Proceedings of the 1990 Winter Simulation Conference
Osman Balci, Randall P. Sadowski, Richard E. Nance (eds.)

CHARACTERIZING DISTRIBUTIONS OF DISCRETE BIVARIATE RANDOM VARIABLES FOR SIMULATION AND EVALUATION OF SOLUTION METHODS

Bridget A. Moore
Jeffrey A. Peterson

American Airlines
Decision Technologies
P.O. Box 619616
DFW Airport, Texas 75261-9616

Charles H. Reilly

Department of Industrial
and Systems Engineering
The Ohio State University
Columbus, Ohio 43210-1271

ABSTRACT

Results are presented from empirical evaluations of the performance of solution procedures on random binary knapsack and weighted set covering problems in which correlation is induced between the objective function and constraint coefficients. It is concluded that the performance of the solution procedures degrades as the correlation induced among the test problem parameters is increased and that test problems with structured dependence should be used for many empirical evaluations of solution methods.

Since structured dependence in discrete optimization test problems is desirable, ways to characterize the joint distribution of a discrete bivariate random variable for any feasible correlation are needed. The problem of finding a characterization with the maximum value of the smallest probability for any possible point is formulated as a linear program. An efficient algorithm based on the Northwest Corner Rule and a simple probability reallocation scheme is presented and demonstrated on a small example.

1. INTRODUCTION

When a new solution procedure for some discrete optimizaton problem is developed, questions about the method's performance arise. In the case of an exact (or optimizing) procedure, these questions are most likely to deal with execution time or the number of iterations performed by the procedure, while in the case of heuristic methods, additional questions dealing with solution quality often arise. In some cases, analytical results that describe the worst-case behavior of the solution method can be provided. Another type of analytical result might describe the average performance. Worst-case results often apply to unusual problem instances, and average-case results are frequently based on restrictive assumptions. Consequently, empirical evaluations of solution methods are sometimes conducted on test problems that are assumed to be representative of the problems the procedure might be called upon to solve.

In most empirical evaluations of algorithms and heuristics, there are not enough real examples of problems to provide a satisfactory evaluation. As a result, many random test problems are generated and solved in an effort to provide an "adequate" evaluation. In most cases when the random test problems are generated, iid observations from some discrete uniform distribution are used to provide the objective function coefficients, with the constraint coefficients normally generated in the same manner. This standard approach may be inadequate for evaluating solution procedures because the assumptions made in constructing the test problems (e.g., independence, uniformly-distributed parameters, etc.) may be violated in real problems and because the test problems generated may not be sufficiently difficult to adequately challenge the solution methods.

In the next section, the effect of correlation among problem parameters in random binary knapsack and weighted set covering problems is investigated. Computational results indicate that strong positive correlation between objective function and constraint coefficients in these problems substantially degrades the performance of implicit enumeration routines and heuristics. This implies that more rigorous empirical evaluations of solution procedures could be conducted if strong positive correlation were induced among the parameters in test problems, instead of randomly generating the parameters independently.

In §3, it is shown that the problem of finding an "optimal" characterization of a discrete bivariate distribution with a specified correlation can be formulated as a bottleneck transportation problem with a side constraint (BTPSC). Here, a characterization of a discrete bivariate distribution is considered to be optimal if the smallest joint probability value is as large as possible, given the correlation specified. In §4, bivariate distributions that result from the solution of BTPSC are compared to bivariate distributions that result from probabilistic mixing, or composition. An efficient algorithm for BTPSC is presented in §5. Suggestions for additional research in the areas of evaluation of solution procedures and characterizations of discrete multivariate distributions are discussed in §6. The results presented here and the suggestions for future research should be of interest to those who evaluate solution procedures and to those interested in simulation.

2. CORRELATION AND SOLUTION PROCEDURE PERFORMANCE

It is reasonable to expect that the coefficients in discrete optimization problems are not independent of one another. Two empirical studies of the performance of exact and heuristic procedures on random test problems in which dependence was induced among the test problem parameters are reported below. Some remarks about the computational results follow.

2.1 Binary Knapsack Problems

Consider a binary knapsack problem where the value of the items that are included in a knapsack is to be maximized, subject to a single restriction on weight:

Maximize

$$\sum_{j=1}^{n} v_j x_j$$

Subject to

$$\sum_{j=1}^{n} w_j x_j \leq W$$

$$x_j = 0 \text{ or } 1, \quad j = 1, 2, ..., n,$$

where:

$$x_j = \begin{cases} 1 & \text{if item } j \text{ is included in the knapsack;} \\ 0 & \text{otherwise;} \end{cases}$$

$v_j > 0$ is the value of item j; $w_j > 0$ is the weight of item j; W is the capacity of the knapsack (in weight); and n is the number of items considered for inclusion in the knapsack. It is easy to imagine a direct, although not perfectly direct, relationship between the value of an item and its weight.

