Learning in Continuous Time

Duarte Gonçalves

University College London

Topics in Economic Theory

When Is the Evidence Good Enough?

A consumer is deciding whether a product is right for them.

Read another review? Compare another attribute? Think a little longer?

More evidence may prevent a bad purchase, or a missed opportunity.

Gathering and interpreting it takes time.

What can response time tell us: about the decision problem? about the choices made?

Learning and Stopping

Lecture 1: gamble unfolding over time; stop when option value no compensates its cost.

Lecture 2: strategic interaction over time; identify state, derive generator, solve for stopping boundaries.

Now must first find *observable state of stopping problem*: posterior belief.

Filtering: translate observations into beliefs.

Coupling and time change: compare learning across environments.

Girsanov and Brownian bridges: understand selection into stopping at given time.

Overview

1. A Pioneer from Cognitive Sciences
2. Speed, Accuracy, and Complexity
3. Evidence and Beliefs
4. Exogenous Stopping and Problem Complexity
5. Optimal Stopping
6. Comparative Statics on Incentives and Revealing Complexity
7. Fast and Slow Errors

Overview

1. A Pioneer from Cognitive Sciences
2. Speed, Accuracy, and Complexity
3. Evidence and Beliefs
4. Exogenous Stopping and Problem Complexity
5. Optimal Stopping
6. Comparative Statics on Incentives and Revealing Complexity
7. Fast and Slow Errors

A Pioneer from Cognitive Sciences

Ratcliff (1978 PsyRev): highly influential paper in cognitive sciences, 6k+ cites.

Original Idea: differentiating items is based on noisy comparisons that take time.

Motivation: Juola, Fischler, Wood, & Atkinson (1971).

Participants told to memorise set of 16 words.

Presented with words and asked whether words were in the set.

Speed of decision related to 'discriminability values': ease of recall/distinguish from words in set.

A Benchmark Model of Noisy Learning

Action $\alpha = \pm 1$; unknown state $\theta \pm 1$.

Matching state is better: $u(1,1) > u(-1,1)$ and $u(-1,-1) > u(1,-1)$.

Prior $p_0 := \mathbb{P}(\theta = 1) \in (0,1)$.

A Benchmark Model of Noisy Learning

Action $\alpha = \pm 1$; unknown state $\theta \pm 1$.

Matching state is better: $u(1,1) > u(-1,1)$ and $u(-1,-1) > u(1,-1)$.

Prior $p_0 := \mathbb{P}(\theta = 1) \in (0,1)$.

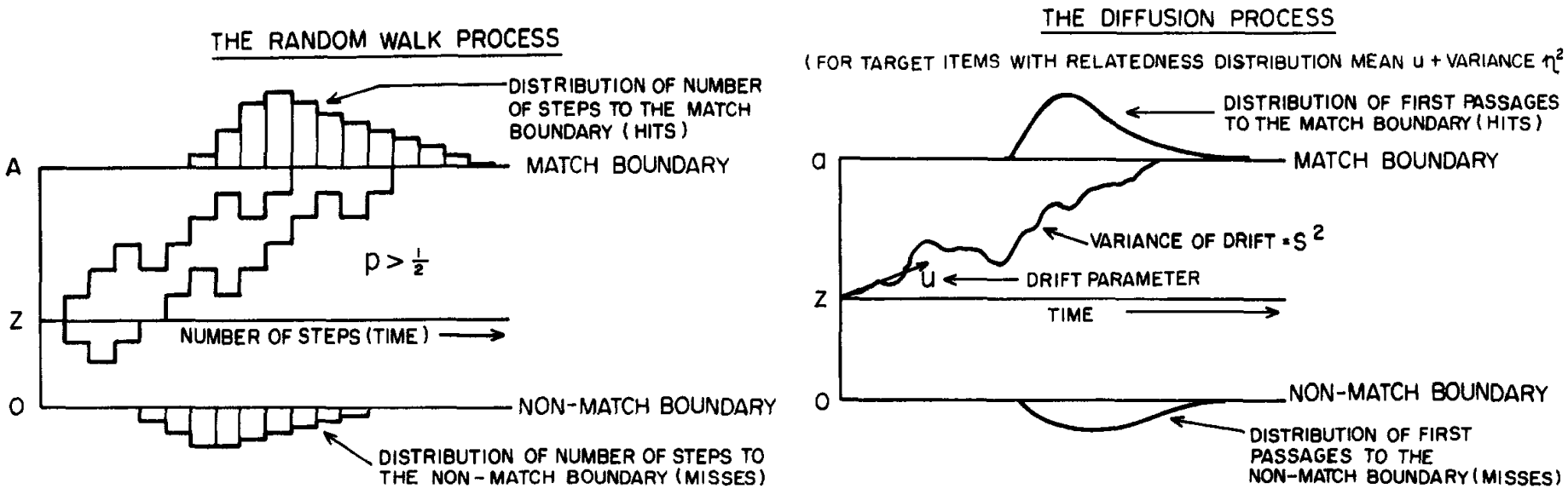
Evidence $X_t := \mu\theta t + \sigma B_t$; $\mu, \sigma > 0$, and B is Brownian motion independent of θ.

Direction is informative about which action is better.

Stopping time τ. Choices upon stopping α_τ.

Rich set of predictions: joint distribution state, actions, stopping time $(\theta, \tau, \alpha_\tau)$.

Noisy Recollection



Recalling/deliberating takes time and is noisy process: model as random walk.

Stop when process hits boundary and choose accordingly.

Continuous-time limit: Brownian motion with drift.

Difficulty and Speed

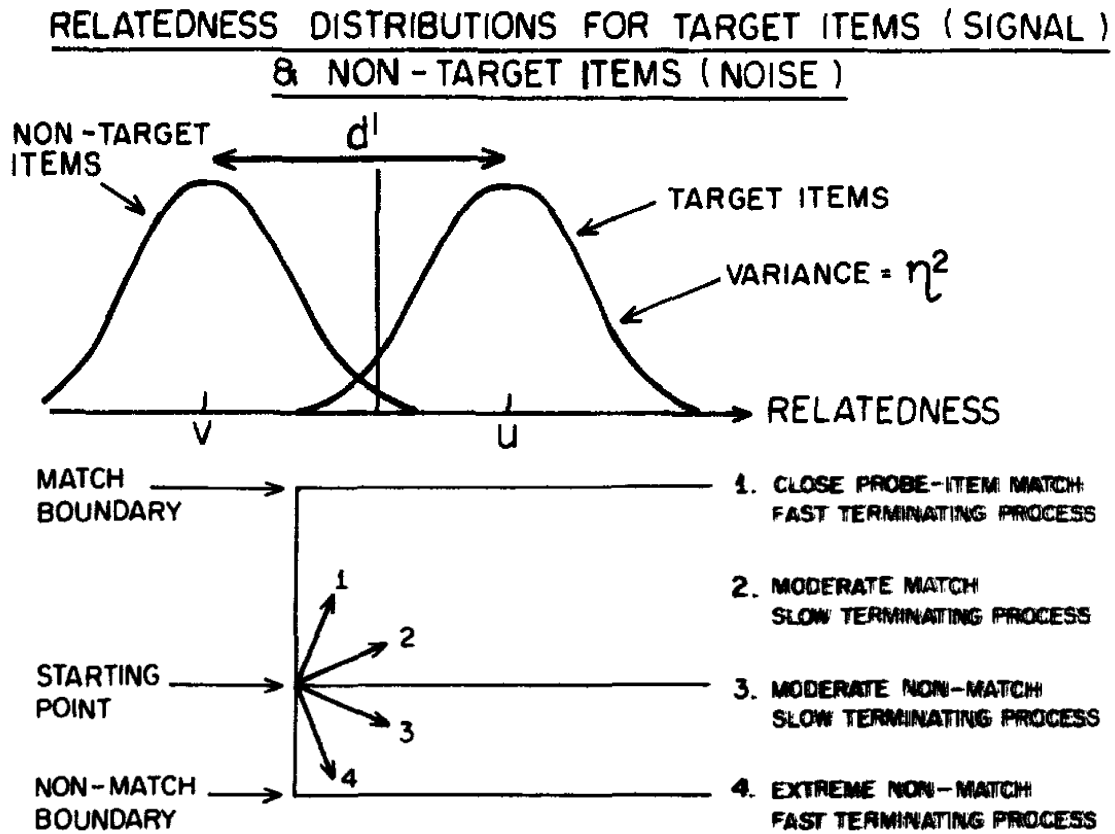


Figure 2. The relation between relatedness and amount of match in the diffusion random walk process. (Relatedness varies from high [Process 1] to low [Process 4].)

Words easier to differentiate/discriminate ($\uparrow d'$) $\Longrightarrow$ faster drift.

Noisy Recollection and Parallel Comparison

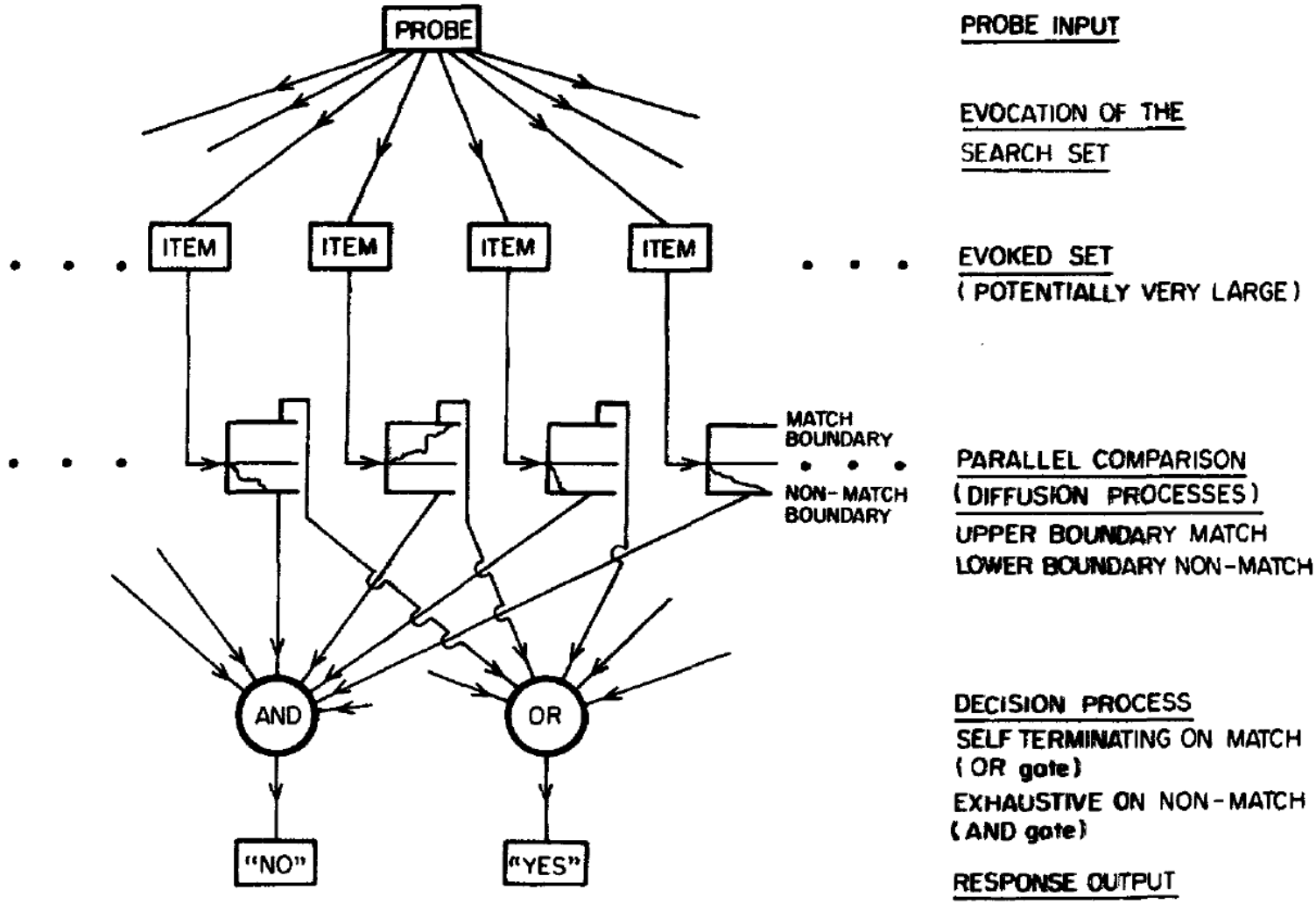
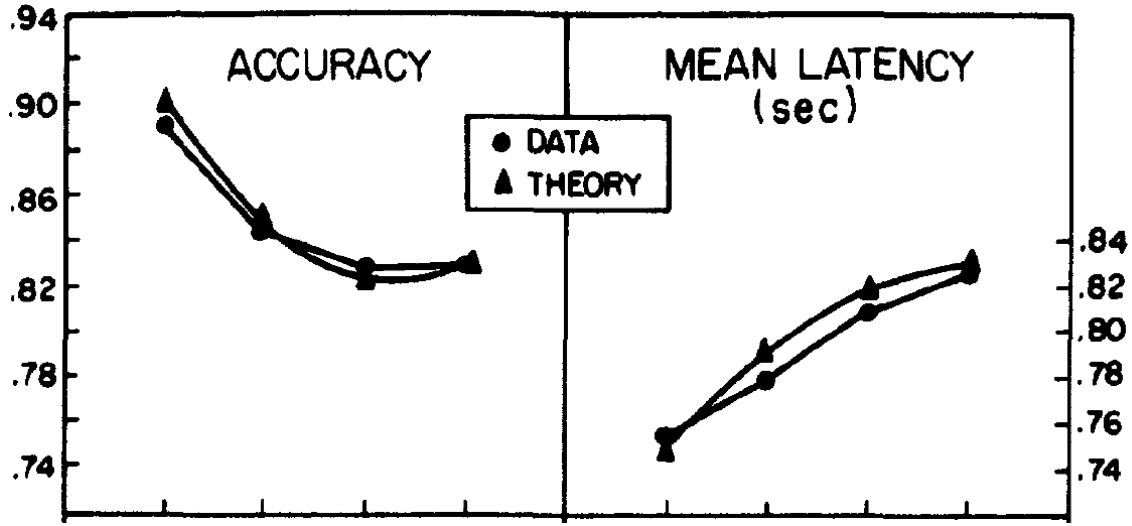


Figure 1. An overview of the item recognition model.

Noisy Recollection and Parallel Comparison



Item recognition

Participants shown 16 words, 1 at a time.

Then presented with 32 words, 1 at a time, for 1 sec; self-paced.

