}(\mathbf{x}) = \sum_{\mathbf{x} \in \Psi} \frac{f_{a_1|\bar{a}_1}(\mathbf{x})}{\tilde{v}_{\mathbf{x}_{\bar{a}_1}}} \\ &= \sum_{\mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1}} \frac{1}{\tilde{v}_{\mathbf{x}_{\bar{a}_1}}} \sum_{\substack{\mathbf{y} \in \Psi \\ \mathbf{y}_{\bar{a}_1} = \mathbf{x}_{\bar{a}_1}}} f_{a_1|\bar{a}_1}(\mathbf{y}) = \sum_{\mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1}} \frac{1}{\tilde{v}_{\mathbf{x}_{\bar{a}_1}}}. \end{aligned}$$

Therefore,

$$\tilde{u}_\oplus = \tilde{u}_+^{(2)} = \tilde{u}_+^{(3)} = \dots = \tilde{u}_+^{(k)},$$

where

$$\tilde{u}_\oplus = \sum_{\mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1}} \frac{1}{\tilde{v}_{\mathbf{x}_{\bar{a}_1}}} \quad \text{and} \quad \forall_{j \in \{2, \dots, k\}} \tilde{u}_+^{(j)} = \sum_{\mathbf{x}_{\bar{a}_j} \in \Psi_{\bar{a}_j}} \tilde{u}_{\mathbf{x}_{\bar{a}_j}}^{(j)}.$$

Letting $\mathbf{v} \equiv \tilde{v}_\oplus \tilde{\mathbf{v}}$ and $\mathbf{u}^{(j)} \equiv \tilde{\mathbf{u}}^{(j)} / \tilde{u}_+^{(j)}$ for all $j \in \{2, \dots, k\}$, we conclude the proof. \square

Theorem 3. \mathbb{F} is compatible if and only if the SRM $R^{[1;2,\dots,k]}$ has a ROPE matrix.

Once compatibility is confirmed, one may be interested in identifying the associated joint distribution(s). Notice that $R^{[1;2,\dots,k]}$ may have many ROPE matrices. By Lemma 2, we can obtain the marginal pdf of $\mathbf{X}_{\bar{a}_i}$ for each $i \in \{1, \dots, k\}$ from the $k - 1$ probability vectors $\mathbf{u}^{(j)}$ or the inverted probability vector \mathbf{v} associated with the chosen ROPE matrix E . Multiplying any marginal pdf of $\mathbf{X}_{\bar{a}_i}$ with its corresponding conditional pdf $f_{\bar{a}_i|\bar{a}_i}(\mathbf{x})$ given in the model \mathbb{F} , we then get the joint pdf associated with E . Hence, we could have k different ways to obtain this joint pdf. The following theorem, which can be proved by Lemma 2, provides additional formulas for finding the marginal pdfs of $\mathbf{X}_{\bar{a}_1}, \dots, \mathbf{X}_{\bar{a}_k}$ directly in terms of the entries in the chosen ROPE matrix E . Here, we write E as

$$E = \begin{pmatrix} E_2 \\ \vdots \\ E_k \end{pmatrix}$$

where $\mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1}$ and for each $j \in \{2, \dots, k\}$, $E_j = [e_{\mathbf{y}_{\bar{a}_j}, \mathbf{x}_{\bar{a}_1}}], \mathbf{y}_{\bar{a}_j} \in \Psi_{\bar{a}_j}$.

Theorem 4. For the chosen ROPE matrix E , the $\mathbf{X}_{\bar{a}_1}$ -marginal pdf can be expressed as

$$(F1) \pi_{\bar{a}_1}(\mathbf{x}_{\bar{a}_1}) = \left(\sum_{\mathbf{y}_{\bar{a}_j} \in \Psi_{\bar{a}_j}} e_{\mathbf{y}_{\bar{a}_j}, \mathbf{x}_{\bar{a}_1}} \right)^{-1}, \quad \mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1},$$

for all $j \in \{2, \dots, k\}$. In addition, whatever $j \in \{2, \dots, k\}$, the $\mathbf{X}_{\bar{a}_j}$ -marginal pdf can be expressed as

$$(F2) \pi_{\bar{a}_j}(\mathbf{x}_{\bar{a}_j}) = \left(\sum_{\mathbf{y}_{\bar{a}_1} \in \Psi_{\bar{a}_1}} \frac{1}{e_{\mathbf{x}_{\bar{a}_j}, \mathbf{y}_{\bar{a}_1}}} \right)^{-1}, \quad \mathbf{x}_{\bar{a}_j} \in \Psi_{\bar{a}_j}.$$

Moreover, the joint pdf based on E can be obtained from $f_{\bar{a}_j|\bar{a}_j} \pi_{\bar{a}_j}$ for any $j \in \{1, \dots, k\}$.

