Dynamic Programming

Chapter 2

Thomas J. Sargent and John Stachurski

2023

Introduction

Summary of this lecture:

- Introduction to partial orders
- Pointwise orders
- Order-preserving maps
- Fixed points and order
- Conjugate maps
- Convergence rates and gradient-based methods

Order

The next few slides give a quick introduction to **order theory**

One of the foundational subjects of maths, on par with

- algebra
- geometry
- topology
- number theory
- set theory

But not commonly taught in foundational math courses

Why?

Rarely used in

- physics
- chemistry
- biology, etc.

Math courses are biased toward these subjects!

But not commonly taught in foundational math courses

Why?

Rarely used in

- physics
- chemistry
- biology, etc.

Math courses are biased toward these subjects!

But **very important for econ and related fields**

Examples.

- Does consumer X <u>prefer</u> good A or good B?
- Is welfare <u>greater</u> under policy A or policy B?
- Does R & D <u>increase</u> profits?
- How can firm Y <u>minimize</u> costs?

For these lectures, we need order for

- studying optimality
- fixed point results

Partial orders

Let P be a nonempty set

A **partial order** on a P is a binary relation $\preceq$ on $P \times P$ satisfying, for any p, q, r in P,

$p \preceq p$,

$p \preceq q$ and $q \preceq p$ implies $p = q$ and

$p \preceq q$ and $q \preceq r$ implies $p \preceq r$

(Reflexivity, antisymmetry, transitivity)

We call $(P, \preceq)$ (or just P) a **partially ordered set**

Ex.

1. Show that the usual order $\leqslant$ on $\mathbb{R}$ is a partial order on $\mathbb{R}$
2. Given set M, show that $\subset$ is a partial order on $\wp(M)$

Proof for 2: Clearly, for all $A, B, C \subset M$,

- $A \subset A$ holds
- $A \subset B$ and $B \subset A$ implies $A = B$
- $A \subset B$ and $B \subset C$ implies $A \subset C$

Ex.

1. Show that the usual order $\leqslant$ on $\mathbb{R}$ is a partial order on $\mathbb{R}$
2. Given set M, show that $\subset$ is a partial order on $\wp(M)$

Proof for 2: Clearly, for all $A, B, C \subset M$,

- $A \subset A$ holds
- $A \subset B$ and $B \subset A$ implies $A = B$
- $A \subset B$ and $B \subset C$ implies $A \subset C$

A partial order $\preceq$ on P is called a **total order** if

$$\text{either } p \preceq q \text{ or } q \preceq p \text{ for all } p, q \in P$$

Example. $\leqslant$ is a total order on $\mathbb{R}$

Ex. Prove: $\subset$ is not a total order on $\wp(M)$ when $|M| > 1$

Proof: If M has more than two elements, then we can take nonempty $A, B \subset M$ with $A \cup B = \varnothing$

But then $A \subset B$ and $B \subset A$ both fail

A partial order $\preceq$ on P is called a **total order** if

$$\text{either } p \preceq q \text{ or } q \preceq p \text{ for all } p, q \in P$$

Example. $\leqslant$ is a total order on $\mathbb{R}$

Ex. Prove: $\subset$ is not a total order on $\wp(M)$ when $|M| > 1$

Proof: If M has more than two elements, then we can take nonempty $A, B \subset M$ with $A \cup B = \varnothing$

But then $A \subset B$ and $B \subset A$ both fail

Pointwise Partial Orders

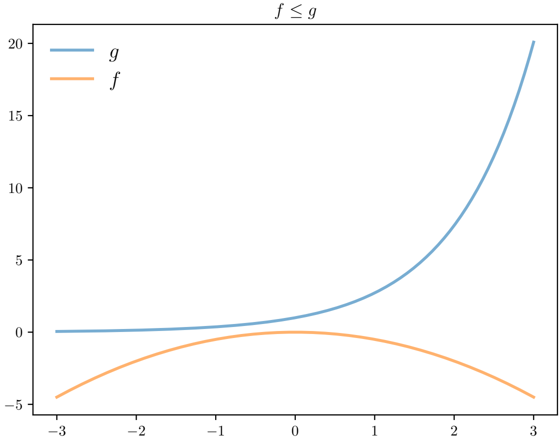
Let

- M be any set and
- let $\mathbb{R}^M$ be all $f \colon M \to \mathbb{R}$

The **pointwise partial order** over $\mathbb{R}^M$ is writen as $\leqslant$ and defined as follows:

- Given f, g in $\mathbb{R}^M$, we set

$$f \leqslant g \iff f(x) \leqslant g(x) \text{ for all } x \in M$$


$f \leq g$
20
15
10
5
0
−5
−3
−2
−1
0
1
2
3
g
f

Ex. Show $\leqslant$ is a partial order on $\mathbb{R}^M$

Proof:

Let's just check antisymmetry

Fix $f, g \in \mathbb{R}^M$ and suppose $f \leqslant g$ and $g \leqslant f$

Pick any $x \in M$

By definition, $f(x) \leqslant g(x)$ and $g(x) \leqslant f(x)$

Therefore, $f(x) = g(x)$

Since x was arbitrary, we have $f = g$

Ex. Show $\leqslant$ is a partial order on $\mathbb{R}^M$

Proof:

Let's just check antisymmetry

Fix $f, g \in \mathbb{R}^M$ and suppose $f \leqslant g$ and $g \leqslant f$

Pick any $x \in M$

By definition, $f(x) \leqslant g(x)$ and $g(x) \leqslant f(x)$

Therefore, $f(x) = g(x)$

Since x was arbitrary, we have $f = g$

Let's define the **pointwise partial order for matrices**

Let $\mathbb{M}^{n\times k} :=$ all $n \times k$ matrices

For $A = (a_{ij})$ and $B = (b_{ij})$ in $\mathbb{M}^{n\times k}$, we set

$$A \leqslant B \iff a_{ij} \leqslant b_{ij} \text{ for all } i, j$$

Example.

$$\begin{pmatrix} 1 & 2 \\ -2 & 0 \end{pmatrix} \leqslant \begin{pmatrix} 10 & 20 \\ 0 & 10 \end{pmatrix}$$

Ex. Show that $\leqslant$ is a partial order on $\mathbb{M}^{n\times k}$

Special case: **pointwise order for vectors**

Recall $[n] := \{1, \ldots, n\}$

For $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, we write

$$x \leqslant y \quad \iff \quad x_i \leqslant y_i \text{ for all } i \in [n]$$

Pointwise partial order $\leqslant$ on $\mathbb{R}^2$:

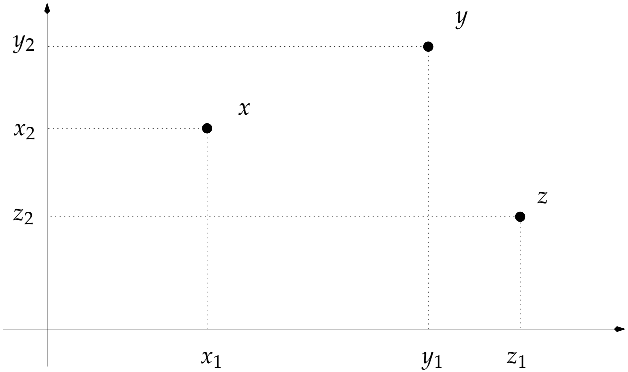


Figure: $x \leqslant y$ but neither x, z nor y, z are comparable

Ex. Prove: for $a, b \in \mathbb{R}^n$ and sequence (x_k) in $\mathbb{R}^n$, we have

$$a \leqslant x_k \leqslant b \text{ for all } k \in \mathbb{N} \text{ and } x_k \to x \text{ implies } a \leqslant x \leqslant b$$

Proof: Fix $i \in [n]$

Let a^i be the i-th element of a, etc.