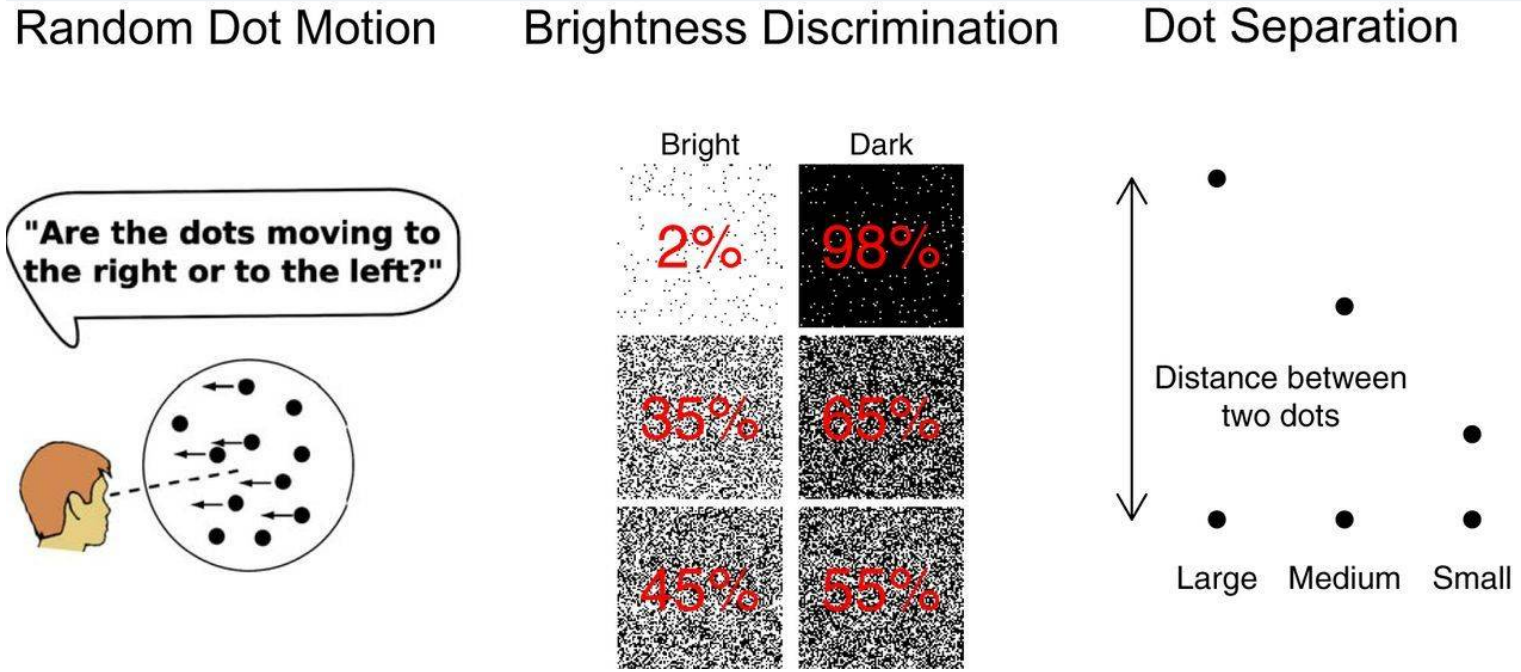
Choice: word in/not in starting 16 word set.

Ticks: serial position 2-8, 9-16, 17-24, 25-32.

Speed-accuracy trade-off: $\uparrow$ time elapsed $\Longrightarrow \downarrow$ accuracy, $\uparrow$ RT.

Also: Three types of 'wrong' answers: neutral, synonyms, and homophones; decisions slower homophones > synonyms > neutral.

A Pioneer from Cognitive Sciences



Psychology/Neuroscience:

Model used in many settings: memory, perception, reasoning.

Use $(\theta, \alpha_\tau, \tau)$ to estimate drift/boundaries. Study correlation with problem (difficulty, time pressure), setting (working memory load, fatigue), participant (IQ, age, clinical conditions), brain activation.

Survey: Forstmann, Ratcliff, & Wagenmakers (2016 Ann. Rev. Psy.).

Textbook: Luce (1986).

Different Approaches

Economics: different questions, same underlying idea.

State determining which action is optimal.

Learning is noisy/takes time.

DM decides when to stop learning and take an action.

Examples:

- consumers evaluating products;
- employers assessing employee productivity;
- enterpreneurs identifying profitability of potential investment;
- policy makers determining which policy to deploy.

Underlying process: learning/processing data, introspecting, researching.

Differences in approach: not fitting model + correlations in data,
but comparative statics from the model.

Overview

1. A Pioneer from Cognitive Sciences
2. Speed, Accuracy, and Complexity
3. Evidence and Beliefs
4. Exogenous Stopping and Problem Complexity
5. Optimal Stopping
6. Comparative Statics on Incentives and Revealing Complexity
7. Fast and Slow Errors

Learning in Continuous time

Gonçalves (WP) Speed, Accuracy, and Complexity (SAC).

DM can learn about θ before taking action α.

Action $\alpha = \pm 1$; unknown state $\theta \pm 1$.

Matching state is better: $u(1,1) > u(-1,1)$ and $u(-1,-1) > u(1,-1)$.

Assume for now: $u(\alpha,\theta) = 1_{\{\alpha=\theta\}}$. Prior $p_0 := \mathbb{P}(\theta > 0) = 1/2$. (Relaxed later.)

Evidence $X_t := m\theta t + \sigma B_t$; $m \geq 0, \sigma > 0$, and B is BM independent of θ, m.

Direction is informative about which action is better.

Observing X_t and X_t/σ is equivalent. Beliefs only depend on m/σ, normalise $\sigma = 1$.

Write instead: $X_t := m\theta t + B_t$.

Posterior $p_t := \mathbb{P}(\theta > 0 \mid \mathcal{F}_t^X)$.

Stopping time τ adapted to $\mathcal{F}_t^X$. Action accuracy $A_\tau := 1_{\{\text{sign } \alpha_\tau = \text{sign } \theta\}}$.

Higher m, higher drift, more information, simpler complex decision problem.

How does accuracy and stopping time vary with problem complexity?

(When) can time be a reliable indicator of problem complexity?

Learning in Continuous time

How does accuracy and stopping time vary with problem complexity?

(When) can time be a reliable indicator of problem complexity?

Plan:

Two cases: (a) m known, (b) m unknown $m \sim \mu$.

1. Law of motion for beliefs.
2. Case (a) Comparative statics with exogenous vs optimal stopping.
3. Case (b) Comparative statics wrt realised m vs wrt μ.

We'll start off with (a).

Overview

1. A Pioneer from Cognitive Sciences

2. Speed, Accuracy, and Complexity

3. Evidence and Beliefs
   - Girsanov's Theorem + Itô's Lemma
   - General Filtering Results
   - Properties of Belief Volatility

4. Exogenous Stopping and Problem Complexity

5. Optimal Stopping

6. Comparative Statics on Incentives and Revealing Complexity

7. Fast and Slow Errors

Evolving Beliefs

First step in learning problems: derive law of motion for posterior beliefs p_t.

Evolving Beliefs

First step in learning problems: derive law of motion for posterior beliefs p_t.

Let $\mathbb{P}_{\pm}(\cdot) := \mathbb{P}(\cdot \mid \theta = \pm 1)$.

1. Apply Bayes's rule: $\dfrac{p_t}{1-p_t} = \dfrac{\mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)}{\mathbb{P}(\theta = -1 \mid \mathcal{F}_t^X)} = \dfrac{p_0}{1-p_0} \dfrac{\mathrm{d}\mathbb{P}_+}{\mathrm{d}\mathbb{P}_-}\Big|_{\mathcal{F}_t^X}$.
2. Derive $\frac{\mathrm{d}\mathbb{P}_+}{\mathrm{d}\mathbb{P}_-}\big|_{\mathcal{F}_t^X}$.
3. Apply Itô's lemma: Obtain law of motion for p_t (will still depend on X).
4. Express law of motion using Brownian motion: Use fact that p_t is $\mathcal{F}^X$-martingale.

Girsanov's Theorem

Theorem (Novikov's Condition)

Let B be a Brownian motion under $\mathbb{P}$, μ_t be progressively measurable, and $M_t := \int_0^t \mu_s \, dB_s$. If $\mathbb{E}e^{\frac{1}{2}[M]_T} = \mathbb{E}e^{\frac{1}{2}\int_0^T \mu_t^2 \, dt} < \infty$, then $Z_t := e^{M_t - \frac{1}{2}[M]_t}$ is a martingale on $[0, T]$.

Girsanov's Theorem

Theorem (Novikov's Condition)

Let B be a Brownian motion under $\mathbb{P}$, μ_t be progressively measurable, and $M_t := \int_0^t \mu_s \, dB_s$. If $\mathbb{E}e^{\frac{1}{2}[M]_T} = \mathbb{E}e^{\frac{1}{2}\int_0^T \mu_t^2 \, dt} < \infty$, then $Z_t := e^{M_t - \frac{1}{2}[M]_t}$ is a martingale on $[0, T]$.

Theorem (Girsanov's Theorem)

Let B be a Brownian motion under $\mathbb{P}$ and μ_t be progressively measurable with respect to $\mathcal{F}$, and let $M_t := \int_0^t \mu_s \, dB_s$. Suppose $Z_t := e^{M_t - \frac{1}{2}[M]_t}$ is a martingale on $[0, T]$. Define $\mathbb{Q}$ such that $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = Z_T$.

Then, for every continuous local martingale N under $\mathbb{P}$, $\widetilde{N}_t = N_t - [N, M]_t$ is a continuous local martingale under $\mathbb{Q}$ on $[0, T]$.

Moreover, $\widetilde{B}_t := B_t - [B, M]_t$ is a $\mathbb{Q}$-Brownian-motion on $[0, T]$, where $[B, M]_t = \int_0^t \mu_s \, ds$.

Applying Girsanov's Theorem

Under $\theta = -1$, $B^- := X_t + mt$ is $\mathbb{P}_-$-Brownian-motion.

Under $\theta = +1$, $B^+ := X_t - mt$ is $\mathbb{P}_+$-Brownian-motion.

Applying Girsanov's Theorem

Under $\theta = -1$, $B^- := X_t + mt$ is $\mathbb{P}_-$-Brownian-motion.

Under $\theta = +1$, $B^+ := X_t - mt$ is $\mathbb{P}_+$-Brownian-motion.

Change drift from $-m$ to $+m$, apply Girsanov with $\mu_t = 2m$ on $[0, T]$ for any $T < \infty$:

Define $M_t = \int_0^t 2m \, dB_s^- = 2mB_t^-$ and $Z_t := e^{2mB_t^- - 2m^2 t}$. Z_t is a $\mathbb{P}_-$-martingale on $[0, T]$.

Let $\mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = Z_T$.

By Girsanov's theorem: $B^- - 2mt = X_t - mt$ is $\mathbb{Q}$-Brownian-motion on $[0, T]$.

Applying Girsanov's Theorem

Under $\theta = -1$, $B^- := X_t + mt$ is $\mathbb{P}_-$-Brownian-motion.

Under $\theta = +1$, $B^+ := X_t - mt$ is $\mathbb{P}_+$-Brownian-motion.

Change drift from $-m$ to $+m$, apply Girsanov with $\mu_t = 2m$ on $[0, T]$ for any $T < \infty$:

Define $M_t = \int_0^t 2m \, dB_s^- = 2mB_t^-$ and $Z_t := e^{2mB_t^- - 2m^2 t}$. Z_t is a $\mathbb{P}_-$-martingale on $[0, T]$.

Let $\mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = Z_T$.

By Girsanov's theorem: $B^- - 2mt = X_t - mt$ is $\mathbb{Q}$-Brownian-motion on $[0, T]$.

Since on $[0, T]$, law of X is the same under $\mathbb{Q}$ and $\mathbb{P}+$, conclude $\mathbb{Q}|_{\mathcal{F}_T^X} = \mathbb{P}+|_{\mathcal{F}_T^X}$.

Applying Girsanov's Theorem

Under $\theta = -1$, $B^- := X_t + mt$ is $\mathbb{P}_-$-Brownian-motion.

Under $\theta = +1$, $B^+ := X_t - mt$ is $\mathbb{P}_+$-Brownian-motion.

Change drift from $-m$ to $+m$, apply Girsanov with $\mu_t = 2m$ on $[0, T]$ for any $T < \infty$:

Define $M_t = \int_0^t 2m \, dB_s^- = 2mB_t^-$ and $Z_t := e^{2mB_t^- - 2m^2 t}$. Z_t is a $\mathbb{P}_-$-martingale on $[0, T]$.

Let $\mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = Z_T$.

By Girsanov's theorem: $B^- - 2mt = X_t - mt$ is $\mathbb{Q}$-Brownian-motion on $[0, T]$.

Since on $[0, T]$, law of X is the same under $\mathbb{Q}$ and $\mathbb{P}+$, conclude $\mathbb{Q}|_{\mathcal{F}_T^X} = \mathbb{P}+|_{\mathcal{F}_T^X}$.

Hence, $\frac{d\mathbb{P}_+}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = \frac{d\mathbb{Q}}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = Z_T = e^{2mB_t^- - 2m^2 t} = e^{2mX_t}$.

Applying Girsanov's Theorem

Under $\theta = -1$, $B^- := X_t + mt$ is $\mathbb{P}_-$-Brownian-motion.

Under $\theta = +1$, $B^+ := X_t - mt$ is $\mathbb{P}_+$-Brownian-motion.

Change drift from $-m$ to $+m$, apply Girsanov with $\mu_t = 2m$ on $[0, T]$ for any $T < \infty$:

Define $M_t = \int_0^t 2m \, dB_s^- = 2mB_t^-$ and $Z_t := e^{2mB_t^- - 2m^2 t}$. Z_t is a $\mathbb{P}_-$-martingale on $[0, T]$.

Let $\mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = Z_T$.

By Girsanov's theorem: $B^- - 2mt = X_t - mt$ is $\mathbb{Q}$-Brownian-motion on $[0, T]$.

Since on $[0, T]$, law of X is the same under $\mathbb{Q}$ and $\mathbb{P}+$, conclude $\mathbb{Q}|_{\mathcal{F}_T^X} = \mathbb{P}_+|_{\mathcal{F}_T^X}$.

Hence, $\frac{d\mathbb{P}_+}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = \frac{d\mathbb{Q}}{d\mathbb{P}_-}\Big|_{\mathcal{F}_T^X} = Z_T = e^{2mB_t^- - 2m^2 t} = e^{2mX_t}$.

Interpretation:

Change of measure: paths receive different probability weights.

Higher drift puts greater weight on paths with higher endpoint.

Here, likelihoods on path space depend only on endpoint.
(More generally, also depend on time.)

(A Confession)

This was a pretext to introduce Girsanov's theorem, which will be used later to 'remove drift' by changing the measure.