Because (F1) holds for any $j \in \{2, \dots, k\}$, (F1) can be reexpressed as

$$(F1') \pi_{\bar{a}_1}(\mathbf{x}_{\bar{a}_1}) = (k - 1) \left(e_{+, \mathbf{x}_{\bar{a}_1}} \right)^{-1},$$

where $e_{+, \mathbf{x}_{\bar{a}_1}}$ is the sum of the $\mathbf{x}_{\bar{a}_1}$ -column of E . In addition, by Lemma 2, for any row of E , say $\mathbf{y}_{\bar{a}_j}$, we have

$$(F1'') \pi_{\bar{a}_1}(\mathbf{x}_{\bar{a}_1}) = \frac{1}{e_{\mathbf{y}_{\bar{a}_j}, \mathbf{x}_{\bar{a}_1}}} \left(\sum_{\mathbf{z}_{\bar{a}_1} \in \Psi_{\bar{a}_1}} \frac{1}{e_{\mathbf{y}_{\bar{a}_j}, \mathbf{z}_{\bar{a}_1}}} \right)^{-1}.$$

From Theorem 4, we see that one ROPE matrix would generate one joint density when \mathbb{F} is compatible. However, as mentioned earlier, the ROPE matrix of a SRM may not be unique. In the next section, we further show that there is a one-to-one correspondence between the set of all ROPE matrices of the SRM being used and the set of all possible joint pdfs, i.e., different ROPE matrices will generate different joint pdfs.

4. Addressing the compatibility issues by using IBD matrix

Given a full conditional model, we have provided a ROPE matrix method for checking its compatibility and finding its associated joint distribution(s) in the previous section. In this section, we will use the IBD (irreducible block diagonal) matrix technique (see Song et al. [16]) of the associated SRM to provide a more efficient procedure for addressing the compatibility issues. More specifically, by resorting to the IBD matrix technique, it will typically be easier to check compatibility and to find all joint pdfs for a given conditional model.

A SRM R is said to be reducible if, after interchanging some rows and/or columns, it can be rearranged as

$$\left(\begin{array}{c|c} T_1 & * \\ * & T_2 \end{array} \right),$$

where the entries off the diagonal block matrices T_1 and T_2 are all $*$. The matrix R is irreducible if it is not reducible. The concepts of reducibility and irreducibility used here are somewhat different from those in matrix theory. We say that T is an IBD matrix of R if T can be obtained by interchanging some rows and/or columns of R such that

$$T = \left(\begin{array}{c|ccc|c} T_1 & * & \cdots & \cdots & * \\ * & T_2 & & & * \\ \vdots & * & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * \\ * & * & \cdots & * & T_L \end{array} \right) \equiv \text{diag}(T_1, \dots, T_L),$$

where the diagonal block matrices T_1, \dots, T_L are irreducible.

The following lemma can be deduced from Lemma 3 and Theorem 6 of Song et al. [16].

Lemma 5. Let $\text{diag}(T_1, \dots, T_L)$ be any IBD matrix of a SRM R . Then (i) R has a ROPE matrix if and only if each of T_1, \dots, T_L has a ROPE matrix. (ii) If R has a ROPE matrix, then the ROPE matrix is unique if and only if $L = 1$.

With Theorem 3 and Lemma 5(i), we have the following corollary.

Corollary 6. Suppose that $\text{diag}(T_1, \dots, T_L)$ is any IBD matrix of $R^{[1;2,\dots,k]}$. Then \mathbb{F} is compatible if and only if each of T_1, \dots, T_L has a ROPE matrix.

The next theorem shows that there is a one-to-one correspondence between the set of all possible joint pdfs and the set of all ROPE matrices of the SRM being used.

Theorem 7. Let \mathfrak{E} be the set of all ROPE matrices of $R^{[1;2,\dots,k]}$, and \mathfrak{F} be the set of all joint pdfs compatible with \mathbb{F} . Then there is a one-to-one correspondence between \mathfrak{E} and \mathfrak{F} .

