Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 53 (3) (2024), 803 – 822
DOI : 10.15672/hujms.1216921

Research Article

Closed-form estimates for missing counts in multidimensional incomplete tables

Sayan Ghosh*[1], Palaniappan Vellaisamy[2]

[1] *Department of Mathematics, Birla Institute of Technology and Sciences Pilani, Hyderabad Campus, Hyderabad 500078, India*

[2] *Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India*

Abstract

A useful technique for analyzing incomplete tables is to model the missing data mechanisms of the variables using log-linear models. In this paper, we use log-linear parametrization and propose estimation methods for arbitrary three-way and n-dimensional incomplete tables. All possible cases in which data on one or more of the variables may be missing are considered. We provide simple closed form estimates of expected cell counts and parameters for the various missing data models. We also obtain explicit boundary estimates under nonignorable nonresponse models. Finally, a real-life dataset is analyzed to illustrate our results for modelling and estimation in multidimensional incomplete tables.

Mathematics Subject Classification (2020). 62H17, 62D10, 62H12

Keywords. Incomplete tables, ML estimation, boundary solutions, log-linear models, NMAR models

1. Introduction

Contingency tables are frequently used for the display and analysis of categorical data. Missing data in such tables pose a common problem in various epidemiological studies, clinical trials and social science studies. The results of analyses that improperly treat missing data can be biased and imprecise obscuring the underlying phenomena. So, the analysis of contingency tables with missing data, also called incomplete tables, is of practical interest. The two types of counts in such tables are (i) fully observed counts and (ii) partially classified margins (nonresponses). A systematic study of missing data involves three types of missingness mechanisms proposed in the literature (see [14]): missing completely at random (MCAR), missing at random (MAR) and not missing at random (NMAR). If the probability (of an observation being missing) is independent of both observed and unobserved data, then a mechanism is said to be MCAR. It is called MAR if conditional on the observed data, the probability is independent of unobserved data, and NMAR if the probability depends only on unobserved data. For likelihood inference, nonresponses are classified as either ignorable (when the missing data mechanism is MAR or MCAR) or nonignorable (when the missing data mechanism is NMAR).

*Corresponding Author.
Email addresses: sayan@hyderabad.bits-pilani.ac.in (S. Ghosh), pv@math.iitb.ac.in (P. Vellaisamy)
Received: 09.12.2022; Accepted: 26.12.2023

According to [14], an incomplete table may be analyzed using mainly the following techniques: complete case analysis (using only the fully observed counts), weighting, imputation and modelling. Various types of models for analyzing such incomplete tables are available in the literature, for example, the pattern-mixture model (see [10, 13, 17]), and the selection model (see [3, 9, 14]). Log-linear models have generally been used to study missing data mechanisms in incomplete tables (see [3–7, 20]). Some of the estimation methods used are weighted least squares, maximum likelihood (ML) and Bayesian techniques.

Baker et al. [4] used log-linear models for analyzing a two-way incomplete table with data missing in both variables, and obtained closed-form estimates of missing counts. In this paper, we adopt the hierarchical log-linear parametrization for arbitrary three-dimensional and n-dimensional incomplete tables in general (see [12]). We focus on log-linear models with main effects and two-way interactions among variables and their missing indicators. This is because higher order interactions are difficult to interpret and models with such parameters may become non-identifiable. We consider all possible cases when data on one or more of the variables are missing.

We derive explicit, closed-form formulae for estimates of expected cell counts under various missing data models in the above tables. The formulae involve only observed cell counts or their sums, which simplifies the fitting of the models. Closed-form estimates are important since they provide a compact, simplified algebraic expression unlike iterative algorithms. The convergence of an iterative algorithm is an increasing function of the input size, that is, sample size. So, larger the sample size (total cell count in the incomplete table), the longer it takes to obtain the estimates. However, evaluating closed-form expressions is independent of the sample size and requires constant time. Hence, closed-form estimates are usually faster to compute than iterative solutions thereby reducing computational burden. This is especially true when lots of iterations are required to compute the estimates of the cell probabilities if some of them are zero under nonignorable (NMAR) missing data models.

Incomplete tables with data missing in at least one variable are common in the social sciences and medical fields. For example, in the analysis of survey data, the gender of each respondent is usually known. Suppose we are interested in the association between two partially missing variables (say, income and education level), stratified by gender. This is an example of a three-way incomplete table with data on two variables missing. While the EM algorithm (see [8]) is available for such settings, it does not automatically produce asymptotic covariance matrices for the parameter estimates so that estimation of standard errors of the estimates becomes difficult. The rate of convergence of the EM algorithm also depends on the proportion of missing information for each parameter. So Meng and Rubin [16] proposed a componentwise EM procedure, which is computationally expensive for covariance estimation. In this paper, we explicitly model the missing data mechanism of each variable which leads to a full likelihood specification and use ML estimation to obtain the parameter estimates. Unlike the EM algorithm, covariance estimates of the parameters can be calculated in the usual way by inverting the Fisher information while any of the common fit statistics can be used to compare the fits of different models. Besides estimating missing cell counts, we obtain closed-form estimates of joint, marginal and conditional probabilities of the variables and their missingness under various models. Also, estimates of the marginal odds ratios and their asymptotic variances are provided for each model.

The problem of boundary solutions occurs in nonignorable models while using ML estimation. Such solutions occur when the MLE's of nonresponse cell probabilities are all zeros for certain levels of a variable, that is, they lie on the boundary of the parameter space. Some references to this problem for various incomplete tables include [4, 11, 12, 18]. In this paper, we provide explicit closed-form MLE's of expected cell counts and other parameters if boundary solutions occur under nonignorable models for some three-way incomplete tables.

The remaining part of the paper is organized as follows. In Section 2, we provide log-linear parametrizations and discuss estimation methods for three-way incomplete tables with data missing in one variable, two variables and all variables. We also discuss boundary solutions and their occurrence under NMAR models in each of the above tables. Section 3 extends the methodology and results in Section 2 to arbitrary n-dimensional incomplete tables. A real-life dataset is analyzed in Section 4 to illustrate the results in Section 3. Section 5 provides some concluding remarks.

2. Log-linear parametrization for 3-dimensional incomplete tables

For studying missing data mechanisms in an $I \times J \times 2 \times 2$ incomplete table, [4] considered nine identifiable log-linear models. In this section, we use such hierarchical log-linear models (see [12]) for three-way contingency tables where data on at least one of the variables may be missing. Partially classified (supplementary) margins of a table are assumed to be positive.

2.1. Missing in one of the variables

Without loss of generality (WLOG), let data on Y_1 be missing. Denote the missing indicator for Y_1 by R, where $R = 1$ if Y_1 is observed and $R = 2$ otherwise. Then we have an $I \times J \times K \times 2$ table corresponding to Y_1, Y_2, Y_3 and R with cell counts $\mathbf{y} = \{y_{ijkx}\}$, where $1 \le i \le I$, $1 \le j \le J$, $1 \le k \le K$ and $x = 1, 2$. Denote the vector of observed counts by $\mathbf{y_{obs}} = (\{y_{ijk1}\}, \{y_{+jk2}\})$, where $\{y_{ijk1}\}$ are the fully observed counts, $\{y_{+jk2}\}$ are the supplementary margins and '+' means summation over levels of the corresponding variable. Let $\pi = \{\pi_{ijkx}\}$ be the vector of cell probabilities, $\mu = \{\mu_{ijkx}\}$ be the vector of expected counts and $N = \sum_{i,j,k,x} y_{ijkx}$ be the total cell count. For $I = J = K = 2$, the $2 \times 2 \times 2 \times 2$ incomplete table is shown below. (Table 1).

Table 1. $2 \times 2 \times 2 \times 2$ Incomplete Table

			$Y_3 = 1$	$Y_3 = 2$
$R = 1$	$Y_1 = 1$	$Y_2 = 1$	y_{1111}	y_{1121}
		$Y_2 = 2$	y_{1211}	y_{1221}
	$Y_1 = 2$	$Y_2 = 1$	y_{2111}	y_{2121}
		$Y_2 = 2$	y_{2211}	y_{2221}
$R = 2$	Missing	$Y_2 = 1$	y_{+112}	y_{+122}
		$Y_2 = 2$	y_{+212}	y_{+222}

The log-linear model (with no three-way interactions) for this case is given by

$$\begin{aligned} \log \mu_{ijkx} \;=\; & \lambda + \lambda_{Y_1}(i) + \lambda_{Y_2}(j) + \lambda_{Y_3}(k) + \lambda_R(x) + \lambda_{Y_1Y_2}(i,j) + \lambda_{Y_1Y_3}(i,k) + \lambda_{Y_2Y_3}(j,k) \\ & + \lambda_{Y_1R}(i,x) + \lambda_{Y_2R}(j,x) + \lambda_{Y_3R}(k,x). \end{aligned} \tag{2.1}$$

Each log-linear parameter in (2.1) satisfies the constraint that the sum over each of its arguments is 0, for example, $\sum_i \lambda_{Y_1Y_3}(i,k) = \sum_k \lambda_{Y_1Y_3}(i,k) = 0$. Define $a_{ijk} = \frac{P(R=2|Y_1=i,Y_2=j,Y_3=k)}{P(R=1|Y_1=i,Y_2=j,Y_3=k)} = \frac{\pi_{ijk2}}{\pi_{ijk1}} = \frac{\mu_{ijk2}}{\mu_{ijk1}}$, which describes the missing data mechanism of Y_1. It is the odds of Y_1 being missing. Then $\mu_{ijk2} = a_{ijk}\mu_{ijk1}$. Also, $\sum_{i,j,k} \mu_{ijk1}(1 + a_{ijk}) = N$ and the joint probability $\pi_{ijk+} = \mu_{ijk1}(1 + a_{ijk})/N$, from which the marginals may be derived. Note that under (2.1), $a_{ijk} = \exp[-2\{\lambda_R(1) + \lambda_{Y_1R}(i,1) + \lambda_{Y_2R}(j,1) + \lambda_{Y_3R}(k,1)\}]$. Denote a_{ijk} by $\alpha_{i..}$ or $\alpha_{.j.}$ or $\alpha_{..k}$ or $\alpha_{...}$ if it depends on only i or j or k or none of these, respectively. From [12], we have the following definition.

Definition 2.1. The missing mechanism of Y_1 under (2.1) is NMAR if $a_{ijk} = \alpha_{i..}$, MAR if $a_{ijk} = \alpha_{.j.}$ or $\alpha_{..k}$ and MCAR if $a_{ijk} = \alpha_{...}$.

Under Poisson sampling for observed cell counts, the log-likelihood of μ is

$$l(\mu; \mathbf{y_{obs}}) = \sum_{i,j,k} y_{ijk1} \log \mu_{ijk1} + \sum_{j,k} y_{+jk2} \log \mu_{+jk2} - \sum_{i,j,k,x} \mu_{ijkx} + \Delta, \tag{2.2}$$

where Δ is some constant. The various missing data models and the MLE's under them are given as follows :

1. $\alpha_{i..}$ (NMAR for Y_1).
We have $\hat{\mu}_{ijk1} = y_{ijk1}$ and $\hat{\alpha}_{i..}$ satisfies $\sum_i \hat{\mu}_{ijk1}\hat{\alpha}_{i..} = y_{+jk2}\ \forall\ 1 \le j \le J, 1 \le k \le K$.
2. $\alpha_{.j.}$ (MAR for Y_1).
We have $\hat{\mu}_{ijk1} = \frac{y_{ijk1}y_{+jk+}y_{+j+1}}{y_{+jk1}y_{+j++}}$ and $\hat{\alpha}_{.j.} = \frac{y_{+j+2}}{y_{+j+1}}$.
3. $\alpha_{..k}$ (MAR for Y_1).
We have $\hat{\mu}_{ijk1} = \frac{y_{ijk1}y_{+jk+}y_{++k1}}{y_{+jk1}y_{++k+}}$ and $\hat{\alpha}_{..k} = \frac{y_{++k2}}{y_{++k1}}$.
4. $\alpha_{...}$ (MCAR for Y_1).
We have $\hat{\mu}_{ijk1} = \frac{y_{ijk1}y_{+jk+}y_{+++1}}{y_{+jk1}y_{++++}}$ and $\hat{\alpha}_{...} = \frac{y_{+++2}}{y_{+++1}}$.