It suffices to show that

$$a^i \leqslant x^i \leqslant b^i \tag{1}$$

Note $x_k \to x$ implies $x_k^i \to x^i$

Moreover, $a^i \leqslant x_k^i \leqslant b^i$ for all k

Weak inequalities in $\mathbb{R}$ are preserved under limits, so (1) holds

Ex. Prove: for $a, b \in \mathbb{R}^n$ and sequence (x_k) in $\mathbb{R}^n$, we have

$$a \leqslant x_k \leqslant b \text{ for all } k \in \mathbb{N} \text{ and } x_k \to x \text{ implies } a \leqslant x \leqslant b$$

<u>Proof</u>: Fix $i \in [n]$

Let a^i be the i-th element of a, etc.

It suffices to show that

$$a^i \leqslant x^i \leqslant b^i \tag{1}$$

Note $x_k \to x$ implies $x_k^i \to x^i$

Moreover, $a^i \leqslant x_k^i \leqslant b^i$ for all k

Weak inequalities in $\mathbb{R}$ are preserved under limits, so (1) holds

In other words, the pointwise partial order $\leqslant$ is preserved under limits

As a result, these sets are **closed**

- $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : 0 \leqslant x\}$
- $[a, b] := \{x \in \mathbb{R}^n : a \leqslant x \leqslant b\}$
- etc.

A key connection between order and topology!

Ex. Prove: If B is $m \times k$ and $B \geqslant 0$, then

$$|Bx| \leqslant B|x| \text{ for all } k \times 1 \text{ column vectors } x$$

<u>Proof</u>: Fix $B \in \mathbb{M}^{m \times k}$ with $b_{ij} \geqslant 0$ for all i, j

Fix $i \in [m]$ and $x \in \mathbb{R}^k$

By the triangle inequality, we have $|\sum_j b_{ij} x_j| \leqslant \sum_j b_{ij} |x_j|$

Stacking these inequalities yields

$$|Bx| \leqslant B|x|$$

Ex. Prove: If B is $m \times k$ and $B \geqslant 0$, then

$$|Bx| \leqslant B|x| \text{ for all } k \times 1 \text{ column vectors } x$$

Proof: Fix $B \in \mathbb{M}^{m\times k}$ with $b_{ij} \geqslant 0$ for all i, j

Fix $i \in [m]$ and $x \in \mathbb{R}^k$

By the triangle inequality, we have $|\sum_j b_{ij}x_j| \leqslant \sum_j b_{ij}|x_j|$

Stacking these inequalities yields

$$|Bx| \leqslant B|x|$$

Lemma. Given a finite set M and f, g in $\mathbb{R}^M$, we have

$$\left|\max_{x\in M} f(x) - \max_{x\in M} g(x)\right| \leqslant \max_{x\in M} |f(x) - g(x)|$$

<u>Proof</u>: Fixing $f, g \in \mathbb{R}^M$, we have

$$f = f - g + g \leqslant |f - g| + g \qquad \text{(pointwise)}$$

$$\therefore \quad \max f \leqslant \max(|f - g| + g) \leqslant \max |f - g| + \max g$$

$$\therefore \quad \max f - \max g \leqslant \max |f - g|$$

Reversing the roles of f and g proves the claim

Order-preserving maps

Let

- $(P, \preceq)$ and $(Q, \trianglelefteq)$ be partially ordered sets
- $T \colon P \to Q$

T is called **order-preserving** if, for all $x, y \in P$,

$$x \preceq y \implies Tx \trianglelefteq Ty$$

- Meaning: If x goes up then Tx goes up
- Very important concept for dynamic programming

Example. Let $(P, \preceq) = (\mathcal{C}, \leqslant)$ where

- $\mathcal{C}$ is all continuous functions from $[a, b]$ to $\mathbb{R}$
- $\leqslant$ is the pointwise partial order

