Outline

1. Stochastic Dominance
   - Definition
   - Characterization of Stochastic Dominance by Lorenz Functions

2. Dominance Constrained Optimization

3. Conjugate Function Method for Equal Probabilities

4. Quantile Cutting Plane Methods for General Probability Spaces
   - The Scaled Methods
   - The Unscaled Method

5. Numerical Experience
   - Performance Comparison
   - Implied Utility Functions
   - Implied Risk Measures
Stochastic Dominance

Distribution Functions

\[
F_1(X; \eta) = \int_{-\infty}^{\eta} P_X(dt) = P\{X \leq \eta\} \text{ for all } \eta \in \mathbb{R}
\]

\[
F_k(X; \eta) = \int_{-\infty}^{\eta} F_{k-1}(X; t) \, dt \text{ for all } \eta \in \mathbb{R}, \quad k = 2, 3, \ldots
\]

kth order Stochastic Dominance

\[
X \succeq_{(k)} Y \iff F_k(X, \eta) \leq F_k(Y, \eta) \text{ for all } \eta \in \mathbb{R}
\]
Second-Order Stochastic Dominance

\[ F_2(X, \eta) = \int_{-\infty}^{\eta} F_1(X, t) \, dt = \mathbb{E}(\eta - X)_+ \text{ for all } \eta \in \mathbb{R} \]
The Lorenz Function (O. Lorenz, 1905)

Quantile function

\[ F_{(-1)}(X; p) = \inf \{ \eta : F_1(X; \eta) \geq p \} \]

Absolute Lorenz function

\[ F_{(-2)}(X; p) = \int_0^p F_{(-1)}(X; t) \, dt \quad \text{for} \quad 0 < p \leq 1, \]

\[ F_{(-2)}(X; 0) = 0 \quad \text{and} \quad F_{(-2)}(X; p) = +\infty \quad \text{for} \quad p \notin [0, 1]. \]
Integrated Distribution Function and the Lorenz function

\[ F_2(X; \eta) \]

\[ \eta - \mu_x \]

\[ 0 \]

\[ F_{(-2)}(X; p) \]

Fenchel conjugate function

\[ F^*(p) = \sup_u \{ pu - F(u) \}. \]

Ogryczak - Ruszczyński (2002)

\[ F_{(-2)}(X; :) = [F_2(X; :)^* \quad \text{and} \quad F_2(X; :) = [F_{(-2)}(X; :)^*} \]
Characterization of Stochastic Dominance by Lorenz Functions

\[
X \succeq_{(2)} Y \iff F_{(-2)}(X; p) \geq F_{(-2)}(Y; p) \quad \text{for all } 0 \leq p \leq 1.
\]
Given $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ - benchmark random outcome

Inverse Stochastic Dominance Constraints

$$\max f(z)$$

$$\text{s.t. } F_{(-2)}(G(z); p) \geq F_{(-2)}(Y; p), \quad \forall \ p \in [\alpha, \beta],$$

$$z \in Z$$

$Z$ is a closed subset of a vector space $\mathcal{L}$, $[\alpha, \beta] \subset (0, 1)$

$G : \mathcal{L} \rightarrow \mathcal{L}_1(\Omega, \mathcal{F}, P)$ and $f : \mathcal{L} \rightarrow \mathbb{R}$ are continuous.

Equivalent formulation

$$\max_{z, \eta(\cdot)} f(z)$$

$$\text{s.t. } \eta(p) - \frac{1}{p} \mathbb{E} \max(0, \eta(p) - G(z)) \geq \frac{1}{p} F_{(-2)}(Y; p), \quad \forall \ p \in [\alpha, \beta]$$

$\eta(\cdot)$ - nondecreasing, $z \in Z$

with $\eta(p)$ a $p$-quantile of $G(z)$. 

Darinka Dentcheva and Andrzej Ruszczynski
Discrete Equiprobable Probability Space

Assumption

\[ \Omega = \{\omega_1, \ldots, \omega_N\} \text{ with probabilities } p_i = \mathbb{P}\{\omega_i\} = 1/N, \ i = 1, \ldots, N. \]

The Lorenz curves \( F_{(-2)}(G(z); \cdot) \) and \( F_{(-2)}(Y; \cdot) \) are convex and piecewise linear with break points at \( \pi_i = i/N \).

