Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

A SIMULATION-BASED COMPARISON OF MAXIMUM ENTROPY AND COPULA METHODS FOR CAPTURING NON-LINEAR PROBABILITY DEPENDENCE

Ehsan Salimi

Daniel J. Epstein Department of Industrial and Systems Engineering

University of Southern California
3710 McClintock Ave, RTH 322
Los Angeles, CA 90089, USA

Ali E. Abbas

Daniel J. Epstein Department of Industrial and Systems Engineering and Price School of Public Policy
University of Southern California
36500 McClintock Ave, OHE 310
Los Angeles, CA 90089, USA

ABSTRACT

The modeling of complex service systems entails capturing many sub-components of the system, and the dependencies that exist among them in the form of a joint probability distribution. Two common methods for constructing joint probability distributions from experts using partial information include maximum entropy methods and copula methods. In this paper we explore the performance of these methods in capturing the dependence between random variables using correlation coefficients and lower-order pairwise assessments. We focus on the case of discrete random variables, and compare the performance of these methods using a Monte Carlo simulation when the variables exhibit both independence and non-linear dependence structures. We show that the maximum entropy method with correlation coefficients and the Gaussian copula method perform similarly, while the maximum entropy method with pairwise assessments performs better particularly when the variables exhibit non-linear dependence.

1 INTRODUCTION

Decision making in today's complex and integrated environment is an ever-challenging task. Constructing an appropriate joint probability distribution for a complex system, is an important part of this decision-making task (Howard and Abbas 2015). Bayes' rule offers a systematic way to construct the joint probabilities. In this method, the joint probabilities are computed based on the assessed marginal distributions and conditional probabilities. The exponential growth of number of assessments with the number of variables, makes the use of this method difficult for complex systems (Zacks 1971, Ferguson 1973). To overcome this issue, many methods have been introduced to construct joint distributions using lower order assessments. The maximum entropy methods and the copula methods have gained recent popularity because they require fewer assessments than the typical marginal-conditional approach.

The maximum entropy methods (Jaynes 1957) select the probability distribution that has maximum entropy subject to the available information. The information provided can be of many forms (Abbas 2003, Abbas 2005, Abbas 2006, Abbas and Aczél 2010).

This paper discusses the accuracy of approximations of joint probability distributions when capturing the linear and non-linear dependence between the variables. It compares the performance of the maximum entropy method and the Gaussian copula method using Monte Carlo simulation. We show that extra effort

is needed when constructing joint probability distributions in the presence of non-linear dependence structures, and discuss the accuracy of the approximations.

The remainder of this paper is organized as follows. Section 2 reviews maximum entropy and copula methods for constructing joint distributions. Sections 3, 4 and 5 discuss the simulation steps for comparing the performance of each method using uniform sampling, independent sampling and non-linear dependence sampling. Section 6 compares of the accuracy of the methods using conditional probability distributions. Section 7 summarizes the results and main conclusions.

2 REVIEW OF MAXIMUM ENTROPY AND COPULA METHODS

2.1 The Maximum Entropy Method

For a discrete random variable X, the entropy (Shannon 1949) is defined as:

$$H(x) = -\sum_{i=1}^{n} p(x_i) \log p(x_i)$$

where $p(x_i)$ represents the probability of outcome x_i, for $i = 1, \dots, n$.
The Kullback-Leibler divergence (Kullback and Leibler 1951) determines the amount by which the entropy increases when a reference distribution Q is estimated with another distribution P:

$$K(P{:}\,Q) = \sum_{i=1}^{n} p(x_i) \log \frac{p(x_i)}{q(x_i)}$$

where $p(x_i)$ and $q(x_i)$ represent the probabilities for outcome x_i, for $i = 1, \dots, n$, for distributions P and Q respectively. The Kullback-Leibler measure is non-negative and is equal to zero if and only if the distributions P and Q are identical.