If $I \colon \mathcal{C} \to \mathbb{R}$ is defined by

$$Ig := \int_a^b g(x)dx \qquad (g \in \mathcal{C})$$

then I is order-preserving on $\mathcal{C}$

(Larger functions have larger integrals)

Example. Let $\leqslant$ denote the pointwise partial order on $\mathbb{R}^n$

Let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be defined by $Tx = Ax + b$

If $A \geqslant 0$, then T is order preserving on $\mathbb{R}^n$

Proof: Fix $x \leqslant y$

Then $0 \leqslant y - x$

$$\therefore \quad 0 \leqslant A(y-x) \leqslant Ay - Ax$$

$$\therefore \quad Ax \leqslant Ay$$

$$\therefore \quad Tx \leqslant Ty$$

Example. Let $\leqslant$ denote the pointwise partial order on $\mathbb{R}^n$

Let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be defined by $Tx = Ax + b$

If $A \geqslant 0$, then T is order preserving on $\mathbb{R}^n$

Proof: Fix $x \leqslant y$

Then $0 \leqslant y - x$

$$\therefore \quad 0 \leqslant A(y - x) \leqslant Ay - Ax$$

$$\therefore \quad Ax \leqslant Ay$$

$$\therefore \quad Tx \leqslant Ty$$

Special Case: Real-Valued Functions

Special case: maps from $(P, \preceq)$ into $(\mathbb{R}, \leqslant)$

Then "order-preserving" $=$ "increasing"

In particular, we also call $h \in \mathbb{R}^P$

- **increasing** if $x \preceq y$ implies $h(x) \leqslant h(y)$ and
- **decreasing** if $x \preceq y$ implies $h(x) \geqslant h(y)$

Let P be partially ordered by $\preceq$

We write $i\mathbb{R}^P$ **for the increasing functions in $\mathbb{R}^P$**

Thus,

$$h \in i\mathbb{R}^P \quad \iff \quad x, y \in P \text{ and } x \preceq y \text{ implies } h(x) \leqslant h(y)$$

Example. Let $P = \{1, \ldots, n\}$ and let $\preceq$ be the usual order $\leqslant$ on $\mathbb{R}$

Then

- $x \mapsto 2x$ and $x \mapsto \mathbb{1}\{2 \leqslant x\}$ are in $i\mathbb{R}^P$
- $x \mapsto -x$ and $x \mapsto \mathbb{1}\{x \leqslant 2\}$ are not

Ex. Prove the following:

If $f, g \in i\mathbb{R}^P$, then

- $\alpha f + \beta g \in i\mathbb{R}^P$ when $\alpha, \beta \geqslant 0$
- $f \vee g \in i\mathbb{R}^P$
- $f \wedge g \in i\mathbb{R}^P$

Ex. Given finite P, show that $i\mathbb{R}^P$ is closed in $\mathbb{R}^P$

Proof: Take $(f_k)_{k \geqslant 1}$ in $i\mathbb{R}^P$ and $f \in \mathbb{R}^P$ with $f_k \to f$

Since $f_k \to f$ we have $f_k(z) \to f(z)$ for all $z \in P$

- norm convergence implies pointwise convergence

Fix $x, y \in P$ with $x \preceq y$

From $(f_k) \subset i\mathbb{R}^P$ we have $f_k(x) \leqslant f_k(y)$ for all k

Since weak inequalities are preserved under limits, $f(x) \leqslant f(y)$

Hence $f \in i\mathbb{R}^P$

Strict inequalities

We write

- $f \ll g$ if $f(x) < g(x)$ for all $x \in$ some given set M
- $x \ll y$ if $x_i < y_i$ for all $i \in [n]$
- $A \ll B$ if $a_{ij} < b_{ij}$ for all i, j

These are not partial orders

Ex. Why is $f \ll g$ not a partial order on $\mathbb{R}^M$?