Moore (1989) has studied the performance of an implicit enumeration (branch-and-bound) routine on binary knapsack problems with several different correlations between the objective function and constraint coefficients, including the uncorrelated and independent case. (This study was suggested by an exercise in Bratley, Fox, and Schrage (1983).) In this investigation, a set of ten, 20-variable knapsack problems was generated for each of the following correlations between objective function coefficients, v_j, and constraint coefficients, w_j: -0.99, -0.5, 0, +0.5, +0.99. The objective function coefficients were generated uniformly over the integers from 1 to 100, and the constraint coefficients were generated uniformly over the integers from 1 to 50. The right-hand side constants, W, were given by $\left\lceil \sum_{j=1}^{20} w_j/2 \right\rceil$. A probabilistic mixing method that appears in Schmeiser and Lal (1982) was used to generate the problem parameters. Common random numbers were synchronized across correlation values.

See Table 1 for a summary of the implicit enumeration iteration counts observed. These results suggest that the number of iterations required to solve a knapsack problem to optimality is an exponential function of the correlation of the problem parameters. Let η_i denote the number of implicit enumeration iterations and ρ_i denote the target correlation among the parameters in test problem i. Moore fit the following regression line to the data summarized in Table 1:

$$\ln \eta_i = 8.7 + 2.6\rho_i.$$

The coefficient of determination for this regression line is 0.9137.

This phenomenon is explainable to at least a limited extent. When the parameters of a knapsack problem are highly positively correlated, many variables are likely to have similar v_j/w_j ratios. Consequently, many items are likely to seem about equally attractive for inclusion in the knapsack. Conversely, when the problem parameters are highly negatively correlated, the items corresponding to those variables with relatively large v_j and relatively small w_j are the principal candidates for inclusion in the knapsack; the other items (variables) with small v_j and large w_j are almost certain to be excluded from the optimal knapsack contents. The fact that the number of iterations increases as the target correlation increases from near -1 to near $+1$ is not surprising. However, it was not anticipated that the number of iterations would be an exponential function of the target correlation.

2.2 Weighted Set Covering Problems

Another investigation has been carried out by Moore (1990) to evaluate the performance of two greedy heuristics and implicit enumeration on the weighted set covering problem. The weighted set covering problem has the following form:

Minimize

$$\sum_{j=1}^{n} c_j x_j$$

Subject to

$$\sum_{j=1}^{n} a_{ij} x_j \geq 1, \qquad i = 1, 2, ..., m$$

$$x_j = 0 \text{ or } 1, \quad j = 1, 2, ..., n,$$

where:

$$x_j = \begin{cases} 1 & \text{if set } j \text{ is included in the cover;} \\ 0 & \text{otherwise;} \end{cases}$$

$$a_{ij} = \begin{cases} 1 & \text{if } i \text{ is an element of set } j; \\ 0 & \text{otherwise;} \end{cases}$$

$c_j > 0$ is the cost of set j; and n is the number of sets considered for inclusion in the cover.

If one thinks of a set covering problem as pertaining to the

Table 1. Statistics on Iterations to Optimality for Knapsack Problems

Statistic	$\rho = -0.99$	$\rho = -0.50$	$\rho = 0$ (indep.)	$\rho = +0.5$	$\rho = +0.99$
Mean	985.00	1,675.40	5,332.20	27,601.20	117,400.40
Median	585	1,320	3,465	19,661	95,943
Maximum	2,595	3,437	16,111	56,869	259,789
Minimum	185	473	929	3,769	28,001
Std. Error	285.22	333.15	1,467.44	5,844.99	21,899.32

construction of warehouses, each with certain capacities that would enable them to serve specified retail establishments, then the constraints would guarantee that each retail outlet is served by at least one warehouse. The problem is to build the cheapest collection of warehouses that can serve all of the retail outlets. A warehouse that would be able to serve many retail establishments (i.e., one with a large column sum, $\sum_i a_{ij}$) would require much capacity and, therefore, be relatively costly to build. Again, it seems that standard approaches for evaluating solution procedures would not include enough realistic weighted set covering examples in which there is clearly dependence among the problem parameters.

All of the problems in this study had 50 variables and 20 constraints. The objective function coefficients were generated uniformly over the integers from 1 to 100. The distributions of the column sums (constraint coefficients) were chosen so that the target constraint-matrix densities were 0.1, 0.2, and 0.3. Both discrete uniform and binomial distributions were used for the column sums: U[1,3], U[1,7], U[1,11], Bin(20,0.1), Bin(20,0.2), and Bin(20,0.3). (Actually, zero-valued binomial variates were rejected and replaced with positive variates. As a result, the expected constraint-matrix densities were slightly higher than the target densities.) Let ρ_{min} and ρ_{max} be the minimum and maximum achievable correlations between the objective function coefficients and the column sums. Five test problems were randomly generated using synchronized common random numbers and probabilistic mixing for each combination of column-sum distribution and the following correlation values: ρ_{min}, $\rho_{min}/2$, 0, $\rho_{max}/2$, and ρ_{max}.

See Table 2 for a summary of the iteration counts observed when implicit enumeration is used to solve the random weighted set covering problems. The implicit enumeration program used for the set covering problems was not the same program that was used for the knapsack problems. Rather, it was specifically intended for set covering problems and, therefore, permitted the economical solution of problems with more variables. The implicit enumeration results obtained are consistent with those observed for the knapsack problem: the number of iterations tends to increase exponentially as the target correlation between the objective function coefficients and the sums of the binary constraint coefficients increases. However, it seems that the problems with target correlation $\rho_{max}/2$ were not as difficult to solve as those with independent parameters ($\rho = 0$).

