Working Paper 238

On Solving Some Stochastic Discrete Optimization Problems Under General Regret Function

By

Shubhabrata Das
Diptesh Ghosh
Pranab K. Mandal

July 2005

Please address all correspondence to:

Prof. Shubhabrata Das
QM&IS Area
Indian Institute of Management – Bangalore
Bannerghatta Road, Bangalore 560 076, India
Email: shubho@iimb.ernet.in
Phone: +91-80-6993150, Fax: +91-80-2658 4050

Prof. Diptesh Ghosh
P&QM Area,
Indian Institute of Management Ahmedabad
Vastrapur, Ahmedabad 380 015, India

Prof. Pranab K. Mandal
Dept. of Applied Mathematics, Faculty of EEMCS,
University of Twente, P.O. Box 217, 7500AE Enschede,
The Netherlands

On Solving Some Stochastic Discrete Optimization Problems Under General Regret Function

Shubhabrata Das *

QM&IS Area, IIM Bangalore, Bannerghatta Road, Bangalore 560076, INDIA.

Diptesh Ghosh

P&QM Area, IIM Ahmedabad, Vastrapur, Ahmedabad 380015, INDIA.

Pranab K. Mandal

Dept. of Applied Mathematics, Faculty of EEMCS, University of Twente, P.O.Box 217, 7500AE Enschede, The NETHERLANDS.

July 13, 2005

Abstract

In this paper we consider stochastic discrete optimization problems (DOP) in which feasible solutions remain feasible irrespective of the randomness of the problem parameters. We introduce the concept of the risk associated with a solution and define optimal solution in terms of having least possible risk. We show that a least risk solution can be obtained by solving a non-stochastic discrete optimization problem similar to the stochastic problem in certain problems and present results regarding the generation of the non-stochastic problem in terms of finding the parameter of the distribution which may act as surrogate for the random element in its non-stochastic counterpart. While this surrogate is the mean for a linear regret function, the situation is complex under general regret. Our results show that the above result continues to hold (in general) if the DOP has only one random element having symmetric distribution. We obtain some bounds for this parameter for certain group of asymmetric distributions and study its limiting behaviour under two asymptotic setup. We establish through various examples, that the results from the uni-dimensional case cannot be extended to stochastic DOP with multiple random element with any reasonable generality. However, we characterize a finite number of solutions which will include the optimal solution in this case. An heuristic based on local search type algorithm is also devised when the number of random elements is too high, and we study the performance of this algorithm through simulation.

*Corresponding author

1 Introduction

In discrete optimization problems (DOPs), some of the problem parameters are often stochastic in nature. In these situations, the traditional notion of optimality (e.g. least cost solutions for minimization problems) needs to be revised. In this article we consider optimality in terms of regret function, leading to the notion of least risk solutions (formally introduced in Section 2). We assume knowledge about the stochasticity of the parameters (which is often the case in practice) and try to find least risk solutions. The method we adopt is to formulate a non-stochastic DOP, such that an optimal solution to the non-stochastic DOP is a least risk solution to the stochastic DOP. The non-stochastic DOP is essentially the stochastic DOP with the random parameters pegged at values based on our knowledge of their randomness as well as other problem-specific parameters.

In stochastic DOPs, the notion of optimality is not unique. A well considered criterion is to maximize the probability that the (random) objective function reaches a prespecified threshold level (Frank (1969) [7] in the context of shortest path problem and Henig (1990) [9], Carraway et al (1993) [2], in the context of stochastic knapsack problem). Another closely related notion is to find the solution thats leads to the optimal thershold value satisfying the constraint that the probability of the objective function reaching the threshold value is at most a pre-specified value *alpha*. See, e.g., Henig 1990 [9] (knapsack problem), Ishii et al (1981) [12] (minimum spanning tree problem).

The most prevalent in practice/literature, though, is the *expected utility* criterion of Von Neumann and Morgenstern (see Fishburn (1968) [6]). With this criterion one maximizes the decision maker's expected utility. See, e.g., Murthy et al (1998) [16], Loui (1983) [15] (shortest path problem); Dean et al (2004) [5] (knapsack problem)

On the other hand, behavioral notion of (posterior) *regret* (introduced by Savage (1954) [18]) for not achieving what the decision maker could have achieved with another choice, plays an important role in making a decision. See, for example, Kaliszewski et al (1998) [13]. This motivates us to work with a notion of optimality involving regret function. To the best of authors' knowledge not much has been done with this notion of optimality.

The remainder of this paper is organized as follows. Section 2 contains all the notation and definitions to be used throughout the article along with the precise description of the nature of the stochastic DOP we consider. Section 3 is focussed on DOPs with one random element. In Section 3.1 we show that knowledge of the mean of the distribution functions of the costs of the random elements is sufficient to obtain an optimal solution for these problems provided the distribution of the random element is symmetric irrespective of the choice of the regret function. The general scenario with random elément having homogeneously skewed distribution is taken up in Section 3.2. The case of multiple random element is discussed in Section 4, where we identify the candidate list of optimal solutions. However, it is also shown that unless the regret function is linear, it is not possible to extend the result of the earlier sections, even if the symmetrically distributed random elements are independent. In Section 5, we describe the algorithms that we have implemented to solve discrete optimization problems with random variables, and report our experiences with these algorithms on instances of the binary knapsack problem with six to eight random elements, whose joint distributions are discrete. Our results suggest that when the number of random elements are in this range, tabu search is a more efficient choice than complete enumeration for solving problems.

2 Notation and Definitions

In this section, we describe the notation that we use in this paper, and also provide the relevant definitions.

Definition 1 A *discrete optimization problem (DOP)* is denoted by $\pi = (G, \mathbb{S}, z)$, where G is a finite ground set, with each element $e \in G$ having an associated value c_e (often referred to as the cost of e). The set, $\mathbb{S}$, of feasible solutions is a subset of the power set of G and is usually described by a set of rules that each $S \in \mathbb{S}$ must satisfy. The function $z : \mathbb{S} \to \Re$ is referred to as the objective function (or the cost function), and the optimization problem is one of finding a member of $\arg\min_{S \in \mathbb{S}}\{z(S)\}$.

Definition 2 An element $e \in G$ in $\pi = (G, \mathbb{S}, z)$ is called *random* (alternatively *fixed*) if c_e is random valued (alternatively constant).

Definition 3 A *stochastic DOP* (SDOP) is one in which the costs of some of the elements in G are random.

Consider, for example, a symmetric traveling salesperson's problem (TSP) where one has to cover n cities, with many possible intercity routes to choose from. Suppose the costs associated with some of the intercity routes are random. The goal is to find an optimal route in terms of cost of travelling. Another example is the binary knapsack problem where one is to select items (each having an associated weight w and value v) maximizing the total value so that the total weight does not exceed a predetermined capacity c. If some of the values (v) are random then this problem would turn into an SDOP.

In this work, we restrict ourselves to SDOPs where all feasible solutions remain feasible, irrespective of the random parameter values. In the TSP setup described above this is satisfied if we assume that all routes are (always) feasible. On the other hand, in the knapsack problem, we need to have deterministic (nonrandom) weights. If the budget or the weights are random, then the feasibility of a set of items depends on the randomness and such a problem is beyond the scope of the current work; see Das and Ghosh (2003) [4] for treatment of such problems.

Further, we shall consider *min-sum* objective functions, that is, $z(S) = \sum_{e \in S} c_e$. Also, The probability distributions of the random elements are assumed to be known and unimodal.

Definition 4 Given any fixed set of values for c_e's, the regret associated with a solution $S \in \mathbb{S}$ is defined by

$$\text{regret}(S) = r(z(S) - Z^*),$$

where Z^* is the minimum possible value of the objective function for given values of c_e's (and hence is a function of these c_e's) and $r(\cdot)$ is an increasing continuous function on $[0, \infty)$, such that $r(0) = 0$.

Obviously, with some of the c_e's being random, the regret associated with any feasible solution S is also a random variable. In practice, it would not be desirable to adopt a new course of action with every alteration of the c_e's, especially if we deal with $\mathcal{NP}$-hard problems. So, we need to find a solution which would be "good" regardless of the realization of the costs of the random elements. With this in mind, we define the risk associated with a solution in the following manner:

Definition 5 The *risk* associated with a solution $S \in \mathbb{S}$ is given by

$$R(S) = \mathbb{E}\,\text{regret}(S) = \mathbb{E}\, r(z(S) - Z^*),$$

where Z^* is the cost of the least cost solution at specific values of the random elements, and hence is random itself. The expectation is taken with respect to the costs of the random elements. The $r(\cdot)$ function is as in Definition 4.

Definition 6 For a DOP with random elements, an *optimal* solution (also referred to as a least risk solution) is defined as a feasible solution with minimum risk among all feasible solutions.

Notice that if all the elements of the instance are fixed, the minimum risk solution corresponds to the traditional concept of an optimal solution, i.e., a least cost solution.

We will need the following notion in our later analysis.

Definition 7 The set of feasible solutions $\mathbb{S}$ in a (stochastic) DOP $\pi = (G, \mathbb{S}, z)$ is said to be *balanced* (equivalently the DOP is called balanced) if

$$S(\subseteq G) \in \mathbb{S}, |S| = m \Rightarrow \tilde{S} \in \mathbb{S} \text{ for any } \tilde{S} \subseteq G \text{ with } |\tilde{S}| = m.$$

Recall that the probability distributions of the random costs under consideration are assumed to be unimodal and thus they can be either symmetric or skewed. In part of this work, particular attention is given to a specific class of skewed distributions which we refer to as homogeneously skewed distributions.

Definition 8 Suppose a unimodal distribution with mode M has density $h(\cdot)$, that is, $h(\cdot)$ is increasing in $(-\infty, M]$ and decreasing in $[M, \infty)$. It is said to be *homogeneously right-skewed* if

$$h(M + x) \geq h(M - x) \quad \text{for almost all } x > 0. \tag{1}$$

It is called homogeneously left-skewed if

$$h(M + x) \leq h(M - x) \quad \text{for almost all } x > 0. \tag{2}$$

A unimodal distribution is said to be *homogeneously skewed* provided it is either homogeneously right-skewed or homogeneously left-skewed. It is convenient to formally define the skewness function and a measure of skewness for homogeneously skewed distributions.

Definition 9 The *skewness function* of a homogeneously skewed distribution with mode M and p.d.f. $h(\cdot)$ is defined as

$$\gamma_h(x) = h(M + x) - h(M - x), \quad x > 0. \tag{3}$$

Definition 10 The *measure of homogeneous skewness* of a homogeneously skewed distribution with mode M and p.d.f. $h(\cdot)$ is defined as

$$\tau_h = \int_0^\infty \gamma_h(x)dx = \int_0^\infty \{h(M + x) - h(M - x)\}dx. \tag{4}$$

3 DOPs with one random element

As mentioned already, we confine ourselves to DOPs with min-sum objective functions. Let $\pi = (G, \mathbb{S}, z)$ be a DOP instance with a single random element $e \in G$. First, we study the least cost objective function value (Z^*) as a function of c_e.

Let us denote the cost of the random element e by a random variable X. Let X have a (cumulative) distribution function $H(.)$ with mean μ, i.e. $H(x) = P(X \leq x)$, and $\mu = \int x dH(x)$.

We split the set of all feasible solutions $\mathbb{S}$ into $\mathbb{S}_e$ and $\mathbb{S}^e$, respectively consisting of all solutions containing e, and of all solutions not containing e. Let S_e be a least cost solution in $\mathbb{S}_e$ and S^e be a least cost solution in $\mathbb{S}^e$. We note that while S_e and S^e need not be unique, they remain least cost solutions in their respective groups regardless of the value of c_e. This is because, a change in c_e does not affect the cost of any solution in $\mathbb{S}^e$, while it affects all solutions in $\mathbb{S}_e$ by the same amount.

For extreme possible low values of c_e, typically, $z(S_e) < z(S^e)$. (Otherwise, the randomness or otherwise of c_e is not an issue at all, since e would not be included in the optimal solution in any case.) When c_e increases, the cost of all solutions in $\mathbb{S}_e$ increase while the cost of all solution in $\mathbb{S}^e$ remain the same. So S_e remains optimal until c_e increases to become larger than some threshold value, say ω, when $z(S_e)$ becomes equal to $z(S^e)$. If c_e increases further, $z(S_e) > z(S^e)$, and S^e becomes a new optimal solution. Clearly, no further increase in c_e will make S^e suboptimal. We see therefore, that $Z^*(c_e)$ is a continuous function with a slope of 1 when $c_e < \omega$ and a slope of 0 when $c_e > \omega$ (see Figure 1).

Note that

$$z(S_e) = c_e + \sum_{e' \in S_e \setminus \{e\}} c_{e'}. \tag{5}$$

It follows from the discussion above that

$$\omega = z(S^e) - \sum_{e' \in S_e \setminus \{e\}} c_{e'}. \tag{6}$$

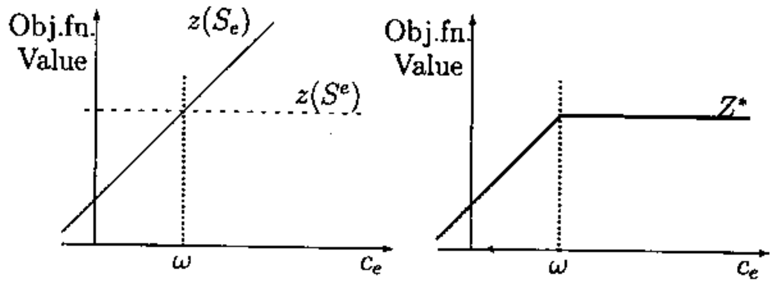

Figure 1: $z(S_e)$, $z(S^e)$, and Z^* as a function of c_e (min-sum objective)

Further, since S_e is optimal in the least cost sense if $c_e \leq \omega$, in adopting S_e as a solution, one incurs a regret equal to $r(x-\omega)$ if $c_e(=x) > \omega$. Similarly, by taking S^e as a solution, there is a regret of $r(\omega - x)$ when $x < \omega$. Thus, the risk of these two solutions are

$$R(S_e) = \int_{\omega}^{\infty} r(x-\omega)dH(x); \quad R(S^e) = \int_{-\infty}^{\omega} r(\omega - x)dH(x). \tag{7}$$

Recall that our objective in this paper is to replace the cost of the random element in the DOP with a fixed value, such that the least cost solution to the modified DOP is the least risk solution to the original stochastic DOP. To that end, let us define the function

$$\Psi_{r,X}(t) = \int_{t}^{\infty} r(x-t)dH(x) \; - \; \int_{-\infty}^{t} r(t-x)dH(x). \tag{8}$$

It is easy to see from (8) that $\Psi_{r,X}(\cdot)$ is a decreasing function for any increasing $r(\cdot)$. (See also Figure 2 below.) Naturally, $\Psi_{r,X}(\cdot)$ depends on X through its distribution function $H(\cdot)$. In the notation, either or both of the suffixes in Ψ may be suppressed, if obvious from the context.

Now we state the main result of this section.

Theorem 1 *A least risk solution to a SDOP with one random element can be obtained by solving a non-stochastic DOP obtained by replacing the random cost by θ, where θ is the solution to*

$$\Psi(t) = 0, \tag{9}$$

and $\Psi(t)$ is as defined in (8).

Proof: It follows from (7) and (8) that $\Psi(\omega) = R(S_e) - R(S^e)$. Hence

$$\begin{aligned} & R(S^e) \leq R(S_e) \\ \Leftrightarrow \quad & \Psi(\omega) \geq 0 \equiv \Psi(\theta) \\ \Leftrightarrow \quad & \omega \leq \theta, \qquad \text{since } \Psi(\cdot) \text{ is decreasing} \\ \Leftrightarrow \quad & z(S^e) \leq \theta + \sum_{e' \in S_e \setminus \{e\}} c_{e'}, \qquad \text{by (6).} \end{aligned}$$

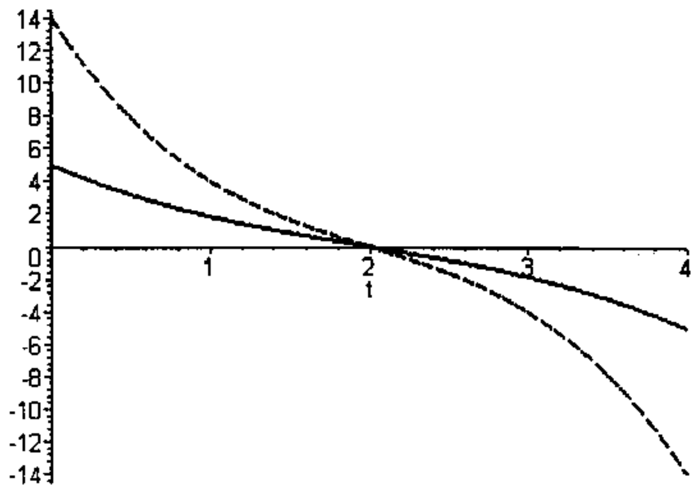

Figure 2: $\Psi(t)$ when $c_e \sim N(2,1)$ and regret function is either quadratic (solid line) or cubic (dashed lines).

But from (5), the RHS of the last inequality $\theta + \sum_{e' \in S_e \setminus \{e\}} c_{e'}$ is equal to $z(S_e)$ when $c_e = \theta$. Hence S^e is a least risk solution if and only if S^e is the least *cost* solution to the (nonstochastic) DOP when the random cost is replaced by θ. ∎

In the remainder of this section, we study the properties of the solution θ of (9).

3.1 Random Element with Symmetric Distribution

Suppose the random element has a symmetric distribution. Then, as proved in the following theorem, the optimal solution to the one dimensional SDOP under a general non-decreasing regret function $r(\cdot)$ may be obtained by replacing the random element with its central value.

Theorem 2 *Let π be a DOP with a single random element e with cost $c_e = X$ having a symmetric density function around its measure of location μ, and π_1 be the same DOP but with c_e fixed at μ. Then a least cost solution to π_1 is a least risk solution to π.*

Proof: It is easy to observe that the $\Psi(\cdot)$ function introduced in (8) may be alternatively written as

$$\Psi(t) = \mathbb{E}\left[r(X-t)\mathbf{I}_{\{X \geq t\}}\right] - \mathbb{E}\left[r(t-X)\mathbf{I}_{\{X \leq t\}}\right]. \tag{10}$$

Now X being symmetric, $X - \mu$ has the same distribution as $\mu - X$. Hence,

$$\mathbb{E}\left[r(X-\mu)\mathbf{I}_{\{X-\mu \geq 0\}}\right] = \mathbf{E}\left[r(\mu - X)\mathbb{I}_{\{\mu - X \geq 0\}}\right],$$

and hence

$$\Psi(\mu) = 0. \tag{11}$$

The result then follows from Theorem 1. ∎

Theorem 2 implies that for DOPs with one random element, a knowledge of the mean (median) of the distribution function of the cost of the random element is enough to compute an optimal solution provided c_e follows a distribution that is symmetric in nature.

If the regret function is linear, then $\Psi_{\text{linear}}(\mu) = 0$ for symmetric as well as asymmetric distributions. This result holds for the case with an arbitrary number of random elements as shown later through Theorem 8.

In many situations, c_e is restricted to have non-negative values. However, with proper choice of parameters of the distribution one can achieve nonnegativity for all practical purposes. Also, from Theorem 6 in Section 3.3, it is apparent that if the distributional assumption is even approximately valid, then the result will continue to hold with minor deviations in the critical value θ. Therefore modeling c_e using distributions that allow negative values (say Gaussian) does not pose a big problem.