In this particular problem, we could have derived $\frac{d\mathbb{P}_+}{d\mathbb{P}_-}\big|_{\mathcal{F}_t^X}$ directly:

Since, conditional on θ, $X_t \sim N(m\theta t, t)$,

$$\frac{p_t}{1-p_t} = \frac{p_0}{1-p_0}\frac{d\mathbb{P}_+}{d\mathbb{P}_-}\Big|_{\mathcal{F}_t^X} = \frac{p_0}{1-p_0}\frac{\phi((X_t - mt)/\sqrt{t})}{\phi((X_t + mt)/\sqrt{t})} = \frac{p_0}{1-p_0}e^{2mX_t}.$$

We'll encounter Girsanov's theorem later. Also,

(A Confession)

This was a pretext to introduce Girsanov's theorem, which will be used later to 'remove drift' by changing the measure.

In this particular problem, we could have derived $\frac{d\mathbb{P}_+}{d\mathbb{P}_-}\big|_{\mathcal{F}_t^X}$ directly:

Since, conditional on θ, $X_t \sim N(m\theta t, t)$,

$$\frac{p_t}{1-p_t} = \frac{p_0}{1-p_0}\frac{d\mathbb{P}_+}{d\mathbb{P}_-}\Big|_{\mathcal{F}_t^X} = \frac{p_0}{1-p_0}\frac{\phi((X_t - mt)/\sqrt{t})}{\phi((X_t + mt)/\sqrt{t})} = \frac{p_0}{1-p_0}e^{2mX_t}.$$

We'll encounter Girsanov's theorem later. Also,

A famous application: Black–Scholes–Merton.

A Famous Application: Black–Scholes–Merton

Suppose stock follows $\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t$, and the risk-free interest rate is r.

Girsanov's theorem allows us to change to an equivalent *risk-neutral measure* $\mathbb{Q}$, under which $\mathrm{d}S_t = rS_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t^{\mathbb{Q}}$.
Discounted stock prices are then martingales.

European option: financial contract giving right (not obligaton) to buy/sell asset at predetermined strike price K on specified future maturity date T.

Insight: because we can continuously trade stock and risk-free asset, we can construct portfolio that exactly replicates payoff of european option.
No-arbitrage requires option to have same price as replicating portfolio.

Hence, european option with payoff $h(S_T)$ has price $V_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[h(S_T) \mid \mathcal{F}_t]$.

Remarkably, price does not depend on stock's actual expected return μ.
$\mathbb{Q}$ does depend on μ, but distribution of S_t under $\mathbb{Q}$ does not!

Applying Itô's Lemma

Back to our agenda: deriving law of motion of beliefs.

Let $\pi(x) := \frac{p_0 e^{mx}}{p_0 e^{mx} + (1-p_0)e^{-mx}}$. Then, $p_t = \pi(X_t)$.

Note:

$$\pi'(X_t) = 2m\, p_t(1 - p_t),$$

$$\pi''(X_t) = 4m^2 p_t(1 - p_t)(1 - 2p_t),$$

$$[X]_t = t.$$

Apply Itô's lemma to π:

$$\mathrm{d}p_t = \pi'(X_t)\,\mathrm{d}X_t + \frac{1}{2}\pi''(X_t)\,\mathrm{d}t = 2m\, p_t(1 - p_t)(\mathrm{d}X_t - m(2p_t - 1)\,\mathrm{d}t).$$

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Belief increments have zero conditional mean: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Belief increments have zero conditional mean: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s - 1)\, ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Note $\mathbb{E}[\theta \mid \mathcal{F}_t^X] = p_t(1) + (1 - p_t)(-1) = 2p_t - 1$. Then,

$$\mathbb{E}[\widetilde{B}_{t'} \mid \mathcal{F}_t^X] = \mathbb{E}[X_{t'} - \int_0^{t'} m(2p_s - 1)\, ds \mid \mathcal{F}_t^X]$$

Law of Motion

$dp_t = 2m\,p_t(1-p_t)(dX_t - m(2p_t-1)\,dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Belief increments have zero conditional mean: $\mathbb{E}[dX_t - m(2p_t-1)\,dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s-1)\,ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Note $\mathbb{E}[\theta \mid \mathcal{F}_t^X] = p_t(1) + (1-p_t)(-1) = 2p_t - 1$. Then,

$$\mathbb{E}[\widetilde{B}_{t'} \mid \mathcal{F}_t^X] = \mathbb{E}[X_{t'} - \int_0^{t'} m(2p_s-1)\,ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + \mathbb{E}[X_{t'} - X_t - \int_t^{t'} m(2p_s-1)\,ds \mid \mathcal{F}_t^X]$$

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Belief increments have zero conditional mean: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s - 1)\, ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Note $\mathbb{E}[\theta \mid \mathcal{F}_t^X] = p_t(1) + (1 - p_t)(-1) = 2p_t - 1$. Then,

$$\mathbb{E}[\widetilde{B}_{t'} \mid \mathcal{F}_t^X] = \mathbb{E}[X_{t'} - \int_0^{t'} m(2p_s - 1)\, ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + \mathbb{E}[X_{t'} - X_t - \int_t^{t'} m(2p_s - 1)\, ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + \mathbb{E}[m\theta(t' - t) + B_{t'} - B_t - \int_t^{t'} m(2p_s - 1)\, ds \mid \mathcal{F}_t^X]$$

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Belief increments have zero conditional mean: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s - 1)\, ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Note $\mathbb{E}[\theta \mid \mathcal{F}_t^X] = p_t(1) + (1 - p_t)(-1) = 2p_t - 1$. Then,

$$\mathbb{E}[\widetilde{B}_{t'} \mid \mathcal{F}_t^X] = \mathbb{E}[X_{t'} - \int_0^{t'} m(2p_s - 1)\, ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + \mathbb{E}[X_{t'} - X_t - \int_t^{t'} m(2p_s - 1)\, ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + \mathbb{E}[m\theta(t' - t) + B_{t'} - B_t - \int_t^{t'} m(2p_s - 1)\, ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + m\mathbb{E}[\theta \mid \mathcal{F}_t^X](t' - t) + \mathbb{E}[B_{t'} - B_t \mid \mathcal{F}_t^X] - \int_t^{t'} m(2\mathbb{E}[p_s \mid \mathcal{F}_t^X] - 1)\, ds$$

Law of Motion

$dp_t = 2m\, p_t(1-p_t)(dX_t - m(2p_t-1)\,dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Belief increments have zero conditional mean: $\mathbb{E}[dX_t - m(2p_t-1)\,dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s-1)\,ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Note $\mathbb{E}[\theta \mid \mathcal{F}_t^X] = p_t(1) + (1-p_t)(-1) = 2p_t - 1$. Then,

$$\mathbb{E}[\widetilde{B}_{t'} \mid \mathcal{F}_t^X] = \mathbb{E}[X_{t'} - \int_0^{t'} m(2p_s-1)\,ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + \mathbb{E}[X_{t'} - X_t - \int_t^{t'} m(2p_s-1)\,ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + \mathbb{E}[m\theta(t'-t) + B_{t'} - B_t - \int_t^{t'} m(2p_s-1)\,ds \mid \mathcal{F}_t^X]$$

$$= \widetilde{B}_t + m\mathbb{E}[\theta \mid \mathcal{F}_t^X](t'-t) + \mathbb{E}[B_{t'} - B_t \mid \mathcal{F}_t^X] - \int_t^{t'} m(2\mathbb{E}[p_s \mid \mathcal{F}_t^X] - 1)\,ds$$

$$= \widetilde{B}_t + m(2p_t-1)(t'-t) - \int_t^{t'} m(2p_t-1)\,ds = \widetilde{B}_t.$$

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Changes in beliefs are now predictable: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Changes in beliefs are now predictable: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s - 1)\, ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Moreover, $[\widetilde{B}]_t = [X]_t = [B]_t = t$.

Since $p_t \in [0, 1]$, $\| \int_t^{t'} m(2p_s - 1)\, ds \| \leq m|t - t'|$ and X has continuous paths, then $\widetilde{B}$ has continuous paths.

$\widetilde{B}$ has $\mathcal{F}^X$-continuous martingale and $[\widetilde{B}]_t = t$ iff $\widetilde{B}$ is $\mathcal{F}^X$-Brownian-motion.

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Changes in beliefs are now predictable: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s - 1)\, ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Moreover, $[\widetilde{B}]_t = [X]_t = [B]_t = t$.

Since $p_t \in [0, 1]$, $\| \int_t^{t'} m(2p_s - 1)\, ds \| \leq m|t - t'|$ and X has continuous paths, then $\widetilde{B}$ has continuous paths.

$\widetilde{B}$ has $\mathcal{F}^X$-continuous martingale and $[\widetilde{B}]_t = t$ iff $\widetilde{B}$ is $\mathcal{F}^X$-Brownian-motion.

Conclude: $dp_t = 2m\, p_t(1 - p_t)\, d\widetilde{B}_t$.

Law of Motion

$dp_t = 2m\, p_t(1 - p_t)(dX_t - m(2p_t - 1)\, dt)$. Then, $p_t = \pi(X_t)$.

Law of motion still depends on X. But we know that beliefs are a martingale:

$\mathbb{E}[p_{t'} \mid \mathcal{F}_t^X] = p_t$ for any $t' > t$.

Changes in beliefs are now predictable: $\mathbb{E}[dX_t - m(2p_t - 1)\, dt \mid \mathcal{F}_t^X] = 0$.

Innovations process: $\widetilde{B}_t := X_t - \int_0^t m(2p_s - 1)\, ds$. $\widetilde{B}_t$ is $\mathcal{F}^X$-martingale!

Moreover, $[\widetilde{B}]_t = [X]_t = [B]_t = t$.

Since $p_t \in [0, 1]$, $\| \int_t^{t'} m(2p_s - 1)\, ds \| \leq m|t - t'|$ and X has continuous paths, then $\widetilde{B}$ has continuous paths.

$\widetilde{B}$ has $\mathcal{F}^X$-continuous martingale and $[\widetilde{B}]_t = t$ iff $\widetilde{B}$ is $\mathcal{F}^X$-Brownian-motion.

Conclude: $dp_t = 2m\, p_t(1 - p_t)\, d\widetilde{B}_t$.

This is a filtering problem: observe X_t and want to infer something about underlying θ.

We'll now see a more powerful way to get at dp_t.

A General Filtering Problem

Hidden state θ and **observed signal** X satisfy

$$d\theta_t = b(t, \theta_t, X_t)\, dt + a(t, \theta_t, X_t)\, dW_t, \qquad dX_t = \mu(t, \theta_t, X_t)\, dt + \sigma(t, X_t)\, dB_t.$$

W, B are independent Brownian motions, independent of θ_0; a, b, μ, σ are known.

Importantly: σ does not depend on θ.

Want to characterise law of motion for beliefs about some function of θ, $f(\theta_t)$.

A General Filtering Problem

Hidden state θ and **observed signal** X satisfy

$$d\theta_t = b(t, \theta_t, X_t)\, dt + a(t, \theta_t, X_t)\, dW_t, \qquad dX_t = \mu(t, \theta_t, X_t)\, dt + \sigma(t, X_t)\, dB_t.$$

W, B are independent Brownian motions, independent of θ_0; a, b, μ, σ are known.

Importantly: σ does not depend on θ.

Want to characterise law of motion for beliefs about some function of θ, $f(\theta_t)$.

Filter: $\phi_t(f) := \mathbb{E}[f(\theta_t) \mid \mathcal{F}_t^X]$.

Hidden-state generator: $(\mathcal{L}_t^x f)(\theta) = b(t, \theta, x) f'(\theta) + \frac{1}{2} a(t, \theta, x)^2 f''(\theta)$.

Write $\mu_t := \mu(t, \theta_t, X_t)$, $\sigma_t := \sigma(t, X_t)$, and $\mathrm{Cov}_t(\cdot) := \mathrm{Cov}(\cdot \mid \mathcal{F}_t^X)$.

A General Filtering Problem

Hidden state θ and **observed signal** X satisfy

$$d\theta_t = b(t,\theta_t,X_t)\,dt + a(t,\theta_t,X_t)\,dW_t, \qquad dX_t = \mu(t,\theta_t,X_t)\,dt + \sigma(t,X_t)\,dB_t.$$

W, B are independent Brownian motions, independent of θ_0; a, b, μ, σ are known.

Importantly: σ does not depend on θ.

Want to characterise law of motion for beliefs about some function of θ, $f(\theta_t)$.

Filter: $\phi_t(f) := \mathbb{E}[f(\theta_t) \mid \mathcal{F}_t^X]$.

Hidden-state generator: $(\mathcal{L}_t^x f)(\theta) = b(t,\theta,x)f'(\theta) + \frac{1}{2}a(t,\theta,x)^2 f''(\theta)$.

Write $\mu_t := \mu(t,\theta_t,X_t)$, $\sigma_t := \sigma(t,X_t)$, and $\mathrm{Cov}_t(\cdot) := \mathrm{Cov}(\cdot \mid \mathcal{F}_t^X)$.

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$d\phi_t(f) = \phi_t(\mathcal{L}_t^{X_t} f)\,dt + \frac{\phi_t(f\mu_t) - \phi_t(f)\phi_t(\mu_t)}{\sigma_t}\,d\widetilde{B}_t = \phi_t(\mathcal{L}_t^{X_t} f)\,dt + \frac{\mathrm{Cov}_t(f(\theta_t),\mu_t)}{\sigma_t}\,d\widetilde{B}_t,$$

where $d\widetilde{B}_t := \dfrac{dX_t - \phi_t(\mu_t)\,dt}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

A General Filtering Problem

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$\mathrm{d}\phi_t(f) = \underbrace{\phi_t(\mathcal{L}_t^{X_t} f)\,\mathrm{d}t}_{\text{Evolution of hidden state}} + \underbrace{\frac{\mathrm{Cov}_t(f(\theta_t), \mu_t)}{\sigma_t}\,\mathrm{d}\widetilde{B}_t}_{\text{Learning}}.$$

where $\mathrm{d}\widetilde{B}_t := \dfrac{\mathrm{d}X_t - \phi_t(\mu_t)\,\mathrm{d}t}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

Some loose sufficient regularity conditions: a, b, μ, σ: Borel measurable, bounded, Lipschitz continuous in spatial arguments; σ bounded away from zero; independent square-integrable initial conditions; $f \in \mathcal{C}_b^2$.

Extends to allowing for jumps.

References: Germ and Gyöngy (2025) 'On partially observed jump diffusions I: the filtering equations'. Also (less general): Liptser and Shiryaev (2001, Ch. 8-10).

A General Filtering Problem

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$\mathrm{d}\phi_t(f) = \underbrace{\phi_t(\mathcal{L}_t^{X_t} f)\,\mathrm{d}t}_{\text{Evolution of hidden state}} + \underbrace{\frac{\mathrm{Cov}_t(f(\theta_t), \mu_t)}{\sigma_t}\,\mathrm{d}\widetilde{B}_t}_{\text{Learning}}.$$

where $\mathrm{d}\widetilde{B}_t := \dfrac{\mathrm{d}X_t - \phi_t(\mu_t)\,\mathrm{d}t}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

Prediction: hidden state may evolve even without new observations.

Updating: unexpected evidence changes estimate in proportion to its conditional covariance with the signal drift.

