Department of Applied Mathematics
Faculty of EEMCS

University of Twente
The Netherlands

P.O. Box 217
7500 AE Enschede
The Netherlands

Phone: +31-53-4893400
Fax: +31-53-4893114

Email: memo@math.utwente.nl
www.math.utwente.nl/publications

Memorandum No. 1784

On solving discrete optimization problems with one random element under general regret functions

D. Ghosh[1], P.K. Mandal and S. Das[2]

December, 2005

ISSN 0169-2690

[1]P&QM Area, IIM Ahmedabad, Vastrapur, Ahmedabad 380015, INDIA
[2]QM&IS Area, IIM Bangalore, Bannerghatta Road, Bangalore 560076, INDIA

On Solving Discrete Optimization Problems with One Random Element Under General Regret Functions

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.

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

Abstract

In this paper we consider the class of stochastic discrete optimization problems in which the feasibility of a solution does not depend on the particular values the random elements in the problem take. Given a regret function, we introduce the concept of the risk associated with a solution, and define an optimal solution as one having the least possible risk. We show that for discrete optimization problems with one random element and with min-sum objective functions a least risk solution for the stochastic problem can be obtained by solving a non-stochastic counterpart where the latter is constructed by replacing the random element of the former with a suitable parameter. We show that the above surrogate is the mean if the stochastic problem has only one symmetrically distributed random element. We obtain bounds for this parameter for certain classes of asymmetric distributions and study the limiting behavior of this parameter in details under two asymptotic frameworks.

AMS Subject Classification: 90C31
Key words: stochastic discrete optimization problems, min-sum, regret, risk

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) does not remain unique.

A well-considered optimization criterion for such problems is to maximize the probability that the (random) objective function reaches a pre-specified threshold level (see, e.g., Frank [6] in the context of the shortest path problem and Henig [7], Carraway et al. [2] in the context of the stochastic knapsack problem). Another closely related notion is to find a solution that leads to the optimal threshold value satisfying the constraint that the probability of the objective function reaching the threshold value is at least a pre-specified value α (see, e.g., Henig [7] on the knapsack problem, and Ishii et al. [8] on the minimum spanning tree problem). The most prevalent notion in practice and literature, though, is the *expected utility* criterion of Von Neumann and Morgenstern (see Fishburn [5]). With this criterion one maximizes the decision maker's expected utility (see, e.g., Murthy et al. [12], Loui [10] on the shortest path problem, and Dean et al. [4] on the knapsack problem).

On the other hand, behavioral notion of (posterior) *regret* (introduced in Savage [13]) for not achieving what the decision maker could have achieved with another choice, plays an important role in making a decision (see, e.g., Kaliszewski et al. [9]). This motivates us to work with a notion of optimality involving regret function. To the best of the authors' knowledge this notion of optimality has not been explored in the stochastic discrete optimization literature so far.

In this paper, we assume knowledge about the stochasticity of the parameters (which is often the case in practice). The optimality in terms of regret functions leads us to the notion of least risk solutions (formally introduced in Section 2). The method we adopt to find the least risk solutions to a stochastic DOP is to formulate a non-stochastic DOP, such that an optimal least cost solution of the latter is a desired least risk solution of the former. The non-stochastic DOP is essentially the stochastic DOP with the random parameter pegged at a value based on our knowledge of its randomness as well as other problem-specific parameters.

The remainder of this paper is organized as follows. We introduce in Section 2 the notations and definitions to be used throughout the article including a precise description of the class of the stochastic DOP we consider. In Section 3 we prove our main result about solving a stochastic DOP with one random element by solving an auxiliary non-stochastic DOP. We show in Subsection 3.1 that if the probability distribution of the cost of the random element is symmetric then knowledge of the mean (or median) of the distribution is sufficient to obtain the auxiliary DOP. This is irrespective of the choice of the regret function. Subsection 3.2 deals with the case where the cost of the random element follows a homogeneously skewed distribution. Some asymptotic results are provided in Section 4. Finally, we summarize our contributions in Section 5.

2 Notations and Definitions

In this section, we describe the notations 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 π is called *random* (alternatively *fixed*) if the associated cost c_e is random valued (alternatively constant).