Consider for example, a case in which we model the randomness of c_e using a Gaussian distribution with mean mu and standard deviation 1; however, in reality c_e has a truncated Gaussian distribution supported on $[0, \infty)$. Obviously, the true distribution is not symmetric and hence θ does not coincide with the modeled mean μ. According to the discussion above, the error $(\theta - \mu)$ is small, as shown in the Figures 3 and 4.

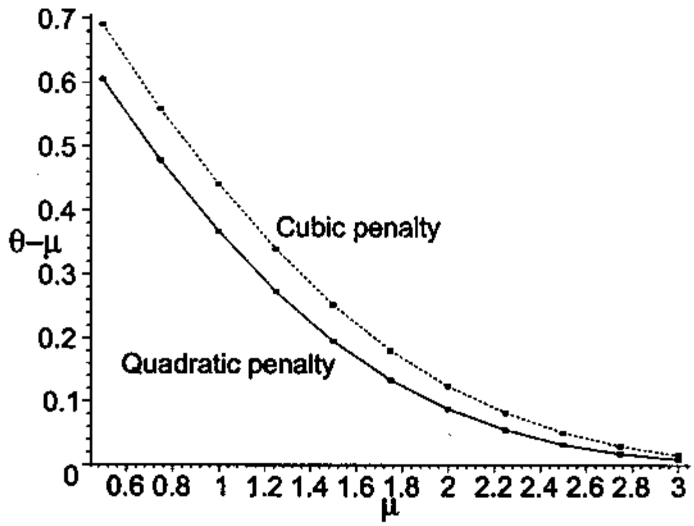

Figure 3: Plot of error vs. μ when non-negative values are not allowed

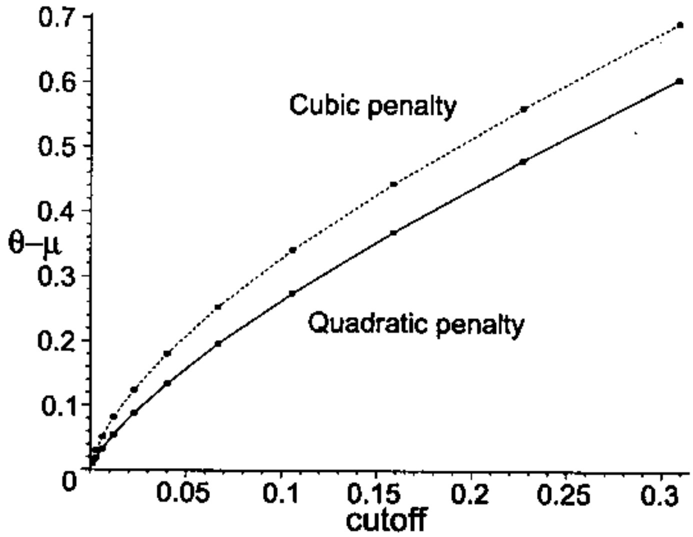

Figure 4: Plot of error vs. $\int_{-\infty}^{0} \phi(x)dx$ when nonnegative values are not allowed

3.2 Random Element with Skewed Distribution

In this section we deal with the case where the random element follows a skewed distribution. Theorem 2 does not hold true in general without the assumption of symmetry. For example, if c_e has a Beta distribution with parameters 1 and 5, then $\mu = \frac{1}{6}$, and with the squared error regret function ($r(t) = t^2$), θ turns out to be close to 0.19.

Suppose the cost of the random element has continuous distribution with density $h(x)$, it is unimodal with mode M, and homogeneously right skewed as defined in (1). We also assume that the density function has requisite finite moments so that the optimal solutions have finite risk as per the choice of the regret function. Then the following holds.

Theorem 3 *Consider a stochastic DOP with a single random element which has a homogeneously right-skewed cost distribution with mode M. Then $\theta \geq M$.*

Proof: Note that

$$\begin{aligned}
\Psi(M) &= \int_{M}^{\infty} r(x-M)h(x)dx - \int_{-\infty}^{M} r(M-x)h(x)dx \\
&= \int_{0}^{\infty} r(y)h(M+y)dy + \int_{\infty}^{0} r(y)h(M-y)dy \\
&= \int_{0}^{\infty} r(y)[h(M+y) - h(M-y)]dy \geq 0, \quad \text{(by (1))}.
\end{aligned}$$

The result follows from the fact that $\Psi(\cdot)$ is a non-increasing function. ∎

Remark 1 The result holds even when the random element has a finite support (say $[L, U]$). The proof follows along similar lines by noting that $U - M \geq M - L$ as a consequence of $h(\cdot)$ being homogeneously right-skewed.

In case the random element has a non-increasing density function, the result in Theorem 3 is not useful. In such cases the following theorem provides an upper bound to the value of θ.

Theorem 4 *Consider a stochastic DOP with a single random element which has a non-increasing density function $h(\cdot)$ supported on $[L, U]$. Then for any general non-decreasing regret function $r(\cdot)$, $\theta \leq \frac{L+U}{2}$.*

Proof: Using steps similar to those in the proof of Theorem 3 we obtain

$$\Psi\Big(\frac{L+U}{2}\Big) = \int_0^{\frac{U-L}{2}} r(y)\Big\{h\Big(y + \frac{L+U}{2}\Big) - h\Big(\frac{L+U}{2} - y\Big)\Big\}dy$$

which is non-positive since $h(\cdot)$ is non-increasing. The result follows since $\Psi(\cdot)$ is non-increasing. ∎

Remark 2 Natural analogues to Theorems 3 and 4 exist for situations in which the distribution of c_e is homogeneously left skewed (see Definition 8) or has a non-decreasing density.

We now investigate the behavior of θ under certain distributions for the random element. In particular, two types of density functions,viz. Triangular and Beta have been considered here. They are defined on a finite support, $[0, 1]$ having the functional form:

$$\text{Beta distribution: } h(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise; and} \end{cases}$$

$$\text{Triangular distribution: } h(x) = \begin{cases} \frac{2x}{M} & \text{for } 0 \leq x \leq M \\ \frac{2(1-x)}{(1-M)} & \text{for } M \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Within this framework, the parameters have been varied to incorporate various degrees of skewness for the random element. The optimal solution is sought under the regret functions of the form $r(t) = (1+t)^n - 1$, for illustration. Note that this regret is interesting as by increasing the value of n, we obtain a set of regret functions which impose penalties of increasing strictness for the same amount of deviation. The associated variations in the θ are shown in Figures 5 through 7. In Figure 6, the τ refers to the measure of homogenous skewness as defined in Definition 4. The exact values of θ for Figures 5 through 7 are reported in Appendix Tables 6 through 8.

Key observations from these computational results are objectively reinforced through the asymptotic results of the next subsection.

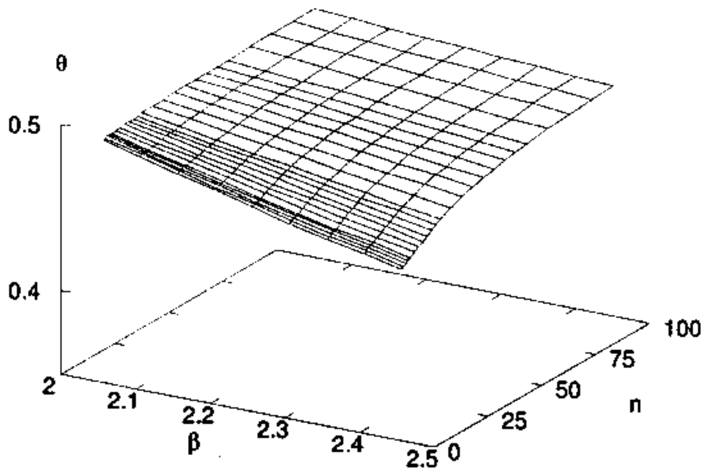

Figure 5: Plot of θ values against β and n when $c_e \sim Beta(2, \beta)$ and $r(t) = (1+t)^n - 1$

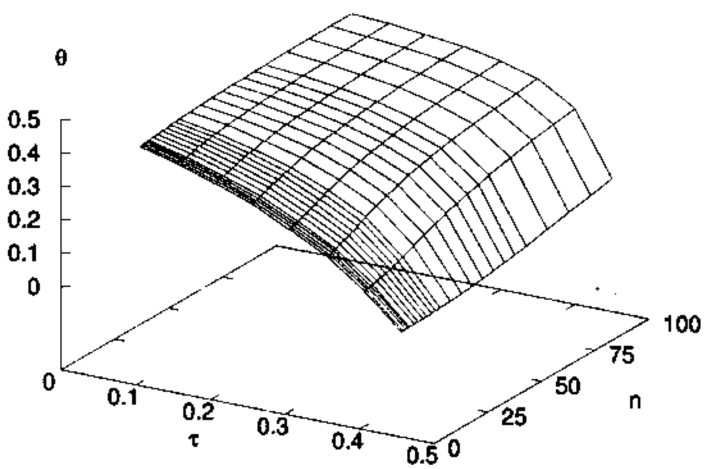

Figure 6: Plot of θ values against τ and n when $c_e \sim Beta(2, \beta)$ and $r(t) = (1+t)^n - 1$

3.3 Asymptotic Results with Single Random Element

Recall that from our discussion in the beginning of this section that the value of θ (as defined through (9)) is critical in solving the stochastic DOP through a solution of its non-stochastic counterpart. In this section, we study the limiting behaviour of θ under two asymptotic scenario. These results, proved under various regularity conditions, show that the theorems of Sections 3.1and 3.2 would be (approximately) valid if the requisite conditions on the probability distribution is more or less true.

Theorem 5 *Consider a stochastic DOP with a single random element (X) supported on the interval $[L, U]$. For a sequence of (strictly increasing) regret functions*

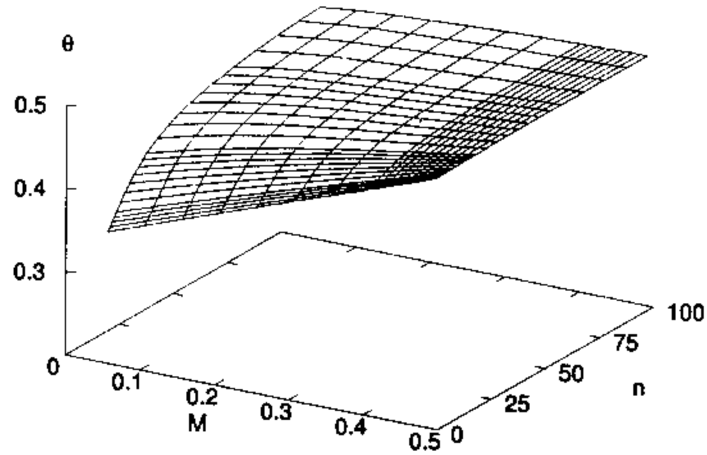

Figure 7: Plot of θ values against M and n when the distribution of c_e is triangular with mode M and $r(t) = (1+t)^n - 1$

$\{r_n(\cdot),\ n \geq 1\}$ satisfying:

$$\lim_{n\to\infty} \frac{r_n(t_1)}{r_n(t_2)} = \infty,\ \forall t_1 > t_2, \tag{12}$$

define $\Psi_n(t) = \Psi_{r_n,X}(t)$ as in (10) and let θ_n be the solution of $\Psi_n(t) = 0$. Then

$$\theta_n \to \frac{L+U}{2} \quad \textit{as } n \to \infty,$$

so long as the the RHS of the limit makes sense, i.e. at least one of L and U is finite.

Remark 3 Condition (12) of the theorem above is satisfied by commonly used regret functions such as $r_n(t) = t^n$; $r_n(t) = (1+t)^n - 1$ and $r_n(t) = \exp(\lambda_n t) - 1$, where $\lambda_n \to \infty$.

Proof of Theorem 5 First let us consider both L and U to be finite. Note that for any $\delta > 0$, $[L, L+\delta]$ and $[U-\delta, U]$, will have positive probabilities. It is enough to show that, for any given $\epsilon > 0$,

$$\frac{L+U}{2} - \epsilon \leq \theta_n \leq \frac{L+U}{2} + \epsilon, \tag{13}$$

for 'large' n. To prove (13), it suffices ($\Psi_n(\cdot)$ being a decreasing function) to show that

$$\Psi_n\left(\frac{L+U}{2} - \epsilon\right) \geq 0 \text{ and } \Psi_n\left(\frac{L+U}{2} + \epsilon\right) \leq 0.$$

Suppose $L < t < \frac{L+U}{2}$. Denote $\delta_1 := \frac{L+U}{2} - t > 0$. Then

$$\begin{aligned}
\Psi_n(t) &= \mathbb{E}\left[r_n(X-t)\mathbb{I}_{\{t\le X<U-\delta_1\}}\right] + \mathbb{E}\left[r_n(X-t)\mathbb{I}_{\{U-\delta_1\le X\le U\}}\right] \\
&\qquad\qquad\qquad\qquad - \mathbb{E}\left[r_n(t-X)\mathbb{I}_{\{L\le X\le t\}}\right] \\
&\ge \mathbb{E}\left[r_n(X-t)\mathbb{I}_{\{U-\delta_1\le X\le U\}}\right] - \mathbb{E}\left[r_n(t-X)\mathbb{I}_{\{L\le X\le t\}}\right] \\
&\ge r_n(U-\delta_1-t)\mathbb{P}(U-\delta_1\le X\le U) - r_n(t-L)\cdot\mathbb{P}(L\le X\le t) \\
&= r_n(\beta_1)\cdot a_1 - r_n(\gamma_1)\cdot b_1, \qquad (14)
\end{aligned}$$

where

$$\begin{aligned}
&a_1 = \mathbb{P}(U-\delta_1\le X\le U) > 0; \qquad b_1 = \mathbb{P}(L\le X\le t) > 0; \\
&\beta_1 = U-\delta_1-t = \tfrac{U-L}{2} > 0; \qquad \gamma_1 = t-L > 0;
\end{aligned}$$

Since $\beta_1 - \gamma_1 = \delta_1 > 0$ we have $\lim_{n\to\infty}\frac{r_n(\beta_1)}{r_n(\gamma_1)} = \infty$. Further, noting that $a_1, b_1 > 0$, it follows from (14) that there exists $N^* > 0$ such that

$$\Psi_n\left(\frac{L+U}{2} - \epsilon\right) \ge r_n(\beta_1)\cdot a_1 - r_n(\gamma_1)\cdot b_1 > 0 \qquad \text{for} \quad n \ge N^*.$$

In a similar manner it can be shown that there exists $N^{**} > 0$ such that

$$\Psi_n\left(\frac{L+U}{2} + \epsilon\right) \le 0 \qquad \text{for} \quad n \ge N^{**}.$$

This completes the proof of (13).

Now suppose $U = \infty$ and L is finite. Note that in this case $\mathbb{P}(X \ge \xi) > 0$ for any given ξ. Consider any fixed $M > L$. Choose $\xi > 2M - L$. Then we have

$$\begin{aligned}
\Psi_n(M) &= \mathbb{E}\left[r_n(X-M)\mathbb{I}_{\{X\ge M\}}\right] - \mathbb{E}\left[r_n(M-X)\mathbb{I}_{\{L\le X\le M\}}\right] \\
&\ge \mathbb{E}\left[r_n(X-M)\mathbb{I}_{\{X\ge \xi\}}\right] - \mathbb{E}\left[r_n(M-X)\mathbb{I}_{\{L\le X\le M\}}\right] \\
&\ge r_n(\xi-M)\mathbb{P}(X\ge\xi) - r_n(M-L)\cdot\mathbb{P}(L\le X\le M) \\
&= r_n(\beta_3)\cdot a_3 - r_n(\gamma_3)\cdot b_3, \qquad (15)
\end{aligned}$$

where

$$\begin{aligned}
&a_3 = \mathbb{P}(X\ge\xi) > 0; \qquad b_3 = \mathbb{P}(L\le X\le M) > 0; \\
&\beta_3 = \xi - M; \qquad \gamma_3 = M - L; \qquad \text{and} \quad \beta_3 > \gamma_3.
\end{aligned}$$

Therefore for any given M, $\Psi_n(M) > 0$ for large n, implying $\theta_n \ge M$. This proves that $\theta_n \to \infty$.

Similarly one can prove that $\theta_n \to -\infty$, if $L = -\infty$ and U is finite. ∎

Theorem 6 *Consider a sequence of stochastic DOPs π_n with a single random element having cost X_n. Suppose X_n converges to a X in some suitable sense to be specified below. Define $\Psi_n(t) = \Psi_{r,X_n}(t)$ and $\Psi(t) = \Psi_{r,X}(t)$ for any fixed (increasing) regret function $r(\cdot)$. Let $\Psi_n(\theta_n) = 0$ and $\Psi(\theta) = 0$. Then*

$$\theta_n \to \theta, \quad \text{as } n \to \infty. \tag{16}$$

under any of the following regularity conditions involving X_n, its convergence and/or the regret function $r(\cdot)$.

(A) *The random variables X_n, $n \geq 1$ and X are discrete taking the* identical set of values *with X_n converging to X weakly (in distribution). If the set of values of the random variables is an infinite collection, then $r(\cdot)$ is required to be a bounded function.*

(B) *The random variables X_n, $n \geq 1$ and X are discrete taking the* same (finite) number of distinct values *with X_n converging to X weakly (in distribution) and $r(\cdot)$ is a continuous function.*

(C) *The random variables X_n, $n \geq 1$ and X are continuous with the densities $h_n(x) \to h(x)$ for each x and $r(\cdot)$ is bounded.*

(D) *The random variables X_n, $n \geq 1$ and X are continuous with the densities h_n's and h having identical finite support $[L, U]$, $h_n(x) \to h(x)$ for each x and $r(\cdot)$ is continuous.*

(E) *The random variables X_n, $n \geq 1$ and X are continuous with the densities $h_n(x)$ converging to $h(x)$ uniformly in x and r is integrable.*

Remark 4 If the limiting random variable X is symmetric, then θ_n converges to μ, the mean of the limiting distribution.

Remark 5 In the regularity conditions A and D, the 'identical' support constraint may be relaxed to indicate that the supports of X and X_n (for large n) are contained in a finite interval.

Proof of Theorem 6. For any given $\epsilon > 0$, we shall show that

$$\Psi_n(\theta - \epsilon) \geq 0 \quad \text{and} \quad \Psi_n(\theta + \epsilon) \leq 0, \quad \text{for all 'large' } n,$$

implying that $\theta - \epsilon \leq \theta_n \leq \theta + \epsilon$. Note that

$$\begin{aligned} \Psi_n(t) &= \Psi(t) + \left(\mathbb{E}\left[r(X_n - t)\mathbb{I}_{\{X_n \geq t\}}\right] - \mathbb{E}\left[r(X - t)\mathbb{I}_{\{X \geq t\}}\right]\right) \\ &\qquad + \left(\mathbb{E}\left[r(t - X)\mathbb{I}_{\{X \leq t\}}\right] - \mathbb{E}\left[r(t - X_n)\mathbb{I}_{\{X_n \leq t\}}\right]\right) \\ &= \Psi(t) + a_n(t) + b_n(t), \quad \text{say.} \end{aligned} \tag{17}$$

We will show below that, under any of the regularity conditions (A) – (E),

$$\lim_{n\to\infty} a_n(t) = 0 \quad \text{and} \quad \lim_{n\to\infty} b_n(t) = 0, \quad \forall t. \tag{18}$$

By definition of θ and property of Ψ, note that $\Psi(\theta-\epsilon) = \delta > 0$ and $\Psi(\theta+\epsilon) = -\gamma < 0$. Then from (17) and (18) we have for large n,

$$\Psi_n(\theta-\epsilon) = \delta + a_n(\theta-\epsilon) + b_n(\theta-\epsilon) > 0$$

and

$$\Psi_n(\theta+\epsilon) = -\gamma + a_n(\theta+\epsilon) + b_n(\theta+\epsilon) < 0$$

completing the proof of the theorem. ∎

The proof of (18) under the regularity conditions: We would provide the proof for the $\{a_n\}$ sequence only, as the same for the $\{b_n\}$ would follow similarly.