Parametric Monotonicity

Let $(P, \preceq)$ be a partially ordered set

Given two self-maps S and T on P, we set

$$S \preceq T \iff Sx \preceq Tx \text{ for every } x \in P$$

We say that T **dominates** S on P

Ex. Show that $\preceq$ is a partial order on

$$\mathcal{S}_P := P^P := \text{ set of all self-maps on } P$$

Proof of antisymmetry of $\preceq$ on $\mathcal{S}_P$:

Let $(P, \preceq)$ and $S, T \in \mathcal{S}_P$ be as defined above

Suppose $S \preceq T$ and $T \preceq S$

Fix any $x \in P$

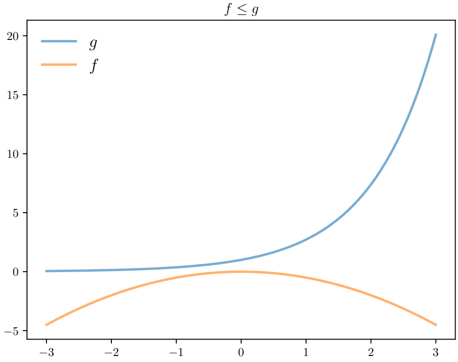
We have $Sx \preceq Tx$ and $Tx \preceq Sx$

Since $\preceq$ is antisymmetric on P, we have $Sx = Tx$

Since p was arbitrary, $S = T$

Hence $\preceq$ is antisymmetric on $\mathcal{S}_P$

Example. If $(\preceq, P) = (\leqslant, \mathbb{R})$, then $\leqslant$ is the pointwise partial order over functions



Example. Consider $\mathbb{R}^n_+$ with the pointwise partial order $\leqslant$

- Called the **positive cone** in $\mathbb{R}^n$

Let

- $Sx = Ax + b$
- $Tx = Bx + b$

Ex. Show that $0 \leqslant A \leqslant B \implies T$ dominates S on $\mathbb{R}^n_+$

Proof: Fixing $x \in \mathbb{R}^n_+$, suffices to show that $Sx \leqslant Tx$

Since $A \leqslant B$ and $x \geqslant 0$, we have $Ax \leqslant Bx$

Hence $Sx \leqslant Tx$

Example. Consider $\mathbb{R}^n_+$ with the pointwise partial order $\leqslant$

- Called the **positive cone** in $\mathbb{R}^n$

Let

- $Sx = Ax + b$
- $Tx = Bx + b$

Ex. Show that $0 \leqslant A \leqslant B \implies T$ dominates S on $\mathbb{R}^n_+$

Proof: Fixing $x \in \mathbb{R}^n_+$, suffices to show that $Sx \leqslant Tx$

Since $A \leqslant B$ and $x \geqslant 0$, we have $Ax \leqslant Bx$

Hence $Sx \leqslant Tx$

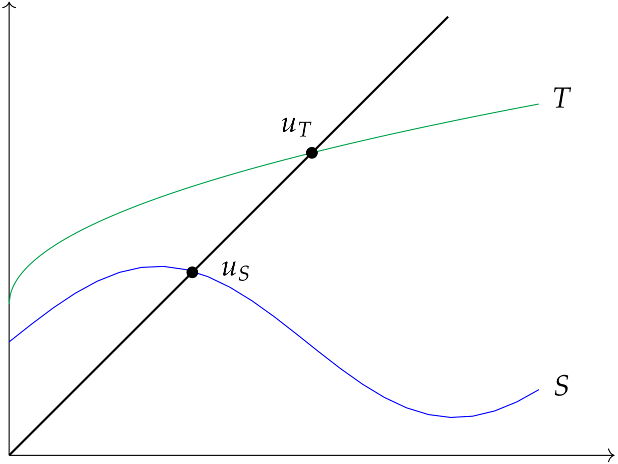
Conjecture: If $S \leqslant T$, then the fixed points of T will be larger

This is not true in general...