Proof. Define a mapping $H : \mathfrak{E} \rightarrow \mathfrak{F}$ by $H(E) = \pi$, where π is the joint pdf obtained by the inverted probability vector $\mathbf{v} = (v_{\mathbf{x}_{\bar{a}_1}} : \mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1})^\top$, which is associated with E in Lemma 2, through the following equation

$$\pi(\mathbf{x}) = f_{a_1|\bar{a}_1}(\mathbf{x}) \frac{1}{v_{\mathbf{x}_{\bar{a}_1}}}, \quad \mathbf{x} \in \Psi.$$

We claim that H is bijective.

H is surjective: For each $\pi \in \mathfrak{F}$, we can set $E \in \mathfrak{E}$ as Eq. (1) by letting $\mathbf{u}^{(j)} = (\pi_{\bar{a}_j}(\mathbf{x}_{\bar{a}_j}) : \mathbf{x}_{\bar{a}_j} \in \Psi_{\bar{a}_j})^\top$, for all $j \in \{2, \dots, k\}$ and $\mathbf{v} = (1/\pi_{\bar{a}_1}(\mathbf{x}_{\bar{a}_1}) : \mathbf{x}_{\bar{a}_1} \in \Psi_{\bar{a}_1})^\top$. Hence, H is surjective.

H is injective: Suppose that $H(E) = H(F) = \pi$, where

$$E = \begin{pmatrix} \mathbf{u}^{(2)} \\ \vdots \\ \mathbf{u}^{(k)} \end{pmatrix} \mathbf{v}^\top \quad \text{and} \quad F = \begin{pmatrix} \mathbf{s}^{(2)} \\ \vdots \\ \mathbf{s}^{(k)} \end{pmatrix} \mathbf{t}^\top.$$

We have

$$f_{a_1|\bar{a}_1}(\mathbf{x}) \frac{1}{v_{\mathbf{x}_{\bar{a}_1}}} = \pi(\mathbf{x}) = f_{a_1|\bar{a}_1}(\mathbf{x}) \frac{1}{t_{\mathbf{x}_{\bar{a}_1}}}.$$

This implies that $\mathbf{v} = \mathbf{t}$. By Lemma 2, we also have

$$\forall_{j \in \{2, \dots, k\}} f_{a_j|\bar{a}_j}(\mathbf{x}) u_{\mathbf{x}_{\bar{a}_j}}^{(j)} = \pi(\mathbf{x}) = f_{a_j|\bar{a}_j}(\mathbf{x}) s_{\mathbf{x}_{\bar{a}_j}}^{(j)}.$$

This implies that $\mathbf{u}^{(j)} = \mathbf{s}^{(j)}$. We then have $E = F$. Therefore, H is injective. \square

With Lemma 5(ii) and Theorem 7, we have the following corollary.

Corollary 8. Suppose that $\text{diag}(T_1, \dots, T_L)$ is an IBD matrix of $R^{[1;2,\dots,k]}$. Then the following statements are equivalent. (i) The associated joint distribution is unique. (ii) The ROPE matrix of $R^{[1;2,\dots,k]}$ is unique. (iii) $L = 1$ and T_1 has a ROPE matrix.

Multiplying the given conditional pdf $f_{a_1|\bar{a}_1}$ by all possible marginal pdfs of $\mathbf{X}_{\bar{a}_1}$, we could obtain all possible joint pdfs for the conditional model. Naturally comes the question: how do we find all possible marginal pdfs of $\mathbf{X}_{\bar{a}_1}$? By using the IBD matrix technique, we give a potentially easier way to accomplish it in the next theorem.

Theorem 9. Suppose that $\text{diag}(T_1, \dots, T_L)$ is an IBD matrix of $R^{[1;2,\dots,k]}$ and for all $\ell \in \{1, \dots, L\}$, T_ℓ , with size $I_\ell \times J_\ell$ has ROPE matrix $\tilde{T}_\ell = [t_{ij}^{(\ell)}]$. For each \mathbf{x} , with \bar{a}_1 -component $\mathbf{x}_{\bar{a}_1}$, in Ψ , assume that the column, labeled $\mathbf{x}_{\bar{a}_1}$, of $R^{[1;2,\dots,k]}$ has been interchanged to, say, the j th column of T_ℓ . Then, all possible joint pdfs at \mathbf{x} can be found by either

$$f_{a_1|\bar{a}_1}(\mathbf{x})(k-1)p_\ell \left(\sum_{i=1}^{I_\ell} t_{ij}^{(\ell)} \right)^{-1} \quad \text{or} \quad f_{a_1|\bar{a}_1}(\mathbf{x}) \frac{p_\ell}{t_{ij}^{(\ell)}} \left(\sum_{j^*=1}^{J_\ell} \frac{1}{t_{ij^*}^{(\ell)}} \right)^{-1}$$

for any $i \in \{1, \dots, I_\ell\}$, where $p_1 > 0, \dots, p_L > 0$ are arbitrary with $p_1 + \dots + p_L = 1$.