From [12], boundary solutions occur if $\hat{\alpha}_{i..} \le 0$ for at least one and at most $(I-1)$ values of Y_1. If any $\hat{\alpha}_{i..} < 0$, then boundary estimates are obtained by setting $\hat{\alpha}_{i..} = 0$ in (2.2). For example, if Y_1 is binary with levels 1 and 2, and $\hat{\alpha}_{1..} = 0$ under Model 1, then the MLE's are

$$\hat{\alpha}_{2..} = \frac{y_{+++2}}{y_{2++1}},\ \hat{\mu}_{1jk1} = y_{1jk1},\ \hat{\mu}_{2jk1} = \frac{(y_{2jk1} + y_{+jk2})y_{2++1}}{y_{+++2}}.$$

A perfect fit model is one for which the estimated expected counts are equal to the observed counts. Consider now the hypotheses H_0: the proposed model (among Models 1-4 mentioned above) fits the data, and H_1: the perfect fit model fits the data. Let L_0 and L_1 denote the maximized log-likelihood functions under the proposed and perfect fit models respectively. Then the likelihood ratio statistic for testing H_0 against H_1 is given by

$$\begin{aligned} G^2 &= -2(L_0 - L_1) \\ &= -2\left[\sum_{i,j,k} y_{ijk1} \ln\left(\frac{\hat{\mu}_{ijk1}}{y_{ijk1}}\right) + \sum_{j,k} y_{+jk2} \ln\left(\frac{\sum_i \hat{\mu}_{ijk1}\hat{a}_{ijk}}{y_{+jk2}}\right) - \sum_{i,j,k} \hat{\mu}_{ijk1}(1+\hat{a}_{ijk}) + N\right]. \end{aligned} \tag{2.3}$$

Note that G^2 follows χ^2_ν asymptotically, where $\nu = (I+1)JK$ (number of observed counts) $-$ number of free estimable parameters under the proposed model. If Y_1 is binary and boundary solutions occur under Model 1, then the boundary MLE's are obtained for the level of Y_1 corresponding to which G^2 is minimum.

2.2. Missing in two of the variables

WLOG, suppose data on Y_1 and Y_2 are missing. For $i = 1, 2$, denote the missing indicator for Y_i by R_i such that $R_i = 1$ if Y_i is observed and $R_i = 2$ otherwise. Then we have an $I \times J \times K \times 2 \times 2$ table corresponding to Y_1, Y_2, Y_3, R_1 and R_2 with cell counts $\mathbf{y} = \{y_{ijkxs}\}$, where $1 \le i \le I$, $1 \le j \le J$, $1 \le k \le K$ and $x, s = 1, 2$. Denote the vector of observed counts by $\mathbf{y_{obs}} = (\{y_{ijk11}\}, \{y_{+jk21}\}, \{y_{i+k12}\}, \{y_{++k22}\})$. Also, let $\pi = \{\pi_{ijkxs}\}$ be the vector of cell probabilities, $\mu = \{\mu_{ijkxs}\}$ be the vector of expected counts and N be the total cell count. For $I = J = K = 2$, the $2 \times 2 \times 2 \times 2 \times 2$ incomplete table is shown below (Table 2).

Table 2. $2 \times 2 \times 2 \times 2 \times 2$ Incomplete Table

				$Y_3 = 1$	$Y_3 = 2$
$R_1 = 1$	$Y_1 = 1$	$R_2 = 1$	$Y_2 = 1$	y_{11111}	y_{11211}
			$Y_2 = 2$	y_{12111}	y_{12211}
		$R_2 = 2$	Missing	y_{1+112}	y_{1+212}
	$Y_1 = 2$	$R_2 = 1$	$Y_2 = 1$	y_{21111}	y_{21211}
			$Y_2 = 2$	y_{22111}	y_{22211}
		$R_2 = 2$	Missing	y_{2+112}	y_{2+212}
$R_1 = 2$	Missing	$R_2 = 1$	$Y_2 = 1$	y_{+1121}	y_{+1221}
			$Y_2 = 2$	y_{+2121}	y_{+2221}
		$R_2 = 2$	Missing	y_{++122}	y_{++222}

The log-linear model (without three-way or higher order interactions) is given by

$$\begin{aligned} \log \mu_{ijkxs} &= \lambda + \lambda_{Y_1}(i) + \lambda_{Y_2}(j) + \lambda_{Y_3}(k) + \lambda_{R_1}(x) + \lambda_{R_2}(s) + \lambda_{Y_1Y_2}(i,j) + \lambda_{Y_1Y_3}(i,k) \\ &\quad + \lambda_{Y_2Y_3}(j,k) + \lambda_{Y_1R_1}(i,x) + \lambda_{Y_2R_1}(j,x) + \lambda_{Y_3R_1}(k,x) \\ &\quad + \lambda_{Y_1R_2}(i,s) + \lambda_{Y_2R_2}(j,s) + \lambda_{Y_3R_2}(k,s) + \lambda_{R_1R_2}(x,s). \end{aligned} \tag{2.4}$$

Each log-linear parameter in (2.4) satisfies the constraint that the sum over each of its arguments is 0. Define the following quantities

$$\begin{aligned} a_{ijk} &= \frac{P(R_1 = 2, R_2 = 1 \mid Y_1 = i, Y_2 = j, Y_3 = k)}{P(R_1 = 1, R_2 = 1 \mid Y_1 = i, Y_2 = j, Y_3 = k)} = \frac{\pi_{ijk21}}{\pi_{ijk11}} = \frac{\mu_{ijk21}}{\mu_{ijk11}}, \\ b_{ijk} &= \frac{P(R_1 = 1, R_2 = 2 \mid Y_1 = i, Y_2 = j, Y_3 = k)}{P(R_1 = 1, R_2 = 1 \mid Y_1 = i, Y_2 = j, Y_3 = k)} = \frac{\pi_{ijk12}}{\pi_{ijk11}} = \frac{\mu_{ijk12}}{\mu_{ijk11}}. \end{aligned}$$

Then the missing data mechanisms of Y_1 and Y_2 are described by a_{ijk} and b_{ijk}, respectively. Note that a_{ijk} is the conditional odds of Y_1 being missing given Y_2 is observed, while b_{ijk} is the conditional odds of Y_2 being missing given Y_1 is observed. The odds ratio between R_1 and R_2 is

$$\begin{aligned} \theta &= \frac{P(R_1 = 1, R_2 = 1 \mid Y_1 = i, Y_2 = j, Y_3 = k)P(R_1 = 2, R_2 = 2 \mid Y_1 = i, Y_2 = j, Y_3 = k)}{P(R_1 = 1, R_2 = 2 \mid Y_1 = i, Y_2 = j, Y_3 = k)P(R_1 = 2, R_2 = 1 \mid Y_1 = i, Y_2 = j, Y_3 = k)} \\ &= \frac{\pi_{ijk11}\pi_{ijk22}}{\pi_{ijk12}\pi_{ijk21}} = \frac{\mu_{ijk11}\mu_{ijk22}}{\mu_{ijk12}\mu_{ijk21}}. \end{aligned}$$

If $\theta = 1$, then the missingness patterns of Y_1 and Y_2, that is, R_1 and R_2 are independent. Also, $\mu_{ijk21} = a_{ijk}\mu_{ijk11}, \mu_{ijk12} = b_{ijk}\mu_{ijk11}$, $\mu_{ijk22} = \mu_{ijk11}a_{ijk}b_{ijk}\theta$ and $N = \sum_{i,j,k} \mu_{ijk11}(1 + a_{ijk} + b_{ijk} + a_{ijk}b_{ijk}\theta)$. The joint probability is $\pi_{ijk++} = \mu_{ijk11}(1 + a_{ijk} + b_{ijk} + a_{ijk}b_{ijk}\theta)/N$, from which the marginals can be obtained. The conditional probability of Y_1 being missing given that Y_2 is observed is

$$\phi_{1|2}(i,j,k) = P(R_1 = 2 \mid R_2 = 1, Y_1 = i, Y_2 = j, Y_3 = k) = \frac{a_{ijk}}{1 + a_{ijk}}.$$

Similarly, the conditional probability of Y_2 being missing given that Y_1 is observed is

$$\phi_{2|1}(i,j,k) = P(R_2 = 2 \mid R_1 = 1, Y_1 = i, Y_2 = j, Y_3 = k) = \frac{b_{ijk}}{1 + b_{ijk}}.$$

Under (2.4), we have $a_{ijk} = \exp[-2\{\lambda_{R_1}(1) + \lambda_{Y_1R_1}(i,1) + \lambda_{Y_2R_1}(j,1) + \lambda_{Y_3R_1}(k,1) + \lambda_{R_1R_2}(1,1)\}]$, $b_{ijk} = \exp[-2\{\lambda_{R_2}(1) + \lambda_{Y_1R_2}(i,1) + \lambda_{Y_2R_2}(j,1) + \lambda_{Y_3R_2}(k,1) + \lambda_{R_1R_2}(1,1)\}]$ and
$\theta = \exp[4\lambda_{R_1R_2}(1,1)]$. If each of a_{ijk} and b_{ijk} depends on only one of i, j, k or none of these, then let $a_{ijk} \in \{\alpha_{i..}, \alpha_{.j.}, \alpha_{..k}, \alpha_{...}\}$ and $b_{ijk} \in \{\beta_{i..}, \beta_{.j.}, \beta_{..k}, \beta_{...}\}$. The next definition is due to [12].

Definition 2.2. The missing mechanism of Y_1 under (2.4) is NMAR if $a_{ijk} = \alpha_{i..}$, MAR if $a_{ijk} = \alpha_{.j.}$ or $\alpha_{..k}$ and MCAR if $a_{ijk} = \alpha_{...}$, respectively. Similarly, the missing mechanism of Y_2 is NMAR if $b_{ijk} = \beta_{.j.}$, MAR if $b_{ijk} = \beta_{i..}$ or $\beta_{..k}$ and MCAR if $b_{ijk} = \beta_{...}$.

Under Poisson sampling, the log-likelihood kernel of μ is

$$\begin{aligned} l(\mu; \mathbf{y_{obs}}) &= \sum_{i,j,k} y_{ijk11} \log \mu_{ijk11} + \sum_{j,k} y_{+jk21} \log \mu_{+jk21} + \sum_{i,k} y_{i+k12} \log \mu_{i+k12} \\ &\quad + \sum_{k} y_{++k22} \log \mu_{++k22} - \sum_{i,j,k,x,s} \mu_{ijkxs}. \end{aligned} \tag{2.5}$$

There are 16 identifiable models in this case. The various models and the MLE's under them are given in the Appendix. From [12], boundary solutions occur under at least one of the following cases.

1. $\hat{\alpha}_{i..} \leq 0$ for at least one and at most $(I-1)$ values of Y_1,
2. $\hat{\beta}_{.j.} \leq 0$ for at least one and at most $(J-1)$ values of Y_2.

They occur in models for which the missing mechanism of at least one of the variables is NMAR. If any $\hat{\alpha}_{i..} < 0$ or any $\hat{\beta}_{.j.} < 0$, then boundary estimates can still be obtained by setting $\hat{\alpha}_{i..} = 0$ or $\hat{\beta}_{.j.} = 0$ in (2.5) for relevant models. Now suppose Y_1 and Y_2 are binary variables, each with levels 1 and 2. Then we have a $2 \times 2 \times K \times 2 \times 2$ incomplete contingency table. The boundary MLE's obtained when $\hat{\alpha}_{1..} = 0$ or $\hat{\beta}_{.2.} = 0$ (say) under various NMAR models are shown below.

