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

Christopher J. Donohue
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

1This work has been partially supported by the National Science Foundation under Grant DDM-9215921 and the Great Lakes Center for Trucking and Transit Research
2Christopher 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
2 Development of the Upper Bound

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

\[
\begin{align*}
x_i^L &= \sup \{ x_i : F(x) \leq 0 \}, \\
x_i^H &= \inf \{ x_i : F(x) \leq 1 \}.
\end{align*}
\]  

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

\[
M_0 \equiv \sum_{i_1 \in \{L, H\}} \cdots \sum_{i_n \in \{L, H\}} (p_i^{i_1} \ast \cdots \ast p_n^{i_n}) \ast \sigma(z_1^{i_1}, \ldots, z_n^{i_n}) \geq \bar{\sigma},
\]  

where

\[
p_i^L \equiv \frac{x_i^H - x_i}{x_i^H - x_i^L} \quad \text{and} \quad p_i^H \equiv \frac{x_i - x_i^L}{x_i^H - x_i^L}.
\]

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 \( \sigma(X) \) 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 \( \bar{\sigma} \) 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 $(X_i, X_j), i = 1, \ldots, n, j = 1, \ldots, n$ if for all $1 \leq i \leq n, 1 \leq j \leq n, \lambda \geq 0, \delta \geq 0$,

$$f(x_1, \ldots, x_j + \delta, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)$$

$$\leq f(x_1, \ldots, x_i + \lambda, \ldots, x_j + \delta, \ldots, x_n) - f(x_1, \ldots, x_i + \lambda, \ldots, x_j, \ldots, x_n).$$

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 $p_i^L \equiv \text{prob}\{X_i = x^L_i\}$ and $p_i^H \equiv \text{prob}\{X_i = x^H_i\}$ for all $i$ be such that

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

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

$$E[f(X)] \leq p^L f(x_1^L, x_2^L, \ldots, x_n^L) + p^H f(x_1^H, x_2^H, \ldots, x_n^H) \equiv H_{L0}.$$  

(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,

\[
E[f(X)] \leq p^L_1 E[f(x^L_1, X_2, \ldots, X_n)] + p^H_1 E[f(x^H_1, X_2, \ldots, X_n)]
\]

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,

\[
E[f(x_1, X_2, \ldots, X_n)] \leq p^L f(x_1, x^L_2, \ldots, x^L_n) + p^H f(x_1, x^H_2, \ldots, x^H_n),
\]

where \( p^L \equiv \max\{ p^L_1, p^L_2, \ldots, p^L_n \} \) and \( p^H \equiv 1 - p^L \). Thus,

\[
E[f(X)] \leq p^L_1 p^L f(x^L_1, x^L_2, \ldots, x^L_n) + p^H_1 p^H f(x^L_1, x^H_2, \ldots, x^H_n) + p^H_1 p^L f(x^H_1, x^L_2, \ldots, x^L_n) + \]

\[
p^H_1 p^H f(x^H_1, x^H_2, \ldots, x^H_n). \tag{8}
\]

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 \( p^L_1 \geq p^L \). To simplify notation, let

\[
f(L, L) \equiv f(x^L_1, x^L_2, \ldots, x^L_n), \quad f(L, H) \equiv f(x^L_1, x^H_2, \ldots, x^H_n),
\]

\[
f(H, L) \equiv f(x^H_1, x^L_2, \ldots, x^L_n), \quad f(H, H) \equiv f(x^H_1, x^H_2, \ldots, x^H_n).
\]

Then, by construction.
\[ E[f(X)] \leq p_L f(L, L) + p_H f(H, H) + p_L f(L, H) + p_H f(H, L) \]
\[ + p_L p_H f(H, H), \]
\[ \leq p_L f(L, L) + p_H f(H, H) + p_L f(L, H) + p_H f(H, L) + f(H, H) - f(L, L) + p_H f(H, H). \]
\[ \leq (p_L + p_H) f(L, L) + (p_H - p_L) f(L, H) + p_H f(H, H), \]
\[ \leq (p_L + p_H) f(L, L) + p_H f(H, H), \]
\[ = p_L f(L, L) + p_H f(H, H). \]