Proof. For each $\ell \in \{1, \dots, L\}$, let $\tilde{T}_\ell = \alpha_\ell \beta_\ell^\top$ for some positive vectors $\alpha_\ell = (\alpha_{\ell 1}, \dots, \alpha_{\ell, I_\ell})^\top$ and $\beta_\ell = (\beta_{\ell 1}, \dots, \beta_{\ell, J_\ell})^\top$. Then

$$\left\{ \begin{pmatrix} q_1 \alpha_1 \\ \vdots \\ q_L \alpha_L \end{pmatrix} (q_1^{-1} \beta_1^\top, \dots, q_L^{-1} \beta_L^\top) : q_\ell > 0, 1 \leq \ell \leq L \right\}$$

is the set of all possible ROPE matrices of $\text{diag}(T_1, \dots, T_L)$. It yields from (F1') that

$$\begin{aligned} \pi(\mathbf{x}) &= f_{a_1|\bar{a}_1}(\mathbf{x})(k-1) \left(\sum_{m=1}^L \sum_{i=1}^{I_m} \frac{q_m}{q_\ell} \alpha_{mi} \beta_{\ell j} \right)^{-1} \\ &= f_{a_1|\bar{a}_1}(\mathbf{x})(k-1) \frac{\sum_{i=1}^{I_\ell} q_\ell \alpha_{\ell i}}{\sum_{m=1}^L \sum_{i=1}^{I_m} q_m \alpha_{mi}} \left(\sum_{i=1}^{I_\ell} \alpha_{\ell i} \beta_{\ell j} \right)^{-1} \\ &= f_{a_1|\bar{a}_1}(\mathbf{x})(k-1) p_\ell \left(\sum_{i=1}^{I_\ell} t_{ij}^{(\ell)} \right)^{-1}, \end{aligned}$$

where

$$p_\ell = \left(\sum_{i=1}^{I_\ell} q_\ell \alpha_{\ell i} \right) / \left(\sum_{m=1}^L \sum_{i=1}^{I_m} q_m \alpha_{mi} \right).$$

Similarly, it yields from (F1''), for arbitrary $i \in \{1, \dots, I_\ell\}$, that

$$\begin{aligned} \pi(\mathbf{x}) &= f_{a_1|\bar{a}_1}(\mathbf{x}) \frac{1}{\alpha_{\ell i} \beta_{\ell j}} \left\{ \sum_{m=1}^L \sum_{j^*=1}^{J_m} \left(\frac{q_\ell}{q_m} \alpha_{\ell i} \beta_{mj^*} \right)^{-1} \right\}^{-1} \\ &= f_{a_1|\bar{a}_1}(\mathbf{x}) \frac{\sum_{j^*=1}^{J_\ell} (q_\ell^{-1} \beta_{\ell j^*})^{-1}}{\alpha_{\ell i} \beta_{\ell j} \sum_{m=1}^L \sum_{j^*=1}^{J_m} (q_m^{-1} \beta_{mj^*})^{-1}} \left\{ \sum_{j^*=1}^{J_\ell} (\alpha_{\ell i} \beta_{\ell j^*})^{-1} \right\}^{-1} \\ &= f_{a_1|\bar{a}_1}(\mathbf{x}) \frac{p_\ell}{t_{ij}^{(\ell)}} \left(\sum_{j^*=1}^{J_\ell} \frac{1}{t_{ij^*}^{(\ell)}} \right)^{-1}, \end{aligned}$$

where

$$p_\ell = \sum_{j^*=1}^{J_\ell} (q_\ell^{-1} \beta_{\ell j^*})^{-1} / \sum_{m=1}^L \sum_{j^*=1}^{J_m} (q_m^{-1} \beta_{mj^*})^{-1}$$

Thus the argument is complete. \square

Observe that, when applying the IBD matrix technique to address the compatibility issues, we usually only need to be concerned with the column labels, and consequently the row labels can be omitted for convenience.