(A) Let the distinct set of values of X_n and X be $\mathcal{S} = \{s_1 < s_2 < s_3 < \cdots\}$. First note that, weak convergence of X_n to X implies:

$$P(X_n = s_k) \to P(X = s_k), \quad P(X_n > s_k) \to P(X > s_k) \quad \forall k. \tag{19}$$

From (17), we have

$$a_n(t) = \sum_{s_k \geq t} r(s_k - t)[P(X_n = s_k) - P(X = s_k)] \tag{20}$$

and hence if $\mathcal{S}$ is finite, the result follows immediately by finite summmation of limits. To prove the result when $\mathcal{S}$ is infinite, assume that the regret function $r(\cdot)$ is bounded by B. Given any $\epsilon > 0$, find K such that

$$P(X > s_K) < \epsilon; \tag{21}$$

this would also imply that

$$P(X_n > s_K) < 2\epsilon, \tag{22}$$

for large n. Now from (20),

$$\begin{aligned} a_n(t) &= \sum_{t \leq s_k \leq s_K} r(s_k - t)[P(X_n = s_k) - P(X = s_k)] \\ &\qquad + \sum_{k=K+1}^{\infty} r(s_k - t)[P(X_n = s_k) - P(X = s_k)] \\ &\leq \sum_{t \leq s_k \leq s_K} r(s_k - t)[P(X_n = s_k) - P(X = s_k)] + B \times 2\epsilon, \end{aligned}$$

for large n, by (21), (22) and boundedness of $r(\cdot)$. Since ϵ is arbitrary it follows from (19) that $a_n(t) \to 0$.

(B) Suppose X_n takes $K(<\infty)$ values : $s_{n1} < \cdots < s_{nK}$ and the K values of X are $s_1 < \cdots < s_K$. Then from the weak convergence of X_n to X, it follows that for each $k = 1, 2, \ldots K$,

$$s_{nk} \to s_k \qquad \text{and} \qquad \mathbb{P}(X_n = s_{nk}) \to \mathbb{P}(X = s_k) \qquad \text{as } n \to \infty.$$

Then the continuity of $r(\cdot)$ results in

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[r(X_n - t)\mathbb{I}_{\{X_n \geq t\}}\right] &= \lim_{n\to\infty} \sum_{k=1}^{K} r(s_{nk} - t)P(X_n = s_{nk})\mathbb{I}_{[t,\infty)}(s_{nk}) \\
&= \sum_{k=1}^{K} \lim_{n\to\infty} \left[r(s_{nk} - t)P(X_n = s_{nk})\mathbb{I}_{[t,\infty)}(s_{nk})\right] \\
&= \sum_{k=1}^{K} r(s_k - t)P(X = s_k)\mathbb{I}_{[t,\infty)}(s_k) \;=\; \mathbb{E}\left[r(X - t)\mathbb{I}_{\{X \geq t\}}\right],
\end{aligned}$$

completing the proof that $a_n(t) \to 0$.

In the continuous case note that

$$\begin{aligned}
|a_n(t)| &= \left|\mathbb{E}\left[r(X_n - t)\mathbb{I}_{\{X_n \geq t\}}\right] - \mathbb{E}\left[r(X - t)\mathbb{I}_{\{X \geq t\}}\right]\right| \\
&= \left|\int_t^\infty r(x - t)\left[h_n(x) - h(x)\right] dx\right|. \qquad (23)
\end{aligned}$$

(C) Suppose $r(\cdot)$ is bounded by B. Then from (23) we have

$$\lim_{n\to\infty} |a_n(t)| \;\leq\; B \lim_{n\to\infty} \int_{-\infty}^{\infty} |h_n(y) - h(y)|\, dy \;=\; 0.$$

The last equality follows from the fact that h_n's and h are density functions and hence pointwise converges implies L^1 convergence (see, for example, [1] Theorem 16.11).

(D) Let the (identical) finite support h_n's and h be $[L, U]$. Since $r(\cdot)$ is continuous, it is bounded by B, say, on $[L, U]$. Hence from (23) we have

$$\lim_{n\to\infty} |a_n(t)| \;\leq\; B \lim_{n\to\infty} \int_L^U |h_n(y) - h(y)|\, dy \;=\; 0.$$

(E) In this case we have, from (23),

$$|a_n(t)| \;\leq\; \sup_x |h_n(x) - h(x)| \int_0^\infty r(y)\, dy \;\to\; 0,$$

by uniform convergence of h_n and integrability of $r(\cdot)$.

The following conjecture is a possible result in the the same direction as in Theorem 6, when the measure of homogenous skewness vanishes in the limit.

Conjecture 1 *Consider a sequence of stochastic DOPs, each having a single random element with cost X_n. Let the random element be homogeneously right-skewed with the density function $(h_n(\cdot))$ and mode M_n (converging to M) with decreasing order of skewness (partial order determined by the skewness function), i.e.*

$$\gamma_n(x) \geq \gamma_{n+1}(x), \quad \forall x, \forall n, \quad \text{where} \quad \gamma_n(x) = h_n(M_n + x) - h_n(M_n - x).$$

Then $\forall x$, the sequence $\gamma_n(x)$ has a (nonnegative) limit, say $\gamma(x)$, and if the limit of the homogeneous skewness measure (which has to exist)

$$\tau = \lim \int \gamma_n(x)dx = \int \gamma(x)dx$$

is 0, then θ_n, satisfying $\Psi_{r,X_n}(\theta_n) = 0$, must converge to M as $n \to \infty$. Recall that by Theorem 3, $\theta_n \geq M_n$.

4 DOPs with Multiple random elements

We now consider a DOP instance π with r (more than one) random elements. Accordingly, we partition G into $G_R = \{e_1, \ldots, e_r\}$ of random elements, and $G_F = \{e_{r+1}, \ldots, e_{r+f}\}$ of fixed elements. Let $X_1, \ldots, X_r$ be the random variables denoting the values of $c_{e_1}, \ldots, c_{e_r}$ and $H(x_1, \ldots, X_r)$ denote $Pr(X_1 \leq x_1, \ldots, X_r \leq X_r)$. We represent the objective function value of any solution S as

$$z(S) = F(S) + \sum_{i:e_i \in S \cap G_R} X_i \tag{24}$$

where $F(S) = \sum_{e \in S \cap G_F} c_e$ is the fixed component of the cost $z(S)$.

Let $K_1, \ldots, K_{2^r}$ be the 2^r subsets of $K = \{1, \ldots, r\}$. For $i = 1, \ldots, 2^r$, let

$$\mathbb{S}_i = \{S : S \in \mathbb{S}; \quad e_j \in S \quad \forall j \in K_i; \quad e_j \notin S \quad \forall j \in K \setminus K_i,\} \tag{25}$$

constitute a partition of $\mathbb{S}$. In certain problem situations, some of the $\mathbb{S}_i$'s may be empty.

Lemma 1 *If $S^1, S^2 \in \mathbb{S}_i$, for some i, then $z(S^1) - z(S^2)$ is non-random.*

Proof: By construction (25), S^1 and S^2 have the same set of random elements and hence by (24), $z(S^1) - z(S^2) = F(S^1) - F(S^2)$ which is non-random. ∎

Definition 11 For $i = 1, \ldots, 2^r$, denote the least cost (risk) solution within $\mathbb{S}_i$ as S_i, i.e.,

$$z(S_i) = \min_{S \in \mathbb{S}_i} z(S).$$

Remark 6 While the S_i in Definition 11 need not be unique, its existence (whenever $\mathbb{S}_i \neq \emptyset$) is ensured by Lemma 1, since the least cost solution within $\mathbb{S}_i$ can not change with randomness of $(X_1, \ldots, X_r)$.

Definition 12 We now introduce the sets $\{\mathcal{R}_i; 1 \leq i \leq 2^r\}$ in the r-dimensional Euclidean space ($\Re^r$) as follows. $\mathcal{R}_i = \emptyset$ if $\mathbb{S}_i = \emptyset$, otherwise

$$\mathcal{R}_i = \{(x_1, \ldots x_r) : S_i \text{ is a least cost solution at } (x_1, \ldots x_r)\}. \tag{26}$$

We define a partition of $\Re^r$ through $\{P_i; 1 \leq i \leq 2^r\}$ where

$$P_1 = \mathcal{R}_1, \quad \text{and } P_i = \mathcal{R}_i \setminus (\bigcup_{j<i} P_j) \quad i = 2, \ldots, 2^r. \tag{27}$$

Notice that for all $i = 1, \ldots, 2^r$, $P_i \subseteq \mathcal{R}_i$.

We are now in a position to prove the main theorem for DOPs with more than one random elements.

Theorem 7 *Consider a DOP with r random elements $e_1, e_2, \ldots, e_r$ having costs $X_1, \ldots, X_r$ which are random variables. Then the least risk solution (under general loss function) of the DOP will be one of the S_i's, as introduced in Definition 11, and their risks are given by*

$$R(S_i) = \sum_{j=1}^{2^r} \int_{P_j} r(z(S_i) - z(S_j))dH(.), \tag{28}$$

where P_j's are as defined in (27).

Proof: We first show that,

$$R(S) \geq \min_j\{R(S_j)\} \quad \text{for any } S \in \mathbb{S}. \tag{29}$$

Consequently, at least one among S_1 through S_{2^r}, is an optimal solution in the minimum risk sense. To prove (29), note that $\exists\ j$ such that $S \in \mathbb{S}_j$. Then

$$R(S) = \mathbf{E}\, r(z(S) - Z^*) = \mathbf{E}\, r(z(S) - z(S_j) + z(S_j) - Z^*) \geq \mathbf{E}\, r(z(S_j) - Z^*) = R(S_j),$$

with the inequality following from the fact that $r(\cdot)$ is increasing and $z(S) - z(S_j)$ is non-random and nonnegative.

The expression (28) follows from the definition of risk and the fact that S_j is the least cost solution when the random cost is in P_j. ∎

Remark 7 By (24),

$$\begin{aligned} z(S_i) - z(S_j) &= F(S_i) + \sum_{m \in K_i} x_m - \Big[F(S_j) + \sum_{m \in K_j} x_m\Big] \\ &= \Big[\sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m\Big] + F(S_i) - F(S_j). \end{aligned} \tag{30}$$

If S_i is a least cost solution at $(x_1, \ldots X_r)$, then for this set of costs, $z(S_i) \leq z(S_j), \quad \forall j = 1, \ldots, 2^r$. Hence, an alternative characterization of P_i is

$$P_i = \Big\{(x_1, \ldots, X_r) : \sum_{m \in K_i \setminus K_j} x_m - \sum_{m \in K_j \setminus K_i} x_m \leq F(S_j) - F(S_i), \quad j = 1, \ldots, 2^r\Big\} \tag{31}$$

Theorem 8 *Consider a stochastic DOP with r random elements having costs $X_1, \ldots, X_r$. If X_i's are random variables having finite means μ_i, $i = 1, \ldots, r$ respectively, then the least cost solution of the non-stochastic DOP corresponding to the costs of $c_{e_1}, \ldots, c_{e_r}$ fixed at $\mu_1, \ldots, \mu_r$, will be an optimal solution of the stochastic DOP in the least risk sense for any linear regret function of the form*

$$r(t) = \alpha + \beta t, \text{where } \beta > 0.$$

Proof: Under the linear regret, minimizing $R(S)$ is equivalent to minimizing $\mathbb{E}z(S)$ which by (24) reduces to minimizing

$$F(S) + \sum_{i: e_i \in S \cap G_R} \mu_i.$$

This is the same as finding the least cost feasible solution of the DOP with random costs being replaced by their mean values. ∎

However, for non-linear regret function, it is not enough to replace the random costs by their respective averages, as shown even in the single random case. It is natural to explore if Theorem 2 can be extended for symmetrically distributed multiple random elements. In an attempt to provide reasonably complete answer to such a possible extensions, we now explore a series of examples and associated results from simulation exercises that lead to partial results and conjecture. In these examples, the (r) random elements are assumed to have costs $X_1, \ldots, X_r$ (with X_i having mean μ_i) as before, while the f fixed elements have costs $c_1 < \ldots < c_f$. We also use the notation $S_{i_1, \ldots, i_l}^{j_1, \ldots, j_t}$ to denote the least cost (risk) solution among solutions containing $X_{i_1}, \ldots, X_{i_l}$, but not containing any of the $X_{j_1}, \ldots, X_{j_t}$; the suffix or superfix in such notation for a solution may be omitted, if it is obvious from the context.

In addition, in many of the examples, we consider a balanced DOP (refer to Definition 7), where the minimum cardinality of a feasible solution is k. Note that any optimal solution (in any sense) would have cardinality exactly k in that case.

Example 1 Consider a balanced stochastic DOP with $r = 2, f = 1 = k$. Suppose now that X_1 has a V-shaped distribution, supported on (0,1) having density

$$f(x) = \begin{cases} 4|x - 0.5| & \text{if } 0 < x < 1, \\ 0 & \text{otherwise }, \end{cases} \tag{32}$$

while $X_2 = 1 - X_1$. Thus the two random costs are identically and symmetrically distributed random variables, which are strongly dependent. The three candidates for an optimal (least risk) solution S_1, S_2 and $S^{1,2}$ have costs X_1, X_2 and c_1 respectively. For $c_1 = 0.5$, $Z^* = \min(X_1, X_2)$. It is easy to check that, with the regret function $r(t) = t^2$, the risks of the three solutions are as follows:

$$R(S_1) = R(S_2) = 0.25, \quad R(S^{1,2}) = 0.125.$$

If c_1 is marginally higher, say equal to 0.501, the risk the random edge solutions stay (almost) unchanged while the $S^{1,2}$ now has a risk of 0.125668. Thus, the least cost solution to the corresponding non-stochastic DOP (with the random edges replaced by fixed edges with respective average costs) is no longer the least risk solution of the stochastic DOP. Indeed, one would continue to observe this 'counter-example' for $c_1 \in (0.5, 0.65)$.

The above illustration is possibly more pronounced because of somewhat unusual feature of the V-distribution. However this is hardly specific to the distribution considered. For example, if we consider a symmetric triangular distribution on (0,1), i.e. having a density:

$$f(x) = \begin{cases} 4x & \text{if } 0 < x \leq 0.5, \\ 4(1-x) & \text{if } 0.5 < x < 1, \\ 0 & \text{otherwise }, \end{cases} \tag{33}$$

we observe the same phenomenon for $c_1 \in (0.5, 0.6)$.

Example 2 Consider a balanced stochastic DOP with $r = f = k = 2$. The two random elements are assumed to have costs that are independently distributed, with X_1 following the V-distribution and X_2 following the Triangular distribution as specified in (32) and (33), or both X_1 and X_2 following the Triangular distribution. Tables 9 through 10 report the risks of these four candidate solutions, for some randomly generated values of c_1, c_2 for the regret function $r(t) = t^2$.

These examples show that for SDOP's with multiple random elements, extension of Theorem 2 is, in general, not true, i.e., it is not enough to replace the random costs by the respective average costs to arrive at the optimal solution irrespective of whether the symmetrically distributed random elements are

- dependent or independent of each other;

- identically distributed or otherwise.

The results from these simulation exercises, however, indicate possibility of a partial result that can be considered as a weak extension of Theorem 2 in the multidimensional case. Towards that direction, we put forward the following theorem in the two-dimensional case (2 random elements).

Theorem 9 *Consider a 2-dimensional balanced stochastic DOP ($r = 2$). Suppose the random costs X_1, and X_2 are independent and symmetrically distributed with mean μ_1 and μ_2, respectively. Further assume that the regret function satisfies the following growth condition*

$$r(b+\beta) - r(b) \geq r(a+\alpha) - r(a) \quad \textit{for } b \geq a \geq 0 \textit{ and } \beta \geq \alpha \geq 0. \tag{34}$$

Then in each of the following situations

$$(a)\ f = 1 = k \qquad (b)\ f = 1,\ k = 2 \qquad (c)\ f = 2 = k$$

the following holds:

(i) If $(\mu_1, \mu_2) \in \mathcal{R}^{1,2}$ then $S^{1,2}$ is a least risk solution.

(ii) If $(\mu_1, \mu_2) \in \mathcal{R}_{1,2}$ then $S_{1,2}$ is a least risk solution.

Proof: The proof is given in the Appendix 5.4. ■

Remark 8 The theorem is true in a more general setup, namely when $U_i := X_i - \mu_i,\ i = 1,2$ satisfy

$$(U_1, U_2) \stackrel{D}{=} (-U_1, U_2) \stackrel{D}{=} (U_1, -U_2) \stackrel{D}{=} (-U_1, -U_2) \tag{35}$$

which is true when X_1, and X_2 are independent and symmetrically distributed with mean μ_1 and μ_2.

Remark 9 The growth condition (34) is satisfied by commonly used regret functions such as $r(t) = t^n$; $r(t) = (1+t)^n - 1$ and $r(t) = \exp(\lambda t) - 1,\quad$ where $\lambda > 0$.

Remark 10 The reverse statement of the theorem is not true and consequently

$(\mu_1, \mu_2) \in P_1^2 \not\Rightarrow S_1^2$ is a least risk solution, and

$(\mu_1, \mu_2) \in P_2^1 \not\Rightarrow S_2^1$ is a least risk solution.

These can be seen from the last instances in Tables 9 through 10.

The conclusion of Theorem 9 does not hold true if the two random elements are dependent on each other. This is evident from the simulation study reported in the Appendix Table 11. This is in the framework of Example 2 with (X_1, X_2) following Bivariate Normal distribution with identical mean 10 and standard deviation 2 having a correlation of 0.5. At the same time, simulation result supports the validity of the result when the variables are independent and normal as seen in the Appendix Table 12. The theorem leads us to the following conjecture which is also supported by additional computations via simulation, as detailed below.

Conjecture 2 *Consider a balanced stochastic DOP with r random elements, having costs $X_1, \ldots, X_r$ that are independent and symmetric random variables with finite means μ_i, $i = 1, \ldots, r$ respectively, and f fixed elements with costs $c_1 \le c_2 \le \ldots \le c_f$. Then the least risk solution will consist only of the*

- *fixed elements (when feasible)* *if* $\min_{1 \le i \le r} \mu_i > c_f$,
- *random elements (when feasible)* *if* $\max_{1 \le i \le r} \mu_i < c_1$.

Remark 11 Without the 'balanced' condition the conjecture need not be true. To see this consider a DOP with 3 elements R, F_1, F_2, the first of them being random while the other two are fixed; let their costs be X (random), c_1 and c_2 respectively, such that $c_1 + c_2 > \mu(= E(X)) > c_2 > c_1$. Suppose $S = \{\{R\}, \{F_1, F_2\}\}$. For symmetric X the least risk solution is then $\{R\}$ (see Theorem 1 in section 3) which violates the conclusion of Conjecture 2. This stochastic DOP is not balanced.