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 \( M_1 \) on the expectation of \( \phi(X) \) can be found by applying the Edmundson-Madansky bound on each of the subintervals \([x^L_1, x^M_1], [x^M_1, x^H_1]\) as follows:

\[ M_1 \equiv c_L^f \left( \left( \frac{x^M_1 - x^L_1}{x^M_1 - x^L_1} \right) \phi(x^L_1) + \left( \frac{x^H_1 - x^L_1}{x^H_1 - x^M_1} \right) \phi(x^M_1) \right) + c_H^f \left( \left( \frac{x^H_1 - x^M_1}{x^H_1 - x^M_1} \right) \phi(x^M_1) + \left( \frac{x^H_1 - x^M_1}{x^H_1 - x^M_1} \right) \phi(x^H_1) \right). \]

where

\[ c_L^f \equiv \int_{x^L_1}^{x^M_1} dF(x_1) > 0, \quad \text{and} \quad c_H^f \equiv 1 - c_L^f > 0. \]

\[ \mu_L^2 \equiv \int_{x^L_1}^{x^M_1} x_1 dF(x_1) \]
\[ \mu_H^2 \equiv \int_{x^M_1}^{x^H_1} x_1 dF(x_1) \]
\[ \mid \frac{x^H_1 - x^M_1}{c_H^f}, \]
\[ \mid \frac{x^H_1 - x^M_1}{c_H^f}. \]

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 $\text{HL}_0$ 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 $M_1$ 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

\[
\hat{\phi} \leq M_1 \leq p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) \leq M_0
\]

for $0 \leq p^M \leq \left(c_i^L \left(\frac{u^L_i - z^L_i}{z^L_i - x_i^L}\right) + c_i^H \left(\frac{u^H_i - z^H_i}{z^H_i - x_i^H}\right)\right)$, where $c_i^L, c_i^H, \mu_i^L$ and $\mu_i^H$ 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 $X$, 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 $p_i^L \equiv \text{prob}(X_i = x_i^L), p_i^M \equiv \text{prob}(X_i = x_i^M)$ and $p_i^H \equiv \text{prob}(X_i = x_i^H)$ for all $i$ be such that

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

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 $p^L \equiv \max\{\frac{\phi(x_i^L)}{p_i^L + p_i^H} : i = 1, \ldots, n\}$ and $p^H = 1 - p^L$. Further, let $p^M = \min\{c_i^L \left(\frac{x_i^L - x_i^M}{z_i^L - x_i^L}\right) + c_i^H \left(\frac{x_i^H - x_i^M}{z_i^H - x_i^H}\right) : i = 1, \ldots, n\}$. where $c_i^L, c_i^H, \mu_i^L$, and $\mu_i^H$ are as defined in equations (9) and (10). Then

\[
E[f(X)] \leq p^L f(x_1^L, \ldots, x_n^L) + p^M f(x_1^M, \ldots, x_n^M) + p^H f(x_1^H, \ldots, x_n^H) \equiv \text{HL}_1,
\]

where

\[
p^L = p^L \ast (1 - p^M), \quad \text{and} \quad p^H = p^H \ast (1 - p^M).
\]
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,

