An Upper Bound on the Expected Value of a Non-Increasing Convex Function with Convex Marginal Return Functions[1]

Christopher J. Donohue[2]
and
John R. Birge
Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor MI 48109-2117

Technical Report 95-6

An Upper Bound on the Expected Value of a Non-Increasing Convex Function with Convex Marginal Return Functions[1]

Christopher J. Donohue[2]
Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor, Michigan 48109

John R. Birge
Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor, Michigan 48109

Abstract: In this note, we show that if a convex function is non-increasing and has a special property we call convex marginal return functions, an effective upper bound can be established using only two function evaluations. Further, we show that this bound can be refined in such a way that the number of function evaluations needed grows linearly with the number of refinements performed.

Keywords: Upper Bound, Convex Function, Approximation for Stochastic Programming

[1] This work has been partially supported by the National Science Foundation under Grant DDM-9215921 and the Great Lakes Center for Trucking and Transit Research

[2] Christopher J. Donohue, Department of Industrial and Operations Engineering, 1205 Beal Ave., University of Michigan, Ann Arbor, Michigan 48109-2117

1 Introduction

Bounding the expected value of a convex function of a multivariate random variable is a problem which has many applications in mathematical programming, in particular stochastic programming. Unfortunately, the effort for finding upper and lower bounds generally favors the lower bound, both in computational difficulty and effectiveness. Jensen's inequality gives the classic lower bound which only requires one function evaluation. This lower bound is generally effective, and as was shown in Huang, Ziemba and Ben-Tal [7], this bound can be replicated over finer and finer partitions of the probability space until the bound converges to the expected value. Edmundson and Madansky [8] developed the classic upper bound which requires 2^n function evaluations, where n is the number of random variables. This bound requires the random components to be independent and to have bounded support. Since the number of function evaluations grows exponentially with the number of random components, other methods for finding an upper bound have been developed.

Gassmann and Ziemba [6] developed an upper bound which requires solving a linear program. This bound applies to random variables with arbitrary convex domains which may be bounded or unbounded. Frauendorfer [5] obtained an extension of the Edmundson-Madansky bound for the case of dependent random components. Further, upper bounds have been developed for specific convex functions with special attributes. Birge and Wets [3] considered the recourse function of a two-stage stochastic linear program and showed that an upper bound could be obtained using separable sublinear functions. Wallace [10] showed that the network recourse function could be bounded above by solving three network flow problems. Birge and Wallace [2] extended this result for general linear recourse function.

In this paper, we consider a special structure that obtains an upper bound with only two function evaluations. We first observed this structure in minimum cost network flow problems with stochastic link capacities. We refer to the general property as convex marginal return functions. This property can be proven to exist for various aspects of minimum cost network flow problems with stochastic link capacities (see Donohue and Birge [4] for details). Applications include vehicle allocation problems and other logistical models.

2 Development of the Upper Bound

Let $X = (X_1, X_2, \ldots, X_n)$ be a multivariate random variable on the probability space (Ω, Σ, P) having distribution function F and finite mean $\bar{x}$. Assume that the components of X are independently distributed, each with bounded support. For $i = 1, \ldots, n$, let

$$x_i^L = \sup\{x_i : F(x) \leq 0\}, \tag{1}$$

$$x_i^H = \inf\{x_i : F(x) \leq 1\}. \tag{2}$$

Let $\phi(X)$ be a bounded convex function of $X \in \left[x_i^L, x_i^H\right]$, $i = 1, \ldots, n$. Then the Edmundson-Madansky inequality $\mathbf{M_0}$ gives the classic upper bound on $E\left[\phi(X)\right] \equiv \bar{\phi}$, as follows:

$$\mathbf{M_0} \equiv \sum_{i_1 \in \{L,H\}} \cdots \sum_{i_n \in \{L,H\}} (p_1^{i_1} * \cdots * p_n^{i_n}) * \phi(x_1^{i_1}, \cdots, x_n^{i_n}) \geq \bar{\phi}, \tag{3}$$

where

$$p_i^L \equiv \frac{x_i^H - \bar{x}_i}{x_i^H - x_i^L} \quad \text{and} \quad p_i^H \equiv \frac{\bar{x}_i - x_i^L}{x_i^H - x_i^L}. \tag{4}$$

This corresponds to considering only the corner points of the n-dimensional rectangle defined by the values of x_i^L and x_i^H for all i.

For a random vector X of high dimensions, this bound still requires the function values for 2^n realizations of X. For functions which are not directly calculable (such as the objective function of a mathematical program), this may be computationally prohibitive. The ability to reduce the number of realizations of X that need to be calculated would therefore be useful. The following paper shows that if $\phi(X)$ is a non-increasing, convex function of X with convex marginal return functions with respect to any pair (X_i, X_j), $i = 1, \ldots, n$, then an upper bound on $\bar{\phi}$ can be established using only two realizations of X.

