1 First Order Stochastic Dominance

Let \( F, G : \mathbb{R} \to [0, 1] \) be cumulative probability distributions.

1 Theorem The following are equivalent.
   \[ \forall t \quad F(t) \leq G(t) \] \hspace{1cm} (1)
   For every nondecreasing function \( u : \mathbb{R} \to \mathbb{R} \), \[ \int u \, dF \geq \int u \, dG. \] \hspace{1cm} (2)

2 Definition If either (1) or (2), we say that \( F \) (first order) stochastically dominates \( G \), written \( F \succeq_1 G \).

3 Theorem If \( F \succeq_1 G \) and \( F \neq G \), then for any strictly increasing function \( u \),
   \[ \int u \, dF > \int u \, dG, \]
   provided \( \int u \, dG < \infty \). Consequently, \( F \succeq_1 G \) and \( G \succeq_1 F \) imply that \( F = G \).

Some hints on why this is true: Consider the case where the distributions \( F \) and \( G \) have support in the finite set \( \{ x_1 < \ldots < x_n \} \). Say \( F \) assigns value \( x_i \) probability \( p_i \) (which may be zero), \( i = 1, \ldots, n \), with \( \sum_{i=1}^n p_i = 1 \); and \( G \) assigns point \( x_i \) probability \( q_i > 0 \), \( i = 1, \ldots, n \), with \( \sum_{i=1}^n q_i = 1 \).

For a function \( u \), \( u_i = u(x_i) \). Then the expected utility of \( u \) under \( F \) is \( \sum_{i=1}^n u_i p_i \) and under \( G \) it is \( \sum_{i=1}^n u_i q_i \). Next rewrite this using Abel’s famous formula for “summation by parts.” That is,

\[
\begin{align*}
   u_1p_1 + u_2p_2 + \cdots + u_np_n &= p_1(u_1 - u_2) \\
   &\quad + (p_1 + p_2)(u_2 - u_3) \\
   &\quad + (p_1 + p_2 + \cdots + p_{n-1})(u_{n-1} - u_n) \\
   &\quad + (p_1 + p_2 + \cdots + p_n)u_n \\
   &= u_n - \sum_{i=1}^{n-1} (p_1 + \cdots + p_i)(u_{i+1} - u_i) \\
   &= u_n - \sum_{i=1}^{n-1} F(x_i)(u_{i+1} - u_i)
\end{align*}
\]
Likewise the expected utility of $u$ under $G$ is

$$u_n - \sum_{i=1}^{n-1} G(x_i)(u_{i+1} - u_i).$$

Now if $u$ is nondecreasing, since $x_{i+1} > x_i$ we have that $u_{i+1} - u_i > 0$. So if $F \leq G$, it is clear that $\int u \, dF \geq \int u \, dG$.

By considering $u$ of the form $u(x) = 0$ for $x < x_k$ and $u(x) = 1$ for $x \geq x_k$, we see that $F(x_k) \leq G(x_k)$ is necessary for (2) to hold.

**An integration by parts argument:** (1) $\implies$ (2) Let $u$ be nondecreasing. Assume $F$ and $G$ have common support $[a, b]$. If $u$ is right continuous, then we can integrate by parts to get

$$\int_a^b u(x) \, dF(x) = u(x)F(x) \bigg|_a^b - \int F(x) \, du(x) = u(b) - \int F(x^{-}) \, du(x).$$

Likewise

$$\int_a^b u(x) \, dG(x) = u(b) - \int_a^b G(x^{-}) \, du(x).$$

But (1) implies $\int_a^b G(x^{-}) \, du(x) \leq \int_a^b F(x^{-}) \, du(x)$, so

$$\int_a^b u(x) \, dF(x) \geq \int_a^b u(x) \, dG(x).$$

**A dominance argument:** (1) $\implies$ (2) By Proposition 13 below, setting $X(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}$, $Y(t) = \inf\{x \in \mathbb{R} : G(x) \geq t\}$, we have that $X$ has cdf $F$ and $Y$ has cdf $G$, so

$$\int_0^1 u(X(t)) \, d\lambda(t) = \int u \, dF, \quad \int_0^1 u(Y(t)) \, d\lambda(t) = \int u \, dG.$$

Now observe that (1) implies $X(t) \geq Y(t)$ for each $t$. Thus $\int u \, dF \geq \int u \, dG$.

