NC STATE UNIVERSITY

Bounds on the Expectation of a Convex Function of a Random Variable: With Applications to Stochastic Programming

C. C. Huang, W. T. Ziemba, A. Ben-Tal
Operations Research, 1977

Yue Li
Ye Tian

OUTLINE

- Jensen’s lower bound and its generalization
- E-M upper bound and its generalization
- Convergence
- Application

  Distribution problem

  SP with resource problem

  Numerical integration

JENSEN'S LOWER BOUND

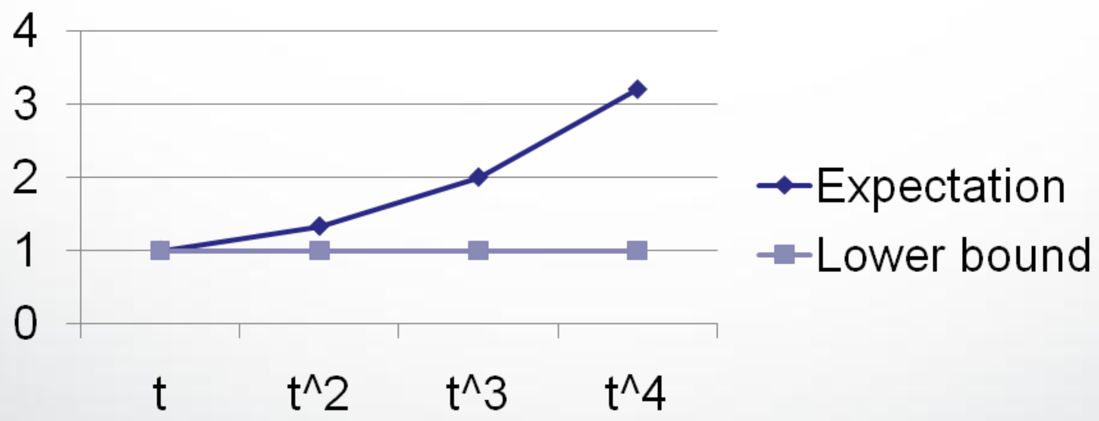
- Suppose $\phi(t):(a,b) \to R$ is an integrable convex function. Then Jensen's inequality:

$$\phi(\mu_0) \le \int_a^b \phi(t)\, dF(t) \equiv \bar{\phi}$$

provides a lower bound to the expected value of ϕ

EXAMPLE

t~U(0,2)	$\overline{\phi}$	$\phi(\mu_0)$	difference
$\phi(t)=t$	1	1	0
$\phi(t)=t^2$	1.33	1	0.33
$\phi(t)=t^3$	2	1	1
$\phi(t)=t^4$	3.2	1	2.2



GENERALIZATION OF JENSEN'S LOWER BOUND (1)

THEOREM 1. *Let $\phi(t)$ be an integrable convex function defined on the convex set $C \subset R^n$ and t be an integrable random vector such that $P[t \in C] = 1$. Then*

$Et \in C$ and

$$\phi(Et) \underset{(b)}{\leqq} P_A\phi(E(t \mid t \in A)) + P_{C/A}\phi(E(t \mid t \in C/A) \underset{(a)}{\leqq} E\phi(t),$$

for all measurable sets $A \subset C$, such that A and C/A are convex.

Proof. (a) $\phi(Et) \le E(\phi(t)) \Longrightarrow$

$$\phi(E(t|t \in A)) \le E(\phi(t)|t \in A) \qquad \times P_A$$
$$\phi(E(t|t \in C|A)) \le E(\phi(t)|t \in C|A) \qquad \times P_{C|A}$$

$$P_A\phi(E(t|t \in A)) + P_{C|A}\phi(E(t|t \in C|A)) \le E\phi(t).$$

(b) $Et = P_A E(t|t \in A) + P_{C|A}E(t|t \in C|A)$

ϕ convex $\Rightarrow \phi(Et) \le P_A\phi(E(t|t \in A)) + P_{C|A}\phi(E(t|t \in C|A)).$