Next, we give an example of a full conditional model with $n = 4$ and $k = 3$ to briefly illustrate our method. Consider a conditional model $\mathbb{F} = \{f_{12|34}, f_{31|24}, f_{24|13}\}$ as follows. Here, for simplicity, $f_{12|34}$ is used to denote the conditional pdf $f_{\{1,2\}|\{3,4\}}$ and analogous notations for the others.

x_1	1	2	1	1	2	3	3	3
x_2	1	1	1	2	2	2	1	2
x_3	1	1	2	1	1	2	3	3
x_4	1	1	1	2	2	2	2	2
$f_{12 34}$	1/3	2/3	1	3/4	1/4	1	3/4	1/4
$f_{31 24}$	1/5	1	4/5	1	1	2/3	1	1/3
$f_{24 13}$	1/4	2/3	1	3/4	1/3	1	3/4	1/4

The SRM of $f_{12|34}$ over $\{f_{3|124}, f_{24|13}\}$ and an associated IBD matrix are given below on the left and right, respectively.

	(1, 1)(1, 2)(2, 1)(2, 2)(3, 2)		(1, 1)(1, 2)(2, 1)(2, 2)(3, 2)
(1, 1, 1)	5/3 * 5/4 * *	(1, 1, 1)	5/3 * 5/4 * *
(2, 1, 1)	2/3 * * * *	(2, 1, 1)	2/3 * * * *
(1, 2, 2)	* 3/4 * * *	(1, 1)	4/3 1 * * *
(2, 2, 2)	* 1/4 * * *	(2, 1)	1 3/4 * * *
(3, 1, 2)	* * * * 3/4	(1, 2, 2)	* 3/4 * * *
(3, 2, 2)	* * * 3/2 3/4	(2, 2, 2)	* 1/4 * * *
(1, 1)	4/3 1 * * *	(1, 2)	* * 1 * *
(2, 1)	1 3/4 * * *	(3, 1, 2)	* * * * 3/4
(1, 2)	* * 1 * *	(3, 2, 2)	* * * 3/2 3/4
(3, 2)	* * * 1 *	(3, 2)	* * * 1 *
(3, 3)	* * * * 1	(3, 3)	* * * * 1

The ROPE matrices of the two diagonal block matrices are given in the following table.

	(1, 1)(1, 2)(2, 1)(2, 2)(3, 2)
(1, 1, 1)	5/3 5/4 5/4 * *
(2, 1, 1)	2/3 1/2 1/2 * *
(1, 1)	4/3 1 1 * *
(2, 1)	1 3/4 3/4 * *
(1, 2, 2)	1 3/4 3/4 * *
(2, 2, 2)	1/3 1/4 1/4 * *
(1, 2)	4/3 1 1 * *
(3, 1, 2)	* * * 3/2 3/4
(3, 2, 2)	* * * 3/2 3/4
(3, 2)	* * * 1 1/2
(3, 3)	* * * 2 1

By Corollary 6, \mathbb{F} is compatible. In addition, the associated joint pdf is not unique by Corollary 8. Using Theorem 9, all possible joint pdfs can be expressed as

x_1	1	2	1	1	2	3	3	3
x_2	1	1	1	2	2	2	1	2
x_3	1	1	2	1	1	2	3	3
x_4	1	1	1	2	2	2	2	2
π_{1234}	$p_1/11$	$2p_1/11$	$4p_1/11$	$3p_1/11$	$p_1/11$	$p_2/3$	$p_2/2$	$p_2/6$

where $p_1 + p_2 = 1$.

5. Almost compatible joint distributions

In practice, it is likely that a full conditional model of observed conditional pdfs is not compatible even if they are sampled from the same joint distribution, since they may be contaminated by sampling errors. In this section, we will focus on the near compatibility issue and provide a method to construct almost compatible joint distributions for incompatible conditional models.

Recall that if $\mathbb{F} = \{f_{a_i|\bar{a}_i} : 1 \leq i \leq k\}$ is a compatible model, then the corresponding SRM, by Theorem 3, has a ROPE matrix. This ROPE matrix can be expressed, by Lemma 2, as the product of a column vector, which is composed of $k - 1$ probability vectors, and an inverted probability row vector. We can then construct a joint pdf by using any one of these $k - 1$ probability vectors or the inverted probability vector together with its corresponding conditional pdf.