(a) $(\alpha_{i..}, \beta_{...})$ (NMAR for Y_1, MCAR for Y_2) :
If $\hat{\alpha}_{1..} = 0$, then the MLE's are

$$\hat{\alpha}_{2..} = \frac{y_{+++21}y_{+++1+}}{y_{+++11}y_{2++1+}},\ \hat{\beta}_{...} = \frac{y_{+++12}}{y_{+++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},$$
$$\hat{\mu}_{1jk11} = \frac{y_{1jk11}y_{1++1+}y_{+++11}}{y_{1++11}y_{+++1+}},\ \hat{\mu}_{2jk11} = \frac{y_{+++11}y_{2++1+}(y_{2jk11} + y_{+jk21})}{y_{+++1+}(y_{2++11} + y_{+++21})}.$$

(b) $(\alpha_{i..}, \beta_{i..})$ (NMAR for Y_1, MAR for Y_2) :
If $\hat{\alpha}_{1..} = 0$, then the MLE's are

$$\hat{\alpha}_{2..} = \frac{y_{+++21}}{y_{2++11}},\ \hat{\beta}_{i..} = \frac{y_{i++12}}{y_{i++11}},\ \hat{\theta} = \frac{y_{2++11}y_{+++22}}{y_{2++12}y_{+++21}},$$
$$\hat{\mu}_{1jk11} = y_{1jk11},\ \hat{\mu}_{2jk11} = \frac{y_{2++11}(y_{2jk11} + y_{+jk21})}{y_{2++11} + y_{+++21}}.$$

(c) $(\alpha_{i..}, \beta_{.j.})$ (NMAR for both Y_1 and Y_2) :
(i) If $\hat{\alpha}_{1..} = 0$, then the MLE's are

$$\hat{\alpha}_{2..} = \frac{y_{+++21}}{y_{2++11}},\ \hat{\theta} = \frac{y_{2++11}y_{+++22}}{y_{2++12}y_{+++21}},\ \hat{\mu}_{1jk11} = y_{1jk11},\ \hat{\mu}_{2jk11} = \frac{y_{2++11}(y_{2jk11} + y_{+jk21})}{y_{2++11} + y_{+++21}}.$$

Also, $\hat{\beta}_{.j.}$ satisfies $\sum_j \hat{\mu}_{ijk11}\hat{\beta}_{.j.} = y_{i+k12}$.
(ii) If $\hat{\beta}_{.2.} = 0$, then the MLE's are

$$\hat{\beta}_{.1.} = \frac{y_{+++12}}{y_{+1+11}},\ \hat{\theta} = \frac{y_{+1+11}y_{+++22}}{y_{+1+12}y_{+++21}},\ \hat{\mu}_{i1k11} = \frac{y_{+1+11}(y_{i1k11} + y_{i+k12})}{y_{+1+11} + y_{+++21}},\ \hat{\mu}_{i2k11} = y_{i2k11}.$$

Also, $\hat{\alpha}_{i..}$ satisfies $\sum_i \hat{\mu}_{ijk11}\hat{\alpha}_{i..} = y_{+jk21}$.
(d) $(\alpha_{...}, \beta_{.j.})$ (NMAR for Y_2, MCAR for Y_1) :
If $\hat{\beta}_{.2.} = 0$, then the MLE's are

$$\hat{\beta}_{.1.} = \frac{y_{+++12}y_{++++1}}{y_{+++11}y_{+1++1}},\ \hat{\alpha}_{...} = \frac{y_{+++12}}{y_{+++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},$$
$$\hat{\mu}_{i1k11} = \frac{y_{+++11}y_{+1++1}(y_{i1k11} + y_{i+k12})}{y_{++++1}(y_{+1+11} + y_{+++12})},\ \hat{\mu}_{i2k11} = \frac{y_{i2k11}y_{+2++1}y_{+++11}}{y_{+2+11}y_{++++1}}.$$

(e) $(\alpha_{.j.}, \beta_{.j.})$ (NMAR for Y_2, MAR for Y_1) :
If $\hat{\beta}_{.2.} = 0$, then the MLE's are

$$\begin{aligned}
\hat{\beta}_{.1.} &= \frac{y_{+++12}}{y_{+1+11}},\ \hat{\alpha}_{.j.} = \frac{y_{+jk21}}{y_{+jk11}},\ \hat{\theta} = \frac{y_{+1+11}y_{+++22}}{y_{+1+12}y_{+++21}}, \\
\hat{\mu}_{i1k11} &= \frac{y_{+1+11}(y_{i1k11} + y_{i+k12})}{y_{+1+11} + y_{+++21}},\ \hat{\mu}_{i2k11} = y_{i2k11}.
\end{aligned}$$

The above method for obtaining closed-form boundary MLE's can be generalized to non-binary variables also. Next consider testing the hypotheses H_0: the proposed model (among Models 1-16 in the Appendix) fits the data against H_1: the perfect fit model fits the data. Let L_0 and L_1 denote the maximized log-likelihood functions under the proposed and perfect fit models respectively. Then the likelihood ratio statistic for testing H_0 against H_1 is given by

$$\begin{aligned}
G^2 &= -2(L_0 - L_1) \\
&= -2\left[\sum_{i,j,k} y_{ijk11} \ln\left(\frac{\hat{\mu}_{ijk11}}{y_{ijk11}}\right) + \sum_{j,k} y_{+jk21} \ln\left(\frac{\sum_i \hat{\mu}_{ijk11}\hat{a}_{ijk}}{y_{+jk21}}\right)\right. \\
&+ \sum_{i,k} y_{i+k12} \ln\left(\frac{\sum_j \hat{\mu}_{ijk11}\hat{b}_{ijk}}{y_{i+k12}}\right) + \sum_{k} y_{++k22} \ln\left(\frac{\sum_{i,j} \hat{\mu}_{ijk11}\hat{a}_{ijk}\hat{b}_{ijk}\hat{\theta}}{y_{++k22}}\right) \\
&\left. - \sum_{i,j,k} \hat{\mu}_{ijk11}(1 + \hat{a}_{ijk} + \hat{b}_{ijk} + \hat{a}_{ijk}\hat{b}_{ijk}\hat{\theta}) + N\right]. \qquad (2.6)
\end{aligned}$$

Note that G^2 follows χ^2_ν asymptotically, where $\nu = (I+1)(J+1)K-$ number of free estimable parameters under the proposed model. If Y_1 and Y_2 are binary variables and boundary solutions occur, then the boundary MLE's are obtained for the level of Y_1 or Y_2 (depending on whether $\hat{\alpha}_{i..} < 0$ or $\hat{\beta}_{.j.} < 0$) corresponding to which G^2 is minimum.

Marginal odds ratios. When $Y_3 = k$ is fixed, consider the Y_1Y_2-marginal odds ratios. Let $OR_{..k} = (\hat{\pi}_{ijk..}\hat{\pi}_{i'j'k..})/(\hat{\pi}_{ij'k..}\hat{\pi}_{i'jk..})$ denote an estimated odds ratio on the Y_1Y_2-margin, where $1 \le i < i' \le I$, $1 \le j < j' \le J$ and $1 \le k \le K$. Also, let $OR_{11k} = (y_{ijk11}y_{i'j'k11})/(y_{ij'k11}y_{i'jk11})$ be the estimated odds ratio when $R_1 = R_2 = 1$. From the closed-form MLE's for the models (see Appendix), it can be shown that $OR_{..k} = OR_{11k}$ under Models 2, 4, 9, 12, 13, 14 and 16 *a priori*, and under Models 1, 3, 5, 6, 8, 11 and 15 for non-boundary (interior) estimates. We can derive closed-form expressions for the asymptotic variance of estimated marginal odds ratio in case of non-boundary MLE's. We assume that the data follows Poisson distribution. The asymptotic variance of a statistic $f(\{y_{ijk11}\}, \{y_{i+k12}\}, \{y_{+jk21}\}, y_{++k22})$ for fixed k (see [1]) is

$$\begin{aligned}
Var(f) &= \sum_{i,j}\left(\frac{\partial f}{\partial y_{ijk11}}\right)^2 \hat{\mu}_{ijk11} + \sum_{i}\left(\frac{\partial f}{\partial y_{i+k12}}\right)^2 \hat{\mu}_{i+k12} + \sum_{j}\left(\frac{\partial f}{\partial y_{+jk21}}\right)^2 \hat{\mu}_{+jk21} \\
&+ \left(\frac{\partial f}{\partial y_{++k22}}\right)^2 \hat{\mu}_{++k22}. \qquad (2.7)
\end{aligned}$$

When $OR_{..k} = OR_{11k} = (y_{ijk11}y_{i'j'k11})/(y_{ij'k11}y_{i'jk11})$, we get from (2.7)

$$Var(OR_{..k}) = OR^2_{..k}\left[\frac{\hat{\mu}_{ijk11}}{y^2_{ijk11}} + \frac{\hat{\mu}_{ij'k11}}{y^2_{ij'k11}} + \frac{\hat{\mu}_{i'jk11}}{y^2_{i'jk11}} + \frac{\hat{\mu}_{i'j'k11}}{y^2_{i'j'k11}}\right]. \qquad (2.8)$$

Using (2.8), the asymptotic variances of estimated marginal odds ratios for k fixed under various models (see Appendix) are as follows.

1. Models 2, 3 and 4 :

$$Var(OR_{..k}) = OR_{..k}^2 \frac{y_{++k11}}{y_{++k+1}} \left[\frac{y_{+jk+1}}{y_{+jk11}} \left(\frac{1}{y_{ijk11}} + \frac{1}{y_{i'jk11}} \right) + \frac{y_{+j'k+1}}{y_{+j'k11}} \left(\frac{1}{y_{ij'k11}} + \frac{1}{y_{i'j'k11}} \right) \right]$$

2. Models 5, 9 and 13 :

$$Var(OR_{..k}) = OR_{..k}^2 \frac{y_{++k11}}{y_{++k1+}} \left[\frac{y_{i+k1+}}{y_{i+k11}} \left(\frac{1}{y_{ijk11}} + \frac{1}{y_{ij'k11}} \right) + \frac{y_{i'+k1+}}{y_{i'+k11}} \left(\frac{1}{y_{i'jk11}} + \frac{1}{y_{i'j'k11}} \right) \right]$$

3. Models 6, 8, 11, 12, 14, 15 and 16 :

$$Var(OR_{..k}) = OR_{..k}^2 \left[\frac{1}{y_{ijk11}} + \frac{1}{y_{ij'k11}} + \frac{1}{y_{i'jk11}} + \frac{1}{y_{i'j'k11}} \right]$$

It may be remarked that this variance approximation is based on a Taylor series linearization method (sometimes called the delta method). Alternatively, the variances can be computed from the inverse of the observed information matrix using the method in [2]. Note that if boundary solutions occur under NMAR models, then this method provides a variance estimate given that the closed-form MLE's of the expected cell counts in (2.8) lie on the boundary of the parameter space. However, the bootstrap technique provides an unconditional variance estimate in this case (see [3]).

2.3. Missing in all three variables

For $i = 1, 2, 3$, denote R_i to be the missing indicator of Y_i, where $R_i = 1$ if Y_i is observed and $R_i = 2$ otherwise. Then we have an $I \times J \times K \times 2 \times 2 \times 2$ table corresponding to Y_1, Y_2, Y_3, R_1, R_2 and R_3 with cell counts $\mathbf{y} = \{y_{ijkxsz}\}$, where $1 \le i \le I$, $1 \le j \le J$, $1 \le k \le K$ and $x, s, z = 1, 2$. Also, $\mathbf{y_{obs}} = (\{y_{ijk111}\}, \{y_{+jk211}\}, \{y_{i+k121}\}, \{y_{ij+112}\}, \{y_{++k221}\}, \{y_{+j+212}\}, \{y_{i++122}\}, y_{+++222})$. Let $\pi = \{\pi_{ijkxsz}\}$ be the vector of cell probabilities, N be the total cell count and $\mu = \{\mu_{ijkxsz}\}$ be the vector of expected counts. For $I = J = K = 2$, the $2 \times 2 \times 2 \times 2 \times 2 \times 2$ incomplete table is shown below (Table 3).