Remark 12 Only the following cases are relevant for the conjecture :

$$(a)\ r > f = k \quad (b)\ r = k > f \text{ and } (c)\ r = f = k.$$

To see this, first observe that we may assume, without loss of any generality, $f \le k$. Further, if $r < k$, then the optimal solution would necessarily contain at least $(k - r)$ fixed elements with costs $c_1, \cdots, c_{k-r}$, and consequently the problem can be redefined in terms of a balanced DOP with $r = k > f$. Finally, if $r > k$ and $f < k$, then neither $S_{1,2,\ldots,r}$ nor $S^{1,2\ldots,r}$ is feasible.

Remark 13 In light of Remark 12, conjecture 2 reduces to Theorem 9 when $r = 2$.

The scope and validity of the conjecture has been examined on the basis of fairly extensive simulation study. Appendix Table 13 reports the risks of the candidate solutions in the balanced DOP with $r = 2 = k, f = 1$ with random elements having independent Normal distribution, where as Table 14 and Table 16 respectively report the same for $r = 3, f = 2 = k$ and $r = 3 = k, f = 2$. Appendix Table 15 considers the $r = f = k = 3$ case with the three random elements having three different types of symmetric distributions, but are independent on each other. Each of these tables reports only selected cases of much larger (1000+) simulation runs.

Another (partial) result is also obtained which has significance in the context of a (symmetric) Traveling Salesperson Problem. Suppose The salesperson has to cover n cities in a given order. From every city to the next in order there are two routes. One route has known cost and the other random. An optimal route (in the sense of minimum cost) for the salesperson has to be decided before the journey begins. The following theorem says that if the random costs are independent of each other and symmetrically distributed then the pairwise optimal choices (i.e., the optimal route between the fixed cost one and the random cost one connecting every pair of cities) constitute the optimal route for the whole journey of the salesperson.

Theorem 10 *Consider a stochastic DOP with n random elements having costs X_i's and n fixed elements with costs c_i's, for $i = 1, \ldots, n$. Suppose further that in this DOP, a solution is feasible if and only if it contains one (and only one) from each of the pairs $(X_1, c_1), \ldots, (X_n, c_n)$. If the X_i's are independent random variables having symmetric distributions with mean μ_i's, then a feasible solution containing only those X_i's satisfying $\mu_i < c_i$ is a least risk solution.*

Proof: The proof is given in the Appendix 5.5. ∎

Remark 14 The theorem holds in a more general setup when $U_i := X_i - \mu_i,\ i = 1, 2, \cdots n$ satisfy for any $i_j = 0, 1,\ j = 1, 2, \ldots, n$,

$$((-1)^{i_1}U_1, (-1)^{i_2}U_2, \ldots, (-1)^{i_n}U_n) \overset{D}{=} (U_1, U_2, \ldots, U_n) \tag{36}$$

which is true when X_i's are independent and symmetrically distributed with mean μ_i.

Remark 15 The assumption about symmetrical distribution is critical for the validity of Theorem 10. First of all without symmetry the surrogate parameters are no longer (necessarily) the means μ_i's. Perhaps more critically, as the following example shows, without the symmetry assumption, one may not get an optimal solution by combining the optimal solutions of the (independent) one-dimensional stochastic DOP's (i.e., having single random element). While this may appear to be counter-intuitive, this is a natural manifestation of nonlinear regret function. To illustrate this take $n = 2$ and consider

$$\pi = (G = \{X_1, c_1, X_2, c_2\}, \mathcal{S} = \{\{X_1, X_2\}, \{X_1, c_2\}, \{c_1, X_2\}, \{c_1, c_2\}\}, z)$$

where $X_1 \sim$ Beta$(1, 1.5)$; $c_1 = 0.401$; $X_2 \sim$ Beta$(1, 2)$; $c_2 = 0.35$. Now consider the two sub-DOP's each with single random element, viz.,

$$\pi_i = (G_i = \{X_i, c_i\}, \mathcal{S}_i = \{\{X_i\}, \{c_i\}\}, z),\ i = 1, 2;\ r(t) = t^2.$$

Then

$$R_1(c_1) = 0.0306,\ R_1(X_1) = 0.0380,\ R_1(c_2) = 0.0261,\ R_1(X_2) = 0.0298,$$

i.e., c_1 is better than X_1 and c_2 is better than X_2 but for SDOP π $\{c_1, X_2\}$ is optimal, since

$$R_2(c_1, c_2) = 0.0809,\ R_2(c_1, X_2) = 0.0808,\ R_2(X_1, c_2) = 0.0881,\ R_2(X_1, X_2) = 0.0881.$$

5 Computational Experience

In this section we report our experiences with algorithms that we designed to obtain minimum risk solutions to combinatorial optimization problems. We chose to implement our algorithms to obtain minimum risk solutions to the 0-1 knapsack problem (01KP). The algorithms were coded in C and compiled and run on a 2.8GHz personal computer with 512MB RAM running Linux.

5.1 Description of the Problem Sets

For the 01KP, we chose problems with N random elements, and 10 fixed (non-random) elements. The marginal distributions for each of the N random element were discrete, each having a pre-specified number P support points. Therefore, each of our problems therefore have 2^N candidate optimal solutions, and P^N support points in the joint distribution of the random elements.

In our experiments, we considered two situations, one in which the N random elements were independent, and the second in which they were dependent. To facilitate comparison, for each problem instance in the dependent case, the joint distributions of the random elements were generated keeping the marginal distribution for each random element identical with the marginal distribution of the element with the same index in the corresponding independent case.

We experimented with the (N, P) pairs (6, 6), (6, 8), (6, 10), (8, 4), and (8, 6). Problems smaller than these were too trivial to be interesting computationally, while problems larger than these took exorbitant amounts of solution time. For each of the (N, P) pairs that we chose, we generated ten instances. Each instance consists of the profit and cost values of all the fixed elements, the non-random cost values of the random elements, and the joint distribution of the profits of the random elements. The collection of these ten instances is called a set.The performance of an algorithm on any instance is measured by the suboptimality of the solution it generates — the higher the value of suboptimality, the worse the performance. The performance of an algorithm on any of the sets is measured by the average of the performance of the algorithm on all the ten problems in the set. Table 1 presents the size of the search problem for our chosen values of N and P.

Table 1: Size of the search problem

N	P	2^N	P^N
6	6	64	46656
6	8	64	262144
6	10	64	1000000
8	4	256	65536
8	6	256	1679616

5.2 Description of the Algorithms

For knapsack problems, we do not have an efficient representation of the regions in the solution space in which a particular solution is the minimum cost one. We are therefore unable to make use of (28) in our implementations; instead we adopt one of the two methods described below to compute the risk of any solution.

Generate All Support Points (GASP) In this method, all support points of the joint distribution are generated. The objective function values of each of the candidate optimum solutions are obtained for the support point, and the maximum of the solution values is retained as the best solution value achievable at that point. In order to compute the risk associated with any solution, the objective function value of the solution is computed at each point in the support of the joint distribution, and the suboptimality of the solution at that support point is computed, making use of the retained best solution value at that point. The expected value of suboptimality of that solution is then computed as the risk of the solution.

Monte-Carlo (MC) In this process, a simple random sample of a pre-specified number (s) of points are generated in the support of the joint distribution. As in GASP, the objective function values of each of the candidate optimum solutions are obtained for the points in the sample, and the maximum of the solution values is retained as the best solution value achievable at that point. In order to compute the risk associated with any solution, the objective function value of the solution is computed at each of the sampled points, and the suboptimality of the solution at that point is computed. The suboptimalities of the solution at each sample point are added up and appropriately scaled to provide a measure of the risk of the solution. Needless to say, this is an approximate value of the risk of the solution.

We also implement the following two ways of searching for a least risk solution among the candidate optimal solutions.

Complete Enumeration (CE) In this method, we simply evaluate the risk associated with each of the 2^N candidate optimal solutions, and choose one with the minimum risk value. This method is extremely time consuming, and is appropriate only for very small problems. However, it is also an assured method of finding a least risk solution when combined with GASP, and can be used to benchmark the performance of other algorithms.

Tabu Search (TS) Tabu Search is a well-known method of obtaining high-quality solutions to large combinatorial optimization problem. It is an extension of the local improvement algorithm. The pseudocode below describes the procedure.

Procedure Tabu Search

Step 0 (Initialize) : Choose a solution as the *current* solution. Generate an empty list *TABU*. Let *BestSolution* ← *current*.

Step 1 (Terminate) : If a user-specified termination condition is reached, output *BestSolution* and exit.

Step 2 (Search) : Search the neighborhood of *current*. Let *BestNontabu* be the best solution in the neighborhood that can be reached from the current solution without the aid of any move in the *TABU* list. Also let *BestTabu* be the best solution in the neighborhood that can be reached from the current solution using a move in the *TABU* list.

Step 3a (Aspirate) : If *BestTabu* is better than both *BestNontabu* and *BestSolution* then let *BestSolution* ← *BestTabu*, *current* ← *BestTabu*, and empty the *TABU* list. Go to Step 1.

Step 3b (Move) : Choose a value of *tenure* and add the move from *current* to *BestNontabu* to the *TABU* list for a period of *tenure* iterations. Let *current* ← *BestNontabu*. If *BestNontabu* is better than *BestSolution* then let *BestSolution* ← *BestNontabu*.

Step 4 (Update) : Update the *TABU* list by removing the moves that have already been in the *TABU* list for their prescribed tenure. Go to Step 1.◇

In our implementations, two solutions are said to be neighbors if the sets of random elements in the two solutions differ by at most two elements. The solution chosen as the initial *current* solution is an optimal solution to the 01KP instance obtained by setting the profit value of each of the random elements to the expected value of its marginal distribution. The *tenure* value is chosen as a random integer between $\frac{N}{2}$ and $\frac{2N}{3}$. The termination criterion was based on the execution time alloted for the search, and was set between 15 CPU seconds and 250 CPU seconds depending on the N and P values.

Given that we have two methods for computing the risk of a solution and two methods for search a least risk solution, we define four algorithms, GASP-CE, which uses GASP to compute the risk of each individual solution, and CE to obtain a least risk solution; GASP-TS, which uses GASP to compute the risk of each individual solution, and TS to obtain a least risk solution; MC-CE, which uses MC to compute the risk of each individual solution, and CE to obtain a least risk solution; and MC-TS, which uses MC to compute the risk of each individual solution, and TS to obtain a least risk solution. Of these we recommend GASP-CE for instances with low N and P values, GASP-TS for instances with low P values and moderate N values, MC-CE for instances with moderate N and P values, and MC-TS for instances with high N and P values.

5.3 Results from Computations

We first report our experience with the execution times required by the four algorithms on our problem sets. The time required by the algorithms can be broken up into two components, the time required by the algorithms to generate the support points (i.e., the GASP and the MC components) and the time required to search for the least risk solution among the candidate solutions (i.e., the CE and TS components). Table 2 reports the times taken by the copmponents of the GASP-CE algorithm, while Table 3 reports the times required by the components of the MC-CE algorithm with different cardinalities of the support (denoted by s). The time required by any of the components on an instance in the dependent case was never found to be significantly different from the time required by that component on the corresponding problem instance in the independent case; therefore we report an average of these times in our tables.

Table 2: Execution times required by GASP-CE (in CPU seconds)

N	P	GASP	CE
6	6	1.625	2.475
6	8	9.094	13.871
6	10	34.759	53.005
8	4	11.558	18.641
8	6	280.255	433.200

Table 3: Execution times required by MC-CE (in CPU seconds)

N	P	s=3000		s=4000		s=5000	
		MC	CE	MC	CE	MC	CE
6	6	34.061	0.037	45.127	0.056	56.323	0.078
6	8	37.469	0.037	49.552	0.055	60.799	0.079
6	10	49.333	0.037	64.981	0.057	79.896	0.077
8	4	132.237	0.154	176.485	0.236	220.763	0.330
8	6	159.400	0.155	208.820	0.240	258.795	0.329

In our experiments with GASP-TS and MC-TS, we found that GASP-TS and MC-TS used the same time as GASP-CE and MC-CE respectively to generate all support points and compute the maximum profit solution at each support point. The

TS procedure took less that 0.001 seconds in all cases for GASP-TS and the maximum time limit set for MC-TS to generate the solution that the algorithms finally output.

Note that GASP-CE is an exact algorithm in the sense that it outputs the least risk solution to the problem. The three other algorithms output solutions which are not necessarily minimum risk (but hopefully low risk). The suboptimality of a solution output by any algorithm (say $\mathcal{A}$) to an instance is computed as

$$\text{suboptimality} = \frac{R(x^{\mathcal{A}}) - R(x^*)}{R(x^*)} \qquad (37)$$

where x^* is the least risk solution for the instance, and $x^{\mathcal{A}}$ is the solution output by the algorithm $\mathcal{A}$.

The performance of GASP-TS was very encouraging for the problems we tested it on. Each of the solutions that it output was found to be optimal. Hence we can conclude that, at least for these problem sizes, GASP-TS clearly outperforms GASP-CE in terms of execution times, without sacrificing solution quality.

In case of MC-CE and MC-TS algorithms we chose three values of s for our experiments, viz. 3000, 4000, and 5000. Since we take a random sample for the MC procedure, we performed 25 runs for each problem instance and chose the average of the suboptimality values over all 25 runs as the suboptimality of the MC-CE algorithm for each instance. Tables 4 and 5 contain the results of our experiments

Table 4: Quality of solutions output by MC-CE and MC-TS when random elements are independent

N	P	s	MC-CE		MC-TS	
			mean	s.d.	mean	s.d.
6	6	3000	1.7354	0.0060	0.4826	0.0000
6	8	3000	1.9721	0.0048	0.5169	0.0000
6	10	3000	1.9564	0.0010	0.4205	0.0000
8	4	3000	0.0230	0.0100	0.0770	0.0000
8	6	3000	0.0194	0.0020	0.1775	0.0000
6	6	4000	1.7347	0.0056	0.4588	0.0000
6	8	4000	1.9697	0.0022	0.4964	0.0000
6	10	4000	1.9562	0.0000	0.4205	0.0000
8	4	4000	0.0230	0.0107	0.0659	0.0001
8	6	4000	0.0198	0.0034	0.2039	0.0000
6	6	5000	1.7337	0.0032	0.2672	0.0000
6	8	5000	1.9716	0.0038	0.5169	0.0000
6	10	5000	1.9564	0.0010	0.4205	0.0000
8	4	5000	0.0236	0.0083	0.0770	0.0000
8	6	5000	0.0191	0.0025	0.2039	0.0000

with MC-CE and MC-TS on the problem sets. The mean suboptimality value for

Table 5: Quality of solutions output by MC-CE and MC-TS when random elements are dependent

N	P	s	MC-CE		MC-TS	
			mean	s.d.	mean	s.d.
6	6	3000	1.7395	0.0077	0.4835	0.0000
6	8	3000	1.9724	0.0052	0.5169	0.0000
6	10	3000	1.9563	0.0010	0.4205	0.0000
8	4	3000	0.0221	0.0076	0.0766	0.0000
8	6	3000	0.0188	0.0044	0.1773	0.0000
6	6	4000	1.7380	0.0058	0.4595	0.0000
6	8	4000	1.9700	0.0027	0.4965	0.0000
6	10	4000	1.9561	0.0000	0.4205	0.0000
8	4	4000	0.0206	0.0106	0.0657	0.0001
8	6	4000	0.0197	0.0027	0.2005	0.0000
6	6	5000	1.7376	0.0056	0.2661	0.0000
6	8	5000	1.9717	0.0038	0.5169	0.0000
6	10	5000	1.9563	0.0010	0.4205	0.0000
8	4	5000	0.0226	0.0079	0.0766	0.0000
8	6	5000	0.0182	0.0024	0.2005	0.0000

solutions output by MC-CE for problems with 6 random elements is seen to be high when compared to those for solutions output by MC-TS. On inspection of the results for individual instances, this high mean suboptimality is seen to be caused by one problem instance in the set. If we remove this problem instance from the sets, then the mean suboptimality values of the solutions output by MC-CE are seen to be of the same order as those for the solutions output by MC-TS. Changing the set of points in the sample of support points chosen by MC-CE for this problem however, did not remove this anomaly.

From the standard deviation values seen in the tables, it seems that MC-CE is sensitive to the samples of support points chosen by the Monte-Carlo method, while MC-TS is not. However, on examination of the results for individual problem instances, MC-CE is seen to be sensitive to the choice of sample points in 1 or 2 of the instances in the sets with 6 random elements, and in 5 or 6 of the instances in the sets with 8 random elements.

We do not see any consistent improvement in the quality of solutions when the sample size is increased. This is surprising, although we think that such an improvement will be observed when the sample size increases significantly. The quality of solutions output by the algorithms when the random elements are independent is not significantly different from when the random elements are dependent. This is expected, since the algorithms do not make use of the property of independence (or otherwise) of the marginal distributions.

References

[1] Billingsley, P. (1985), *Probability and Measure*, second edition, John Wiley, New York.

[2] Caraway, R., Schmidt, S., and Weatherford, L. (1993), An Algorithm for Maximizing Target Achievement in the Stochastic Knapsack Problem with Normal Returns, *Naval Research Logistics*, **40**, 161-173.

[3] Chakravarti, N. and Wagelmans, A.P.M. (1998), Calculation of stability radii for combinatorial optimization problems, *Operations Research Letters*, **23**, 1–7.

[4] Das, S. and Ghosh, D. (2003), Binary Knapsack problems with random budgets, *J. of the OR Society*, **54**, 970–983.

[5] Dean, B.C., Goemans, M.X. and Vondrák, J. (2004), Approximating the Stochastic Knapsack Problem: The Benefit of Adaptivity, Proceedings of the 45th Anuual IEEE Symposium on *Foundations of Computer Science*, 208–217.

[6] Fishburn, P. C. (1968), Utility theory, *Management Sci.*, **14**(5, Theory Series), 335-378.

[7] Frank, H. (1969), Shortest paths in probabilistic graphs. *Operations Research*, **17**, 583–599.

[8] Greenberg, H.J. (1998), An annotated bibliography for post-solution analysis in mixed integer programming and combinatorial optimization, in *Advances in Computational and Stochastic Optimization, Logic Programming, and Heuristic Search* (Gal and Greenberg eds.), 97–148, Kluwer Academic Publishers.

[9] Henig, M.I. (1990), Risk criteria in a stochastic knapsack problem, *Operations Research*, **38**(**5**), 820–825.

[10] Klein Haneveld, W.K. and Van der Vlerk, M.H. (2000), *Stochastic Programming* (Lecture Notes), Available on request, email *m.h.van.der.vlerk@eco.rug.nl*

[11] Van Hoesel, S. and Wagelmans, A.P.M. (1999), On the complexity of postoptimality analysis of 0/1 programs, *Discrete Applied Mathematics*, **91**, 251–263.

[12] Ishii, H., Shiode, S., and Nishida, T. (1981), Stochastic spanning tree problem, *Discrete Applied Mathematics*, **3**, 263–273.

[13] Kaliszewski, I. and Michalowski, W. (1998), Establishing Regret Attitude of a Decision Maker within the MCDM Modeling Framework, Interim Report IR-98-070/September *International Institute for Applied Systems Analysis.*

[14] Kravchenko, S.A., Sotskov, Y.N. and Werner, F. (1995), Optimal schedules with infinitely large stability radius, *Optimization*, **33**, 271–280.