First the property of convex marginal return functions with respect to any pair of X components must be defined.

Definition: A function $f(X_1, X_2, \ldots, X_n)$ has *convex marginal return functions with respect to any pair* $(X_i, X_j), i = 1, \ldots, n, \quad j = 1, \ldots, n$ if for all $1 \le i \le n, 1 \le j \le n, \lambda \ge 0, \delta \ge 0$,

$$f(x_1, \ldots, x_j + \delta, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n) \tag{5}$$

$$\le f(x_1, \ldots, x_i + \lambda, \ldots, x_j + \delta, \ldots, x_n) - f(x_1, \ldots, x_i + \lambda, \ldots, x_j, \ldots, x_n). \tag{6}$$

Note that if $f(X)$ is a differentiable function over all possible values of X, this property is equivalent to requiring that

$$\frac{\partial f(X_1, \ldots, X_n)}{\partial X_j}$$

be a non-decreasing function of X_i.

Let X be an n-dimensional random vector, where each X_i is independently distributed with a bounded support and finite mean. Further, let x_i^L, x_i^H be the realizations of X_i as defined in equations (1) and (2), so that $x_i^L < x_i^H$ for all i. Let $p_i^L \equiv prob\{X_i = x_i^L\}$ and $p_i^H \equiv prob\{X_i = x_i^H\}$ for all i be such that

$$p_i^L x_i^L + p_i^H x_i^H = E[X_i], \quad p_i^L + p_i^H = 1, \quad p_i^L, p_i^H \ge 0.$$

Let $f(X)$ be a non-increasing convex function of X with convex marginal return functions with respect to any pair $(X_i, X_j), i = 1, \ldots, n, \ j = 1 \ldots, n$. Then the following shows that an upper bound on $E[f(X)]$ can be established with only two function evaluations. While this bound may not be as effective as the Edmundson-Madansky bound, the loss in tightness may be more than outweighed by the savings in computation time.

Theorem 2.1 *Let X and $f(X)$ be as defined above. Let $p^L \equiv \max\{p_1^L, p_2^L, \ldots, p_n^L\}$ and $p^H \equiv 1 - p^L$. Then*

$$E[f(X)] \le p^L f(x_1^L, x_2^L, \ldots, x_n^L) + p^H f(x_1^H, x_2^H, \ldots, x_n^H) \equiv \mathbf{HL_0}. \tag{7}$$

Proof: By induction. Let $n = 1$. The result holds by Edmundson-Madansky.

Suppose the result is true for $n = t \geq 1$, and consider the case where $n = t+1$. Then, by Edmundson-Madansky,

$$E[f(X)] \leq p_1^L E[f(x_1^L, X_2, \ldots, X_n)] + p_1^H E[f(x_1^H, X_2, \ldots, X_n)]$$

For a particular value of $X_1 = x_1$, $f(x_1, X_2, \ldots, X_n)$ is a function of t variables. By the induction hypothesis,

$$E[f(x_1, X_2, \ldots, X_n)] \leq \hat{p}^L f(x_1, x_2^L, \ldots, x_n^L) + \hat{p}^H f(x_1, x_2^H, \ldots, x_n^H),$$

where $\hat{p}^L \equiv \max\{p_2^L, p_3^L, \ldots, p_n^L\}$ and $\hat{p}^H \equiv 1 - \hat{p}^L$. Thus,

$$\begin{aligned} E[f(X)] \quad \leq \quad & p_1^L \hat{p}^L f(x_1^L, x_2^L, \ldots, x_n^L) + p_1^L \hat{p}^H f(x_1^L, x_2^H, \ldots, x_n^H) + p_1^H \hat{p}^L f(x_1^H, x_2^L, \ldots, x_n^L) + \\ & p_1^H \hat{p}^H f(x_1^H, x_2^H, \ldots, x_n^H). \end{aligned} \tag{8}$$

Note that now the set of random variables $(X_2, X_3, \ldots, X_n)$ are perfectly correlated and therefore can be replaced by a single random variable.