The performance of two set covering heuristics was also investigated by Moore (1990). She used a greedy primal method that begins with the maximal cover, i.e., all $x_j = 1$, and selectively removes redundant elements from the cover until removing any additional elements would result in an infeasible solution. This method is referred to here as PRIMAL. She also used a dual procedure, an extension of the method in Chvátal (1979) and referred to here as DUAL, that begins with all $x_j = 0$, adds elements until a cover is formed, and finally calls PRIMAL to remove redundant elements.

See Table 3 for a summary of the performance of the two set covering heuristics. The results for the two set covering heuristics indicate that the same problems that require many implicit enumeration iterations are the most difficult problems for the heuristics as well: fewer optimal solutions and poorer approximate solutions are found when there is high positive correlation between the objective function coefficients and the constraint-matrix column sums. This is true for both heuristics, although DUAL seems to perform better than PRIMAL. As was observed with the implicit enumeration results, the performance of the heuristics is not a nondecreasing function of the target correlation. In fact, the problems with uncorrelated and independent parameters seem reasonably challenging, especially when the target constraint-matrix density is 0.1.

2.3 Discussion of Empirical Results

The results described for knapsack and weighted set covering problems suggest that empirical evaluations of solution procedures should include problems with structured dependence, at least when a reasonable argument for dependence can be made. The outcome will be that the evaluation of the solution procedures will be conducted on test problems more similar to those that may be encountered in practice. By evaluating the performance of solution procedures on problems with independently distributed parameters, the performance of the procedures being scrutinized may not be adequately evaluated.

Table 2. Average Iterations to Optimality for Weighted Set Covering Problems

Column-Sum Distribution	$\rho = \rho_{min}$	$\rho = \rho_{min}/2$	$\rho = 0$ (indep.)	$\rho = \rho_{max}/2$	$\rho = \rho_{max}$
U[1,3]	157.2	541.6	3,090.0	674.4	9,310.6
U[1,7]	63.6	80.8	280.2	212.8	5,873.8
U[1,11]	13.6	24.0	34.6	245.4	6,757.8
Bin(20;0.1)	113.8	284.2	1130.0	337.0	1,663.0
Bin(20;0.2)	89.8	54.4	590.2	607.8	4,103.8
Bin(20;0.3)	14.4	41.6	117.4	90.6	2,499.8

Table 3. Average Relative Error for Set Covering Heuristics (%)

Column-Sum Distribution	Method	$\rho = \rho_{min}$	$\rho = \rho_{min}/2$	$\rho = 0$ (indep.)	$\rho = \rho_{max}/2$	$\rho = \rho_{max}$
U[1,3]	PRIMAL	0.95	1.51	15.55	7.52	19.00
	DUAL	0.42	0.06	6.57	1.11	4.81
U[1,7]	PRIMAL	0.71	7.20	4.75	17.19	19.80
	DUAL	6.33	45.40	2.83	7.76	9.96
U[1,11]	PRIMAL	0	6.60	9.62	10.68	16.28
	DUAL	0.61	2.22	1.96	8.28	5.07
Bin(20,0.1)	PRIMAL	1.99	15.67	12.59	14.72	20.33
	DUAL	1.03	4.68	4.55	2.73	14.99
Bin(20,0.2)	PRIMAL	0	22.04	8.54	5.44	13.92
	DUAL	0	3.75	7.64	3.89	9.94
Bin(20,0.3)	PRIMAL	5.31	1.92	7.76	23.50	12.83
	DUAL	0	1.50	2.42	0	10.25

The results for the weighted set covering problems are quite interesting. It seems that the same problems can be used to test both algorithms and heurisitcs. Although this is fairly common practice, these results provide some justification for this convention. The pattern of performance degradation for the solution methods for the weighted set covering problem is not as clear as that for the implicit enumeration routine applied to the knapsack problem. Perhaps this is due to the fact that there were 20 constraints in the the weighted set covering problems and only one in the knapsack problems. The interaction of the set covering constraints, which could be difficult to measure, may be an important factor in solution method performance.

For the investigations described in this section, observations of correlated random variables were generated with probabilistic mixing (see, for example, Schmeiser and Lal (1982)). A disadvantage of the mixing procedure is that, for a specified correlation value, only one characterization of a bivariate random variable is possible. Furthermore, when one is interested in sampling uncorrelated random variables, one samples from independent random variables by default. In the next section, a new, optimization approach for the characterization of a discrete bivariate random variable is proposed. By changing the objective function and/or constraints in the optimization model, different characterizations of a discrete bivariate random variable can be found for a single correlation value. Furthermore, if the characterizations found with this approach were used to generate observations, uncorrelated random variables would not necessarily be independent. Additional discussion of alternative characterizations can be found in §4.

3. OPTIMAL CHARACTERIZATIONS OF BIVARIATE DISTRIBUTIONS

It is useful for evaluators of solution procedures to be able to randomly generate test problems with specified dependence structures. In most cases, there is an infinite number of ways to characterize a joint distribution of two discrete random variables with a particular correlation value. Here, the concern is with finding an optimal characterization of such a joint distribution. A characterization of a joint distribution is considered to be optimal if the smallest joint probability for any possible realization is maximized.