2.1.1 The Maximum Entropy Formulation with Correlation Coefficients

This model assumes that the marginal distributions and the pairwise correlation coefficients are available. The objective function is Shannon's entropy for the joint probability distribution. The marginal distributions and the pairwise correlation coefficients form the constraints. For three discrete random variables, X, Y and Z, the assessed marginals are $P(X), P(Y)$ and $P(Z)$. The ρ_{xy}, ρ_{xz} and ρ_{yz} represent the pairwise correlations between the pairs, (X, Y), (X, Z) and (Y, Z) respectively.

$$p^* = argmax - \sum\nolimits_{x,y,z} p(x,y,z) \log\big(p(x,y,z)\big)$$

Subject to

$$\sum\nolimits_{x,y,z} p(x,y,z) = 1$$
$$\sum\nolimits_{y,z} p(x,y,z) = P(X)$$
$$\sum\nolimits_{x,z} p(x,y,z) = P(Y)$$
$$\sum\nolimits_{x,y} p(x,y,z) = P(Z)$$
$$Corr(X,Y) = \rho_{xy}, \;\; Corr(X,Z) = \rho_{xz}, \;\; Corr(Y,Z) = \rho_{yz}$$

where $p(x,y,z)$ represents the joint probability for (X,Y,Z) and $Corr(\cdot,\cdot)$ is the Pearson's product-moment correlation coefficient between the variables. For simplicity, we refer to this model as ME-CC.

2.1.2 Maximum Entropy Formulation with Pairwise Probability Assessments

The maximum entropy formulation with the pairwise assessments (ME-PA) is as follows

$$p^* = argmax - \sum_{x,y,z} p(x,y,z)\log\big(p(x,y,z)\big)$$

Subject to

$$\sum_{x,y,z} p(x,y,z) = 1$$
$$\sum_{z} p(x,y,z) = P(x,y) \;\; for\; \forall\, x,y$$
$$\sum_{y} p(x,y,z) = P(x,z) \;\; for\; \forall\, x,y$$
$$\sum_{x} p(x,y,z) = P(y,z) \;\; for\; \forall\, x,y$$

2.2. Copula Methods

A multivariate function, C, expresses the joint probability distribution of n random variables, $F(x_1, x_2, \dots, x_n)$, in terms of their marginal distributions. It is defined as:

$$F(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)).$$

where $F_i(x_i)$ represents the marginal distribution for the i^{th} random variable. In this paper, we focus on the Gaussian copula method for constructing the joint probability distribution.

2.1.3 Gaussian Copula Method

The Gaussian copula (GC), representing the joint probability distribution of $F(x_1, x_2, \dots, x_n)$, can be written as:

$$C\big(F_1(x_1), F_2(x_2), \dots, F_n(x_n)\big) = \Phi_R(\Phi^{-1}\big(F_1(x_1)\big), \Phi^{-1}\big(F_2(x_2)\big), \dots, \Phi^{-1}(F_n(x_n))).$$

where Φ_R represents the joint cumulative distribution function of a multivariate normal distribution with mean vector zero and the correlation matrix R and Φ^{-1} is the inverse cumulative distribution function of a standard normal distribution.

To construct the joint probability distribution using the Gaussian copula for discrete random variables, we follow (Wang and Dyer 2013). We assume that random variables X, Y and Z each has three outcomes $\{-1,0,1\}$ with associated probabilities of $\{p_{\cdot}(-1), p_{\cdot}(0), p_{\cdot}(1)\}$ where the "." represents one of the three random variables. The closed-form solution for joint probabilities is:

$$F(X=i,Y=j,Z=k)=\int_{lbx}^{upx}\int_{lby}^{upy}\int_{lbz}^{upz}\phi(z_1,z_2,z_3,\Sigma_z)\,dz_1dz_2dz_3.$$

where $i,j,k\in\{-1,0,1\}$ and z_1,z_2 and z_3 represent the standard normal variables with Σ_z as the covariance matrix for the Gaussian copula. It is important to note that the correlation between the original variables is r_{ij}. It should be transformed to the correlation for the Gaussian copula ρ_{ij} in order to be used. The correlation matching algorithm described in (Cario and Nelson 1997) can be used for this purpose. In this paper we use r_{ij} as a rough approximation of ρ_{ij}. The upper bounds and lower bounds for the integration can be calculated from the following expressions: (the same expression can be used for Y and Z.)

$$lbx=\begin{cases}\Phi^{-1}\big(p_x(-1)+p_x(0)\big) & if\ X=1\\ \Phi^{-1}\big(p_x(-1)\big) & if\ X=0\\ -\infty & if\ X=-1\end{cases}$$

and

$$ubx=\begin{cases}\infty & if\ X=1\\ \Phi^{-1}\big(p_x(-1)+p_x(0)\big) & if\ X=0\\ \Phi^{-1}\big(p_x(-1)\big) & if\ X=-1\end{cases}$$