Now, without loss of generality, assume that $p_1^L \geq \hat{p}^L$. To simplify notation, let

$$f(L, L) \equiv f(x_1^L, x_2^L, \ldots, x_n^L), \quad f(L, H) \equiv f(x_1^L, x_2^H, \ldots, x_n^H),$$

$$f(H, L) \equiv f(x_1^H, x_2^L, \ldots, x_n^L), \quad f(H, H) \equiv f(x_1^H, x_2^H, \ldots, x_n^H).$$

Then, by construction,

$$
\begin{aligned}
E[f(X)] &\leq p_1^L\hat{p}^L f(L,L) + p_1^L\hat{p}^H f(L,H) + p_1^H\hat{p}^L f(H,L) \\
&\quad + p_1^H\hat{p}^H f(H,H), && \text{from (8),} \\
&\leq p_1^L\hat{p}^L f(L,L) + p_1^L\hat{p}^H f(L,H) + p_1^H\hat{p}^L (f(L,L) \\
&\quad + f(H,H) - f(L,H)) + p_1^H\hat{p}^H f(H,H), && \text{since } f \text{ has convex marginal return functions, } f(H,L) \leq f(L,L) + f(H,H) - f(L,H), \\
&= (p_1^L\hat{p}^L + p_1^H\hat{p}^L) f(L,L) \\
&\quad + (p_1^L\hat{p}^H - p_1^H\hat{p}^L) f(L,H) + p_1^H f(H,H), \\
&\leq (p_1^L\hat{p}^L + p_1^H\hat{p}^L - p_1^H\hat{p}^L) f(L,L) + p_1^H f(H,H), && \text{since } f \text{ is non-increasing, } f(L,H) \leq f(L,L), \text{ and by assumption, } p_1^L\hat{p}^H - p_1^H\hat{p}^L \geq 0, \\
&= p_1^L f(L,L) + p_1^H f(H,H). \quad \square
\end{aligned}
$$

In [7], it was shown that upper and lower bounds on the expected value of a convex function of a multivariate random variable (with independent components) could be refined to arbitrary accuracy by repeating the bound over sharper and sharper partitions of the probability space of the random variable. For example, for a one-dimensional random variable X, a sharper bound $\mathbf{M_1}$ on the expectation of $\phi(X)$ can be found by applying the Edmundson-Madansky bound on each of the subintervals $[x_1^L, x_1^M], [x_1^M, x_1^H]$ as follows:

$$
\mathbf{M_1} \equiv c_1^L \left(\left(\frac{x_1^M - \mu_1^L}{x_1^M - x_1^L} \right) \phi(x_1^L) + \left(\frac{\mu_1^L - x_1^L}{x_1^M - x_1^L} \right) \phi(x_1^M) \right) + c_1^H \left(\left(\frac{x_1^H - \mu_1^H}{x_1^H - x_1^M} \right) \phi(x_1^M) + \left(\frac{\mu_1^H - x_1^M}{x_1^H - x_1^M} \right) \phi(x_1^H) \right),
$$

where

$$
c_1^L \equiv \int_{x_1^L}^{x_1^M} dF(x_1) > 0, \qquad \text{and} \qquad c_1^H \equiv 1 - c_1^L > 0, \tag{9}
$$

$$
\mu_1^L \equiv \frac{\int_{x_1^L}^{x_1^M} x_1 dF(x_1)}{c_1^L}, \qquad \text{and} \qquad \mu_1^H \equiv \frac{\int_{x_1^M}^{x_1^H} x_1 dF(x_1)}{c_1^H}. \tag{10}
$$

Clearly, this method can be used to refine the bound shown in Theorem 2.1, as long as each partition is rectangular. By applying this method to increasing partitions, we again may have many function evaluations. The bound in Theorem 2.1 forces the individual random variables to become perfectly correlated. Simply reapplying this bound over various rectangular partitions of the probability space changes that correlation. However, the following shows that we may refine the $\mathbf{HL_0}$ bound and keep the perfect correlation. This helps primarily with keeping the number of function evaluations small.

First we state without proof a rather obvious extension of the bound $\mathbf{M_1}$ given above.

Lemma 2.1 *If $p^L x^L + p^M x^M + p^H x^H = E[x]$, $p^L + p^M + p^H = 1$, and $p^L, p^M, p^H \geq 0$, then*

$$\bar{\phi} \leq \mathbf{M_1} \leq p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) \leq \mathbf{M_0}$$

for $0 \leq p^M \leq \left(c_1^L \left(\frac{\mu_1^L - x^L}{x^M - x^L}\right) + c_1^H \left(\frac{\mu_1^H - x^M}{x^H - x^M}\right)\right)$, where c_1^L, c_1^H, μ_1^L and μ_1^H are as defined in equations (9) and (10).