[15] Loui, R.P. (1983), Optimal paths in graphs with stochastic or multidimensional weights, *Commun. ACM*, **26**, 670–676.

[16] Murthy, I. and Sarkar, S. (1998), Stochastic Shortest Path Problems with Piecewise-Linear Concave Utility Functions, *Management Science*, 44, 125–136.

[17] Ramaswamy, R. and N. Chakravarti (1995), Complexity of determining exact tolerances for min-sum and min-max combinatorial optimization problems, *WPS-247/95*, Indian Institute of Management, Calcutta, India.

[18] Savage, L.J. (1954), *The Foundations of Statistics*, John Wiley, New York.

[19] Sotskov, Y.N. (1989), Stability of high-speed optimal schedules, *Computational Mathematics and Mathematical Physics*, **29**, 57–63.

[20] Sotskov, Y.N. (1991), Stability of an optimal schedule, *European Journal of Operational Research*, **55**, 91–102.

[21] Sotskov, Y.N., Leontev, V.K., and Gordeev, E.N. (1995), Some concepts of stability analysis in combinatorial optimization, *Discrete Applied Mathematics*, **58**, 169–190.

[22] Sotskov, Y.N., Sotskova, N.Y., and Werner, F. (1997), Stability of an optimal schedule in a job shop, *Omega*, **25**, 397–414.

Appendix

5.4 Proof of Theorem 9

We shall give the proof under the general condition (35) of Remark 8. It is sufficient to compare $R(S^{1,2}), R(S_1^2), R(S_2^1)$, and $R(S_{1,2})$. Without loss of generality assume that $\mu_1 \le \mu_2$.

First consider the situation (a). Note that in this case
$\mathcal{R}_{1,2} = \emptyset,\ \mathcal{R}_1^2 = \{x_1 \le \min(c, x_2)\},\ \mathcal{R}_2^1 = \{x_2 \le c, x_2 \le x_1\},\ \mathcal{R}^{1,2} = \{x_1 \ge c, x_2 \ge c\}$.
Since $\mathcal{R}_{1,2} = \emptyset$, only statement (ii) is to be proved. So assume that the fixed cost $c \le \mu_1 \le \mu_2$. We need to show that $R(S^{1,2}) \le \min\{R(S_1^2), R(S_2^1)\}$. Note that in this case the partitions of $\mathbf{R}^2$, given in Definition 12, can be taken to be
$P_{1,2} = \emptyset,\ P_1^2 = \{x_1 \le \min(c, x_2)\},\ P_2^1 = \{x_2 \le c, x_2 < x_1\},\ P^{1,2} = \{x_1 > c, x_2 > c\}$.
Letting $u_i = x_i - \mu_i$ we can rewrite the partitions w.r.t. (u_1, u_2)-space as follows:

$$
\begin{aligned}
Q_1^2(u_1,u_2) &= \{u_1 \le \min(c-\mu_1, u_2+\mu_2-\mu_1)\},\\
Q_2^1(u_1,u_2) &= \{u_2 \le c-\mu_2, u_2 < u_1+\mu_1-\mu_2\},\\
Q^{1,2}(u_1,u_2) &= \{u_1 > c-\mu_1, u_2 > c-\mu_2\}.
\end{aligned}
$$

From (28) and (35) we then have,

$$\begin{aligned} R(S_1^2) &= E[r(X_1 - X_2)\mathbb{I}_{P_2^1}(X_1, X_2)] + E[r(X_1 - c)\mathbb{I}_{P^{1,2}}(X_1, X_2)] \\ &= E[r(U_1 + \mu_1 - U_2 - \mu_2)\mathbb{I}_{Q_2^1(U_1,U_2)}] + E[r(U_1 + \mu_1 - c)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] \\ &= E[r(U_1 + U_2 + \mu_1 - \mu_2)\mathbb{I}_{Q_2^1(U_1,-U_2)}] + E[r(U_1 + \mu_1 - c)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] \end{aligned}$$

and similarly

$$\begin{aligned} R(S_2^1) &= E[r(X_2 - X_1)\mathbb{I}_{P_1^2}(X_1, X_2)] + E[r(X_2 - c)\mathbb{I}_{P^{1,2}}(X_1, X_2)] \\ &= E[r(U_1 + U_2 + \mu_2 - \mu_1)\mathbb{I}_{Q_1^2(-U_1,U_2)}] + E[r(U_2 + \mu_2 - c)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] \\ R(S^{1,2}) &= E[r(c - X_1)\mathbb{I}_{P_1^2}(X_1, X_2)] + E[r(c - X_2)\mathbb{I}_{P_2^1}(X_1, X_2)] \\ &= E[r(U_1 + c - \mu_1)\mathbb{I}_{Q_1^2(-U_1,U_2)}] + E[r(U_2 + c - \mu_2)\mathbb{I}_{Q_2^1(U_1,-U_2)}]. \end{aligned}$$

Then the difference $R(S_1^2) - R(S^{1,2})$

$$\begin{aligned} &= E[r(U_1 + \mu_1 - c)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] - E[r(U_1 + c - \mu_1)\mathbb{I}_{Q_1^2(-U_1,U_2)}] \\ &\quad + E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\mathbb{I}_{Q_2^1(U_1,-U_2)}]. \end{aligned}$$

Note that (see Figure 8)

$$\begin{aligned} Q^{1,2}(U_1, U_2) &= \{U_1 > c - \mu_1, U_2 > c - \mu_2\} \\ &= \{c - \mu_1 < U_1 < \mu_1 - c, U_2 > c - \mu_2\} \cup \{U_1 \geq \mu_1 - c, U_2 > c - \mu_2\} \\ &= A \cup B, \text{ say,} \end{aligned}$$

$$\begin{aligned} Q_1^2(-U_1, U_2) &= \{U_1 \geq \max(\mu_1 - c, -U_2 - \mu_2 + \mu_1)\} \\ &= B \cup \{U_2 \leq c - \mu_2, U_1 \geq -U_2 - \mu_2 + \mu_1\} = B \cup C^{(2)}, \text{ say,} \end{aligned}$$

$$\begin{aligned} Q_2^1(U_1, -U_2) &= \{U_2 \geq \mu_2 - c, U_2 > -U_1 - \mu_1 + \mu_2\} \\ &= \{U_1 < c - \mu_1, U_2 > -U_1 - \mu_1 + \mu_2\} \\ &\quad \cup \{c - \mu_1 \leq U_1 \leq \mu_1 - c, U_2 \geq \mu_2 - c, U_1 + U_2 > -\mu_1 + \mu_2\} \\ &\quad \cup \{U_1 > \mu_1 - c, U_2 > U_1 - \mu_1 + \mu_2\} \cup \{U_2 \geq \mu_2 - c, U_1 \geq U_2 - \mu_2 + \mu_1\} \\ &= D^{(1)} \cup E \cup D \cup C, \text{ say.} \end{aligned}$$

Then $R(S_1^2) - R(S^{1,2})$ can be rewritten as

$$\begin{aligned} &= E[r(U_1 + \mu_1 - c)\mathbb{I}_A] + E[\{r(U_1 + \mu_1 - c) - r(U_1 + c - \mu_1)\}\mathbb{I}_B] \\ &\quad - E[r(U_1 + c - \mu_1)\mathbb{I}_{C^{(2)}}] \\ &\quad + E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\{\mathbb{I}_{D^{(1)}} + \mathbb{I}_E + \mathbb{I}_D + \mathbb{I}_C\}]. \end{aligned}$$

Clearly $E[r(U_1 + \mu_1 - c)\mathbb{I}_A] \geq 0$. Also, since r is increasing and $c \leq \mu_1$, we have $r(U_1 + \mu_1 - c) \geq r(U_1 + c - \mu_1)$ and hence $E[\{r(U_1 + \mu_1 - c) - r(U_1 + c - \mu_1)\}\mathbb{I}_B] \geq 0$. Now note that the region $C^{(2)}$ reflected along u_1-axis (i.e., replacing U_2 by $-U_2$) is

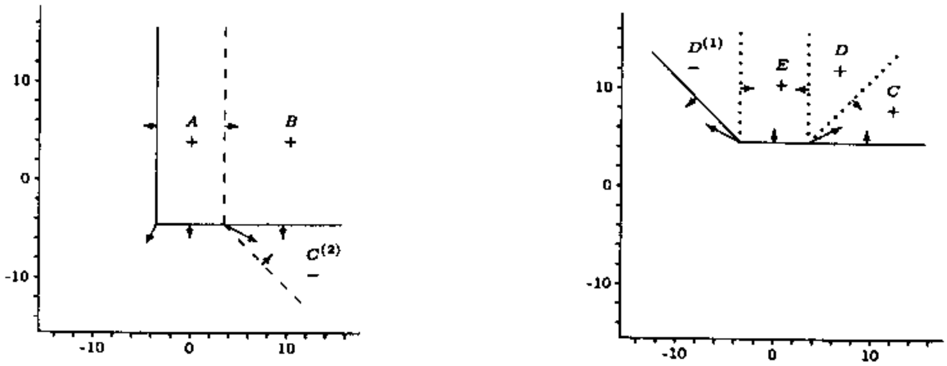

Figure 8: Decompositions of sets of integrations for $R(S_1^2) - R(S^{1,2})$. First plot is for sets involving $r(U_1 + \cdots)$ and the second for the (common) set involving $r(U_2 + \cdots)$ and $r(U_1 + U_2 + \cdots)$. Arrows indicate the (non)inclusion of boundary lines/points. Symbols '+'/'−' indicate the contribution from the integral over the corresponding set to the difference $R(S_1^2) - R(S^{1,2})$.

same as C and similarly, $D^{(1)}$ reflected along u_2-axis (i.e., replacing U_1 by $-U_1$) is same as D. Then using symmetry, we get

$$E[r(U_1 + c - \mu_1)\mathbb{I}_{C^{(2)}}] = E[r(U_1 + c - \mu_1)\mathbb{I}_C], \quad \text{and}$$

$$\begin{aligned} & E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\mathbb{I}_{D^{(1)}}] \\ = \; & E[\{r(-U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\mathbb{I}_D] \end{aligned}$$

Thus

$$\begin{aligned} & E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\mathbb{I}_C] - E[r(U_1 + c - \mu_1)\mathbb{I}_{C^{(2)}}] \\ = \; & E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2) - r(U_1 + c - \mu_1)\}\mathbb{I}_C] \geq 0 \end{aligned}$$

because, on $C = \{U_2 \geq \mu_2 - c, U_1 \geq U_2 - \mu_2 + \mu_1\}$,

$$\begin{aligned} & r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2) - r(U_1 + c - \mu_1) \\ = \; & [r(\underbrace{U_1 + U_2 + \mu_1 - \mu_2}_{b+\beta}) - r(\underbrace{U_2 + c - \mu_2}_{b})] - [r(\underbrace{U_1 + c - \mu_1}_{a+\alpha}) - r(\underbrace{0}_{a})] \\ \geq \; & 0 \quad \text{since} \quad a \geq 0, \quad b - a = U_2 + c - \mu_2 \geq 0; \\ & \text{and} \quad \alpha = U_1 + c - \mu_1 \geq 0, \quad \beta - \alpha = 2(\mu_1 - c) \geq 0. \end{aligned}$$

Further $E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\{\mathbb{I}_{D^{(1)}} + \mathbf{I}_D\}]$

$$\begin{aligned} = \; & E[\{[r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)] \\ & + [r(-U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)]\}\mathbf{I}_D] \geq 0 \end{aligned}$$

because on $D = \{U_1 > \mu_1 - c, U_2 > U_1 - \mu_1 + \mu_2\}$,

$$[r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)] + [r(-U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)]$$

$$
\begin{aligned}
&= [r(\underbrace{U_1 + U_2 + \mu_1 - \mu_2}_{b+\beta}) - r(\underbrace{U_2 + c - \mu_2}_{b})] \\
&\quad - [r(\underbrace{U_2 + c - \mu_2}_{a+\alpha}) - r(\underbrace{-U_1 + U_2 + \mu_1 - \mu_2}_{a})] \\
&\geq 0 \quad \text{since} \quad a = U_2 - U_1 + \mu_1 - \mu_2 \geq 0, \quad b - a = U_1 + c - \mu_1 \geq 0; \\
&\qquad \text{and} \quad \alpha = U_1 + c - \mu_1 \geq 0, \quad \beta - \alpha = 2(\mu_1 - c) \geq 0.
\end{aligned}
$$

Finally note that $E[\{r(U_1 + U_2 + \mu_1 - \mu_2) - r(U_2 + c - \mu_2)\}\mathbb{I}_E] \geq 0$ because on $E = \{c - \mu_1 \leq U_1 \leq \mu_1 - c, U_2 \geq \mu_2 - c\}$, $U_1 + \mu_1 - c \geq 0$ and r is increasing.

Hence $R(S_1^2) - R(S^{1,2}) \geq 0$. In a similar manner one can prove that $R(S_2^1) - R(S^{1,2}) \geq 0$, completing the proof of Theorem for situation (a).

Note that in situation (b), $\mathcal{R}^{1,2} = \emptyset$ and thus only statement (i) needs to be proved, which can be done similarly as for situation (a) by showing the differences $R(S_1^2) - R(S_{1,2})$ and $R(S_2^1) - R(S_{1,2})$ to be nonnegative.

For situation (c) assume without loss of generality that $c_3 \leq c_4$. Note that, in this case

$$
\begin{aligned}
&\mathcal{R}^{1,2} = \{x_1 \geq c_4,\ x_2 \geq c_4\}, && \mathcal{R}_2^1 = \{x_1 \geq c_3,\ x_2 \leq \min(c_4,\ x_1)\}, \\
&\mathcal{R}_1^2 = \{x_1 \leq c_4,\ x_2 \geq \max(c_3,\ x_1)\}, && \mathcal{R}_{1,2} = \{x_1 \leq c_3,\ x_2 \leq c_3\}.
\end{aligned}
$$

One corresponding partition can be taken to be

$$
\begin{aligned}
&P^{1,2} = \{x_1 \geq c_4,\ x_2 \geq c_4\}, && P_2^1 = \{x_1 \geq c_3,\ x_2 < \min(c_4,\ x_1)\}, \\
&P_1^2 = \{x_1 < c_4,\ x_2 \geq \max(c_3,\ x_1)\}, && P_{1,2} = \{x_1 < c_3,\ x_2 < c_3\}.
\end{aligned}
$$

As before denoting the partitions w.r.t. the (u_1, u_2)-space by $Q_{(\cdot)}^{(\cdot)}$ from (28) and (35) we have

$$
\begin{aligned}
R(S^{1,2}) &= E[r(z(S^{1,2}) - z^*)] \\
&= E[r(c_4 - X_2)\mathbb{I}_{P_2^1}] + E[r(c_4 - X_1)\mathbb{I}_{P_1^2}] + E[r(c_3 + c_4 - X_1 - X_2)\mathbb{I}_{P_{1,2}}] \\
&= E[r(c_4 - \mu_2 - U_2)\mathbb{I}_{Q_2^1(U_1,U_2)}] + E[r(c_4 - \mu_1 - U_1)\mathbb{I}_{Q_1^2(U_1,U_2)}] \\
&\qquad + E[r(c_3 + c_4 - \mu_1 - \mu_2 - U_1 - U_2)\mathbb{I}_{Q_{1,2}(U_1,U_2)}] \\
&= E[r(U_2 + c_4 - \mu_2)\mathbb{I}_{Q_2^1(U_1,-U_2)}] + E[r(U_1 + c_4 - \mu_1)\mathbb{I}_{Q_1^2(-U_1,U_2)}] \\
&\qquad + E[r(U_1 + U_2 + c_3 + c_4 - \mu_1 - \mu_2)\mathbb{I}_{Q_{1,2}(-U_1,-U_2)}] \\
&= E[r(U_1 + c_4 - \mu_1)\mathbb{I}_{\Theta^{1,2}(1)}] + E[r(U_2 + c_4 - \mu_2)\mathbb{I}_{\Theta^{1,2}(2)}] + E[r(U_1 + U_2 + c_3 + c_4 - \mu_1 - \mu_2)\mathbb{I}_{\Theta^{1,2}(1,2)}],
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
R(S_2^1) &= E[r(X_2 - c_4)\mathbb{I}_{P^{1,2}}] + E[r(X_2 - X_1)\mathbb{I}_{P_1^2}] + E[r(c_3 - X_1)\mathbb{I}_{P_{1,2}}] \\
&= E[r(U_2 + \mu_2 - c_4)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] + E[r(U_2 + U_1 + \mu_2 - \mu_1)\mathbb{I}_{Q_1^2(-U_1,U_2)}] \\
&\qquad + E[r(c_3 + U_1 - \mu_1)\mathbb{I}_{Q_{1,2}(-U_1,-U_2)}] \\
&= E[r(U_1 + c_3 - \mu_1)\mathbb{I}_{\Theta_2^1(1)}] + E[r(U_2 + \mu_2 - c_4)\mathbb{I}_{\Theta_2^1(2)}] + E[r(U_1 + U_2 + \mu_2 - \mu_1)\mathbb{I}_{\Theta_2^1(1,2)}],
\end{aligned}
$$

$$
R(S_1^2) = E[r(X_1 - c_4)\mathbb{I}_{P^{1,2}}] + E[r(X_1 - X_2)\mathbb{I}_{P_2^1}] + E[r(c_3 - X_2)\mathbb{I}_{P_{1,2}}]
$$

$$
\begin{aligned}
&= E[r(U_1+\mu_1-c_4)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] + E[r(U_1+U_2+\mu_1-\mu_2)\mathbb{I}_{Q_2^1(U_1,-U_2)}] \\
&\qquad + E[r(U_2+c_3-\mu_2)\mathbb{I}_{Q_{1,2}(-U_1,-U_2)}] \\
&= E[r(U_1+\mu_1-c_4)\mathbb{I}_{\Theta_1^2(1)}] + E[r(U_2+c_3-\mu_2)\mathbb{I}_{\Theta_1^2(2)}] + E[r(U_1+U_2+\mu_1-\mu_2)\mathbb{I}_{\Theta_1^2(1,2)}],
\end{aligned}
$$

$$
\begin{aligned}
R(S_{1,2}) &= E[r(X_1+X_2-c_3-c_4)\mathbb{I}_{P^{1,2}}] + E[r(X_1-c_3)\mathbb{I}_{P_2^1}] + E[r(X_2-c_3)\mathbb{I}_{P_1^2}] \\
&= E[r(U_1+U_2+\mu_1+\mu_2-c_3-c_4)\mathbb{I}_{Q^{1,2}(U_1,U_2)}] \\
&\qquad + E[r(U_1+\mu_1-c_3)\mathbb{I}_{Q_2^1(U_1,-U_2)}] + E[r(U_2+\mu_2-c_3)\mathbb{I}_{Q_1^2(-U_1,U_2)}] \\
&= E[r(U_1+\mu_1-c_3)\mathbb{I}_{\Theta_{1,2}(1)}] + E[r(U_2+\mu_2-c_3)\mathbb{I}_{\Theta_{1,2}(2)}] + E[r(U_1+U_2+\mu_1+\mu_2-c_3-c_4)\mathbb{I}_{\Theta_{1,2}(1,2)}],
\end{aligned}
$$

where

$$
\begin{array}{lll}
\Theta^{1,2}(1) = Q_1^2(-U_1,U_2) & \Theta^{1,2}(2) = Q_2^1(U_1,-U_2) & \Theta^{1,2}(1,2) = Q_{1,2}(-U_1,-U_2) \\
\Theta_2^1(1) = Q_{1,2}(-U_1,-U_2) & \Theta_2^1(2) = Q^{1,2}(U_1,U_2) & \Theta_2^1(1,2) = Q_1^2(-U_1,U_2) \\
\Theta_1^2(1) = Q^{1,2}(U_1,U_2) & \Theta_1^2(2) = Q_{1,2}(-U_1,-U_2) & \Theta_1^2(1,2) = Q_2^1(U_1,-U_2) \\
\Theta_{1,2}(1) = Q_2^1(U_1,-U_2) & \Theta_{1,2}(2) = Q_1^2(-U_1,U_2) & \Theta_{1,2}(1,2) = Q^{1,2}(U_1,U_2).
\end{array}
$$

To prove part (i) assume that $(\mu_1,\mu_2) \in \mathcal{R}^{1,2}$, i.e., $c_3 \le c_4 \le \mu_1 \le \mu_2$. We need to show that $R(S^{1,2}) \le \min(R(S_2^1), R(S_1^2), R(S_{1,2}))$. We shall use the same technique used for situations (a) and (b). Namely, we consider each of the differences $R(S_2^1) - R(S^{1,2})$, $R(S_1^2) - R(S^{1,2})$, and $R(S_{1,2}) - R(S^{1,2})$; split the difference into sums of expectations by decomposing the sets of integration, then regroup the expectations in such a way that the random variables, expectations of which are being taken, become nonnegative, thus making the difference nonnegative. We prove the nonegativity of $R(S_2^1) - R(S^{1,2})$ only. The other cases can be proved similarly.