The problem of choosing a preferred characterization of a discrete multivariate random variable with a specified correlation structure can be formulated as a linear program (Peterson, 1990). Suppose one wishes to generate observations from a discrete bivariate random variable, (Y_1, Y_2), where Y_i is distributed over the n_i values y_{ij_i} according to the pmf $f_i(y_{ij_i})$, $i = 1,2$; $j_i = 1,2,...,n_i$, and $\text{Corr}(Y_1, Y_2) = \rho$. The decision variables in this linear program are: $x_{j_1,j_2} = \Pr(Y_1 = y_{1j_1}, Y_2 = y_{2j_2})$ and $\theta = \min_{j_1,j_2}\{x_{j_1,j_2}\}$. The complete formulation is:

Maximize

$$\theta \qquad (1)$$

Subject to

$$x_{j_1,j_2} - \theta \geq 0, \qquad \forall j_1, j_2 \qquad (2)$$

$$\sum_{j_2=1}^{n_2} x_{j_1,j_2} = f_1(y_{1j_1}), \qquad \forall j_1 \qquad (3)$$

$$\sum_{j_1=1}^{n_1} x_{j_1,j_2} = f_2(y_{2j_2}), \qquad \forall j_2 \qquad (4)$$

$$\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} y_{1j_1}y_{2j_2}x_{j_1,j_2} = \text{E}(Y_1Y_2) \qquad (5)$$

$$x_{j_1,j_2} \geq 0, \qquad \forall j_1, j_2, \qquad (6)$$

where $\text{E}(Y_1Y_2) = \rho(\text{Var}(Y_1)\text{Var}(Y_2))^{\frac{1}{2}} + \text{E}(Y_1)\text{E}(Y_2)$.

The objective function (1) maximizes the smallest probability value assigned to any realization (y_{1j_1}, y_{2j_2}) of (Y_1, Y_2). Constraint set (2) enforces the definition of θ. The constraint sets (3) and (4) insure that the bivariate probabilities conform to the marginal distributions. Constraint (5) enforces a specified correlation, ρ, between Y_1 and Y_2. Finally, since each decision variable represents a probability, constraint set (6) is included to guarantee that all of the probabilities are nonnegative. In this formulation, there are $n_1 + n_2 + 1$ structural constraints in n_1n_2 variables. The structural constraint sets (3) and (4) are the constraints for a transportation problem, one of the easiest types of linear programming problems to solve. The linear program above can be classified as a bottleneck transportation problem with a side constraint (BTPSC), where the side constraint is constraint (5). This approach to representing probability assignment problems as transportation problems was used by Roach and Wright (1977) to find optimal antithetic sampling plans.

Evans (1984) has shown that the Northwest Corner Rule (NWCR) can be used to find an optimal solution to the following transportation problem:

Minimize

$$\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} y_{1j_1} y_{2j_2} x_{j_1,j_2} \qquad (7)$$

Subject to (3), (4), and (6),

when $y_{11} \geq y_{12} \geq \cdots \geq y_{1n_1} \geq 0$ and $0 \leq y_{21} \leq y_{22} \leq \cdots \leq y_{2n_2}$. Peterson (1990) has shown that the nonnegativity restriction for the $y_{ij_i}s$ is not necessary for Evans' result.

Note the objective function (7) gives the minimum achievable value for $E(Y_1Y_2)$ when a valid probability assignment, i.e., one that satisfies (3), (4), and (6), is made to the ordered pairs (y_{1j_1}, y_{2j_2}). This is the minimum-correlation probability assignment as well. Denote the value of $E(Y_1, Y_2)$ associated with this solution as K_{min}.

Suppose that $y_{11} \leq y_{12} \leq \cdots \leq y_{1n_1}$ instead and the objective function (7) is maximized. In this case, NWCR can be used to find the probability assignment with the maximum achievable correlation between Y_1 and Y_2 (Peterson, 1990). The value of $E(Y_1, Y_2)$ associated with this solution is denoted by K_{max}. Note that the minimum- and maximum-correlation probability assignments for (Y_1, Y_2) can nearly be found by inspection.

A tight upper bound on the solution value to BTPSC for any correlation value can be found.

Lemma 1. (Peterson, 1990) For all feasible values of ρ,

$$\theta \leq \theta^* = \min\left\{\min_{j_1}\left\{\frac{f_1(y_{1j_1})}{n_2}\right\}, \min_{j_2}\left\{\frac{f_2(y_{2j_2})}{n_1}\right\}\right\}.$$

Proof: Since $\sum_{j_1=1}^{n_1} x_{j_1,j_2} = f_2(y_{2j_2}),\ \forall j_2$, and $\theta \leq x_{j_1,j_2}$, $\forall j_1, j_2$, $\theta \leq f_2(y_{2j_2})/n_1$, $\forall j_2$. Therefore, $\theta \leq \min_{j_2}\left\{\frac{f_2(y_{2j_2})}{n_1}\right\}$. A similar argument yields $\theta \leq \min_{j_1}\left\{\frac{f_1(y_{1j_1})}{n_2}\right\}$. When the bounds on θ are combined, the desired result is obtained. □

Let $x'_{j_1,j_2} = x_{j_1,j_2} - \theta$. NWCR can be used to solve the following transportation problem for the minimum-correlation probability assignment for (Y_1, Y_2) in which the smallest joint probability is θ:

Minimize

$$\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} y_{1j_1} y_{2j_2}(x'_{j_1,j_2} + \theta) \qquad (8)$$