For \( \eta = (\eta_1, \ldots, \eta_N) \), and \( g_m(z) = [G(z)](\omega_m) \), set

\[
q_i = \frac{1}{\pi_i} F_{(-2)}(Y; \pi_i) = \frac{1}{i} \sum_{k=1}^{i} y[k].
\]

Equivalent formulation

\[
\max_{z, \eta} f(z)
\]

subject to

\[
\eta_i - \frac{1}{i} \sum_{m=1}^{N} \max(0, \eta_i - g_m(z)) \geq q_i \quad \forall \ i = 1, \ldots, N
\]

\[z \in Z\]
Conjugate Function Method

Step 0: Set $k = 0$.

Step 1: Solve the problem and obtain $(\hat{z}^k, \hat{\eta}^k)$:

$$\max_{z, \eta} f(z)$$

subject to

$$\eta_{ij} - \frac{1}{i_j} \sum_{m \in A_j} (\eta_{ij} - g_m(z)) \geq \varrho_{ij} \quad \forall j = 1, \ldots, k$$

$$z \in Z$$

Step 2: Let $\delta_k = \max_{1 \leq i \leq N} \left\{ \varrho_i - \hat{\eta}_{ik}^k + \frac{1}{i} \sum_{m=1}^{N} \max(0, \hat{\eta}_{ik}^k - g_m(\hat{z}^k)) \right\}$. If $\delta_k \leq 0$, stop; otherwise, continue.

Step 3: Let $i_k$ be the index at which the maximum in Step 2 is achieved, and let $A_k = \{ m : \hat{\eta}_{ik}^k > g_m(\hat{z}^k) \}$.

Step 4: Increase $k$ by one, and go to Step 1.
New Dominance Characterization for General Probability Spaces

Theorem

Suppose $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$. Then $X \succeq_{(2)} Y$ if and only if

$$\mathbb{E}[X|A] \geq \frac{1}{P(A)} F_{(-2)}(Y; P(A)) \quad \forall A \in \mathcal{F}: P(A) > 0.$$

Corollary

$X \succeq_{(2)} Y$ if and only if

$$\mathbb{E}[X|X \leq t] \geq \frac{1}{F(X; t)} F_{(-2)}(Y; F(X; t)) \quad \forall t \in \mathbb{R}, F(X; t) > 0.$$

Corollary

Suppose $X$ has a discrete distribution with $N$ realizations, and $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$. Then $X \succeq_{(2)} Y$ if and only if

$$\sum_{k=1}^{j} \pi_k x_k \geq F_{(-2)}(Y; \sum_{k=1}^{j} \pi_k), \quad j = 1, \ldots, N.$$
Quantile Cutting Plane Methods: The Scaled Method

Step 0: Set $k = 0$.

Step 1: Solve the problem to obtain $z^k$ and $X^k = G(z^k)$:

$$\max f(z)$$

s.t. $\mathbb{E}[G(z)|S^i] \geq \frac{1}{P(S^i)} F(-2)(Y; P(S^j)) \quad j = 1, \ldots, k$

$z \in Z$

Step 2: Consider the sets $A_t^k = \{X^k \leq t\}$ and let

$$\delta_k = \sup_t \left\{ \frac{1}{P(A_t^k)} F(-2)(Y; P(A_t^k)) - \mathbb{E}[X^k|A_t^k] : P(A_t^k) > 0 \right\}.$$

If $\delta_k \leq 0$, stop; otherwise, continue.

Step 3: Find $t_k$ such that $P(X^k \leq t_k) > 0$ and

$$\mathbb{E}[X^k|A_{t_k}^k] - \frac{1}{P(A_{t_k}^k)} F(-2)(Y; P(A_{t_k}^k)) \leq -\frac{\delta_k}{2}.$$

Step 4: Set $S^{k+1} = A_{t_k}^k$, increase $k$ by one, and go to Step 1.
Theorem. Assume that $Z \subset \mathbb{R}^n$ is compact, $f(\cdot)$ is continuous, and the operator $G(\cdot) : \mathbb{R}^n \to \mathcal{L}_\infty(\Omega, \mathcal{F}, P)$ is Lipschitz continuous on $Z$. If problem $(\mathcal{P}_{-2})$ has a nonempty feasible set, then either the scaled method stops at a solution of it, or every accumulation point of the sequence $\{z^k\}$ generated by the method is a solution of $(\mathcal{P}_{-2})$.

In finite probability space $\Omega$, the scaled method converges in finitely many iterations.

Modification

If the operator $G(\cdot) : \mathbb{R}^n \to \mathcal{L}_1(\Omega, \mathcal{F}, P)$ is Lipschitz continuous on $Z$, then we can modify the method to ensure convergence.
Assumption

The operator \( G(\cdot) : \mathbb{R}^n \rightarrow \mathcal{L}_1(\Omega, \mathcal{F}, P) \) is Lipschitz continuous on \( Z \).

