AJPS Research Article

# Concurrent Generation of Binary, Ordinal, and Count Data with Specified Marginal and Associational Quantities in Pharmaceutical Sciences

Hakan DEMIRTAS[1*], Kubra COSAR[2], Mutlu ALTUNTAS[3],

[1*] Division of Epidemiology and Biostatistics (MC923), University of Illinois at Chicago, Chicago, USA

[2] Department of Econometrics, Ankara Haci Bayram Veli University, Ankara, TURKEY

[3] Department of Statistics, Sinop University, Sinop, TURKEY

**ABSTRACT:** This manuscript is concerned with establishing a unified framework for concurrently generating data sets that include three major kinds of variables (i.e., binary, ordinal, and count) when the marginal distributions and a feasible association structure are specified for simulation purposes. The simulation paradigm has been commonly utilized in pharmaceutical practice. A central aspect of every simulation study is the quantification of the model components and parameters that jointly define a scientific process. When this quantification goes beyond the deterministic tools, researchers often resort to random number generation (RNG) in finding simulation-based solutions to address the stochastic nature of the problem. Although many RNG algorithms have appeared in the literature, a major limitation is that most of them were not devised to simultaneously accommodate all variable types mentioned above. Thus, these algorithms provide only an incomplete solution, as real data sets include variables of different kinds. This work represents an important augmentation of the existing methods as it is a systematic attempt and comprehensive investigation for mixed data generation. We provide an algorithm that is designed for generating data of mixed marginals; illustrate its operational, logistical, and computational details; and present ideas on how it can be extended to span more sophisticated distributional settings in terms of a broader range of marginal features and associational quantities.

**Key Words:** Biserial correlation, phi coefficient, simulation, tetrachoric correlation, random number generation, mixed data

## 1 INTRODUCTION

Stochastic simulation is an indispensable part and major focus of scientific inquiry. Model building, estimation, and testing typically require verification via simulation to assess the reliability, validity, and plausibility of inferential techniques, to evaluate how well the implemented models capture the specified true population values, and how reasonably these models respond to departures from underlying assumptions, among other things. Describing a real notion by creating mirror images and imperfect proxies of the perceived underlying truth, iteratively refining and occasionally redefining the empirical truth to decipher the

*Corresponding Author: Hakan DEMİRTAŞ
E-mail: demirtas@uic.edu

mechanism by which the process under consideration is assumed to operate in a repeated manner allows researchers to study the performance of their methods through simulated data replicates that mimic the real data characteristics of interest in any given setting. Accuracy and precision measures for the model parameters signal if the procedure works properly; and may suggest remedial action to minimize the discrepancies between expectation and reality.

Simulation studies have been commonly employed in a broad range of disciplines in order to better comprehend and solve today's increasingly intricate issues. A core component of every simulation study is the quantification of the model components and parameters that jointly define a scientific phenomenon. Deterministic tools are typically inadequate to quantify complex situations, leading researchers to harness RNG techniques in finding simulation-based solutions to address the stochastic behavior of the problems that generally involve variables of many different types on a structural level; i.e., causal and correlational interdependencies are a function of a mixture of binary, ordinal, and count variables, which jointly act to characterize the mechanisms that collectively delineate a paradigm. In modern times, we are unequivocally moving from small data to big data, from mechanistical to empirical thinking, from exact solutions to simulation-driven solutions, from mathematical perfection to reasonable approximation to reality; the ideas presented herein are consequential in the sense that the basic mixed-data generation setup can be augmented to handle a large spectrum of situations that can be encountered in many areas.

This work is concerned with building the basics of a unified skeleton for concurrently generating data sets that include three major kinds of variables (i.e., binary, ordinal, and count) when the marginal distributions and a feasible association structure in the form of Pearson correlations are specified for simulation purposes. Although many RNG algorithms have appeared in the literature, a fundamental restriction is that they were not designed for a mix of all prominent types of data. The current paper is a systematic attempt and compendious investigation for mixed data generation; it represents a substantial augmentation of the existing methods, and it has potential to advance scientific research and knowledge in a meaningful way. The broader impact of this framework is that it can assist data analysts, practitioners, theoreticians, and methodologists across many disciplines to simulate mixed data with relative ease. The proposed algorithm constitutes a comprehensive set of computational tools that offers promising potential for building enhanced computing infrastructure for research and education.

We propose an RNG algorithm that encompasses three major variable types, building upon our previous work in generation of multivariate ordinal data [1], joint generation of binary and normal data [2], ordinal and normal data [3], count and normal data [4] with the specification of marginal and associational parameters along with other related work [5-10]. Equally importantly, we discuss the extensions on how to incorporate continuous data to the mix via power polynomials that can handle the overwhelming majority of continuous shapes [11-15], count data that are amenable to over- and under-dispersion via generalized Poisson distribution [16], broader measures of associations such as Spearman's rank correlation, and the specification of higher order product moments. Conceptual, algorithmic, operational, and procedural details will be conveyed throughout the paper.

The organization of the manuscript is as follows: In Sect. 2, the algorithm for simultaneous generation of binary, ordinal, and count data is given. The essence of the algorithm is finding the correlation structure of underlying multivariate normal (MVN) data that form a basis for the subsequent discretization in the binary and ordinal cases, and correlation mapping using inverse cumulative distribution functions (cdf's) in the count data case, where modeling the correlation transitions for different distributional pairs is discussed in detail. Sect. 3 presents some logistical details and illustrative examples through an R package that implements the algorithm, demonstrating how well the proposed technique works. Sect. 4 includes discussion on extensions, limitations, future directions, and concluding remarks.

## 2 ALGORITHM

The algorithm is designed for concurrently generating binary, ordinal, and count data, with the added utility that normal variables can potentially be incorporated to the system; for generality we present a version that includes normal components. The count part is assumed to follow the Poisson distribution. While binary is a special case of ordinal, for the purpose of exposition, the steps are presented separately. Skipped patterns are allowed for ordinal variables. The marginal characteristics (the proportions for the binary and ordinal part, the rate parameters for the count part, and the means and variances for the normal part) and a feasible Pearson correlation matrix need to be specified by the users. The algorithmic skeleton establishes the basic foundation, extensions to more general and complicated situations will be discussed in Sect. 4.

The operational engine of the algorithm hinges upon computing the correlation matrix of underlying MVN data that serve as an intermediate tool in the sense that binary and ordinal variables are obtained via dichotomization and ordinalization, respectively, through the threshold concept, and count variables are retrieved by correlation mapping using inverse cdf matching. The procedure entails modeling the correlation transformations that result from discretization and mapping.

In what follows, let B, O, C, and N denote binary, ordinal, count, and normal variables, respectively. Let  be the specified Pearson correlation matrix which is comprised of ten submatrices that correspond to all possible variable-type combinations.

Required parameter values are p's for binary and ordinal variables, λ's for count variables, $\left(\mu,\sigma^2\right)$ pairs for normal variables, and the entries of the correlation matrix. These quantities are either specified or estimated from a real data set that is to be mimicked.

1. Check if  is positive definite.

2. Find the upper and lower correlation bounds for all pairs by the sorting method of [6]. It is well-known that correlations are not bounded between -1 and +1 in most bivariate settings as different upper and/or lower bounds may be imposed by the marginal distributions. These restrictions apply to discrete variables as well as continuous ones. Let  (F, G) be the set of cdf's *H* on $R^2$ having marginal cdf's F and G. It can be proven that in  (F, G), there exist cdf's $H_L$ and $H_U$, called the lower and upper bounds, having minimum and maximum correlation. For all (x, y)  $R^2$, $H_L$ (x, y) = max[F(x) + G(y) - 1,0] and $H_U$ (x, y) = min[F(x), G(y)]. For any H  (F, G) and all (x, y)  $R^2$, HL (x, y)  H (x, y)  < $H_U$ (x, y). If $\delta_L$, $\delta_U$, and $\delta$ denote the Pearson correlation coefficients for $H_L$, $H_U$, and H, respectively, then $\delta_L \le \delta \le \delta_U$. One can infer that if V is uniform in [0,1], then F-1 (V) and G-1(V) are maximally correlated; and $F^{-1}$ (V) and $G^{-1}$(V) are maximally anticorrelated. In practical terms, generating X and Y independently with a large number of data points before sorting them in the same or opposite direction give the approximate upper and lower correlation bounds, respectively. Make sure all elements of Σ are within the plausible range.

3. Perform logical checks such as binary proportions are between 0 and 1, probabilities add up to 1 for ordinal variables, the Poisson rates are positive for count variables, variances

for normal variables are positive, the mean, variance, proportion and rate vectors are consistent with the number of variables, $\Sigma$ is symmetric and its diagonal entries are 1, to prevent obvious misspecification errors.

4. For B-B combinations, find the tetrachoric (pre-dichotomization) correlation given the specified phi coefficient (post-dichotomization correlation). Let $X_1$, $X_2$ represent binary variables such that $E[Xj] = pj$ and $Cor(X_1, X_2) = \delta_{12}$, where $pj$ ($j$ = 1, 2) and $\delta_{12}$ (phi coefficient) are given.

Let $\Phi[t_1, t_2, \rho_{12}]$ be the cdf for a standard bivariate normal random variable with correlation coefficient $\rho_{12}$ (tetrachoric correlation).Naturally,

$$\Phi[t_1, t_2, \rho_{12}] = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} f(z_1, z_2, \rho_{12})\, dz_1 dz_2 ,$$

where

$$f(z_1, z_2, \rho_{12}) = \left[2\pi\left(1-\rho_{12}^{\;2}\right)^{1/2}\right]^{-1} \times exp\left[\frac{-\left(z_1^{\;2} - 2\rho_{12} z_1 z_2 + z_2^{\;2}\right)}{\left(2\left(1-\rho_{12}^{\;2}\right)\right)}\right]$$

The connection between $\delta_{12}$ and $\rho_{12}$ is reflected in the equation

$$\Phi\left[z(p_1), z(p_2), \rho_{12}\right] = \delta_{12}\left(p_1 q_1 p_2 q_2\right)^{1/2} + p_1 p_2$$

Solve for $\rho_{12}$ where $z(p_j)$ denotes the $p_j^{\;th}$ quantile of the standard normal distribution, and $q_j = 1 - p_j$. Repeat this process for all B-B pairs.