2 Increasing Risk

Suppose the supports of $F$ and $G$ lie in $[a, b]$. That is, $F(a) = G(a) = 0$ and $F(b) = G(b) = 1$.

**4 Theorem** The following are equivalent.

$$\forall s \in [a, b] \quad \int_a^s F(t) \, dt \leq \int_a^s G(t) \, dt \quad \& \quad \int_a^b F(t) \, dt = \int_a^b G(t) \, dt \quad (3)$$

For every concave $u$, $\int u \, dF \geq \int u \, dG. \quad (4)$
Proof that (4) implies (3): Let \( s \in [a, b] \). Integrating by parts,

\[
\int_a^s F(t) \, dt = tF(t) \bigg|_a^s - \int_a^s t \, dF(t).
\]

\[
= sF(s) - \int_a^s t \, dF(t)
\]

\[
= \int_a^s (s-t) \, dF(t)
\]

\[
= \int_a^b (s-t)^+ \, dF(t).
\]

Similarly

\[
\int_a^s G(t) \, dt = \int_a^b (s-t)^+ \, dG(t).
\]

Since \( (s-t)^+ \) is a convex function of \( t \), (4) implies

\[
\int_a^s F(t) \, dt = \int_a^b (s-t)^+ \, dF(t) \leq \int_a^b (s-t)^+ \, dG(t) = \int_a^s G(t) \, dt.
\]

When \( s = b \), this becomes

\[
\int_a^b F(t) \, dt = \int_a^b (b-t) \, dF(t).
\]

Now \( b-t \) is both convex and concave in \( t \), so we must have

\[
\int_a^b F(t) \, dt = \int_a^b G(t) \, dt.
\]

\[\blacksquare\]

Figure 1. The nonincreasing convex function \((s-t)^+\).

"Proof" that (3) implies (4): Define \( \Phi(s) = \int_a^s F(t) \, dt \) and \( \Gamma(s) = \int_a^s G(t) \, dt \). Let \( u \) be concave. Then

\[
\int_a^b u(t) \, dF(t) = u(t)F(t) \bigg|_a^b - \int_a^b F(t) u'(t) \, dt
\]

\[
= u(b) - \int_a^b F(t) u'(t) \, dt
\]

\[
= u(b) - \left( \Phi(t)u'(t) \right) \bigg|_a^b - \int_a^b \Phi(t) u''(t) \, dt
\]

\[
= u(b) - \Phi(b)u'(b) + \int_a^b \Phi(t) u''(t) \, dt.
\]

Likewise,

\[
\int_a^b u(t) \, dG(t) = u(b) - \Gamma(b)u'(b) + \int_a^b \Gamma(t) u''(t) \, dt.
\]
But (3) implies \( \Gamma(b) = \Phi(b) \) and \( \Gamma(t) \geq \Phi(t) \) for all \( t \). Since \( u \) is concave, \( u''(t) \leq 0 \) for all \( t \). Thus \( \Phi(t)u''(t) \geq \Gamma(t)u''(t) \) for all \( t \), so

\[
\int_a^b u(t) dF(t) = u(b) - \Phi(b)u'(b) + \int_a^b \Phi(t)u''(t) dt \\
= u(b) - \Gamma(b)u'(b) + \int_a^b \Phi(t)u''(t) dt \\
\geq u(b) - \Gamma(b)u'(b) + \int_a^b \Gamma(t)u''(t) dt \\
= \int_a^b u(t) dG(t).
\]

5 Definition If either (3) or (4) holds, we say that \( G \) is riskier than \( F \).

6 Theorem If \( G \) is riskier than \( F \) and \( F \neq G \), then for any strictly concave \( u \) on \( [a, b] \),

\[
\int u \, dF > \int u \, dG.
\]

Consequently, if \( G \) is riskier than \( F \) and \( F \) is riskier than \( G \), then \( F = G \).