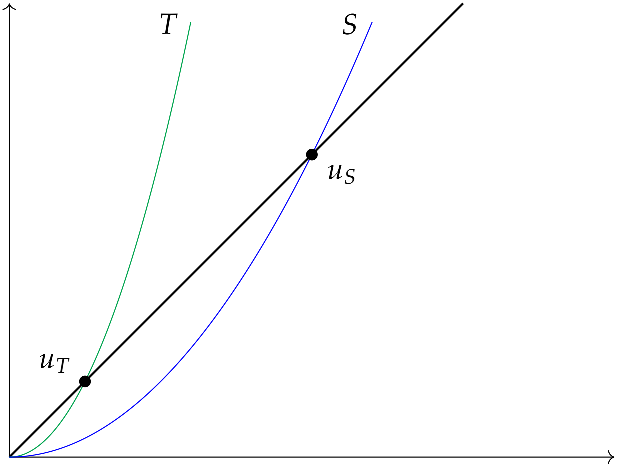
Conjecture: If $S \leqslant T$, then the fixed points of T will be larger

This is <u>not</u> true in general...

Sometimes true:



And sometimes false:



One difference: in the first case, T is globally stable

This leads us to our next result

Proposition. Let

- S and T be self-maps on $M \subset \mathbb{R}^n$
- $\leqslant$ be the pointwise partial order on M

If

1. T dominates S on M and
2. T is order-preserving and globally stable on M,

then the unique fixed point of T dominates any fixed point of S

Proof: Assume the conditions

Let

- u_T be the unique fixed point of T and
- u_S be any fixed point of S

Since $S \leqslant T$, we have $u_S = Su_S \leqslant Tu_S$

Applying T to both sides of $u_S \leqslant Tu_S$ gives

$$u_S \leqslant Tu_S \leqslant T^2 u_S$$

Continuing in this fashion yields $u_S \leqslant T^k u_S$ for all $k \in \mathbb{N}$

Since $\leqslant$ is preserved under limits and T is globally stable,

$$u_S \leqslant \lim_k T^k u_S = u_T$$

Example. Recall that, in the job search model,

$$h^* = c + \beta \sum_{w'} \max\left\{\frac{w'}{1-\beta}, h^*\right\} \varphi(w')$$

We found h^* as the fixed point of $g \colon \mathbb{R}_+ \to \mathbb{R}_+$ defined by

$$g(h) = c + \beta \sum_{w'} \max\left\{\frac{w'}{1-\beta}, h\right\} \varphi(w')$$

In the exercise, you showed that g is a contraction map on $\mathbb{R}_+$

Ex. Prove that the optimal continuation value h^* is increasing in β

Proof: Fix $\beta_1 \leqslant \beta_2$ and let

- $h_i^* :=$ fixed point corresponding to β_i
- $g_i :=$ fixed point map corresponding to β_i

Since $\beta_1 \leqslant \beta_2$, we have $g_1(h) \leqslant g_2(h)$ for all $h \in \mathbb{R}_+$

In addition, g_2 is

1. a contraction (so globally stable) and
2. increasing (order-preserving)

Hence $h_1^* \leqslant h_2^*$

Ex. Prove that the optimal continuation value h^* is increasing in β

Proof: Fix $\beta_1 \leqslant \beta_2$ and let

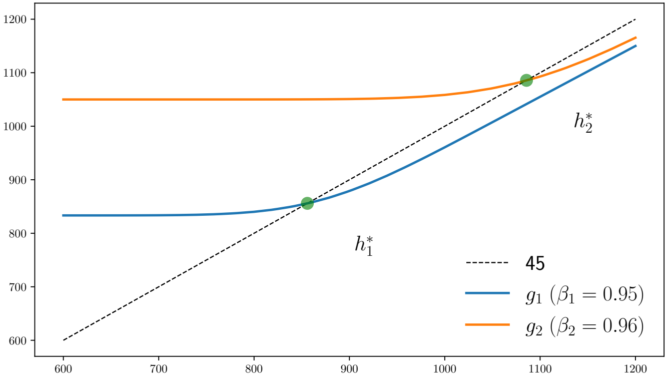
- $h_i^* :=$ fixed point corresponding to β_i
- $g_i :=$ fixed point map corresponding to β_i