3 UNIFORM SAMPLING ON THE SIMPLEX

Monte Carlo simulation will be implemented in order to compare the performance of three mentioned methods. We use more general sampling method, uniform sampling on simplex (Abbas 2006) in this section. Linear and non-linear dependence sampling will be discussed in Sections 4 and 5. We consider the case of three discrete random variables with outcomes of $\{-1,\ 0, 1\}$ for each random variable. We denote the random variables X, Y and Z. We generate a test distribution, based on uniform sampling method, at each iteration of the simulation and then reconstruct the joint probabilities using ME-PA, ME-CC and GC methods.

3.1 Uniform Sampling on the Simplex

For generating a test joint distribution for n random variables with 3 outcomes, 3^n joint probabilities are needed. We follow the uniform sampling on the simplex method as follows

1. Generate (3^n-1) independent samples from a uniform [0, 1] distribution, $x_1, x_2, \dots, x_{3^n-1}$
2. Sort the generated samples in decreasing order, $u_1, u_2, \dots, u_{3^n-1}$
3. Take the difference between each two consecutive elements of the ordered sample:
$$u_1-0, u_2-u_1, \dots, 1-\ u_{3^n-1}$$
The differences form the joint probability for n random variables with 3 outcomes.

3.2 Monte Carlo Simulation

The steps of the Monte Carlo simulation that we implement in this section can be summarized as follows:

1. Generate a joint probability distribution using the uniform sampling for 3 random variables with three outcomes $\{-1,0,1\}$ is generated.
2. Calculate the marginal distribution for each random variable.
3. Calculate the correlation coefficients between the random variables.
4. Reconstruct an approximate distribution using the calculated marginal distributions and pairwise correlation coefficients once using the ME-CC; once using the GC, and once using the ME-PA.
5. Calculate the Kullback-Leibler divergence between the reconstructed distribution and the original distribution.
6. Calculate the total variation between the reconstructed distribution and original distribution. The total variation between distributions P and Q is:

$$\sum_{i=1}^{n} |p_i - q_i|.$$

where p_i and q_i represent the probabilities associated with the i-th outcome of P and Q respectively.

The ***fmincon*** function in MATLAB was used to solve the optimization problem for the maximum entropy formulations. The simulation will run for 1000 times (1000 different samples).

3.3 Simulation Result for Uniform Sampling on the Simplex

The simulation results based on uniform sampling are shown in Table 1. Two main observations can be made from Table:

1. The results suggest that the ME-PA performs better than the ME-CC and the GC methods in case of the uniform sampling, with respect to the defined divergence measures. This should not come as a surprise because pairwise assessments include more information than merely the marginal distributions and correlation coefficients. Therefore, the approximated joint probability constructed by this method is closer to the test distribution than approximation constructed by the ME-CC and the GC methods.
2. Another observation from Table 1 is that both the GC and the ME-CC methods perform relatively close in the case of uniform sampling on the simplex. This observation can be explained with the fact that both methods use the information about the pairwise correlation coefficients and the marginal distributions. Hence, the approximations are close.

Table 1: Simulation results for uniform sampling.

Divergence Measure	GC vs. Simulated Distribution	ME-CC vs. simulated Distribution	ME-PA vs. Simulated Distribution
K-L Divergence	Average Deviation: 0.367 Standard Deviation: 0.140	Average Deviation: 0.355 Standard Deviation: 0.133	Average Divergence: 0.134 Standard Deviation: 0.080
Total Deviation	Average Deviation: 0.579 Standard Deviation: 0.100	Average Deviation: 0.548 Standard Deviation: 0.098	Average Deviation: 0.303 Standard Deviation: 0.092

3.4 A Note on the Sampling Size

The Monte Carlo simulation described in Section 3.2, samples (iterated) 1000 times (for each sampling method discussed later). Although this number seems small for this type of simulation, we observed that the both Kullback-Leibler divergence and the total deviation converge to a constant value after about 300

iterations. Figures 1 and 2 depict the average K-L divergence and the total deviation for the ME-CC method and ME-PA method respectively. Although, figures present the result for uniform sampling, but the same holds for the independent sampling and the non-linear sampling discussed in Section 4 and Section 5.