Using this lemma, the bound for non-increasing convex functions with convex marginal return functions can be refined. Let X be an n-dimensional random vector, where each X_i is independently distributed with bounded support and finite mean. Further, let x_i^L, x_i^H be the realizations of X_i as defined in equations (1) and (2), so that $x_i^L < x_i^H$ for all i. Further, let $x_i^L < x_i^M < x_i^H$ for all i. Let $p_i^L \equiv prob\{X_i = x_i^L\}, p_i^M \equiv prob\{X_i = x_i^M\}$ and $p_i^H \equiv prob\{X_i = x_i^H\}$ for all i be such that

$$p_i^L x_i^L + p_i^M x_i^M + p_i^H x_i^H = E[X_i], \quad p_i^L + p_i^M + p_i^H = 1, \quad p_i^L, p_i^M, p_i^H \geq 0.$$

Let $f(X)$ be a non-increasing convex function of X with convex marginal return functions with respect to any pair $(X_i, X_j), i = 1, \ldots, n,\ j = 1 \ldots, n$.

Theorem 2.2 *Let X and $f(X)$ be as defined above. Let $p^L \equiv \max\{\frac{p_i^L}{p_i^L + p_i^H} : i = 1, \ldots, n\}$ and $p^H = 1 - p^L$. Further, let $p^M = \min\{c_i^L \left(\frac{\mu_i^L - x_i^L}{x_i^M - x_i^L}\right) + c_i^H \left(\frac{x_i^H - \mu_i^H}{x_i^H - x_i^M}\right) : i = 1, 2, \ldots, n\}$, where c_i^L, c_i^H, μ_i^L, and μ_i^H are as defined in equations (9) and (10). Then*

$$E[f(X)] \leq p^L f(x_1^L, \ldots, x_n^L) + p^M f(x_1^M, \ldots, x_n^M) + \bar{p}^H f(x_1^H, \ldots, x_n^H) \equiv \mathbf{HL_1},$$

where

$$\bar{p}^L = p^L * (1 - p^M), \quad \textit{and} \quad p^H = p^H * (1 - p^M).$$

Proof: By induction. Let $n = 1$. The result holds by the Huang, Ziemba, and Ben-Tal extension of the Edmundson-Madansky bound.

Assume the result is true for $n = t \geq 1$, and consider the case where $n = t+1$. Then, by the extension of the Edmundson-Madansky bound,

$$E[f(X)] \leq p_1^L E[f(x_1^L, X_2, \ldots, X_n)] + p_1^M E[f(x_1^M, X_2, \ldots, X_n)] + p_1^H E[f(x_1^H, X_2, \ldots, X_n)].$$

However, for a particular value of $X_1 = x_1$, $f(x_1, X_2, \ldots, X_n)$ is a function of t variables, and so, by the induction hypothesis,

$$E[f(x_1, X_2, \ldots, X_n)] \leq \tilde{p}_L f(x_1, x_2^L, \ldots, x_n^L) + \tilde{p}^M f(x_1, x_2^M, \ldots, x_n^M) + \tilde{p}_H f(x_1, x_2^H, \ldots, x_n^H),$$

where $\tilde{p}^M = \min\{c_i^L \left(\frac{\mu_i^L - x_i^L}{x_i^M - x_i^L}\right) + c_i^H \left(\frac{x_i^H - \mu_i^H}{x_i^H - x_i^M}\right) : i = 2, 3, \ldots, n\}$, $\tilde{p}^L = \max\{\frac{p_i^L}{p_i^L + p_i^H} : i = 2, \ldots, n\} * (1 - \tilde{p}^M)$ and $\tilde{p}^H = (1 - \max\{\frac{p_i^L}{p_i^L + p_i^H} : i = 2, \ldots, n\} * (1 - \tilde{p}^M)$. Thus, from above,

$$\begin{aligned} E[f(X)] \leq\; & p_1^L \tilde{p}^L f(x_1^L, x_2^L, \ldots, x_n^L) + p_1^L \tilde{p}^M f(x_1^L, x_2^M, \ldots, x_n^M) + p_1^L \tilde{p}^H f(x_1^L, x_2^H, \ldots, x_n^H) \\ & + p_1^M \tilde{p}^L f(x_1^M, x_2^L, \ldots, x_n^L) + p_1^M \tilde{p}^M f(x_1^M, x_2^M, \ldots, x_n^M) + p_1^M \tilde{p}^H f(x_1^M, x_2^H, \ldots, x_n^H) \\ & + p_1^H \tilde{p}^L f(x_1^H, x_2^L, \ldots, x_n^L) + p_1^H \tilde{p}^M f(x_1^H, x_2^M, \ldots, x_n^M) + p_1^H \tilde{p}^H f(x_1^H, x_2^H, \ldots, x_n^H). \end{aligned}$$

Note that now the set of random variables $(X_2, \ldots, X_n)$ are perfectly correlated and, therefore, can again be replaced by a single random variable. To simplify the notation, let $y^k \equiv (x_2^k, \ldots, x_n^k)$, for $k = L, M, H$.