5. For B-O and O-O combinations, implement an iterative procedure that finds the polychoric (pre-discretization) correlation given the ordinal phi coefficient (post-discretization correlation). Suppose **Z** = $(Z_1, Z_2)$ ~ $N\left(0, \Delta_{Z_1Z_2}\right)$, where **Z** denotes the bivariate standard normal distribution with correlation matrix $\Delta_{Z_1Z_2}$ whose off-diagonal entry is $\delta_{Z_1Z_2}$. Let **X** = $(X_1,X_2)$ be the bivariate ordinal data where underlying **Z** is discretized based on corresponding normal quantiles given the marginal proportions, with a correlation matrix $\Delta_{Z_1Z_2}$. If we need to sample from a random vector $(X_1, X_2)$ whose marginal cdf's are $F_1$, $F_2$ tied together via a Gaussian copula, we generate a sample $(z_1, z_2)$ from **Z** ~ N $\left(0, \Delta_{Z_1Z_2}\right)$, then set $\mathbf{x} = (x_1, x_2) = \left(F_1^{-1}(u_1), F_2^{-1}(u_2)\right)$ when $\mathbf{u} = (u_1, u_2) = \left(\Phi(z_1), \Phi(z_2)\right)$, where $\Phi$ is the cdf of the standard normal distribution. The correlation matrix of **X**, denoted by $\Delta_{Z_1Z_2}$ (with an off-diagonal entry $\delta_{X_1X_2}$) obviously differs from $\Delta_{Z_1Z_2}$ due to discretization. More specifically, $\left|\delta_{X_1X_2}\right| < \left|\delta_{Z_1Z_2}\right|$ in large samples. The relationship between $\delta_{X_1X_2}$ and $\delta_{Z_1Z_2}$ can be established via the following algorithm [9]:

a. Generate standard bivariate normal data with the correlation $\delta^0_{Z_1Z_2}$ where $\delta^0_{Z_1Z_2} = \delta_{X_1X_2}$ (Here, $\delta^0_{Z_1Z_2}$ is the initial polychoric correlation).

b. Discretize $Z_1$ and $Z_2$, based on the cumulative probabilities of the marginal distribution $F_1$ and $F_2$, to obtain $X_1$ and $X_2$, respectively.

c. Compute $\delta^1_{X_1X_2}$ through $X_1$ and $X_2$ (Here, $\delta^1_{X_1X_2}$ is the ordinal phi coefficient after the first iteration).

d. Execute the following loop as long as $\left|\delta^v_{X_1X_2} - \delta_{X_1X_2}\right| > \epsilon$ and $1 \le v \le v_{max}$ ($v_{max}$ and $\epsilon$ are the maximum number of iterations and the maximum tolerated absolute error, respectively, both quantities are set by the users):

i. Update $\delta^v_{Z_1Z_2}$ by $\delta^v_{Z_1Z_2} = \delta^{v-1}_{Z_1Z_2} g(v)$, where $g(v) = \delta_{X_1X_2} / \delta^v_{X_1X_2}$. Here, $g(v)$ serves as a correction coefficient, which ultimately converges to 1.

ii. Generate bivariate normal data with $\delta^v_{Z_1Z_2}$ and compute $\delta^{v+1}_{X_1X_2}$ after discretization.

Again, one should repeat this process for each B-O and O-O pair.

6. For C-C combinations, compute the corresponding normal-normal correlations (pre-mapping) given the specified count-count correlations (post-mapping) via the inverse cdf method in Yahav and Shmueli that was proposed in the context of correlated count data generation [10]. Their method utilizes a slightly modified version of the NORTA (Normal to Anything) approach [17], which involves generation of MVN variates with given univariate marginals and the correlation structure (*RN*), and then transforming it into any desired distribution using the inverse cdf. In the Poisson case, NORTA can be implemented by the following steps:

a. Generate a *k*-dimensional normal vector $\boldsymbol{Z}_N$ from *MVN* with mean vector **0** and a correlation matrix $R_N$.

b. Transform $\boldsymbol{Z}_N$ to a Poisson vector $\boldsymbol{X}_C$ as follows:

i. For each element $z_i$ of $\boldsymbol{Z}_N$, calculate the Normal cdf, $\Phi(z_i)$.

ii. For each value of $\Phi(z_i)$, calculate the Poisson inverse cdf with a desired corresponding marginal rate $\lambda_i$, $\Psi^{-1}_{\lambda_i}(\Phi(z_i))$; where

$$\Psi_{\lambda_i}(x) = \sum_{i=0}^{x} \frac{e^{-\lambda}\lambda^i}{i!}$$

c. $\boldsymbol{X}_C = \left[\Psi^{-1}_{\lambda_i}(\Phi(z_i)), \ldots, \Psi^{-1}_{\lambda_k}(\Phi(z_k))\right]^T$ is a draw from the desired multivariate count data with correlation matrix *RPOIS*.

An exact theo-retical connection

between $R_N$ and $R_{POIS}$ has not been established to date. However, it has been shown that a feasible range of correlation between a pair of Poisson variables after the inverse cdf transformation is within

$$\underline{\rho} = \left[ \begin{array}{l} Cor\left(\Psi_{\lambda_i}^{-1}(U), \Psi_{\lambda_j}^{-1}(1-U)\right), \\ \bar{\rho} = Cor\left(\Psi_{\lambda_i}^{-1}(U), \Psi_{\lambda_j}^{-1}(U)\right) \end{array} \right],$$

where $\lambda_i$ and $\lambda_j$ are the marginal rates, and $U \sim Uniform\,(0,1)$ Yahav and Shmueli proposed a conceptually simple method to approximate the relationship between the two correlations [10]. They have demonstrated that *RPOIS* can be approximated as an exponential function of $R_N$ where the coefficients are the functions of $\underline{\rho}$ and $\bar{\rho}$.

7. For B-N/O-N combinations, find the biserial/polyserial correlation (before discretization of one of the variables) given the point-biserial/point-polyserial correlation (after discretization) by the linearity and constancy arguments proposed by [7]. Suppose that $X$ and $Y$ follow a bivariate normal distribution with a correlation of $\delta_{XY}$. Without loss of generality, we may assume that both $X$ and $Y$ are standardized to have a mean of 0 and a variance of 1. Let $X_D$ be the binary variable resulting from a split on $X$, $X_D = I(X \geq k)$. Thus, $E[X_D] = p$ and $V[X_D] = pq$ where $q = 1 - p$. The correlation between $X_D$ and $X$, $\delta_{X_D X}$ can be obtained in a simple way, namely,

$$\begin{aligned} \delta_{X_D X} &= \frac{Cov[X_D, X]}{\sqrt{V[X_D]V[X]}} \\ &= E[X_D X]/\sqrt{pq} \\ &= E[X \mid X \geq k]/\sqrt{pq} \end{aligned}.$$

We can also express the relationship between $X$ and $Y$ via the following linear regression model:

$$Y = \delta_{XY} X + \epsilon \quad (1)$$

where $\epsilon$ is independent of $X$ and $Y$, and follows $N \sim (0, 1 - \delta_{XY}^2$. When we generalize this to nonnormal $X$ and/or $Y$ (both centered and scaled), the same relationship can be assumed to hold with the exception that the distribution of $\epsilon$ follows a nonnormal distribution. As long as Eq. 1 is valid,

$$\begin{aligned} Cov[X_D, Y] &= Cov[X_D, \delta_{XY} X + \epsilon] \\ &= Cov[X_D, \delta_{XY} X] + Cov[X_D, \epsilon] \\ &= \delta_{XY} Cov[X_D, X] + Cov[X_D, \epsilon]. \quad (2) \end{aligned}$$

Since $\epsilon$ is independent of $X$, it will also be independent of any deterministic function of $X$ such as $X_D$, and thus $Cov[X_D, \epsilon]$ will be 0. As $E[X] = E[Y] = 0$, $V[X] = V[Y] = 1$, $Cov[X_D, Y] = \delta_{X_D Y}\sqrt{pq}$ and $Cov[X, Y] = \delta_{XY}$, Eq. 2 reduces to

$$\delta_{X_D Y} = \delta_{XY}\delta_{X_D X}. \quad (3)$$

In the bivariate normal case, $\delta_{X_D X} = h/\sqrt{pq}$ where $h$ is the ordinate of the normal curve at the point of dichotomization.

Eq. 3 indicates that the linear association between $X_D$ and $Y$ is assumed to be fully explained by their mutual association with $X$ [7]. The ratio, $\delta_{X_D Y}/\delta_{XY}$ is equal to

$$\delta_{X_D X} = E\left[X_D X\right]/\sqrt{pq}$$
$$= E\left[X \mid X \geq k\right]/\sqrt{pq}.$$

It is a constant given $p$ and the distribution of $(X, Y)$. These correlations are invariant to location shifts and scaling, $X$ and $Y$ do not have to be centered and scaled, their means and variances can take any finite values. Once the ratio $(\delta_{X_D X})$ is found, one can compute the biserial correlation when the point-biserial correlation is specified. When $X$ is ordinalized to obtain $XO$, the fundamental ideas remain unchanged. If the assumptions of Eqs. 1 and 3 are met, the method is equally applicable to the ordinal case in the context of the relationship between the polyserial (before ordinalization) and point-polyserial (after ordinalization) correlations. The easiest way of computing $\delta_{X_O X}$ is to generate $X$ with a large number of data points, then ordinalize it to obtain $XO$, and then compute the sample correlation between $XO$ and $X$. $X$ could follow any continuous univariate distribution. However, here $X$ is assumed to be a part of MVN data before discretization.

8. For C-N combinations, use the count version of Eq. 3, which is $\delta_{X_C Y} = \delta_{XY}\delta_{X_C X}$ is valid. The only difference is that we use the inverse cdf method rather than discretization via thresholds as in the binary and ordinal cases.

9. For B-C and O-C combinations, suppose that there are two identical standard normal variables, one underlies the binary/ordinal variable before discretization, the other underlies the count variable before inverse cdf matching. One can find *Cor*(*O*, *N*) by the method of [7]. Then, assume *Cor*(*C*, *O*) = *Cor*(*C*, *N*)* *Cor*(*O*, *N*). *Cor*(*C*, *O*) is specified and *Cor*(*O*, *N*) is calculated. Solve for *Cor*(*C*, *N*). Then, find the underlying N-N correlation by Item 8 above [4,7].

10. Construct an overall, intermediate correlation matrix, $\Sigma^*$ using the results from Steps 4 through 9, in conjunction with the N-N part that remains untouched when we compute $\Sigma^*$ from $\Sigma$.

11. Check if $\Sigma^*$ is positive definite. If it is not, find the nearest positive definite correlation matrix by the method of Higham [18].

12. Generate multivariate normal data with a mean vector of (0, ..., 0) and correlation matrix of $\Sigma^*$, which can easily be done by using the Cholesky decomposition of $\Sigma^*$ and a vector of univariate normal draws. The Cholesky decomposition of $\Sigma^*$ produces a lower-triangular matrix $A$ for which $AA^T = \Sigma^*$. If

$z = (z_1, \ldots, z_d)$ are $d$ independent standard normal random variables, then $Z = Az$ is a random draw from this distribution.

13. Dichotomize binary, ordinalize ordinal by respective quantiles, go from normal to count by inverse cdf matching.

The assessment of the algorithm performance in terms of commonly accepted accuracy and precision measures in RNG and imputation settings as well as in other simulated environments can be carried out through the evaluation metric developed in Demirtas [19-34].

## 3 SOME OPERATIONAL DETAILS AND ILLUSTRATIVE SIMULATED EXAMPLES

The software implementation of the algorithm has been done in PoisBinOrd package within R environment [35,36]. The package consists of ten functions. The functions *validation.bin*, *validation.ord*, and *validation.corr* validate the specified quantities to prevent users from committing obvious specification errors. The function *correlation.limits* returns the lower and upper bounds of the pairwise correlation of Poisson-Poisson, Poisson-Binary, Poisson-Ordinal, Binary-Binary, Binary-Ordinal, and Ordinal-Ordinal combinations given their marginal distributions, i.e. returns the range of feasible pairwise correlations. The function *correlation.bound.check* checks the validity of the values of pairwise correlations. The functions *intermediate.corr.PP*, *intermediate.corr.BO*, and *intermediate.corr.PBO* computes intermediate correlation matrix for Poisson-Poisson combinations, binary/ordinal and binary/ordinal combinations, and Poisson and binary/ordinal combinations, respectively. The function *overall.corr.mat* assembles the final correlation matrix. The engine function *gen.PoisBinOrd* generates mixed data in accordance with the specified marginal and correlational quantities. Throughout the package, variables are supposed to be inputted in a certain order, namely, first count variables, next binary variables, and then ordinal variables should be placed.