So consider $R(S_2^1) - R(S^{1,2})$

$$
\begin{aligned}
&= E[r(U_1+c_3-\mu_1)\mathbb{I}_{\Theta_2^1(1)}] - E[r(U_1+c_4-\mu_1)\mathbb{I}_{\Theta^{1,2}(1)}] \\
&\quad + E[r(U_2+\mu_2-c_4)\mathbb{I}_{\Theta_2^1(2)}] - E[r(U_2+c_4-\mu_2)\mathbb{I}_{\Theta^{1,2}(2)}] \\
&\quad + E[r(U_1+U_2+\mu_2-\mu_1)\mathbb{I}_{\Theta_2^1(1,2)}] - E[r(U_1+U_2+c_3+c_4-\mu_1-\mu_2)\mathbb{I}_{\Theta^{1,2}(1,2)}].
\end{aligned}
$$

The decomposition of the sets of integrations are as follows (see figure 9).

$$
\begin{aligned}
\Theta_2^1(1) &= \Theta^{1,2}(1,2) = \{U_1 > \mu_1 - c_3,\ U_2 > \mu_2 - c_3\} = A_1, \text{ say,} \\
\Theta^{1,2}(1) &= \Theta_2^1(1,2) = A_1 \cup \{\mu_1 - c_4 < U_1 \le \mu_1 - c_3,\ U_2 > U_1 + \mu_2 - \mu_1\} \\
&\qquad \cup \{\mu_2 - c_4 < U_2 \le \mu_2 - c_3,\ U_1 \ge U_2 + \mu_1 - \mu_2\} \cup \{U_1 > \mu_1 - c_4,\ c_4 - \mu_2 \le U_2 \le \mu_2 - c_4\} \\
&\qquad \cup \{c_3 - \mu_2 \le U_2 < c_4 - \mu_2,\ U_1 \ge -U_2 + \mu_1 - \mu_2\} \\
&= A_1 \cup B_1 \cup C_1 \cup D_1 \cup C_1^{(2)}, \quad \text{say,} \\
\Theta_2^1(2) &= \{U_1 \ge c_4 - \mu_1,\ U_2 > \mu_2 - c_4\} \cup \{U_1 \ge c_4 - \mu_1,\ c_4 - \mu_2 \le U_2 \le \mu_2 - c_4\} = E_1 \cup F_1, \text{ say,} \\
\Theta^{1,2}(2) &= E_1 \cup \{c_3 - \mu_1 \le U_1 < c_4 - \mu_1,\ U_2 > \mu_2 - \mu_1 - U_1\} = E_1 \cup B_1^{(1)}, \text{ say.}
\end{aligned}
$$

Then $R(S_2^1) - R(S^{1,2})$

$$
\begin{aligned}
&= E[(r(U_1+c_3-\mu_1) - r(U_1+c_4-\mu_1) + r(U_1+U_2+\mu_2-\mu_1) - r(U_1+U_2+c_3+c_4-\mu_1-\mu_2))\mathbf{I}_{A_1}] \\
&\quad + E[(r(U_1+U_2+\mu_2-\mu_1) - r(U_1+c_4-\mu_1))\mathbf{I}_{B_1}] + E[(r(U_1+U_2+\mu_2-\mu_1) - r(U_1+c_4-\mu_1))\mathbf{I}_{C_1}] \\
&\quad + E[(r(U_1+U_2+\mu_2-\mu_1) - r(U_1+c_4-\mu_1))\mathbf{I}_{D_1}] + E[(r(U_1+U_2+\mu_2-\mu_1) - r(U_1+c_4-\mu_1))\mathbf{I}_{C_1^{(2)}}] \\
&\quad + E[(r(U_2+\mu_2-c_4) - r(U_2+c_4-\mu_2))\mathbf{I}_{E_1}] + E[r(U_2+\mu_2-c_4)\mathbf{I}_{F_1}] - E[r(U_2+c_4-\mu_2)\mathbf{I}_{B_1^{(1)}}] \\
&= I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8, \quad \text{say.}
\end{aligned}
$$

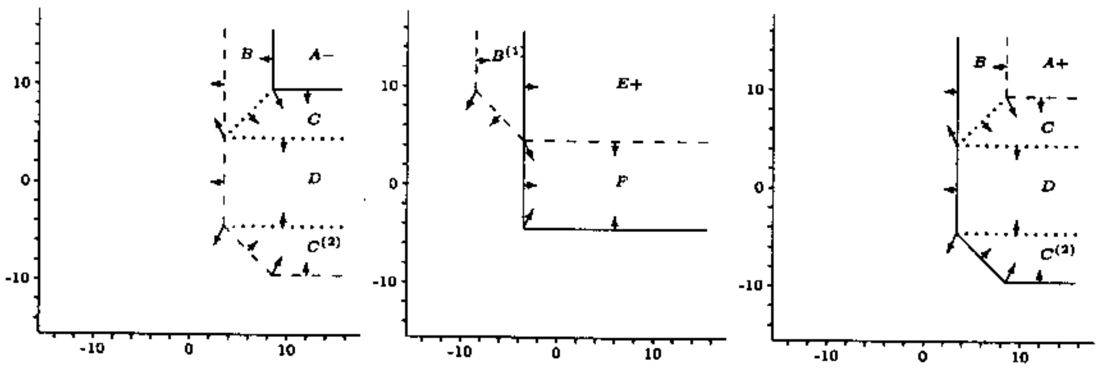

Figure 9: Decompositions of sets of integrations for $R(S_2^1)$ (solid line) and $R(S^{1,2})$ (dashed line). First plot is for sets involving $r(U_1+\cdots)$, second for sets involving $r(U_2+\cdots)$ and third for sets involving $r(U_1+U_2+\cdots)$. Arrows indicate the (non)inclusion of boundary lines/points. Symbols '+'/'−' indicate the contribution from the integral over the corresponding set to the difference $R(S_2^1)-R(S^{1,2})$.

From growth condition (34) it follows that $I_1 \ge 0$ because on $A_1=\{U_1>\mu_1-c_3,\ U_2>\mu_2-c_3\}$,

$$\begin{aligned}
&r(U_1+c_3-\mu_1)-r(U_1+c_4-\mu_1)+r(U_1+U_2+\mu_2-\mu_1)-r(U_1+U_2+c_3+c_4-\mu_1-\mu_2)\\
&= [r(\underbrace{U_1+U_2+\mu_2-\mu_1}_{b+\beta})-r(\underbrace{U_1+U_2+c_3+c_4-\mu_1-\mu_2}_{b})]-[r(\underbrace{U_1+c_4-\mu_1}_{a+\alpha})-r(\underbrace{U_1+c_3-\mu_1}_{a})]\\
&\ge 0,\quad \text{since}\quad \alpha=c_4-c_3\ge 0,\quad \beta-\alpha=2\mu_2-c_3-c_4-(c_4-c_3)=2(\mu_2-c_4)\ge 0\\
&\qquad \text{and}\quad a=U_1+c_3-\mu_1\ge 0,\quad b-a=U_2+c_4-\mu_2\ge U_2+c_3-\mu_2\ge 0.
\end{aligned}$$

Also, noting that $B_1^{(1)}$ reflected w.r.t. the y-axis is same as B_1 and $C_1^{(2)}$ reflected w.r.t. the x-axis is same as C_1, from (35) we have

$$\begin{aligned}
I_8 &= E[r(U_2+c_4-\mu_2)\mathbf{I}_{B_1^{(1)}}] &&= E[r(U_2+c_4-\mu_2)\mathbf{I}_{B_1}]\\
I_5 &= E[(r(U_1+U_2+\mu_2-\mu_1)-r(U_1+c_4-\mu_1))\mathbf{I}_{C_1^{(2)}}] &&= E[(r(U_1-U_2+\mu_2-\mu_1)-r(U_1+c_4-\mu_1))\mathbf{I}_{C_1}]
\end{aligned}$$

Then $I_2+I_8\ge 0$ because on $B_1=\{\mu_1-c_4<U_1\le\mu_1-c_3,\ U_2>U_1+\mu_2-\mu_1\}$,

$$\begin{aligned}
&(r(U_1+U_2+\mu_2-\mu_1)-r(U_1+c_4-\mu_1))-r(U_2+c_4-\mu_2)\\
&= [r(\underbrace{U_1+U_2+\mu_2-\mu_1}_{b+\beta})-r(\underbrace{U_1+c_4-\mu_1}_{b})]-[r(\underbrace{U_2+c_4-\mu_2}_{a+\alpha})-r(\underbrace{0}_{a})]\\
&\ge 0,\quad \text{since}\quad \alpha=U_2+c_4-\mu_2\ge 0,\quad \beta-\alpha=U_2+\mu_2-c_4-(U_2+c_4-\mu_2)=2(\mu_2-c_4)\ge 0\\
&\qquad \text{and}\quad a\ge 0,\quad b-a=U_1+c_4-\mu_1\ge 0.
\end{aligned}$$

Similarly, $I_3+I_5\ge 0$ because on $C_1=\{\mu_2-c_4<U_2\le\mu_2-c_3,\ U_1\ge U_2+\mu_1-\mu_2\}$,

$$\begin{aligned}
&(r(U_1+U_2+\mu_2-\mu_1)-r(U_1+c_4-\mu_1))-(r(U_1-U_2+\mu_2-\mu_1)-r(U_1+c_4-\mu_1))\\
&= [r(\underbrace{U_1+U_2+\mu_2-\mu_1}_{b+\beta})-r(\underbrace{U_1+c_4-\mu_1}_{b})]-[r(\underbrace{U_1+c_4-\mu_1}_{a+\alpha})-r(\underbrace{U_1-U_2+\mu_2-\mu_1}_{a})]\\
&\ge 0,\quad \text{since}\quad \alpha=U_2+c_4-\mu_2\ge 0,\quad \beta-\alpha=U_2+\mu_2-c_4-(U_2+c_4-\mu_2)=2(\mu_2-c_4)\ge 0\\
&\qquad \text{and}\quad a=U_1-U_2+\mu_2-\mu_1\ge 0,\quad b-a=U_2+c_4-\mu_2\ge 0.
\end{aligned}$$

Furthermore, $I_4\ge 0$ because on $D_1=\{U_1>\mu_1-c_4,\ c_4-\mu_2\le U_2\le\mu_2-c_4\}$,

$$r(\underbrace{U_1+U_2+\mu_2-\mu_1}_{b})-r(\underbrace{U_1+c_4-\mu_1}_{a})\ge 0,\qquad \text{since}\qquad b-a=U_2+\mu_2-c_4\ge 0;$$

$I_6 \geq 0$ because on $E_1 = \{U_1 \geq c_4 - \mu_1,\ U_2 > \mu_2 - c_4\}$,

$$r(\underbrace{U_2 + \mu_2 - c_4}_{b}) - r(\underbrace{U_2 + c_4 - \mu_2}_{a}) \geq 0, \qquad \text{since} \qquad b - a = 2(\mu_2 - c_4) \geq 0;$$

and $I_7 \geq 0$ because on $F_1 = \{U_1 \geq c_4 - \mu_1,\ c_4 - \mu_2 \leq U_2 \leq \mu_2 - c_4\}$,

$$r(U_2 + \mu_2 - c_4) \geq 0, \qquad \text{since} \qquad U_2 + \mu_2 - c_4 \geq 0.$$

Thus $R(S_2^1) - R(S^{1,2}) = I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8 \geq 0$.

Part (ii) of the theorem is also proved similarly: assuming that $(\mu_1, \mu_2) \in \mathcal{R}_{1,2}$, i.e., $\mu_1 \leq \mu_2 \leq c_3 \leq c_4$ and further showing that the differences $R(S_{1,2}) - R(S^{1,2})$, $R(S_{1,2}) - R(S_2^1)$ and $R(S_{1,2}) - R(S_1^2)$ are nonnegative.

5.5 Proof of Theorem 10

We prove the theorem under the general setup of Remark 14, i.e., under (36). Note that the set of feasible solutions $\mathcal{S} = \{S_I : I \subseteq \{1, 2, \ldots, n\} =: \mathbb{N}\}$, where $X_i \in S_I$, if $i \in I$, otherwise $c_i \in S_I$. Then $z_I := z(S_I) = X^I + c^{\bar{I}}$, where $\bar{I} = \mathbb{N} \setminus I$, the complement of I, and we use a^I to denote $\sum_{i \in I} a_i$, if $I \neq \phi$ and 0, otherwise.

Then defining $d_i = c_i - \mu_i$ we have

$$\begin{aligned}
\mathcal{R}_I &:= \{z_I = \min_{J \subset \mathbb{N}} z_J\} = \cap_J \{z_I \leq z_J\} = \cap_J \{X^I + c^{\bar{I}} \leq X^J + c^{\bar{J}}\} \\
&= \cap_J \{X^I - X^J \leq c^{\bar{J}} - c^{\bar{I}} = c^I - c^J\} = \cap_J \{U^I - U^J \leq d^I - d^J\} \\
&= \cap_J \{U^{I \setminus J} - U^{J \setminus I} \leq d^{I \setminus J} - d^{J \setminus I}\} = \{U_i \leq d_i, i \in I \text{ and } U_i \geq d_i, i \notin I\}.
\end{aligned}$$

One set of corresponding partition can be then taken to be

$$P_I = \{U_i < d_i, i \in I \text{ and } U_i \geq d_i, i \notin I\}, \quad I \subset \mathbb{N}.$$

Then from (36), we have

$$\begin{aligned}
R_I &:= R(S_I) = \sum_{J \neq I} E[r(z_I - z_J)\mathbb{I}_{P_J}] = \sum_{J \neq I} E[r(U^{I \setminus J} - U^{J \setminus I} - (d^{I \setminus J} - d^{J \setminus I}))\mathbb{I}_{P_J}] \\
&= \sum_{J \neq I} E[r(U^{I \setminus J} + U^{J \setminus I} - (d^{I \setminus J} - d^{J \setminus I}))\mathbb{I}_{Q_J}] = \sum_{J \neq I} E[r(U^{I \Delta J} - (d^{I \setminus J} - d^{J \setminus I}))\mathbb{I}_{Q_J}] \\
&= \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[r\left(U^K - \left(d^{KI} - d^{K\bar{I}}\right)\right)\mathbb{I}_{Q_{I \Delta K}}\right],
\end{aligned}$$

where

$$Q_I = \{U_i > -d_i, i \in I \text{ and } U_i \geq d_i, i \notin I\}, \quad I \subset \mathbb{N}.$$

Now suppose that $\mu_i < c_i, i \in M$ and $\mu_i \geq c_i, i \notin M$, i.e., $d_i > 0, i \in M$ and $d_i \leq 0, i \notin M$. For the theorem we need to show that $R(S_M) \leq R(S_I)$ for all $I \subset \mathbb{N}$.

Note that for any $I \subset \mathbb{N}$,

$$R(S_I) - R(S_M) = \sum_{\substack{K\subset\mathbb{N}\\K\neq\phi}} E\left[r\left(U^K - d^{KI} + d^{K\bar{I}}\right)\mathbb{I}_{Q_{I\Delta K}} - r\left(U^K - d^{KM} + d^{K\bar{M}}\right)\mathbb{I}_{Q_{M\Delta K}}\right]$$

$$= \sum_{\substack{K\subset\mathbb{N}\\K\neq\phi}} E\left[\left\{r\left(U^K - d^{KI} + d^{K\bar{I}}\right) - r\left(U^K - d^{KM} + d^{K\bar{M}}\right)\right\}\mathbb{I}_{Q_{I\Delta K}\cap Q_{M\Delta K}}\right]$$

$$+ \sum_{\substack{K\subset\mathbb{N}\\K\neq\phi}} E\left[r\left(U^K - d^{KI} + d^{K\bar{I}}\right)\mathbb{I}_{Q_{I\Delta K}\setminus Q_{M\Delta K}}\right]$$

$$- \sum_{\substack{K\subset\mathbb{N}\\K\neq\phi}} E\left[r\left(U^K - d^{KM} + d^{K\bar{M}}\right)\mathbb{I}_{Q_{M\Delta K}\setminus Q_{I\Delta K}}\right]$$

Note that

$$\begin{aligned}(U^K - d^{KI} + d^{K\bar{I}}) - (U^K - d^{KM} + d^{K\bar{M}}) &= d^{K\bar{I}} - d^{KI} + d^{KM} - d^{K\bar{M}} \\ &= d^{K\bar{I}M} + d^{K\bar{I}\bar{M}} - d^{KIM} - d^{KI\bar{M}} + d^{KIM} + d^{K\bar{I}M} - d^{KI\bar{M}} - d^{K\bar{I}\bar{M}} \\ &= 2(d^{K\bar{I}M} - d^{KI\bar{M}}) \geq 0, \text{ since } d_i > 0, i \in M, \text{ and } d_i \leq 0, i \notin M.\end{aligned}$$

Hence the first sum in $R(S_I) - R(S_M)$ is nonnegative. Now let us look at the sets $Q_{I\Delta K} \setminus Q_{M\Delta K}$ and $Q_{M\Delta K} \setminus Q_{I\Delta K}$. Below we describe the structure of these sets.

for	$Q_{I\Delta K}$	$Q_{M\Delta K}$	$Q_{I\Delta K} \setminus Q_{M\Delta K}$	$Q_{M\Delta K} \setminus Q_{I\Delta K}$
$i \in KMI$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$
$i \in KM\bar{I}$	$-d_i < u_i$	$d_i \leq u_i$	"$-d_i < u_i < d_i$"	$d_i \leq u_i$
$i \in K\bar{M}I$	$d_i \leq u_i$	$-d_i < u_i$	"$d_i \leq u_i \leq -d_i$"	$-d_i < u_i$
$i \in K\bar{M}\bar{I}$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$
$i \in \bar{K}MI$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$	$-d_i < u_i$
$i \in \bar{K}M\bar{I}$	$d_i \leq u_i$	$-d_i < u_i$	$d_i \leq u_i$	"$-d_i < u_i < d_i$"
$i \in \bar{K}\bar{M}I$	$-d_i < u_i$	$d_i \leq u_i$	$-d_i < u_i$	"$d_i \leq u_i \leq -d_i$"
$i \in \bar{K}\bar{M}\bar{I}$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$	$d_i \leq u_i$