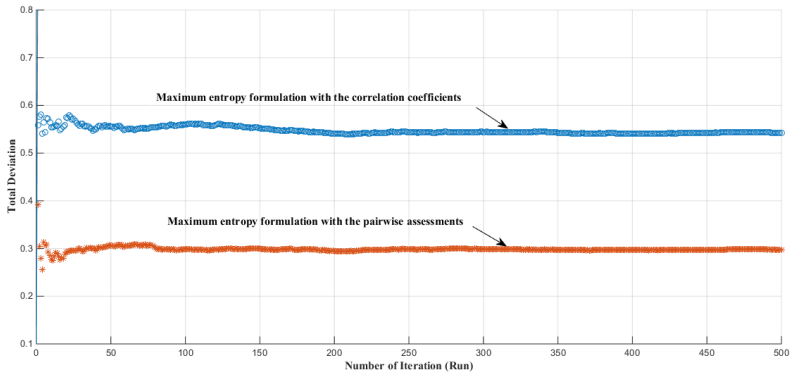

Figure 1: The average KL-measure vs. number of iterations.

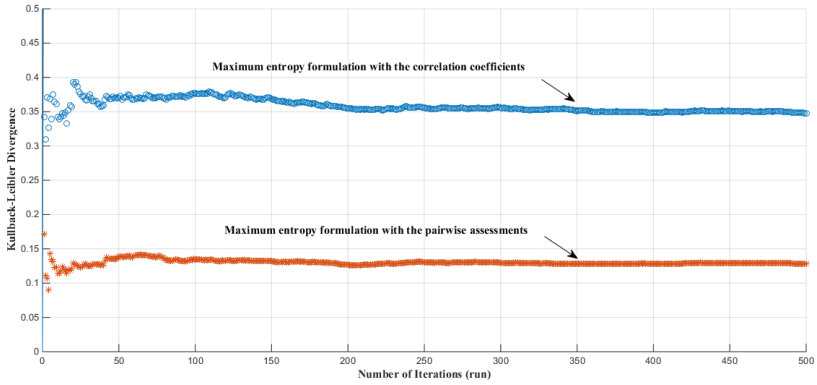

Figure 2: The average total deviation measure vs. number of iterations.

4 INDEPNDENCE STRUCTURE

As discussed in the previous section, uniform sampling on simplex, is a general sampling method that samples the whole space of distributions. In this section we investigate the case where there is no dependence between the variables and random variables, X, Y and Z are independent. We study this case as a simple baseline to demonstrate the accuracy of the calculations and illustrate the case where the reconstructed distribution is the same as the test distribution.

4.1 Independence Structure Sampling

Three random variables are said to be independent if:

$$P(X = i, Y = j, Z = k) = P(X = i)P(Y = j)P(Z = k).$$

In this section we focus on the cases when the three random variables, X, Y and Z, are independent in the test distribution. We generate joint probability distributions where the underlying random variables are independent. The algorithm for generating such a distribution can be summarized as following:

1. Generate 3-outcome probabilities for X using uniform sampling (refer to Section 3.1).
2. Generate 3-outcome probabilities for Y using uniform sampling.
3. Generate 3-outcome probabilities for Z, using uniform sampling.
4. For every outcome of the joint probability distribution:

$$P(X = i, Y = j, Z = k) = P(X = i)P(Y = j)P(Z = k), \forall i, j, k \in \{-1,0,1\}.$$

4.2 Simulation Results for Independence Structure Sampling

In order to examine the performance of three methods, the ME-PA, the ME-CC and the GC, we follow the Monte Carlo simulation described in Section 3.2: sample joint probability distributions are generated using the independence sampling and reconstructed using the ME-PA, the ME-CC and the GC methods. The divergence measures defined in Section 3.2 are then calculated between the simulated joint probability distributions and the reconstructed one for each method. Table 2 summarizes the simulation results for 1000 samples. Comparing the results in Table 5 reveals that all three methods, the ME-PA, the ME-CC and the GC perform very close to each other. This is expected as all the three methods are expected to recover the true underlying joint probability distribution. The magnitude of the divergence measures, suggests that all the approximations by three methods are very close to the test distribution. Also, the comparison between Table 1 and Table 2 suggests that all three methods perform much better in case of the independence sampling than case of the uniform sampling on the simplex. The previous argument holds for this observation as well: the joint probability distribution with the independence structure only needs marginal distributions to be recovered exactly. Hence, the three methods perform better in this case than any other dependence structure.