Table 3. $2 \times 2 \times 2 \times 2 \times 2 \times 2$ Incomplete Table

				$R_3 = 1$ $Y_3 = 1$	 $Y_3 = 2$	$R_3 = 2$ Missing
$R_1 = 1$	$Y_1 = 1$	$R_2 = 1$	$Y_2 = 1$	y_{11111}	y_{112111}	y_{11+112}
			$Y_2 = 2$	y_{121111}	y_{122111}	y_{12+112}
		$R_2 = 2$	Missing	y_{1+1121}	y_{1+2121}	y_{1++122}
	$Y_1 = 2$	$R_2 = 1$	$Y_2 = 1$	y_{211111}	y_{212111}	y_{21+112}
			$Y_2 = 2$	y_{212111}	y_{222111}	y_{22+112}
		$R_2 = 2$	Missing	y_{2+1121}	y_{2+2121}	y_{2++122}
$R_1 = 2$	Missing	$R_2 = 1$	$Y_2 = 1$	y_{+11211}	y_{+12211}	y_{+1+212}
			$Y_2 = 2$	y_{+21211}	y_{+22211}	y_{+2+212}
		$R_2 = 2$	Missing	y_{++1221}	y_{++2221}	y_{+++222}

The log-linear model in this case is

$$\begin{aligned}
\log \mu_{ijkxsz} \quad = \quad & \lambda + \lambda_{Y_1}(i) + \lambda_{Y_2}(j) + \lambda_{Y_3}(k) + \lambda_{R_1}(x) + \lambda_{R_2}(s) + \lambda_{R_3}(z) + \lambda_{Y_1Y_2}(i,j) \\
& + \lambda_{Y_1Y_3}(i,k) + \lambda_{Y_2Y_3}(j,k) + \lambda_{Y_1R_1}(i,x) + \lambda_{Y_2R_1}(j,x) + \lambda_{Y_3R_1}(k,x) \\
& + \lambda_{Y_1R_2}(i,s) + \lambda_{Y_2R_2}(j,s) + \lambda_{Y_3R_2}(k,s) + \lambda_{Y_1R_3}(i,z) + \lambda_{Y_2R_3}(j,z) \\
& + \lambda_{Y_3R_3}(k,z) + \lambda_{R_1R_2}(x,s) + \lambda_{R_1R_3}(x,z) + \lambda_{R_2R_3}(s,z). \qquad (2.9)
\end{aligned}$$

Each log-linear parameter in (2.9) satisfies the constraint that the sum over each of its arguments is 0. Define the following quantities

$$
\begin{aligned}
a_{ijk} &= \frac{P(R_1=2,R_2=1,R_3=1 \mid Y_1=i,Y_2=j,Y_3=k)}{P(R_1=1,R_2=1,R_3=1 \mid Y_1=i,Y_2=j,Y_3=k)} = \frac{\pi_{ijk211}}{\pi_{ijk111}} = \frac{\mu_{ijk211}}{\mu_{ijk111}},\\
b_{ijk} &= \frac{P(R_1=1,R_2=2,R_3=1 \mid Y_1=i,Y_2=j,Y_3=k)}{P(R_1=1,R_2=1,R_3=1 \mid Y_1=i,Y_2=j,Y_3=k)} = \frac{\pi_{ijk121}}{\pi_{ijk111}} = \frac{\mu_{ijk121}}{\mu_{ijk111}},\\
c_{ijk} &= \frac{P(R_1=1,R_2=1,R_3=2 \mid Y_1=i,Y_2=j,Y_3=k)}{P(R_1=1,R_2=1,R_3=1 \mid Y_1=i,Y_2=j,Y_3=k)} = \frac{\pi_{ijk112}}{\pi_{ijk111}} = \frac{\mu_{ijk112}}{\mu_{ijk111}}.
\end{aligned}
$$

Then a_{ijk}, b_{ijk} and c_{ijk} describe the missing data mechanisms of Y_1, Y_2 and Y_3, respectively. Here a_{ijk} is the conditional odds of Y_1 being missing given both Y_2 and Y_3 are observed, b_{ijk} is the conditional odds of Y_2 being missing given both Y_1 and Y_3 are observed, and c_{ijk} is the conditional odds of Y_3 being missing given both Y_1 and Y_2 are observed. Let the conditional odds ratio between R_1 and R_2 given Y_3 is observed be

$$
\begin{aligned}
\theta_{12} &= \frac{P(R_1=1,R_2=1,R_3=1|Y_1=i,Y_2=j,Y_3=k)}{P(R_1=1,R_2=2,R_3=1|Y_1=i,Y_2=j,Y_3=k)}\\
&\quad\times\frac{P(R_1=2,R_2=2,R_3=1|Y_1=i,Y_2=j,Y_3=k)}{P(R_1=2,R_2=1,R_3=1|Y_1=i,Y_2=j,Y_3=k)}\\
&= \frac{\pi_{ijk111}\pi_{ijk221}}{\pi_{ijk121}\pi_{ijk211}} = \frac{\mu_{ijk111}\mu_{ijk221}}{\mu_{ijk121}\mu_{ijk211}}.
\end{aligned}
$$

Similarly, define θ_{13} to be the conditional odds ratio between R_1 and R_3 given Y_2 is observed, and θ_{23} to be the conditional odds ratio between R_2 and R_3 given Y_1 is observed. Also, define

$$
\begin{aligned}
\theta_{123} &= \frac{P(R_1=2,R_2=2,R_3=2 \mid Y_1=i,Y_2=j,Y_3=k)}{P(R_1=1,R_2=1,R_3=1 \mid Y_1=i,Y_2=j,Y_3=k)}\\
&= \frac{\pi_{ijk222}}{\pi_{ijk111}} = \frac{\mu_{ijk222}}{\mu_{ijk111}}.
\end{aligned}
$$

Here, θ_{12}, θ_{13} and θ_{23} describe the conditional associations between the missing mechanisms of Y_1 and Y_2, Y_1 and Y_3, and Y_2 and Y_3 respectively. For $i \neq j \neq k = 1,2,3$, if $\theta_{ij} = 1$, then the missing mechanisms of Y_i and Y_j are conditionally independent given that Y_k is observed. Note that θ_{123} denotes the joint odds of Y_1, Y_2 and Y_3 simultaneously missing. The joint probability is $\pi_{ijk+++} = \mu_{ijk111}(1 + a_{ijk} + b_{ijk} + c_{ijk} + a_{ijk}b_{ijk}\theta_{12} + a_{ijk}c_{ijk}\theta_{13} + b_{ijk}c_{ijk}\theta_{23} + \theta_{123})/N$, from which the marginals can be obtained. Under (2.9), we have

$$
\begin{aligned}
a_{ijk} &= \exp[-2\{\lambda_{R_1}(1) + \lambda_{Y_1R_1}(i,1) + \lambda_{Y_2R_1}(j,1) + \lambda_{Y_3R_1}(k,1) + \lambda_{R_1R_2}(1,1) + \lambda_{R_1R_3}(1,1)\}],\\
b_{ijk} &= \exp[-2\{\lambda_{R_2}(1) + \lambda_{Y_1R_2}(i,1) + \lambda_{Y_2R_2}(j,1) + \lambda_{Y_3R_2}(k,1) + \lambda_{R_1R_2}(1,1) + \lambda_{R_2R_3}(1,1)\}],\\
c_{ijk} &= \exp[-2\{\lambda_{R_3}(1) + \lambda_{Y_1R_3}(i,1) + \lambda_{Y_2R_3}(j,1) + \lambda_{Y_3R_3}(k,1) + \lambda_{R_1R_3}(1,1) + \lambda_{R_2R_3}(1,1)\}],
\end{aligned}
$$

$$
\begin{aligned}
\theta_{12} &= \exp[4\lambda_{R_1R_2}(1,1)],\ \theta_{13} = \exp[4\lambda_{R_1R_3}(1,1)],\ \theta_{23} = \exp[4\lambda_{R_2R_3}(1,1)],\\
\theta_{123} &= \exp[-2\{\lambda_{R_1}(1) + \lambda_{R_2}(1) + \lambda_{R_3}(1) + \lambda_{Y_1R_1}(i,1) + \lambda_{Y_2R_1}(j,1) + \lambda_{Y_3R_1}(k,1)\\
&+ \lambda_{Y_1R_2}(i,1) + \lambda_{Y_2R_2}(j,1) + \lambda_{Y_3R_2}(k,1) + \lambda_{Y_1R_3}(i,1) + \lambda_{Y_2R_3}(j,1) + \lambda_{Y_3R_3}(k,1)\}].
\end{aligned}
$$

Based on the assumption in the previous case regarding the missing mechanism of a variable, $a_{ijk} \in \{\alpha_{...}, \alpha_{i..}, \alpha_{.j.}, \alpha_{..k}\}$, $b_{ijk} \in \{\beta_{...}, \beta_{i..}, \beta_{.j.}, \beta_{..k}\}$ and $c_{ijk} \in \{\gamma_{...}, \gamma_{i..}, \gamma_{.j.}, \gamma_{..k}\}$ (say). For the definition below, see [12].

Definition 2.3. The missing mechanism of Y_1 under (2.9) is NMAR if $a_{ijk} = \alpha_{i..}$, MAR if $a_{ijk} = \alpha_{.j.}$ or $\alpha_{..k}$ and MCAR if $a_{ijk} = \alpha_{...}$. Similarly, the missing mechanism of Y_2 is NMAR if $b_{ijk} = \beta_{.j.}$, MAR if $b_{ijk} = \beta_{i..}$ or $\beta_{..k}$ and MCAR if $b_{ijk} = \beta_{...}$. Finally, the missing mechanism of Y_3 is NMAR if $c_{ijk} = \gamma_{..k}$, MAR if $c_{ijk} = \gamma_{i..}$ or $\gamma_{.j.}$ and MCAR if $c_{ijk} = \gamma_{...}$.

We have 64 possible identifiable models which are mixtures of the various missing mechanisms of the variables. Under Poisson sampling, the log-likelihood can be wriiten as a function of $a_{ijk}, b_{ijk}, c_{ijk}, \theta_{12}, \theta_{13}, \theta_{23}$ and θ_{123} which is then maximized to obtain closed-form MLE's of μ_{ijkxsz} under various missing data models. Note that from [12], boundary solutions occur if at least one of the following holds.

1. $\hat{\alpha}_{i..} \leq 0$ for at least one and at most $(I-1)$ values of Y_1,
2. $\hat{\beta}_{.j.} \leq 0$ for at least one and at most $(J-1)$ values of Y_2,
3. $\hat{\gamma}_{..k} \leq 0$ for at least one and at most $(K-1)$ values of Y_3.

The boundary estimates are obtained by setting $\hat{\alpha}_{i..} = 0$ or $\hat{\beta}_{.j.} = 0$ or $\hat{\gamma}_{..k} = 0$ in the log-likelihood for relevant models. The likelihood ratio statistic G^2 for testing the goodness of fit of a missing data model can be obtained as in the previous case. Here G^2 follows χ^2_ν asymptotically, where $\nu = (I+1)(J+1)(K+1)-$ number of free estimable parameters under the proposed model.

Remark 2.4. For all the above cases, perfect fit solutions for fully observed counts occur under the following types of models:

(i) non-boundary cases of NMAR only models for one or more variables,
(ii) non-boundary cases of a mixture of NMAR and MAR models for the variables,
(iii) MAR only models for two or more variables.

However, if the missing mechanism is MCAR for at least one of the variables, then perfect fit solutions don't occur.

WLOG, consider models in which the missing mechanism is NMAR for Y_1. Then we have the following observations.