Example

$$\phi(t) = t^3 \quad t \sim U(0,2)$$

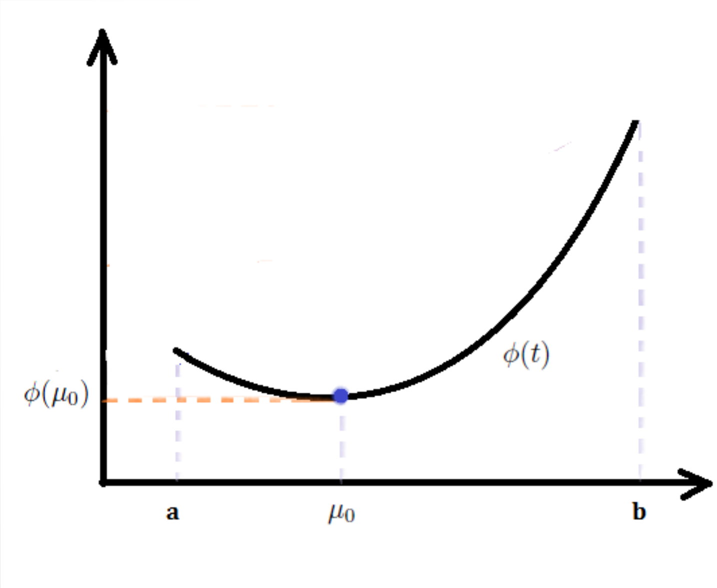
$$A = (0,1) \quad \phi(1) \leq 0.5\phi(0.5) + 0.5\phi(1.5) \leq \bar{\phi}$$

$$\Rightarrow \quad 1 \leq 1.75 \leq 2$$

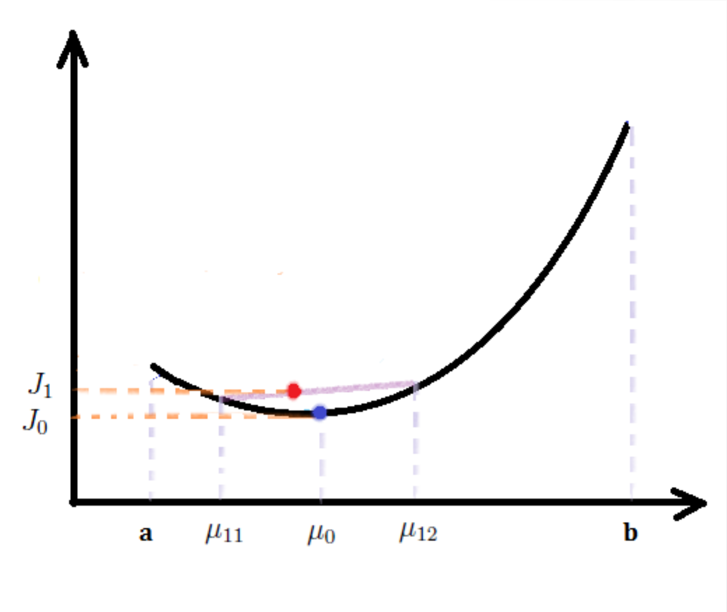
$$A = (0,0.5) \quad \phi(1) \leq 0.25\phi(0.25) + 0.75\phi(1.25) \leq \bar{\phi}$$

$$\Rightarrow \quad 1 \leq 1.47 \leq 2$$

GENERALIZATION OF JENSEN'S LOWER BOUND



$\Longrightarrow$



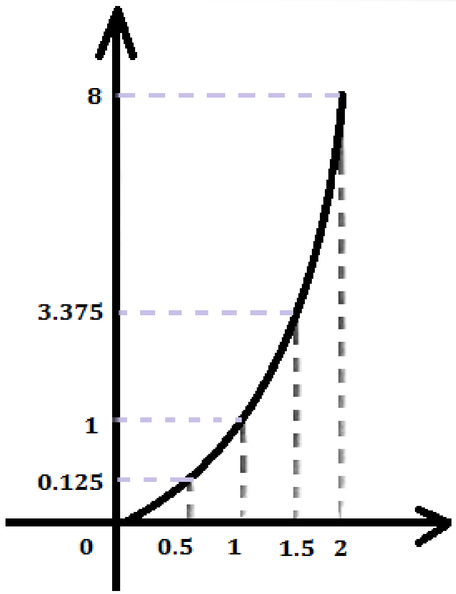
GENERALIZATION OF JENSEN'S LOWER BOUND (2)

