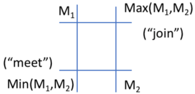
Insights on Monotone Methods in Economics You Must Know

Lones Smith

Madison, 2025

Join, Meet, Lattice

- A **poset** is a set X and a partial order $\succeq$
- The **join** $x \vee x'$ is the supremum of x, x'

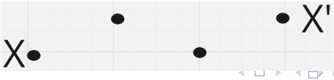


- The **meet** $x \wedge x'$ the infimum of x and x'
- A **lattice** is a poset that contains all meets and joins
- We restrict to Euclidean lattices $X \subset \mathbb{R}^n$, where

$$\mathbf{x} \vee \mathbf{x}' = (\max\{x_1, x_1'\}, ..., \max\{x_N, x_N'\})$$
$$\mathbf{x} \wedge \mathbf{x}' = (\min\{x_1, x_1'\}, ..., \min\{x_N, x_N'\})$$

- **Strong Set Order (SSO)**, denoted $\sqsupseteq$
- $X \sqsupseteq X'$ if for all $x \in X, x' \in X'$, $x \vee x' \in X$ & $x \wedge x' \in X'$.

- Prove $X' \sqsupseteq X$ fails here:



- $F : X \to \mathbb{R}$ is **supermodular** (SPM) if for all $x, x' \in X$

$$F(x \wedge x') + F(x \vee x') \geq F(x) + F(x')$$

- Fact: A function on a totally ordered set (*chain*) is SPM
- If $F(x, \theta)$ is SPM, then F has **increasing differences** (ID) in (x, θ) if $F(x_2, \theta) - F(x_1, \theta)$ increases in θ, for all $x_2 > x_1$
- If $F : \mathbb{R}^n \to \mathbb{R}$ is C^2, then F is SPM iff $\frac{\partial^2 F}{\partial x_i \partial x_j} \geq 0$ for all x
- Addition: If $F, G : X \to \mathbb{R}$ are SPM, then $F + G$ is SPM

Lemma (Maximization Preserves SPM)

F SPM on the lattice $X \times Y \Rightarrow G(x) = \sup_y F(x, y)$ SPM on X.

- *Proof:* Let $y, y' \in Y$ and $x, x' \in X$. Since F is SPM:

$$\begin{aligned} F(x', y') + F(x, y) &\leq F(x \vee x', y' \vee y) + F(x \wedge x', y' \wedge y) \\ &\leq G(x' \vee x) + G(x' \wedge x) \end{aligned}$$

- So $G(x' \vee x) + G(x' \wedge x)$ is an upper bound for the LHS.
- Maximizing the left side over all y, y', we get:

$$G(x') + G(x) \leq G(x' \vee x) + G(x' \wedge x) \qquad \square$$

Comparative Statics

- Let $\mathbf{X}^*(\theta)$ be the set of solutions to the problem

$$\max_{\mathbf{x} \in X} F(\mathbf{x}, \theta)$$

- **Topkis Theorem (1978):** *Let X be a lattice, and Θ a poset. If $F : X \times \Theta \to \mathbb{R}$ has ID in (x, θ) and is SPM in x, then $X^*(\theta)$ is monotone in the SSO.*
- *Proof:* Let $\theta' \succ \theta''$ and $x' \in X^*(\theta')$ and $x'' \in X^*(\theta'')$.

$$\begin{aligned}
0 &\geq F(x' \vee x'', \theta') - F(x', \theta') && \text{by } x' \in X^*(\theta') \\
&\geq F(x' \vee x'', \theta'') - F(x', \theta'') && \text{by ID in } (x, \theta) \\
&\geq F(x'', \theta'') - F(x' \wedge x'', \theta'') && \text{by SPM in } x \\
&\geq 0 && \text{by } x'' \in X^*(\theta'')
\end{aligned}$$

- All inequalities are therefore equalities
- Then $x' \vee x'' \in X^*(\theta')$ and $x' \wedge x'' \in X^*(\theta')$
- So $X^*(\theta)$ is increasing in the SSO. $\square$

Quasi-supermodularity

- $F\colon X \to \mathbb{R}$ is **quasi-supermodular** (QSPM) if $\forall x, x' \in X$:

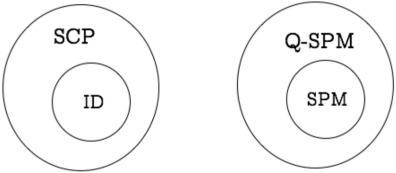
$$\begin{aligned} F(x) \geq F(x \wedge x') &\quad\Rightarrow\quad F(x \vee x') \geq F(x') \\ F(x) > F(x \wedge x') &\quad\Rightarrow\quad F(x \vee x') > F(x') \end{aligned}$$

- The contrapositive of each yields the equivalent:

$$\begin{aligned} F(x) < F(x \wedge x') &\quad\Leftarrow\quad F(x \vee x') < F(x') \\ F(x) \leq F(x \wedge x') &\quad\Leftarrow\quad F(x \vee x') \leq F(x') \end{aligned}$$

- If $F(x, \theta)$ is QSPM, then F obeys the **single crossing property** in (x, θ) if for all $x_2 \succ x_1$ and $\theta_2 \succ \theta_1$

$$\begin{aligned} F(x_2, \theta_1) \geq F(x_1, \theta_1) &\quad\Rightarrow\quad F(x_2, \theta_2) \geq F(x_1, \theta_2) \\ F(x_2, \theta_1) > F(x_1, \theta_1) &\quad\Rightarrow\quad F(x_2, \theta_2) > F(x_1, \theta_2) \end{aligned}$$



Ordinal Comparative Statics

- ***Milgrom-Shannon Theorem (1994):*** *Let X be a lattice, and Θ a poset. If $F\colon X \times \Theta \to \mathbb{R}$ obeys the SCP in (x, θ) and is QSPM in x, then $X^*(\theta)$ is monotone in the SSO.*
- Proof: Let $\theta' \succeq \theta$ with $x \in X^*(\theta)$ and $x' \in X^*(\theta')$.
- $x \vee x' \in X^*(\theta')$ since