Table 2: Simulation results for independence structure sampling.

Divergence Measure	GC vs. simulated Distribution	ME-CC vs. simulated Distribution	ME-PA vs. Simulated Distribution
K-L Divergence	Average Deviation: 1.91×10^{-10} Standard Deviation: 2.37×10^{-9}	Average Deviation: 1.64×10^{-7} Standard Deviation: 1.08×10^{-6}	Average Divergence: 1.0×10^{-6} Standard Deviation: 9.4×10^{-6}
Total Deviation	Average Deviation: 2.97×10^{-10} Standard Deviation: 2.09×10^{-9}	Average Deviation: 1.89×10^{-5} Standard Deviation: 4.49×10^{-5}	Average Deviation: 4.62×10^{-5} Standard Deviation: 0.00039

5 NON-LINEAR DEPENDENCE STRUCTURE

In this section we focus on the performance of the approximation methods when the test distributions exhibit non-linear dependence between the variables.

5.1 Non-Linear Dependence Sampling

The non-linear dependence structure we aim to impose in this section has a simple structure. We sort the conditional probabilities, $P(Y|X)$ and $P(Z|X, Y)$ in decreasing and increasing orders alternatively. This

way, we make sure that the random variables, do not exhibit any particular linear dependence (positive or negative correlation) between them. Figure 3, is an example of a non-linear dependence structure between 3 random variables, shown on a probability tree. The probability distributions which satisfy the increasing order, are distinguished by dashed circle around them. The rest, follow the decreasing order. To clarify what we mean by decreasing order, we make sure that as the value of the outcome increases for one random variable, its conditional likelihood decreases. The increasing order is the reverse order relation between the variables. This structure, will guarantee that throughout the simulation, that no pair of two random variables have any particular linear dependence (see Figure 3).

The following steps describe the algorithm for generating the joint probability distribution with non-linear dependence structure (resembling the one shown in Figure 3):

1. Generate 3 outcome probabilities for X using uniform sampling.
2. Sort the generated values in decreasing order: $P(X=1) < P(X=0) < P(X=-1)$
3. Generate 3 outcome probabilities for $P(Y|X)$ using uniform sampling for each outcome of X; We will need to generate 9 different outcomes for this step
4. For three outcomes of X, $\{-1,0,1\}$, we sort the generated values of $P(Y|X)$ in Step 3 in decreasing, increasing and decreasing order alternatively.
5. Generate 3 outcome probabilities for $P(Z|X,Y)$ using uniform sampling for each pair of $(X=i, Y=j)$ for $, j=-1,0,1$; The total number of outcomes needed to be generated will be 27 outcomes.
6. Sort the generated values of $P(Z|X,Y)$ in increasing order for the pairs $(X=i, Y=0)$, $i=-1,0,1$, and the rest in decreasing order. For example the following relations hold (refer to Figure 3):

$$P(Z=-1|X=-1,Y=-1) > P(Z=0|X=-1,Y=-1) > P(Z=1|X=-1,Y=-1),$$
$$P(Z=-1|X=-1,Y=0) < P(Z=0|X=-1,Y=0) < P(Z=1|X=-1,Y=0).$$

5.2 Simulation Result for Non-Linear Dependence Sampling

We use the Monte Carlo simulation, described in Section 3.2 to evaluate the performance of the ME-PA, the ME-CC and the GC in approximating the test joint probability distribution. Same as previous simulations, it samples (iterates) 1000 times. Results for this simulation are presented in Table 3. The results are similar to previous simulations:

1. The ME-PA method performs better than both the ME-CC and the GC methods in the case of non-linear dependence structure sampling. This is an expected result as this method is using more information in order to reconstruct the joint probability distribution than other two methods, the ME-CC and the GC method.
2. Also, we notice that ME-CC and GC methods perform close in this case as well. This is explained with the fact that the pairwise correlation coefficients and the marginal distributions are used in both methods in order to approximate the test distribution.
3. Comparing the results of Table 1, 2 and 3 shows that performance of all three methods is the best with the independence sampling. This is due to the fact that in this method, all three methods are expected to reconstruct the true underlying joint probability distribution with respect to the available information.