### 3.1 Simulation Settings and Parameters of Interest

All simulated scenarios have been implemented via the use of the package **PoisBinOrd**. We assumed that there were six variables (two count, two binary, and two ordinal) in the system. We have chosen two levels of Poisson rates, two combinations of binary proportions, and two combinations for each of the ordinal variables, leading to 16 scenarios in total. Within each scenario, two sets of correlation structures and two levels of sample sizes (100 and 10,000) have been investigated. More specifically, for count variables, the Poisson rates ($\lambda_1$, $\lambda_2$) considered were (2, 7) and (0.2, 0.7); for binary variables, proportions ($p_1, p_2$) were chosen as (0.45, 0.50) and (0.80, 0.90); for ordinal variables, four levels (1, 2, 3, 4) were assumed with corresponding probabilities (0.25, 0.25, 0.25, 0.25) and (0.65, 0.15, 0.10, 0.10) for the first one, (0.20, 0.25, 0.30, 0.25) and (0.50, 0.30, 0.10, 0.10) for the second one. Table 1 summarizes all the marginal specifications. Note that the cumulative thresholds for each level (except for the last level) of ordinal variables are given rather than the marginal probabilities. Furthermore, the letter *P* is used for count variables rather than *C* that appeared in the algorithm to better reflect the fact that these variables follow the Poisson distribution. In regard to the association structure, two correlation matrices were employed. Each element of the correlation matrices was randomly sampled from a uniform distribution with range (-0.20, +0.20) in a way such that the resulting matrix is positive definite. Non-redundant entries for each of the two correlation matrices are given in Table 2 in a column format. From a marginal distribution standpoint, four distinct settings were considered: large Poisson rate/balanced binary and ordinal data (scenarios 1-4), large Poisson rate/imbalanced binary and ordinal data (scenarios 5-8), small Poisson rate/balanced binary and ordinal data (scenarios 9-12), small Poisson rate/imbalanced binary and ordinal data (scenarios 13-16). As far as the Poisson rates go, our small and large dichotomy was based on the arbitrary threshold 1, and what we mean by ’imbalanced’ in what follows is about proximity to 0.5 for binary variables (farther from 0.5 is labeled ’imbalanced’), and degree of deviations from uniform probabilities (0.25) for ordinal variables (farther from 0.25) is labeled ’imbalanced’).

**Table 1.** The 16 scenarios considered in simulations

| scenario | | | 1-4 | 5-8 | 9-12 | 13-16 |
|---|---|---|---|---|---|---|
| Poisson | $P_1$ | $\lambda_1$ | 2.00 | 2.00 | 0.20 | 0.20 |
| | $P_2$ | $\lambda_2$ | 7.00 | 7.00 | 0.70 | 0.70 |
| Binary | $B_1$ | $p_1$ | 0.45 | 0.80 | 0.45 | 0.80 |
| | $B_2$ | $p2$ | 0.50 | 0.90 | 0.50 | 0.90 |
| Ordinal | $O_1$ | $t_1$ | 0.25 | 0.65 | 0.25 | 0.65 |
| | | $t_2$ | 0.50 | 0.80 | 0.50 | 0.80 |
| | | $t_3$ | 0.75 | 0.90 | 0.75 | 0.90 |
| | $O_2$ | $t_1$ | 0.20 | 0.50 | 0.20 | 0.50 |
| | | $t_2$ | 0.45 | 0.80 | 0.45 | 0.80 |
| | | $t_3$ | 0.75 | 0.90 | 0.75 | 0.90 |

**Table 2.** The correlation structures of the 16 scenarios

| scenario | 1,2,5,6,9,10,13,14 | 3,4,7,8,11,12,15,16 |
|---|---|---|
| $\rho\ P_1P_2$ | 0.0789 | 0.0172 |
| $\rho\ P_1B_1$ | 0.1907 | -0.0619 |
| $\rho\ P_1B_2$ | -0.0488 | 0.0460 |
| $\rho\ P_1O_1$ | 0.0559 | -0.0962 |
| $\rho\ P_1O_2$ | -0.0784 | -0.1050 |
| $\rho\ P_2B_1$ | -0.0882 | -0.0405 |
| $\rho\ P_2B_2$ | 0.1942 | -0.0672 |
| $\rho\ P_2O_1$ | 0.0321 | 0.1549 |
| $\rho\ P_2O_2$ | 0.1450 | 0.1468 |
| $\rho\ B_1B_2$ | 0.1509 | 0.0897 |
| $\rho\ B_1O_1$ | 0.1148 | 0.0209 |
| $\rho\ B_1O_2$ | -0.0618 | -0.1080 |
| $\rho\ B_2O_1$ | -0.0373 | 0.0681 |
| $\rho\ B_2O_2$ | 0.1502 | 0.1476 |
| $\rho\ O_1O_2$ | 0.1219 | -0.0526 |

### 3.2 Evaluation Criteria and Results

In each scenario, 1,000 simulated data sets were generated to evaluate the performance of the method. Let the true parameter be $\theta$, and the estimated value be $\hat{\theta}$. The evaluation metrics include two accuracy measures: the relative bias (RB), defined as $E\left[\frac{\hat{\theta}-\theta}{\theta}\right]\times 100$ %, and the standardized bias (SB), defined as $E\left[\frac{\left|\hat{\theta}-\theta\right|}{SD\left(\hat{\theta}\right)}\right]\times 100\,\%$. RB is the deviation of the average estimates from the expected value with respect to $\theta$, whereas SB takes overall uncertainty in the system as the assessment base. The standard deviation (SD) of estimates across all simulation replicates is a pure precision quantity. Furthermore, the root mean square error (RMSE) of $\theta$, defined as $\sqrt{E\left[\hat{\theta}-\theta\right]^2}$, which is arguably the best integrated measure of accuracy and precision, and the coverage rate (CR), which is the percentage of times that $\theta$ is contained within a 95% confidence interval, are reported.

Tables 3-10 given in the Appendix show the true values (TV), average estimates (AE), SD, RB, SB, RMSE, and CR that are calculated across the 1,000 replications. Throughout these results, the discrepancies between the specified and empirically computed correlations are indiscernibly small and the deviations are within an acceptable range that can be expected in any stochastic process. For all marginal and associational quantities considered, relative and standardized biases as well as coverage rates and RMSEs demonstrate a close agreement with a nearly perfectly functioning procedure, lending a suggestive and compelling support to the presented methodology. Important relevant references in this context include Amatya and Demirtas (2015b, 2015c, 2016, 2017), Demirtas (2009b, 2009c, 2010b, 2014, 2017c, 2019); Demirtas and Gao (2022), Demirtas et al. (2014, 2017), Gao and Demirtas (2023), and Li et al. [37-53].

## 4 LOOKING AHEAD/FUTURE DIRECTIONS

The significance of the current study stems from the three major reasons: First, data analysts and practitioners across many different disciplines including pharmaceutical sciences can simulate multivariate data of mixed types with relative ease using this approach. Second, the proposed work can serve as a milestone for the development of more sophisticated simulation, computation, and data analysis techniques in the digital information, massive data era. Capability of generating many variables of different distributional types, nature, and dependence structures may be a contributing factor for better grasping the operational characteristics of today's intensive data trends. Overall, it provides a comprehensive and useful set of computational tools whose generality and flexibility offer promising potential for building enhanced statistical computing infrastructure for research and education.

While this work represents a decent step forward in mixed data generation, it may not be sufficiently complex for real-life applications in the sense that real count and continuous data are typically more complicated than what Poisson and normal distributions accommodate, and it is likely that specification of parameters that control the first two moments and the second order product moment is inadequate. To address these concerns, we plan on building a more inclusive structural umbrella, whose ingredients are as follows: First, the continuous part will be extended to encompass nonnormal continuous variables by the operational utility of the third order power polynomials. This approach is a moment- matching procedure where any given continuous variable in the system is expressed by the sum of linear combinations of powers of a standard normal variate, which requires the specification of the first four moments [11,12,14]. A more elaborate version in the form of the fifth order system will be implemented in an attempt to control for higher order moments to cover a larger area in the skewness-elongation plane and to provide a better approximation to the probability density functions of the continuous variables; and the count data part can be augmented through the generalized Poisson distribution that allows under- and over-dispersion, which is usually encountered in most applications, via an additional dispersion parameter [13,15,16]. Second, although the Pearson correlation may not be the best association quantity in every situation, it is the most widespread measure of association; and generality of the methods proposed herein with different kinds of variables requires the broadest possible framework. For further broadening the scale, scope, and applicability of the ideas presented in this paper, the proposed RNG technique can be extended to allow the specification of the Spearman's rho, which is more popular for discrete and heavily skewed continuous distributions, could be incorporated into the algorithm for concurrently generating all major types of variables. For the continuous-continuous pairs, the connection between the Pearson and Spearman correlations is given in Headrick through the power coefficients, and these two correlations are known to be equal for the binary-binary pairs [13]. The relationship can be derived for all other variable type combinations. Inclusion of Spearman's rho as an option will allow us to specify nonlinear associations whose monotonic components are reflected in the rank correlation. Third, the expanded fifth order polynomial system could be further augmented to accommodate L-moments and L-correlations that are based on expectations of certain linear combinations of order statistics. The marginal and product L-moments are known to be more robust to outliers than their conventional counterparts in the sense that they suffer less from the effects of sampling variability, and they enable more secure inferences to be made from small samples about an underlying probability distribution. On a related note, further expansions can be designed to handle more complex associations that involve higher order product moments.

The salient advantages of the proposed algorithm and its augmented versions are as follows: (1) Individual components are well-established. (2) Given their computational simplicity, generality, and flexibility, these methods are likely to be widely used by researchers, methodologists, and practitioners in a wide spectrum of scientific disciplines, especially in the big data era. (3) They could be very useful in graduate-level teaching of statistics courses that involve computation and simulation, and in training graduate students. (4) A specific set of moments for each variable is fairly rare in practice, but a specific distribution that would lead to these moments is very common; so having access to these methods is needed by potentially a large group of people. (5) Simulated variables can be treated as outcomes or predictors in subsequent statistical analyses as the variables are being generated jointly. (6) Required quantities can either be specified or estimated from a real data set. (7) The improved product after all these extensions will allow the specification of two prominent types of correlations (Pearson and Spearman correlations) and one emerging type (L-correlations) provided that they are within the limits imposed by marginal distributions. This makes it feasible to generate linear and a broad range of nonlinear associations. (8) The continuous part can include virtually any shape (skewness, low or high peakedness, mode at the boundary, multimodality, etc.) that is spanned by power polynomials; the count data part can be under- or over-dispersed. (9) Ability to jointly generate different types of data may facilitate comparisons among existing data analysis and computation methods in assessing the extent of conditions under which available methods work properly, and foster the development of new tools, especially in contexts where correlations play a significant role (e.g., longitudinal, clustered, and other multilevel settings). (10) The approaches presented here can be regarded as a variant of multivariate Gaussian copula-based methods as (a) the binary and ordinal variables are assumed to have a latent normal distribution before discretization; (b) the count variables go through a correlation mapping procedure via the anything-to-normal approach; and (c) the continuous variables consist of polynomial terms involving normals. To the best of our knowledge, existing multivariate copulas are not designed to have the generality of encompassing all these variable types simultaneously. (11) As the mixed data generation routine is involved with latent variables that are subsequently discretized, it should be possible to see how the correlation structure changes when some variables in a multivariate continuous setting are dichotomized/ordinalized [7,8,54]. An important by-product of this research will be a better understanding of the nature of discretization, which may have significant implications in interpreting the coefficients in regression-type models when some predictors are discretized. On a related note, this could be useful in meta-analysis when some studies discretize variables and some do not. (12) Availability of a general mixed data generation algorithm can markedly facilitate simulated power-sample size calculations for a broad range of statistical models.

## APPENDIX

The results that come out of a comprehensive simulation study that spans a wide range of parameter value combinations are given in Tables 3-10 below.

## 5 CONFLICT OF INTEREST

Authors declare that there is no conflict of interest.

## 6 REFERENCES

[1] Demirtas H. A method for multivariate ordinal data generation given marginal distributions and correlations. *Journal of Statistical Computation and Simulation*, 2006; **76**: 1017- 1025.