Now, assume that $\mathbb{F} = \{f_{a_i|\bar{a}_i} : 1 \leq i \leq k\}$ is an incompatible model and the SRM in use is $R = R^{[1;2,\dots,k]}$. We will find a vector $\mathbf{u}^\top = (\mathbf{u}^{(2)\top}, \dots, \mathbf{u}^{(k)\top})$, with each $\mathbf{u}^{(i)}$ a probability vector, and an inverted probability vector \mathbf{v} so that the following squared quasi-Frobenius norm L is minimized:

$$L = \|R^{[1;2,\dots,k]} - \mathbf{u}\mathbf{v}^\top\|_{qF}^2 = \sum_{b=2}^k \sum_{(i,j) \in \Lambda_b} (r_{ij}^{(b)} - u_{bi}v_j)^2.$$

Here, $r_{ij}^{(b)}$ is the (i, j) th entry of the sub-matrix $R^{[1;b]}$ of the SRM $R^{[1;2,\dots,k]}$, $\Lambda_b = \{(i, j) | r_{ij}^{(b)} \neq *\}$, $\mathbf{u}^\top = (\mathbf{u}^{(2)\top}, \dots, \mathbf{u}^{(k)\top})$, $\mathbf{u}^{(b)} = (u_{b1}, \dots, u_{b,|\Psi_{\bar{a}_b}|})^\top$ for all $b \in \{2, \dots, k\}$ and $\mathbf{v} = (v_1, \dots, v_{|\Psi_{\bar{a}_1}|})^\top$.

We call any minimizer (\mathbf{u}, \mathbf{v}) of L as a quasi-Frobenius solution of R . Hence, we can find k approximated joint distributions for an incompatible model via any quasi-Frobenius solution of R . Note that when the model is compatible, the above quasi-Frobenius norm is zero and the set of vectors $\mathbf{u}^{(2)}, \dots, \mathbf{u}^{(k)}$ and \mathbf{v} in Lemma 2 form a quasi-Frobenius solution of R .

Since we may use any one of k conditional pdfs as the reference conditional pdf to construct a SRM, we can construct k different SRMs. Overall, we can have k^2 approximated joint distributions.

Let $f_{a_i|\bar{a}_i}^{[n_i]}$ be the sample conditional pdf of $f_{a_i|\bar{a}_i}$ with sample size n_i . Then the sample conditional model

$$\mathbb{F}_{\mathbf{n}} = \{f_{a_i|\bar{a}_i}^{[n_i]} : 1 \leq i \leq k\},$$

where $\mathbf{n} = (n_1, \dots, n_k)$, may not be compatible and its corresponding SRM matrix $R_{\mathbf{n}}$ may not have a ROPE matrix. As mentioned above, we can find k approximated joint distributions for $\mathbb{F}_{\mathbf{n}}$ via any quasi-Frobenius solution of $R_{\mathbf{n}}$. For convenience, we use $\mathbf{n} \rightarrow \infty$ to mean that each $n_i \rightarrow \infty$. Next, we shall show that each of these k approximated joint pdfs has the consistency property, in the sense that when the sample sizes $\mathbf{n} \rightarrow \infty$, each of the obtained approximated joint pdfs would approach the joint pdf associated with \mathbb{F} , when its SRM is irreducible.

Lemma 10. *If the matrix $R_{\mathbf{n}} = R_{\mathbf{n}}^{[1;2,\dots,k]}$ is the SRM of sample conditional model $\mathbb{F}_{\mathbf{n}} = \{f_{a_i|\bar{a}_i}^{[n_i]} : 1 \leq i \leq k\}$, which is from the compatible conditional model $\mathbb{F} = \{f_{a_i|\bar{a}_i} : 1 \leq i \leq k\}$ with sample sizes $\mathbf{n} = (n_1, \dots, n_k)$, then any quasi-Frobenius solution $(\mathbf{u}^{[n]}, \mathbf{v}^{[n]})$ of $R_{\mathbf{n}}$ has the following property: For each entry of R without $*$,*

$$\lim_{\mathbf{n} \rightarrow \infty} \mathbf{u}^{[n]} \mathbf{v}^{[n]\top} = R \text{ a.s.},$$

where $R = R^{[1;2,\dots,k]}$ is the SRM for \mathbb{F} and a.s. means almost surely.