THEOREM 2. (a) *Suppose* (a, b) *is subdivided at arbitrary points* $d_0, \cdots, d_m$, *where* $a = d_0 < \cdots < d_m = b$. *Let* $J^m \equiv \sum_{i=1}^{i=m} \alpha_i \phi(\beta_i)$, *denote the m-fold generalized Jensen bound, where* $\alpha_i \equiv \int_{d_{i-1}}^{d_i} dF(t) > 0$, $\beta_i \equiv \alpha_i^{-1} \int_{d_{i-1}}^{d_i} t\, dF(t)$, $i = 1, \cdots, m$. *Then, assuming that the partition corresponding to* $k + 1$ *is at least as fine as that corresponding to* k *for* $k = 1, \cdots, m - 1$, *we obtain* $J_0 \equiv J^1 \leqq \cdots \leqq J^m \leqq \bar{\phi}$.

(b) *Suppose* (a, b) *is subdivided* n *times on the basis of the partial means of the previous subintervals. Let* $J_k \equiv \sum_{i=1}^{i=2^k} c_{ki} \phi(\mu_{ki})$, $k = 0, 1, \cdots, n$, *denote the generalized Jensen bound obtained from the kth subdivision. Then* $J_0 \leqq J_1 \leqq \cdots \leqq J_n \leqq \bar{\phi}$, *where* $c_{01} = 1$, $\mu_{01} = \mu_0$ *and the* i^{th} *interval of the kth subdivision is denoted by* $[a_{ki}, b_{ki}]$, $c_{ki} \equiv \int_{a_{ki}}^{b_{ki}} dF(t) > 0$, $\mu_{ki} \equiv \int_{a_{ki}}^{b_{ki}} t\, dF(t) / c_{ki}$, where $c_{k+1,2i-1} \equiv \int_{a_{ki}}^{\mu_{ki}} dF(t) > 0$ and $c_{k+1,2i} \equiv \int_{\mu_{ki}}^{b_{ki}} dF(t) > 0$.

Example

$$\bar{\phi} = 2$$



$$\phi(t) = t^3,\ t \sim U(0,2)$$

(a)

$$\begin{array}{ll}
(0,2) & J_0 = J^1 = \phi(1) = 1 \\
(0,\tfrac{1}{2})(\tfrac{1}{2},2) & J^2 = \tfrac{1}{4}\phi(\tfrac{1}{4}) + \tfrac{3}{4}\phi(\tfrac{5}{4}) \approx 1.47 \\
(0,\tfrac{1}{2})(\tfrac{1}{2},1)(1,2) & J^3 = \tfrac{1}{4}\phi(\tfrac{1}{4}) + \tfrac{1}{4}\phi(\tfrac{3}{4}) + \tfrac{1}{2}\phi(\tfrac{3}{2}) \approx 1.80 \\
(0,\tfrac{1}{2})(\tfrac{1}{2},1)(1,\tfrac{3}{2})(\tfrac{1}{2},2) & J^4 = \tfrac{1}{4}\phi(\tfrac{1}{4}) + \tfrac{1}{4}\phi(\tfrac{3}{4}) + \tfrac{1}{4}\phi(\tfrac{5}{4}) + \tfrac{1}{4}\phi(\tfrac{7}{4}) \approx 1.94
\end{array}$$

(b)

$$\begin{array}{ll}
(0,2) & J_0 = J^1 = \phi(1) = 1 \\
(0,1)(1,2) & J_1 = \tfrac{1}{2}\phi(\tfrac{1}{2}) + \tfrac{1}{2}\phi(\tfrac{3}{2}) = 1.75 \\
(0,\tfrac{1}{2})(\tfrac{1}{2},1)(1,\tfrac{3}{2})(\tfrac{1}{2},2) & J_2 = J^4 \approx 1.94
\end{array}$$

THE EDMUNDSON-MADANSKY UPPER BOUND

For a convex function, we can use Jensen's inequality to get its lower bound. But how can we find its upper bound, if it is bounded in the whole domain?