Subject to

$$\sum_{j_2=1}^{n_2} x'_{j_1,j_2} = f_1(y_{1j_1}) - n_2\theta, \qquad \forall j_1$$

$$\sum_{j_1=1}^{n_1} x'_{j_1,j_2} = f_2(y_{2j_2}) - n_1\theta, \qquad \forall j_2$$

$$x'_{j_1,j_2} \geq 0, \qquad \forall j_1, j_2,$$

when $y_{11} \geq y_{12} \geq \cdots \geq y_{1n_1}$ and $y_{21} \leq y_{22} \leq \cdots \leq y_{2n_2}$. The value of $E(Y_1, Y_2)$ associated with this solution is denoted by K^*_{min}. If $y_{11} \leq y_{12} \leq \cdots \leq y_{1n_1}$ instead and the objective function (8) is maximized, then the maximum-correlation probability assignment in which the smallest joint probability is θ can be found using NWCR. Denote the value of $E(Y_1, Y_2)$ associated with this solution as K^*_{max}.

Note that $\theta = \theta^*$ for all $K \in [K^*_{min}, K^*_{max}]$.

Proposition 2. (Peterson, 1990) If $f_1(y_{1j_1}) = n_1^{-1}$, $\forall j_1$, and $f_2(y_{2j_2}) = n_2^{-1}$, $\forall j_2$, then $K^*_{min} = K^*_{max} = E(Y_1)E(Y_2)$.

Proof: From Lemma 1, $\theta^* = (n_1n_2)^{-1}$. If $\theta = (n_1n_2)^{-1}$, then $x_{j_1,j_2} = (n_1n_2)^{-1}$, $\forall j_1, j_2$. Since this is clearly the only solution to BTPSC for which $\theta = \theta^*$, it must be that $K^*_{min} = K^*_{max}$. Furthermore, this solution represents the probability assignment for the case when Y_1 and Y_2 are independent. Consequently, $K^*_{min} = K^*_{max} = E(Y_1)E(Y_2)$. □

The proof of Proposition 2 is different from the one which appears in Peterson (1990). This propositon illustrates that there are special cases of the results presented here for certain distributions, like the uniform distribution. Other results for distributions that satisfy certain symmetry properties can be found in Peterson (1990).

Consider the following example:

$$f_1(1) = 0.2,\ f_1(2) = 0.38,\ f_1(3) = 0.32,\ f_1(4) = 0.1$$

$$f_2(2) = 0.1,\ f_2(3) = 0.3,\ f_2(5) = 0.2,\ f_2(6) = 0.2,\ f_2(7) = 0.2.$$

For this example, $E(Y_1) = 2.32$, $Var(Y_1) = 0.8176$, $E(Y_2) = 4.7$, $Var(Y_2) = 3.01$, and $\theta^* = 0.02$. The NWCR solutions for $E(Y_1Y_2) = K_{min}$ ($\theta = 0$), K_{max} ($\theta = 0$), K^*_{min} ($\theta = \theta^*$), and K^*_{max} ($\theta = \theta^*$) are shown in Figures 1, 2, 3, and 4, respectively. For this example, $K_{min} = 9.40$, $K^*_{min} = 10.26$, $K^*_{max} = 11.48$, and $K_{max} = 12.30$. This means that the minimum and maximum achievable correlations for any joint distribution of Y_1 and Y_2 are -0.95 and +0.88, respectively, and the minimum and maximum achievable correlations when the probability of any possible realization is at least 0.02 are -0.41 and +0.36.

Y_1 \ Y_2	2	3	5	6	7	
4	0.10	0				0.10
3		0.30	0.02			0.32
2			0.18	0.20	0	0.38
1					0.20	0.20
	0.10	0.30	0.20	0.20	0.20	

Figure 1. NWCR Solution for $E(Y_1Y_2) = K_{min}$

4. COMPARISON TO PROBABILISTIC MIXING

The probabilistic mixture that appears in Schmeiser and Lal (1982) is:

$$(Y_1, Y_2) = \begin{cases} (1 - \rho/\rho_{max})(Y_1^I, Y_2^I) + \rho/\rho_{max}(Y_1^+, Y_2^+), & \text{if } \rho \geq 0; \\ (1 - \rho/\rho_{min})(Y_1^I, Y_2^I) + \rho/\rho_{min}(Y_1^-, Y_2^-), & \text{if } \rho < 0; \end{cases}$$

Y_1 \ Y_2	2	3	5	6	7	
1	0.10	0.10				0.20
2		0.20	0.18			0.38
3			0.02	0.20	0.10	0.32
4					0.10	0.10
	0.10	0.30	0.20	0.20	0.20	

Figure 2. NWCR Solution for $E(Y_1Y_2) = K_{max}$

Y_1 \ Y_2	2	3	5	6	7	
4	0					0
3	0.02	0.20				0.22
2		0.02	0.12	0.12	0.02	0.28
1					0.10	0.10
	0.02	0.22	0.12	0.12	0.12	

Figure 3. NWCR Solution for $E(Y_1Y_2) = K^{\cdot}_{min}$

Y_1 \ Y_2	2	3	5	6	7	
1	0.02	0.08				0.10
2		0.14	0.12	0.02		0.28
3				0.10	0.12	0.22
4					0	0
	0.02	0.22	0.12	0.12	0.12	

Figure 4. NWCR Solution for $E(Y_1Y_2) = K^{\cdot}_{max}$

where ρ_{max} (ρ_{min}) is the maximum (minimum) possible correlation between Y_1 and Y_2. Y_1^I and Y_2^I are independent, and Y_1^+ (Y_1^-) and Y_2^+ (Y_2^-) have the maximum possible positive (negative) correlation. Note that if $\rho = 0$, then Y_1 and Y_2 are independent.