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

As an example of SDOPs, consider the traveling salesperson problem, which is one of deciding a round trip through several cities, visiting each city exactly once before returning to the city of origin, with an aim to reduce the total travel time. The problem becomes a

SDOP (called stochastic traveling salesperson problem) if the time(s) required to traverse one (or more) intercity route (routes) is (are) random. Another example would be the stochastic binary knapsack problem. In this problem, one is given a set of projects, each of which incurs some cost and generates some benefit. While the costs of the projects are known and fixed, the benefits of some of the projects are random. One is required to choose a subset of the projects such that the total cost of implementing these projects is within a given budget, and the expected benefit of the basket is as high as possible.

In this work, we restrict ourselves to SDOPs where all feasible solutions remain feasible, irrespective of the randomness involved. In the traveling salesperson problem setup described above this is satisfied. On the other hand, in the knapsack problem, we need to have deterministic (nonrandom) costs and budget for this condition to hold. If the budget or the costs 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 [3] for treatment of such problems.

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

$$regret(S) = r(z(S) - Z^*),$$

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, if some of the c_e's are random, the regret associated with any feasible solution S would also be 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

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

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

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 are fixed, the minimum risk solution is an optimal solution in the traditional sense, i.e., a least cost solution.

Though all the definitions and notions above are for general objective function z, in our analysis henceforth 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.

In parts of this work, particular attention is given to a specific class of skewed unimodal distributions which we refer to as homogeneously skewed distributions.

Definition 7 Suppose a unimodal distribution with density $h(\cdot)$ has mode M, that is, $h(\cdot)$ is non-decreasing in $(-\infty, M]$ and non-increasing 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 8 The *skewness function* of a homogeneously skewed distribution with mode M and density function $h(\cdot)$ is defined as

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

Definition 9 The *measure of homogeneous skewness* of a homogeneously skewed distribution with mode M and density function $h(\cdot)$ 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 Solving Stochastic DOP with Single 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 assume that the cost of the random element e has a (cumulative) distribution function $H(\cdot)$ with mean μ, i.e.

$$H(x) = P(c_e \leq x), \text{ and } \mu = \int x dH(x).$$

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 of c_e becomes irrelevant, since e would not be included in the optimal solution in any case.) When c_e increases, the costs of all solutions in $\mathbb{S}_e$ increase while the costs of all solutions in $\mathbb{S}^e$ remain the same. So S_e remains optimal until c_e increases to some threshold value, say ω, at which point $z(S_e)$ 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).

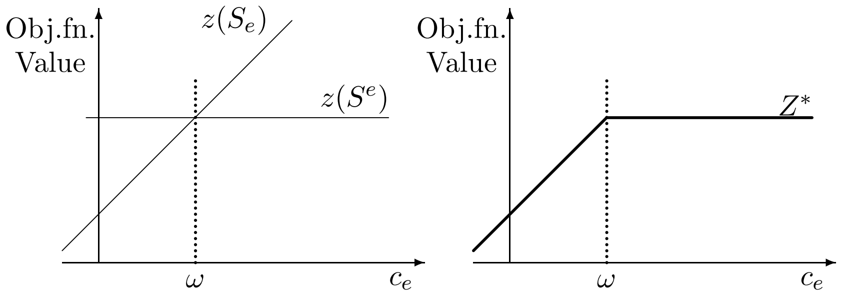

Figure 1: $z(S_e)$, $z(S^e)$, and Z^* as a function of c_e (min-sum objective)

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

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

Further, by adopting S_e as a solution one incurs a regret equal to $r(c_e - \omega)$ if $c_e > \omega$, since S_e is optimal in the least cost sense if $c_e \leq \omega$. Similarly, by taking S^e as a solution, there is a regret of $r(\omega - c_e)$ when $c_e < \omega$. 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 SDOP with a fixed value, such that the least cost solution to the DOP thus obtained is the least risk solution to the original SDOP. To that end, we define for a random variable X and a regret function $r(\cdot)$, 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 $r(\cdot)$. Naturally, $\Psi_{r,X}(\cdot)$ depends on X through its distribution function $H(\cdot)$. For notational convenience, either or both of the suffixes of 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 stochastic DOP 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

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

But from (5), the right hand side 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 (non-stochastic) DOP when the random cost is replaced by θ.

In the remainder of this section, we study the behavior of θ, the solution to equation (9), under various distributions of the random element.