THE EDMUNDSON-MADANSKY UPPER BOUND

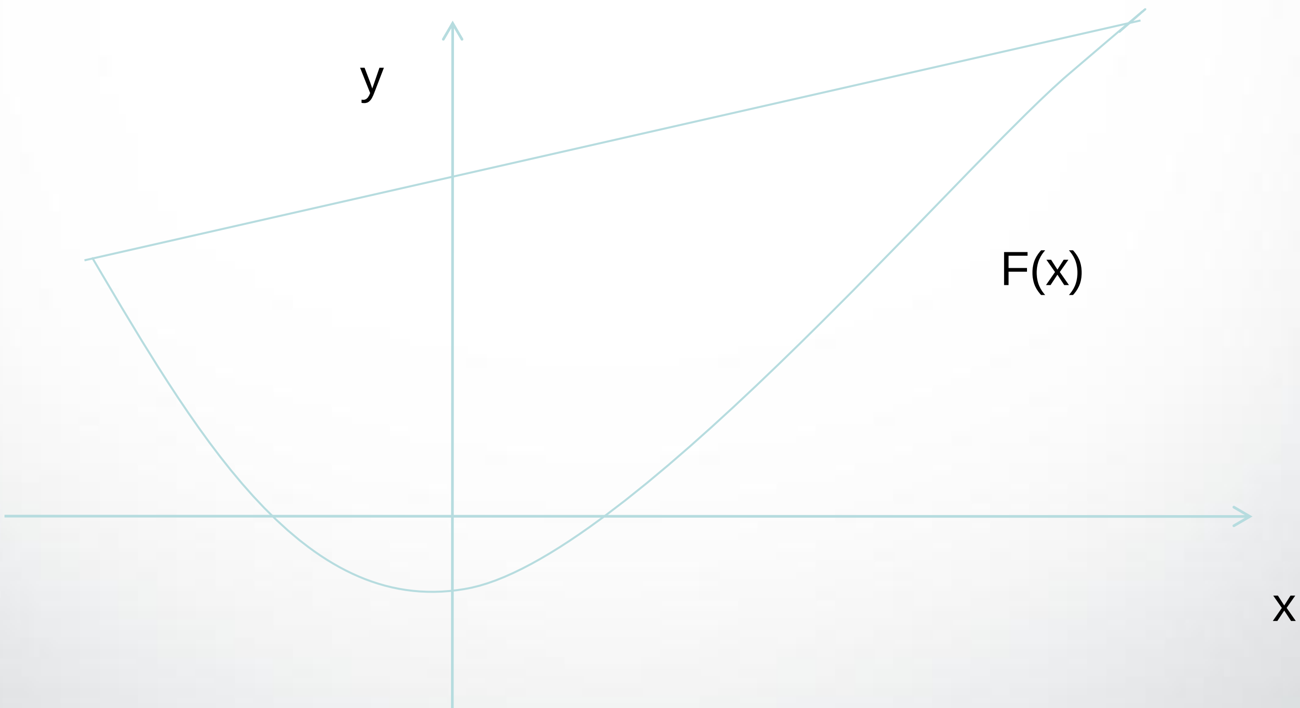
In nonlinear programming, conjugate dual play important role.

Let $f : S \subset E^n \to R$. Its conjugate transform is a function h with domain $\Omega = \{y \in E^n | Sup_{x \in S}[x^T y - f(x)] < +\infty\}$.
and
$h(y) = Sup_{x \in S}[x^T y - f(x)] = -inf_{x \in S}[f(x) - x^T y], \forall y \in \Omega.$

One property of conjugate dual problem is that :
For a concave function f, suppose its conjugate dual function is h. Then the conjugate dual of h is the convex hull function of f, which is a lower bound function of f.

THE EDMUNDSON-MADANSKY UPPER BOUND

For E-M method, they follow the same idea and find the upper bound for convex function.