We modify the method to ensure convergence by changing the following steps

**Step 2a:** Consider the sets \( A^k_t = \{ X^k \leq t \} \) and let

\[
\delta_k = \sup_t \left\{ \frac{1}{P(A^k_t)} F(-2)(Y; P(A^k_t)) - \mathbb{E}[X^k | A^k_t] : P(A^k_t) \geq \varepsilon_k \right\}.
\]

**Step 2b:** If \( \delta_k \leq 0 \), replace \( \varepsilon_k \) by \( \varepsilon_k / 2 \) and go to Step 2a; otherwise, continue.

**Step 4:** If \( \delta_k < \varepsilon_k \) then set \( \varepsilon_{k+1} = \min\{\delta_k, \varepsilon_k\} / 2 \); otherwise set \( \varepsilon_{k+1} = \varepsilon_k \). Set \( S^{k+1} = A^k_{t_k} \), increase \( k \) by one, and go to Step 1.

Observe that it is possible for the method to cycle between Steps 2a and 2b, without increasing the iteration index \( k \).
Quantile Cutting Plane Methods: The Unscaled Method

The method uses cutting planes in a different form.

**Step 2:** Consider the sets \( A_t^k = \{ X^k = G(z^k) \leq t \} \) and let

\[
\delta_k = \sup_{t} \left\{ F(-2)(Y; P(A_t^k)) - F(-2)(X^k; P(A_t^k)) : P(A_t^k) > 0 \right\}.
\]

If \( \delta_k \leq 0 \), stop; otherwise, continue.

**Step 3:** Find \( t_k \) such that \( P(X^k \leq t_k) > 0 \) as well as

\[
F(-2)(X^k; P(A_{t_k}^k)) - F(-2)(Y; P(A_{t_k}^k)) \leq -\frac{\delta_k}{2}.
\]

**Difference:** no normalization by \( P(A_t^k) \).

**Theorem.** Assume that \( Z \subset \mathbb{R}^n \) is compact, \( f(\cdot) \) is continuous, and the operator \( G(\cdot) : \mathbb{R}^n \rightarrow \mathcal{L}_1(\Omega, \mathcal{F}, P) \) is Lipschitz continuous on \( Z \). If problem \((\mathcal{P}_{-2})\) has a nonempty feasible set, then either the unscaled method stops at a solution of \((\mathcal{P}_{-2})\), or every accumulation point of the sequence \( \{z^k\} \) is a solution of \((\mathcal{P}_{-2})\).
Assets $j = 1, \ldots, n$ with random return rates $R_j$
Reference return rate $Y$ (e.g. index, existing portfolio, etc.)
Decision variables $z_j$, $j = 1, \ldots, n$, $Z$ -polyhedral set
Portfolio return rate $R(z) = \sum_{j=1}^{n} z_j R_j$

$$
\begin{align*}
\max & \quad f(z) \\
\text{s.t.} & \quad \sum_{j=1}^{n} z_j R_j \succeq Y \\
& \quad z \in Z
\end{align*}
$$

$f(x) = \mathbb{E}[R(x)]$ or $f(x) = \varrho[R(x)]:$ measure of risk.
Linear Programming Formulation

\[
\text{maximize } \sum_{i=1}^{N} p_i \sum_{k=1}^{n} r_{ik} z_k \\
\text{s.t. } \sum_{k=1}^{n} r_{ik} z_k + s_{ij} \geq y_j, \quad i = 1, \ldots, N, \quad j = 1, \ldots, N, \\
\sum_{i=1}^{N} p_i s_{ij} \leq F_2(Y; y_j), \quad j = 1, \ldots, N, \\
s \geq 0, \quad z \in \mathbb{Z}.
\]
### Table: Dimensions of the three formulations.