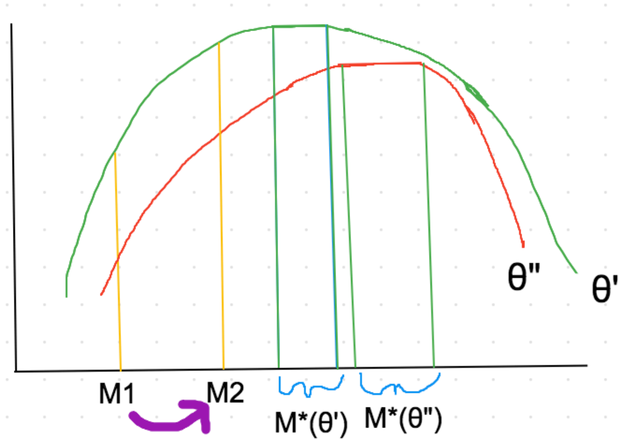
$$\begin{array}{rrcll} & F(x,\theta) & \geq & F(x \wedge x', \theta) & \text{by } x \in X^*(\theta) \\ \Rightarrow & F(x \vee x', \theta) & \geq & F(x', \theta) & \text{by QSPM} \\ \Rightarrow & F(x \vee x', \theta') & \geq & F(x', \theta') & \text{by SCP} \end{array}$$

- Next, $x \wedge x' \in X^*(\theta)$ since:

$$\begin{array}{rrcll} & F(x',\theta') & \geq & F(x \vee x', \theta') & \text{by } x' \in X^*(\theta') \\ \Rightarrow & F(x \wedge x', \theta') & \geq & F(x', \theta') & \text{by QSPM} \\ \Rightarrow & F(x \wedge x', \theta) & \geq & F(x', \theta) & \text{by SCP} \end{array}$$

- We applied the contrapositive forms of QSPM and SCP

Graphical Intuition for the Single Crossing Property



- Since the reals are a totally ordered set, any function on the reals is automatically SPM.

Ordinal Comparative Statics without a Lattice

- Let X and Θ be posets.
- The correspondence $\mathcal{X} : \Theta \to X$ is **nowhere decreasing** if $x_2 \in \mathcal{X}(\theta_1)$ and $x_1 \in \mathcal{X}(\theta_2)$ with $x_2 \succeq x_1$ and $\theta_2 \succeq \theta_1$ imply $x_1 \in \mathcal{X}(\theta_1)$ and $x_2 \in \mathcal{X}(\theta_2)$.
- So the correspondence does not fall anywhere
- ***Nowhere Decreasing Optimizers (2018):*** *Let X and Θ be posets. If $F\colon X \times \Theta \to \mathbb{R}$ obeys the single crossing property, then $X^*(\theta) \equiv \arg\max_{x\in X} F(x,\theta)$ is nowhere decreasing in θ.*
- If $\theta_2 \succeq \theta_1$, $x_2 \in \mathcal{X}(\theta_1)$, $x_1 \in \mathcal{X}(\theta_2)$, and $x_2 \succeq x_1$, optimality and the single crossing property give $x_2 \in \mathcal{X}(\theta_2)$, since:

$$F(x_2,\theta_1) \geq F(x_1,\theta_1) \quad \Rightarrow \quad F(x_2,\theta_2) \geq F(x_1,\theta_2)$$

- Exercise: Prove that $x_1 \in \mathcal{X}(\theta_1)$

Upcrossing Functions

- Let Θ be a poset. Then $\Upsilon : \Theta \to \mathbb{R}$ is *upcrossing* if
    - $\Upsilon(\theta) \geq 0 \Rightarrow \Upsilon(\theta') \geq 0$ for all $\theta' > \theta$
    - $\Upsilon(\theta) > 0 \Rightarrow \Upsilon(\theta') > 0$ for all $\theta' > \theta$
- Υ is *downcrossing* if $-\Upsilon$ is upcrossing
- Υ is *one-crossing* if it is upcrossing or downcrossing.

Upcrossing Preservation

- ***Karlin and Rubin (1956):*** *If Υ is upcrossing, and $\lambda > 0$ is nondecreasing, and μ is a measure, then*

$$\int_{-\infty}^{\infty} \Upsilon(s) d\mu(s) \geq (>) 0 \Rightarrow \int_{-\infty}^{\infty} \Upsilon(s)\lambda(s) d\mu(s) \geq (>) 0$$

- Proof: Let Υ first upcross at t_0. The right side equals

$$\begin{aligned} & \textstyle\int_{-\infty}^{t_0} \Upsilon(s)\lambda(s)d\mu(s) + \int_{t_0}^{\infty} \Upsilon(s)\lambda(s)d\mu(s) \\ \geq (>) \quad & \textstyle\lambda(t_0)\int_{-\infty}^{t_0} \Upsilon(s)d\mu(s) + \lambda(t_0)\int_{t_0}^{\infty} \Upsilon(s)d\mu(s) \\ = \quad & \textstyle\lambda(t_0)\int_{-\infty}^{\infty} \Upsilon(s)d\mu(s) \geq 0 \end{aligned}$$

- Weakening the assumptions on Υ has led to key papers

Monotone Stochastic Dominance

- Let X have cdf F and Y have cdf G.
- *First Order Stochastic Dominance:* $F \succsim_{FOD} G$ if $F(x) \leq G(x)\ \forall x$, iff survivors obey $\bar{F}(x) \geq \bar{G}(x)\ \forall x$
- ***Monotone Ranking Theorem.*** *If $F \succsim_{FOD} G$, then any mean of a monotone function is higher for X than Y.*
- Proof: Intuitively, every increasing function can be thought as the limit of the sum of step functions $\mathbb{I}\{x \geq a\}$.
- Partition the domain $[0, 1]$ with $0 = a_0 < \cdots < a_N = 1$.
- Pick $0 < w_0 < w_2 < \cdots < w_N$.
- Define the weighted sum of step functions:

$$u_N(x) = \sum_{k=0}^{N} w_k \mathbb{I}\{x \geq a_k\}$$

- The mean of $u(\cdot)$ is higher under F than G:

$$\mathbb{E}u_N(X) = \sum_{k=0}^{N} w_k[1 - F(a_k)] \geq \sum_{k=0}^{N} w_k[1 - G(a_k)] = \mathbb{E}u_N(Y)$$

- Take the limit as the mesh vanishes $\Rightarrow Eu(X) \geq Eu(Y)$.
- Which utility function makes the ranking theorem iff?