Remark 2.5. The systems of equations $\sum_i \hat{\mu}_{ijk1}\hat{\alpha}_{i..} = y_{+jk2}$, $\sum_i \hat{\mu}_{ijk11}\hat{\alpha}_{i..} = y_{+jk21}$ and $\sum_i \hat{\mu}_{ijk111}\hat{\alpha}_{i...} = y_{+jk211}$ for $I \times J \times K \times 2$, $I \times J \times K \times 2 \times 2$ and $I \times J \times K \times 2 \times 2 \times 2$ tables respectively are overdetermined (underdetermined) if $I < JK$ $(I > JK)$.

Remark 2.6. Let the matrix of coefficients be $A = (\hat{\mu}_{ijk1})$ or $A = (\hat{\mu}_{ijk11})$ or $A = (\hat{\mu}_{ijk111})$ for $I \times J \times K \times 2$ or $I \times J \times K \times 2 \times 2$ or $I \times J \times K \times 2 \times 2 \times 2$ tables, respectively.
(a) Note that A is of order $JK \times I$ and hence square if $I = JK$ and rectangular otherwise from Remark 2.5. If A is square and non-singular, then unique MLE's of $\alpha_{i..}$, $\beta_{.j.}$ and $\gamma_{..k}$ exist.
(b) For overdetermined systems in Remark 2.5, if rank$(A) = I$ (full rank), then the left inverse of A exists and is given by $A^{-1}_{\text{left}} = (A^T A)^{-1} A^T$. Also, the unique solutions (MLE's of $\alpha_{i..}$, $\beta_{.j.}$ and $\gamma_{..k}$) are obtained using the method of ordinary least squares (see [21]).
(c) For underdetermined systems in Remark 2.5, if rank$(A) = JK$ (full rank), then the right inverse of A exists and is given by $A^{-1}_{\text{right}} = A^T(AA^T)^{-1}$. Also, the unique solutions (MLE's of $\alpha_{i..}$, $\beta_{.j.}$ and $\gamma_{..k}$) are obtained using the method of minimum norm least squares (see [15]).

3. n-dimensional incomplete table

In this section, we extend the discussions and results in the previous sections to n-dimensional incomplete tables.

3.1. Log-linear parametrization

Let $Y_1, \ldots, Y_n$ be n categorical variables with $I_1, \ldots, I_n$ levels respectively. Assume data on k of these variables are missing, while data on the remaining $(n-k)$ variables are always observed, where $1 \leq k \leq n$. For $1 \leq i \leq k$, denote R_i to be the missing indicator for Y_i, where $R_i = 1$ if data on Y_i is observed and $R_i = 2$ otherwise. Accordingly, there are a variety of incomplete tables, from the $I_1 \times \ldots \times I_n \times 2$ table (where one variable is missing) to the $I_1 \times \ldots \times I_n \times 2^n$ table (where all n variables are missing). There are $\binom{n}{k}$ ways in which data on k variables may be missing. WLOG, we assume data on $Y_1, \ldots, Y_k$ are missing. Then we have an $I_1 \times \ldots \times I_n \times 2^k$ table. The vector of observed counts is

$$\begin{aligned}\mathbf{y}_{\text{obs}} &= (\{y_{i_1\ldots i_n 1\ldots 1}\}, \{y_{i_1+\ldots+i_{k+1}\ldots i_n 12\ldots 21\ldots 1}\}, \ldots, \{y_{+\ldots+i_k i_{k+1}\ldots i_n 2\ldots 211\ldots 1}\}, \ldots,\\ &\quad \{y_{+\ldots+i_{k-1} i_k i_{k+1}\ldots i_n 2\ldots 2111\ldots 1}\}, \ldots, \{y_{i_1\ldots i_{k-1}+i_{k+1}\ldots i_n 1\ldots 121\ldots 1}\}, y_{+\ldots+2\ldots 2}).\end{aligned}$$

Note that there are a total of $\prod_{k=1}^n I_k$ fully observed counts and $(2^k - 1)$ supplementary margins. Let $\mu_{i_1\ldots i_n r_1\ldots r_k} = E(Y_{i_1\ldots i_n r_1\ldots r_k})$ denote the expected cell frequency. Then the log-linear model is given by

$$\log \mu_{i_1\ldots i_n r_1\ldots r_k} = \lambda + \sum_{p=1}^{n} \lambda_{Y_p}(i_p) + \sum_{p\neq q=1}^{n} \lambda_{Y_pY_q}(i_p, i_q) + \sum_{p=1}^{n}\sum_{q=1}^{k} \lambda_{Y_pR_q}(i_p, r_q) + \sum_{p\neq q=1}^{k} \lambda_{R_pR_q}(r_p, r_q), \tag{3.1}$$

where $1 \le i_l \le I_l,\ 1 \le l \le n,\ r_j = 1, 2,\ 1 \le j \le k.$

Three-way and higher order associations are assumed to be zero in (3.1) as they are difficult to interpret. Also, closed-form MLE's of parameters become difficult to obtain along with issues of non-identifiability. Note that association terms among Y_i's and those among R_i's are not involved in studying the missing data mechanisms of Y_i's in (3.1). Hence, there is no need to include three-way or higher order interactions among the outcome variables such as $Y_1Y_2Y_3$ or the missing indicators such as $R_1R_2R_3$. It is assumed that the MAR mechanism of a variable depends on any one of the other variables so that interaction terms like $Y_iY_jR_k$ for $i \neq j \neq k$ are excluded from (3.1). The missingness mechanism of a variable cannot be NMAR and MAR simultaneously, which excludes terms with $Y_iY_jR_i$ for $i \neq j$ in (3.1). Interactions such as $Y_iR_kR_l$ for $i \neq k \neq l$ are absent in (3.1) since their interpretation is unclear. Also, they are redundant for the derivation of closed-form estimates of the expected cell counts. The following constraints are required for identifiability of (3.1) :

$$\begin{aligned}\sum_{i_p} \lambda_{Y_p}(i_p) &= \sum_{i_p} \lambda_{Y_pY_q}(i_p, i_q) = \sum_{i_q} \lambda_{Y_pY_q}(i_p, i_q) = \sum_{i_p} \lambda_{Y_pR_q}(i_p, r_q) = \sum_{r_q} \lambda_{Y_pR_q}(i_p, r_q)\\ &= \sum_{r_p} \lambda_{R_pR_q}(r_p, r_q) = \sum_{r_q} \lambda_{R_pR_q}(r_p, r_q) = 0, \quad p \neq q.\end{aligned}$$

Next, we introduce some parameters to study the missingness mechanisms of $Y_1, \ldots, Y_k$. Let $\overline{k} = \{1, \ldots, k\}$ and $\{R_{\overline{k}\setminus\{p\}} = 1\} = \{R_i = 1 \mid i \neq p\}$. Define

$$\phi^p_{i_1\ldots i_n} = \frac{P(\{R_{\overline{k}\setminus\{p\}} = 1\}, R_p = 2 \mid Y_1 = i_1, \ldots, Y_n = i_n)}{P(\{R_{\overline{k}\setminus\{p\}} = 1\}, R_p = 1 \mid Y_1 = i_1, \ldots, Y_n = i_n)}, \quad 1 \le p \le k,$$

which is the conditional odds of Y_p being missing given the other Y_i's are observed and hence describes the missing data mechanism of Y_p. There are k such odds. Next define

$$\begin{aligned}\theta_{ij} &= \frac{P(R_i = 1, R_j = 1, \{R_{\overline{k}\setminus\{i,j\}} = 1\} | Y_1 = i_1, \ldots, Y_n = i_n)}{P(R_i = 1, R_j = 2, \{R_{\overline{k}\setminus\{i,j\}} = 1\} | Y_1 = i_1, \ldots, Y_n = i_n)}\\ &\quad \times \frac{P(R_i = 2, R_j = 2, \{R_{\overline{k}\setminus\{i,j\}} = 1\} | Y_1 = i_1, \ldots, Y_n = i_n)}{P(R_i = 2, R_j = 1, \{R_{\overline{k}\setminus\{i,j\}} = 1\} | Y_1 = i_1, \ldots, Y_n = i_n)},\end{aligned}$$

which is the conditional odds ratio between R_i and R_j. If $\theta_{ij} = 1$, then the missingness patterns of Y_i and Y_j, that is, R_i and R_j are conditionally independent given that the remaining variables are observed. There are $\binom{k}{2}$ such ratios. Let $A \subseteq \bar{k} = \{1, \ldots, k\}$ such that $|A| \ge 3$. There are $(2^k - (k+1) - \binom{k}{2})$ such sets. Let $R_A = \{R_i | i \in A\}$. Then $\{R_A = 1\} = \{R_i = 1 | i \in A\}$ and $\{R_{\bar{k}\setminus A} = 1\} = \{R_i = 1 | i \notin A\}$. Also, let

$2_A = \{r_i = 2|i \in A\}$, $1_A = \{r_i = 1|i \in A\}$, $1_{\bar{k}\backslash A} = \{r_i = 1|i \notin A\}$, $2_{\bar{k}\backslash A} = \{r_i = 2|i \notin A\}$, $Y_A = \{Y_i|i \in A\}$ and $Y_{\bar{k}\backslash A} = \{Y_i|i \notin A\}$. Now define

$$\theta_A = \frac{P(\{R_A = 2\}, \{R_{\bar{k}\backslash A} = 1\}|Y_1 = i_1, \ldots, Y_n = i_n)}{P(\{R_A = 1\}, \{R_{\bar{k}\backslash A} = 1\}|Y_1 = i_1, \ldots, Y_n = i_n)} = \frac{\pi_{i_1\ldots i_n 2_A 1_{\bar{k}\backslash A}}}{\pi_{i_1\ldots i_n 1_A 1_{\bar{k}\backslash A}}} = \frac{\mu_{i_1\ldots i_n 2_A 1_{\bar{k}\backslash A}}}{\mu_{i_1\ldots i_n 1_A 1_{\bar{k}\backslash A}}},$$

which is the conditional odds of Y_A being missing given that $Y_{\bar{k}\backslash A}$ are observed. Then for $1 \le p \le k$ and $R_p = 2, \{R_{\bar{k}\backslash\{p\}} = 1\}$, we have $\mu_{i_1\ldots i_n 1\ldots 2\ldots 1} = \phi^p_{i_1\ldots i_n}\mu_{i_1\ldots i_n 1\ldots 1}$. Also, $\mu_{i_1\ldots i_n 1\ldots 1}\phi^r_{i_1\ldots i_n}\phi^s_{i_1\ldots i_n}\theta_{rs} = \mu_{i_1\ldots i_n 2_{\{r,s\}}1_{\bar{k}\backslash\{r,s\}}}$ for $r \neq s = 1, \ldots, k$ and $\mu_{i_1\ldots i_n 1\ldots 1}\theta_A = \mu_{i_1\ldots i_n 2_A 1_{\bar{k}\backslash A}}$. Note that the joint probability

$$\pi_{i_1\ldots i_n+\ldots+} = \mu_{i_1\ldots i_n 1\ldots 1}(1 + \sum_{p=1}^{k}\phi^p_{i_1\ldots i_n} + \sum_{r\neq s=1}^{k}\phi^r_{i_1\ldots i_n}\phi^s_{i_1\ldots i_n}\theta_{rs} + \{\theta_A|A \subseteq \bar{k}, |A| \ge 3\})/N,$$

from which the marginals can be obtained. The total count N is obtained by summing both sides of the above equation over $i_1, \ldots, i_n$. Under (3.1), the parameters are given as follows.

$$\begin{aligned} \phi^t_{i_1\ldots i_n} &= \exp\left[-2\left\{\lambda_{R_t}(1) + \sum_{p=1}^{n}\lambda_{Y_pR_t}(i_p, 1) + \sum_{p\neq t=1}^{k}\lambda_{R_pR_t}(1,1)\right\}\right], \quad 1 \le t \le k, \\ \theta_{ij} &= \exp\left[4\lambda_{R_iR_j}(1,1)\right], \quad i \neq j = 1, \ldots, k, \\ \theta_A &= \exp\left[-2\left\{\sum_{p=1}^{k}\lambda_{R_p}(1) + \sum_{p=1}^{n}\sum_{q=1}^{k}\lambda_{Y_pR_q}(i_p, 1)\right\}\right], \quad A \subseteq \bar{k}, |A| \ge 3. \end{aligned}$$

The following definition (see [12]) gives the various missing data mechanisms of a variable under (3.1).