It is interesting to note that solutions to BTPSC actually represent a new mixture of bivariate distributions that is an alternative to the probabilistic mixture above:

$$(Y_1, Y_2) = n_1 n_2 \theta (U_1, U_2) + (1 - n_1 n_2 \theta)(Z_1, Z_2),$$

where U_1 and U_2 are independently distributed uniform random variables with n_1 and n_2 possible values, respectively, and (Z_1, Z_2) is a bivariate distribution with at most $n_1 + n_2 - 1$ possible points. This observation may provide insight for the construction of an efficient composition-based variate generation procedure.

Recall the example that was introduced in §3. Figure 5 shows the joint pmf of Y_1 and Y_2 that is obtained using probabilistic mixing for $\rho = -0.64$. The joint pmf of Y_1 and Y_2 that results from the solution of BTPSC with $\rho = -0.64$ is shown in Figure 6. Generally speaking, the corresponding joint probability values are similar in magnitude in these two cases. This may be explained by the fact that as ρ is decreased toward ρ_{min}, the two approaches will give increasingly similar joint pmfs. When $\rho = \rho_{min}$, both approaches yield the same pmf, that of (Y_1^-, Y_2^-).

Y_1 \ Y_2	2	3	5	6	7	
1	0.0067	0.0199	0.0133	0.0133	0.1468	0.20
2	0.0126	0.0379	0.1454	0.1588	0.0253	0.38
3	0.0106	0.2322	0.0346	0.0213	0.0213	0.32
4	0.0701	0.0100	0.0067	0.0066	0.0066	0.10
	0.10	0.30	0.20	0.20	0.20	

Figure 5. Probabilistic Mixture of Y_1 and Y_2 ($\rho = -0.64$)

Greater differences in the joint pmfs that result with these two approaches can be seen if we consider the case where $\rho = 0$. Figure 7 shows the joint pmf for probabilistic mixing, that of (Y_1^I, Y_2^I), and Figure 8 shows the joint pmf that results from the solution of BTPSC. The probability values look more dissimilar in this case. The minimum probability value for the BTPSC pmf is twice as large as the minimum probability value for the other pmf.

Note that most characterizations of Y_1 and Y_2 have finite probabilities for every possible point. The only cases where a joint probability value is zero occur when $\rho = \rho_{max}$ or $\rho = \rho_{min}$.

5. SOLUTION PROCEDURE

Peterson (1990) has devised an algorithm that allows one to find the probability assignment for (Y_1, Y_2) that maximizes (1) for any feasible correlation. (It is not possible to achieve perfect

Y_1 \ Y_2	2	3	5	6	7	
1	0.012	0.012	0.012	0.012	0.152	0.20
2	0.012	0.016	0.164	0.164	0.024	0.38
3	0.024	0.260	0.012	0.012	0.012	0.32
4	0.052	0.012	0.012	0.012	0.012	0.10
	0.10	0.30	0.20	0.20	0.20	

Figure 6. BTPSC Mixture of Y_1 and Y_2 ($\rho = -0.64$)

Y_1 \ Y_2	2	3	5	6	7	
1	0.020	0.060	0.040	0.040	0.040	0.20
2	0.038	0.114	0.076	0.076	0.076	0.38
3	0.032	0.096	0.064	0.064	0.064	0.32
4	0.010	0.030	0.020	0.020	0.020	0.10
	0.10	0.30	0.20	0.20	0.20	

Figure 7. Joint p.m.f. - Y_1 and Y_2 Independent

Y_1 \ Y_2	2	3	5	6	7	
1	0.0400	0.0755	0.0200	0.0200	0.0445	0.20
2	0.0200	0.0400	0.1400	0.1400	0.0400	0.38
3	0.0200	0.1645	0.0200	0.0200	0.0955	0.32
4	0.0200	0.0200	0.0200	0.0200	0.0200	0.10
	0.10	0.30	0.20	0.20	0.20	

Figure 8. Joint p.m.f. - Y_1 and Y_2 Uncorrelated

correlations (i.e., ± 1) for all pairs of random variables Y_1 and Y_2.) This method begins by finding the four solutions mentioned in §3 with NWCR. If any of the K_{min}, K^*_{min}, K^*_{max}, or K_{max} solutions corresponds to the specified correlation value, the desired instance of BTPSC is solved. If not, the procedure reallocates probability in such a way that (3), (4), and (6) are always satisfied and terminates when the reallocation process produces a probability assignment that satisfies (5) for the specified correlation value.

Suppose that the BTPSC solution is sought for $E(Y_1Y_2) = K$, where $K \in (K^*_{min}, K^*_{max})$. In this case, there is no change in the optimal solution value, θ^*. If one starts from the K^*_{min} solution to BTPSC, one simply has to reassign the excess probabilities, i.e., the x'_{j_1,j_2}s, from realizations which contribute little to

$$\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} y_{1j_1}y_{2j_2}(x'_{j_1,j_2}+\theta^*) = E(Y_1Y_2)$$

to those realizations that contribute more to this expectation, until this expectation equals K.