Proof. Because \mathbb{F} is compatible, there exists $(\mathbf{u}^*, \mathbf{v}^*)$ so that $\|R - \mathbf{u}^* \mathbf{v}^{*\top}\|_{qF} = 0$. Hence,

$$\inf_{\mathbf{u}, \mathbf{v}} \|R_{\mathbf{n}} - \mathbf{u} \mathbf{v}^{\top}\|_{qF} \leq \|R_{\mathbf{n}} - \mathbf{u}^* \mathbf{v}^{*\top}\|_{qF} \leq \|R_{\mathbf{n}} - R\|_{qF} + \|R - \mathbf{u}^* \mathbf{v}^{*\top}\|_{qF} = \|R_{\mathbf{n}} - R\|_{qF}.$$

Since $(\mathbf{u}^{[n]}, \mathbf{v}^{[n]})$ is a quasi-Frobenius solution of $R_{\mathbf{n}}$, i.e.,

$$\inf_{\mathbf{u}, \mathbf{v}} \|R_{\mathbf{n}} - \mathbf{u} \mathbf{v}^{\top}\|_{qF} = \|R_{\mathbf{n}} - \mathbf{u}^{[n]} \mathbf{v}^{[n]\top}\|_{qF},$$

then $\|R_{\mathbf{n}} - \mathbf{u}^{[n]} \mathbf{v}^{[n]\top}\|_{qF} \leq \|R_{\mathbf{n}} - R\|_{qF}$. By the previous inequality and the fact that

$$\|R - \mathbf{u}^{[n]} \mathbf{v}^{[n]\top}\|_{qF} \leq \|R - R_{\mathbf{n}}\|_{qF} + \|R_{\mathbf{n}} - \mathbf{u}^{[n]} \mathbf{v}^{[n]\top}\|_{qF},$$

we have $\|R - \mathbf{u}^{[n]} \mathbf{v}^{[n]\top}\|_{qF} \leq 2\|R - R_{\mathbf{n}}\|_{qF}$. However, as the sample sizes \mathbf{n} tend to infinity, $R_{\mathbf{n}}$ converges to R a.s., for each entry of R without $*$. That is, $\|R - R_{\mathbf{n}}\|_{qF} \rightarrow 0$ a.s. when $\mathbf{n} \rightarrow \infty$. Therefore, $\lim_{\mathbf{n} \rightarrow \infty} \mathbf{u}^{[n]} \mathbf{v}^{[n]\top} = R$ a.s., for each entry without $*$. \square

Next, we give the following consistency result.

Theorem 11. *Under the same notations as those in Lemma 10, we further assume that the SRM R is irreducible. Then all approximated joint pdfs based on $(\mathbf{u}^{[n]}, \mathbf{v}^{[n]})$, which is a quasi-Frobenius solution of $R_{\mathbf{n}}$, converge to the true joint pdf associated with \mathbb{F} a.s. when the sample sizes \mathbf{n} tend to infinity.*

Proof. Since R is irreducible, there is a unique ROPE matrix of R , say E . By Lemma 2, E can be expressed as

$$E = \begin{pmatrix} \mathbf{u}^{(2)} \\ \mathbf{u}^{(3)} \\ \vdots \\ \mathbf{u}^{(k)} \end{pmatrix} \mathbf{v}^{\top} = \mathbf{u} \mathbf{v}^{\top},$$

where each probability vector $\mathbf{u}^{(i)}$ gives the values of the marginal pdf of $\mathbf{X}_{\bar{a}_i}$ for $i \in \{2, \dots, k\}$, and the inverted probability vector \mathbf{v} gives the marginal pdf of $\mathbf{X}_{\bar{a}_1}$. Furthermore, by Lemma 10, we have, for each entry of E

$$\lim_{\mathbf{n} \rightarrow \infty} \mathbf{u}^{[n]} \mathbf{v}^{[n]\top} = E \text{ a.s.}$$

It follows that $\lim_{\mathbf{n} \rightarrow \infty} \mathbf{v}^{[n]} = \mathbf{v}$ a.s. and $\lim_{\mathbf{n} \rightarrow \infty} \mathbf{u}^{[n]} = \mathbf{u}$ a.s. In addition, $\lim_{n_i \rightarrow \infty} f_{a_i|\bar{a}_i}^{[n_i]}(\mathbf{x}) = f_{a_i|\bar{a}_i}(\mathbf{x})$ a.s. Hence, we can conclude that all approximated joint pdfs based on $(\mathbf{u}^{[n]}, \mathbf{v}^{[n]})$ converge to the true joint pdf associated with \mathbb{F} a.s. when the sample sizes \mathbf{n} tend to infinity. This concludes the argument. \square

As mentioned earlier, we can have k^2 approximated joint distributions for any full incompatible conditional model. A mixture of these k^2 approximated joint distributions is also an approximated joint distribution. If all mixing weights are constants, then this new approximated joint distribution would also have the consistency property. For simplicity, we take the equal mixing weight for this mixture distribution, and call it the almost compatible joint distribution. Of course, one may adopt other mixing weights, e.g., using the divergence measure given by Chen et al. [9] to decide the mixing weights, to find a new approximated joint distribution.