Now, the main idea behind this part of the proof is to break the domain of X into various rectangular regions, eliminate points within each of those regions and continue this process until the only points remaining are the desired points. Assume without loss of generality that

$$p_1^L/(p_1^L + p_1^H) = max\{p_i^L/(p_i^L + p_i^H) : i = 1, \ldots, n\}.$$

Further, let

$$p_1^k = (1 - p^M) * p_1^k/(p_1^L + p_1^H) \quad (k = L, H), \quad \text{and} \quad \check{p}^k = \tilde{p}^k * (1 - p^M)/(1 - \tilde{p}^M) \quad (k = L, H).$$

Note that by construction, $\breve{p}^L = max\{p_i^L/(p_i^L + p_i^H) : i = 2, \ldots, n\} * (1 - p^M)$. Then by Lemma 2.1,

$$\begin{aligned}
E[f(X)] &\leq p_1^L \tilde{p}^L f(x_1^L, y^L) + p_1^L \tilde{p}^M f(x_1^L, y^M) + p_1^L \tilde{p}^H f(x_1^L, y^H) + \\
&\quad p_1^M \tilde{p}^L f(x_1^M, y^L) + p_1^M \tilde{p}^M f(x_1^M, y^M) + p_1^M \tilde{p}^H f(x_1^M, y^H) + \\
&\quad p_1^H \tilde{p}^L f(x_1^H, y^L) + p_1^H \tilde{p}^M f(x_1^H, y^M) + p_1^H \tilde{p}^H f(x_1^H, y^H) \\
&\leq \hat{p}_1^L \breve{p}^L f(x_1^L, y^L) + \hat{p}_1^L p^M f(x_1^L, y^M) + \hat{p}_1^L \breve{p}^H f(x_1^L, y^H) + \\
&\quad p^M \breve{p}^L f(x_1^M, y^L) + p^M p^M f(x_1^M, y^M) + p^M \breve{p}^H f(x_1^M, y^H) + \\
&\quad \hat{p}_1^H \breve{p}^L f(x_1^H, y^L) + \hat{p}_1^H p^M f(x_1^H, y^M) + \hat{p}_1^H \breve{p}^H f(x_1^H, y^H).
\end{aligned} \tag{11}$$

Now, consider the region of the domain of X in which $x_1^L \leq x_1 \leq x_1^M$ and $y^L \leq y \leq y^M$. By Theorem 2.1, since $\hat{p}_1^L \geq \breve{p}^L$, we obtain:

$$\begin{aligned}
&\hat{p}_1^L \breve{p}^L f(x_1^L, y^L) + \hat{p}_1^L p^M f(x_1^L, y^M) + p^M \breve{p}^L f(x_1^M, y^L) + p^M p^M f(x_1^M, y^M) \\
&\qquad \leq \hat{p}_1^L (\breve{p}^L + p^M) f(x_1^L, y^L) + p^M (p^M + \breve{p}^L) f(x_1^M, y^M).
\end{aligned} \tag{12}$$

Next consider the region of the domain of X in which $x_1^M \leq x_1 \leq x_1^H$ and $y^M \leq y \leq y^H$. Again by Theorem 2.1, since $\breve{p}^H \geq \hat{p}_1^H$, we obtain:

$$\begin{aligned}
&p^M p^M f(x_1^M, y^M) + p^M \breve{p}^H f(x_1^M, y^H) + \hat{p}_1^H p^M f(x_1^H, y^M) + \hat{p}_1^H \breve{p}^H f(x_1^H, y^H) \\
&\qquad \leq p^M (p^M + \breve{p}^H) f(x_1^M, y^M) + \hat{p}_1^H (\breve{p}^H + p^M) f(x_1^H, y^H).
\end{aligned} \tag{13}$$

Finally, looking over the entire domain of X and only considering the corner points, since $\hat{p}_1^L \geq \breve{p}^L$, we can use Theorem 2.1 to obtain:

$$\begin{aligned}
&\hat{p}_1^L \breve{p}^L f(x_1^L, y^L) + \hat{p}_1^L \breve{p}^H f(x_1^L, y^H) + \hat{p}_1^H \breve{p}^L f(x_1^H, y^L) + \hat{p}_1^H \breve{p}^H f(x_1^H, y^H) \\
&\qquad \leq \hat{p}_1^L (\breve{p}^L + \breve{p}^H) f(x_1^L, y^L) + \hat{p}_1^H (\breve{p}^H + \breve{p}^L) f(x_1^H, y^H).
\end{aligned} \tag{14}$$

Combining equations (11)-(14) gives the desired result. $\square$

Note that the only restrictions on the choice of x_i^M are that $x_i^L < x_i^M < x_i^H$ and that x_i^M be such that

$0 < c_i^L < 1$. Therefore, the middle points can be chosen so that the p_i^M values are all relatively close. This aids the effectiveness of the refined bound.