Table 3: Simulation results for non-linear sampling.

Divergence Measure	GC vs. Simulated Distribution	ME-CC vs. simulated Distribution	ME-PA vs. Simulated Distribution
K-L Divergence	Average Deviation: 0.531 Standard Deviation: 0.227	Average Deviation: 0.522 Standard Deviation: 0.229	Average Divergence: 0.015 Standard Deviation: 0.013
Total Deviation	Average Deviation: 0.642 Standard Deviation: 0.143	Average Deviation: 0.627 Standard Deviation: 0.147	Average Deviation: 0.073 Standard Deviation: 0.038

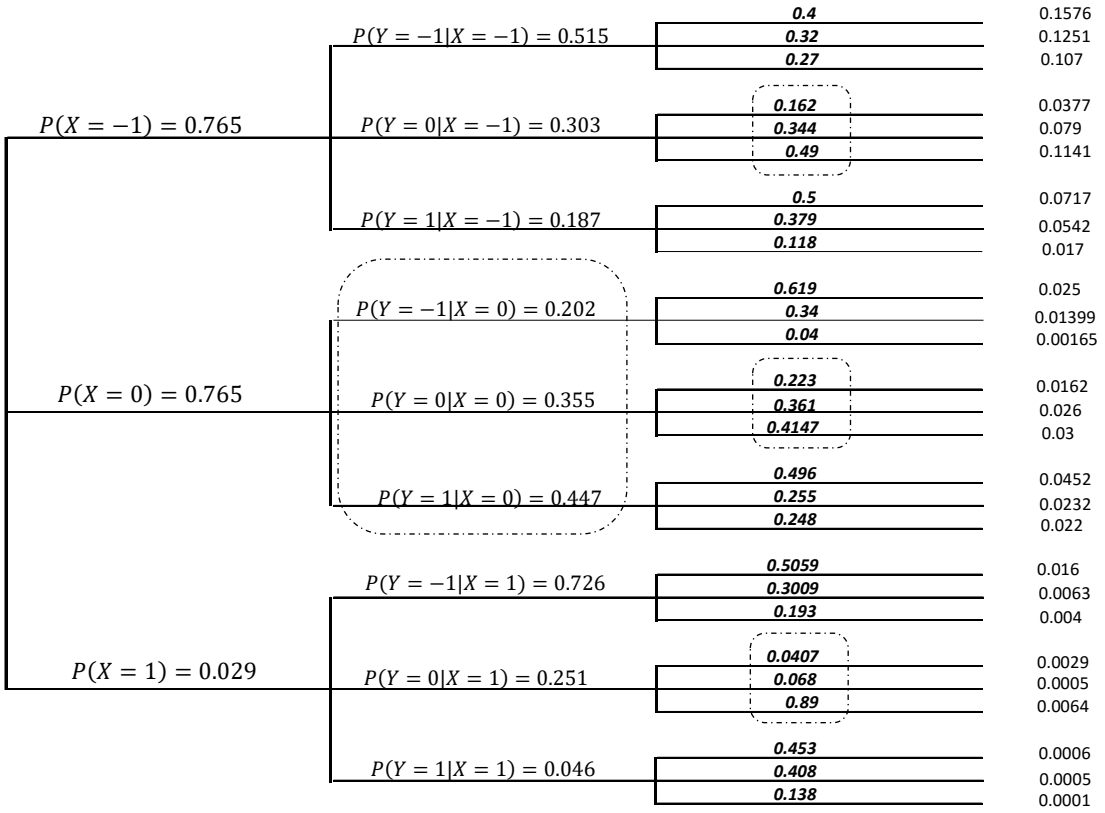

Figure 3: Probability tree with non-linear dependence structure.

6 ACCURACY OF THE ESTIMATED CONDITIONAL PROBABILITIES: $P(Z|X,Y)$

The comparison made in the previous sections, considered the divergence between the final values of the approximated and simulated joint probabilities. The objective of this paper is to compare the performance of methods on their ability to capture the underlying dependence structure, conditional probabilities in specific. In this section, we take a closer look at the conditional probabilities, specifically $P(Z|X,Y)$. We calculated the Kullback-Leibler divergence and the total deviation of $P(Z|X,Y)$ for the test distribution and the reconstructed ones.