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 SDOP with one random element 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 Stochastic DOP with a single random element e with cost $c_e \equiv X$ having a symmetric distribution 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

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

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

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

and hence

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

The result then follows from Theorem 1.

Remark 1 Theorem 2 also holds for asymmetric distributions if the regret function is linear. In fact, if the regret function is linear, then the result holds even in the general case (with multiple random elements which are not necessarily symmetric) as shown in Mandal et al. [11].

Remark 2 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.

In the following section we deal with the case where the cost of the random element follows a homogeneously skewed distribution.

3.2 Random Element with Skewed Distribution

Suppose the cost of the random element has a continuous distribution with density $h(x)$ which is unimodal with mode M, and homogeneously right-skewed. We also assume that the density function has the requisite finite moments so that the optimal solutions have finite risk as per the choice of the regret function. Then the following theorem 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 3 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, though true, 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

$$\Psi\Big(\frac{L+U}{2}\Big) = \int_0^{\frac{U-L}{2}} r(y)\Big\{h\Big(\frac{L+U}{2}+y\Big) - h\Big(\frac{L+U}{2}-y\Big)\Big\}dy \leq 0$$

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

Remark 4 Natural analogues to Theorems 3 and 4 exist when the distribution of c_e is homogeneously left-skewed or has a non-decreasing density.

We now illustrate the behavior of θ under certain distributions for the random element. In particular, we consider two types of density functions,viz. Triangular and Beta. They are defined on a finite support $[0, 1]$ and have the following functional forms:

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

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

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. The associated variations in the θ are shown in Figures 2 through 4. In Figure 3, the τ refers to the measure of homogeneous skewness as defined in Definition 9.

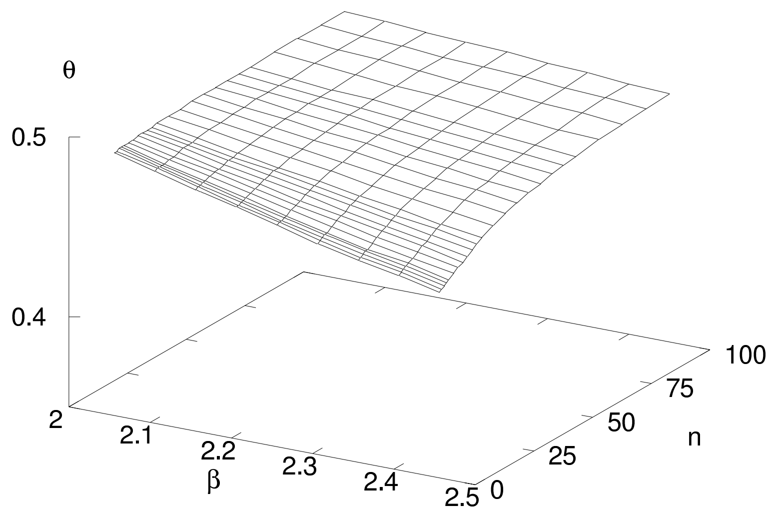

Figure 2: Plot of θ values against β and n when $c_e \sim Beta(2, \beta)$ and $r(t) = (1+t)^n - 1$

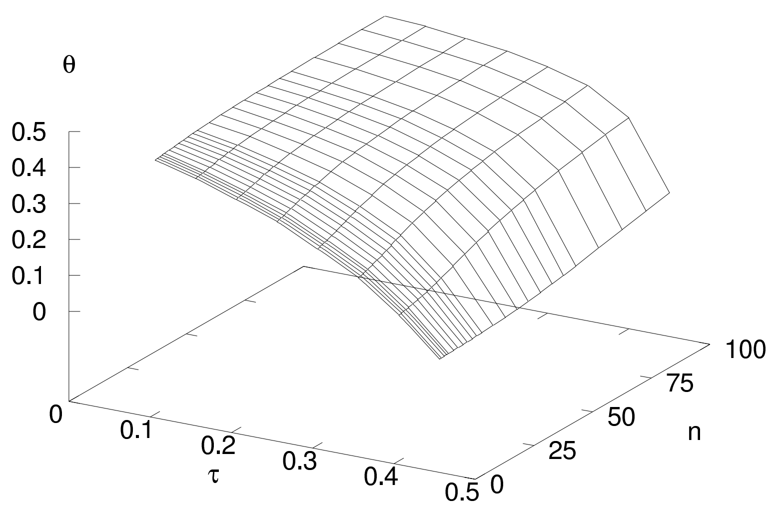

Figure 3: Plot of θ values against τ and n when $c_e \sim Beta(2, \beta)$ and $r(t) = (1+t)^n - 1$

The asymptotic behavior of θ observed from these computational results are formalized through the results in Section 4.