Monotone Concave Stochastic Dominance

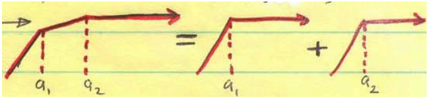
- *Second Order Stochastic Dominance:* $F \succsim_{SOSD} G$ if

$$\int_0^x F(t)dt \leq \int_0^x G(t)dt \quad \forall x$$

- ***Monotone Concave Stochastic Order.*** *If $F \succsim_{SOSD} G$, then any mean of a monotone concave utility function is higher for F than G.*
- Proof: We prove this for "ramp" functions $u_a = \min\{a, x\}$.
- Suppose that $0 \leq X, Y \leq M$. Then:

$$\begin{aligned} \mathbb{E}_F u_a(X) &= \int_0^a x dF(x) + \int_a^M a dF(x) \\ &= xF(x)\Big|_0^a - \int_0^a F(x)dx + a(1 - F(a)) \\ &= a - \int_0^a F(x)dx \end{aligned}$$

- So $\mathbb{E}_F u_a(X) \geq \mathbb{E}_G u_a(Y)$ iff $-\int_0^a F(x)dx \geq -\int_0^a G(x)dx$
- Intuitively, concave functions through the origin can be approximated as the limit of weighted sums of ramps
- So $\mathbb{E}_F u(\cdot) \geq \mathbb{E}_G u(\cdot)$. $\square$



- **Stochastic Dominance on Closed Cones**
- A **convex cone** is a vector space subset closed under positive linear combinations with positive coefficients.
- If the mean of F exceeds G on a set of functions V, then this holds on the convex cone $U = cc(V \cup \{\pm 1\})$.
- Example: If $U = \{$ concave functions$\}$ and $V = \{\min\langle 0, x-a\rangle\} \cup \{\pi(x) = -x\}$ then $U = cc(V \cup \pm 1)$

$$\mathbb{E}_F(\min\langle 0, X-a) \geq \mathbb{E}_F(\min\langle 0, X-a) \ \forall a \quad \text{and} \quad \mathbb{E}_F(-X) \geq \mathbb{E}_G(-X)$$

$\Rightarrow$ F, G have same mean $\Rightarrow \int_0^1 [1-F(t)]dt = \int_0^1 [1-G(t)]dt$

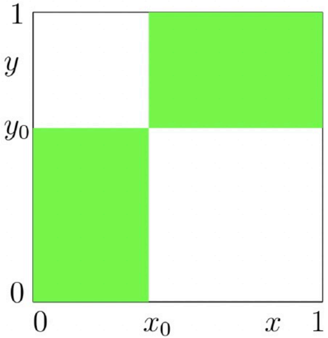
- *Mean Preserving Spread:* G is a MPS of F on $[0,1]$ if

$$\int_0^x F(t)dt \leq \int_0^x G(t)dt \quad \forall x, \text{ with equality at } x = 1$$

- ***Concave Stochastic Order.*** *If $F \succsim_{MPS} G$, then any mean of a concave utility function is higher for G than F.*

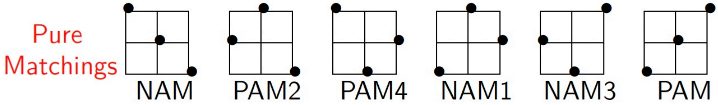
  EXAMPLE. If F is a MPS of G, then $\sigma_F^2 \geq \sigma_G^2$.

PQD: Increased Sorting in Pairwise Matches



- *Positive quadrant dependence* (PQD) partially orders bivariate probability distributions $M \in \mathcal{M}(G, H)$
- *Sorting increases in the PQD order* if the mass in every northeast and southwest quadrant increases.
- So $M_2 \succeq_{PQD} M_1$ *iff* $M_2(x, y) \geq M_1(x, y)$ *for all* x, y
- We call M_2 *more sorted than* M_1

PQD Order with Three Types



$$\text{NAM} \preceq_{PQD} [\text{PAM2}, \text{PAM4}] \preceq_{PQD} [\text{NAM1}, \text{NAM3}] \preceq_{PQD} \text{PAM}$$

- Check that this is not a lattice!
  - For PAM2 or PAM4, each of NAM1, NAM3, and PAM are upper bounds, but there is no least upper bound
  - For NAM1 or NAM3, each of PAM2. PAM4, and NAM are lower bounds, but there is no greatest upper bound
  - Hence, PQD partial order is not a lattice on three types
  - Maybe we are missing mixed matchings that restore the lattice property. There is not.

PQD - SPM Stochastic Dominance Theorem

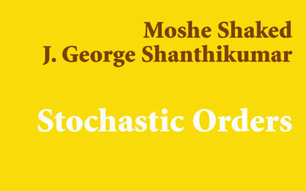
- This is missing from first year micro PhD curriculum:

Lemma (PQD Stochastic Dominance Theorem)

The PQD and SPM orders coincide on R^2, i.e. increases in the PQD order raise (lower) the total output for any SPM (SBM) function f, and conversely:

$$M_2 \succeq_{PQD} M_1 \Leftrightarrow \int \phi(x,y) M_2(dx,dy) \geq \int \phi(x,y) M_1(dx,dy)$$

- Hence, PQD is called the supermodular order
- Method of cones intuition: a SPM function is in the cone of indicator functions $[x,\infty) \times [y,\infty) \cup (-\infty,x] \times (-\infty,y]$



Economics of the PQD Order

Lemma (PQD is Economically Relevant)

If sorting increases in the PQD order,
(a) the average distance between matched types falls;
(b) the covariance / correlation of matched pairs rises, and
(c) the coefficient in a linear regression of men's type on matched women's type increases.