If $K \in (K_{min}, K^*_{min})$ or $K \in (K^*_{max}, K_{max})$, the objective function value will continually change as K is decreased below K^*_{min} or increased above K^*_{max}. A reallocation process that reduces θ by some $\Delta > 0$ and reallocates probability to all of the possible realizations can be used in this case. Essentially, this process works like a parametric analysis for the parameter K.

NWCR finds a basic feasible solution to a transportation problem. Define the NWCR path,

$$\{p_1, p_2, \ldots, p_{n_1+n_2-1}\},$$

to be the ordered list of cells in the transportation tableau that correspond to the basic variables chosen by NWCR, and let the cells in the path be referenced by row and column numbers. The first cell in the path is always $(1,1)$, and the last cell in the path is always (n_1, n_2).

Let r_{j_1,j_2} be the reallocation coefficient for the cell in row j_1 and column j_2. The r_{j_1,j_2}s are integer and denote the multiple of Δ that is to be added to each cell when probability is redistributed. For all cells (j_1, j_2) that are not on the NWCR path, $r_{j_1,j_2} = 0$. If $p_2 = (1,2)$, then $r_{11} = n_1$. Otherwise, $r_{11} = n_2$. If $p_{n_1+n_2-2} = (n_1, n_2 - 1)$, then $r_{n_1n_2} = n_1$. Otherwise, $r_{n_1n_2} = n_2$. The remainder of the reallocation coefficients are determined so that:

$$\sum_{j_1=1}^{n_1} r_{j_1,j_2} = n_2, \quad \forall j_2$$

and

$$\sum_{j_2=1}^{n_2} r_{j_1,j_2} = n_1, \quad \forall j_1.$$

If $K \in (K_{min}, K^*_{min})$ ($K \in (K^*_{max}, K_{max})$), start the reallocation process from the solution for $K = K^*_{min}$ (K^*_{max}). Let $\tau = \sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} y_{1j_1}y_{2j_2}x_{j_1,j_2}$ be the working value of K. The nominal reallocated probability value is:

$$\Delta = \min\left\{\frac{K-\tau}{\delta}, \min_{r_{j_1,j_2}<0}\left\{\frac{-x'_{j_1,j_2}}{r_{j_1,j_2}}\right\}\right\}.$$

If $\Delta = (K-\tau)/\delta$, then the BTPSC solution is updated and the solution procedure terminates. Otherwise, the BTPSC solution is updated, the solution path (basic feasible solution) is modified by flipping the cell that provided the limiting value for Δ over the solution path, and the process is repeated until the solution for the desired value of K is found.

The algorithm can be stated as follows:

1. Find K_{min} and K_{max} solutions with NWCR.
   If $K < K_{min}$ or $K > K_{max}$, stop as the problem is infeasible. If $K = K_{min}$ or $K = K_{max}$, stop.
   Otherwise, continue.

2. Find K^{-}_{min} and K^{*}_{max} solutions with NWCR.
   If $K = K^{*}_{min}$ or $K = K^{-}_{max}$, stop.
   If $K \in (K^{-}_{min}, K^{*}_{max})$, go to Step 3.
   If $K \in (K_{min}, K^{*}_{min})$, go to Step 6.
   If $K \in (K^{-}_{max}, K_{max})$, go to Step 7.

3. Begin with the NWCR solution for $K = K^{-}_{min}$.
   Let $(e, f) = p_1$ and $(r, s) = p_{m+n-1}$.
   $\tau = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} y_{1j_1} y_{2j_2} x_{j_1,j_2}$.

4. $\phi = (y_{2f} - y_{2s})(y_{1r} - y_{1e})$.

5. $\beta = \min\{(K - \tau)/\phi, x'_{ef}, x'_{rs}\}$.
   $x'_{rf} \leftarrow x'_{rf} + \beta$. $x'_{ef} \leftarrow x'_{ef} - \beta$.
   $x'_{es} \leftarrow x'_{es} + \beta$. $x'_{rs} \leftarrow x'_{rs} - \beta$.
   If $\beta = (K - \tau)/\phi$, then $x_{j_1,j_2} = x'_{j_1,j_2} + \theta, \quad \forall j_1, j_2$, and stop. Otherwise, $\tau \leftarrow \tau + \beta\phi$.
   If $\beta = x'_{ef}$, then $(e, f) = p_{e+f}$. Otherwise, $(r, s) = p_{r+s-2}$.
   Go to Step 4.

6. Begin with the NWCR solution for $K = K^{-}_{min}$.
   $\tau = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} y_{1j_1} y_{2j_2} x_{j_1,j_2}$. Go to Step 8.

7. Begin with the NWCR solution for $K = K^{-}_{max}$.
   $\tau = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} y_{1j_1} y_{2j_2} x_{j_1,j_2}$.

8. Calculate reallocation coefficients, r_{j_1,j_2}, $\forall j_1, j_2$.

9. $\delta = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} y_{j_1} y_{j_2} (r_{j_1,j_2} - 1)$.
   $\Delta = \min\left\{\frac{K-\tau}{\delta}, \min_{r_{j_1,j_2}<0}\left\{\frac{-x'_{j_1,j_2}}{r_{j_1,j_2}}\right\}\right\}$.