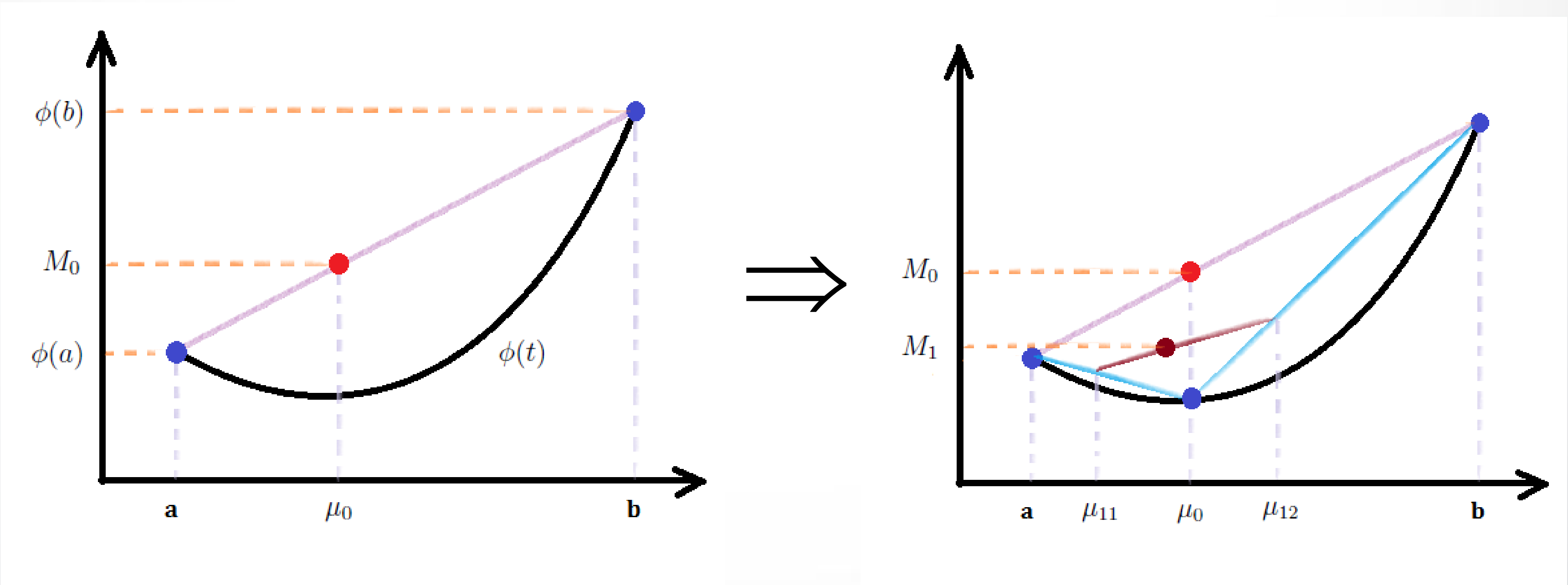


THE EDMUNDSON-MADANSKY UPPER BOUND

Let $t \in [a, b] \subset R$ have distribution function F and finite mean μ_0. Suppose ϕ is a bounded convex function of $t \in [a, b]$. The classic upper bound is the *Edmundson-Madansky inequality* $M_0 \equiv [(b - \mu_0)/(b - a)]\phi(a) + [(\mu_0 - a)/(b - a)]\phi(b) \geqq \bar{\phi}$.

Proof. Since ϕ is convex and $[a(b - t) + (t - a)b]/(b - a) \equiv t$, for all $t \in [a, b]$, we have $[(b - t)/(b - a)]\phi(a) + [(t - a)/(b - a)]\phi(b) \geqq \phi(t)$. Integrating gives the result.

GENERALIZATION OF THE EDMUNDSON-MADANSKY UPPER BOUND (1)



GENERALIZATION OF THE EDMUNDSON-MADANSKY UPPER BOUND (2)