[2] Demirtas H, Doganay B. Simultaneous generation of binary and normal data with specified marginal and association structures. *Journal of Biopharmaceutical Statistics*, 2012; **22**: 223-236.

[3] Demirtas H, Yavuz Y. Concurrent generation of ordinal and normal data. *Journal of Biopharmaceutical Statistics*, 2015; **25**: 635-650.

[4] Amatya A, Demirtas H. Simultaneous generation of multivariate mixed data with Poisson and normal marginals. *Journal of Statistical Computation and Simulation*, 2015a; **85**: 3129-3139.

[5] Emrich JL, Piedmonte MR. A method for generating high-dimensional multivariate binary variates. *The American Statistician*, 1991; **45**: 302—304.

[6] Demirtas H, Hedeker D. A practical way for computing approximate lower and upper correlation bounds. *The American Statistician*, 2011; **65**: 104-109.

[7] Demirtas H, Hedeker D. Computing the point-biserial correlation under any underlying continuous distribution. *Communications in Statistics-Simulation and Computation*, 2016;

**45**: 2744-2751.

[8] Demirtas H, Ahmadian R, Atis S, Can FE, Ercan I. A nonnormal look at polychoric correlations: Modeling the change in correlations before and after discretization. *Computational Statistics*, 2016; **31**: 1385-1401.

[9] Ferrari PA, Barbiero A. Simulating ordinal data. *Multivariate Behavioral Research*, 2012; **47**: 566-589.

[10] Yahav I, Shmueli G. On generating multivariate Poisson data in management science applications. *Applied Stochastic Models in Business and Industry*, 2012; **28**: 91—102.

[11] Fleishman AI. A method for simulating non-normal distributions. *Psychometrika*, 1978; **43**: 521-532.

[12] Vale CD, Maurelli VA. Simulating multivariate nonnormal distributions. *Psychometrika*, 1983; **48**: 465-471.

[13] Headrick T.C. *Power Method Polynomials and Other Transformations*. Boca Raton. 2010; FL: Chapman and Hall/CRC.

[14] Demirtas H, Hedeker D, Mermelstein J.M. Simulation of massive public health data by power polynomials. *Statistics in Medicine*, 2012; **31**: 3337-3346.

[15] Demirtas H. Concurrent generation of binary and nonnormal continuous data through fifth order power polynomials. *Communications in Statistics-Simulation and Computation*, 2017a; **46**: 344-357.

[16] Demirtas H. On accurate and precise generation of generalized Poisson vari- ates. *Communications in Statistics-Simulation and Computation*, 2017b; **46**: 489-499.

[17] Nelsen RB. *An Introduction to Copulas*. 2006; Berlin, Germany: Springer.

[18] Higham N.J. Computing the nearest correlation matrix - a problem from finance. *IMA Journal of Numerical Analysis*, 2002; **22**: 329-343.

[19] Demirtas H. Simulation-driven inferences for multiply imputed longitudinal datasets. *Statistica Neerlandica*, 2004a; **58**: 466-482.

[20] Demirtas H. Assessment of relative improvement due to weights within gen- eralized estimating equations framework for incomplete clinical trials data. *Journal of Biopharmaceutical Statistics*, 2004b; **14**: 1085-1098.

[21] Demirtas H. Multiple imputation under Bayesianly smoothed pattern-mixture models for non-ignorable drop-out. *Statistics in Medicine*, 2005a; **24**: 2345-2363.

[22] Demirtas H. Bayesian analysis of hierarchical pattern-mixture models for clinical trials data with attrition and comparisons to commonly used ad-hoc and model- based approaches. *Journal of Biopharmaceutical Statistics*, 2005b; **15**: 383-402.

[23] Demirtas H. Practical advice on how to impute continuous data when the ultimate interest centers on dichotomized outcomes

through pre-specified thresholds. *Communications in Statistics-Simulation and Computation*, 2007a; **36**: 871-889.

[24] Demirtas H. The design of simulation studies in medical statistics. *Statistics in Medicine*, 2007b; **26**: 3818-3821.

[25] Demirtas H. On imputing continuous data when the eventual interest pertains to ordinalized outcomes via threshold concept. *Computational Statistics and Data Analysis*, 2008; **52**: 2261-2271.

[26] Demirtas H. Multiple imputation under the generalized lambda distribution. *Journal of Biopharmaceutical Statistics*, 2009a; **19**: 77-89.

[27] Demirtas H. A distance-based rounding strategy for post-imputation ordinal data. *Journal of Applied Statistics*, 2010a; **37**: 489-500.

[28] Demirtas H, Hedeker D. Gaussianization-based quasi-imputation and expan- sion strategies for incomplete correlated binary responses. *Statistics in Medicine*, 2007; **26**: 782-799.

[29] Demirtas H, Hedeker D. An imputation strategy for incomplete longitudinal ordinal data. *Statistics in Medicine*, 2008a; **27**: 4086-4093.

[30] Demirtas H, Hedeker D. Multiple imputation under power polynomials. *Communications in Statistics-Simulation and Computation*, 2008b; 37: 1682-1695.

[31] Demirtas H, Hedeker D. Imputing continuous data under some non-Gaussian distributions. Statistica *Neerlandica*, 2008c; **62**: 193-205.

[32] Demirtas H, Schafer JL. On the performance of random-coefficient pattern-mixture models for non-ignorable drop-out. *Statistics in Medicine*, 2003; **22**: 2253-2575.

[33] Demirtas H, Arguelles LM, Chung H, Hedeker D. On the performance of bias-reduction techniques for variance estimation in approximate Bayesian bootstrap imputation. *Computational Statistics and Data Analysis*, 2007; **51**: 4064-4068.

[34] Demirtas H, Freels SA, Yucel RM. Plausibility of multivariate normality assumption when multiply imputing non-Gaussian continuous outcomes: A simulation assessment. *Journal of Statistical Computation and Simulation*, 2008; **78**: 69-84.

[35] Inan G, Demirtas H, Gao R. *PoisBinOrd: Data Generation with Poisson, Bi- nary, and Ordinal Components*, 2021; URL:http://CRAN.Rproject.org/package=PoisBinOrd

[36] R Development Core Team. *R: A Language and Environment for Statistical Computing.* Foundation for Statistical Computing. 2023: Vienna, Austria.

[37] Amatya A, Demirtas H. MultiOrd: An R package for generating correlated ordinal data. *Communications in Statistics-Simulation and Computation*, 2015b; 44: 1683-1691.

[38] Amatya A, Demirtas H. OrdNor: An R

package for concurrent generation of ordinal and normal data. *Journal of Statistical Software*, 2015c; **68**: 1-14.

[39] Amatya A, Demirtas H. Concurrent generation of multivariate mixed data with variables of dissimilar types. *Journal of Statistical Computation and Simulation*, 2016; **86**: 3595-3607.

[40] Amatya A, Demirtas H. PoisNor: An R package for generation of multivariate data with Poisson and normal marginals. *Communications in Statistics-Simulation and Computation*, 2017; **46**: 2241-2253.

[41] Demirtas H. Rounding strategies for multiply imputed binary data. *Biometrical Journal*, 2009b; **51**: 677-688.

[42] Demirtas H. Multiple imputation for longitudinal data under a Bayesian multilevel model. *Communications in Statistics-Theory and Methods*, 2009c; **38**: 2812-2828.

[43] Demirtas H. An application of multiple imputation under the two generalized parametric families. *Journal of Data Science*, 2010b; **8**: 443-455.

[44] Demirtas H. Joint generation of binary and nonnormal continuous data. *Journal of Biometrics and Biostatistics*, 2014; **5**: 1-9.

[45] Demirtas H. A multiple imputation framework for massive multivariate data of different variable types: A Monte-Carlo technique. *In ICSA Book Series in Statistics, John Dean Chen and Ding-Geng (Din) Chen (Eds): Monte-Carlo Simulation- Based Statistical Modeling*. Singapore: Springer, 2017c: 143-162.

[46] Demirtas H. Inducing any feasible level of correlation to bivariate data with any marginals. *The American Statistician*, 2019; **73**: 273-277.

[47] Demirtas H, Gao R. Mixed data generation packages and related computational tools in R. *Communications in Statistics-Simulation and Computation,* 2022; **51**: 4520-4563.

[48] Demirtas H, Amatya A, Doganay B. BinNor: An R package for concur- rent generation of binary and normal data. *Communications in Statistics-Simulation and Computation*, 2014; **43**: 569-579.

[49] Demirtas H, Allozi R, Hu Y, Inan G, Ozbek L. Joint generation of binary, ordinal, count, and normal data with specified marginal and association structures in Monte-Carlo simulations. In ICSA Book Series in Statistics, John Dean Chen and Ding-Geng (Din) Chen (Eds): Monte-Carlo Simulation-Based Statistical Modeling. Singapore: Springer, 2017: 3-15.

[50] Gao R, Demirtas H. CorrToolBox: an R package for modeling correlational transformations in discretization contexts. *Communications in Statistics-Simulation and Computation*, 2023.

[51] Li H, Demirtas H, Chen R. RNGforGPD: An R package for generation of univariate and multivariate generalized Poisson data. The R

*Journal*, 2020; **12**: 120-133.

[52] Demirtas H, Amatya A, Pugach O, Cursio J, Shi F, Morton D, Doganay B. Accuracy versus convenience: a simulation-based comparison of two continuous imputation models for incomplete ordinal longitudinal clinical trials data. *Statistics and Its Interface*, 2009; **2**: 449-456.

[53] Demirtas H, Vardar-Acar, C. Anatomy of correlational magnitude transforma- tions in latency and discretization contexts in Monte-Carlo studies. *In ICSA Book Series in Statistics, John Dean Chen and Ding-Geng (Din) Chen (Eds): Monte-Carlo Simulation-Based Statistical Modeling*. Singapore: Springer, 2017; 59-84.

[54] Demirtas H. A note on the relationship between the phi coefficient and the tetrachoric correlation under nonnormal underlying distributions. *The American Statistician*, 2016; **70**: 143-148.