- PROOF OF (*a*)
  - Claim: $\phi(x, y)$ is SBM for all $\gamma \geq 1$.
  - $E[\phi(X, Y)] = |X - Y|^\gamma$ over matched X, Y falls if $\gamma \geq 1$
- PROOF OF (*b*)
  - xy SPM $\Rightarrow$ covariance $E_M[XY] - E[X]E[Y]$ increases
  - Marginal distributions on X and Y are invariant to M.
  - $E[X^2]$ and $E[Y^2]$ fixed in match measure M

    $\Rightarrow$ correlation coefficient increases too
- PROOF OF (*c*): You try it! It's not hard!

Beyond the SCP: Interval Dominance Order

- At this point, detour to Quah's Interval Dominance Order slides
- We apply cone logic to relax Milgrom and Shannon's SCP premise

Theorem

The maximizer set $\arg\max_x V(x,t)$ *increases in* t *provided:*

$$\exists \alpha > 0 \textit{ nondecreasing: } V_x(x|t_2) \geq \alpha(x) V_x(x|t_1) \; \forall t_2 > t_1 \quad (\bigstar)$$

- Let $t_2 > t_1$ and $x_i \in \arg\max_x V(x|t_i)$ for $i = 1, 2$
- **Claim 1**: $\max(x_1, x_2) \in \arg\max_x V(x|t_2)$
- True if $x_2 \geq x_1$. Assume $x_1 > x_2$.

$$V(x_1|t_2) - V(x_2|t_2) = \int_{x_2}^{x_1} V_x(x|t_2)dx \geq \int_{x_2}^{x_1} \alpha(x) V_x(x|t_1)dx \quad (\ddagger)$$

- $x_1 \in \arg\max_x V(x,t_1) \Rightarrow \int_y^{x_1} V_x(x|t_1)dx \geq 0 \; \forall y \in [x_2, x_1]$.

$\Rightarrow$ $\int_{x_2}^{x_1} \alpha(x) V_x(x|t_1)dx \geq 0$ by integral SCP

- By $(\ddagger)$, $V(x_1|t_2) \geq V(x_2|t_2)$
- Altogether, $\max(x_1, x_2) \in \arg\max V(x, t_2)$

Topkis without the Single Crossing Property

- **Claim 2**: $\min(x_1, x_2) \in \arg\max V(x|t_1)$.
- True if $x_1 \leq x_2$. Assume $x_1 > x_2$.
- For a contradiction, assume that $V(x_1|t_1) > V(x_2|t_1)$.
- Then $\int_{x_2}^{x_1} V_x(x|t_1)dx > 0$.
- $x_1 \in \arg\max V(x, t_1) \Rightarrow \int_y^{x_1} V_x(x|t_1)dx \geq 0 \ \forall y \in [x_2, x_1]$.

$\Rightarrow$ $\int_{x_2}^{x_1} \alpha(x) V_x(x|t_1)dx > 0$ by strict integral SCP

- By (‡), $V(x_1|t_2) - V(x_2|t_2) > 0$
- This contradicts $x_2 \in \arg\max V(x|t_2)$.

$\Rightarrow$ $V(x_1|t_1) = V(x_2|t_1)$

$\Rightarrow$ $\min(x_1, x_2) \in \arg\max V(x|t_1)$.

- PS: This proof is far more general than in Quah and Strulovici, since it uses the method of cones

Beyond Karlin and Rubin: Integral Single Crossing Property

Corollary (Integral Single Crossing Property)

If $\alpha(x) \geq 0$ is nondecreasing, then (if all integrals are finite)

$$\int_{[y,\infty)\cap X} f(x)dx \geq 0 \quad \text{for all } y \qquad \Rightarrow \qquad \int_X f(x)\alpha(x)dx \geq 0$$

Inequality is strict if $\int_X f(x)dx > 0$ and $\exists m > 0$ s.t. $\alpha(x) \geq m$

- Instead of f upcrossing, we assume $\int f$ upcrossing
- Idea: $\alpha \geq 0$ non$\downarrow$ $\Rightarrow \int_X f(x)\alpha(x)dx$ lies in cone of $(\int_{[y,\infty)\cap X} f(x)dx)$ $\forall y$
- This is clear for α a step function, or a sum of step functions, etc.
- Formal Proof (read on your own):
    - Since α is monotone, its upper sets are $U = [y, \infty)$
    - Fix $M > 0$ very big
    - Let $\alpha_M = M$ for $x \in U(M)$ and $\alpha_M(x) = \alpha(x)$ otherwise
    - Banks Lemma ($m > 0$ on next slide) $\Rightarrow \int_X f(x)\alpha_M(x)dx \geq 0$
    - If $M \uparrow \infty$, get $\int_X f(x)\alpha(x)dx \geq 0$ by monotone convergence theorem

Offline: Dallas Banks Integral Inequality

- Beesack (1957), "A note on an integral inequality"
- *upper set* $U(y) = \{x \in X \subset \mathbb{R}, \alpha(x) \geq y\}$ of function α

Lemma (Banks Lemma, 1963)

If $m \leq \alpha(x) \leq M < \infty \ \forall x \in X$ *then*

$$\int_X f(x)\alpha(x)dx = m\int_X f(x)dx + \int_m^M \left(\int_{U(y)} f(x)dx\right) dy \quad (\dagger)$$