In the table if a region is within quotation marks, this implies that some (but not necessarily all) of these type of boundary conditions would appear in the corresponding set. The exact construction is as follows. For nonempty K^*, $K \subset \mathbb{N}$,

$$Q_{I\Delta K^*} \setminus Q_{M\Delta K^*} = \bigcup_{J^*\subset K^*M\bar{I}} \bigcup_{L^*\subset K^*\bar{M}I} \left\{\begin{array}{ll} d_i \leq u_i & i \in K^*MI \cup \bar{K}^*\bar{M}\bar{I} \\ -d_i < u_i & i \in K^*\bar{M}\bar{I} \cup \bar{K}^*MI \\ -d_i < u_i < d_i & i \in J^* \\ d_i \leq u_i & i \in \bar{J}^*K^*M\bar{I} \cup \bar{K}^*M\bar{I} \\ d_i \leq u_i \leq -d_i & i \in L^* \\ -d_i < u_i & i \in \bar{L}^*K^*\bar{M}I \cup \bar{K}^*\bar{M}I \end{array}\right\}$$

$$= \bigcup_{J^* \subset K^* M \bar{I}} \bigcup_{L^* \subset K^* \bar{M} I} Q(I \setminus M, K^*, J^*, L^*), \text{ say,}$$

$$Q_{M\Delta K} \setminus Q_{I\Delta K} = \bigcup_{J \subset \bar{K} M \bar{I}} \bigcup_{L \subset \bar{K} \bar{M} I} \left\{ \begin{array}{ll} d_i \le u_i & i \in KMI \cup \bar{K}\bar{M}\bar{I} \\ -d_i < u_i & i \in K\bar{M}\bar{I} \cup \bar{K}MI \\ -d_i < u_i < d_i & i \in J \\ d_i \le u_i & i \in \bar{J}\bar{K}M\bar{I} \cup KM\bar{I} \\ d_i \le u_i \le -d_i & i \in L \\ -d_i < u_i & i \in \bar{L}\bar{K}\bar{M}I \cup K\bar{M}I \end{array} \right\}$$

$$= \bigcup_{J \subset \bar{K} M \bar{I}} \bigcup_{L \subset \bar{K} \bar{M} I} Q(M \setminus I, K, J, L).$$

Then

$$\begin{aligned} & R(S_I) - R(S_M) \\ & \ge \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[r\left(U^K - d^{KI} + d^{K\bar{I}}\right) \mathbb{I}_{Q_{I\Delta K} \setminus Q_{M\Delta K}}\right] - \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} E\left[r\left(U^K - d^{KM} + d^{K\bar{M}}\right) \mathbb{I}_{Q_{M\Delta K} \setminus Q_{I\Delta K}}\right] \\ & = \sum_{\substack{K^* \subset \mathbb{N} \\ K^* \neq \phi}} \sum_{J^* \subset K^* M \bar{I}} \sum_{L^* \subset K^* \bar{M} I} E\left[r\left(U^{K^*} - d^{K^*I} + d^{K^*\bar{I}}\right) \mathbb{I}_{Q(I \setminus M, K^*, J^*, L^*)}\right] \\ & \quad - \sum_{\substack{K \subset \mathbb{N} \\ K \neq \phi}} \sum_{J \subset \bar{K} M \bar{I}} \sum_{L \subset \bar{K} \bar{M} I} E\left[r\left(U^K - d^{KM} + d^{K\bar{M}}\right) \mathbb{I}_{Q(M \setminus I, K, J, L)}\right] \end{aligned}$$

We shall show that for each negative term in the 2nd sum above there is a positive term in the 1st sum so that the net contribution from these two terms becomes nonnegative, and thus proving that $R(S_I) \ge R(S_M)$. So fix a negative term, i.e., fix $\phi \neq K \subset \mathbb{N}$, $J \subset \bar{K}M\bar{I}$, $L \subset \bar{K}\bar{M}I$. Consider $K^* = K \cup J \cup L$, $J^* = J$, $L^* = L$.

First note that $J \subset M\bar{I}$ and $J \subset K^* \Rightarrow J^* = J \subset K^* M \bar{I}$, and similarly $L^* = L \subset K^* \bar{M} I$.

Next $J \cup L \subset (I \Delta M) \Rightarrow$

$$\bar{K}^* \bar{M} \bar{I} = \bar{K} \cap \overline{(J \cup L)} \cap \bar{M}\bar{I} = \bar{K}\bar{M}\bar{I}, \qquad \bar{K}^* MI = \bar{K} \cap \overline{(J \cup L)} \cap MI = \bar{K}MI,$$

$$K^* MI = KMI \cup [(J \cup L) \cap MI] = KMI \qquad K^* \bar{M} \bar{I} = K\bar{M}\bar{I} \cup [(J \cup L) \cap \bar{M}\bar{I}] = K\bar{M}\bar{I}.$$

Further, note that $J \subset \bar{K}M\bar{I} \Rightarrow \bar{J} \supset K \cup \bar{M} \cup I \supset KM\bar{I}$. Then

$$L \subset \bar{M}I \quad \Rightarrow \quad \bar{J}^* K^* M \bar{I} = \bar{J} \cap (J \cup K \cup L) \cap M\bar{I} = \bar{J} \cap K \cap M\bar{I} = KM\bar{I}.$$

In a similar fashion one can see that $\bar{L}^* K^* \bar{M} I = K\bar{M}I$. Finally note that

$$L \subset \bar{M}I \quad \Rightarrow \quad \bar{L} \supset M\bar{I} \quad \Rightarrow \quad \bar{K}^* M \bar{I} = \bar{K} \cap \bar{J} \cap \bar{L} \cap M\bar{I} = \bar{J}\bar{K}M\bar{I},$$

and similarly $\bar{K}^* \bar{M} I = \bar{L}\bar{K}\bar{M}I$.

Hence $Q(M \setminus I, K, J, L) \;=\; Q(I \setminus M, K^*, J^*, L^*)$, and

$$
\begin{aligned}
&E\left[r\left(U^{K^*} - d^{K^*I} + d^{K^*\bar{I}}\right)\mathbb{I}_{Q(I\setminus M, K^*, J^*, L^*)}\right] - E\left[r\left(U^{K} - d^{KM} + d^{K\bar{M}}\right)\mathbb{I}_{Q(M\setminus I, K, J, L)}\right] \\
&= \; E\left[\left\{r\left(\underbrace{U^{K^*} - d^{K^*I} + d^{K^*\bar{I}}}_{b}\right) - r\left(\underbrace{U^{K} - d^{KM} + d^{K\bar{M}}}_{a}\right)\right\}\mathbb{I}_{Q(M\setminus I, K, J, L)}\right] \\
&\geq \; 0,
\end{aligned}
$$

because from $K^*MI = KMI$, $K^*\bar{M}\bar{I} = K\bar{M}\bar{I}$, and the facts that $d_i > 0$, $i \in M$; $d_i \leq 0$, $i \notin M$, we have

$$
\begin{aligned}
b - a &= U^{K^*} - d^{K^*I} + d^{K^*\bar{I}} - U^K + d^{KM} - d^{K\bar{M}} \\
&= \quad U^K + U^J + U^L \; - (d^{K^*MI} + d^{K^*\bar{M}I}) \; + (d^{K^*M\bar{I}} + d^{K^*\bar{M}\bar{I}}) \\
&\quad\; - U^K \qquad\qquad\quad + (d^{KMI} + d^{KM\bar{I}}) \quad - (d^{K\bar{M}I} + d^{K\bar{M}\bar{I}}) \\
&= U^J + U^L - d^{K^*\bar{M}I} + d^{K^*M\bar{I}} + d^{KM\bar{I}} - d^{K\bar{M}I} \\
&= U^J + U^L - d^L - d^{\bar{L}K^*\bar{M}I} + d^J + d^{\bar{J}K^*M\bar{I}} + d^{KMI} - d^{K\bar{M}I} \\
&= (U^J + d^J) + (U^L - d^L) + (d^{\bar{J}K^*M\bar{I}} + d^{KM\bar{I}}) - (d^{\bar{L}K^*\bar{M}I} + d^{K\bar{M}I}) \\
&= \sum_{i\in J}(u_i + d_i) + \sum_{i\in L}(u_i - d_i) + \sum_{i\in(\bar{J}K^*M\bar{I}\cup KM\bar{I})\subset M} d_i - \sum_{i\in(\bar{L}K^*\bar{M}I\cup K\bar{M}I)\subset\bar{M}} d_i \\
&\geq 0,
\end{aligned}
$$

if $(U_1, \ldots, U_n) \in Q(M \setminus I, K, J, L)$.

Finally, it is clear that this mapping from negative term to positive term is injective. Hence we do not use the same positive term for two different negative terms, i.e., each negative term is compensated by a separate positive term. Hence the result is proved.

5.6 Tables

Table 6: Data for Figure 5

n	β value								
	2.05	2.1	2.15	2.2	2.25	2.3	2.35	2.4	2.45
2	0.494	0.488	0.482	0.477	0.471	0.465	0.460	0.456	0.451
3	0.494	0.489	0.483	0.477	0.472	0.466	0.461	0.457	0.452
4	0.495	0.489	0.483	0.478	0.472	0.467	0.462	0.458	0.453
5	0.495	0.489	0.484	0.478	0.473	0.468	0.463	0.459	0.454
6	0.495	0.489	0.484	0.479	0.473	0.468	0.463	0.460	0.455
8	0.495	0.490	0.485	0.480	0.475	0.470	0.465	0.461	0.457
10	0.495	0.490	0.486	0.481	0.476	0.472	0.467	0.463	0.459
12	0.496	0.491	0.486	0.482	0.477	0.473	0.469	0.465	0.460
14	0.496	0.492	0.487	0.483	0.479	0.474	0.470	0.466	0.462
16	0.496	0.492	0.488	0.484	0.480	0.476	0.472	0.468	0.464
18	0.496	0.492	0.488	0.485	0.481	0.477	0.473	0.470	0.466
20	0.497	0.493	0.489	0.485	0.482	0.478	0.475	0.471	0.468
25	0.497	0.494	0.490	0.487	0.484	0.480	0.477	0.474	0.471
30	0.497	0.494	0.491	0.488	0.485	0.482	0.479	0.477	0.474
35	0.498	0.495	0.492	0.489	0.487	0.484	0.481	0.479	0.476
40	0.498	0.495	0.493	0.490	0.488	0.485	0.483	0.480	0.478
45	0.498	0.496	0.493	0.491	0.489	0.486	0.484	0.482	0.479
50	0.498	0.496	0.494	0.492	0.489	0.487	0.485	0.483	0.481
55	0.498	0.496	0.494	0.492	0.490	0.488	0.486	0.484	0.482
60	0.498	0.496	0.494	0.493	0.491	0.489	0.487	0.485	0.483
70	0.499	0.497	0.495	0.493	0.492	0.490	0.488	0.486	0.485
80	0.499	0.497	0.496	0.494	0.492	0.491	0.489	0.488	0.486
90	0.499	0.497	0.496	0.494	0.493	0.492	0.490	0.489	0.487
100	0.499	0.498	0.496	0.495	0.494	0.492	0.491	0.489	0.488

Table 7: Data for Figure 6

n	τ values (corresponding β values)							
	0.10 (2.40)	0.15 (2.70)	0.20 (3.15)	0.25 (3.75)	0.30 (4.80)	0.35 (6.50)	0.40 (11.00)	0.45 (35.00)
2	0.456	0.427	0.390	0.350	0.296	0.237	0.155	0.054
3	0.457	0.428	0.392	0.352	0.298	0.239	0.156	0.054
4	0.458	0.429	0.394	0.354	0.300	0.241	0.158	0.055
5	0.459	0.431	0.395	0.356	0.303	0.244	0.159	0.055
6	0.460	0.432	0.398	0.359	0.305	0.246	0.161	0.055
8	0.461	0.435	0.402	0.364	0.311	0.251	0.164	0.056
10	0.463	0.438	0.406	0.369	0.317	0.256	0.168	0.056
12	0.465	0.442	0.410	0.374	0.323	0.262	0.171	0.057
14	0.466	0.444	0.415	0.379	0.329	0.268	0.175	0.057
16	0.468	0.447	0.418	0.384	0.335	0.275	0.180	0.058
18	0.470	0.450	0.421	0.389	0.341	0.282	0.185	0.059
20	0.471	0.452	0.425	0.393	0.346	0.288	0.191	0.060
25	0.474	0.457	0.432	0.403	0.359	0.303	0.205	0.062
30	0.477	0.460	0.438	0.412	0.371	0.316	0.219	0.065
35	0.479	0.463	0.443	0.419	0.379	0.328	0.233	0.068
40	0.480	0.466	0.447	0.423	0.387	0.338	0.246	0.073
45	0.482	0.469	0.451	0.428	0.394	0.347	0.257	0.078
50	0.483	0.471	0.454	0.433	0.401	0.356	0.267	0.083
55	0.484	0.472	0.457	0.437	0.406	0.363	0.277	0.088
60	0.485	0.474	0.459	0.440	0.411	0.370	0.286	0.095
70	0.486	0.476	0.462	0.446	0.419	0.380	0.302	0.109
80	0.488	0.478	0.466	0.450	0.425	0.389	0.315	0.123
90	0.489	0.480	0.468	0.454	0.430	0.397	0.327	0.136
100	0.489	0.482	0.470	0.457	0.435	0.404	0.336	0.149

Table 8: Data for Figure 7

n	Mode M																
	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.49
2	0.354	0.371	0.387	0.403	0.419	0.436	0.452	0.468	0.471	0.474	0.478	0.481	0.484	0.487	0.490	0.494	0.497
3	0.358	0.375	0.390	0.406	0.422	0.438	0.454	0.469	0.472	0.475	0.478	0.482	0.485	0.488	0.491	0.494	0.497
4	0.363	0.379	0.394	0.410	0.425	0.440	0.455	0.470	0.473	0.476	0.479	0.482	0.485	0.488	0.491	0.494	0.497
5	0.367	0.383	0.398	0.413	0.428	0.443	0.457	0.472	0.474	0.477	0.480	0.483	0.486	0.489	0.492	0.494	0.497
6	0.372	0.387	0.402	0.417	0.431	0.445	0.459	0.473	0.476	0.478	0.481	0.484	0.486	0.489	0.492	0.495	0.497
8	0.381	0.395	0.410	0.423	0.437	0.450	0.463	0.475	0.478	0.480	0.483	0.485	0.488	0.490	0.493	0.495	0.498
10	0.389	0.403	0.417	0.430	0.442	0.455	0.466	0.478	0.480	0.482	0.484	0.487	0.489	0.491	0.493	0.496	0.498
12	0.397	0.411	0.424	0.436	0.448	0.459	0.470	0.480	0.482	0.484	0.486	0.488	0.490	0.492	0.494	0.496	0.498
14	0.405	0.418	0.430	0.442	0.452	0.463	0.472	0.482	0.484	0.486	0.487	0.489	0.491	0.493	0.495	0.496	0.498
16	0.411	0.424	0.436	0.446	0.457	0.466	0.475	0.484	0.485	0.487	0.489	0.490	0.492	0.494	0.495	0.497	0.498
18	0.417	0.429	0.440	0.451	0.460	0.469	0.477	0.485	0.487	0.488	0.490	0.491	0.493	0.494	0.496	0.497	0.499
20	0.422	0.434	0.445	0.455	0.463	0.472	0.479	0.486	0.488	0.489	0.490	0.492	0.493	0.495	0.496	0.497	0.499
25	0.433	0.444	0.454	0.462	0.470	0.477	0.483	0.489	0.490	0.491	0.492	0.493	0.494	0.496	0.497	0.498	0.499
30	0.441	0.451	0.460	0.468	0.474	0.480	0.486	0.491	0.492	0.493	0.493	0.494	0.495	0.496	0.497	0.498	0.499
35	0.447	0.457	0.465	0.472	0.478	0.483	0.488	0.492	0.493	0.494	0.494	0.495	0.496	0.497	0.498	0.498	0.499
40	0.452	0.462	0.469	0.475	0.480	0.485	0.489	0.493	0.494	0.494	0.495	0.496	0.496	0.497	0.498	0.499	0.499
45	0.457	0.466	0.473	0.478	0.483	0.487	0.490	0.494	0.494	0.495	0.496	0.496	0.497	0.497	0.498	0.499	0.499
50	0.460	0.469	0.475	0.480	0.484	0.488	0.491	0.494	0.495	0.495	0.496	0.497	0.497	0.498	0.498	0.499	0.499
55	0.463	0.471	0.477	0.482	0.486	0.489	0.492	0.495	0.495	0.496	0.496	0.497	0.497	0.498	0.498	0.499	0.500
60	0.466	0.474	0.479	0.483	0.487	0.490	0.493	0.495	0.496	0.496	0.497	0.497	0.498	0.498	0.499	0.499	0.500
70	0.470	0.477	0.482	0.486	0.489	0.491	0.494	0.496	0.496	0.497	0.497	0.498	0.498	0.498	0.499	0.499	0.500
80	0.474	0.480	0.484	0.487	0.490	0.492	0.494	0.496	0.497	0.497	0.497	0.498	0.498	0.499	0.499	0.499	0.500
90	0.476	0.482	0.486	0.489	0.491	0.493	0.495	0.497	0.497	0.497	0.498	0.498	0.498	0.499	0.499	0.499	0.500
100	0.479	0.484	0.487	0.490	0.492	0.494	0.496	0.497	0.497	0.498	0.498	0.498	0.499	0.499	0.499	0.499	0.500

Table 9: Table for Example 2: Least Risk solutions in balanced DOP with $r = 2, f = 2, k = 2$.DOP symmetric when both random elements have Triangular distribution on (0,1).

		Risk of candidate solutions				Optimal solution	
c_1	c_2	S_{12}	S_1^2	S_2^1	S^{12}	stochastic	non-stochastic
0.5020	0.5216	0.0516	0.0541	0.0541	0.0635	1	1
0.5160	0.5304	0.0461	0.0548	0.0548	0.0695	1	1
0.4260	0.4350	0.1010	0.0587	0.0587	0.0292	4	4
0.4630	0.4831	0.0723	0.0545	0.0545	0.0450	4	4
0.4771	0.5888	0.0521	0.0491	0.0491	0.0870	2 or 3	2 or 3
0.4872	0.5235	0.0562	0.0528	0.0528	0.0616	2 or 3	2 or 3
0.4937	0.5844	0.0471	0.0504	0.0504	0.0881	1	2 or 3
0.4069	0.5816	0.0806	0.0461	0.0461	0.0743	2 or 3	2 or 3
0.4052	0.5064	0.0931	0.0501	0.0501	0.0462	4	2 or 3
0.4920	0.5945	0.0467	0.0499	0.0499	0.0926	1	2 or 3
0.4381	0.5541	0.0704	0.0483	0.0483	0.0663	2 or 3	2 or 3
0.4739	0.5049	0.0639	0.0532	0.0532	0.0532	4	2 or 3

Among the candidate solutions, '1' stands for $S_{1,2}$, '2' for S_1^2, '3' for S_2^1, '4' for $S^{1,2}$.