# APPENDIX

**Table 3.** Scenario 01/02, large Poisson rate and balanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 2.0000 | 1.9952 | 0.1399 | -0.2380 | 3.4020 | 0.1399 | 0.9520 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 6.9990 | 0.2593 | -0.0143 | 0.3856 | 0.2592 | 0.9540 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4520 | 0.0486 | 0.4333 | 4.0104 | 0.0486 | 0.9510 |
| | $B_2$ | $p_2$ | 0.5000 | 0.4998 | 0.0501 | -0.0340 | 0.3390 | 0.0501 | 0.9450 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2481 | 0.0438 | -0.7600 | 4.3394 | 0.0438 | 0.9360 |
| | | $t_2$ | 0.5000 | 0.4987 | 0.0512 | -0.2580 | 2.5207 | 0.0512 | 0.9410 |
| | | $t_3$ | 0.7500 | 0.7490 | 0.0439 | -0.1387 | 2.3689 | 0.0439 | 0.9200 |
| | $O_2$ | $t_1$ | 0.2000 | 0.1990 | 0.0415 | -0.4800 | 2.3127 | 0.0415 | 0.9430 |
| | | $t_2$ | 0.4500 | 0.4488 | 0.0511 | -0.2711 | 2.3882 | 0.0511 | 0.9370 |
| | | $t_3$ | 0.7500 | 0.7492 | 0.0442 | -0.1000 | 1.6979 | 0.0442 | 0.9390 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0797 | 0.1004 | 1.0134 | 0.7960 | 0.1004 | 0.9430 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1932 | 0.0975 | 1.3132 | 2.5692 | 0.0974 | 0.9360 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0473 | 0.1011 | -3.0294 | 1.4631 | 0.1010 | 0.9480 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0566 | 0.1002 | 1.4139 | 0.7879 | 0.1002 | 0.9520 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0775 | 0.1002 | -1.1388 | 0.8911 | 0.1002 | 0.9470 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0848 | 0.0996 | -3.9237 | 3.4750 | 0.0996 | 0.9520 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1956 | 0.0962 | 0.7283 | 1.4696 | 0.0962 | 0.9470 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0311 | 0.0984 | -3.1446 | 1.0256 | 0.0983 | 0.9610 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1449 | 0.0988 | -0.0781 | 0.1147 | 0.0987 | 0.9470 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1499 | 0.1017 | -0.6450 | 0.9569 | 0.1016 | 0.9420 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1153 | 0.1003 | 0.4452 | 0.5095 | 0.1003 | 0.9490 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0652 | 0.0979 | 5.4187 | 3.4206 | 0.0979 | 0.9560 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0318 | 0.1015 | -14.6723 | 5.3869 | 0.1016 | 0.9420 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1518 | 0.0987 | 1.0362 | 1.5761 | 0.0987 | 0.9450 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1181 | 0.0975 | -3.1432 | 3.9308 | 0.0975 | 0.9490 |
| 10000 | $P_1$ | $\lambda_1$ | 2.0000 | 2.0008 | 0.0141 | 0.0391 | 5.5376 | 0.0141 | 0.9490 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 7.0003 | 0.0271 | 0.0038 | 0.9786 | 0.0271 | 0.9430 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4499 | 0.0050 | -0.0262 | 2.3706 | 0.0050 | 0.9510 |
| | $B_2$ | $p_2$ | 0.5000 | 0.5000 | 0.0051 | 0.0053 | 0.5145 | 0.0051 | 0.9390 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2500 | 0.0045 | 0.0095 | 0.5310 | 0.0045 | 0.9440 |
| | | $t_2$ | 0.5000 | 0.4999 | 0.0049 | -0.0216 | 2.2044 | 0.0049 | 0.9480 |
| | | $t_3$ | 0.7500 | 0.7501 | 0.0044 | 0.0180 | 3.0753 | 0.0044 | 0.9540 |
| | $O_2$ | $t_1$ | 0.2000 | 0.2001 | 0.0040 | 0.0346 | 1.7165 | 0.0040 | 0.9500 |
| | | $t_2$ | 0.4500 | 0.4501 | 0.0050 | 0.0179 | 1.6071 | 0.0050 | 0.9560 |
| | | $t_3$ | 0.7500 | 0.7501 | 0.0043 | 0.0127 | 2.2259 | 0.0043 | 0.9460 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0792 | 0.0101 | 0.4393 | 3.4248 | 0.0101 | 0.9430 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1914 | 0.0096 | 0.3846 | 7.6666 | 0.0096 | 0.9490 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0489 | 0.0097 | 0.2336 | 1.1773 | 0.0097 | 0.9530 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0561 | 0.0100 | 0.5117 | 2.8578 | 0.0100 | 0.9520 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0785 | 0.0101 | 0.1621 | 1.2605 | 0.0101 | 0.9450 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0882 | 0.0102 | 0.0256 | 0.2223 | 0.0101 | 0.9490 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1939 | 0.0096 | -0.1627 | 3.2895 | 0.0096 | 0.9560 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0316 | 0.0098 | -1.3955 | 4.5498 | 0.0098 | 0.9560 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1452 | 0.0100 | 0.1111 | 1.6182 | 0.0100 | 0.9420 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1508 | 0.0098 | -0.0272 | 0.4181 | 0.0098 | 0.9480 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1149 | 0.0100 | 0.0710 | 0.8151 | 0.0100 | 0.9430 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0622 | 0.0103 | 0.7170 | 4.3077 | 0.0103 | 0.9440 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0373 | 0.0096 | 0.0031 | 0.0121 | 0.0096 | 0.9640 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1501 | 0.0097 | -0.0600 | 0.9303 | 0.0097 | 0.9540 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1219 | 0.0095 | -0.0130 | 0.1670 | 0.0095 | 0.9620 |

**Table 4.** Scenario 03/04, large Poisson rate and balanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 2.0000 | 1.9988 | 0.1380 | -0.0595 | 0.8624 | 0.1379 | 0.9580 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 6.9979 | 0.2739 | -0.0296 | 0.7556 | 0.2738 | 0.9260 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4504 | 0.0505 | 0.0822 | 0.7326 | 0.0505 | 0.9330 |
| | $B_2$ | $p_2$ | 0.5000 | 0.4976 | 0.0525 | -0.4700 | 4.4735 | 0.0526 | 0.9240 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2504 | 0.0433 | 0.1800 | 1.0392 | 0.0433 | 0.9490 |
| | | $t_2$ | 0.5000 | 0.5008 | 0.0498 | 0.1700 | 1.7074 | 0.0498 | 0.9480 |
| | | $t_3$ | 0.7500 | 0.7517 | 0.0428 | 0.2227 | 3.9052 | 0.0428 | 0.9280 |
| | $O_2$ | $t_1$ | 0.2000 | 0.1996 | 0.0406 | -0.2100 | 1.0344 | 0.0406 | 0.9500 |
| | | $t_2$ | 0.4500 | 0.4492 | 0.0491 | -0.1733 | 1.5871 | 0.0491 | 0.9550 |
| | | $t_3$ | 0.7500 | 0.7502 | 0.0425 | 0.0227 | 0.3997 | 0.0425 | 0.9450 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0177 | 0.1021 | 2.9406 | 0.4959 | 0.1021 | 0.9420 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0608 | 0.0990 | -1.6395 | 1.0248 | 0.0989 | 0.9550 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0427 | 0.0962 | -7.1895 | 3.4363 | 0.0963 | 0.9560 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0950 | 0.0965 | -1.2477 | 1.2429 | 0.0965 | 0.9550 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.1015 | 0.1003 | -3.3185 | 3.4752 | 0.1003 | 0.9400 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0407 | 0.0990 | 0.4266 | 0.1748 | 0.0989 | 0.9450 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0669 | 0.1021 | -0.4419 | 0.2911 | 0.1020 | 0.9400 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1605 | 0.0991 | 3.5906 | 5.6119 | 0.0992 | 0.9400 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1424 | 0.1004 | -3.0004 | 4.3875 | 0.1005 | 0.9420 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0907 | 0.0966 | 1.1309 | 1.0498 | 0.0966 | 0.9510 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0189 | 0.1008 | -9.3592 | 1.9405 | 0.1007 | 0.9430 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1107 | 0.0955 | 2.4979 | 2.8238 | 0.0955 | 0.9560 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0689 | 0.0985 | 1.2663 | 0.8748 | 0.0985 | 0.9480 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1500 | 0.0969 | 1.6017 | 2.4390 | 0.0969 | 0.9420 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0494 | 0.0970 | -6.0600 | 3.2878 | 0.0970 | 0.9600 |
| 10000 | $P_1$ | $\lambda_1$ | 2.0000 | 2.0000 | 0.0143 | -0.0017 | 0.2337 | 0.0143 | 0.9410 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 6.9992 | 0.0265 | -0.0118 | 3.1271 | 0.0265 | 0.9550 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4496 | 0.0050 | -0.0848 | 7.6435 | 0.0050 | 0.9460 |
| | $B_2$ | $p_2$ | 0.5000 | 0.5001 | 0.0050 | 0.0106 | 1.0594 | 0.0050 | 0.9490 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2502 | 0.0043 | 0.0942 | 5.5233 | 0.0043 | 0.9570 |
| | | $t_2$ | 0.5000 | 0.5002 | 0.0049 | 0.0326 | 3.3310 | 0.0049 | 0.9510 |
| | | $t_3$ | 0.7500 | 0.7500 | 0.0045 | 0.0058 | 0.9617 | 0.0045 | 0.9540 |
| | $O_2$ | $t_1$ | 0.2000 | 0.1997 | 0.0039 | -0.1278 | 6.4889 | 0.0039 | 0.9360 |
| | | $t_2$ | 0.4500 | 0.4497 | 0.0048 | -0.0581 | 5.4616 | 0.0048 | 0.9480 |
| | | $t_3$ | 0.7500 | 0.7500 | 0.0043 | -0.0061 | 1.0474 | 0.0043 | 0.9430 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0171 | 0.0099 | -0.9153 | 1.5889 | 0.0099 | 0.9490 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0621 | 0.0097 | 0.4678 | 2.9869 | 0.0097 | 0.9550 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0465 | 0.0099 | 1.1459 | 5.3300 | 0.0099 | 0.9530 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0961 | 0.0102 | -0.0280 | 0.2642 | 0.0102 | 0.9380 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.1049 | 0.0100 | -0.1259 | 1.3154 | 0.0100 | 0.9490 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0409 | 0.0099 | 0.8270 | 3.3707 | 0.0099 | 0.9440 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0672 | 0.0101 | 0.0090 | 0.0603 | 0.0101 | 0.9450 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1548 | 0.0096 | -0.1035 | 1.6654 | 0.0096 | 0.9580 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1467 | 0.0094 | -0.0796 | 1.2392 | 0.0094 | 0.9620 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0899 | 0.0099 | 0.2498 | 2.2610 | 0.0099 | 0.9550 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0210 | 0.0096 | 0.3741 | 0.8131 | 0.0096 | 0.9530 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1081 | 0.0101 | 0.1271 | 1.3557 | 0.0101 | 0.9460 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0678 | 0.0102 | -0.3375 | 2.2604 | 0.0102 | 0.9450 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1469 | 0.0097 | -0.4828 | 7.3258 | 0.0097 | 0.9510 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0525 | 0.0101 | -0.1716 | 0.8922 | 0.0101 | 0.9380 |