Definition 3.1. If $\phi^p_{i_1\ldots i_n}$ under (3.1) depends on i_p (denoted by $\phi^p_{\ldots i_p \ldots}$), then we have a NMAR missingness mechanism for Y_p. If it depends on i_q for $p \neq q$ (denoted by $\phi^p_{\ldots i_q \ldots}$), then the missingness mechanism for Y_p is MAR, while if it depends on none of $i_1, \ldots, i_n$ (denoted by $\phi^p_{\ldots}$), then the missingness mechanism for Y_p is MCAR.

Since there are $(n+1)$ possible realizations of $\phi^p_{i_1\ldots i_n}$ for each $p = 1, \ldots, k$, we have a total of $(n+1)^k$ possible models which may be categorized as follows:

- B1. MCAR model - the missingness mechanism of each of $Y_1, \ldots, Y_k$ is constant (1 case),
- B2. NMAR model - the missingness mechanism of each of $Y_1, \ldots, Y_k$ depends only on itself (1 case),
- B3. MAR model - the missingness mechanism of each of $Y_1, \ldots, Y_k$ depends on any one of the remaining $(n-1)$ variables ($(n-1)^k$ cases),
- B4. Mixture of MCAR and NMAR models - the missingness mechanism of each of $Y_1, \ldots, Y_k$ may be MCAR or NMAR, but all variables cannot have the same mechanism ($(2^k - 2)$ cases),
- B5. Mixture of MCAR and MAR models - the missingness mechanism of each of $Y_1, \ldots, Y_k$ may be MCAR or MAR, but all variables cannot have the same mechanism ($(n^k - (n-1)^k - 1)$ cases),
- B6. Mixture of NMAR and MAR models - the missingness mechanism of each of $Y_1, \ldots, Y_k$ may be NMAR or MAR, but all variables cannot have the same mechanism ($(n^k - (n-1)^k - 1)$ cases),
- B7. Mixture of NMAR, MAR and MCAR models - the missingness mechanism of each of $Y_1, \ldots, Y_k$ may be NMAR or MAR or MCAR, but all variables cannot have the same mechanism ($((n+1)^k + (n-1)^k - 2(n^k - 1) - 2^k)$ cases).

The log-likelihood kernel under Poisson sampling is

$$
\begin{aligned}
l(\mu;\mathbf{y_{obs}}) = &\sum_{i_1,\dots,i_n} y_{i_1\dots i_n 1\dots 1}\log\mu_{i_1\dots i_n 1\dots 1} + \sum_{i_2,\dots,i_n} y_{+i_2\dots i_n 21\dots 1}\log\mu_{+i_2\dots i_n 21\dots 1}\\
&+ \sum_{i_1,\dots,i_{k-1},i_{k+1},\dots,i_n} y_{i_1\dots i_{k-1}+i_{k+1}\dots i_n 1\dots 121\dots 1}\log\mu_{i_1\dots i_{k-1}+i_{k+1}\dots i_n 1\dots 121\dots 1} + \dots\\
&+ \sum_{i_{k+1},\dots,i_n} y_{+\dots+i_{k+1}\dots i_n 2\dots 21\dots 1}\log\mu_{+\dots+i_{k+1}\dots i_n 2\dots 21\dots 1} - \sum_{i_1,\dots,i_n,r_1,\dots,r_k}\mu_{i_1\dots i_n r_1\dots r_k}.
\end{aligned}
\tag{3.2}
$$

Rewriting (3.2) in terms of the parameters ϕ's and θ's, we can obtain closed-form MLE's of the parameters and the expected cell counts under the models described above. Perfect fits for fully observed counts are obtained for categories B2, B3 and B6 of models. From Ghosh and Vellaisamy (2016), boundary solutions occur if the MLE of any of the parameters ϕ's < 0, which are then set to zero to obtain boundary estimates. Note that for at least one $p \in \{1,\dots,k\}$, we have $\hat{\phi}^p_{\dots i_p\dots} = 0$ for at least one and at most $(I_p - 1)$ values of Y_p in case of boundary solutions.

Consider the hypotheses H_0: the proposed model (among models in categories B1 to B7 mentioned above) fits the data, and H_1: the perfect fit model fits the data. Let L_0 and L_1 denote the maximized log-likelihood functions under the proposed and perfect fit models respectively. Then the likelihood ratio statistic for testing H_0 against H_1 is

$$
\begin{aligned}
G^2 &= -2(L_0 - L_1)\\
&= -2\Bigg[\sum_{i_1,\dots,i_n}\ln\left(\frac{\hat{\mu}_{i_1\dots i_n 1\dots 1}}{y_{i_1\dots i_n 1\dots 1}}\right) + \sum_{i_2,\dots,i_n} y_{+i_2\dots i_n 21\dots 1}\ln\left(\frac{\sum_{i_1}\hat{\mu}_{i_1\dots i_n 1\dots 1}\hat{\phi}^1{}_{i_1\dots i_n}}{y_{+i_2\dots i_n 21\dots 1}}\right)\\
&+\dots+\sum_{i_1,\dots,i_{k-1},i_{k+1},\dots,i_n} y_{i_1\dots i_{k-1}+i_{k+1}\dots i_n 1\dots 121\dots 1}\ln\left(\frac{\sum_{i_k}\hat{\mu}_{i_1\dots i_n 1\dots 1}\hat{\phi}^k{}_{i_1\dots i_n}}{y_{i_1\dots i_{k-1}+i_{k+1}\dots i_n 1\dots 121\dots 1}}\right)\\
&+\dots+\sum_{i_{k+1},\dots,i_n} y_{+\dots+i_{k+1}\dots i_n 2\dots 21\dots 1}\ln\left(\frac{\hat{\mu}_{+\dots+i_{k+1}\dots i_n 1\dots 1}\hat{\theta}_{1\dots k}}{y_{+\dots+i_{k+1}\dots i_n 2\dots 21\dots 1}}\right)\\
&-\sum_{i_1,\dots i_n}\hat{\mu}_{i_1\dots i_n 1\dots 1}\left(1+\sum_{p=1}^{k}\hat{\phi}^p_{i_1\dots i_n} + \sum_{r\neq s=1}^{k}\hat{\phi}^r_{i_1\dots i_n}\hat{\phi}^s_{i_1\dots i_n}\hat{\theta}_{rs} + \{\hat{\theta}_A | A\subseteq \bar{k}, |A|\geq 3\}\right) + N\Bigg].
\end{aligned}
\tag{3.3}
$$

Note that $G^2 \sim \chi^2_\nu$ asymptotically, where $\nu = (\prod_{p=k+1}^{n} I_p)\prod_{r\neq p=1}^{k}(1+I_r)-$ number of free estimable parameters under the proposed model.

4. Data analysis

In this section, we illustrate our results in Section 3 using a real-life example from [19]. Table 4 (a $2\times2\times2\times2\times2\times2$ table) below shows the Slovenian public opinion (SPO) survey dataset classified by the variables Secession (Y_1), Attendance (Y_2) and Independence (Y_3), each having two levels Yes (1) and No (2). Here "Missing" denotes the "Don't know" category (missing margins) for each variable. Note that from (2.3), (2.6) and (3.3), we observe that G^2 becomes undefined if any of the fully observed counts is 0. So, the count 0 is replaced by 2 in the full table. WLOG, consider the subtable of Table 4 in which data on Y_1 is missing as shown below. To determine the missing data mechanism, we fit Models 1-4 (see Section 2.1) to the data in Table 5. The system of equations for Model 2 (NMAR for Y_1) yields $\hat{\alpha}_{1..} = 0.0721$ and $\hat{\alpha}_{2..} = 0.0258$ implying that boundary solutions do not occur. We use the closed-form MLE's in Section 2.1 to fit the above models. Let G^2 denote the likelihood ratio statistic for testing the goodness of fit of each of the Models 1-4 against

Table 4. Data from the SPO survey

				$R_3 = 1$		$R_3 = 2$
				$Y_3 = 1$	$Y_3 = 2$	Missing
$R_1 = 1$	$Y_1 = 1$	$R_2 = 1$	$Y_2 = 1$	1191	8	21
			$Y_2 = 2$	8	2	4
		$R_2 = 2$	Missing	107	3	9
	$Y_1 = 2$	$R_2 = 1$	$Y_2 = 1$	158	68	29
			$Y_2 = 2$	7	14	3
		$R_2 = 2$	Missing	18	43	31
$R_1 = 2$	Missing	$R_2 = 1$	$Y_2 = 1$	90	2	109
			$Y_2 = 2$	1	2	25
		$R_2 = 2$	Missing	19	8	96

Table 5. Subtable Y_1 of Table 4

			$Y_3 = 1$	$Y_3 = 2$
$R = 1$	$Y_1 = 1$	$Y_2 = 1$	1191	8
		$Y_2 = 2$	8	2
	$Y_1 = 2$	$Y_2 = 1$	158	68
		$Y_2 = 2$	7	14
$R = 2$	Missing	$Y_2 = 1$	90	2
		$Y_2 = 2$	1	2

the perfect fit model. The table below gives the G^2 values, p-values and degrees of freedom (d.f.) for the tests. We usually don't consider perfect fit models (see the example in [4])

Table 6. Comparison of fit among models

Model	**Boundary solution**	$\mathbf{G^2}$	p**-value**	**d.f.**
$\alpha_{i..}$	No	0	1	2
$\alpha_{.j.}$	No	2.4622	0.2920	2
$\alpha_{..k}$	No	2.0949	0.3508	2
$\alpha_{...}$	No	2.8538	0.4147	3

for model selection so that $\alpha_{i..}$ is discarded. From Table 6, based on p-values, the plausible models for the data in Table 5 are $\alpha_{...}$, $\alpha_{.j.}$ and $\alpha_{..k}$. However, we deduce that the best fit model is $\alpha_{..k}$ (MAR for Y_1) based on minimum G^2 value $= 2.0949$. This implies that the missingness in the variable 'Secession' depends on the observed variable 'Independence'. This dependence is expected because if one is unsure about voting for Slovenian's secession from Yugoslavia, then one is also most likely decided about Slovenian independence. Note that 'Secession' differs from 'Independence' here since independence without secession was also possible with the formation of a new internal state.
The table of expected cell counts using the closed-form estimates (see Section 2.1) is given below (Table 7).