3 Second Order Stochastic Dominance

Suppose the supports of \( F \) and \( G \) lie in \( [a, b] \).

7 Theorem The following are equivalent.

\[
\forall s \in [a, b] \quad \int_a^s F(t) dt \leq \int_a^s G(t) dt \tag{5}
\]

For all nondecreasing concave \( u \) defined on \( [a, b] \),

\[
\int u \, dF \geq \int u \, dG. \tag{6}
\]

Proof: The proof of this is virtually identical to that of Theorem 4, taking note of the fact that (6) is equivalent to

For all nonincreasing convex \( u \) defined on \( [a, b] \),

\[
\int u \, dF \leq \int u \, dG,
\]

and the fact that \( (s - t)^+ \) is a nonincreasing convex function of \( t \). Note that \( b - t \) is not a nondecreasing concave function of \( t \), so we cannot conclude \( \int_a^b F(t) dt \geq \int_a^b G(t) dt \), so the two integrals need not be equal.

8 Definition If either (5) or (6) holds, then we say that \( F \) second order stochastically dominates \( G \), written \( F \succ_{2} G \).
9 Theorem If \( u \) is strictly increasing and strictly concave and \( F \succcurlyeq_2 G \) and \( F \neq G \), then
\[
\int u \, dF > \int u \, dG.
\]

Thus \( F \succcurlyeq_2 G \) and \( G \succcurlyeq_2 F \) imply \( F = G \).

Now drop the assumption that \( F \) and \( G \) have bounded support.

10 Theorem The following are equivalent.
\[
\forall s \in \mathbb{R} \quad \int (x \wedge s) \, dF(x) \geq \int (x \wedge s) \, dG(x). \tag{7}
\]

For all nondecreasing concave \( u \) defined on the support of both \( F \) and \( G \),
\[
\int u \, dF \geq \int u \, dG \tag{8}
\]

11 Definition In this case we still say \( F \succcurlyeq_2 G \).

12 Theorem If \( u \) is strictly increasing and strictly concave and defined on the support of both \( F \) and \( G \), and \( F \succcurlyeq_2 G \) and \( F \neq G \), then
\[
\int u \, dF > \int u \, dG.
\]

A Constructing a random variable with a given cdf

13 Proposition Given any function \( F: \mathbb{R} \to [0,1] \) that is nondecreasing, right continuous, and satisfies \( \lim_{t \to -\infty} F(t) = 0 \) and \( \lim_{t \to \infty} F(t) = 1 \), there is a random variable \( X \) on the standard probability space \((\mathbb{R}, \mathcal{B}, \lambda)\) with \( F = F_X \).

Proof: Given such an \( F \), define \( X : [0,1] \to \mathbb{R} \) by
\[
X(t) = \inf \{ x \in \mathbb{R} : F(x) \geq t \}.
\]
(This makes \( X(0) = -\infty \), but that’s okay since \( \lambda\{0\} = 0 \).) When \( F \) is strictly increasing and maps onto \([0,1]\), then \( X \) is just \( F^{-1} \). More generally, flat spots in \( F \) correspond to jumps in \( X \) and vice-versa. See Figure 2.

First note that \( X \) is nondecreasing, and therefore Borel measurable (inverse images of intervals are intervals). To see this, let \( t < s \). Then
\[
\{ z \in \mathbb{R} : F(z) \geq s \} \subset \{ z \in \mathbb{R} : F(z) \geq t \},
\]
so
\[
X(t) = \inf \{ z \in \mathbb{R} : F(z) \geq t \} \leq \inf \{ z \in \mathbb{R} : F(z) \geq s \} = X(s).
\]
Thus \( X \) is a random variable on the probability space \([0,1]\). Since \( F \) is right continuous, another key property is that
\[
X(t) \leq y \iff t \leq F(y),
\]
which implies that
\[ \{ t \in [0, 1] : X(t) \leq y \} = [0, F(y)], \]
so
\[ \lambda \{ t \in [0, 1] : X(t) \leq y \} = F(y). \]
In other words, \( F \) is the cdf of \( X \).