10. $\theta \leftarrow \theta - \Delta$. $\tau \leftarrow \tau + \delta\Delta$.
    $x'_{j_1,j_2} \leftarrow x'_{j_1,j_2} + r_{j_1,j_2}\Delta, \quad \forall j_1, j_2$.
    If $\tau = K$, then $x_{j_1,j_2} = x'_{j_1,j_2} + \theta, \quad \forall j_1, j_2$, and stop.
    Otherwise, go to Step 11.

11. Modify the NWCR solution path and go to Step 8.

Suppose that $\delta = +\infty$ for all $K \in [K^{*}_{min}, K^{-}_{max}]$. Peterson (1990) points out that $-1/\delta$ is the shadow price of the side constraint (5).

Recall the example problem presented earlier and suppose a probability assignment with $E(Y_1Y_2) = 9.90$, or $\rho = -0.64$, is sought. The algorithm begins with the NWCR solution for $K = K^{-}_{min}$ (see Figure 3). The reallocation coefficients are shown in Figure 9. After one iteration, the desired solution is found. This solution is displayed in Figure 10.

Y_1 \ Y_2	2	3	5	6	7	
4	5					5
3	-1	6				5
2		-2	4	4	-1	5
1					5	5
	4	4	4	4	4	

Figure 9. r_{j_1,j_2}s for $E(Y_1Y_2) = K^{-}_{min}$ Solution

Y_1 \ Y_2	2	3	5	6	7	
4	0.040					0.040
3	0.012	0.248				0.260
2		0.004	0.152	0.152	0.012	0.320
1					0.140	0.140
	0.052	0.252	0.152	0.152	0.152	

Figure 10. Optimal Solution (x'_{j_1,j_2}s) for $E(Y_1Y_2) = 9.90$

6. SUMMARY AND CONCLUSIONS

From the empirical evaluations of the performance of implicit enumeration routines on binary knapsack and weighted set covering problems and of greedy heuristic procedures on the weighted set covering problem, it seems that standard approaches for empirically evaluating solution procedures do not sample enough difficult and realistic test problems; i.e., those problems in which there is high positive correlation between the objective function and constraint coefficients. In order to conduct better evaluations of solution procedures, methods for characterizing discrete bivariate distributions that provide more realisitic and more challenging test problems are needed.

The problem of finding the characterization of a discrete bivariate random variable for which the smallest joint probability value is as large as possible was shown to be a bottleneck transportation problem with a side constraint, BTPSC. An algorithm, based on NWCR and a simple probability redistribution scheme, was devised to solve BTPSC for any correlation value.

It is suspected that this new characterization, regardless of correlation value, would provide more challenging test problems for empirical evaluations of solution procedures for integer programs than conventional problem generation methods. Evaluations of solution methods would be conducted over the same space of problems as in the case where problem parameters are distributed independently. But, the problems generated would be increasingly likely to have some dependence structure as $|\rho| \rightarrow +1$.

There are many opportunities for more research in this area. For example, the studies by Moore (1989, 1990) could be repeated for more positive correlation values to learn more precisely how correlation affects the performance of implicit enumeration and greedy heuristics. Peterson (1990) has shown how the problem of maximizing the smallest joint probability value for a multivariate random variable with any feasible correlation structure can be formulated as a linear program. Perhaps an algorithm can be devised to solve this linear program. The linear programming approach might be extendable to include distributions with a countable number of possible values (e.g., the Poisson distribution). It would be interesting to consider other objective functions that would lead to different characterizations of discrete bivariate and multivariate distributions. Also, an empirical evaluation of the performance of solution procedures when test problems are generated using the bivariate characterization based on probabilistic mixing and the one based on BTPSC mixing should be undertaken. Finally, efficient schemes for generating observations of discrete bivariate random variables characterized by the approach described here should be developed.

ACKNOWLEDGEMENT

The authors are grateful to Dr. Barry L. Nelson for participating in many discussions about this research.

REFERENCES

Bratley, P., Fox, B.L., and L.E. Schrage (1983), *A Guide to Simulation*, Springer-Verlag, Inc., New York, NY.

Chvátal, V. (1979), "A Greedy Heuristic for the Set-Covering Problem," *Mathematics of Operations Research 4*, 3, 233-235.

Evans, J.R. (1984), "The Factored Transportation Problem," *Management Science 30* 8, 1021-1024.

Moore, B.A. (1989), "Correlated 0-1 Knapsack Problems," IND ENG 854 Course Project, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, OH.

Moore, B.A. (1990), "The Effect of Correlation on the Performance of Exact and Heuristic Procedures for the Weighted Set Covering Problem," M.S. Thesis, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, OH.

Peterson, J.A. (1990), "A Parametric Analysis of a Bottleneck Transportation Problem Applied to the Characterization of Correlated Discrete Random Variables," M.S. Thesis, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, OH.

Roach, W., and R. Wright (1977), "Optimal Antithetic Sampling Plans," *Journal of Statistical Computation and Simulation 5*, 1, 99-114.

Schmeiser, B.W., and R. Lal (1982), "Bivariate Gamma Random Variables," *Operations Research 30* 2, 355-374.