THEOREM 3. (a) *Suppose* $[a, b]$ *is subdivided at arbitrary points* $d_0, \cdots, d_m$, *where* $a = d_0 < \cdots < d_m = b$. *Let* $M^m \equiv \sum_{i=0}^{i=m} \delta_i \phi(d_i)$ *denote the m-fold generalized E-M bound, where* $\delta_i \equiv \alpha_i[(\beta_i - d_{i-1})/(d_i - d_{i-1})] + \alpha_{i+1}[(d_{i+1} - \beta_{i+1})/(d_{i+1} - d_i)]$, $i = 0, 1, \cdots, m$, $\alpha_i \equiv \int_{d_{i-1}}^{d_i} dF(t) > 0$, $\beta_i \equiv \alpha_i^{-1} \int_{d_{i-1}}^{d_i} t\, dF(t)$, $i = 1, \cdots, m$, $\alpha_0 = \alpha_{m+1} \equiv 0$. *Then* $M_0 \equiv M^1 \geqq \cdots \geqq M^m \geqq \bar{\phi}$, *assuming that the partition corresponding to* $k + 1$ *is at least as fine as that corresponding to* k *for* $k = 1, \cdots, m - 1$.

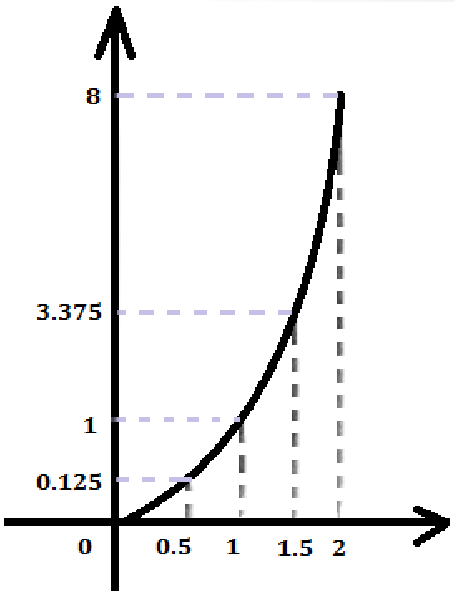
(b) *Suppose* $[a, b]$ *is subdivided* n *times on the basis of the partial means of the previous subintervals. Let*

$$M_k \equiv \sum_{i=1}^{i=2^k} c_{ki}[((b_{ki} - \mu_{ki})/(b_{ki} - a_{ki}))\phi(a_{ki}) + ((\mu_{ki} - a_{ki})/(b_{ki} - a_{ki}))\phi(b_{ki})], \qquad k = 0, 1, \cdots, n,$$

denote the generalized E-M bound obtained from the k^{th} *subdivision. Then* $M_0 \geqq M_1 \geqq \cdots \geqq M_n \geqq \bar{\phi}$, *where the* c_{ki}, μ_{ki}, a_{ki} *and* b_{ki} *are defined in Theorem 2.*

Example

$\bar{\phi} = 2$



$\phi(t) = t^3,\ t \sim U(0,2)$

(a)

$(0,2)$ $\quad M_0 = M^1 = \frac{2-1}{2-0}\phi(0) + \frac{1-0}{2-0}\phi(2) = 4$

$(0,\frac{1}{2})(\frac{1}{2},2)$ $\quad M^2 = \frac{1}{8}\phi(0) + \frac{1}{2}\phi(\frac{1}{2}) + \frac{3}{8}\phi(2) \approx 3.06$

$(0,\frac{1}{2})(\frac{1}{2},1)(1,\frac{3}{2})(\frac{1}{2},2)$ $\quad M^4 = \frac{1}{8}\phi(0) + \frac{1}{4}\phi(\frac{1}{2}) + \frac{1}{4}\phi(1) + \frac{1}{4}\phi(\frac{3}{2}) + \frac{1}{8}\phi(2) = 2.125$

(b)

$(0,2)$ $\quad M_0 = M^1 = 4$

$(0,1)(1,2)$ $\quad M_1 = \frac{1}{4}\phi(0) + \frac{1}{2}\phi(1) + \frac{1}{4}\phi(2) = 2.5$

$(0,\frac{1}{2})(\frac{1}{2},1)(1,\frac{3}{2})(\frac{1}{2},2)$ $\quad M_2 = M^4 = 2.125$

CONVERGENCE

THEOREM 4. Suppose $\phi(t)$ is a continuous convex function on $[a, b]$, and Et is finite. Then $\bar{\phi}$ is finite, $J^n \rightarrow \bar{\phi} \leftarrow M^n$, assuming that each subinterval becomes arbitrary small as $n \rightarrow \infty$, and $J_n \rightarrow \bar{\phi} \leftarrow M_n$.