- See "layer cake" integral notion in wikipedia, but swap $\alpha(x)$ and $f(x)$
- Proof: Define $F(y) = \int_{U(y)} f(x)dx$ for $y \in [m, M)$, and $F(M) = 0$
- Layer Cake Claim: $\int_X f(x)\alpha(x)dx = -\int_m ydF(y)$
  - Proof: Take a partition $m = y_0 < y_1 < \cdots < y_n = M$
  - $\int_m ydF(y) \sim \sum_{k=1}^n y_i[F(y_{k-1}) - F(y_k)] \sim \sum_{k=1}^n \alpha(x_k)f(x_k)\Delta x_k$ since $y_k \leq \alpha(x) \leq y_{k+1}$ on $U(y_{k-1}) \setminus U(y_k)$
- Integrate Layer Cake Claim by parts to get $(\dagger)$

$$\int_X f(x)\alpha(x)dx = -yF(y)|_m^M + \int_m^M F(y)dy$$

Bayes Rule

- Imagine we are trying to learning about the *state of the world* $\theta \in \Theta$ with a prior density $g(\theta)$ and cdf, if $\theta \in \mathbb{R}$
- Typical case: $\Theta = \{L, H\}$.
- A signal is r.v. X whose density $f(x|\theta)$ depends on θ
- A *signal* is a family of r.v.s $\{f(x|\theta), \theta \in \Theta\}$, for every state
- Standard abuse of terminology: the "signal realization" x is often called the "signal"
- By Bayes rule, upon seeing x, the posterior density is

$$g(\theta|x) = \frac{pdf(\theta \text{ and } x)}{pdf(x)} = \frac{g(\theta)f(x|\theta)}{\int_{-\infty}^{\infty} f(x|s)g(s)} \propto g(\theta)f(x|\theta)$$

Odds Formulation of Bayes' Rule

- To eliminate the messy denominator, we often use odds
- Posterior odds = (prior odds) $\times$ (likelihood ratio)

$$\frac{g(\theta_2|x)}{g(\theta_1|x)} = \frac{g(\theta_2)f(x|\theta_2)}{g(\theta_1)f(x|\theta_1)}$$

- Example: A test to detect AIDS, whose prevalence is $\frac{1}{1000}$, has a false positive rate of 5%.
- Given a + result, what is the chance one is infected?
- Roughly: Posterior odds against infection are

$$\frac{999}{1} \times \frac{1}{19} \approx 1000/20 = 50$$

- One is infected with chance 2%

More Favorable Signals

- Recall first order stochastic dominance (FOSD): $G^2 \succ_{FSD} G^1$ if $G^2(s) \leq G^1(s)$ for all s
- We will prove this later with fancy method of cones.
- ***First Ranking Theorem***: *$G^2 \succeq_{FOSD} G^1$ iff $E_{G^2}\psi(X) \geq E_{G^1}\psi(X)$ for all nondecreasing functions ψ*
- Signal realization x is *more favorable than* y if $G(\cdot|x) \succeq_{FSD} G(\cdot|y)$ for all nondegenerate priors G
- Idea: you to think θ is bigger after seeing x vs. y
- If g is discrete, with $g(\theta_1) = g(\theta_2) = \frac{1}{2}$, then x is more favorable than y only if

$$\theta_2 > \theta_1 \Leftrightarrow \frac{f(x|\theta_2)}{f(x|\theta_1)} > \frac{f(y|\theta_2)}{f(y|\theta_1)}$$

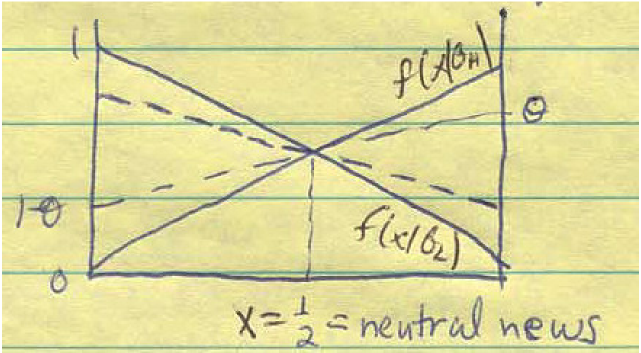
- Soon: This is iff for any prior on state spaces $\Theta \subset \mathbb{R}$

Monotone Likelihood Ratio Property (MLRP)

- Signal $\{f(x|\theta)\}$ obeys the MLRP iff x is more favorable than any $y < x$
- $f(x, \theta)$ is *affiliated* if $f(x_1|\theta_1)f(x_2|\theta_2) \geq f(x_1|\theta_2)f(x_2|\theta_1)$
- Classic Signal Families with MLRP
  1. exponential: $f(x|\theta) = (1/\theta)e^{-x/\theta}$, $x \geq 0$
  2. skewed uniform: $f(x|\theta) = nx^{n-1}/\theta^n$, $0 \leq x \leq \theta$
  3. binomial: $f(x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$, $x = 0, 1, ..., n$
- Signal outcomes x and y are *equivalent* if $f(x|\theta_2)f(y|\theta_1) = f(x|\theta_1)f(y|\theta_2)\ \forall\theta_1, \theta_2$.
- Signal outcome x is *neutral news* if $f(x|\theta_1) = f(x|\theta_2)\ \forall\theta_1, \theta_2$ $\Rightarrow G \equiv G(\cdot|x)$
- Signal outcome x is *good news* if it is *more favorable than neutral news*, i.e. iff $f(x|\theta)$ is increasing in θ (*bad news* it if is less favorable).