Cases of pure learning: If $b = 0$ (hidden state does not drift) and (a) $a = 0$ (hidden state fixed) or (b) f is affine, then $\phi_t(\mathcal{L}_t^{X_t} f) = 0$.

A General Filtering Problem

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$\mathrm{d}\phi_t(f) = \underbrace{\phi_t(\mathcal{L}_t^{X_t} f)\,\mathrm{d}t}_{\text{Evolution of hidden state}} + \underbrace{\frac{\mathrm{Cov}_t(f(\theta_t), \mu_t)}{\sigma_t}\,\mathrm{d}\widetilde{B}_t}_{\text{Learning}}.$$

where $\mathrm{d}\widetilde{B}_t := \dfrac{\mathrm{d}X_t - \phi_t(\mu_t)\,\mathrm{d}t}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

Prediction: hidden state may evolve even without new observations.

Updating: unexpected evidence changes estimate in proportion to its conditional covariance with the signal drift.

Cases of pure learning: If $b = 0$ (hidden state does not drift) and (a) $a = 0$ (hidden state fixed) or (b) f is affine, then $\phi_t(\mathcal{L}_t^{X_t} f) = 0$.

Now you know how to derive $\mathrm{d}p_t$ for basically any problem!

A General Filtering Problem

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$\mathrm{d}\phi_t(f) = \underbrace{\phi_t(\mathcal{L}_t^{X_t} f)\,\mathrm{d}t}_{\text{Evolution of hidden state}} + \underbrace{\frac{\mathrm{Cov}_t(f(\theta_t), \mu_t)}{\sigma_t}\,\mathrm{d}\widetilde{B}_t}_{\text{Learning}}.$$

where $\mathrm{d}\widetilde{B}_t := \dfrac{\mathrm{d}X_t - \phi_t(\mu_t)\,\mathrm{d}t}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

Cases of pure learning: If $a = 0$ (hidden state does not drift) and

(a) $b = 0$ (hidden state fixed) or (b) f is affine, then $\phi_t(\mathcal{L}_t^{X_t} f) = 0$.

A General Filtering Problem

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$\mathrm{d}\phi_t(f) = \underbrace{\phi_t(\mathcal{L}_t^{X_t} f)\,\mathrm{d}t}_{\text{Evolution of hidden state}} + \underbrace{\frac{\mathrm{Cov}_t(f(\theta_t), \mu_t)}{\sigma_t}\,\mathrm{d}\widetilde{B}_t}_{\text{Learning}}.$$

where $\mathrm{d}\widetilde{B}_t := \dfrac{\mathrm{d}X_t - \phi_t(\mu_t)\,\mathrm{d}t}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

Cases of pure learning: If $a = 0$ (hidden state does not drift) and

(a) $b = 0$ (hidden state fixed) or (b) f is affine, then $\phi_t(\mathcal{L}_t^{X_t} f) = 0$.

In our problem with known m:

$a = b = 0$, choose $f(\theta) = 1_{\{\theta=1\}} = (1+\theta)/2$ ($\because \theta = \pm 1$). And since $\mu_t = m\theta$ and $\sigma_t = 1$,

A General Filtering Problem

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$\mathrm{d}\phi_t(f) = \underbrace{\phi_t(\mathcal{L}_t^{X_t} f)\,\mathrm{d}t}_{\text{Evolution of hidden state}} + \underbrace{\frac{\mathrm{Cov}_t(f(\theta_t), \mu_t)}{\sigma_t}\,\mathrm{d}\widetilde{B}_t}_{\text{Learning}}.$$

where $\mathrm{d}\widetilde{B}_t := \dfrac{\mathrm{d}X_t - \phi_t(\mu_t)\,\mathrm{d}t}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

Cases of pure learning: If $a = 0$ (hidden state does not drift) and

(a) $b = 0$ (hidden state fixed) or (b) f is affine, then $\phi_t(\mathcal{L}_t^{X_t} f) = 0$.

In our problem with known m:

$a = b = 0$, choose $f(\theta) = 1_{\{\theta=1\}} = (1+\theta)/2$ ($\because \theta = \pm 1$). And since $\mu_t = m\theta$ and $\sigma_t = 1$,

$$\mathrm{d}p_t = \mathrm{Cov}_t\left(\frac{1+\theta}{2}, m\theta\right)\mathrm{d}\widetilde{B}_t = \frac{m}{2}\mathrm{Var}_t(\theta)\,\mathrm{d}\widetilde{B}_t = 2mp_t(1-p_t)\,\mathrm{d}\widetilde{B}_t.$$

A General Filtering Problem

Theorem (Kushner–Stratonovich Equation)

Under suitable regularity conditions,

$$\mathrm{d}\phi_t(f) = \underbrace{\phi_t(\mathcal{L}_t^{X_t} f)\,\mathrm{d}t}_{\text{Evolution of hidden state}} + \underbrace{\frac{\mathrm{Cov}_t(f(\theta_t), \mu_t)}{\sigma_t}\,\mathrm{d}\widetilde{B}_t}_{\text{Learning}}.$$

where $\mathrm{d}\widetilde{B}_t := \dfrac{\mathrm{d}X_t - \phi_t(\mu_t)\,\mathrm{d}t}{\sigma_t}$ is the innovations Brownian motion under $\mathcal{F}^X$.

Cases of pure learning: If $a = 0$ (hidden state does not drift) and

(a) $b = 0$ (hidden state fixed) or (b) f is affine, then $\phi_t(\mathcal{L}_t^{X_t} f) = 0$.

In our problem with known m:

$a = b = 0$, choose $f(\theta) = 1_{\{\theta=1\}} = (1+\theta)/2$ ($\because \theta = \pm 1$). And since $\mu_t = m\theta$ and $\sigma_t = 1$,

$$\mathrm{d}p_t = \mathrm{Cov}_t\left(\frac{1+\theta}{2}, m\theta\right)\mathrm{d}\widetilde{B}_t = \frac{m}{2}\mathrm{Var}_t(\theta)\,\mathrm{d}\widetilde{B}_t = 2mp_t(1-p_t)\,\mathrm{d}\widetilde{B}_t.$$

Now you know how to derive $\mathrm{d}p_t$ for basically any problem!

Filtering with (Potentially) Unknown Drift

Results in new version of SAC. Will make our lives very convenient.

Filtering with (Potentially) Unknown Drift

Results in new version of SAC. Will make our lives very convenient.

Theorem (Gonçalves (WP) SAC)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Suppose B, θ, m are independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$ and $m \sim \mu$, where $\mu(\mathbb{R}_{++}) > 0$.
Let $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$ and define $\mathcal{M} := \{\mu \in \Delta(\mathbb{R}_+) \mid \mu(\mathbb{R}_{++}) > 0\}$, endowed with the weak-convergence topology. Then, $\exists \pi(t,x)$ smooth on $\mathbb{R}_{++} \times \mathbb{R}$ such that $t > 0, p \in (0,1)$, strictly increasing in x, such that $p_t = \pi(t, X_t)$.

From process to beliefs and back: write $p_t = \pi(t, X_t)$ and $X_t = \chi(t, p_t)$, where $\chi(t, \cdot) : p = \pi(t, \chi(t,p))$.

Known m: When $\mu(\{m\}) = 1$, $\operatorname{logit} p_t = 2mX_t + \operatorname{logit} p_0$, where $\operatorname{logit}(p) := \ln(p/(1-p))$.
Therefore, $\chi(t,p) = (\operatorname{logit} p - \operatorname{logit} p_0)/(2m)$.

Filtering with (Potentially) Unknown Drift

Results in new version of SAC. Will make our lives very convenient.

Theorem (Gonçalves (WP) SAC)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Suppose B, θ, m are independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$ and $m \sim \mu$, where μ has finite mean and $\mu(\mathbb{R}_{++}) > 0$.

Let $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$ and define $\mathcal{M} := \{\mu \in \Delta(\mathbb{R}_+) \mid \mu(\mathbb{R}_{++}) > 0\}$, endowed with the weak-convergence topology. Then, $\exists \sigma_\mu(t,p)$ such that p_t is a strong solution of $dp_t = \sigma_\mu(t,p)\, d\widetilde{B}_t$ and satisfies

(P1) $\sigma_\mu \in \mathcal{C}^\infty(\mathbb{R}_+ \times (0,1))$.

(P2) $\sigma_\mu(t,p)$ is continuous in $(t,p,\mu) \in \mathbb{R}_{++} \times (0,1) \times \mathcal{M}$.

(P3) $\sigma_\mu > 0$ on $\mathbb{R}_+ \times (0,1)$.

(P4) $\partial_t \sigma_\mu(t,p) \leq 0$, $\forall t \geq 0$, with a strict inequality iff μ is non-degenerate.

We now know a bunch of properties for the law of motion of p_t that hold in general.

Known m: $\sigma(t,p) = 2mp(1-p)$ as we've derived.

Filtering with (Potentially) Unknown Drift

Results in new version of SAC. Will make our lives very convenient.

Proposition (Gonçalves (WP) SAC)

Define $\bar{x} = \sup\left\{a \geq 0 : \int_0^\infty e^{am}\, d\mu(m) < \infty\right\}$. Suppose (a) μ has bounded support or (b) μ admits a log-concave density and either $\bar{x} = \infty$ or $\bar{x} < \infty$ and $\lim_{m\to\infty} e^{\bar{x}m}\, d\mu(m) = \infty$. Then,
(P5) $\sigma_\mu \in \mathcal{C}^0(\mathbb{R}_+ \times [0,1])$ and $\sigma_\mu = 0$ on $\mathbb{R}_+ \times \{0,1\}$ and $\lim_{t\to 0} \sup_{p\in[0,1]} |\sigma_\mu(t,p) - \sigma_\mu(0,p)| = 0$.

Under general-enough conditions, σ admits a continuous extension to $p \in \{0,1\}$ and $t = 0$. This is not for free! We'll assume this away.

Filtering with (Potentially) Unknown Drift

The following claims will allow us to establish 'learn arbitrarily fast' as the distribution of m shifts up and 'learn nothing' as distribution of m collapses to 0.

Filtering with (Potentially) Unknown Drift

The following claims will allow us to establish 'learn arbitrarily fast' as the distribution of m shifts up and 'learn nothing' as distribution of m collapses to 0.

Proposition (Gonçalves (WP) SAC)

Let $\{\mu_n\}_n$ have log-concave densities. The following claims hold:

$\lim_n \mathbb{E}_{M\sim\mu_n}[M] = 0$ if and only if $\lim_n \sup_{t\geq 0, p\in[0,1]} \sigma_{\mu_n}(t,p) = 0$.

If $\lim_n \mu_n([0,R]) = 0$ for every $R > 0$, then, for any compact $K \subset (0,1)$ and $t > 0$, $\lim_n \int_0^t \inf_{p\in K} \sigma_{\mu_n}(s,p)^2 \, ds = \infty$.

Proposition (Gonçalves (WP) SAC)

The following claims hold:

If $\{\mu_n\}_n$ satisfy $\operatorname{ess\,sup}_{M\sim\mu_n} M \to 0$, then, $\lim_n \sup_{t\geq 0, p\in[0,1]} \sigma_{\mu_n}(t,p) = 0$.

If $\{\mu_n\}_n$ satisfy $\operatorname{ess\,inf}_{M\sim\mu_n} M \to \infty$, then, for any compact $K \subset (0,1)$ and $t > 0$, $\lim_n \int_0^t \inf_{p\in K} \sigma_{\mu_n}(s,p)^2 \, ds = \infty$.

Overview

1. A Pioneer from Cognitive Sciences

2. Speed, Accuracy, and Complexity

3. Evidence and Beliefs

4. Exogenous Stopping and Problem Complexity
   - Belief Thresholds
   - Evidence Thresholds

5. Optimal Stopping

6. Comparative Statics on Incentives and Revealing Complexity

7. Fast and Slow Errors

Exogenous Stopping

Stopping thresholds: $(\underline{p}_t, \overline{p}_t)$. Stopping time: $\tau := \inf\{t \geq 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$.

Choice: $\alpha_\tau = 1$ if belief hits upper threshold $p_\tau = \overline{p}_\tau$; ow $\alpha_\tau = -1$ if $p_\tau = \underline{p}_\tau$.

E.g., stopping upon reaching pre-determined time-dependent level of certainty, or securing satisficing level of expected payoff.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.
Let $\overline{p}, \underline{p} : \mathbb{R}_+ \to (0,1)$ be $\mathcal{C}^1$ and satisfy: (a) $\overline{p}_t' \leq 0 \leq \underline{p}_t'$, (b) $1/2 \in (\underline{p}_t, \overline{p}_t)$ whenever non-empty. Define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = \text{sign}(p_t - 1/2)$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.
Let $\overline{p}, \underline{p} : \mathbb{R}_+ \to (0,1)$ be $\mathcal{C}^1$ and satisfy: (a) $\overline{p}_t' \leq 0 \leq \underline{p}_t'$, (b) $1/2 \in (\underline{p}_t, \overline{p}_t)$ whenever non-empty. Define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = \text{sign}(p_t - 1/2)$.

Suppose, for $\underline{m} > 0$, (c) $-\rho(\underline{p}_t)\overline{p}_t'$, $\rho(\overline{p}_t)\underline{p}_t' \leq 1$, where $\rho(q) := \int_{\underline{p}_t}^{\overline{p}_t} 2\frac{p-q}{\overline{p}_t - \underline{p}_t} \frac{1}{\sigma(t,p,\underline{m})^2}\, dp$ and $\sigma(t,p,m) := 2mp(1-p)$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.

Let $\overline{p}, \underline{p} : \mathbb{R}_+ \to (0,1)$ be $\mathcal{C}^1$ and satisfy: (a) $\overline{p}_t' \leq 0 \leq \underline{p}_t'$, (b) $1/2 \in (\underline{p}_t, \overline{p}_t)$ whenever non-empty. Define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = \text{sign}(p_t - 1/2)$.

Suppose, for $\underline{m} > 0$, (c) $-\rho(\underline{p}_t)\overline{p}_t', \rho(\overline{p}_t)\underline{p}_t' \leq 1$, where $\rho(q) := \int_{\underline{p}_t}^{\overline{p}_t} 2\frac{p-q}{\overline{p}_t - \underline{p}_t} \frac{1}{\sigma(t,p,\underline{m})^2}\, dp$ and $\sigma(t,p,m) := 2mp(1-p)$.

Then, (1) $\mathbb{E}[\tau]$ decreases in $m \geq \underline{m}$, and (2) $\mathbb{P}(\alpha_\tau = \text{sign}(\theta))$ increases in m.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.
Let $\overline{p}, \underline{p} : \mathbb{R}_+ \to (0,1)$ be $\mathcal{C}^1$ and satisfy: (a) $\overline{p}_t' \leq 0 \leq \underline{p}_t'$, (b) $1/2 \in (\underline{p}_t, \overline{p}_t)$ whenever non-empty. Define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = \text{sign}(p_t - 1/2)$.
Suppose, for $\underline{m} > 0$, (c) $-\rho(\underline{p}_t)\overline{p}_t', \rho(\overline{p}_t)\underline{p}_t' \leq 1$, where $\rho(q) := \int_{\underline{p}_t}^{\overline{p}_t} 2\frac{p-q}{\overline{p}_t - \underline{p}_t} \frac{1}{\sigma(t,p,\underline{m})^2}\, dp$ and $\sigma(t,p,m) := 2mp(1-p)$.
Then, (1) $\mathbb{E}[\tau]$ decreases in $m \geq \underline{m}$, and (2) $\mathbb{P}(\alpha_\tau = \text{sign}(\theta))$ increases in m.