<table>
<thead>
<tr>
<th>Events</th>
<th>Direct LP</th>
<th>Conjugate Formulation</th>
<th>Quantile Formulations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 500$ and $Y$ is the return rate of the S&amp;P 500 index.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>Variables</td>
<td>Constraints</td>
<td>Variables</td>
</tr>
<tr>
<td>50</td>
<td>3000</td>
<td>2551</td>
<td>550</td>
</tr>
<tr>
<td>100</td>
<td>10500</td>
<td>10101</td>
<td>600</td>
</tr>
<tr>
<td>150</td>
<td>23000</td>
<td>22651</td>
<td>650</td>
</tr>
<tr>
<td>200</td>
<td>40500</td>
<td>40201</td>
<td>700</td>
</tr>
<tr>
<td>250</td>
<td>63000</td>
<td>62751</td>
<td>750</td>
</tr>
<tr>
<td>300</td>
<td>90500</td>
<td>90301</td>
<td>800</td>
</tr>
</tbody>
</table>

### Table: Performance on problems with equal probabilities of elementary events.

<table>
<thead>
<tr>
<th>Events</th>
<th>Direct Linear Programming</th>
<th>Conjugate Function Method</th>
<th>Quantile Cutting Plane Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>CPU</td>
<td>Iter.</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>2.94</td>
<td>562</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>6.61</td>
<td>3170</td>
</tr>
<tr>
<td>150</td>
<td></td>
<td>21.23</td>
<td>9676</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td>393.20</td>
<td>15765</td>
</tr>
<tr>
<td>250</td>
<td></td>
<td>905.11</td>
<td>34911</td>
</tr>
<tr>
<td>300</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table: Performance on problems with unequal probabilities of elementary events.

<table>
<thead>
<tr>
<th>Events</th>
<th>Direct Linear Programming</th>
<th></th>
<th></th>
<th>Quantile Cutting Plane Methods</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>CPU</td>
<td>Iter.</td>
<td>CPU</td>
<td>Cuts</td>
<td>Iter.</td>
<td>CPU</td>
</tr>
<tr>
<td>50</td>
<td>1.391</td>
<td>1044</td>
<td>0.718</td>
<td>23</td>
<td>98</td>
<td>0.749</td>
</tr>
<tr>
<td>100</td>
<td>7.953</td>
<td>4927</td>
<td>1.31</td>
<td>41</td>
<td>231</td>
<td>1.89</td>
</tr>
<tr>
<td>150</td>
<td>23.907</td>
<td>11548</td>
<td>1.313</td>
<td>41</td>
<td>219</td>
<td>2.309</td>
</tr>
<tr>
<td>200</td>
<td>144.171</td>
<td>21321</td>
<td>1.968</td>
<td>58</td>
<td>587</td>
<td>2.938</td>
</tr>
<tr>
<td>250</td>
<td>1928.98</td>
<td>34278</td>
<td>2.027</td>
<td>51</td>
<td>384</td>
<td>1.955</td>
</tr>
<tr>
<td>300</td>
<td>-</td>
<td>-</td>
<td>2.592</td>
<td>69</td>
<td>190</td>
<td>2.812</td>
</tr>
</tbody>
</table>
For any two random variables $X, Y \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$

Risk-Averse Consistency via Expected Utility

$X \succeq_{(2)} Y \iff E_u(X) \geq E_u(Y) \quad \forall$ nondecreasing concave $u : \mathbb{R} \to \mathbb{R}$.

Hadar and Russell (1969)

Risk-Averse Consistency via Rank Dependent Utility

$X \succeq_{(2)} Y$ holds true if and only if for all nondecreasing concave functions $w : [0, 1] \to \mathbb{R}$ that are subdifferentiable at 0

\[
\int_0^1 F_{(-1)}(X; p) \, dw(p) \geq \int_0^1 F_{(-1)}(Y; p) \, dw(p).
\]

Dentcheva and Ruszczyński (2006)
The Implied Rank Dependent Utility Function

Uniform inverse dominance condition (UIDC) for $(\Psi_{-2})$

\[ \exists \tilde{z} \in Z \text{ such that } \inf_{p \in [\alpha, \beta]} \left\{ F_{-2}(G(\tilde{z}); p) - F_{-2}(Y; p) \right\} > 0. \]

Lagrangian-like functional

\[ \Phi(z, w) = f(z) + \int_{0}^{1} F_{-1}(G(z); p) \, dw(p) - \int_{0}^{1} F_{-1}(Y; p) \, dw(p) \]

\( \mathcal{W}([\alpha, \beta]) \) contains all concave and nondecreasing functions

\( w : [0, 1] \rightarrow \mathbb{R} \) such that \( w(p) = 0 \) for all \( p \in [\beta, 1] \) and

\( w(p) = w(\alpha) + c(p - \alpha) \) with some \( c > 0 \), for all \( p \in [0, \alpha] \).

The dual functional \( \Psi(w) = \sup_{z \in Z} \Phi(z, w) \).