Good News and Bad News

- Most signal distributions have no neutral news signal outcomes
- To see why neutral news is rare, consider this example: $\theta \sim U(0,1)$, and $f(x|\theta) = 2(\theta x + (1-\theta)(1-x))$.
- So $f(\frac{1}{2}|\theta) = 2(\frac{1}{2}\theta + \frac{1}{2}(1-\theta)) = 1$ for all states θ



MLRP: What is a Good Signal? (Milgrom, 1981)

- ***Theorem**: x is more favorable than y iff $\forall \theta_2 > \theta_1$, $f(x|\theta_1) f(y|\theta_2) \geq f(x|\theta_1) f(y|\theta_2)$.*
- Proof: Assume positive signal densities at $x > y$.
- Be careful about dummy variables of integration!
- Inequality: $f(x|s)/f(x|\theta) \geq f(y|s)/f(y|\theta)$, if $\theta < \theta_2 < s$

$$
\begin{aligned}
\frac{G(\theta_2|x)}{1 - G(\theta_2|x)} &= \frac{\int_{-\infty}^{\theta_2} f(x|s)dG(s)}{\int_{\theta_2}^{\infty} f(x|s)dG(s)} \\
&= \int_{-\infty}^{\theta_2} \frac{1}{\int_{\theta_2}^{\infty} [f(x|s)/f(x|\theta)]dG(s)} dG(\theta) \\
&\leq \int_{-\infty}^{\theta_2} \frac{1}{\int_{\theta_2}^{\infty} f(y|s)/f(y|\theta)dG(s)} dG(\theta) \\
&= \frac{G(\theta_2|y)}{1 - G(\theta_2|y)}
\end{aligned}
$$

Application: Contract Theory to Moral Hazard

- *Principal-Agent Problem (Holmstrom, 1979)*
- Agent expends effort θ, influencing stochastic profit π
- Profit π has density $f(\pi|\theta)$ given effort θ
- Agent's payoff to wealth w is $U(w) - \theta$, where $U' > 0 > U''$
- Principal has utility $V(\cdot)$, where $V' > 0 \geq V''$
- Principal and Agent are weakly/strictly risk-averse
- Optimal sharing rule: Principal gives the agent a profit share $s(\pi)$, where $\frac{V'(\pi - s(\pi))}{U'(s(\pi))} = b + c\frac{f_\theta(\pi|\theta)}{f(\pi|\theta)}$ $(c > 0)$
    - I won't prove this, but it solves the principal's optimization subject to agent obeying his IC constraints
- Proof that sharing rule rises when $f(\pi|\theta)$ has the MLRP:

$$\frac{f(\pi|\theta_2)}{f(\pi|\theta_1)} = \exp\left\{\int_{\theta_1}^{\theta_2} \frac{\partial}{\partial\theta}[\log f(\pi|\theta)]d\theta\right\} = \exp\left\{\int_{\theta_1}^{\theta_2} \frac{f_\theta(\pi|\theta)}{f(\pi|\theta)}d\theta\right\}$$

- Intuition: Profits are a good signal of effort, so that $1 \geq s'(\pi) > 0$, if $\frac{f_\theta(\pi|\theta)}{f(\pi|\theta)}$ increases in π (the MLRP)

Logsupermodularity (LSPM)

- If $f > 0$, then f *logsupermodular* iff $\log f$ supermodular

$$f(x_1|\theta_1)f(x_2|\theta_2) \geq f(x_1|\theta_2)f(x_2|\theta_1) \quad \Leftrightarrow$$

$$\log f(x_1|\theta_1) + \log f(x_2|\theta_2) \geq \log f(x_1|\theta_2) + \log f(x_2|\theta_1)$$

- Auction theorists call f *affiliated* iff f is logsupermodular
- Logsupermodularity on a lattice is defined without logs:

$$f(x)f(y) \leq f(x \vee y)f(x \wedge y)$$

- Multiplication preserves LSPM: f, g LSPM $\Rightarrow$ fg LSPM
- Addition need not preserve LSPM!
- An indicator function: $f(x, y) = 1$ if $x \geq y$ and 0 if $x < y$
- Prove that $\boxed{\text{indicator functions are LSPM}}$.

Signaling is Transitive

- Say X is signals Y if $f(x|y)$ is LSPM
- If X is signals Y (*unobserved*) and Y signals Z, does X signal Z?

Theorem

If $f(x, y)$ and $g(x, z)$ are LSPM, then so is $h(x, z) \equiv \int f(x, y)g(y, z)d\mu(y)$, for any positive measure μ.

- *LSPM is preserved under partial integration*
- Is this surprise? For addition does not preserve LSPM
- Milgrom and Weber (1982), "A Theory of Auctions and Competitive Bidding" (2020 Nobel Prize) repeatedly uses this property

Logsupermodularity, & the Preservation Lemma

- Ahlswede and Daykin (1979) proved the next result.
  - Karlin and Rinott (1980) wonderfully proved it

Lemma (Preservation)

Let $f_1, f_2, f_3, f_4 \geq 0$ on $\mathbb{R}^n$. Then

$$\boxed{PREMISE} \quad f_1(s)f_2(s') \leq f_3(s \vee s')f_4(s \wedge s') \quad \forall s, s' \in \mathbb{R}^n$$

$$\Longrightarrow$$

$$\int f_1(s)d\mu(s) \int f_2(s)d\mu(s) \leq \int f_3(s)d\mu(s) \int f_4(s)d\mu(s) \tag{1}$$

Preservation Lemma Proof (Easy Part)

- Use induction on the dimensionality of $\mathbb{R}^n$!
- We prove $n = 1$ case. The PREMISE with $s' = s$ gives:

$$f_1(s)f_2(s) \leq f_3(s)f_4(s) \tag{2}$$

- Since $\int f_i(s)ds \int f_j(s)ds = \int \int f_i(x)f_j(y)dxdy$, we need

$$\iint\limits_{x<y} [f_1(x)f_2(y) + f_1(y)f_2(x)]\, dxdy \leq \iint\limits_{x<y} [f_3(x)f_4(y) + f_3(y)f_4(x)]\, dxdy$$

- $a = f_1(x)f_2(y)$, $b = f_1(y)f_2(x)$, $c = f_3(x)f_4(y)$, $d = f_3(y)f_4(x)$
- It suffices to show that $a + b \leq c + d$.
- Claim 1: $d \geq a, b$.
  - Proof: $d = f_3(y)f_4(x) = f_3(x \vee y)f_4(x \wedge y)$ since $x < y$
  - PREMISE: $d \geq a$ by $s = x, s' = y$ & $d \geq b$ by $s = y, s' = x$.
- Claim 2: $ab \leq cd$.
  - Proof: multiply (2) at $s = x$ and $s = y$
    $\Rightarrow (c + d) - (a + b) = [(d - a)(d - b) + (cd - ab)]\,/d \geq 0$