Proof. Since ϕ is continuous and F is nondecreasing and hence of bounded variation, $\bar{\phi}$ exists (see Apostol [1, pp. 168, 211]). Let $U_k \equiv \sum_{i=1}^{i=2^k} c_{ki} M_{ki}$ and $L_k \equiv \sum_{i=1}^{i=2^k} c_{ki} m_{ki}$, where $M_{ki} \equiv \sup\{\phi(t) \mid t \in [a_{ki}, b_{ki}]\}$ and $m_{ki} \equiv \inf\{\phi(t) \mid t \in [a_{ki}, b_{ki}]\}$. By construction $L_n \leqq J^n \leqq \bar{\phi} \leqq M^n \leqq U_n$ and $L_n \leqq J_n \leqq \bar{\phi} \leqq M_n \leqq U_n$. Since F is nondecreasing on $[a, b]$, existence of $\bar{\phi}$ implies that $L_n \rightarrow \bar{\phi} \rightarrow U_n$ (see Apostol [1, p. 206]). Hence the result.

GENERALIZATION (1)

E-M bounds are presented for the case when

t is a bounded random variable.

If either $a = -\infty$, $b = +\infty$ or both, then, all the bounds exist (except for M_0 in the case $(-\infty, \infty)$) if one assumes that for:

$(-\infty, b]$ $\lim_{t \to -\infty} \phi(t)/t \equiv \alpha$, exists and is finite;

$[a, \infty)$ $\lim_{t \to \infty} \phi(t)/t \equiv \beta$, exists and is finite; and

$(-\infty, \infty)$ $\lim_{t \to \pm\infty} \phi(t)/t$ exists and is finite.

GENERALIZATION (2)

Suppose $t_1, \cdots, t_m$ are independent random variables, where each t_k is distributed on $[a_k, b_k]$ with distribution function F_k and finite mean μ_k. Let $\{d_i^k\}$, $i = 0, \cdots, n_k$ be a partition of $[a_k, b_k]$ such that $a_k = d_0^k < d_1^k < \cdots < d_{n_k-1}^k < d_{n_k}^k = b^k$,

$$\alpha_i^k \equiv \int_{d_{i-1}^k}^{d_i^k} dF_k(t) > 0,\ \beta_i^k \equiv (1/\alpha_i^k) \int_{d_{i-1}^k}^{d_i^k} t\, dF(t), \qquad i = 1, \cdots, n_k,$$

$$\delta_i^k \equiv \alpha_i^k [(\beta_i^k - d_{i-1}^k)/(d_i^k - d_{i-1}^k)] + \alpha_{i+1}^k [(d_{i+1}^k - \beta_{i+1}^k)/(d_{i+1}^k - d_i^k)], \qquad i = 0, \cdots, n_k,$$

where $\alpha_0^k = \alpha_{n_k+1}^k \equiv 0$. Suppose $\phi(t_1, \cdots, t_m)$ is bounded and convex on $\prod_{k=1}^{k=m} [a_k, b_k]$. A conditioning argument yields

$$\sum_{i_1=1}^{i_1=n_1} \cdots \sum_{i_m=1}^{i_m=n_m} \left(\prod_{k=1}^{k=m} \alpha_{i_k}^k\right) \phi(\beta_{i_1}^1, \cdots, \beta_{i_m}^m) \leq E\phi(t_1, \cdots, t_m) \leq \sum_{i_1=0}^{i_1=n_1} \cdots \sum_{i_m=0}^{i_m=n_m} \left(\prod_{k=1}^{k=m} \delta_{i_k}^k\right) \phi(d_{i_1}^1, \cdots, d_{i_m}^m).$$

GENERALIZATION (3)

In the dependent multivariate case lower bounds are available using the conditional form of Jensen's inequality. For example, suppose $\bigcup_{k=1}^{k=K} B_k = C$ is a partition of C, where each B_k is convex and $\neq \phi$ and $B_k \cap B_l = \phi$ if $k \neq l$. For $t \in B_k$ we have

$$p(B_k)\phi\left(\int_{B_k} (t/p(B_k))\, dF(t)\right) \leqq \int_{B_k} \phi(t)\, dF(t), \quad p(B_k) > 0.$$

Then the conditional form of Jensen's inequality may be written as

$$\sum_{k=1}^{k=K} p(B_k)\phi\left(\int_{B_k} (t/p(B_k))\, dF(t)\right) \leqq \sum_{k=1}^{k=K} \int_{B_k} \phi(t)\, dF(t).$$

One then develops a bound by determing K, the B_k, $p(B_k)$, $\phi\left(\int_{B_k} (t/p(B_k))\, dF(t)\right)$, and adding.