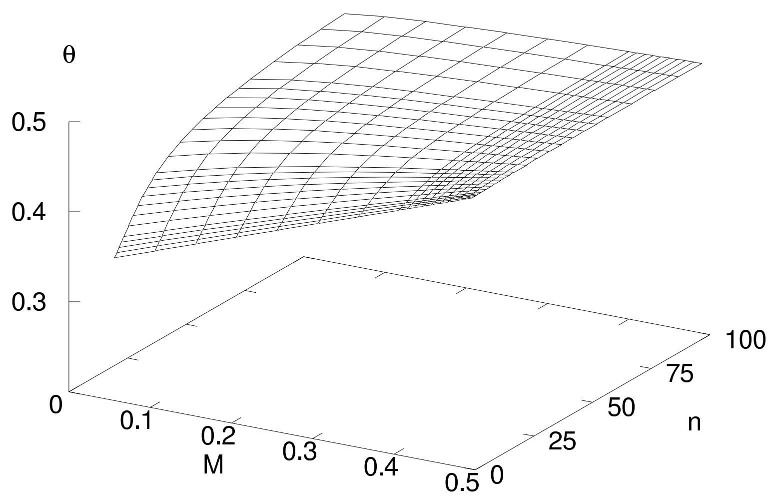

Figure 4: Plot of θ values against M and n when the distribution of c_e is triangular with mode M and $r(t) = (1+t)^n - 1$

In the remainder of this section, we will consider a case of homogeneously skewed distributions that occurs due to practical considerations in a problem. This is a case in which the cost of an element follows an essentially symmetric distribution, but is constrained not to be below a certain value. As an example, consider the stochastic traveling salesperson problem, described in Section 2, where the time taken to traverse a certain road segment follows a Gaussian distribution, but it is constrained to assume non-negative values. One can consider this scenario as a minor deviation from one having a symmetric random cost.

We model the randomness of c_e using a Gaussian distribution with mean μ and standard deviation 1 truncated at zero, so that the distribution is supported on $[0, \infty)$. Obviously, the true distribution is not symmetric any more and θ does not coincide with μ. The errors $(\theta - \mu)$ are observed to be small in magnitude, and reduce as μ increases (see Figure 5). As expected, they increase in magnitude when the cutoff probability, i.e. the Gaussian probability below zero, increases (see Figure 6). This observation is formalized in Theorem 6 in the next section, which makes it apparent that if the distributional assumption of symmetry is even approximately valid, then the deviation of the observed critical value from μ (the critical value for symmetric distributions) will be minor.

4 Asymptotic Results

As we have seen in the previous section, the value of θ, the solution to equation (9), is critical in solving the SDOP through a solution of its non-stochastic counterpart. In this section, we study the limiting behavior of θ under two asymptotic scenarios. These results, proved under various regularity conditions, show that the theorems of Sections 3.1 and 3.2 would be (approximately) valid if the requisite conditions on the probability distribution are more or less true.

Theorem 5 *Consider a stochastic DOP with a single random element (X) supported on*