6.1 Simulation Results for Accuracy of $P(Z|X,Y)$

We examine the divergence of conditional probabilities, between the approximated and simulated distribution, using Monte Carlo simulation explained in Section 3.2. We use three sampling methods described before in this paper. Then, for each sampling method, we reconstruct the joint probabilities with three mentioned method: the ME-PA, the ME-CC and the GC. We then calculate the conditional probabilities of $P(Z|X,Y)$ for each approximation and compare them with the simulated ones. It is important to note that $P(Z|X,Y)$ comprised of 9 different probability distributions as there are 9 possible $(X=i, Y=j)$ pairs for $i,j=-1,0,1$. Therefore, reported values are much larger than previous reported values. Tables 4 and 5 summarize the simulation results for uniform and non-linear dependence sampling respectively.

The results reflect our expectation: the ME-PA performs better in reconstructing the conditional probabilities than the ME-CC and the GC in all the sampling methods. The divergence values are smaller for the ME-PA. It again shows the importance of providing more information, in the form of pairwise probabilities assessment, in approximating the joint probability distribution. Also the same observation can be seen here as in the previous sections: the GC method and the ME-CC method have similar performance with respect to divergence measures. We dropped the result for the case of independence sampling as all three methods are capturing the true conditional probabilities in this case.

Table 4: Deviation measures for conditional probabilities, $P(Z|X,Y)$, using uniform sampling.

Divergence Measure	GC vs. Simulated Distribution	ME-CC vs. simulated Distribution	ME-PA vs. Simulated Distribution
K-L Divergence	Average Deviation: 3.112 Standard Deviation: 1.175	Average Deviation: 2.887 Standard Deviation: 1.107	Average Divergence: 1.661 Standard Deviation: 0.923
Total Deviation	Average Deviation: 4.8565 Standard Deviation: 0.8522	Average Deviation: 4.632 Standard Deviation: 0.813	Average Deviation: 3.268 Standard Deviation: 0.869

Table 5: Deviation measures for conditional probabilities, $P(Z|X,Y)$, using non-linear sampling.

Divergence Measure	GC vs. Simulated Distribution	ME-CC vs. simulated Distribution	ME-PA vs. Simulated Distribution
K-L Divergence	Average Deviation: 3.58 Standard Deviation: 1.31	Average Deviation: 3.70 Standard Deviation: 1.36	Average Divergence: 0.67 Standard Deviation: 0.45
Total Deviation	Average Deviation: 5.03 Standard Deviation: 0.65	Average Deviation: 5.06 Standard Deviation: 0.697	Average Deviation: 1.72 Standard Deviation: 0.524

7 CONCLUSIONS

In this paper, we studied the performance of three common methods for reconstructing the joint probability distribution using partial information: the maximum entropy formulation with pairwise probability assessments (ME-PA), the Gaussian copula method (GC), and the maximum entropy formulation with pairwise correlation coefficients and marginal distributions (ME-CC) were the studied methods. We explored the approximations when different types of dependence structures were present. For the purpose of study, we assumed the case of three discrete random variables with three possible outcomes, {-1,0,1}.

As a recurring observation throughout our simulation, we noticed that both the ME-CC and the GC methods perform quite close to each other. One explanation can be that both method rely on the pairwise correlation coefficients and the marginal distributions in order to reconstruct the joint probability distributions. Although the dependence structure, imposed by the sampling methods, played an important role in the performance of each method, we demonstrated that ME-CC and the GC approximate the underlying joint probability distribution similarly in each sampling scenario.

Also, we observed that the ME-PA performs better than both the GC and the ME-CC methods in reconstructing the joint probability distributions, with regard to defined divergence measures, in both the uniform sampling on simplex and the non-linear dependence structure scenarios. The reason the ME-PA performed better, lies within the fact that this method utilizes more information to reconstruct the joint probability distribution than both the ME-CC and the GC methods. It uses the experts assessment of the pairwise probabilities, which provides more information than the pairwise correlation coefficients. Although, providing more information by experts means spending more time and resource, but the results of this paper suggest that it might be a worthwhile effort in some cases.