Condition (b) needed for (2) (accuracy) not for (1).

Condition (c) needed for (1) (time) not for (2).

DM knows m, but stopping depends on some potentially suboptimal thresholds.

Thresholds *in belief space* held fixed as m changes.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.

Let $\overline{p}, \underline{p} : \mathbb{R}_+ \to (0,1)$ be $\mathcal{C}^1$ and satisfy: (a) $\overline{p}_t' \leq 0 \leq \underline{p}_t'$, (b) $1/2 \in (\underline{p}_t, \overline{p}_t)$ whenever non-empty. Define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = \text{sign}(p_t - 1/2)$.

Suppose, for $\underline{m} > 0$, (c) $-\rho(\underline{p}_t)\overline{p}_t', -\rho(\overline{p}_t)\underline{p}_t' \leq 1$, where $\rho(q) := \int_{\underline{p}_t}^{\overline{p}_t} 2\frac{p-q}{\overline{p}_t - \underline{p}_t} \frac{1}{\sigma(t,p,\underline{m})^2}\, dp$ and $\sigma(t,p,m) := 2mp(1-p)$.

Then, (1) $\mathbb{E}[\tau]$ decreases in $m \geq \underline{m}$, and (2) $\mathbb{P}(\alpha_\tau = \text{sign}(\theta))$ increases in m.

Interpretation? Why do we care about exogenous stopping rules?

(i) Potentially accommodates weird payoff functions.

(ii) Contrast with optimal stopping: are comparative statics general or due to optimality?

(iii) Results generalise to general cases when $\theta \sim G \in \Delta(\mathbb{R})$ using only properties of $\sigma(t,p,m)$. Then, $|\theta|$ captures uncertainty about drift rate and m is belief shifter.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.
Let $\overline{p}, \underline{p} : \mathbb{R}_+ \to (0,1)$ be $\mathcal{C}^1$ and satisfy: (a) $\overline{p}_t' \leq 0 \leq \underline{p}_t'$, (b) $1/2 \in (\underline{p}_t, \overline{p}_t)$ whenever non-empty. Define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = \text{sign}(p_t - 1/2)$.
Suppose, for $\underline{m} > 0$, (c) $-\rho(\underline{p}_t)\overline{p}_t', -\rho(\overline{p}_t)\underline{p}_t' \leq 1$, where $\rho(q) := \int_{\underline{p}_t}^{\overline{p}_t} 2\frac{p-q}{\overline{p}_t - \underline{p}_t} \frac{1}{\sigma(t,p,\underline{m})^2}\, dp$ and $\sigma(t,p,m) := 2mp(1-p)$.
Then, (1) $\mathbb{E}[\tau]$ decreases in $m \geq \underline{m}$, and (2) $\mathbb{P}(\alpha_\tau = \text{sign}(\theta))$ increases in m.

Proof guide:

Proof of (1) expresses $\mathbb{E}^{t,p}[\tau \wedge T - t]$ as solution to boundary value problem and relies on repeated application of maximum principle (later). Details beyond scope of this course.

Proof of (2) next.

Exogenous Stopping

Proof Idea

Belief about the state conditional on choice = Stopping posterior.

$\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta | \alpha_\tau = 1) = \mathbb{E}[\overline{p}_\tau \mid \alpha_\tau = 1]$ and $\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta | \alpha_\tau = -1) = \mathbb{E}[1 - \underline{p}_\tau \mid \alpha_\tau = -1]$.

Exogenous Stopping

Proof Idea

Belief about the state conditional on choice = Stopping posterior.

$\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta | \alpha_\tau = 1) = \mathbb{E}[\overline{p}_\tau \mid \alpha_\tau = 1]$ and $\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta | \alpha_\tau = -1) = \mathbb{E}[1 - \underline{p}_\tau \mid \alpha_\tau = -1]$.

Accuracy = 1/2 + expected distance stopping posterior to 1/2.

$$\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta) = \mathbb{E}[1_{\{\text{sign}\,\alpha_\tau = \text{sign}\,\theta\}}] = \mathbb{E}[1_{\{\alpha_\tau = 1\}}\overline{p}_\tau + 1_{\{\alpha_\tau = -1\}}(1 - \underline{p}_\tau)]$$

Exogenous Stopping

Proof Idea

Belief about the state conditional on choice = Stopping posterior.

$\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta | \alpha_\tau = 1) = \mathbb{E}[\overline{p}_\tau \mid \alpha_\tau = 1]$ and $\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta | \alpha_\tau = -1) = \mathbb{E}[1 - \underline{p}_\tau \mid \alpha_\tau = -1]$.

Accuracy = 1/2 + expected distance stopping posterior to 1/2.

$$\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta) = \mathbb{E}[1_{\{\text{sign}\,\alpha_\tau = \text{sign}\,\theta\}}] = \mathbb{E}[1_{\{\alpha_\tau = 1\}}\overline{p}_\tau + 1_{\{\alpha_\tau = -1\}}(1 - \underline{p}_\tau)]$$

$$= \frac{1}{2} + \mathbb{E}[1_{\{\alpha_\tau = 1\}}(\overline{p}_\tau - \frac{1}{2}) + 1_{\{\alpha_\tau = -1\}}(\frac{1}{2} - \underline{p}_\tau)]$$

Exogenous Stopping

Proof Idea

Belief about the state conditional on choice = Stopping posterior.

$\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta|\alpha_\tau = 1) = \mathbb{E}[\overline{p}_\tau \mid \alpha_\tau = 1]$ and $\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta|\alpha_\tau = -1) = \mathbb{E}[1 - \underline{p}_\tau \mid \alpha_\tau = -1]$.

Accuracy = 1/2 + expected distance stopping posterior to 1/2.

$$\begin{aligned}\mathbb{P}(\text{sign}\,\alpha_\tau = \text{sign}\,\theta) &= \mathbb{E}[1_{\{\text{sign}\,\alpha_\tau = \text{sign}\,\theta\}}] = \mathbb{E}[1_{\{\alpha_\tau=1\}}\overline{p}_\tau + 1_{\{\alpha_\tau=-1\}}(1-\underline{p}_\tau)] \\ &= \frac{1}{2} + \mathbb{E}[1_{\{\alpha_\tau=1\}}(\overline{p}_\tau - \frac{1}{2}) + 1_{\{\alpha_\tau=-1\}}(\frac{1}{2} - \underline{p}_\tau)] \\ &= \frac{1}{2} + \mathbb{E}[1_{\{\alpha_\tau=1\}}|\overline{p}_\tau - \frac{1}{2}| + 1_{\{\alpha_\tau=-1\}}|\frac{1}{2} - \underline{p}_\tau|] = \frac{1}{2} + \mathbb{E}[|p_\tau - \frac{1}{2}|]\end{aligned}$$

Here you see why it is important that $1/2 \in (\underline{p}_t, \overline{p}_t)$ for the proof.

Exogenous Stopping

Proof Idea

Key intuitions:

(1) higher m, diffuse more at every (t, p), beliefs are somewhat like 'MPS';

(2) continuation region shrinking, so exit earlier means more extreme beliefs;

hence, expect larger $\mathbb{E}[|p_\tau - \frac{1}{2}|]$.

Exogenous Stopping

Proof Idea

Key intuitions:

(1) higher m, diffuse more at every (t, p), beliefs are somewhat like 'MPS';

(2) continuation region shrinking, so exit earlier means more extreme beliefs;

hence, expect larger $\mathbb{E}[|p_\tau - \frac{1}{2}|]$.

Key hurdles: (a) exit times not FOSD ordered;

(b) thresholds not symmetric;

(c) prior not necessarily 1/2;

(d) $|p_\tau - \frac{1}{2}|$ is submartingale, so taking longer is beneficial.

Exogenous Stopping

Proof Idea

Key intuitions:

(1) higher m, diffuse more at every (t, p), beliefs are somewhat like 'MPS';

(2) continuation region shrinking, so exit earlier means more extreme beliefs;

hence, expect larger $\mathbb{E}[|p_\tau - \frac{1}{2}|]$.

Key hurdles: (a) exit times not FOSD ordered;

(b) thresholds not symmetric;

(c) prior not necessarily 1/2;

(d) $|p_\tau - \frac{1}{2}|$ is submartingale, so taking longer is beneficial.

Proof strategy: Couple exit times on common time scale with fixed belief realisations;

There, higher m translates to taking *longer* (on this different time scale).

Conclude via optional stopping theorem.

Exogenous Stopping

Theorem (Dambis–Dubins–Schwarz)

Let M be a continuous local martingale with $M_0 = 0$. Define $A_t := [M]_t$.
(On possibly enlarged probability space) $\exists$ standard Brownian motion W, wrt suitable time-changed filtration, s.t. $M_t = W_{A_t}$, $t \geq 0$, a.s.
If $A_\infty = \infty$ a.s., no enlargement is needed and $W_u = M_{T_u}$, $T_u := \inf\{t \geq 0 : A_t > u\}$.

See Karatzas–Shreve (1998 Ch. 3).

Exogenous Stopping

Theorem (Dambis–Dubins–Schwarz)

Let M be a continuous local martingale with $M_0 = 0$. Define $A_t := [M]_t$.
(On possibly enlarged probability space) $\exists$ standard Brownian motion W, wrt suitable time-changed filtration, s.t. $M_t = W_{A_t}$, $t \geq 0$, a.s.
If $A_\infty = \infty$ a.s., no enlargement is needed and $W_u = M_{T_u}$, $T_u := \inf\{t \geq 0 : A_t > u\}$.

See Karatzas–Shreve (1998 Ch. 3).

Proof

Write P_t^m for the belief process fixing $m > 0$ and $C_t := (\underline{p}_t, \overline{p}_t)$.

Using Dambis-Dubins-Schwarz:

Let q be BM with $q_0 = p_0$ and absorbed at $\{0, 1\}$.

By DDS: $A_t^m := [P^m]_t = \int_0^t \sigma(s, P_s^m, m)^2 \, ds$ and $P^m := q_{A^m}$.
Defining $\tilde{B}_t := \int_0^t \sigma(t, P_t^m, m)^{-1} \, dP_t^m$, we have $[\tilde{B}]_t = t$ and $\tilde{B}$ is a BM.
That's fine, but want to tie-in realisations for P^m and $P^{m'}$ on same filtration.
We'll do this with random clock based on q.

Proof

Couple belief processes for different m.

Write ρ^m for generalised inverse of A^m: $A^m_{\rho^m(u)} = u \iff \rho^m(A^m_t) = t$.

Proof

Couple belief processes for different m.

Write ρ^m for generalised inverse of A^m: $A^m_{\rho^m(u)} = u \iff \rho^m(A^m_t) = t$.

Differentiate ρ^m, conclude ρ^m solves ODE $\dfrac{\mathrm{d}}{\mathrm{d}u}\rho^m(u) = \dfrac{1}{\sigma(\rho^m(u), q_u, m)^2}$, $\rho^m(0) = 0$.

Proof

Couple belief processes for different m.

Write ρ^m for generalised inverse of A^m: $A^m_{\rho^m(u)} = u \iff \rho^m(A^m_t) = t$.

Differentiate ρ^m, conclude ρ^m solves ODE $\dfrac{\mathrm{d}}{\mathrm{d}u}\rho^m(u) = \dfrac{1}{\sigma(\rho^m(u), q_u, m)^2}$, $\rho^m(0) = 0$.

Define $\tau_{\text{abs}} := \inf\{u > 0 : q_u \in \{0, 1\}\}$.

Let ρ^m solve random ODE above on $[0, \tau_{\text{abs}})$; set $\rho^m(u) = \infty$ for $u \geq \tau_{\text{abs}}$ (this deals with never getting beyond 0 and 1).

Proof

Couple belief processes for different m.

Write ρ^m for generalised inverse of A^m: $A^m_{\rho^m(u)} = u \iff \rho^m(A^m_t) = t$.

Differentiate ρ^m, conclude ρ^m solves ODE $\frac{\mathrm{d}}{\mathrm{d}u}\rho^m(u) = \frac{1}{\sigma(\rho^m(u), q_u, m)^2}$, $\rho^m(0) = 0$.

Define $\tau_{\text{abs}} := \inf\{u > 0 : q_u \in \{0, 1\}\}$.

Let ρ^m solve random ODE above on $[0, \tau_{\text{abs}})$; set $\rho^m(u) = \infty$ for $u \geq \tau_{\text{abs}}$ (this deals with never getting beyond 0 and 1).

Define $\widetilde{P}^m_t := q_{A^m_t}$. $\implies [\widetilde{P}^m]_t = A^m_t$ and $A^m_t = \int_0^t \sigma(s, \widetilde{P}^m_s, m)^2 \,\mathrm{d}s$.

$\implies \widetilde{P}^m$ solves $\mathrm{d}\widetilde{P}^m_t = \sigma(t, \widetilde{P}^m_t, m)\,\mathrm{d}\widetilde{B}_t$ for some Brownian motion $\widetilde{B}$.

Conclude $\widetilde{P}^m \overset{d}{=} P^m$ (skipping discussion on strong/weak solutions).

Proof

Couple belief processes for different m.

Write ρ^m for generalised inverse of A^m: $A^m_{\rho^m(u)} = u \iff \rho^m(A^m_t) = t$.

Differentiate ρ^m, conclude ρ^m solves ODE $\frac{\mathrm{d}}{\mathrm{d}u}\rho^m(u) = \frac{1}{\sigma(\rho^m(u), q_u, m)^2}$, $\rho^m(0) = 0$.

Define $\tau_{\text{abs}} := \inf\{u > 0 : q_u \in \{0, 1\}\}$.

Let ρ^m solve random ODE above on $[0, \tau_{\text{abs}})$; set $\rho^m(u) = \infty$ for $u \geq \tau_{\text{abs}}$ (this deals with never getting beyond 0 and 1).

Define $\widetilde{P}^m_t := q_{A^m_t}$. $\implies [\widetilde{P}^m]_t = A^m_t$ and $A^m_t = \int_0^t \sigma(s, \widetilde{P}^m_s, m)^2 \,\mathrm{d}s$.

$\implies \widetilde{P}^m$ solves $\mathrm{d}\widetilde{P}^m_t = \sigma(t, \widetilde{P}^m_t, m)\,\mathrm{d}\widetilde{B}_t$ for some Brownian motion $\widetilde{B}$.

Conclude $\widetilde{P}^m \overset{d}{=} P^m$ (skipping discussion on strong/weak solutions).