3 Example Functions

In this section, we consider two examples to illustrate the advantage of this new bound. The first example represents a common logarithmic utility function with two commodities. The second is taken from a vehicle allocation problem.

Example #1: $\phi(x_1, x_2) = -\ln(x_1^2 + 8x_2)$, $\quad 1 \le x_1 \le 25,\ 0 \le x_2 \le 20$.

Since all second derivatives of $\phi(X_1, X_2)$ are nonnegative over the range of all possible values of X_1 and X_2, $\phi(X_1, X_2)$ is a convex function of X_1 and X_2 over that range (See, e. g. , Rockafellar [9]). Since the argument of $\ln(\cdot)$ always increases as X_1 and/or X_2 increases, and since $-\ln(x)$ is a strictly decreasing function of $x \ge 0$, $\phi(X_1, X_2)$ is also a non-increasing function of X_1 and X_2. Finally, since

$$\frac{\partial \phi(x_1, x_2)}{\partial x_2} = \frac{-8}{x_1^2 + 4x_2}$$

is a non-decreasing function of X_1, $\phi(X_1, X_2)$ has convex marginal return functions with respect to X_1 and X_2. Thus, the results from Section 2 can be used to generate upper bounds on $E[\phi(X_1, X_2)] \equiv \bar{\phi}$.

Let the random variables X_1 and X_2 have the following probability density functions.

$$\begin{array}{llll}
X_1 : & f_1(x_1) = .0567 - .0011574 * x_1 & 1 \le x_1 \le 25 & \\
& \Rightarrow \quad E[X_1] = 9.4967 & \Rightarrow & \text{E-M weights: } (.6460, .3540); \\
X_2 : & f_2(x_2) = .11565 * \exp^{\frac{-1}{10} x_2} & 0 \le x_2 \le 20 & \\
& \Rightarrow \quad E[X_2] = 6.870 & \Rightarrow & \text{E-M weights: } (.6565, .3435).
\end{array}$$

Theorem 2.1 can be applied to give the following upper bound on $\bar{\phi}$,

$$\begin{aligned}
\mathbf{HL_0} &= .6565 * \phi(1,0) + .3435 * \phi(25,20) \\
&= (.6565) * (-\ln(1)) + (.3435) * (-\ln(785)) \\
&= -2.28966 \ \geq \ \bar{\phi}.
\end{aligned}$$

compared to the Edmundson-Madansky bound, which is

$$\begin{aligned}
\mathbf{M_0} &= .6460 * .6565 * \phi(1,0) + .6460 * .3435 * \phi(1,20) \\
&\quad + .3540 * .6565 * \phi(25,0) + .3540 * .3435 * \phi(25,20) \\
&= -3.43439 \ \geq \ \bar{\phi}.
\end{aligned}$$

Then by Jensen's inequality, a lower bound can be established.

$$\bar{\phi} \geq \phi(9.4967, 6.870) = -\ln(145.1467) = -4.977439$$

Finally, the upper bound can be tightened using Theorem 2.2. Let $x_1^M = E[X_1]$, and $x_2^M = E[X_2]$. Then $p^M = .51365$. The improved upper bound is:

$$\begin{aligned}
\mathbf{HL_1} &= .6565 * (1 - .51365) * \phi(1,0) + .51365 * \phi(9.4967, 6.870) + .3435 * (1 - .51365) * \phi(25,20) \\
&= -3.67024 \ \geq \ \bar{\phi}.
\end{aligned}$$

Note that this bound improves on the Edmundson-Madansky bound and requires fewer function evaluations.

Example #2: Consider the network flow problem shown in figure 1.

To the left of Nodes 1 and 2 is the available supply for those nodes, and to the right of Nodes 3 through 7 is the demand requirements for each of those nodes. The link number, the link's upper capacity, and the cost per unit flow is listed along each link. Each link is assumed to have lower capacity of zero. Note that Links 1,2,3, and 4 have random variables ξ_1, ξ_2, ξ_3, and ξ_4 respectively for their upper capacity. For a particular realization of $\xi = (\xi_1, \xi_2, \xi_3, \xi_4)$, let $f(\xi)$ denote the value of minimizing costs in this network. Letting x_i denote the flow along Link i, then $f(\xi)$ can be written as follows:

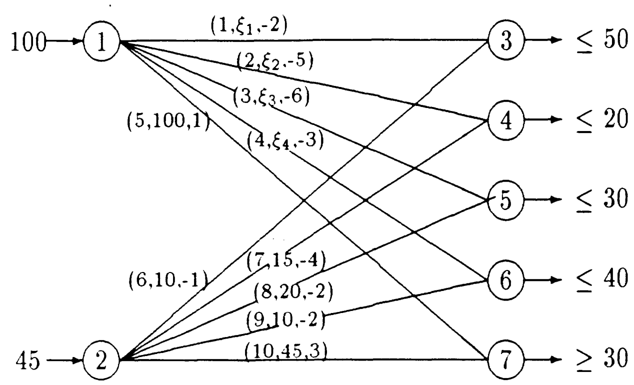

Figure 1: (Link number, upper capacity, cost per unit flow)

$$
\begin{array}{llllllllllllll}
f(\xi) = \min & -2x_1 & -5x_2 & -6x_3 & -3x_4 & +x_5 & -x_6 & -4x_7 & -2x_8 & -2x_9 & +3x_{10} & & \\
\text{s.t.} & x_1 & +x_2 & +x_3 & +x_4 & +x_5 & & & & & & = & 100 \\
 & & & & & & x_6 & +x_7 & +x_8 & +x_9 & +x_{10} & = & 45 \\
 & x_1 & & & & & +x_6 & & & & & \leq & 50 \\
 & & x_2 & & & & & +x_7 & & & & \leq & 20 \\
 & & & x_3 & & & & & +x_8 & & & \leq & 30 \\
 & & & & x_4 & & & & & +x_9 & & \leq & 40 \\
 & & & & & x_5 & & & & & +x_{10} & \geq & 30 \\
 & x_1 & & & & & & & & & & \leq & \xi_1 \\
 & & x_2 & & & & & & & & & \leq & \xi_2 \\
 & & & x_3 & & & & & & & & \leq & \xi_3 \\
 & & & & x_4 & & & & & & & \leq & \xi_4 \\
 & & & & & & x_6 & & & & & \leq & 10 \\
 & & & & & & & x_7 & & & & \leq & 15 \\
 & & & & & & & & x_8 & & & \leq & 20 \\
 & & & & & & & & & x_9 & & \leq & 10 \\
 & & & & & & x & & & & & \geq & 0.
\end{array}
$$

Suppose ξ_1 and ξ_3 have the following distributions:

x	11	12	13	14	15	16	17	18	19	20
$Pr(x)$	.02	.04	.05	.05	.06	.06	.06	.07	.07	.08

x	21	22	23	24	25	26	27	28	29	30
$Pr(x)$	.07	.07	.06	.05	.05	.04	.04	.03	.02	.01

And ξ_2 and ξ_4 have the following distributions:

x	21	22	23	24	25	26	27	28	29	30
$Pr(x)$	.01	.03	.05	.06	.07	.08	.09	.09	.10	.10

x	31	32	33	34	35	36	37	38	39	40
$Pr(x)$	.11	.08	.05	.02	.01	.01	.01	.01	.01	.01

This gives the following:

$$\Rightarrow \quad E[\xi_1] = E[\xi_3] = 19.8 \quad \Rightarrow \quad \text{E-M weights: } (.5368, .4632);$$

$$\Rightarrow \quad E[\xi_2] = E[\xi_4] = 28.69 \quad \Rightarrow \quad \text{E-M weights: } (.5953, .4147).$$

Solving each of the 20^4 possible versions of this linear program could be very difficult, making the value of effective bounds quite useful. This mathematical program, viewed as a function of the random upper capacities ξ_1, ξ_2, ξ_3, and ξ_4 is a convex function (see e.g., Birge and Louveaux [1]), a non-increasing function since any solution feasible for $\xi^{\mathbf{1}} = (\xi_1^1, \xi_2^1, \xi_3^1, \xi_4^1)$ is also feasible for any $\xi^{\mathbf{2}} = (\xi_1^2, \xi_2^2, \xi_3^2, \xi_4^2)$ such that $\xi_i^2 \geq \xi_i^1$ $(i = 1, \ldots, 4)$, and has convex marginal return functions with respect to each of these components by [4].

Briefly, the explanation for convex marginal return functions in this case is as follows. Suppose that the network flow problem is solved for particular values of ξ_1, ξ_2, ξ_3, and ξ_4. Suppose then that the value of ξ_1 is increased by one, while all other capacities remain unchanged. Suppose that the new problem is solved by a network algorithm which recursively deletes negative cost cycles, such as the out-of-kilter method, and the solution to the previous problem is used as the starting solution. The only possible negative cost cycle would have to include an increase of flow on Link 1, since the previous solution was optimal and the only changed aspect of the problem is the upper capacity of Link 1. But that implies that the solution to the updated problem either decreases or leaves the flow along Link 2 unchanged. In either event, if ξ_2 is now increased by one, the effect of that increase is less than if ξ_2 had been increased before ξ_1.