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

where \( \hat{p}_M = \min\{c_i^L \left( \frac{y_i^L - y_i^L}{x_i^L} \right) + c_i^H \left( \frac{y_i^H - y_i^M}{x_i^H - x_i^M} \right) : i = 2, 3, \ldots, n \}, \hat{p}_L = \max\{\frac{p_i^L}{p_i^L + p_i^H} : i = 2, \ldots, n \} \ast (1 - \hat{p}_M) \) and \( \hat{p}_H = (1 - \max\{\frac{p_i^L}{p_i^L + p_i^H} : i = 2, \ldots, n \} \ast (1 - \hat{p}_M). \) Thus, from above,

\[
E[f(X)] \leq p_1^L \hat{p}_L f(x_1^L, x_2^L, \ldots, x_n^L) + p_1^M \hat{p}_M f(x_1^M, x_2^M, \ldots, x_n^M) + p_1^H \hat{p}_H f(x_1^H, x_2^H, \ldots, x_n^H) + p_1^L \hat{p}_L f(x_1^L, x_2^L, \ldots, x_n^L) + p_1^M \hat{p}_M f(x_1^M, x_2^M, \ldots, x_n^M) + p_1^H \hat{p}_H f(x_1^H, x_2^H, \ldots, x_n^H) + p_1^L \hat{p}_L f(x_1^L, x_2^L, \ldots, x_n^L) + p_1^M \hat{p}_M f(x_1^M, x_2^M, \ldots, x_n^M) + p_1^H \hat{p}_H f(x_1^H, x_2^H, \ldots, x_n^H).
\]

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

\[
p_1^k = (1 - p^M) \ast p_i^k/(p_i^L + p_i^H) \quad (k = L, H) \quad \text{and} \quad \hat{p}_k = \hat{p}_L \ast (1 - p^M)/(1 - \hat{p}_M) \quad (k = L, H).
\]
Note that by construction, $\hat{p}_L = \max \{ p^i_L / (p^i_L + p^i_H) : i = 2, \ldots, n \} \ast (1 - p^M)$. Then by Lemma 2.1.

$$E[f(X)] \leq p^L_1 \hat{p}_L f(x^L_1, y^L) + p^M_1 \hat{p}_M f(x^M_1, y^M) + p^H_1 \hat{p}_H f(x^H_1, y^H) + p^M_1 \hat{p}_M f(x^1_M, y^L) + p^M_1 \hat{p}_M f(x^1_M, y^M) + p^M_1 \hat{p}_M f(x^1_M, y^H) + p^H_1 \hat{p}_H f(x^H_1, y^L) + p^H_1 \hat{p}_H f(x^H_1, y^M) + p^H_1 \hat{p}_H f(x^H_1, y^H) \leq \hat{p}_1 \hat{p}_L f(x^L_1, y^L) + \hat{p}_1 \hat{p}_M f(x^L_1, y^M) + \hat{p}_1 \hat{p}_H f(x^L_1, y^H) + p^M \hat{p}_L f(x^1_M, y^L) + p^M \hat{p}_M f(x^1_M, y^M) + p^M \hat{p}_M f(x^1_M, y^H) + p^H \hat{p}_L f(x^H_1, y^L) + p^H \hat{p}_M f(x^H_1, y^M) + p^H \hat{p}_H f(x^H_1, y^H).$$

(11)

Now, consider the region of the domain of $X$ in which $x^L_1 \leq x_1 \leq x^M_1$ and $y^L \leq y \leq y^M$. By Theorem 2.1, since $p^L_1 \geq \hat{p}_L$, we obtain:

$$p^L_1 \hat{p}_L f(x^L_1, y^L) + p^M_1 \hat{p}_M f(x^L_1, y^M) + p^M_1 \hat{p}_M f(x^L_1, y^L) + p^M_1 \hat{p}_M f(x^M_1, y^M) \leq p^L_1 (\hat{p}_L + p^M) f(x^L_1, y^L) + p^M (\hat{p}_M + \hat{p}_L) f(x^L_1, y^M).$$

(12)