Interpretation:

- u is Brownian time (quadratic-variation time).
- $\rho^m(u)$ is the corresponding calendar time.
- A^m_t converts calendar time back to Brownian time.

Exogenous Stopping

Proof

Couple stopping times for different m.

Define $\widetilde{\tau}^{m,C} := \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\}$, and $U^m := A^m_{\widetilde{\tau}^{m,C}}$.

Proof

Couple stopping times for different m.

Define $\widetilde{\tau}^{m,C} := \inf\{t > 0 : \widetilde{P}_t^m \notin C_t\}$, and $U^m := A^m_{\widetilde{\tau}^{m,C}}$.

Since $\rho^m(A_t^m) = t$, then $\rho^m(U^m) = \widetilde{\tau}^{m,C}$ and

$$\widetilde{\tau}^{m,C} = \inf\{t > 0 : \widetilde{P}_t^m \notin C_t\} = \inf\{t > 0 : q_{A_t^m} \notin C_t\}$$

$$= \inf\{\rho^m(A_t^m) > 0 : q_{A_t^m} \notin C_{\rho^m(A_t^m)}\} = \inf\{\rho^m(u) > 0 : q_u \notin C_{\rho^m(u)}\} = \rho^m(U^m).$$

Proof

Couple stopping times for different m.

Define $\widetilde{\tau}^{m,C} := \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\}$, and $U^m := A^m_{\widetilde{\tau}^{m,C}}$.

Since $\rho^m(A^m_t) = t$, then $\rho^m(U^m) = \widetilde{\tau}^{m,C}$ and

$$\widetilde{\tau}^{m,C} = \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\} = \inf\{t > 0 : q_{A^m_t} \notin C_t\}$$

$$= \inf\{\rho^m(A^m_t) > 0 : q_{A^m_t} \notin C_{\rho^m(A^m_t)}\} = \inf\{\rho^m(u) > 0 : q_u \notin C_{\rho^m(u)}\} = \rho^m(U^m).$$

$$\Longrightarrow U^m = \inf\{u > 0 : q_u \notin C_{\rho^m(u)}\}.$$

Intrinsic stopping time uses *same* process q and adjusts continuation region.

Proof

Couple stopping times for different m.

Define $\widetilde{\tau}^{m,C} := \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\}$, and $U^m := A^m_{\widetilde{\tau}^{m,C}}$.

Since $\rho^m(A^m_t) = t$, then $\rho^m(U^m) = \widetilde{\tau}^{m,C}$ and

$$\widetilde{\tau}^{m,C} = \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\} = \inf\{t > 0 : q_{A^m_t} \notin C_t\}$$

$$= \inf\{\rho^m(A^m_t) > 0 : q_{A^m_t} \notin C_{\rho^m(A^m_t)}\} = \inf\{\rho^m(u) > 0 : q_u \notin C_{\rho^m(u)}\} = \rho^m(U^m).$$

$$\Longrightarrow U^m = \inf\{u > 0 : q_u \notin C_{\rho^m(u)}\}.$$

Intrinsic stopping time uses *same* process q and adjusts continuation region.

Since $\widetilde{\tau}^{m,C} = \rho^m(U^m)$ and $\widetilde{P}^m_t = q_{A^m_t}$, then $\widetilde{P}^m_{\widetilde{\tau}^{m,C}} = q_{U^m}$!

Exogenous Stopping

Proof

Couple stopping times for different m.

Define $\widetilde{\tau}^{m,C} := \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\}$, and $U^m := A^m_{\widetilde{\tau}^{m,C}}$.

Since $\rho^m(A^m_t) = t$, then $\rho^m(U^m) = \widetilde{\tau}^{m,C}$ and

$$\widetilde{\tau}^{m,C} = \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\} = \inf\{t > 0 : q_{A^m_t} \notin C_t\}$$

$$= \inf\{\rho^m(A^m_t) > 0 : q_{A^m_t} \notin C_{\rho^m(A^m_t)}\} = \inf\{\rho^m(u) > 0 : q_u \notin C_{\rho^m(u)}\} = \rho^m(U^m).$$

$$\implies U^m = \inf\{u > 0 : q_u \notin C_{\rho^m(u)}\}.$$

Intrinsic stopping time uses *same* process q and adjusts continuation region.

Since $\widetilde{\tau}^{m,C} = \rho^m(U^m)$ and $\widetilde{P}^m_t = q_{A^m_t}$, then $\widetilde{P}^m_{\widetilde{\tau}^{m,C}} = q_{U^m}$!

Comparison principle.

We have $|\widetilde{P}^m_{\widetilde{\tau}^{m,C}} - \frac{1}{2}| = |q_{U^m} - \frac{1}{2}|$. All we need is $m' > m \implies U^{m'} \geq U^m$.

Using $t' > t \implies C_{t'} \subseteq C_t$, we just need that $\rho^{m'} \leq \rho^m$.

Proof

Couple stopping times for different m.

Define $\widetilde{\tau}^{m,C} := \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\}$, and $U^m := A^m_{\widetilde{\tau}^{m,C}}$.

Since $\rho^m(A^m_t) = t$, then $\rho^m(U^m) = \widetilde{\tau}^{m,C}$ and

$$\widetilde{\tau}^{m,C} = \inf\{t > 0 : \widetilde{P}^m_t \notin C_t\} = \inf\{t > 0 : q_{A^m_t} \notin C_t\}$$

$$= \inf\{\rho^m(A^m_t) > 0 : q_{A^m_t} \notin C_{\rho^m(A^m_t)}\} = \inf\{\rho^m(u) > 0 : q_u \notin C_{\rho^m(u)}\} = \rho^m(U^m).$$

$$\implies U^m = \inf\{u > 0 : q_u \notin C_{\rho^m(u)}\}.$$

Intrinsic stopping time uses *same* process q and adjusts continuation region.

Since $\widetilde{\tau}^{m,C} = \rho^m(U^m)$ and $\widetilde{P}^m_t = q_{A^m_t}$, then $\widetilde{P}^m_{\widetilde{\tau}^{m,C}} = q_{U^m}$!

Comparison principle.

We have $|\widetilde{P}^m_{\widetilde{\tau}^{m,C}} - \frac{1}{2}| = |q_{U^m} - \frac{1}{2}|$. All we need is $m' > m \implies U^{m'} \geq U^m$.

Using $t' > t \implies C_{t'} \subseteq C_t$, we just need that $\rho^{m'} \leq \rho^m$.

For this, we'll learn a new tool: *comparison principle*.

Comparison Principle

Let $F : [0, U] \times I \to \mathbb{R}$ be continuous and locally Lipschitz in their second argument, where $I \subseteq \mathbb{R}$ is an interval. Consider the ODE

$$x'(u) = F(u, x(u)), \quad x(0) = x_0. \tag{P}$$

Definition (Sub/Supersolution)

A differentiable function $\underline{x}$ is **subsolution** of (P) if $\underline{x}'(u) \leq F(u, \underline{x}(u))$.

A differentiable function $\overline{x}$ is **supersolution** of (P) if $\overline{x}'(u) \geq F(u, \overline{x}(u))$.

A differentiable function x is **solution** of (P) if it is both a sub- and a supersolution.

Comparison Principle

Let $F : [0, U] \times I \to \mathbb{R}$ be continuous and locally Lipschitz in their second argument, where $I \subseteq \mathbb{R}$ is an interval. Consider the ODE

$$x'(u) = F(u, x(u)), \quad x(0) = x_0. \tag{P}$$

Definition (Sub/Supersolution)

A differentiable function $\underline{x}$ is **subsolution** of (P) if $\underline{x}'(u) \leq F(u, \underline{x}(u))$.

A differentiable function $\overline{x}$ is **supersolution** of (P) if $\overline{x}'(u) \geq F(u, \overline{x}(u))$.

A differentiable function x is **solution** of (P) if it is both a sub- and a supersolution.

Theorem (Comparison Principle for ODEs)

If $\underline{x}$ is subsolution of (P), $\overline{x}$ is a supersolution of (P), and $\underline{x}(0) \leq x_0 \leq \overline{x}(0)$, then, on their common interval of existence,

$$\underline{x}(u) \leq x(u) \leq \overline{x}(u) \quad \text{for all } u.$$

Comparison Principle

Theorem (Comparison Principle for ODEs)

If $\underline{x}$ is subsolution of (P), $\overline{x}$ is a supersolution of (P), and $\underline{x}(\mathbf{0}) \leq x_{\mathbf{0}} \leq \overline{x}(\mathbf{0})$, then, on their common domain,

$$\underline{x}(u) \leq x(u) \leq \overline{x}(u) \quad \text{for all } u.$$

Intuition: Solutions, supersolutions, and subsolutions cannot cross: whenever they meet, subsolution/supersolution has weakly smaller/greater instantaneous growth.

Comparison Principle

Theorem (Comparison Principle for ODEs)

If $\underline{x}$ is subsolution of (P), $\overline{x}$ is a supersolution of (P), and $\underline{x}(\mathbf{0}) \leq x_{\mathbf{0}} \leq \overline{x}(\mathbf{0})$, then, on their common domain,

$$\underline{x}(u) \leq x(u) \leq \overline{x}(u) \quad \text{for all } u.$$

Intuition: Solutions, supersolutions, and subsolutions cannot cross: whenever they meet, subsolution/supersolution has weakly smaller/greater instantaneous growth.

Random ODEs: Same result applies pathwise when F and G depend on common random process, provided assumptions hold a.s.

Comparison Principle

Theorem (Comparison Principle for ODEs)

If $\underline{x}$ is subsolution of (P), $\overline{x}$ is a supersolution of (P), and $\underline{x}(\mathbf{0}) \leq x_{\mathbf{0}} \leq \overline{x}(\mathbf{0})$, then, on their common domain,

$$\underline{x}(u) \leq x(u) \leq \overline{x}(u) \quad \text{for all } u.$$

Intuition: Solutions, supersolutions, and subsolutions cannot cross: whenever they meet, subsolution/supersolution has weakly smaller/greater instantaneous growth.

Random ODEs: Same result applies pathwise when F and G depend on common random process, provided assumptions hold a.s.

Remark: Comparison principles are fundamental tools for ODEs/SDEs (we'll see some more later) and very powerful to get monotonicity results.

Comparison Principle

Theorem (Comparison Principle for ODEs)

If $\underline{x}$ is subsolution of (P), $\overline{x}$ is a supersolution of (P), and $\underline{x}(\mathbf{0}) \leq x_{\mathbf{0}} \leq \overline{x}(\mathbf{0})$, then, on their common domain,

$$\underline{x}(u) \leq x(u) \leq \overline{x}(u) \quad \text{for all } u.$$

Intuition: Solutions, supersolutions, and subsolutions cannot cross: whenever they meet, subsolution/supersolution has weakly smaller/greater instantaneous growth.

Random ODEs: Same result applies pathwise when F and G depend on common random process, provided assumptions hold a.s.

Remark: Comparison principles are fundamental tools for ODEs/SDEs (we'll see some more later) and very powerful to get monotonicity results.

Corollary

Let $F, G : [\mathbf{0}, U] \times I \to \mathbb{R}$ be continuous and locally Lipschitz in their second argument, where $I \subseteq \mathbb{R}$ is an interval. Suppose $x, y : [\mathbf{0}, U] \to I$ solve $x'(u) = F(u, x(u))$, and $y'(u) = G(u, y(u))$.
If (1) $x(\mathbf{0}) \leq y(\mathbf{0})$ and (2) $F(u,z) \leq G(u,z)\ \forall (u,z) \in [\mathbf{0}, U] \times I$, then $x(u) \leq y(u), \forall u \in [\mathbf{0}, U]$.

Exogenous Stopping

Proof

Comparison principle.

Fix $m' > m$. Since $\sigma(t,p,m') > \sigma(t,p,m) \iff 1/\sigma(t,p,m)^2 > 1/\sigma(t,p,m')^2$.

By comparison principle for random ODEs, $\rho^m(u) \geq \rho^{m'}(u), \forall u \geq 0$.

Exogenous Stopping

Proof

Comparison principle.

Fix $m' > m$. Since $\sigma(t,p,m') > \sigma(t,p,m) \iff 1/\sigma(t,p,m)^2 > 1/\sigma(t,p,m')^2$.

By comparison principle for random ODEs, $\rho^m(u) \geq \rho^{m'}(u), \forall u \geq 0$.

Ordering intrinsic stopping time.

C_t weakly decreasing in t wrt subset order $\implies C_{\rho^m(u)} \subseteq C_{\rho^{m'}(u)}$,

$\implies \{q_u \notin C_{\rho^{m'}(u)}\} \subseteq \{q_u \notin C_{\rho^m(u)}\}$,

$\implies U^m \leq U^{m'}$ a.s.

Exogenous Stopping

Proof

Comparison principle.

Fix $m' > m$. Since $\sigma(t,p,m') > \sigma(t,p,m) \iff 1/\sigma(t,p,m)^2 > 1/\sigma(t,p,m')^2$.

By comparison principle for random ODEs, $\rho^m(u) \geq \rho^{m'}(u), \forall u \geq 0$.

Ordering intrinsic stopping time.

C_t weakly decreasing in t wrt subset order $\implies C_{\rho^m(u)} \subseteq C_{\rho^{m'}(u)}$,

$\implies \{q_u \notin C_{\rho^{m'}(u)}\} \subseteq \{q_u \notin C_{\rho^m(u)}\}$,

$\implies U^m \leq U^{m'}$ a.s.

Optional stopping.

q martingale, $|\cdot - 1/2|$ convex, $|q_u - 1/2|$ is submartingale. $C_t \subseteq C_0$, $q_{U^m \wedge u}$ bounded.

Optional stopping: $\mathbb{E}[|q_{U^m} - 1/2|] \leq \mathbb{E}[|q_{U^{m'}} - 1/2|]$

$\implies \mathbb{P}(\operatorname{sign} \alpha_\tau = \operatorname{sign} \theta)$ increases in m.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.
Let $\overline{p}, \underline{p} : \mathbb{R}_+ \to (0,1)$ be $\mathcal{C}^1$ and satisfy: (a) $\overline{p}_t' \leq 0 \leq \underline{p}_t'$, (b) $1/2 \in (\underline{p}_t, \overline{p}_t)$ whenever non-empty. Define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = \text{sign}(p_t - 1/2)$.
Suppose, for $\underline{m} > 0$, (c) $-\rho(\underline{p}_t)\overline{p}_t', \rho(\overline{p}_t)\underline{p}_t' \leq 1$, where $\rho(q) := \int_{\underline{p}_t}^{\overline{p}_t} 2\frac{p-q}{\overline{p}_t - \underline{p}_t} \frac{1}{\sigma(t,p,\underline{m})^2}\, dp$ and $\sigma(t,p,m) := 2mp(1-p)$.
Then, (1) $\mathbb{E}[\tau]$ decreases in $m \geq \underline{m}$, and (2) $\mathbb{P}(\alpha_\tau = \text{sign}(\theta))$ increases in m.

Just restating the result.