B Integration by parts

This is not the most general integration by parts theorem, but it is not bad, and it is easy to prove using Fubini’s Theorem. I have such a proof in a separate handout.

14 Integration by Parts  Suppose \( F \) and \( G \) satisfy \( F(x) = F(a) + \int_a^x f(s) \, ds \) and \( G(x) = G(a) + \int_a^x g(s) \, ds \) for every \( x \) in \([a, b] \), where \( f \) and \( g \) are integrable over \([a, b] \) and \( fg \) is integrable over \([a, b] \times [a, b] \). Then
\[
\int_a^b F(x)g(x) \, dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x) \, dx.
\]

C A Still More General Result

C.1 Finite measures and nondecreasing functions

Let \( \mu \) be a finite (nonnegative) measure on the Borel subsets of \( R \). Define the function \( F_\mu : R \to R_+ \) by
\[
F_\mu(x) = \mu(\{ y \in R : y \leq x \}).
\]
\( F_\mu \) is called the distribution function of \( \mu \), and has the following properties:

1. \( F_\mu \) is nondecreasing.

2. \( F_\mu \) is right continuous. That is, \( F_\mu(x) = \lim_{y \downarrow x} F_\mu(y) \).
3. \( \lim_{x \to -\infty} F_\mu(x) = 0 \).

4. \( \lim_{x \to \infty} F_\mu(x) = \mu(R) \).

5. \( F(b) - F(a) = \mu((a, b]) \) for \( a \leq b \).

Conversely, for any \( F : R \to R_+ \) satisfying

1. \( F \) is nondecreasing.

2. \( F \) is right continuous.

3. \( \lim_{x \to -\infty} F(x) = 0 \).

4. \( \lim_{x \to \infty} F(x) < \infty \).

there is a unique nonnegative Borel measure \( \mu_F \) satisfying \( \mu_F((a, b]) = F(b) - F(a) \) for \( a \leq b \).

Given a distribution function \( F : R \to R_+ \) and a \( \mu_F \)-integrable function \( g \), the Lebesgue–Stieltjes integral

\[
\int g \, dF = \int g \, d\mu_F
\]

by definition.

**15 Integration by Parts for Distribution Functions** Let \( F \) and \( G \) be distribution functions on \( R \). Then

\[
\int_{[a,b]} F(x) \, dG(x) = F(b)G(b) - F(a)G(a) - \int_{[a,b]} G(x^-) \, dF(x),
\]

where \( G(x^-) = \lim_{y \uparrow x} G(y) \).

**Proof:** Define \( A = \{(x, y) \in (a, b]^2 : x \leq y \} \). By Fubini’s Theorem, we have

\[
\int \int I_A \, d(\mu_G \times \mu_F) = \int_{[a,b]} \left( \int_{[a,b]} I_A \, d\mu_G \right) \, d\mu_F
\]

\[
= \int_{[a,b]} \left( \int_{[a,b]} I_A \, d\mu_F \right) \, d\mu_G = \int_{[a,b]} (F(x) - F(a)) \, d\mu_G(x)
\]

\[
= \int_{[a,b]} (G(b) - G(y^-)) \, d\mu_F(y).
\]

Rearrange to get

\[
\int_{[a,b]} (F(x) - F(a)) \, d\mu_G(x) = \int_{[a,b]} (G(b) - G(y^-)) \, d\mu_F(y)
\]

or

\[
\int_{[a,b]} F(x) \, dG(x) - F(a)(G(b) - G(a)) = G(b)(F(b) - F(a)) - \int_{[a,b]} G(y^-) \, d\mu_F(y),
\]

from which the conclusion follows.

**16 Corollary** If either \( F \) or \( G \) is continuous, then

\[
\int_{[a,b]} F(x) \, dG(x) = F(b)G(b) - F(a)G(a) - \int_{[a,b]} G(x) \, dF(x).
\]

**References**