Partial Integration preserves LSPM

Theorem

$g(y, s)$ *LSPM* $\Rightarrow \int g(y, s)d\mu(s)$ *LSPM in* y

Proof:

- $f_1(s) = g(y, s)$, $f_2(s) = g(y', s)$, $f_3(s) = g(y \vee y', s)$, $f_4(s) = g(y \wedge y', s)$.
- Since g is LSPM, $f_1(s)f_2(s') \leq f_3(s \vee s')f_4(s \wedge s')$
- By the Preservation Lemma,

$$\int f_1(s)d\mu(s) \int f_2(s)d\mu(s) \leq \int f_3(s)d\mu(s) \int f_4(s)d\mu(s)$$

- Unwrapping this, we get the desired inequality:

$$\int g(y, s)d\mu(s) \int g(y', s)d\mu(s) \leq \int g(y \vee y', s)d\mu(s) \int g(y \wedge y', s)d\mu(s)$$

Measure Inherits LSPM from Density

- $A \vee B \equiv \cup\{a \vee b,\ \ a \in A, b \in B\}$
- $A \wedge B \equiv \cup\{a \wedge b,\ \ a \in A, b \in B\}$
- Probability measure generated by f is $P(A) = \int_A f(s)ds$
- P is LSPM if $P(A \vee B)P(A \wedge B) \geq P(A)P(B)$
- ***Theorem***: *If f is LSPM, then so is $P(A) = \int_A f(s)ds$*
- Proof: Let $f_1 = \mathbb{I}_A(x)$, $f_2 = \mathbb{I}_B(y)$, $f_3 = \mathbb{I}_{A \vee B}$, $f_4 = \mathbb{I}_{A \wedge B}$.
- Condition (1) holds, since

$$\mathbb{I}_A = 1, \mathbb{I}_B = 1 \Rightarrow \mathbb{I}_{A \vee B} = 1 \text{ and } \mathbb{I}_{A \wedge B} = 1$$

- But $\mathbb{I}_A(x) = 1 \Leftrightarrow x \in A$, and $\mathbb{I}_B(y) = 1 \Leftrightarrow y \in B$.
- But $x \in A$ and $y \in B \Rightarrow x \vee y \in A \vee B$, and $x \wedge y \in A \wedge B$.
- Set $f_1^* = f_1 f$, $f_2^* = f_2 f$, $f_3^* = f_3 f$, $f_4^* = f_4$.
- These obey the premise too! So

$$P(A) = \int_A f_1^* \ \& \ P(B) = \int_B f_2^* \ \& \ P(A \vee B) = \int_{A \vee B} f_3^* \ \& \ P(A \wedge B) = \int_{A \wedge B} f_4^*$$

Logconcavity and LSPM

- $f > 0$ is *log concave* when $f((1-\lambda)x + \lambda y) \geq f(x)^{1-\lambda} f(y)^{\lambda}$
- When $f > 0$ is C^2 on $\mathbb{R}$ when $\log f$ is concave, i.e.

$$(\log f)'' \leq 0 \Leftrightarrow \left(f'/f\right)' \leq 0 \Leftrightarrow ff'' \leq (f')^2$$

Lemma (Logconcavity and LSPM)

If $f > 0$ is log concave then $u(x, y) = f(y - x)$ is LSPM.

- Proof: $u_x = -f'(y-x)$, $u_{xy} = -f''(y-x)$, $u_y = f'(y-x)$.

$$u \text{ LSPM } \Leftrightarrow uu_{xy} \geq u_x u_y \Leftrightarrow -ff'' \geq -(f')^2 \Leftrightarrow f \text{ logconcave}$$

Logconcavity and Prekopa's Theorem (1973)

Theorem (Prekopa)

Let $H(x, y) : \mathbb{R}^{m+n} \to \mathbb{R}$ be log-concave:

$$H\left((1-\lambda)(x_1, y_1) + \lambda(x_2, y_2)\right) \geq H(x_1, y_1)^{1-\lambda} H(x_2, y_2)^{\lambda}$$

for $x_1, x_2 \in \mathbb{R}^m$ and $y_1, y_2 \in \mathbb{R}^n$ and $0 < \lambda < 1$.

Then its marginal $M(y) = \int_{\mathbb{R}^m} H(x, y)\, dx$ is log-concave.

- ***Convolution Corollary***: *If f, g are logconcave on $\mathbb{R}$ then $h(x) \equiv \int_{-\infty}^{\infty} g(x-y) f(y) dy$ is logconcave.*
- Also, the cdf or survivor of a log-concave density is log-concave because the step function is log-concave: $g(x) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{if } x > y \end{cases}$
  1. Normal density: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$ on $\mathbb{R}$
  2. Gamma density: $f(x) = \frac{\lambda^r x^{r-1}}{\Gamma(r)} e^{-\lambda x}$ is logconcave on $\mathbb{R}_+$ iff $r \geq 1$
  3. Beta density: $f(x) \propto x^{a-1}(1-x)^{b-1}$ is logconcave on $[0, 1]$ if $a, b \geq 1$, as with the uniform density