Next, consider WLOG the subtable of Table 4 in which data on Y_1 and Y_2 are missing as shown below. To determine the missing data mechanism, we fit Models 1-16 (see Appendix) to the data in Table 8. On solving the systems of equations in NMAR models for Y_1 or Y_2, we obtain $\hat{\alpha}_{1..} = 0.0721$, $\hat{\alpha}_{2..} = 0.0258$, $\hat{\beta}_{.1.} = 0.073$ and $\hat{\beta}_{.2.} = 2.375$. Hence, there are no boundary solutions. We use the closed-form MLE's in the Appendix to fit the above models . Let G^2 denote the likelihood ratio statistic for testing the goodness

Table 7. Expected cell counts for model $\alpha_{..k}$ using closed-form estimates

			$Y_3 = 1$	$Y_3 = 2$
$R = 1$	$Y_1 = 1$	$Y_2 = 1$	1191.00	7.87
		$Y_2 = 2$	8.00	2.16
	$Y_1 = 2$	$Y_2 = 1$	158.00	66.88
		$Y_2 = 2$	7.00	15.09
$R = 2$	$Y_1 = 1$	$Y_2 = 1$	79.46	0.34
		$Y_2 = 2$	0.53	0.09
	$Y_1 = 2$	$Y_2 = 1$	10.54	2.91
		$Y_2 = 2$	0.47	0.66

Table 8. Subtable Y_1Y_2 of Table 4

				$Y_3 = 1$	$Y_3 = 2$
$R_1 = 1$	$Y_1 = 1$	$R_2 = 1$	$Y_2 = 1$	1191	8
			$Y_2 = 2$	8	2
		$R_2 = 2$	Missing	107	3
	$Y_1 = 2$	$R_2 = 1$	$Y_2 = 1$	158	68
			$Y_2 = 2$	7	14
		$R_2 = 2$	Missing	18	43
$R_1 = 2$	Missing	$R_2 = 1$	$Y_2 = 1$	90	2
			$Y_2 = 2$	1	2
		$R_2 = 2$	Missing	19	8

of fit of each of the Models 1-16 against the perfect fit model. The table below gives the G^2 values, p-values and degrees of freedom (d.f.) for the tests. From Table 9, based on

Table 9. Comparison of fit among models

Model	**Boundary solution**	$\mathbf{G^2}$	**p-value**	**d.f.**
$(\alpha_{...}, \beta_{i..})$	No	48.1188	< 0.0001	6
$(\alpha_{...}, \beta_{.j.})$	No	20.6256	0.0021	6
$(\alpha_{...}, \beta_{..k})$	No	4.5886	0.5975	6
$(\alpha_{i..}, \beta_{...})$	No	75.5003	< 0.0001	6
$(\alpha_{i..}, \beta_{i..})$	No	49.7073	< 0.0001	5
$(\alpha_{i..}, \beta_{.j.})$	No	14.7381	0.0115	5
$(\alpha_{i..}, \beta_{..k})$	No	2.8076	0.7296	5
$(\alpha_{.j.}, \beta_{...})$	No	75.1109	< 0.0001	6
$(\alpha_{.j.}, \beta_{i..})$	No	45.9217	< 0.0001	5
$(\alpha_{.j.}, \beta_{.j.})$	No	15.8222	0.0074	5
$(\alpha_{.j.}, \beta_{..k})$	No	4.2395	0.5155	5
$(\alpha_{..k}, \beta_{...})$	No	82.55	< 0.0001	6
$(\alpha_{..k}, \beta_{.i.})$	No	50.6861	< 0.0001	5
$(\alpha_{..k}, \beta_{.j.})$	No	17.8333	0.0032	5
$(\alpha_{..k}, \beta_{..k})$	No	5.4779	0.3604	5

p-values, the candidate models for the data in Table 8 are $(\alpha_{...}, \beta_{..k})$, $(\alpha_{i..}, \beta_{..k})$, $(\alpha_{.j.}, \beta_{..k})$ and $(\alpha_{..k}, \beta_{..k})$. However, we deduce that the best fit model is $(\alpha_{i..}, \beta_{..k})$ (NMAR for Y_1, MAR for Y_2) based on minimum G^2 value = 2.8076. This implies that the missingness in the variable 'Secession' depends on itself, while the missingness in 'Attendance' depends on the variable 'Independence'. This is due to the fact that if one is unsure about 'Secession', then data on 'Secession' will be missing. Also, if one is unsure about 'Independence', then one may not attend the plebiscite. Hence, data on 'Attendance' will be missing. The table of expected cell counts using the closed-form estimates (see Appendix) is given below (Table 10).

Table 10. Expected cell counts under model $(\alpha_{i..}, \beta_{..k})$ using closed-form estimates

				$Y_3 = 1$	$Y_3 = 2$
$R_1 = 1$	$Y_1 = 1$	$R_2 = 1$	$Y_2 = 1$	1191.00	8.00
			$Y_2 = 2$	8.00	2.00
		$R_2 = 2$	$Y_2 = 1$	109.15	4.00
			$Y_2 = 2$	0.73	1.00
	$Y_1 = 2$	$R_2 = 1$	$Y_2 = 1$	158.00	68.00
			$Y_2 = 2$	7.00	14.00
		$R_2 = 2$	$Y_2 = 1$	14.48	34.00
			$Y_2 = 2$	0.64	7.00
$R_1 = 2$	$Y_1 = 1$	$R_2 = 1$	$Y_2 = 1$	85.93	0.58
			$Y_2 = 2$	0.58	0.14
		$R_2 = 2$	$Y_2 = 1$	21.84	0.80
			$Y_2 = 2$	0.15	0.20
	$Y_1 = 2$	$R_2 = 1$	$Y_2 = 1$	4.07	1.75
			$Y_2 = 2$	0.18	0.36
		$R_2 = 2$	$Y_2 = 1$	1.03	2.43
			$Y_2 = 2$	0.05	0.50

Note that $\hat{\theta} = 2.7738$ for the model $(\alpha_{i..}, \beta_{..k})$, which implies that the missing mechanisms of the variables 'Secession' and 'Attendance' are probably not independent. That is, a realization is more likely to be missing for 'Secession' if it is missing for 'Attendance' or vice-versa. The estimated conditional probability of Y_1 being missing given $Y_2 = 1$ is observed is $\hat{\phi}_{1|2}(1) = \frac{\hat{\alpha}_{1..}}{1+\hat{\alpha}_{1..}} = 0.0673$. Similarly, the estimated conditional probability of Y_1 being missing given $Y_2 = 2$ is observed is $\hat{\phi}_{1|2}(2) = \frac{\hat{\alpha}_{2..}}{1+\hat{\alpha}_{2..}} = 0.0251$. So the estimated probability of nonresponse for 'Secession' is greater when one replies 'No' to attending the plebiscite. Also, the estimated conditional probability of Y_2 being missing given $Y_1 = 1$ is observed is $\hat{\phi}_{2|1}(1) = \frac{\hat{\beta}_{..1}}{1+\hat{\beta}_{..1}} = 0.0839$. Similarly, the estimated conditional probability of Y_2 being missing given $Y_1 = 2$ is observed is $\hat{\phi}_{2|1}(2) = \frac{\hat{\beta}_{..2}}{1+\hat{\beta}_{..2}} = 0.3333$. Hence, the estimated probability of nonresponse for 'Attendance' is greater when one replies 'No' to Slovenia's secession from Yugoslavia.

From the data in Table 8, we have $OR_{..1} = OR_{111} = 6.5957$ and $OR_{..2} = OR_{112} = 0.8235$ for the model $(\alpha_{i..}, \beta_{..k})$. This implies that if none of the responses for the variables is missing, then the estimated odds ratio between 'Secession' and 'Attendance' is greater when the response to 'Independence' is 'Yes' than when it is 'No'. Also, $Var(OR_{..1}) = 11.9646$ and $Var(OR_{..2}) = 0.4823$, that is, for observed data, the estimated odds ratio between 'Secession' and 'Attendance' has greater precision when the response to 'Independence' is 'No' than when it is 'Yes'.

To investigate the occurrence of boundary solutions, we consider subtables of Table 4 in which at least one of Y_1, Y_2 and Y_3 is missing. When we fit perfect fit NMAR models (for fully observed counts) to the data in each subtable, we observe that boundary solutions do not occur in any of them as the MLE's are $\hat{\alpha}_{1..} = 0.0721$, $\hat{\alpha}_{2..} = 0.0258$, $\hat{\beta}_{.1.} = 0.073$, $\hat{\beta}_{.2.} = 2.375$, $\hat{\gamma}_{..1} = 0.0151$ and $\hat{\gamma}_{..2} = 0.3851$, which are positive. So, we modify some fully observed counts in each subtable. Table 11 shows the MLE's under some perfect fit (for fully observed counts) NMAR models in the modified subtables.

Table 11. MLE's in modified subtables of Table 4

Subtable	Changes	NMAR (models) (perfect fit)	MLE's	Boundary solns.
Y_1	$158 \to 1300, 68 \to 28,$ $7 \to 10$	Y_1	$\hat{\alpha}_{1..} = -0.0293, \hat{\alpha}_{2..} = 0.0961$	$\hat{\pi}_{1++2} = 0$
Y_2Y_3	$8 \to 80, 14 \to 10$	Y_2	$\hat{\beta}_{.1.} = -0.1937, \hat{\beta}_{.2.} = 4.2616$	$\hat{\pi}_{+1+2+} = 0$
	$8 \to 80, 14 \to 6$	Y_3	$\hat{\gamma}_{..1} = -0.0132, \hat{\gamma}_{..2} = 0.4588$	$\hat{\pi}_{++1+2} = 0$
	$8 \to 108, 8 \to 108,$ $2 \to 4, 14 \to 2$	Y_2, Y_3	$\hat{\beta}_{.1.} = 0.2138, \hat{\beta}_{.2.} = -1.3706$ $\hat{\gamma}_{..1} = -0.0253, \hat{\gamma}_{..2} = 0.4785$	$\hat{\pi}_{+1+2+} = 0,$ $\hat{\pi}_{++1+2} = 0$
$Y_1Y_2Y_3$	$158 \to 1100, 68 \to 22,$ $7 \to 10$	Y_1	$\hat{\alpha}_{1..} = -0.0346, \hat{\alpha}_{2..} = 0.1193$	$\hat{\pi}_{1++2++} = 0$
	$8 \to 80$	Y_2	$\hat{\beta}_{.1.} = -0.1258, \hat{\beta}_{.2.} = 3.2391$	$\hat{\pi}_{+1++2+} = 0$
	$8 \to 55, 14 \to 6$	Y_3	$\hat{\gamma}_{..1} = -0.0024, \hat{\gamma}_{..2} = 0.4338$	$\hat{\pi}_{++1++2} = 0$
	$1191 \to 3191, 8 \to 48,$ $8 \to 28, 2 \to 4$	Y_1, Y_3	$\hat{\alpha}_{1..} = -0.0291, \hat{\alpha}_{2..} = 0.1828$ $\hat{\gamma}_{..1} = -0.0397, \hat{\gamma}_{..2} = 3.1164$	$\hat{\pi}_{+1++2+} = 0,$ $\hat{\pi}_{++1++2} = 0$

From Table 11, we observe that on fitting perfect fit NMAR models to the modified subtables, boundary solutions occur in each of them since at least one of $\hat{\alpha}_{i..}$, $\hat{\beta}_{.j.}$ and $\hat{\gamma}_{..k}$ is negative. In the last column of Table 11, the boundary solutions under the above models are obtained using the EM algorithm (see the 'ecm.cat' function of the 'cat' package in R software). The forms of boundary solutions under the various models are the same as those mentioned in Section 2. For further discussion on boundary solutions in two and higher dimensional incomplete tables, one could refer to [11] and [12]. The packages 'MASS' and 'cat' in R software are used to perform the data analysis in this paper.

5. Conclusions

In this paper, we have studied missing data mechanisms for variables in $I \times J \times K \times 2$, $I \times J \times K \times 2 \times 2$ and $I \times J \times K \times 2 \times 2 \times 2$ incomplete contingency tables. For this purpose, we have considered hierarchical log-linear models which yield closed-form MLE's of parameters and expected cell counts under various missing data models. Closed-form estimates are also obtained for joint and marginal probabilities, marginal odds ratios, their asymptotic variances and conditional probabilities of missing variables under the models. Note that the methods and results in this paper are applicable for $I \times J \times 2$ and $I \times J \times 2 \times 2$ tables also. Extensions of the models and estimation methods are presented for arbitrary n-dimensional incomplete tables. We have also provided closed-form boundary MLE's under various NMAR models in some incomplete tables. Finally, a real-life data analysis example validates our modelling approach and other results in this paper.

Acknowledgment. The authors express their gratitude to the referees and the associate editor for their valuable comments which significantly improved the quality of the paper.

References

[1] S.G. Baker, *The multinomial-Poisson transformation*, J. R. Stat. Soc. Ser. D. **43**, 495-504, 1994.