**Table 5**. Scenario 05/06, large Poisson rate and imbalanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 2.0000 | 2.0023 | 0.1390 | 0.1165 | 1.6758 | 0.1390 | 0.9520 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 7.0100 | 0.2688 | 0.1431 | 3.7282 | 0.2688 | 0.9470 |
| | $B_1$ | $p_1$ | 0.8000 | 0.7995 | 0.0408 | -0.0675 | 1.3232 | 0.0408 | 0.9330 |
| | $B_2$ | $p_2$ | 0.9000 | 0.9004 | 0.0307 | 0.0444 | 1.3009 | 0.0307 | 0.9240 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6501 | 0.0485 | 0.0154 | 0.2063 | 0.0484 | 0.9360 |
| | | $t_2$ | 0.8000 | 0.7993 | 0.0402 | -0.0850 | 1.6901 | 0.0402 | 0.9470 |
| | | $t_3$ | 0.9000 | 0.9005 | 0.0306 | 0.0556 | 1.6360 | 0.0306 | 0.9390 |
| | $O_2$ | $t_1$ | 0.5000 | 0.5006 | 0.0494 | 0.1100 | 1.1140 | 0.0493 | 0.9140 |
| | | $t_2$ | 0.8000 | 0.8010 | 0.0395 | 0.1200 | 2.4318 | 0.0395 | 0.9280 |
| | | $t_3$ | 0.9000 | 0.9008 | 0.0307 | 0.0844 | 2.4771 | 0.0307 | 0.9320 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0740 | 0.0963 | -6.1466 | 5.0310 | 0.0964 | 0.9540 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1821 | 0.0889 | -4.5045 | 9.6659 | 0.0892 | 0.9650 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0450 | 0.1019 | -7.7583 | 3.7143 | 0.1020 | 0.9380 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0527 | 0.1012 | -5.7095 | 3.1519 | 0.1012 | 0.9490 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0803 | 0.1009 | 2.3651 | 1.8376 | 0.1009 | 0.9370 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0888 | 0.1026 | 0.6346 | 0.5456 | 0.1026 | 0.9410 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1850 | 0.0936 | -4.7374 | 9.8297 | 0.0940 | 0.9560 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0349 | 0.1013 | 8.7773 | 2.7798 | 0.1013 | 0.9430 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1403 | 0.1033 | -3.2978 | 4.6321 | 0.1033 | 0.9290 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1534 | 0.1180 | 1.6891 | 2.1601 | 0.1179 | 0.8800 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1111 | 0.0875 | -3.2815 | 4.3066 | 0.0875 | 0.9690 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0615 | 0.1031 | -0.5192 | 0.3113 | 0.1030 | 0.9390 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0349 | 0.1007 | -6.2317 | 2.3054 | 0.1007 | 0.9490 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1492 | 0.0776 | -0.6747 | 1.3061 | 0.0776 | 0.9760 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1199 | 0.1047 | -1.6597 | 1.9325 | 0.1047 | 0.9320 |
| 10000 | $P_1$ | $\lambda_1$ | 2.0000 | 1.9989 | 0.0142 | -0.0526 | 7.4149 | 0.0142 | 0.9440 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 7.0007 | 0.0258 | 0.0095 | 2.5871 | 0.0258 | 0.9540 |
| | $B_1$ | $p_1$ | 0.8000 | 0.7999 | 0.0040 | -0.0080 | 1.6001 | 0.0040 | 0.9510 |
| | $B_2$ | $p_2$ | 0.9000 | 0.9001 | 0.0030 | 0.0155 | 4.7191 | 0.0030 | 0.9550 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6500 | 0.0048 | 0.0025 | 0.3375 | 0.0048 | 0.9480 |
| | | $t_2$ | 0.8000 | 0.7999 | 0.0041 | -0.0098 | 1.9000 | 0.0041 | 0.9530 |
| | | $t_3$ | 0.9000 | 0.9000 | 0.0031 | -0.0033 | 0.9401 | 0.0031 | 0.9330 |
| | $O_2$ | $t_1$ | 0.5000 | 0.4999 | 0.0054 | -0.0119 | 1.1008 | 0.0054 | 0.9410 |
| | | $t_2$ | 0.8000 | 0.7998 | 0.0039 | -0.0222 | 4.5252 | 0.0039 | 0.9430 |
| | | $t_3$ | 0.9000 | 0.9000 | 0.0029 | -0.0023 | 0.6946 | 0.0029 | 0.9540 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0788 | 0.0099 | -0.0448 | 0.3586 | 0.0099 | 0.9510 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1835 | 0.0091 | -3.7723 | 79.0325 | 0.0116 | 0.9010 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0497 | 0.0100 | 1.8121 | 8.8263 | 0.0101 | 0.9490 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0559 | 0.0102 | 0.1595 | 0.8757 | 0.0102 | 0.9470 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0784 | 0.0097 | -0.0078 | 0.0635 | 0.0097 | 0.9620 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0890 | 0.0095 | 0.8826 | 8.2038 | 0.0095 | 0.9580 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1888 | 0.0089 | -2.7915 | 60.7903 | 0.0104 | 0.9330 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0321 | 0.0095 | -0.0364 | 0.1224 | 0.0095 | 0.9590 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1457 | 0.0102 | 0.4530 | 6.4226 | 0.0102 | 0.9440 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1504 | 0.0120 | -0.3256 | 4.0959 | 0.0120 | 0.8840 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1147 | 0.0085 | -0.0817 | 1.0982 | 0.0085 | 0.9740 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0623 | 0.0107 | 0.8284 | 4.7951 | 0.0107 | 0.9420 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0376 | 0.0107 | 1.0065 | 3.5121 | 0.0107 | 0.9190 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1503 | 0.0074 | 0.0517 | 1.0527 | 0.0074 | 0.9920 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1222 | 0.0108 | 0.2243 | 2.5418 | 0.0108 | 0.9300 |

**Table 6.** Scenario 07/08, large Poisson rate and imbalanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 2.0000 | 2.0031 | 0.1400 | 0.1575 | 2.2506 | 0.1399 | 0.9550 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 6.9920 | 0.2627 | -0.1140 | 3.0375 | 0.2627 | 0.9430 |
| | $B_1$ | $p_1$ | 0.8000 | 0.7992 | 0.0391 | -0.0988 | 2.0230 | 0.0390 | 0.9340 |
| | $B_2$ | $p_2$ | 0.9000 | 0.8986 | 0.0307 | -0.1522 | 4.4619 | 0.0307 | 0.9280 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6504 | 0.0462 | 0.0662 | 0.9304 | 0.0462 | 0.9540 |
| | | $t_2$ | 0.8000 | 0.8000 | 0.0399 | 0.0062 | 0.1253 | 0.0399 | 0.9290 |
| | | $t_3$ | 0.9000 | 0.8996 | 0.0297 | -0.0400 | 1.2135 | 0.0297 | 0.9380 |
| | $O_2$ | $t_1$ | 0.5000 | 0.5013 | 0.0509 | 0.2660 | 2.6108 | 0.0509 | 0.9350 |
| | | $t_2$ | 0.8000 | 0.8008 | 0.0408 | 0.1000 | 1.9600 | 0.0408 | 0.9420 |
| | | $t_3$ | 0.9000 | 0.9006 | 0.0310 | 0.0700 | 2.0292 | 0.0310 | 0.9150 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0160 | 0.1004 | -6.9688 | 1.1954 | 0.1003 | 0.9440 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0632 | 0.1073 | 2.1596 | 1.2452 | 0.1072 | 0.9280 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0473 | 0.0986 | 2.9042 | 1.3552 | 0.0985 | 0.9500 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0950 | 0.0971 | -1.1919 | 1.1799 | 0.0971 | 0.9540 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.1034 | 0.0946 | -1.4833 | 1.6473 | 0.0945 | 0.9620 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0385 | 0.0971 | -4.9782 | 2.0790 | 0.0971 | 0.9470 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0702 | 0.1008 | 4.3457 | 2.8995 | 0.1008 | 0.9420 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1521 | 0.0996 | -1.8303 | 2.8463 | 0.0996 | 0.9390 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1424 | 0.1018 | -2.9973 | 4.3226 | 0.1019 | 0.9460 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0803 | 0.1147 | -10.4572 | 8.1773 | 0.1150 | 0.9050 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0175 | 0.1007 | -16.3906 | 3.4006 | 0.1007 | 0.9480 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1061 | 0.1081 | -1.7358 | 1.7338 | 0.1081 | 0.9200 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0671 | 0.0888 | -1.4635 | 1.1216 | 0.0888 | 0.9720 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1518 | 0.0719 | 2.8645 | 5.8767 | 0.0720 | 0.9920 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0526 | 0.1012 | 0.0605 | 0.0315 | 0.1011 | 0.9470 |
| 10000 | $P_1$ | $\lambda_1$ | 2.0000 | 1.9995 | 0.0143 | -0.0263 | 3.6637 | 0.0143 | 0.9420 |
| | $P_2$ | $\lambda_2$ | 7.0000 | 6.9999 | 0.0258 | -0.0020 | 0.5327 | 0.0258 | 0.9560 |
| | $B_1$ | $p_1$ | 0.8000 | 0.8001 | 0.0039 | 0.0183 | 3.7198 | 0.0039 | 0.9520 |
| | $B_2$ | $p_2$ | 0.9000 | 0.8999 | 0.0031 | -0.0064 | 1.8729 | 0.0031 | 0.9420 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6499 | 0.0047 | -0.0088 | 1.2048 | 0.0047 | 0.9530 |
| | | $t_2$ | 0.8000 | 0.8000 | 0.0040 | -0.0047 | 0.9360 | 0.0040 | 0.9600 |
| | | $t_3$ | 0.9000 | 0.9000 | 0.0031 | -0.0012 | 0.3436 | 0.0031 | 0.9550 |
| | $O_2$ | $t_1$ | 0.5000 | 0.5000 | 0.0050 | -0.0044 | 0.4401 | 0.0050 | 0.9420 |
| | | $t_2$ | 0.8000 | 0.8001 | 0.0039 | 0.0083 | 1.6814 | 0.0039 | 0.9540 |
| | | $t_3$ | 0.9000 | 0.9000 | 0.0030 | 0.0007 | 0.2112 | 0.0030 | 0.9490 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0177 | 0.0099 | 2.8101 | 4.9052 | 0.0099 | 0.9470 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0627 | 0.0104 | 1.3465 | 7.9766 | 0.0105 | 0.9320 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0452 | 0.0099 | -1.7359 | 8.0515 | 0.0099 | 0.9450 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0945 | 0.0095 | -1.6856 | 17.1194 | 0.0096 | 0.9560 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.1035 | 0.0096 | -1.4060 | 15.3555 | 0.0097 | 0.9600 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0405 | 0.0099 | -0.1261 | 0.5188 | 0.0099 | 0.9400 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0684 | 0.0105 | 1.7352 | 11.0966 | 0.0106 | 0.9330 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1562 | 0.0098 | 0.8458 | 13.4212 | 0.0098 | 0.9400 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1472 | 0.0102 | 0.2822 | 4.0702 | 0.0102 | 0.9430 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0893 | 0.0114 | -0.3827 | 3.0228 | 0.0114 | 0.9170 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0213 | 0.0097 | 1.8356 | 3.9361 | 0.0097 | 0.9550 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1081 | 0.0105 | 0.0704 | 0.7274 | 0.0105 | 0.9340 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0683 | 0.0089 | 0.3796 | 2.8931 | 0.0089 | 0.9740 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1475 | 0.0077 | -0.0732 | 1.3938 | 0.0077 | 0.9900 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0533 | 0.0099 | 1.2760 | 6.7619 | 0.0099 | 0.9550 |