Logconcavity and Truncated Means

- Heckman and Honore (1990), Proposition 1
- Let $\underline{f}_0$ be a density, and $\underline{f}_{k+1}(z) \equiv \int_{-\infty}^{z} \underline{f}_k(x)dx$
- Left mean: $\underline{m}(z) = E[X|X \leq z] = \int_{-\infty}^{z} xf_0(x)dx/\underline{f}_1(z)$
- ***Proposition.*** *$\underline{m}'(z) \leq 1$ iff $\underline{f}_2$ is log-concave.*

$$
\begin{aligned}
\text{Proof: } \underline{m}'(z) &= \frac{\underline{f}_1(z)zf_0(z) - \underline{f}_1{}'(z)\int_{-\infty}^{z} xf_0(x)dx}{\left(\underline{f}_1(z)\right)^2} \\
&= \frac{\left(\int_{-\infty}^{z} f_0(x)dx\right) zf_0(z) - f_0(z)\int_{-\infty}^{z} xf_0(x)dx}{\left(\underline{f}_1(z)\right)^2} \\
&= f_0(z)\int_{-\infty}^{z}(z-x)f_0(x)dx/\left(\underline{f}_1(z)\right)^2 \\
&\equiv \underline{f}_2{}''(z)\underline{f}_2(z)/\left(\underline{f}_2{}'(z)\right)^2
\end{aligned}
$$

- This is ≤ 1 iff $\underline{f}_2{}''(z)\underline{f}_2(z) \leq \left(\underline{f}_2(z)\right)^2$, i.e. $\underline{f}_2$ is log-concave

Logconcavity and Truncated Expectations

- Logconcavity is often met and precludes jumps in expectations
- Left variance $\underline{V}(z) = Var(X|X \leq z) \Rightarrow \underline{V}'(z) \leq 1\ \forall z$ iff $\underline{f}_3$ log-concave
- Right mean $\overline{m}(z) = \mathbb{E}[X|X \geq z] \Rightarrow \overline{m}'(z) \leq 1\ \forall z$ iff $\bar{f}_2$ log-concave.
- Right variance $\overline{V}(z) = Var(X|X \leq z) \Rightarrow \overline{V}'(z) \leq 1\ \forall z$ iff $\bar{f}_3$ log-concave
- HW: Prove these results from Prekopa's Theorem, using the fact that a suitable indicator function $\mathbb{I}_B$ on a suitable set B is logconcave.

Total Positivity (Karlin, 1968)

- $u : A \times B \to \mathbb{R}$ is *totally positive of order* k (TP_k, and STP_k if strict) if $\forall m = 1, \ldots, k$ and $x_1 < \cdots < x_m$ in $A \subseteq \mathbb{R}$ and $y_1 < \cdots < y_m$ in $B \subseteq \mathbb{R}$ ($\Longleftarrow$ scalar variables only!)

$$\det \begin{bmatrix} u(x_1, y_1) & \cdots & u(x_1, y_m) \\ \vdots & & \vdots \\ u(x_m, y_1) & \cdots & u(x_m, y_m) \end{bmatrix} \geq 0$$

- TP_1 means nonnegative, and TP_2 is LSPM on $\mathbb{R}^2$
- Easily, $TP_k \Rightarrow TP_{k'}$ $\forall k' \leq k$.
- $u(x, y)$ is TP (or *totally positive*) if it is TP_k $\forall k < \infty$.
- *Lemma: If* $v, w \geq 0$ *on* A *and* B*, and* $u(x, y)$ *is* TP_k*, then* $v(x)w(y)u(x, y)$ *is* TP_k *on* $A \times B$.
- *Lemma: If* v *and* w *are comonotone, and* f *is* TP_k *on* $A \times B$*, then* $u(v(x), w(y))$ *is* TP_k *on* $A \times B$.
  1. $u(x, y) = e^{xy}$ is $STP \Rightarrow e^{-(x-y)^2} = e^{-x^2} e^{-y^2} e^{2xy}$ is STP
  2. $u(x, y) = \frac{1}{x+y}$ is STP
  3. $u(x, y) = C(x, y)$ is TP

Variation Diminishing Property (VDP)

- TP preserves monotonicity and convexity.

Theorem (Monotonicity Preservation)

Let $\int f(x,y)d\mu(y) = 1\ \forall x$. If f is TP_2 and $w(y)$ is monotonic, then $u(x) = \int f(x,y)w(y)d\mu(y)$ is co-monotonic with w.

- Applications: When $f(x,y)$ is a probability density over random outcomes y given x
- Proof: w monotonic $\Leftrightarrow w(y) - \alpha$ is upcrossing $\forall \alpha \in \mathbb{R}$
- Since $\int f(x,y)d\mu(y) = 1$, for any $\alpha \in \mathbb{R}$,

$$u(x) - \alpha = \int f(x,y)(w(y) - \alpha)d\mu(y)$$

- If $w(y) - \alpha$ changes sign $-$ to $+$, then so does $u(x) - \alpha$ by Karlin and Rubin (1956) Upcrossing Preservation, since $f(x,y)$ is LSPM. $\square$

Variation Diminishing Property (VDP)

- Let $S(f)$ be the supremum number of sign changes in $f(t_2) - f(t_1), \ldots, f(t_k) - f(t_{k-1})$ across all sets $t_1 < \cdots < t_k$.
- For a function $w(y)$, define $u(x) \equiv \int f(x, y)w(y)d\mu(y)$.

Theorem (Variation Diminishing Property, Karlin (1968))

Let $f(x, y)$ be TP_k. If $S(w) \leq k - 1$, then $S(u) \leq S(w)$, and u and w have the same arrangement of signs (left to right) in the domain.

- Proof: Obvious for $k = 1$; proven already for $k = 2$.
- For $k > 2$, Karlin's proof is a mess. Andrea Wilson's Induction Proof:
  - Induction step: if $\sum_y f(x, y)w(y)$ is n-crossing, initially $+$ to $-$, and f is TP-$(n+1)$, then $w(y)$ is n-crossing with an initial downcrossing on some $Y' \subset Y$.
  - Let $x_1 < \cdots < x_{n+1}$ and $\alpha_1, \ldots, \alpha_n$ with $(-1)^{j+1}\alpha_j > 0$ with

$$\sum_{i=1}^{n} f(x_j, y_i)w(y_i) = \alpha_j \quad \text{for} \quad j = 1, 2, \ldots, n+1$$

  - She uses Cramer's rule: The TP Determinants are key