Since $\beta_1 \leqslant \beta_2$, we have $g_1(h) \leqslant g_2(h)$ for all $h \in \mathbb{R}_+$

In addition, g_2 is

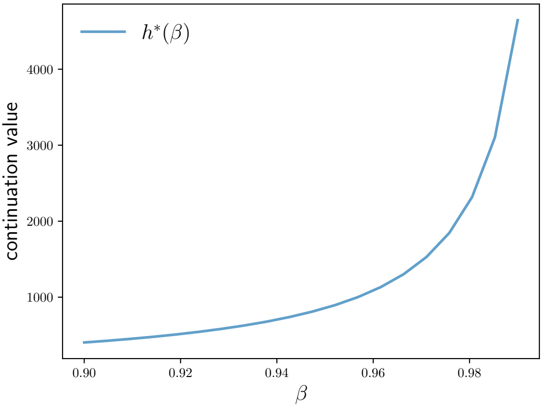
1. a contraction (so globally stable) and
2. increasing (order-preserving)

Hence $h_1^* \leqslant h_2^*$



1200
1100
1000
900
800
700
600
600
700
800
900
1000
1100
1200
h_2^*
h_1^*
45
g_1 $(\beta_1 = 0.95)$
g_2 $(\beta_2 = 0.96)$

Ex. Replicate this figure



(First Order) Stochastic Dominance

In the discussion above we obtained some results from order theory

- Example. parametric monotonicity

To use these results in a stochastic setting, we need to order distributions!

That is, we need a partial order over distributions

The most important of these partial orders is called "first order stochastic dominance"

In this section we define it

To start, let's consider ordering distributions in a special case

Example. The binomial distribution is defined as follows:

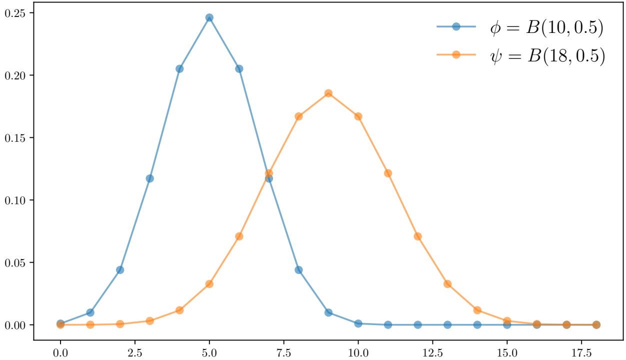
- $X \sim B(n, 0.5)$
- X counts the # of heads in n flips of a fair coin

Suppose $\varphi \overset{d}{=} X \sim B(10, 0.5)$ and $\psi \overset{d}{=} Y \sim B(18, 0.5)$

- Y counts over more flips, so it should be "larger" in some sense

Hence we expect that φ is "$\preceq$" ψ in some sense

Distribution ψ seems "larger than" φ — more mass on higher draws



But how can we make this idea precise?

Let X be a finite set partially ordered by $\preceq$

Fix $\varphi, \psi \in \mathcal{D}(\mathsf{X})$

Write $\langle u, \varphi \rangle$ for $\sum_x u(x)\varphi(x)$, etc.