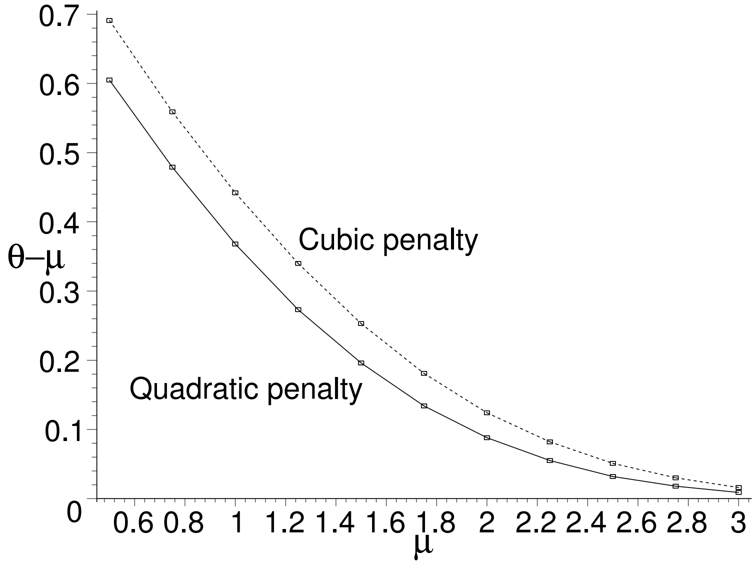

Figure 5: Plot of error vs. μ when non-negative values are not allowed in a Gaussian distribution

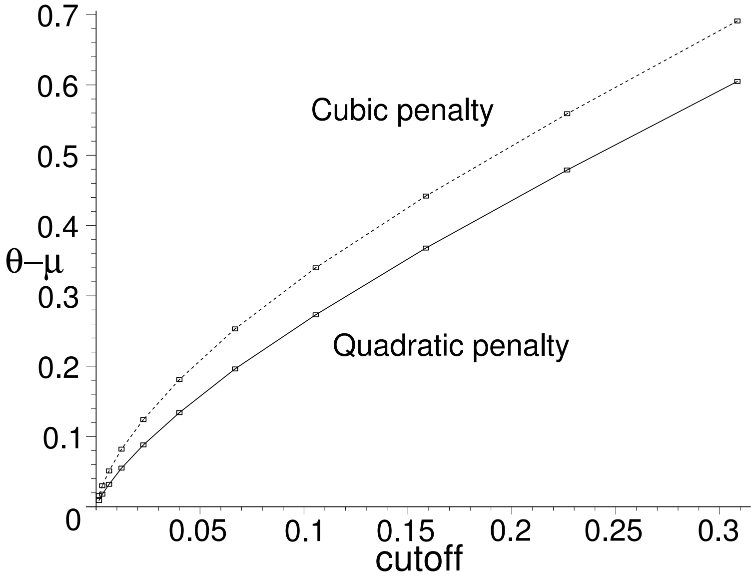

Figure 6: Plot of error vs. cutoff probability when nonnegative values are not allowed in a Gaussian distribution

the interval $[L, U]$*. For a sequence of regret functions* $\{r_n(\cdot),\ n \geq 1\}$ *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*

$$\theta_n \to \frac{L+U}{2} \qquad \text{as } n \to \infty,$$

as long as at least one of L *and* U *is finite.*

Remark 5 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 $\epsilon > 0$,

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

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

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

Suppose $L < t < \frac{L+U}{2}$. Denote $\delta_1 := \frac{L+U}{2} - t > 0$. Then from (10)

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

where

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

Since $\beta_1 - \gamma_1 = \delta_1 > 0$ we have, from (12), $\lim_{n\to\infty} \frac{r_n(\beta_1)}{r_n(\gamma_1)} = \infty$. Further, with $a_1, b_1 > 0$, 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 from (10) we have

$$\Psi_n(M) = \mathbb{E}\left[r_n(X-M)\mathbb{I}_{\{X \ge M\}}\right] - \mathbb{E}\left[r_n(M-X)\mathbb{I}_{\{L \le X \le M\}}\right]$$

$$\begin{aligned} &\geq \mathbb{E}\left[r_n(X-M)\mathbb{I}_{\{X\geq\xi\}}\right] - \mathbb{E}\left[r_n(M-X)\mathbb{I}_{\{L\leq X\leq M\}}\right] \\ &\geq r_n(\xi - M)\mathbb{P}(X \geq \xi) - r_n(M-L)\cdot\mathbb{P}(L \leq X \leq M) \\ &= r_n(\beta_2)\cdot a_2 - r_n(\gamma_2)\cdot b_2, \quad \text{say,} \end{aligned} \tag{15}$$

where

$$\begin{array}{lll} a_2 = \mathbb{P}(X \geq \xi) > 0; & b_2 = \mathbb{P}(L \leq X \leq M) > 0; & \\ \beta_2 = \xi - M; & \gamma_2 = M - L; & \text{and} \quad \beta_2 > \gamma_2. \end{array}$$

Therefore for any given M, $\Psi_n(M) > 0$ for large n, implying $\theta_n \geq M$. 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 (each with a single random element) which are identical to each other except for the first element, which is random, and which has cost X_n. Suppose X_n converges to 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*

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

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 6 If the limiting random variable X in Theorem 6 is symmetric, then from Theorem 2, θ_n converges to μ, the mean of the limiting distribution.