The dual problem

\[ (\mathcal{D}_{-2}) \quad \min_{w \in \mathcal{W}([\alpha, \beta])} \Psi(w). \]
The Implied Rank Dependent Utility Function

Theorem

Under the UIDC if problem \((\mathcal{P}_{-2})\) has an optimal solution, then problem \((\mathcal{D}_{-2})\) has an optimal solution and the optimal values of both problems coincide. The optimal solutions of the dual problem \((\mathcal{D}_{-2})\) are the rank dependent utility functions \(\hat{w} \in \mathcal{W}([\alpha, \beta])\) satisfying the dominance constraint and

\[
\int_0^1 F_{(-1)}(G(\hat{z}); p) \, d\hat{w}(p) = \int_0^1 F_{(-1)}(Y; p) \, d\hat{w}(p)
\]

for an optimal solution \(\hat{z}\) of the problem \(\max_{z \in Z} \Phi(z, \hat{w})\).
The Implied Rank Dependent Utility Function

Darinka Dentcheva and Andrzej Ruszczyński

Inverse Stochastic Dominance Constraints
A measure of risk $\varrho$ assigns to an uncertain outcome $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ a real value $\varrho(X)$ on the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.

A coherent measure of risk is a functional $\varrho : \mathcal{L}_1(\Omega, \mathcal{F}, P) \rightarrow \overline{\mathbb{R}}$ satisfying the axioms:

- **Convexity**: $\varrho(\alpha X + (1 - \alpha) Y) \leq \alpha \varrho(X) + (1 - \alpha) \varrho(Y)$ for all $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ and $\alpha \in [0, 1]$.
- **Monotonicity**: If $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ and $Y(\omega) \geq X(\omega) \ \forall \omega \in \Omega$, then $\varrho(Y) \leq \varrho(X)$.
- **Translation Equivariance**: If $a \in \mathbb{R}$ and $X \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$, then $\varrho(X + a) = \varrho(X) - a$.
- **Positive homogeneity**: If $t > 0$ and $X \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$, then $\varrho(tX) = t\varrho(X)$.

Artzner, Delbaen, Eber and Heath; Ruszczynski, Shapiro
Kusuoka Representation of Measures of Risk

**Definition**

**Average Value at Risk** of $X$ at level $p$ is defined as

$$\text{AVaR}_p(X) = -\frac{1}{p} F_{(-2)}(X; p) = \frac{1}{p} \int_0^p \text{VaR}_t(X) \, dt.$$ 

**Theorem**

For every law invariant, finite-valued coherent measure of risk on $L_\infty(\Omega, \mathcal{F}, P)$ on atomless space $\Omega$, a convex set $\mathcal{M} \subset \mathcal{P}((0, 1])$ exists such that for all $X$

$$\varrho(X) = \sup_{\mu \in \mathcal{M}} \int_0^1 \text{AVaR}_p(X) \, \mu(dp).$$

**Definition**

A measure of risk $\varrho$ is called **spectral** if $\mathcal{M}$ is a singleton.
Mean-Risk Models as a Lagrangian Relaxation

Theorem (necessary optimality conditions):

Under the UIDC, if \( \hat{z} \) is an optimal solution of \( (\Psi_{-2}) \), then a spectral risk measure \( \hat{\varrho} \) and a constant \( \kappa \geq 0 \) exist such that \( G(\hat{z}) \) is also an optimal solution of the problem

\[
\max_{z \in Z} \left\{ f(z) - \kappa \hat{\varrho}(G(z)) \right\} \quad \text{and} \quad (1)
\]

\[
\kappa \hat{\varrho}(G(\hat{z})) = \hat{\varrho}(Y). \quad (2)
\]

If the dominance constraint is active, then condition (2) takes on the form \( \hat{\varrho}(G(\hat{z})) = \hat{\varrho}(Y) \).

The support of the measure \( \mu \) in the spectral representation of \( \hat{\varrho}(\cdot) \) is included in \([\alpha, \beta]\).
The Implied Measure of Risk

\[ \varrho_{100}(X) = 0.1069 \text{AVaR}_{0.1772}(X) + 0.014 \text{AVaR}_{0.3101}(X) + 0.0274 \text{AVaR}_{0.3636}(X) + 0.0577 \text{AVaR}_{0.4093}(X) + 0.3073 \text{AVaR}_{0.4594}(X) + 0.2935 \text{AVaR}_{0.4967}(X) + 0.1077 \text{AVaR}_{0.5081}(X) + 0.0576 \text{AVaR}_{0.557}(X) + 0.0213 \text{AVaR}_{0.5647}(X) + 0.0066 \text{AVaR}_{0.575}(X) \]