**Table 7.** Scenario 09/10, small Poisson rate and balanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 0.2000 | 0.2000 | 0.0432 | 0.0100 | 0.0463 | 0.0432 | 0.9450 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.6944 | 0.0862 | -0.8014 | 6.5107 | 0.0863 | 0.9400 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4543 | 0.0496 | 0.9667 | 8.7622 | 0.0498 | 0.9450 |
| | $B_2$ | $p_2$ | 0.5000 | 0.5032 | 0.0525 | 0.6440 | 6.1379 | 0.0525 | 0.9360 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2520 | 0.0442 | 0.7880 | 4.4605 | 0.0442 | 0.9410 |
| | | $t_2$ | 0.5000 | 0.4994 | 0.0509 | -0.1200 | 1.1796 | 0.0508 | 0.9440 |
| | | $t_3$ | 0.7500 | 0.7486 | 0.0414 | -0.1853 | 3.3607 | 0.0414 | 0.9350 |
| | $O_2$ | $t_1$ | 0.2000 | 0.2008 | 0.0402 | 0.3950 | 1.9637 | 0.0402 | 0.9630 |
| | | $t_2$ | 0.4500 | 0.4503 | 0.0492 | 0.0644 | 0.5891 | 0.0492 | 0.9410 |
| | | $t_3$ | 0.7500 | 0.7488 | 0.0426 | -0.1573 | 2.7685 | 0.0426 | 0.9440 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0841 | 0.1056 | 6.6603 | 4.9753 | 0.1056 | 0.9260 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1897 | 0.0922 | -0.4970 | 1.0280 | 0.0921 | 0.9580 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0514 | 0.0993 | 5.3863 | 2.6483 | 0.0993 | 0.9470 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0565 | 0.1019 | 1.0770 | 0.5906 | 0.1018 | 0.9350 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0811 | 0.0996 | 3.4758 | 2.7373 | 0.0996 | 0.9430 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0851 | 0.0954 | -3.5652 | 3.2950 | 0.0954 | 0.9550 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1989 | 0.0944 | 2.4232 | 4.9867 | 0.0944 | 0.9540 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0349 | 0.0981 | 8.8730 | 2.9033 | 0.0981 | 0.9580 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1418 | 0.0968 | -2.2247 | 3.3324 | 0.0968 | 0.9610 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1539 | 0.1001 | 1.9745 | 2.9764 | 0.1001 | 0.9420 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1137 | 0.0976 | -0.9709 | 1.1421 | 0.0976 | 0.9520 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0602 | 0.0965 | -2.5860 | 1.6559 | 0.0965 | 0.9540 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0370 | 0.1000 | -0.7159 | 0.2667 | 0.1000 | 0.9440 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1457 | 0.1037 | -2.9893 | 4.3315 | 0.1037 | 0.9290 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1235 | 0.0984 | 1.3143 | 1.6287 | 0.0984 | 0.9460 |
| 10000 | $P_1$ | $\lambda_1$ | 0.2000 | 0.1999 | 0.0045 | -0.0696 | 3.0750 | 0.0045 | 0.9420 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.7000 | 0.0082 | 0.0028 | 0.2406 | 0.0082 | 0.9570 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4497 | 0.0048 | -0.0640 | 5.9429 | 0.0049 | 0.9560 |
| | $B_2$ | $p_2$ | 0.5000 | 0.4997 | 0.0049 | -0.0536 | 5.4576 | 0.0049 | 0.9580 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2499 | 0.0044 | -0.0466 | 2.6391 | 0.0044 | 0.9410 |
| | | $t_2$ | 0.5000 | 0.4999 | 0.0048 | -0.0293 | 3.0506 | 0.0048 | 0.9510 |
| | | $t_3$ | 0.7500 | 0.7498 | 0.0042 | -0.0241 | 4.2607 | 0.0042 | 0.9580 |
| | $O_2$ | $t_1$ | 0.2000 | 0.2000 | 0.0038 | -0.0230 | 1.1983 | 0.0038 | 0.9590 |
| | | $t_2$ | 0.4500 | 0.4500 | 0.0048 | 0.0044 | 0.4106 | 0.0048 | 0.9580 |
| | | $t_3$ | 0.7500 | 0.7498 | 0.0042 | -0.0222 | 3.9557 | 0.0042 | 0.9640 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0887 | 0.0108 | 12.5004 | 91.1400 | 0.0146 | 0.8100 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1937 | 0.0095 | 1.5923 | 32.0573 | 0.0099 | 0.9510 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0486 | 0.0099 | -0.5130 | 2.5282 | 0.0099 | 0.9570 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0558 | 0.0099 | -0.0879 | 0.4943 | 0.0099 | 0.9580 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0788 | 0.0099 | 0.4916 | 3.8779 | 0.0099 | 0.9430 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0881 | 0.0104 | -0.1767 | 1.5040 | 0.0104 | 0.9400 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1949 | 0.0093 | 0.3750 | 7.8721 | 0.0093 | 0.9500 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0318 | 0.0102 | -0.9066 | 2.8477 | 0.0102 | 0.9360 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1445 | 0.0095 | -0.3743 | 5.7251 | 0.0095 | 0.9560 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1514 | 0.0098 | 0.3247 | 4.9800 | 0.0098 | 0.9470 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1153 | 0.0097 | 0.4417 | 5.2369 | 0.0097 | 0.9520 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0618 | 0.0101 | -0.0096 | 0.0588 | 0.0101 | 0.9500 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0371 | 0.0100 | -0.4224 | 1.5764 | 0.0100 | 0.9500 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1500 | 0.0097 | -0.1460 | 2.2564 | 0.0097 | 0.9530 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1217 | 0.0100 | -0.1731 | 2.1000 | 0.0100 | 0.9440 |

**Table 8.** Scenario 11/12, small Poisson rate and balanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 0.2000 | 0.2005 | 0.0442 | 0.2500 | 1.1316 | 0.0442 | 0.9470 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.6978 | 0.0827 | -0.3129 | 2.6482 | 0.0827 | 0.9520 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4496 | 0.0491 | -0.0844 | 0.7735 | 0.0491 | 0.9420 |
| | $B_2$ | $p_2$ | 0.5000 | 0.4990 | 0.0502 | -0.2100 | 2.0927 | 0.0502 | 0.9370 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2493 | 0.0442 | -0.2800 | 1.5821 | 0.0442 | 0.9400 |
| | | $t_2$ | 0.5000 | 0.4998 | 0.0510 | -0.0460 | 0.4511 | 0.0510 | 0.9360 |
| | | $t_3$ | 0.7500 | 0.7500 | 0.0435 | 0.0000 | 0.0000 | 0.0434 | 0.9250 |
| | $O_2$ | $t_1$ | 0.2000 | 0.1968 | 0.0399 | -1.5850 | 7.9408 | 0.0400 | 0.9460 |
| | | $t_2$ | 0.4500 | 0.4505 | 0.0508 | 0.1022 | 0.9062 | 0.0507 | 0.9290 |
| | | $t_3$ | 0.7500 | 0.7487 | 0.0456 | -0.1733 | 2.8506 | 0.0456 | 0.9320 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0149 | 0.0995 | -13.7388 | 2.3786 | 0.0994 | 0.9540 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0628 | 0.0994 | 1.5404 | 0.9588 | 0.0993 | 0.9470 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0485 | 0.0978 | 5.5041 | 2.5877 | 0.0978 | 0.9530 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0955 | 0.0950 | -0.7321 | 0.7410 | 0.0950 | 0.9600 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.1002 | 0.0971 | -4.5392 | 4.9074 | 0.0972 | 0.9540 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0417 | 0.0983 | 2.9416 | 1.2135 | 0.0983 | 0.9590 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0638 | 0.1009 | -5.1396 | 3.4248 | 0.1009 | 0.9460 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1551 | 0.1003 | 0.0874 | 0.1350 | 0.1003 | 0.9410 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1483 | 0.0975 | 0.9772 | 1.4717 | 0.0975 | 0.9530 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0921 | 0.1026 | 2.6919 | 2.3528 | 0.1026 | 0.9430 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0179 | 0.0994 | -14.4660 | 3.0401 | 0.0994 | 0.9510 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1052 | 0.0964 | -2.5911 | 2.9033 | 0.0964 | 0.9560 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0701 | 0.1007 | 3.0470 | 2.0592 | 0.1007 | 0.9480 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1439 | 0.0981 | -2.5112 | 3.7770 | 0.0982 | 0.9440 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0520 | 0.1001 | -1.0517 | 0.5527 | 0.1000 | 0.9470 |
| 10000 | $P_1$ | $\lambda_1$ | 0.2000 | 0.2001 | 0.0044 | 0.0662 | 3.0201 | 0.0044 | 0.9550 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.6996 | 0.0086 | -0.0541 | 4.4001 | 0.0086 | 0.9490 |
| | $B_1$ | $p_1$ | 0.4500 | 0.4501 | 0.0050 | 0.0138 | 1.2351 | 0.0050 | 0.9460 |
| | $B_2$ | $p_2$ | 0.5000 | 0.4997 | 0.0049 | -0.0610 | 6.2412 | 0.0049 | 0.9570 |
| | $O_1$ | $t_1$ | 0.2500 | 0.2501 | 0.0045 | 0.0203 | 1.1352 | 0.0045 | 0.9450 |
| | | $t_2$ | 0.5000 | 0.5000 | 0.0051 | -0.0057 | 0.5580 | 0.0051 | 0.9400 |
| | | $t_3$ | 0.7500 | 0.7501 | 0.0041 | 0.0138 | 2.4967 | 0.0041 | 0.9550 |
| | $O_2$ | $t_1$ | 0.2000 | 0.2001 | 0.0040 | 0.0465 | 2.3025 | 0.0040 | 0.9620 |
| | | $t_2$ | 0.4500 | 0.4501 | 0.0050 | 0.0161 | 1.4407 | 0.0050 | 0.9450 |
| | | $t_3$ | 0.7500 | 0.7503 | 0.0043 | 0.0437 | 7.6743 | 0.0043 | 0.9560 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0197 | 0.0101 | 14.5791 | 24.8092 | 0.0104 | 0.9450 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0615 | 0.0092 | -0.5137 | 3.4550 | 0.0092 | 0.9660 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0460 | 0.0102 | -0.1012 | 0.4545 | 0.0102 | 0.9410 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0954 | 0.0100 | -0.8153 | 7.8785 | 0.0100 | 0.9570 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.1053 | 0.0098 | 0.3148 | 3.3805 | 0.0098 | 0.9530 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0402 | 0.0098 | -0.7422 | 3.0615 | 0.0098 | 0.9590 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0670 | 0.0100 | -0.2784 | 1.8779 | 0.0100 | 0.9480 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1550 | 0.0097 | 0.0345 | 0.5497 | 0.0097 | 0.9460 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1459 | 0.0095 | -0.6182 | 9.5337 | 0.0096 | 0.9520 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0891 | 0.0101 | -0.6855 | 6.1165 | 0.0101 | 0.9420 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0212 | 0.0097 | 1.5166 | 3.2714 | 0.0097 | 0.9550 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1085 | 0.0098 | 0.4387 | 4.8277 | 0.0098 | 0.9610 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0686 | 0.0100 | 0.8376 | 5.7179 | 0.0100 | 0.9390 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1473 | 0.0102 | -0.2288 | 3.3156 | 0.0102 | 0.9410 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0526 | 0.0095 | 0.0789 | 0.4347 | 0.0095 | 0.9640 |