Next consider the region of the domain of $X$ in which $x^M_1 \leq x_1 \leq x^H_1$ and $y^M \leq y \leq y^H$. Again by Theorem 2.1, since $\hat{p}_H \geq \hat{p}_H^M$, we obtain:

$$p^M_1 \hat{p}_M f(x^M_1, y^M) + p^M_1 \hat{p}_M f(x^M_1, y^H) + p^H_1 \hat{p}_H f(x^M_1, y^M) + p^H_1 \hat{p}_H f(x^M_1, y^M) \leq p^M_1 (\hat{p}_L + \hat{p}_H) f(x^M_1, y^M) + p^H_1 (\hat{p}_H + p^M) f(x^M_1, y^H).$$

(13)

Finally, looking over the entire domain of $X$ and only considering the corner points, since $p^L_1 \geq \hat{p}_L$, we can use Theorem 2.1 to obtain:

$$p^L_1 \hat{p}_L f(x^L_1, y^L) + p^L_1 \hat{p}_L f(x^L_1, y^H) + p^H_1 \hat{p}_H f(x^H_1, y^L) + p^H_1 \hat{p}_H f(x^H_1, y^H) \leq p^L_1 (\hat{p}_L + \hat{p}_H) f(x^L_1, y^L) + p^H_1 (\hat{p}_H + \hat{p}_L) f(x^H_1, y^H).$$

(14)

Combining equations (11)-(14) gives the desired result. □

Note that the only restrictions on the choice of $x^i_1$ are that $x^L_1 < x^M_1 < x^H_1$ and that $x^M_1$ be such that
0 < c_i^t < 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:** $\phi(x_1, x_2) = -\ln(x_1^2 + 8x_2), \quad 1 \leq x_1 \leq 25, \ 0 \leq x_2 \leq 20.$

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 $x \geq 0$, $\phi(X_1, X_2)$ 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 $E[\phi(X_1, X_2)] \equiv \hat{\phi}$.

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

$$X_1 : \quad f_1(x_1) = .0567 - .0011574 \cdot x_1 \quad 1 \leq x_1 \leq 25$$

$$\Rightarrow \ E[X_1] = 9.4967 \quad \Rightarrow \ \text{E-M weights: (.6460, .3540)};$$

$$X_2 : \quad f_2(x_2) = .11565 \cdot \exp \frac{x_2}{10} \cdot x_2 \quad 0 \leq x_2 \leq 20$$

$$\Rightarrow \ E[X_2] = 6.870 \quad \Rightarrow \ \text{E-M weights: (.6565, .3435)}.$$
Theorem 2.1 can be applied to give the following upper bound on $\bar{\phi}$.

\[
\text{HL}_0 = 0.6565 \cdot \phi(1, 0) + 0.3435 \cdot \phi(25, 20) \\
= (0.6565) \cdot (-\ln(1)) + (0.3435) \cdot (-\ln(785)) \\
= -2.28966 \geq \bar{\phi}.
\]

compared to the Edmundson-Madansky bound, which is

\[
\text{M}_0 = 0.6460 \cdot 0.6565 \cdot \phi(1, 0) + 0.6460 \cdot 0.3435 \cdot \phi(1, 20) \\
+ 0.3540 \cdot 0.6565 \cdot \phi(25, 0) + 0.3540 \cdot 0.3435 \cdot \phi(25, 20) \\
= -3.43439 \geq \bar{\phi}.
\]

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 $p^M = 0.51365$. The improved upper bound is:

\[
\text{HL}_1 = 0.6565 \cdot (1 - 0.51365) \cdot \phi(1, 0) + 0.51365 \cdot \phi(9.4967, 6.870) + 0.3435 \cdot (1 - 0.51365) \cdot \phi(25, 20) \\
= -3.67024 \geq \bar{\phi}.
\]

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 $\xi_1, \xi_2, \xi_3, \xi_4$ 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)

\[
f(\xi) = \text{min} \quad -2x_1 \quad -5x_2 \quad -6x_3 \quad -3x_4 \quad +x_5 \quad -x_6 \quad -4x_7 \quad -2x_8 \quad -2x_9 \quad +3x_{10}
\]

s.t. \quad x_1 + x_2 + x_3 + x_4 + x_5 = 100

x_6 + x_7 + x_8 + x_9 + x_{10} = 45

x_1 \leq 50

x_2 \leq 20

x_3 \leq 30

x_4 \leq 40

x_5 \geq 30

\[x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10} \leq 0.\]
Suppose $\xi_1$ and $\xi_3$ have the following distributions:

<table>
<thead>
<tr>
<th>x</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr(x)</td>
<td>.02</td>
<td>.04</td>
<td>.05</td>
<td>.05</td>
<td>.06</td>
<td>.06</td>
<td>.06</td>
<td>.07</td>
<td>.07</td>
<td>.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr(x)</td>
<td>.07</td>
<td>.07</td>
<td>.06</td>
<td>.05</td>
<td>.05</td>
<td>.04</td>
<td>.04</td>
<td>.03</td>
<td>.02</td>
<td>.01</td>
</tr>
</tbody>
</table>

And $\xi_2$ and $\xi_4$ have the following distributions:

<table>
<thead>
<tr>
<th>x</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr(x)</td>
<td>.01</td>
<td>.03</td>
<td>.05</td>
<td>.06</td>
<td>.06</td>
<td>.07</td>
<td>.08</td>
<td>.09</td>
<td>.09</td>
<td>.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr(x)</td>
<td>.11</td>
<td>.08</td>
<td>.05</td>
<td>.02</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
</tr>
</tbody>
</table>

This gives the following:

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

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

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 $\xi_1, \xi_2, \xi_3,$ and $\xi_4$ is a convex function (see e.g., Birge and Louveaux [1]), a non-increasing function since any solution feasible for $\xi^1 = (\xi^1_1, \xi^1_2, \xi^1_3, \xi^1_4)$ is also feasible for any $\xi^2 = (\xi^2_1, \xi^2_2, \xi^2_3, \xi^2_4)$ such that $\xi^2_i \geq \xi^1_i (i = 1, \ldots, 4)$, 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 $\xi_1, \xi_2, \xi_3,$ and $\xi_4$. Suppose then that the value of $\xi_1$ 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 $\xi_2$ is now increased by one, the effect of that increase is less than if $\xi_2$ had been increased before $\xi_1$.

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,

$$HL_0 = 0.5953 \cdot f(11, 21, 11, 21) + 0.4147 \cdot f(30, 40, 30, 40)$$

$$= (0.5953) \cdot (-270.00) + (0.4147) \cdot (-365.00)$$

$$= -312.0965 \geq E[f(\xi)].$$

The Edmundson-Madansky bound for this function is:

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

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

Finally, Theorem 2.2 can be used to refine the upper bound. Let $\xi^M_1 = \xi^M_3 = 19.8$ and let $\xi^M_2 = \xi^M_4 = 32.5$. Then $p^M = 0.59049$. The improved upper bound is:

$$HL_1 = 0.2470 \cdot f(11, 21, 11, 21) + 0.5851 \cdot f(19.8, 32.5, 19.8, 32.5) + 0.1679 \cdot f(30, 40, 30, 40)$$

$$= -329.4819 \geq E[f(\xi)].$$

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

The difficulty of finding an upper bound for a convex function of a multivariate 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


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
\[
\tilde{\phi} \leq M_1 \leq p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) \leq M_0
\]
for $0 \leq p^M \leq \left( c^L \left( \frac{\mu^L - x^L}{x^L - x^H} \right) + c^H \left( \frac{\mu^H - x^M}{x^M - x^H} \right) \right)$, where $c^L, c^H, \mu^L$, and $\mu^H$ 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 $p^M = \left( c^L \left( \frac{\mu^L - x^L}{x^L - x^H} \right) + c^H \left( \frac{\mu^H - x^M}{x^M - x^H} \right) \right)$, $p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) = M_1$, and at $p_M = 0$, $p^L \phi(x^L) + p^M \phi(x^M) + p^H \phi(x^H) = M_0$. 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. \qed