Comments:

Substance: Under some general-enough conditions, stopping time decreases with m and accuracy increases with m.

More complex $\implies$ slower and worse.

Technique: Along the way, learned:

how to use DDS and random clock changes to couple processes, and

Exogenous Stopping

After quick introspection, you may wonder:

Why are we keeping stopping rule fixed in belief space (which adjust to varying m) and not in evidence space?

Note: since $\text{logit}\, p_t = \text{logit}\, p_0 + 2mX_t$, fixed thresholds in belief space $\overline{p}_t$ imply that stopping thresholds in evidence space shrink as m increases.

Unclear which is more natural comparison.

Does it change anything? Do the results extend?

Exogenous Stopping

After quick introspection, you may wonder:

Why are we keeping stopping rule fixed in belief space (which adjust to varying m) and not in evidence space?

Note: since $\text{logit}\, p_t = \text{logit}\, p_0 + 2mX_t$, fixed thresholds in belief space $\overline{p}_t$ imply that stopping thresholds in evidence space shrink as m increases.

Unclear which is more natural comparison.

Does it change anything? Do the results extend?

Yes, but, in both cases, need to break eggs to make omelettes. Assumptions are general enough that they apply to different problems and non-nested.

Importantly, comparative statics for exog. stopping in evidence space connects cleanly to comparative statics under uncertainty and misspecification.

Proofs are simple enough and use some nice tricks.

Exogenous Stopping

Start by looking at what happens with Brownian motion at a fixed drift rate.

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ increases in $|m|$ for all $t \geq 0$.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ increases in $|m|$ for all $t \geq 0$.

Proof Sketch

Monotonicity on a grid:

Fix any $T > 0$ and partition $0 = t_0 < t_1 < \cdots < t_n = T$, and set $T^n := (t_i)_{i=1,\ldots,n}$.

Define event $E_t^m := \{X_t^m \in (-\overline{x}_t, \overline{x}_t)\}$ and $E^{m,n} := \cap_{t_i \in T^n} E_{t_i}^m$.

Let $B^n := (B_{t_i})_{i=1,\ldots,n}$; note $B^n \sim N(0, \Sigma^n)$. Denote its measure by γ_{Σ^n}.

Let $C^n := \Pi_{i=1}^n (-\overline{x}_{t_i}, \overline{x}_{t_i})$, k-cell symmetric about zero.

Grid approximation: $f_T^n(m) := \mathbb{P}(E^{m,n}) = \mathbb{P}(B^n \in C^n - mT^n) = \int_{\mathbb{R}^n} 1_{C^n}(x + mT^n)\varphi_{\Sigma^n}(x)\,dx$, where φ_{Σ^n} denotes density of B^n and $1_A(\cdot) := 1\{\cdot \in A\}$.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ increases in $|m|$ for all $t \geq 0$.

Proof Sketch

Monotonicity on a grid:

Grid approximation: $f_T^n(m) := \mathbb{P}(E^{m,n}) = \mathbb{P}(B^n \in C^n - mT^n) = \int_{\mathbb{R}^n} 1_{C^n}(x + mT^n)\varphi_{\Sigma^n}(x)\,dx$.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ increases in $|m|$ for all $t \geq 0$.

Proof Sketch

Monotonicity on a grid:

Grid approximation: $f_T^n(m) := \mathbb{P}(E^{m,n}) = \mathbb{P}(B^n \in C^n - mT^n) = \int_{\mathbb{R}^n} 1_{C^n}(x + mT^n)\varphi_{\Sigma^n}(x)\,dx$.

Log-concavity: C^n convex.

$1_{C^n}(x + mT^n)$ log-concave in (x, m) in extended sense.

Gaussian density strictly log-concave.

By Prékopa-Leindler inequality, f_T^n is log-concave in m.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ increases in $|m|$ for all $t \geq 0$.

Proof Sketch

Monotonicity on a grid:

Grid approximation: $f_T^n(m) := \mathbb{P}(E^{m,n}) = \mathbb{P}(B^n \in C^n - mT^n) = \int_{\mathbb{R}^n} 1_{C^n}(x + mT^n)\varphi_{\Sigma^n}(x)\,dx$.

Log-concavity: C^n convex.

$1_{C^n}(x + mT^n)$ log-concave in (x, m) in extended sense.

Gaussian density strictly log-concave.

By Prékopa-Leindler inequality, f_T^n is log-concave in m.

Monotonicity: By Brownian reflection, $f_T^n(m) = f_T^n(-m)$.

Even and log-concave function is single-peaked with max at 0.

$\implies$ f_T^n is non-increasing in $m \geq 0$.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ increases in $|m|$ for all $t \geq 0$.

Proof Sketch

Monotonicity on a grid:

Grid approximation: $f_T^n(m) := \mathbb{P}(E^{m,n}) = \mathbb{P}(B^n \in C^n - mT^n) = \int_{\mathbb{R}^n} 1_{C^n}(x + mT^n)\varphi_{\Sigma^n}(x)\,dx$.

Log-concavity: C^n convex.

$1_{C^n}(x + mT^n)$ log-concave in (x, m) in extended sense.

Gaussian density strictly log-concave.

By Prékopa-Leindler inequality, f_T^n is log-concave in m.

Monotonicity: By Brownian reflection, $f_T^n(m) = f_T^n(-m)$.

Even and log-concave function is single-peaked with max at 0.

$\implies$ f_T^n is non-increasing in $m \geq 0$.

Continuous-time limit: $f_T^n(m) \to f_T(m) := \mathbb{P}(\widehat{\tau}^m > T)$; weak monotonicity.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in m for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Suppose FOSD relation not strict. $\geq_{FOSD}$ partial order $\implies \exists m > m' > 0 : \widehat{\tau}^m \overset{d}{=} \widehat{\tau}^{m'}$.

By monotonicity, $\widehat{\tau}^m \overset{d}{=} \widehat{\tau}^{m''}$ $\forall m'' \in [m', m]$.

Take $T > 0$ such that $f_T(m) = \mathbb{P}(E^m) \geq \varepsilon > 0$, where $E^m := \{\widehat{\tau}^m > T\}$.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in m for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Suppose FOSD relation not strict. $\geq_{FOSD}$ partial order $\implies \exists m > m' > 0 : \widehat{\tau}^m \overset{d}{=} \widehat{\tau}^{m'}$.

By monotonicity, $\widehat{\tau}^m \overset{d}{=} \widehat{\tau}^{m''}$ $\forall m'' \in [m', m]$.

Take $T > 0$ such that $f_T(m) = \mathbb{P}(E^m) \geq \varepsilon > 0$, where $E^m := \{\widehat{\tau}^m > T\}$.

Girsanov's theorem: define $Z_t := e^{mB_t - \frac{1}{2}m^2 t}$ and $\mathbb{Q}^m$ s.t. $\frac{d\mathbb{Q}^m}{d\mathbb{P}}|_{\mathcal{F}_t} = Z_t$, then

$$\underbrace{B_t + mt}_{\text{under } \mathbb{P}} \overset{d}{=} \underbrace{B_t}_{\text{under } \mathbb{Q}^m}.$$

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in m for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Suppose FOSD relation not strict. $\geq_{FOSD}$ partial order $\implies \exists m > m' > 0 : \widehat{\tau}^m \overset{d}{=} \widehat{\tau}^{m'}$.

By monotonicity, $\widehat{\tau}^m \overset{d}{=} \widehat{\tau}^{m''}$ $\forall m'' \in [m', m]$.

Take $T > 0$ such that $f_T(m) = \mathbb{P}(E^m) \geq \varepsilon > 0$, where $E^m := \{\widehat{\tau}^m > T\}$.

Girsanov's theorem: define $Z_t := e^{mB_t - \frac{1}{2}m^2 t}$ and $\mathbb{Q}^m$ s.t. $\frac{\mathrm{d}\mathbb{Q}^m}{\mathrm{d}\mathbb{P}}|_{\mathcal{F}_t} = Z_t$, then

$$\underbrace{B_t + mt}_{\text{under } \mathbb{P}} \overset{d}{=} \underbrace{B_t}_{\text{under } \mathbb{Q}^m}.$$

$$\mathbb{P}(E^m) = \mathbb{Q}^m(E^0) = \mathbb{E}^{\mathbb{P}}\left[1_{E^0}\frac{d\mathbb{Q}^m}{d\mathbb{P}}\right] = \mathbb{E}^{\mathbb{P}}\left[1_{E^0}e^{mB_T - \frac{1}{2}m^2T}\right] = e^{-\frac{1}{2}m^2T}\,\mathbb{E}^{\mathbb{P}}\left[1_{E^0}e^{mB_T}\right].$$

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in $|m|$ for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Girsanov's theorem: $\mathbb{P}(E^m) = e^{-\frac{1}{2}m^2 T}\mathbb{E}^{\mathbb{P}}[1_{E^0} e^{mB_T}]$.

$\mathbb{E}^{\mathbb{P}}\left[1_{E^0} e^{mB_T}\right]$ is moment-generating function of B_T restricted to the event E^0;
it is a real-analytic function of m, and so it f_T.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\overline{x}_t, \overline{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in $|m|$ for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Girsanov's theorem: $\mathbb{P}(E^m) = e^{-\frac{1}{2}m^2 T}\mathbb{E}^{\mathbb{P}}[1_{E^0}e^{mB_T}]$.

$\mathbb{E}^{\mathbb{P}}\left[1_{E^0}e^{mB_T}\right]$ is moment-generating function of B_T restricted to the event E^0;
it is a real-analytic function of m, and so it f_T.

Identity theorem for real analytic functions: real-analytic functions constant on open interval are constant throughout domain $\implies$ f_T constant.

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\bar{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\bar{x}_t, \bar{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in $|m|$ for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Girsanov's theorem: $\mathbb{P}(E^m) = e^{-\frac{1}{2}m^2 T}\mathbb{E}^{\mathbb{P}}[1_{E^0} e^{mB_T}]$.

$\mathbb{E}^{\mathbb{P}}\left[1_{E^0} e^{mB_T}\right]$ is moment-generating function of B_T restricted to the event E^0; it is a real-analytic function of m, and so it f_T.

Identity theorem for real analytic functions: real-analytic functions constant on open interval are constant throughout domain $\implies$ f_T constant.

Contradiction:

$$0 < \varepsilon \leq f_T(m) \leq \mathbb{P}(X_T^m \in (-\bar{x}_T, \bar{x}_T))$$

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\bar{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\bar{x}_t, \bar{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in $|m|$ for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Girsanov's theorem: $\mathbb{P}(E^m) = e^{-\frac{1}{2}m^2 T}\mathbb{E}^{\mathbb{P}}[1_{E^0} e^{mB_T}]$.

$\mathbb{E}^{\mathbb{P}}\left[1_{E^0} e^{mB_T}\right]$ is moment-generating function of B_T restricted to the event E^0; it is a real-analytic function of m, and so it f_T.

Identity theorem for real analytic functions: real-analytic functions constant on open interval are constant throughout domain $\implies$ f_T constant.

Contradiction:

$$0 < \varepsilon \leq f_T(m) \leq \mathbb{P}(X_T^m \in (-\bar{x}_T, \bar{x}_T)) = \mathbb{P}(-\bar{x}_T - mT < B_T < \bar{x}_T - mT)$$

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\bar{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\bar{x}_t, \bar{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in $|m|$ for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Girsanov's theorem: $\mathbb{P}(E^m) = e^{-\frac{1}{2}m^2 T}\mathbb{E}^{\mathbb{P}}[1_{E^0} e^{mB_T}]$.

$\mathbb{E}^{\mathbb{P}}\left[1_{E^0} e^{mB_T}\right]$ is moment-generating function of B_T restricted to the event E^0; it is a real-analytic function of m, and so it f_T.

Identity theorem for real analytic functions: real-analytic functions constant on open interval are constant throughout domain $\implies$ f_T constant.

Contradiction:

$$0 < \varepsilon \leq f_T(m) \leq \mathbb{P}(X_T^m \in (-\bar{x}_T, \bar{x}_T)) = \mathbb{P}(-\bar{x}_T - mT < B_T < \bar{x}_T - mT)$$
$$= \Phi(\bar{x}_T/\sqrt{T} - m\sqrt{T}) - \Phi(-\bar{x}_T/\sqrt{T} - m\sqrt{T})$$

Exogenous Stopping

Proposition (Gonçalves (WP) SAC, Proposition 2)

Let $X_t^m := mt + B_t$, for a constant $m \in \mathbb{R}$, where B is Brownian motion. Suppose $\bar{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\widehat{\tau}^m := \inf\{t > 0 \mid X_t^m \notin (-\bar{x}_t, \bar{x}_t)\}$. Then, $\mathbb{P}(\widehat{\tau}^m \leq t)$ weakly increases in $|m|$ for all $t \geq 0$, strictly so at t such that $\mathbb{P}(\widehat{\tau}^m \leq t) < 1$.

Proof Sketch

Strict FOSD monotonicity:

Girsanov's theorem: $\mathbb{P}(E^m) = e^{-\frac{1}{2}m^2 T}\mathbb{E}^{\mathbb{P}}[1_{E^0} e^{mB_T}]$.

$\mathbb{E}^{\mathbb{P}}\left[1_{E^0} e^{mB_T}\right]$ is moment-generating function of B_T restricted to the event E^0; it is a real-analytic function of m, and so it f_T.

Identity theorem for real analytic functions: real-analytic functions constant on open interval are constant throughout domain $\implies$ f_T constant.

Contradiction:

$$0 < \varepsilon \leq f_T(m) \leq \mathbb{P}(X_T^m \in (-\bar{x}_T, \bar{x}_T)) = \mathbb{P}(-\bar{x}_T - mT < B_T < \bar{x}_T - mT)$$

$$= \Phi(\bar{x}_T/\sqrt{T} - m\sqrt{T}) - \Phi(-\bar{x}_T/\sqrt{T} - m\sqrt{T}) \to 0 \text{ as } m \to \infty$$

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

Let $\theta = \pm 1$, $M \geq 0$, and $X_t := M\theta t + B_t$, where B is a Brownian motion. Let B, θ, M be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$, and $M \sim \mu$, $\mu(\mathbb{R}_{++}) > 0$.
Suppose $\overline{x} \in \mathcal{C}_0(\mathbb{R}_+)$ and $\overline{p}_t := \pi(t, \overline{x}_t), \underline{p}_t := \pi(t, -\overline{x}_t)$ and define $\tau := \inf\{t > 0 \mid p_t \notin (\underline{p}_t, \overline{p}_t)\}$ and $\alpha_t = 1$ if $p_t \geq \overline{p}_t$ and $= -1$ if ow.
Then, (1) $\mathbb{P}(\tau \leq t \mid M = m)$ increases in $m \geq 0$ for any t, and (2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Uncertainty allows us to hold fixed how DM's beliefs depend on evidence despite change in drift rate.

Note: beliefs are martingale wrt $\mathbb{P}$ but not wrt $\mathbb{P}(\cdot \mid M = m)$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(1) $\mathbb{P}(\tau \leq t \mid M = m)$ increases in m for any t.