6. Conclusion remark

Our method is easy to implement and can address the compatible and incompatible tasks for any family of full conditional distributions with/without any pattern of structural zeros.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology, Taiwan (MOST 104-2118-M-390-001 and MOST 103-2118-M-004-002-MY2). The authors thank three referees and the Editor-in-Chief for their comments.

References

- [1] B.C. Arnold, E. Castillo, J.M. Sarabia, Conditionally specified distributions: an introduction (with discussions), *Statist. Sci.* 16 (2001) 249–274.
- [2] B.C. Arnold, E. Castillo, J.M. Sarabia, Exact and near compatibility of discrete conditional distributions, *Comput. Statist. Data Anal.* 16 (2002) 231–252.
- [3] B.C. Arnold, E. Castillo, J.M. Sarabia, Compatibility of partial or complete conditional probability specifications, *J. Statist. Plann. Inference* 123 (2004) 133–159.
- [4] B.C. Arnold, D.V. Gokhale, Distributions most nearly compatible with given families of conditional distributions, *Test* 7 (1998) 377–390.
- [5] B.C. Arnold, S.J. Press, Compatible conditional distributions, *J. Amer. Statist. Assoc.* 84 (1989) 52–156.
- [6] P. Berti, E. Dreassi, P. Rigo, Compatibility results for conditional distributions, *J. Multivariate Anal.* 125 (2014) 190–203.
- [7] H.Y. Chen, Compatibility of conditionally specified models, *Statist. Probab. Lett.* 80 (2010) 670–677.
- [8] S.-H. Chen, E.H. Ip, Behaviour of the Gibbs sampler when conditional distributions are potentially incompatible, *J. Stat. Comput. Simul.* 85 (2015) 3266–3275.
- [9] S.-H. Chen, E.H. Ip, Y.J. Wang, Gibbs ensembles for nearly compatible and incompatible conditional models, *Comput. Statist. Data Anal.* 55 (2011) 1760–1769.
- [10] I. Ghosh, N. Balakrishnan, Study of incompatibility or near compatibility of bivariate discrete conditional probability distributions through divergence measures, *J. Stat. Comput. Simul.* 85 (2015) 117–130.
- [11] I. Ghosh, S. Nadarajah, An alternative approach for compatibility of two discrete conditional distributions, *Comm. Statist. Theory Methods* 45 (2016) 4416–4432.
- [12] E.H. Ip, Y.J. Wang, Canonical representation of conditionally specified multivariate discrete distributions, *J. Multivariate Anal.* 100 (2009) 1282–1290.
- [13] K.-L. Kuo, Y.J. Wang, A simple algorithm for checking compatibility among discrete conditional distributions, *Comput. Statist. Data Anal.* 55 (2011) 2457–2462.
- [14] K.W. Ng, Inversion of Bayes' formula for events, in: M. Lovric (Ed.), *International Encyclopedia of Statistical Science*, Springer, 2011, pp. 690–695.
- [15] A.B. Slavkovic, S. Sullivant, The space of compatible full conditionals is a unimodular toric variety, *J. Symbolic Comput.* 41 (2006) 196–209.
- [16] C.-C. Song, L.-A. Li, C.-H. Chen, T.J. Jiang, K.-L. Kuo, Compatibility of finite discrete conditional distributions, *Statist. Sinica* 20 (2010) 423–440.
- [17] G.-L. Tian, M. Tan, Exact statistical solutions using the inverse Bayes formulae, *Statist. Probab. Lett.* 62 (2003) 305–315.
- [18] G. Tian, M. Tan, K.W. Ng, M. Tang, A unified method for checking compatibility and uniqueness for finite discrete conditional distributions, *Comm. Statist. Theory Methods* 38 (2009) 115–129.
- [19] Y.J. Wang, Comparisons of three approaches for discrete conditional models, *Comm. Statist. Simulation Comput.* 41 (2012) 32–43.
- [20] Y.J. Wang, K.-L. Kuo, Compatibility of discrete conditional distributions with structural zeros, *J. Multivariate Anal.* 101 (2010) 191–199.
- [21] Y.-C. Yao, S.-C. Chen, S.-H. Wang, On compatibility of discrete full conditional distributions: a graphical representation approach, *J. Multivariate Anal.* 124 (2014) 1–9.