Remark 7 In the regularity conditions (A) and (D) of Theorem 6, 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: It suffices to show that for any given $\epsilon > 0$, eventually $\theta - \epsilon \leq \theta_n \leq \theta + \epsilon$, i.e.,

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

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.

Proof of (18) under the regularity conditions: We would provide the proof for the sequence $\{a_n\}$ only, as the same for $\{b_n\}$ 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 summation 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, together with (19), would also imply that

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

for large n. Now from (20),

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

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

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

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

In the continuous case note that

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

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

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

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

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

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

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

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

5 Summary

We consider a broad class of stochastic discrete optimization problems (SDOPs) in which feasible solutions remain feasible regardless of the randomness in the problem parameters, and for which the objective is to minimize the sum of the costs of elements in a solution. Instead of working with the notion of optimality existing in the literature for SDOPs, we follow posterior regret arguments and work with least risk solution as optimal solution.

After describing all the necessary notations and definitions in Section 2, we analyze SDOPs with one random element, both when the distribution of the cost of the random element is symmetric (Section 3.1) and when the distribution is skewed of specific type (Section 3.2). We derive an auxiliary non-stochastic DOP from the original SDOP so that optimal solutions in both problems coincide. The auxiliary DOP is generated (Theorem 1) by pegging the cost of the random element in the SDOP to θ which is defined in terms of the regret function and the distribution of the cost of the random element. We show (Theorem 2) that if the distribution of the cost of the random element is symmetric, then θ coincides with the mean (or median) of the distribution. It appears that an equally simple characterization of θ when the distribution of the cost of the random element is not symmetric is not possible. In Theorems 3 and 4, we provide bounds to the value of θ for such problems.

In Section 4 we examine the limiting behavior of θ under two asymptotic scenarios. In the first scenario, the regret function is made increasingly steeper. In this case we show (through Theorem 5) that θ converges to the mid-point of the support of the distribution of the random cost, whenever the midpoint of the support is defined. In Theorem 6 we show a second type of limiting behavior. It states that if the distribution of the random element converges to a given distribution, then, under certain broad regularity conditions involving the regret function and the distribution of the cost, the θ values converge. In practice, we may not know the exact probability distribution of the cost of the random element and without knowledge of the distribution θ becomes an unknown parameter. Theorem 6 informally suggests that the obtained solution by using θ from an estimated or an approximate distribution should yield a near optimal result.

A natural extension to this paper would be one where more than one costs are random. It can be shown that the results for SDOPs with one random element do not directly generalize to SDOPs with multiple random elements though some partial results exist (see Mandal et al. [11]). However, to keep the current article focused and the size manageable we treat the multiple random element case in a separate article.

References

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

[2] R. Carraway, S. Schmidt, and L. Weatherford,(1993), *An algorithm for maximizing target achievement in the stochastic knapsack problem with normal returns*, Naval Research Logistics **40** (1993), 161-173.

[3] S. Das and D. Ghosh, *Binary knapsack problems with random budgets*, Journal of the Operational Research Society **54** (2003), 970–983.

[4] B. C. Dean, M. X. Goemans, and J. Vondrák, (2004), *Approximating the stochastic knapsack problem: The benefit of adaptivity*, Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (2004), 208–217.

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

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

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

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

[9] I. Kaliszewski and W. Michalowski, *Establishing regret attitude of a decision maker within the MCDM modeling framework*, Interim Report IR-98-070/September, International Institute for Applied Systems Analysis, 1998.

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

[11] P. K. Mandal, D. Ghosh, and S. Das, *On Solving Discrete Optimization Problems with Multiple Random Elements Under General Regret Functions*, Unpublished Manuscript (2005).

[12] I. Murthy and S. Sarkar, *Stochastic shortest path problems with piecewise-linear concave utility functions*, Management Science **44** (1998), 125–136.

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