APPLICATIONS

- The Distribution Problem
- Stochastic Programs with Simple Recourse
- Numerical Integration

The Distribution Problem

In some researches, the distribution properties of the random mathematical program are of great interests. They study the following problem:

$$\phi(c) \equiv \max_x \{\Psi(c, x) | x \in K\}$$

Where ϕ and Ψ are scalar functions, c is a random vector having distribution function G, and x is a decision vector.

The Distribution Problem

We look at the simplest case which studies the mean of that problem.

$$\int \ldots \int \max_x \{\Psi(c,x) | x \in K\} dG(c).$$

The Distribution Problem

In general, in order to mathematically solve the problem above, a large number of mathematical programs are required.

$$\max_x \{\Psi(\overline{c}, x) | x \in K\}$$

The Distribution Problem

Alternatively, they seek approximations to that problem. Under the assumption that the components of the random vector c are independent, they can get the lower and upper bounds as following:

$$J_0 \le \int \Psi(c, \overline{x}(\overline{c}))dG(c) \le \overline{\phi} \le M_0$$

The Distribution Problem

We consider an example

$$\phi(c) \equiv \max_x \{c_1 x_1^2 + c_2^2 x_2^2 | 0 \le x_1 \le 10, 0 \le x_2 \le 5\}$$

c_1 and c_2 have independent uniform distribution on [-0.5, 0.5] and [0,1].

$$\phi(c) = \begin{cases} 100c_1 + 25c_2^2 & \text{if} \quad c_1 \geqq 0 \\ 25c_2^2 & \text{otherwise} \end{cases} \quad \text{and} \quad \bar{\phi} = 20.833.$$

Using the Jensen and E-M bounds, we can get that
6.25 < 8.33 < 20.833 < 37.5
If we subdivide the interval more finer, we can get better bounds 20.70 < 20.80 < 20.83 < 20.90 < 21.09

Stochastic Programs with Simple Recourse

The bounds can also be used to generate approximate solutions to certain stochastic mathematical programs that are difficult to solve.

$$Z \equiv \max_x \{\int \psi(c, x)\, dG(c) \mid x \in K\}$$

We can use Jensen and E-M bounds to bound the optimal value of this problem.

Stochastic Programs with Simple Recourse

For J1 and M1 bounds:

$$\max_x \left[c_{11}\psi(\mu_{11}, x) + c_{12}\psi(\mu_{12}, x) \mid x \in K\right] \leqq Z \leqq$$

$$\max_x \left[c_{11}\{((\mu_0 - \mu_{11})/(\mu_0 - a))\psi(a, x) + ((\mu_{11} - a)/(\mu_0 - a))\psi(\mu_0, x)\}\right.$$

$$+ c_{12}\{((b - \mu_{12})/(b - \mu_0))\psi(\mu_0, x) + ((\mu_{12} - \mu_0)/(b - \mu_0))\psi(b, x)\}$$

$$\left.\mid x \in K\right].$$

Numerical Integration

Because when the subintervals become infinitely small, the bounds converge to the optimal solutions of the problem. Thus these bounds can be used to calculate the expected value of a given convex function of a random variable in an arbitrary given tolerance.

Numerical Integration

Example:

$$\phi(t) = t^2$$

t has uniform distribution on [-10, 10]. Optimal solution is 33.3333.

	J	M
0	0	100
1	25	50
2	31.25	37.5
3	32.8125	34.3750
4	33.2031	33.5938

Numerical Integration

Compared with standard numerical schemes, these bounds an obvious advantage in cases where one wishes to perform calculations for several functions with a give distribution function.

Thank you for listening!