Again, the results from Section 2 can be applied to bound the expected value of $f(\xi)$. The bound from Theorem 2.1 gives an upper bound on the expected value of $f(\xi)$ by solving two versions of this program, instead of the sixteen needed to obtain the E-M bound. Here,

$$\begin{aligned} \mathbf{HL_0} &= .5953 * f(11,21,11,21) + .4147 * f(30,40,30,40) \\ &= (.5953) * (-270.00) + (.4147) * (-365.00) \\ &= -312.0965 \geq E[f(\xi)]. \end{aligned}$$

The Edmundson-Madansky bound for this function is:

$$\mathbf{M_0} = -319.4815 \geq E[f(\xi)].$$

Jensen's Inequality gives an lower bound of -343.09.

Finally, Theorem 2.2 can be used to refine the upper bound. Let $\xi_1^M = \xi_3^M = 19.8$ and let $\xi_2^M = \xi_4^M = 32.5$. Then $p^M = .59049$. The improved upper bound is:

$$\begin{aligned} \mathbf{HL_1} &= .2470 * f(11,21,11,21) + .5851 * f(19.8,32.5,19.8,32.5) + .1679 * f(30,40,30,40) \\ &= -329.4819 \geq E[f(\xi)]. \end{aligned}$$

In this case, the result is again better than the Edmundson-Madansky bound $\mathbf{M_0}$ with only three instead of sixteen function evaluations.

4 Conclusion

The difficulty of finding an upper bound for a convex function of a mulitivariate random variable, where the number of random components is greater than three, makes any special properties of certain classes of convex functions useful in analysis. In this paper, a bound is established for non-increasing convex functions which only requires two function evaluations whenever marginal return functions are convex with respect to any pair of the random components. Further, this bound can be refined in such a way that allows the number of function evaluations needed to grow linearly with the number of refinements performed. The bound has already been effective in establishing upper bounds for certain stochastic network flow problems. In practice, this bound has been found to be consistently as good or better than the Edmundson-Madansky bound after only one refinement of the bound, regardless of the number of random components.

References

[1] J. R. Birge and F. Louveaux, (1995) *Stochastic Programming*, unpublished manuscript.

[2] J. R. Birge and S. W. Wallace, "A Separable Piecewise Linear Upper Bound for Stochastic Linear Programs", *SIAM Journal on Control and Optimization* **26** 725 - 739 (1988).

[3] J. R. Birge and R. J-B. Wets, "Sublinear Upper Bounds for Stochastic Programs with Recourse", *Mathematical Programming* **43** 131 - 149 (1989).

[4] C. J. Donohue and J. R. Birge, "An Upper Bound on the Network Recourse Function," Working Paper, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan (1995).

[5] K. Frauendorfer, "Solving Stochastic Linear Program Recourse Problems with Arbitrary Mulitvariate Distributions - The Dependent Case", *Mathematics of Operations Research* **13** 377 - 394 (1988).

[6] H. Gassmann and W. T. Ziemba, "A Tight Upper Bound for the Expectation of a Convex Function of Multivariate Random Variable", *Mathematical Programming Study* **27** 39 - 53 (1986).

[7] C. C. Huang, W. T. Ziemba and A. Ben-Tal, "Bounds on the Expectation of a Convex Function of a Random Variable: With Applications to Stochastic Programming", *Operations Research* **25** 315 - 325 (1977).

[8] A. Madansky, "Bounds on the Expectation of a Multivariate Random Variable", *Annals of Mathematical Statistics* **30** 743 - 746 (1959).

[9] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, Princeton, N.J, 1970.

[10] S. Wallace, "A Piecewise Linear Upper Bound on the Network Recourse Function", *Mathematical Programming* **38** 133 - 146 (1987).

Lemma 2.1 *If* $p^L x^L + p^M x^M + p^H x^H = E[x]$, $p^L + p^M + p^H = 1$, *and* $p^L, p^M, p^H \geq 0$, *then*

$$\bar{\phi} \leq \mathbf{M_1} \leq p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) \leq \mathbf{M_0}$$

for $0 \leq p^M \leq \left(c_1^L \left(\frac{\mu_1^L - x^L}{x^M - x^L} \right) + c_1^H \left(\frac{\mu_1^H - x^M}{x^H - x^M} \right) \right)$, *where* c_1^L, c_1^H, μ_1^L *and* μ_1^H *are as defined above.*

Proof: The first inequality comes from Huang, Ziemba, and Ben-Tal [7]. The second and third come from noting that at $p^M = \left(c_1^L \left(\frac{\mu_1^L - x^L}{x^M - x^L} \right) + c_1^H \left(\frac{x^H - x^M}{x^H - x^M} \right) \right)$, $p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) = \mathbf{M_1}$, and at $p_M = 0$, $p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) = \mathbf{M_0}$. Between those two values of p^M, $p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H)$ is increasing linearly. $\square$