**Table 9.** Scenario 13/14, small Poisson rate and imbalanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 0.2000 | 0.1983 | 0.0431 | -0.8650 | 4.0120 | 0.0431 | 0.9480 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.6972 | 0.0817 | -0.4029 | 3.4497 | 0.0818 | 0.9580 |
| | $B_1$ | $p_1$ | 0.8000 | 0.7990 | 0.0405 | -0.1188 | 2.3437 | 0.0405 | 0.9260 |
| | $B_2$ | $p_2$ | 0.9000 | 0.9003 | 0.0297 | 0.0356 | 1.0767 | 0.0297 | 0.9350 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6492 | 0.0482 | -0.1262 | 1.7024 | 0.0481 | 0.9420 |
| | | $t_2$ | 0.8000 | 0.7996 | 0.0408 | -0.0475 | 0.9304 | 0.0408 | 0.9360 |
| | | $t_3$ | 0.9000 | 0.8996 | 0.0295 | -0.0411 | 1.2551 | 0.0295 | 0.9480 |
| | $O_2$ | $t_1$ | 0.5000 | 0.4998 | 0.0497 | -0.0500 | 0.5027 | 0.0497 | 0.9400 |
| | | $t_2$ | 0.8000 | 0.7986 | 0.0400 | -0.1775 | 3.5529 | 0.0400 | 0.9290 |
| | | $t_3$ | 0.9000 | 0.8986 | 0.0299 | -0.1556 | 4.6825 | 0.0299 | 0.9430 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0891 | 0.1041 | 12.9317 | 9.7983 | 0.1045 | 0.9410 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1541 | 0.0639 | -19.1716 | 57.1841 | 0.0736 | 0.9800 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0491 | 0.1140 | 0.7002 | 0.2998 | 0.1139 | 0.9210 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0659 | 0.1077 | 18.0311 | 9.3477 | 0.1082 | 0.9260 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0735 | 0.0930 | -6.2267 | 5.2530 | 0.0930 | 0.9660 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0865 | 0.1063 | -1.9815 | 1.6444 | 0.1063 | 0.9380 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1675 | 0.0720 | -13.7434 | 37.0844 | 0.0767 | 0.9850 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0331 | 0.1021 | 3.2181 | 1.0115 | 0.1020 | 0.9410 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1518 | 0.1054 | 4.6654 | 6.4218 | 0.1055 | 0.9330 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1472 | 0.1246 | -2.4195 | 2.9289 | 0.1246 | 0.8690 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1154 | 0.0879 | 0.4896 | 0.6399 | 0.0878 | 0.9720 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0636 | 0.1031 | 2.8783 | 1.7260 | 0.1030 | 0.9340 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0335 | 0.1050 | -10.2231 | 3.6282 | 0.1050 | 0.9450 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1507 | 0.0750 | 0.3500 | 0.7012 | 0.0749 | 0.9910 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1154 | 0.1067 | -5.3959 | 6.1674 | 0.1068 | 0.9210 |
| 10000 | $P_1$ | $\lambda_1$ | 0.2000 | 0.2002 | 0.0045 | 0.1037 | 4.6156 | 0.0045 | 0.9500 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.6997 | 0.0082 | -0.0434 | 3.6874 | 0.0082 | 0.9620 |
| | $B_1$ | $p_1$ | 0.8000 | 0.8000 | 0.0040 | -0.0017 | 0.3442 | 0.0040 | 0.9490 |
| | $B_2$ | $p_2$ | 0.9000 | 0.8999 | 0.0030 | -0.0160 | 4.8281 | 0.0030 | 0.9570 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6500 | 0.0048 | 0.0035 | 0.4754 | 0.0048 | 0.9520 |
| | | $t_2$ | 0.8000 | 0.8002 | 0.0040 | 0.0234 | 4.6355 | 0.0040 | 0.9540 |
| | | $t_3$ | 0.9000 | 0.9001 | 0.0031 | 0.0107 | 3.1142 | 0.0031 | 0.9430 |
| | $O_2$ | $t_1$ | 0.5000 | 0.5003 | 0.0051 | 0.0523 | 5.0880 | 0.0051 | 0.9410 |
| | | $t_2$ | 0.8000 | 0.8002 | 0.0039 | 0.0272 | 5.5610 | 0.0039 | 0.9520 |
| | | $t_3$ | 0.9000 | 0.9001 | 0.0029 | 0.0107 | 3.2758 | 0.0029 | 0.9490 |
| | Correlation | $\rho\ P_1P_2$ | 0.0789 | 0.0884 | 0.0103 | 12.0554 | 91.9103 | 0.0140 | 0.8240 |
| | | $\rho\ P_1B_1$ | 0.1907 | 0.1546 | 0.0061 | -18.9012 | 590.3726 | 0.0366 | 0.0030 |
| | | $\rho\ P_1B_2$ | -0.0488 | -0.0531 | 0.0115 | 8.7482 | 37.2820 | 0.0122 | 0.8810 |
| | | $\rho\ P_1O_1$ | 0.0559 | 0.0585 | 0.0105 | 4.6811 | 24.8621 | 0.0108 | 0.9330 |
| | | $\rho\ P_1O_2$ | -0.0784 | -0.0752 | 0.0094 | -4.0710 | 34.0389 | 0.0099 | 0.9530 |
| | | $\rho\ P_2B_1$ | -0.0882 | -0.0913 | 0.0107 | 3.5480 | 29.3393 | 0.0111 | 0.9180 |
| | | $\rho\ P_2B_2$ | 0.1942 | 0.1666 | 0.0069 | -14.2143 | 400.1801 | 0.0285 | 0.1110 |
| | | $\rho\ P_2O_1$ | 0.0321 | 0.0327 | 0.0103 | 1.8198 | 5.6625 | 0.0103 | 0.9430 |
| | | $\rho\ P_2O_2$ | 0.1450 | 0.1497 | 0.0102 | 3.1951 | 45.5079 | 0.0112 | 0.9210 |
| | | $\rho\ B_1B_2$ | 0.1509 | 0.1508 | 0.0122 | -0.0478 | 0.5889 | 0.0122 | 0.8800 |
| | | $\rho\ B_1O_1$ | 0.1148 | 0.1147 | 0.0088 | -0.0692 | 0.9007 | 0.0088 | 0.9700 |
| | | $\rho\ B_1O_2$ | -0.0618 | -0.0617 | 0.0108 | -0.2162 | 1.2391 | 0.0108 | 0.9280 |
| | | $\rho\ B_2O_1$ | -0.0373 | -0.0371 | 0.0105 | -0.4077 | 1.4502 | 0.0105 | 0.9460 |
| | | $\rho\ B_2O_2$ | 0.1502 | 0.1503 | 0.0072 | 0.0886 | 1.8397 | 0.0072 | 0.9910 |
| | | $\rho\ O_1O_2$ | 0.1219 | 0.1221 | 0.0104 | 0.1381 | 1.6135 | 0.0104 | 0.9370 |

**Table 10.** Scenario 15/16, small Poisson rate and imbalanced binary/ordinal distribution

| Sample Size | Variable | Parameter | TV | AE | SD | RB | SB | RMSE | CR |
|---|---|---|---|---|---|---|---|---|---|
| 100 | $P_1$ | $\lambda_1$ | 0.2000 | 0.2011 | 0.0470 | 0.5550 | 2.3602 | 0.0470 | 0.9220 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.7020 | 0.0825 | 0.2814 | 2.3883 | 0.0825 | 0.9590 |
| | $B_1$ | $p_1$ | 0.8000 | 0.7975 | 0.0396 | -0.3150 | 6.3692 | 0.0396 | 0.9370 |
| | $B_2$ | $p_2$ | 0.9000 | 0.8997 | 0.0304 | -0.0300 | 0.8874 | 0.0304 | 0.9340 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6504 | 0.0468 | 0.0662 | 0.9197 | 0.0467 | 0.9480 |
| | | $t_2$ | 0.8000 | 0.8006 | 0.0400 | 0.0712 | 1.4242 | 0.0400 | 0.9370 |
| | | $t_3$ | 0.9000 | 0.8993 | 0.0292 | -0.0778 | 2.4002 | 0.0292 | 0.9520 |
| | $O_2$ | $t_1$ | 0.5000 | 0.4994 | 0.0491 | -0.1280 | 1.3045 | 0.0490 | 0.9310 |
| | | $t_2$ | 0.8000 | 0.7993 | 0.0404 | -0.0825 | 1.6354 | 0.0403 | 0.9320 |
| | | $t_3$ | 0.9000 | 0.8997 | 0.0301 | -0.0378 | 1.1283 | 0.0301 | 0.9310 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0251 | 0.1063 | 45.4862 | 7.3680 | 0.1065 | 0.9360 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0628 | 0.1052 | 1.5171 | 0.8924 | 0.1051 | 0.9440 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0479 | 0.0868 | 4.1982 | 2.2260 | 0.0867 | 0.9720 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0923 | 0.0884 | -4.0359 | 4.3894 | 0.0885 | 0.9720 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.0993 | 0.0894 | -5.4672 | 6.4216 | 0.0895 | 0.9690 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0471 | 0.1057 | 16.2741 | 6.2461 | 0.1058 | 0.9460 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0756 | 0.1071 | 12.4024 | 7.7895 | 0.1073 | 0.9300 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1627 | 0.1087 | 5.0129 | 7.1481 | 0.1089 | 0.9130 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1524 | 0.1044 | 3.7728 | 5.3068 | 0.1045 | 0.9240 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0870 | 0.1148 | -2.9310 | 2.2900 | 0.1148 | 0.9130 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0176 | 0.0969 | -15.7111 | 3.3865 | 0.0969 | 0.9530 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1067 | 0.1062 | -1.2228 | 1.2431 | 0.1062 | 0.9290 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0662 | 0.0895 | -2.7920 | 2.1246 | 0.0894 | 0.9770 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1452 | 0.0713 | -1.6152 | 3.3456 | 0.0713 | 0.9920 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0492 | 0.0962 | -6.4908 | 3.5478 | 0.0962 | 0.9580 |
| 10000 | $P_1$ | $\lambda_1$ | 0.2000 | 0.1999 | 0.0045 | -0.0501 | 2.2245 | 0.0045 | 0.9490 |
| | $P_2$ | $\lambda_2$ | 0.7000 | 0.7001 | 0.0083 | 0.0165 | 1.3951 | 0.0083 | 0.9500 |
| | $B_1$ | $p_1$ | 0.8000 | 0.8002 | 0.0040 | 0.0300 | 5.9610 | 0.0040 | 0.9490 |
| | $B_2$ | $p_2$ | 0.9000 | 0.9001 | 0.0030 | 0.0113 | 3.3572 | 0.0030 | 0.9410 |
| | $O_1$ | $t_1$ | 0.6500 | 0.6499 | 0.0050 | -0.0095 | 1.2445 | 0.0050 | 0.9410 |
| | | $t_2$ | 0.8000 | 0.7999 | 0.0041 | -0.0092 | 1.8108 | 0.0041 | 0.9620 |
| | | $t_3$ | 0.9000 | 0.8999 | 0.0030 | -0.0129 | 3.8432 | 0.0030 | 0.9530 |
| | $O_2$ | $t_1$ | 0.5000 | 0.4999 | 0.0049 | -0.0242 | 2.4439 | 0.0049 | 0.9560 |
| | | $t_2$ | 0.8000 | 0.7998 | 0.0039 | -0.0191 | 3.8714 | 0.0039 | 0.9400 |
| | | $t_3$ | 0.9000 | 0.8999 | 0.0030 | -0.0081 | 2.4519 | 0.0030 | 0.9550 |
| | Correlation | $\rho\ P_1P_2$ | 0.0172 | 0.0193 | 0.0097 | 12.1725 | 21.6611 | 0.0099 | 0.9520 |
| | | $\rho\ P_1B_1$ | -0.0619 | -0.0650 | 0.0106 | 5.1414 | 29.8670 | 0.0111 | 0.9180 |
| | | $\rho\ P_1B_2$ | 0.0460 | 0.0426 | 0.0088 | -7.3484 | 38.4502 | 0.0094 | 0.9590 |
| | | $\rho\ P_1O_1$ | -0.0962 | -0.0899 | 0.0085 | -6.5609 | 74.2276 | 0.0106 | 0.9370 |
| | | $\rho\ P_1O_2$ | -0.1050 | -0.0993 | 0.0088 | -5.3898 | 64.1506 | 0.0105 | 0.9420 |
| | | $\rho\ P_2B_1$ | -0.0405 | -0.0411 | 0.0102 | 1.2961 | 5.1523 | 0.0102 | 0.9480 |
| | | $\rho\ P_2B_2$ | -0.0672 | -0.0708 | 0.0105 | 5.3360 | 34.0149 | 0.0111 | 0.9180 |
| | | $\rho\ P_2O_1$ | 0.1549 | 0.1623 | 0.0106 | 4.7332 | 68.8938 | 0.0129 | 0.8580 |
| | | $\rho\ P_2O_2$ | 0.1468 | 0.1508 | 0.0107 | 2.7025 | 37.0181 | 0.0114 | 0.9070 |
| | | $\rho\ B_1B_2$ | 0.0897 | 0.0898 | 0.0113 | 0.1579 | 1.2525 | 0.0113 | 0.9170 |
| | | $\rho\ B_1O_1$ | 0.0209 | 0.0212 | 0.0098 | 1.3036 | 2.7856 | 0.0098 | 0.9490 |
| | | $\rho\ B_1O_2$ | -0.1080 | -0.1080 | 0.0103 | -0.0175 | 0.1832 | 0.0103 | 0.9470 |
| | | $\rho\ B_2O_1$ | 0.0681 | 0.0685 | 0.0082 | 0.5891 | 4.8953 | 0.0082 | 0.9790 |
| | | $\rho\ B_2O_2$ | 0.1476 | 0.1476 | 0.0075 | -0.0230 | 0.4497 | 0.0075 | 0.9900 |
| | | $\rho\ O_1O_2$ | -0.0526 | -0.0521 | 0.0099 | -1.0038 | 5.3125 | 0.0099 | 0.9480 |