Table 10: Table for Example 2: Least Risk solutions in balanced DOP with $r = 2, f = 2, k = 2$, when the first random elements has Triangular distribution while the other has V-distribution, both being symmetric and supported on (0,1).

		Risk of candidate solutions				Optimal solution	
c_1	c_2	S_{12}	S_1^2	S_2^1	S^{12}	stochastic	non-stochastic
0.5108	0.5559	0.0938	0.1044	0.1091	0.1446	1	1
0.5000	0.5920	0.0927	0.1000	0.1054	0.1637	1	1
0.4906	0.4954	0.1171	0.1111	0.1100	0.1065	4	4
0.4243	0.4656	0.1641	0.1126	0.1046	0.0801	4	4
0.4563	0.5939	0.1141	0.0967	0.0990	0.1532	2	2 or 3
0.4558	0.5979	0.1137	0.0964	0.0988	0.1556	2	2 or 3
0.4014	0.5473	0.1553	0.0984	0.0961	0.1152	3	2 or 3
0.4893	0.5599	0.1033	0.1022	0.1052	0.1409	2	2 or 3
0.4987	0.5385	0.1029	0.1054	0.1080	0.1311	1	2 or 3
0.4491	0.5334	0.1300	0.1025	0.1015	0.1166	3	2 or 3
0.4564	0.5040	0.1333	0.1073	0.1045	0.1032	4	2 or 3

Among the candidate solutions, '1' stands for $S_{1,2}$, '2' for S_1^2, '3' for S_2^1, '4' for $S^{1,2}$.

Table 11: Table for Example 2: Least Risk solutions in balanced DOP with $r = 2, f = 2, k = 2$, when the random elements have Bivariate Normal distribution with parameters $(10, 10, 2^2, 2^2, 0.5)$

		Risk of candidate solutions				Optimal solution	
c_1	c_2	S_{12}	S_1^2	S_2^1	S^{12}	stochastic	non-stochastic
9.7878	10.5819	5.8811	3.6495	3.6495	7.9734	2 or 3	2 or 3
10.0808	10.9109	4.5136	3.7759	3.7759	10.1913	2 or 3	1
9.7309	10.6532	5.9344	3.5430	3.5430	8.1143	2 or 3	2 or 3
9.3275	10.8678	6.9623	3.0832	3.0832	8.0872	2 or 3	2 or 3
9.0232	10.1056	9.9350	3.5207	3.5207	4.9349	2 or 3	2 or 3
10.2677	10.4189	4.7955	4.3670	4.3670	8.6602	2 or 3	1
9.1540	9.2689	12.2440	4.8715	4.8715	3.1350	4	4
9.5687	9.7831	8.5563	4.2835	4.2835	4.9070	2 or 3	4
9.9648	10.1442	6.2135	4.2314	4.2314	6.8232	2 or 3	2 or 3
9.9922	10.6016	5.2304	3.8595	3.8595	8.5895	2 or 3	2 or 3
9.5665	9.7024	8.8109	4.3991	4.3991	4.6926	2 or 3	4
10.1397	10.3131	5.3387	4.2816	4.2816	7.8793	2 or 3	1
9.9847	10.9146	4.7675	3.6559	3.6559	9.9144	2 or 3	2 or 3
9.0589	10.4022	8.9925	3.2497	3.2497	5.8969	2 or 3	2 or 3
9.2538	9.2613	11.7790	4.9465	4.9465	3.2422	4	4
10.5343	10.7314	3.6777	4.5678	4.5678	10.9126	1	1
9.2382	10.1482	8.8688	3.5946	3.5946	5.3771	2 or 3	2 or 3
9.2667	9.3125	11.5106	4.8525	4.8525	3.3587	4	4
9.9942	10.9693	4.6668	3,6388	3.6388	10.1946	2 or 3	2 or 3
10.5294	10.6980	3.7340	4.5807	4.5807	10.7390	1	1
10.0045	10.8500	4.8045	3.7150	3.7150	9.6824	2 or 3	1
9.0964	9.5372	11.4701	4.3431	4.3431	3.6093	4	4
9.3021	10.2835	8.2582	3.4993	3.4993	5.9027	2 or 3	2 or 3
9.7833	10.3975	6.2512	3.7915	3.7915	7.2698	2 or 3	2 or 3
9.2989	9.3875	11.0653	4.7295	4.7295	3.5533	4	4
10.3386	10.3812	4.6797	4.5062	4.5062	8.7295	2 or 3	1
9.0451	9.1262	13.4294	5.1145	5.1145	2.7558	4	4
10.6072	10.8468	3.3767	4.6307	4.6307	11.7469	1	1

Among the candidate solutions, '1' stands for $S_{1,2}$, '2' for S_1^2, '3' for S_2^1, '4' for $S^{1,2}$.

Table 12: Table for Example 2: Least Risk solutions in balanced DOP with $r = 2, f = 2, k = 2$, when the random elements have independent Normal distributions with different means (10 and 10.5) but common variance (4)

		Risk of candidate solutions				Optimal solution	
c_1	c_2	S_{12}	S_1^2	S_2^1	S^{12}	stochastic	non-stochastic
9.7625	10.0184	7.5375	4.0959	6.4900	4.0194	4	2
10.0000	10.4717	5.8275	3.8730	6.2397	5.6888	2	2
9.7096	10.7309	6.5484	3.5982	5.9328	6.2476	2	2
9.1287	9.6320	11.6021	4.3394	6.7660	2.6561	4	4
10.5744	10.8635	3.6596	4.1358	6.5352	8.2825	1	1
9.0402	9.7990	11.5749	4.1165	6.5154	2.9672	4	4
10.0595	10.2779	5.9188	4.0354	6.4215	5.1306	2	2
10.2182	10.2558	5.4184	4.1572	6.5579	5.2931	2	2
10.0220	10.2892	6.0321	4.0047	6.3872	5.1154	2	2
10.0333	10.7403	5.3659	3.7553	6.1083	6.7243	2	2
10.3668	10.3944	4.7497	4.1774	6.5806	6.0064	2	2
9.6629	10.3751	7.2607	3.7726	6.1278	4.9614	2	2
9.3484	9.6319	10.4384	4.3892	6.8213	2.7768	4	4
10.6629	10.9034	3.4091	4.2156	6.6250	8.6760	1	1
9.9010	10.6388	5.9497	3.7291	6.0790	6.1477	2	2
10.2715	10.5342	4.8527	4.0177	6.4018	6.3379	2	2
9.0644	10.0375	10.7892	3.8743	6.2433	3.5472	4	2
9.4725	10.4940	7.8700	3.6336	5.9726	5.1643	2	2
10.0050	10.7596	5.4393	3.7304	6.0805	6.7564	2	2
9.2382	10.2543	9.4010	3.7292	6.0801	4.2893	2	2
9.4704	10.9830	7.2246	3.4172	5.7311	6.9812	2	2
9.2210	9.4979	11.5401	4.5472	6.9996	2.4416	4	4
9.6162	10.0760	8.0326	3.9846	6.3655	4.0349	2	2
9.8459	10.8350	5.9125	3.6158	5.9525	6.8247	2	2
9.2678	9.4441	11.4710	4.6396	7.1035	2.3666	4	4
9.1721	9.9425	10.4659	3.9866	6.3691	3.3691	4	4
9.6290	10.5430	7.1299	3.6591	6.0008	5.4814	2	2
9.6380	10.6710	6.9102	3.5973	5.9319	5.9443	2	2
9.5055	9.5645	9.8598	4.5269	6.9757	2.7417	4	4

Among the candidate solutions, '1' stands for $S_{1,2}$, '2' for S_1^2, '3' for S_2^1, '4' for $S^{1,2}$.

Table 13: Table Reporting Simulation Results: Least Risk solutions in balanced DOP with $r = 2, f = 1, k = 2$. The random elements have independent Normal distributions with means 10.0 and 10.01 and common variance (1)

	Risk of candidate solutions			Optimal solution	
fixed cost c	S_{12}	S_1^2	S_2^1	stochastic	non-stochastic
10.99468	0.475127	4.435111	4.483167	1	1
9.303764	4.841084	2.165788	2.208434	2	2
9.952451	2.424857	2.454669	2.497831	1	2
9.267555	5.003737	2.159396	2.202032	2	2
10.00178	2.280517	2.496789	2.540035	1	2
9.915653	2.536187	2.425664	2.468769	2	2
9.969233	2.375121	2.468573	2.511763	1	2
9.475918	4.107202	2.206233	2.248941	2	2
10.17492	1.817931	2.677059	2.720683	1	1
10.20032	1.755787	2.708123	2.751814	1	1
10.28028	1.569552	2.814299	2.858226	1	1
10.96328	0.504639	4.335819	4.383602	1	1

Among the candidate solutions, '1' stands for $S_{1,2}$, '2' for S_1^2, '3' for S_2^1.

Table 14: Table Reporting Simulation Results: Least Risk solutions in balanced DOP with $r = 3, f = 2, k = 2$ when the random elements have independent Normal distributions with means 10.0 and 10.1, and 10.11 with identical variance 1

		Risk of candidate solutions							Optimal solution	
c_1	c_2	S_{12}^{3}	S_{13}^{2}	S_{23}^{1}	S_1^{23}	S_2^{13}	S_3^{12}	S^{123}	stoch.	non-stoch.
9.34	10.57	11.47	11.47	11.54	6.16	6.21	6.22	8.45	4	4
9.97	10.77	7.84	7.88	7.94	7.14	7.20	7.25	11.43	4	4
9.05	9.67	14.00	14.01	14.08	6.61	6.66	6.67	3.91	7	7
9.97	10.71	7.85	7.88	7.95	7.16	7.22	7.27	11.01	4	4
9.34	9.67	11.42	11.45	11.52	6.62	6.68	6.70	4.25	7	7
10.31	10.74	6.66	6.72	6.78	8.18	8.25	8.31	12.78	1	1
9.45	10.88	10.49	10.50	10.57	6.14	6.19	6.21	10.54	4	4
9.51	10.31	10.35	10.37	10.43	6.47	6.53	6.55	7.34	4	4
9.16	9.34	12.81	12.84	12.91	6.77	6.83	6.84	2.94	7	7
10.12	10.61	7.25	7.30	7.36	7.59	7.65	7.71	10.99	1	1
9.29	10.79	11.67	11.66	11.73	5.99	6.05	6.06	9.64	4	4
9.63	9.96	9.48	9.51	9.58	6.73	6.79	6.82	5.97	7	7
9.22	10.96	12.05	12.04	12.11	5.84	5.90	5.90	10.57	4	4
10.20	10.66	7.00	7.05	7.11	7.81	7.87	7.93	11.71	1	1
10.55	10.92	6.13	6.20	6.26	9.09	9.16	9.23	15.38	1	1
9.70	10.50	9.17	9.20	9.26	6.66	6.72	6.75	8.87	4	4
9.88	10.59	8.24	8.27	8.34	6.98	7.04	7.09	9.94	4	4
9.20	9.96	12.76	12.76	12.83	6.44	6.49	6.49	5.18	7	7
9.99	10.14	7.63	7.68	7.74	7.31	7.37	7.42	7.86	4	4
10.65	10.70	5.93	5.99	6.05	9.54	9.61	9.69	14.39	1	1
9.82	9.83	8.30	8.36	8.42	6.98	7.04	7.09	5.86	7	7
9.10	9.53	13.47	13.49	13.57	6.69	6.75	6.75	3.48	7	7
9.54	10.04	10.13	10.16	10.23	6.61	6.67	6.69	6.12	7	4
10.15	10.90	7.15	7.20	7.26	7.63	7.69	7.75	13.07	1	1
9.58	9.82	9.71	9.75	9.82	6.71	6.77	6.80	5.26	7	7

Among the candidate solutions, '1' stands for S_{12}^{3}, '2' stands for S_{13}^{2}, '3' stands for S_{23}^{1}, '4' stands for S_1^{23}, '5' stands for S_2^{13}, '6' stands for S_3^{12}, '7' stands for S^{123}.

Table 15: Table Reporting Simulation Results: Least Risk solutions in balanced DOP with $r = 3, f = 3, k = 3$, when the random elements are independent variables having (symmetric) marginal distributions V-shaped (mean 10.0, supported on 10 ± 3) Triangular shaped (mean 10.1, supported on 10.1 ± 3.0) and Normal distribution with mean 10.11 and variance 1.

			Risk of candidate solutions								Optimal solution	
c_1	c_2	c_3	S_{123}	S_{12}^3	S_{13}^2	S_{23}^1	S_1^{23}	S_2^{13}	S_3^{12}	S^{123}	stoch.	non-stoch.
10.13	10.27	10.62	7.98	8.23	8.33	9.26	9.38	10.34	10.44	13.91	1	1
9.55	9.94	10.42	12.84	9.23	9.28	9.96	8.25	8.90	8.95	10.66	5	5
9.65	10.40	10.99	10.16	7.66	7.76	8.66	9.49	10.45	10.55	16.76	2	2
9.64	9.85	10.39	12.72	9.62	9.65	10.31	8.09	8.69	8.73	10.25	5	5
9.60	10.59	10.88	9.88	7.23	7.35	8.32	10.23	11.30	11.42	17.10	2	2
9.13	9.53	9.99	18.94	11.42	11.40	11.86	7.69	8.03	8.00	7.00	8	8
9.69	10.65	10.83	9.23	7.15	7.27	8.27	10.52	11.62	11.76	17.14	2	2
9.37	10.79	10.88	10.81	6.84	6.97	7.99	11.08	12.24	12.39	18.28	2	2
9.54	9.74	9.88	14.69	10.59	10.59	11.09	8.21	8.63	8.63	7.11	8	8
10.40	10.72	10.76	5.51	7.34	7.50	8.70	11.43	12.75	12.93	18.05	1	1
9.52	9.78	10.23	14.03	10.01	10.03	10.63	8.00	8.53	8.55	9.04	5	5
9.35	10.79	10.90	10.96	6.84	6.97	7.99	11.06	12.21	12.36	18.43	2	2
9.22	9.52	9.88	18.45	11.61	11.58	12.00	7.80	8.11	8.07	6.48	8	8
9.20	10.35	10.55	13.44	7.84	7.92	8.76	9.29	10.18	10.27	13.16	2	2
9.08	9.11	9.48	23.85	14.81	14.71	14.87	7.93	7.88	7.76	4.18	8	8
9.40	9.43	9.87	17.62	12.09	12.06	12.46	7.73	7.99	7.95	6.31	8	8
9.53	10.63	10.64	10.28	7.20	7.32	8.29	10.42	11.49	11.62	15.45	2	2
9.61	10.25	11.07	10.89	8.03	8.12	8.96	8.97	9.84	9.93	16.57	2	2
9.29	9.75	9.96	16.36	10.43	10.44	10.95	8.09	8.53	8.53	7.42	8	8
9.60	9.62	10.24	14.40	10.76	10.77	11.32	7.72	8.18	8.18	8.63	5	5
9.32	10.22	10.84	12.91	8.12	8.19	9.00	8.83	9.67	9.74	14.59	2	2
9.33	9.95	10.99	14.03	8.97	9.02	9.73	8.11	8.78	8.84	14.44	5	5
10.05	10.66	10.97	7.24	7.21	7.35	8.41	10.76	11.93	12.08	18.66	2	2
9.73	10.01	10.46	11.31	8.99	9.05	9.78	8.46	9.18	9.23	11.29	5	5
9.23	9.68	10.94	16.28	10.09	10.12	10.72	7.56	8.08	8.10	13.03	5	5

Among the candidate solutions, '1' stands for S_{123}, '2' stands for S_{12}^3, '3' stands for S_{13}^2, '4' stands for S_{23}^1, '5' stands for S_1^{23}, '6' stands for S_2^{13}, '7' stands for S_3^{12}, '8' stands for S^{123}.

Table 16: Table Reporting Simulation Results: Least Risk solutions in balanced DOP with $r = 3, f = 2, k = 3$, when the random elements are independent variables having Normal distribution with variance 1.

Mean of Random elements			fixed costs		Risk of candidate solutions							Optimal solution	
μ_1	μ_2	μ_3	c_1	c_2	S_{123}	S_{12}^3	S_{13}^2	S_{23}^1	S_1^{23}	S_2^{13}	S_3^{12}	stoch.	non-stoch.
9.57	10.10	9.65	10.45	10.64	3.32	7.85	5.24	8.37	11.56	15.74	12.16	1	1
9.04	9.33	10.60	10.96	10.99	2.03	3.67	11.24	13.28	14.25	16.51	29.02	1	1
10.03	9.21	10.12	10.53	10.99	3.01	5.29	10.90	5.82	17.72	11.23	18.60	1	1
10.04	9.82	9.09	10.27	10.55	3.64	10.91	6.25	5.02	16.63	14.76	9.35	1	1
9.57	9.67	9.15	10.35	10.98	2.30	9.04	5.90	6.48	19.11	20.07	15.45	1	1
9.72	10.61	10.89	10.88	10.96	4.82	4.87	6.48	12.41	6.88	12.94	15.25	1	2
9.47	10.02	10.74	10.73	10.89	3.85	3.90	8.02	11.74	8.93	12.82	19.01	1	2
9.48	10.52	10.60	9.58	10.54	10.89	5.22	5.63	12.60	5.15	12.02	12.58	5	2
9.09	10.52	9.19	10.52	10.71	3.07	10.93	3.12	11.68	12.20	25.12	12.99	1	3
9.79	10.44	10.00	9.17	10.20	12.93	8.02	5.37	9.44	6.48	10.81	7.80	3	3
9.19	10.60	9.22	9.19	10.92	9.03	10.02	2.14	10.22	11.38	24.26	11.60	3	3
10.03	10.81	9.13	9.93	10.44	7.99	14.23	3.41	7.83	11.21	17.68	5.47	3	3
9.57	9.56	9.40	9.53	9.94	6.44	7.62	6.62	6.60	10.02	10.00	8.88	1	4
10.67	9.22	9.92	9.72	10.04	9.50	8.59	13.67	4.15	14.74	4.85	9.06	4	4
10.03	10.94	10.77	10.70	10.76	7.01	6.71	5.73	11.72	5.75	11.72	10.50	3	5
10.57	9.91	10.89	9.36	9.76	19.29	7.90	15.38	10.07	7.48	3.68	9.55	6	6
10.29	10.01	10.13	9.26	9.66	15.49	9.50	10.39	8.41	7.58	5.86	6.59	6	6
10.24	10.59	9.87	9.97	10.23	9.43	10.53	5.83	8.01	8.13	10.56	5.88	3	7
9.94	10.76	9.82	9.49	9.93	13.39	11.54	5.34	10.62	6.17	11.62	5.45	3	7
10.36	9.80	9.19	9.44	9.79	10.24	12.62	8.29	4.88	12.66	8.58	4.93	4	7
9.60	10.92	9.26	9.16	9.33	17.12	16.84	5.00	13.87	5.84	14.90	3.95	7	7
9.65	10.44	9.44	9.05	9.25	16.21	13.57	6.48	11.89	5.81	10.91	4.63	7	7

Among the candidate solutions, '1' stands for S_{123}, '2' stands for S_{12}^3, '3' stands for S_{13}^2, '4' stands for S_{23}^1, '5' stands for S_1^{23}, '6' stands for S_2^{13}, '7' stands for S_3^{12}.