[2] S.G. Baker, *A simple method for computing the observed information matrix when using the EM algorithm with categorical data*, J. Comput. Graph. Stat. **1** (1), 63-76, 1992.

[3] S.G. Baker and N.M. Laird, *Regression analysis for categorical variables with outcome subject to nonignorable nonresponse*, J. Am. Stat. Assoc. **83**, 62-69, 1988.

[4] S.G. Baker, W.F. Rosenberger and R. Dersimonian, *Closed-form estimates for missing counts in two-way contingency tables*, Stat. Med. **11**, 643-657, 1992.

[5] P.S. Clarke, *On boundary solutions and identifiability in categorical regression with non-ignorable non-response*, Biom. J. **44**, 701-717, 2002.

[6] P.S. Clarke and P.W.F. Smith, *On maximum likelihood estimation for log-linear models with non-ignorable non-responses*, Stat. Probab. Lett. **73**, 441-448, 2005.

[7] P.S. Clarke and P.W.F. Smith, *Interval estimation for log-linear models with one variable subject to non-ignorable non-response*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **66**, 357-368, 2004.

[8] A.P. Dempster, N.M. Laird and D.B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **39** (1), 1-38, 1977.

[9] R.E. Fay, *Causal models for patterns of nonresponse*, J. Am. Stat. Assoc. **81**, 354-365, 1986.

[10] J.J. Forster and P.W.F. Smith, *Model-based inference for categorical survey data subject to nonignorable nonresponse*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **60**, 57-70, 1998.

[11] S. Ghosh and P. Vellaisamy, *On the occurrence of boundary solutions in two-way incomplete tables*, REVSTAT **48**, 89-108, 2020.

[12] S. Ghosh and P. Vellaisamy, *On the occurrence of boundary solutions in multidimensional incomplete tables*, Stat. Probab. Lett. **119**, 63-75, 2016.

[13] R.J.A. Little, *Pattern-mixture models for multivariate incomplete data*, J. Am. Stat. Assoc. **88**, 125-134, 1993.

[14] R.J.A. Little and D.B. Rubin, *Statistical Analysis with Missing Data.* 2nd ed. New York: Wiley, 2002.

[15] W.R. Madych, *Solutions of underdetermined systems of linear equations*, Spatial statistics and imaging, Institute of Mathematical Statistics, Hayward, CA, 227-238, 1991.

[16] X.L. Meng and D.B. Rubin, *Using EM to obtain asymptotic variancecovariance matrices: the SEM algorithm*, J. Am. Stat. Assoc. **86**, 899-909, 1991.

[17] T. Park and M.B. Brown, *Models for categorical data with nonignorable nonresponse*, J. Am. Stat. Assoc. **89**, 44-52, 1994.

[18] Y.D. Park, D. Kim and S. Kim, *Identification of the occurrence of boundary solutions in a contingency table with nonignorable nonresponse*, Stat. Probab. Lett. **93**, 34-40, 2014.

[19] D.B. Rubin, H.S. Stern and V. Vehovar, *Handling “Don’t know" survey responses : the case of the Slovenian plebiscite*, J. Am. Stat. Assoc. **90**, 822-828, 1995.

[20] P.W.F. Smith, C.J. Skinner and P.S. Clarke, *Allowing for non-ignorable nonresponse in the analysis of voting intention data*, J. R. Stat. Soc. Ser. C. Appl. Stat. **48**, 563-577, 1999.

[21] G. Williams, *Overdetermined systems of linear equations*, Am. Math. Mon. **97**, 511-513, 1990.

Appendix A.

The closed-form estimates of missing counts and other parameters under various missing data models for an $I \times J \times K \times 2 \times 2$ table are as follows.
1. $(\alpha_{...}, \beta_{...})$ (MCAR for both Y_1 and Y_2).
The MLE's are

$$\hat{\alpha}_{...} = \frac{y_{+++21}}{y_{+++11}},\ \hat{\beta}_{...} = \frac{y_{+++12}}{y_{+++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},$$

while the iterates of $\hat{\mu}_{ijk11}$ are

$$\hat{\mu}^{(0)}_{ijk11} = y_{ijk11},\ \hat{\mu}^{(t+1)}_{ijk11} = \frac{y_{+++11}\left(y_{ijk11} + \frac{y_{i+k12}}{\hat{\mu}^{(t)}_{i+k12}}.\hat{\mu}^{(t)}_{ijk11} + \frac{y_{+jk21}}{\hat{\mu}^{(t)}_{+jk21}}.\hat{\mu}^{(t)}_{ijk11}\right)}{y_{++++1} + y_{+++12}}.$$

2. $(\alpha_{...}, \beta_{i..})$ (MCAR for Y_1, MAR for Y_2).
The MLE's are

$$\hat{\alpha}_{...} = \frac{y_{+++21}}{y_{+++11}},\ \hat{\beta}_{i..} = \frac{y_{i++12}}{\hat{\mu}_{i++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},\ \hat{\mu}_{ijk11} = \frac{y_{ijk11}y_{+++11}y_{+jk+1}}{y_{++++1}y_{+jk11}}.$$

3. $(\alpha_{...}, \beta_{.j.})$ (MCAR for Y_1, NMAR for Y_2).
The MLE's are

$$\hat{\alpha}_{...} = \frac{y_{+++21}}{y_{+++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},\ \hat{\mu}_{ijk11} = \frac{y_{ijk11}y_{+++11}y_{+jk+1}}{y_{++++1}y_{+jk11}}.$$

Also, $\hat{\beta}_{.j.}$ satisfies $\sum_j \hat{\mu}_{ijk11}\hat{\beta}_{.j.} = y_{i+k12}$.
4. $(\alpha_{...}, \beta_{..k})$ (MCAR for Y_1, MAR for Y_2).
The MLE's are

$$\hat{\alpha}_{...} = \frac{y_{+++21}}{y_{+++11}},\ \hat{\beta}_{..k} = \frac{y_{++k12}}{\hat{\mu}_{++k11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},\ \hat{\mu}_{ijk11} = \frac{y_{ijk11}y_{+++11}y_{+jk+1}}{y_{++++1}y_{+jk11}}.$$

5. $(\alpha_{i..}, \beta_{...})$ (NMAR for Y_1, MCAR for Y_2).
The MLE's are

$$\hat{\beta}_{...} = \frac{y_{+++12}}{y_{+++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},\ \hat{\mu}_{ijk11} = \frac{y_{ijk11}y_{+++11}y_{i+k1+}}{y_{+++1+}y_{i+k11}}.$$

Also, $\hat{\alpha}_{i..}$ satisfies $\sum_i \hat{\mu}_{ijk11}\hat{\alpha}_{i..} = y_{+jk21}$.
6. $(\alpha_{i..}, \beta_{i..})$ (NMAR for Y_1, MAR for Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\beta}_{i..} = \frac{y_{i++12}}{y_{i++11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_i y_{i++12}\hat{\alpha}_{i..}},$$

where $\hat{\alpha}_{i..}$ satisfies $\sum_i \hat{\mu}_{ijk11}\hat{\alpha}_{i..} = y_{+jk21}$.
7. $(\alpha_{i..}, \beta_{.j.})$ (NMAR for both Y_1 and Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{i,j} y_{ij+11}\hat{\alpha}_{i..}\hat{\beta}_{.j.}},$$

where $\hat{\alpha}_{i..}$ and $\hat{\beta}_{.j.}$ satisfy $\sum_i \hat{\mu}_{ijk11}\hat{\alpha}_{i..} = y_{+jk21}$ and $\sum_j \hat{\mu}_{ijk11}\hat{\beta}_{.j.} = y_{i+k12}$ respectively.
8. $(\alpha_{i..}, \beta_{..k})$ (NMAR for Y_1, MAR for Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\beta}_{..k} = \frac{y_{++k12}}{y_{++k11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{i,k} y_{i+k11}\hat{\alpha}_{i..}\hat{\beta}_{..k}},$$

where $\hat{\alpha}_{i..}$ satisfies $\sum_i \hat{\mu}_{ijk11}\hat{\alpha}_{i..} = y_{+jk21}$.
9. $(\alpha_{.j.}, \beta_{...})$ (MAR for Y_1, MCAR for Y_2).
The MLE's are

$$\hat{\alpha}_{.j.} = \frac{y_{+j+21}}{\hat{\mu}_{+j+11}},\ \hat{\beta}_{...} = \frac{y_{+++12}}{y_{+++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},\ \hat{\mu}_{ijk11} = \frac{y_{ijk11}y_{+++11}y_{i+k1+}}{y_{+++1+}y_{i+k11}}.$$

10. $(\alpha_{.j.}, \beta_{i..})$ (MAR for both Y_1 and Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\alpha}_{.j.} = \frac{y_{+j+21}}{y_{+j+11}},\ \hat{\beta}_{i..} = \frac{y_{i++12}}{y_{i++11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{i,j} y_{ij+11}\hat{\alpha}_{.j.}\hat{\beta}_{i..}}.$$

11. $(\alpha_{.j.}, \beta_{.j.})$ (MAR for Y_1, NMAR for Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\alpha}_{.j.} = \frac{y_{+j+21}}{y_{+j+11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{j} y_{+j+21}\hat{\beta}_{.j.}},$$

where $\hat{\beta}_{.j.}$ satisfies $\sum_j \hat{\mu}_{ijk11}\hat{\beta}_{.j.} = y_{i+k12}$.
12. $(\alpha_{.j.}, \beta_{..k})$ (MAR for both Y_1 and Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\alpha}_{.j.} = \frac{y_{+j+21}}{y_{+j+11}},\ \hat{\beta}_{..k} = \frac{y_{++k12}}{y_{++k11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{j,k} y_{+jk11}\hat{\alpha}_{.j.}\hat{\beta}_{..k}}.$$

13. $(\alpha_{..k}, \beta_{...})$ (MAR for Y_1, MCAR for Y_2).
The MLE's are

$$\hat{\alpha}_{..k} = \frac{y_{++k21}}{\hat{\mu}_{++k11}},\ \hat{\beta}_{...} = \frac{y_{+++12}}{y_{+++11}},\ \hat{\theta} = \frac{y_{+++11}y_{+++22}}{y_{+++12}y_{+++21}},\ \hat{\mu}_{ijk11} = \frac{y_{ijk11}y_{+++11}y_{i+k1+}}{y_{+++1+}y_{i+k11}}.$$

14. $(\alpha_{..k}, \beta_{i..})$ (MAR for both Y_1 and Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\alpha}_{..k} = \frac{y_{++k21}}{y_{++k11}},\ \hat{\beta}_{i..} = \frac{y_{i++12}}{y_{i++11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{i,k} y_{i+k11}\hat{\alpha}_{..k}\hat{\beta}_{i..}}.$$

15. $(\alpha_{..k}, \beta_{.j.})$ (MAR for Y_1, NMAR for Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\alpha}_{..k} = \frac{y_{++k21}}{y_{++k11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{j,k} y_{+jk11}\hat{\alpha}_{..k}\hat{\beta}_{.j.}},$$

where $\hat{\beta}_{.j.}$ satisfies $\sum_j \hat{\mu}_{ijk11}\hat{\beta}_{.j.} = y_{i+k12}$.
16. $(\alpha_{..k}, \beta_{..k})$ (MAR for both Y_1 and Y_2).
The MLE's are

$$\hat{\mu}_{ijk11} = y_{ijk11},\ \hat{\alpha}_{..k} = \frac{y_{++k21}}{y_{++k11}},\ \hat{\beta}_{..k} = \frac{y_{++k12}}{y_{++k11}},\ \hat{\theta} = \frac{y_{+++22}}{\sum_{k} y_{++k12}\hat{\alpha}_{..k}}.$$

Note that closed-form MLE's of m_{jk11} exist for all models except for Model 1. In this case, $\hat{\mu}_{ijk11}$ may be obtained using the EM algorithm (see [8]).