We say that ψ **stochastically dominates** φ and write $\varphi \preceq_{\mathrm{F}} \psi$ if

$$u \in i\mathbb{R}^{\mathsf{X}} \quad \Longrightarrow \quad \langle u, \varphi \rangle \leqslant \langle u, \psi \rangle$$

Example. If

- $\varphi \stackrel{d}{=} X \sim B(10, 0.5)$ and
- $\psi \stackrel{d}{=} Y \sim B(18, 0.5)$,

then $\varphi \preceq_{\mathrm{F}} \psi$

Proof: Fix $u \in i\mathbb{R}^{\mathsf{X}}$ and let

- $\mathsf{X} = \{0, \ldots, 18\}$ and
- $W_1, \ldots, W_{18}$ be IID Bernoulli with $\mathbb{P}\{W_i = 1\} = 0.5$ for all i

Then $X := \sum_{i=1}^{10} W_i \stackrel{d}{=} \varphi$ and $Y := \sum_{i=1}^{18} W_i \stackrel{d}{=} \psi$

Clearly $X \leqslant Y$ with probability one (i.e., for any draw of $\{W_i\}_{i=1}^{18}$)

Hence $u(X) \leqslant u(Y)$

Hence $\mathbb{E}u(X) \leqslant \mathbb{E}u(Y)$

In other words,

$$\langle u, \varphi \rangle \leqslant \langle u, \psi \rangle$$

Example. An agent has preferences over outcomes in X

Preferences are determined by a utility function $u \in \mathbb{R}^{\mathsf{X}}$

The agent prefers more to less, so $u \in i\mathbb{R}^{\mathsf{X}}$

Suppose that the agent ranks lotteries over X according to expected utility

- evaluates $\varphi \in \mathcal{D}(\mathsf{X})$ according to $\sum_x u(x)\varphi(x)$

Then the agent (weakly) prefers ψ to φ whenever $\varphi \preceq_{\mathrm{F}} \psi$

Another Perspective

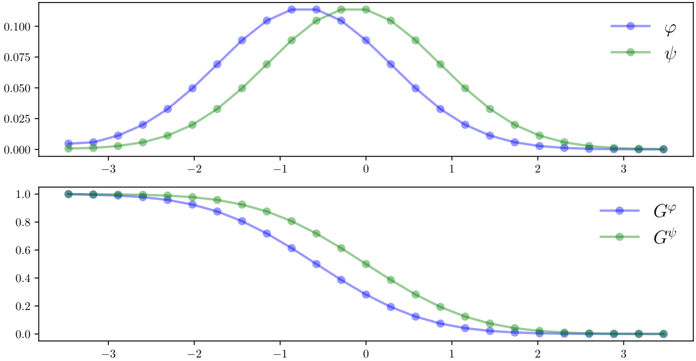
Given $\varphi \in \mathcal{D}(\mathsf{X})$, let

$$G^{\varphi}(y) := \sum_{x \in \mathsf{X}} \mathbb{1}\{y \preceq x\} \varphi(x) \qquad (y \in \mathsf{X})$$

This is the **counter** CDF of φ

Lemma. For each $\varphi, \psi \in \mathcal{D}(\mathsf{X})$, the following statements hold:

1. $\varphi \preceq_{\mathrm{F}} \psi \implies G^{\varphi} \leqslant G^{\psi}$
2. If X is totally ordered by $\preceq$, then $G^{\varphi} \leqslant G^{\psi} \implies \varphi \preceq_{\mathrm{F}} \psi$


0.100
0.075
0.050
0.025
0.000
−3
−2
−1
0
1
2
3
φ
ψ
1.0
0.8
0.6
0.4
0.2
0.0
G^{φ}
G^{ψ}

Lemma. $\preceq_F$ is a partial order on $\mathcal{D}(X)$

Proof:

Let's just prove transitivity

Suppose $f, g, h \in \mathcal{D}(X)$ with $f \preceq_F g$ and $g \preceq_F h$

Fixing $u \in i\mathbb{R}^X$, we have

$$\langle u, f \rangle \leqslant \langle u, g \rangle \quad \text{and} \quad \langle u, g \rangle \leqslant \langle u, h \rangle$$

Hence $\langle u, f \rangle \leqslant \langle u, h \rangle$

Since u was arbitrary in $i\mathbb{R}^X$, we are done