We also noticed that all three method perform very well with the case on independence structure than the uniform sampling and the non-linear dependence structure sampling. This was expected, as only information regarding the marginal distributions is needed to reconstruct the joint probability distribution when the underlying random variables are independent.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation awards CMMI 15-65168, CMMI 16-29752, and CMMI 16-44991 .

REFERENCES

Abbas, Ali E. 2003. "Entropy Methods for Univariate Distributions in Decision Analysis." *Bayesian Inference and Maximum Entropy Methods in Science and Engineering.* American Institute of Physics Conference Proceedings 659. 339-349.

Abbas, Ali E. 2006. "Entropy Methods for Joint Distributions in Decision Analysis." *IEEE Transactions on Engineering Management* 53 (1): 146-159.

Abbas, Ali E. 2005. "Maximum Entropy Distributions Between Upper and Lower bounds." *Bayesian Inference and Maximum Entropy Methods in Science and Engineering.* American Institute of Physics Conference Proceedings 803. 25-42.

Abbas, Ali E, and János Aczél. 2010. "The Role of Some Functional Equations in Decision Analysis." *Decision Analysis* 7 (2): 215-228.

Cario, M.C, and B.L Nelson. 1997. Modeling and Generating Random Vectors with Arbitrary Marginal Distributions and Correlation Matrix. Technical Report, Evanston, Illinois: Department of Industrial Engineering and Management Sciences, Northwestern University.

Clemen, R.T, and T Reilly. 1999. "Correlations and Copulas for Decision and Risk Analysis." *Management Science* 208-224.

David, H.A, and H.N Nagaraja. 1970. *Order Statistics.* Wiley Online Library.

Ferguson, Thomas S. 1973. "A Bayesian Analysis of Some Nonparametric Problems." *The Annals of Statistics* 209-230.

Howard, R.A, and A. E. Abbas. 2015. *Foundations of Decision Analysis.* Prentice Hall.

Jaynes, E.T. 1957. "Information Theory and Statistical Mechanics." *Physical Review* 620.

Kullback, S, and R.A Leibler. 1951. "On Information and Sufficiency." *The Annals of Mathematical Statistics* 79-86.

Nelsen, Roger B. 2013. *An Introduction to Copulas.* Springer Science & Business Media.

Shannon, C.E. 1949. "Communication Theory of Secrecy Systems." *Bell System Technical Journal* 656-715.

Sklar, M. 1959. *Fonctions de Repartition `a n Dimensions et Leurs Marges.* Universite´ Paris 8.

Wang, T, and J. S Dyer. 2012. "A Copulas-Based Approach to Modeling Dependence in Decision Trees." *Operations Research* 225-242.

Wang, Tianyang, James S. Dyer, and John C. Butler. 2015. "Modeling Correlated Discrete Uncertainties in Event Trees with Copulas." *Risk Analysis.*

Zacks, Shelemyahu. 1971. *The Theory of Statistical Inference.* New York: Wiley.

AUTHOR BIOGRAPHIES

EHSAN SALIMI is a Postdoctoral Research Associate in the Daniel J. Epstein Department of Industrial and Systems Engineering at University of Southern California. He holds a Ph.D. in Industrial Engineering from University of Illinois at Urbana-Champaign. His research interests include decision making under uncertainty, discrete-event/discrete-state (DEDS) systems and computational geometry. His email address is esalimi@usc.edu.

ALI E. ABBAS is Professor of Industrial and Systems Engineering and Public Policy at the University of Southern California. He received his Ph.D. in Management Science and Engineering from Stanford University, where he also received his MS degree in Electric Engineering and MS in Engineering Economic Systems and Operations Research. He has directed the National Center for Risk and Economic Analysis of Terrorism Events (CREATE) and currently directs the Center for Interdisciplinary Decisions and Ethics (DECIDE) at the University of Southern California. Professor Abbas is author and coauthor of multiple research papers and is coauthor of "Foundations of Decision Analysis" with Ronald Howard. He is a member of INFORMS and a senior member of the IEEE. Dr. Abbas serves as an Associate Editor for the INFORMS Operations Research and INFORMS Decision Analysis journals and as the Decision Analysis area Editor for IIE Transactions. His email address is aliabbas@ usc.edu