Proof

(1) Follows as corollary:
$\mathbb{P}(\tau > T \mid M = m, \theta) = \mathbb{P}(p_t \in (\underline{p}_t, \overline{p}_t), t \leq T \mid M = m, \theta) = \mathbb{P}(X_t \in (-\overline{x}_t, \overline{x}_t), t \leq T \mid M = m, \theta)$,
is decreasing in m for any $\theta \in \{-1, +1\}$.
Hence, so is $\mathbb{P}(\tau > T \mid M = m)$.

This hinges on symmetry of $\chi(t, \overline{p}_t) = -\chi(t, \underline{p}_t)$. Without it, even $\mathbb{E}[\tau \mid M = m]$ is not necessarily monotone in m (counterexamples exist).

Other sufficient condition (non-nested): $\overline{x}_t$ and $\underline{x}_t$ are constant in t and $p_0 = 1/2$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Coupling processes. Write $X_t^m := mt + B_t$ and $\tau^m := \inf\{t > 0 \mid X_t^m \notin (\underline{x}_t, \overline{x}_t)\}$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Coupling processes. Write $X_t^m := mt + B_t$ and $\tau^m := \inf\{t > 0 \mid X_t^m \notin (\underline{x}_t, \overline{x}_t)\}$.

Reduce to Brownian upper exit. Conditional on $\theta = 1$:

$$\mathbb{P}(\alpha_\tau = 1 \mid \theta = 1, M = m) = \mathbb{P}(p_\tau = \overline{p}_\tau \mid \theta = 1, M = m) = \mathbb{P}(X_\tau = \overline{x}_\tau \mid \theta = 1, M = m) = \mathbb{P}(X_{\tau^m}^m = \overline{x}_{\tau^m}).$$

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Coupling processes. Write $X_t^m := mt + B_t$ and $\tau^m := \inf\{t > 0 \mid X_t^m \notin (\underline{x}_t, \overline{x}_t)\}$.

Reduce to Brownian upper exit. Conditional on $\theta = 1$:

$\mathbb{P}(\alpha_\tau = 1 \mid \theta = 1, M = m) = \mathbb{P}(p_\tau = \overline{p}_\tau \mid \theta = 1, M = m) = \mathbb{P}(X_\tau = \overline{x}_\tau \mid \theta = 1, M = m) = \mathbb{P}(X_{\tau^m}^m = \overline{x}_{\tau^m})$.

Pathwise dominance implies ordered upper exit. For $m' > m$, conditional on $\theta = 1$, $X_t^{m'} = X_t^m + (m' - m)t \geq X_t^m$, pathwise.

Hence, $X_t^{m'} \leq \underline{x}_t \implies X_t^m \leq \underline{x}_t$ and $X_t^m \geq \overline{x}_t \implies X_t^{m'} \geq \overline{x}_t$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Coupling processes. Write $X_t^m := mt + B_t$ and $\tau^m := \inf\{t > 0 \mid X_t^m \notin (\underline{x}_t, \overline{x}_t)\}$.

Reduce to Brownian upper exit. Conditional on $\theta = 1$:

$$\mathbb{P}(\alpha_\tau = 1 \mid \theta = 1, M = m) = \mathbb{P}(p_\tau = \overline{p}_\tau \mid \theta = 1, M = m) = \mathbb{P}(X_\tau = \overline{x}_\tau \mid \theta = 1, M = m) = \mathbb{P}(X_{\tau^m}^m = \overline{x}_{\tau^m}).$$

Pathwise dominance implies ordered upper exit. For $m' > m$, conditional on $\theta = 1$,

$X_t^{m'} = X_t^m + (m' - m)t \geq X_t^m$, pathwise.

Hence, $X_t^{m'} \leq \underline{x}_t \implies X_t^m \leq \underline{x}_t$ and $X_t^m \geq \overline{x}_t \implies X_t^{m'} \geq \overline{x}_t$.

Hence, $\{X_t^m \geq \overline{x}_t \text{ and } X_s^m > \underline{x}_s, s \leq t\} \subset \{X_t^{m'} \geq \overline{x}_t \text{ and } X_s^{m'} > \underline{x}_s, s \leq t\}, \forall t \geq 0$

$\implies \mathbb{P}(X_{\tau^m}^m = \overline{x}_{\tau^m} \text{ and } \tau^m \leq t) \leq \mathbb{P}(X_{\tau^{m'}}^{m'} = \overline{x}_{\tau^{m'}} \text{ and } \tau^{m'} \leq t), \forall t \geq 0$

$\implies \mathbb{P}(X_{\tau^m}^m = \overline{x}_{\tau^m}) \leq \mathbb{P}(X_{\tau^{m'}}^{m'} = \overline{x}_{\tau^{m'}})$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Symmetric argument. Conditional on $\theta = -1$, $X_t^{m'} = X_t^m - (m' - m)t \leq X_t^m$, pathwise, and $\mathbb{P}(\alpha_\tau = -1 \mid \theta = -1, M = m) = \mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m})$.

Symmetric argument yields $\mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m}) \leq \mathbb{P}(X_{\tau^{m'}}^{m'} = \underline{x}_{\tau^{m'}})$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Symmetric argument. Conditional on $\theta = -1$, $X_t^{m'} = X_t^m - (m' - m)t \leq X_t^m$, pathwise, and $\mathbb{P}(\alpha_\tau = -1 \mid \theta = -1, M = m) = \mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m})$.
Symmetric argument yields $\mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m}) \leq \mathbb{P}(X_{\tau^{m'}}^{m'} = \underline{x}_{\tau^{m'}})$.

Monotonicity. Since $\mathbb{P}(\alpha_\tau = \theta \mid M = m, \theta) \leq \mathbb{P}(\alpha_\tau = \theta \mid M = m', \theta)$ for $\theta = \pm 1$, conclude $\mathbb{P}(\alpha_\tau = \theta \mid M = m) \leq \mathbb{P}(\alpha_\tau = \theta \mid M = m')$.

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Symmetric argument. Conditional on $\theta = -1$, $X_t^{m'} = X_t^m - (m' - m)t \leq X_t^m$, pathwise, and $\mathbb{P}(\alpha_\tau = -1 \mid \theta = -1, M = m) = \mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m})$.
Symmetric argument yields $\mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m}) \leq \mathbb{P}(X_{\tau^{m'}}^{m'} = \underline{x}_{\tau^{m'}})$.

Monotonicity. Since $\mathbb{P}(\alpha_\tau = \theta \mid M = m, \theta) \leq \mathbb{P}(\alpha_\tau = \theta \mid M = m', \theta)$ for $\theta = \pm 1$, conclude $\mathbb{P}(\alpha_\tau = \theta \mid M = m) \leq \mathbb{P}(\alpha_\tau = \theta \mid M = m')$.

Strictness. Inequality is strict whenever $x_0 > 0$.
(Idea is to construct path with exit at some T through correct boundary under m' and incorrect boundary with m, argue similar conclusions hold for nearby paths, and appeal to positive measure of such paths.)

Exogenous Stopping

Theorem (Gonçalves (WP) SAC, Theorem 1)

(2) $\mathbb{P}(\alpha_\tau = \theta \mid M = m)$ increases in m.

Proof

Symmetric argument. Conditional on $\theta = -1$, $X_t^{m'} = X_t^m - (m' - m)t \leq X_t^m$, pathwise, and $\mathbb{P}(\alpha_\tau = -1 \mid \theta = -1, M = m) = \mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m})$.
Symmetric argument yields $\mathbb{P}(X_{\tau^m}^m = \underline{x}_{\tau^m}) \leq \mathbb{P}(X_{\tau^{m'}}^{m'} = \underline{x}_{\tau^{m'}})$.

Monotonicity. Since $\mathbb{P}(\alpha_\tau = \theta \mid M = m, \theta) \leq \mathbb{P}(\alpha_\tau = \theta \mid M = m', \theta)$ for $\theta = \pm 1$, conclude $\mathbb{P}(\alpha_\tau = \theta \mid M = m) \leq \mathbb{P}(\alpha_\tau = \theta \mid M = m')$.

Strictness. Inequality is strict whenever $x_0 > 0$.
(Idea is to construct path with exit at some T through correct boundary under m' and incorrect boundary with m, argue similar conclusions hold for nearby paths, and appeal to positive measure of such paths.)

Note that proof holds for arbitrary continuous $\overline{x} \geq \underline{x}$, with $\underline{p}_t := \pi(t, \underline{x}_t)$ and $\overline{p}_t := \pi(t, \overline{x}_t)$.

Taking Stock

With exogenous stopping thresholds, under some general conditions,

Lower m / More complex problems $\implies$ Slower and worse choices.

This matches common intuition. However, it does rely on stopping not being 'adjusted' to problem complexity (m).

Next: we'll see

(1) How exogenous stopping + varying the drift $\equiv$ optimal stopping under drift uncertainty + varying the realised drift;

(2) How optimal stopping + varying the drift / prior distribution gives rise to different comparative statics.

Overview

1. A Pioneer from Cognitive Sciences

2. Speed, Accuracy, and Complexity

3. Evidence and Beliefs

4. Exogenous Stopping and Problem Complexity

5. Optimal Stopping
   - Properties of the Value Function
   - Optimal Stopping and Problem Complexity
   - Reconciling Empirical Evidence

6. Comparative Statics on Incentives and Revealing Complexity

7. Fast and Slow Errors

Value of Learning

Scrap value: $v(p) := \max_{\alpha} u(\alpha, p)$.

Value function:

$$V(p) := \sup_{\tau} \mathbb{E}_p[v(p_\tau) - c\tau].$$

Stopping immediately is feasible; hence $V \geq v$.

Value of Learning

Scrap value: $v(p) := \max_\alpha u(\alpha, p)$.

Value function:

$$V(p) := \sup_\tau \mathbb{E}_p[v(p_\tau) - c\tau].$$

Stopping immediately is feasible; hence $V \geq v$.

Optimal stopping: $\tau := \inf\{t > 0 \mid V(p_t) \leq v(p_t)\}$.

Continuation region: $C := \{V > v\}$.

Indifference: $\tilde{p} : u(1, \tilde{p}) = u(-1, \tilde{p})$. Relative incentives.

Stakes: $\delta := \sum_\theta \sum_{\alpha'} |\max_\alpha u(\alpha, \theta) - u(\alpha', \theta)|$. Absolute incentives.

We have lots of properties that hold generally, even for time inhomogeneous problems.

Properties of Optimal Stopping

Proposition (Gonçalves (WP) SAC, Propositions 7-8+)

Let $\sigma_\mu(t,p)$ satisfy (P1)-(P5). The following properties hold:

- V is convex in p. (General for optimal learning.)
- V is weakly decreasing in t. (Strictly so if σ_μ if μ non-degenerate.)
- V is $\mathcal{C}^{1,2}$ in (t,p) on $\{(t,p) \mid V > v\}$. (HJB holds in interior.)
- V is $\mathcal{C}^1$ in p. (Smooth fit in space.)
- V is $\mathcal{C}^1$ in t on closure of $\{(t,p) \mid V > v\}$. (Smooth fit in time.)
- Continuation region C is connected. (Stopping given by thresholds.)
- $\forall t \geq 0$, $(t,\tilde{p})$ is in $\{V > v\}$. (Always continue at indifference.)
- $\exists \overline{p}, \underline{p} \in \mathcal{C}^1(\mathbb{R}_+) : (t,p) \in C \iff p \in (\underline{p}_t, \overline{p}_t); \overline{p}' \leq 0 \leq \underline{p}'$. (Smooth thresholds.)
- If $\sigma_\mu \leq \sigma_{\mu'}$, then $V^\mu \leq V^{\mu'}$ and $C^\mu \subseteq C^{\mu'}$. (Option value.)

Hold in general learning problems of our class (v convex, piecewise linear).

Results extend to discounting.

Optimal Stopping

Back to our problem:

Recall: $\sigma_m(p) = 2mp(1-p)$. Time homogeneous markovian problem $\Longrightarrow$ time-invariant continuation region $(\underline{p}, \overline{p})$.

Variational Inequality: $0 = \max\{v(p) - V(p), \mathcal{L}V + f\}$.

HJB: $\mathcal{L}V + f = -c + 2m^2p^2(1-p)^2V''(p) = 0$.

Solve HJB in interior. Boundary condition ($V = v$ on $\{\underline{p}, \overline{p}\}$) and smooth fit ($V' = v'$ on $\{\underline{p}, \overline{p}\}$) pin down $(\underline{p}, \overline{p})$.

$$2m^2\delta/c = (A(\overline{p}) + B(\overline{p})) - (A(\underline{p}) + B(\underline{p})) \tag{1}$$

$$\tilde{p} = [C(\overline{p}) - C(\underline{p})]/[(A(\overline{p}) + B(\overline{p})) - (A(\underline{p}) + B(\underline{p}))] \tag{2}$$

$A(p) := \exp(\text{logit}(p)) - \exp(-\text{logit}(p))$,
$B(p) := 2\,\text{logit}(p)$, and
$C(p) := \text{logit}(p) + \exp(\text{logit}(p))$.

Optimal Stopping

Theorem (Gonçalves (WP) SAC, Theorem 2, Proposition 9)

Let $\theta = \pm 1$, $m \geq 0$, and $X_t := m\theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$.
Then, (1) $\mathbb{E}[\tau]$ is non-monotone and quasiconcave in m, and (2) $\mathbb{P}(\alpha_\tau = \theta)$ increases in m when $p_0 = 1/2$ or $V(1/2) > v(1/2)$.

Proof Intuition

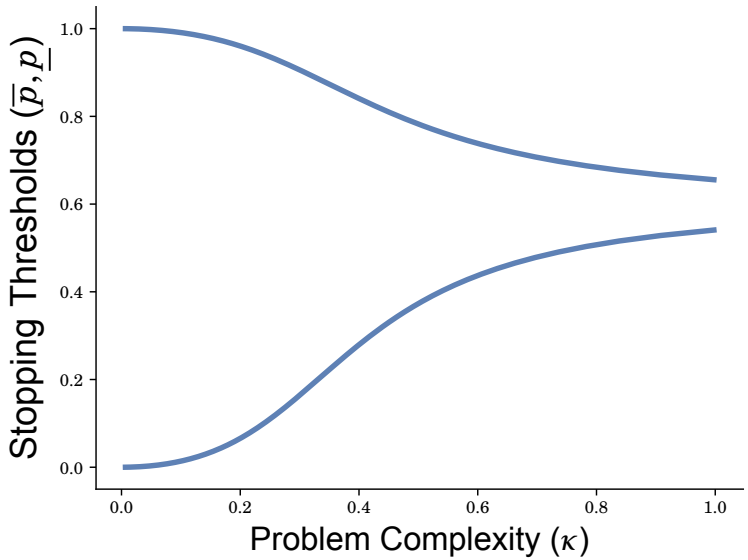
$\mathbb{E}[\tau]$: scale functions + implicit function theorem.

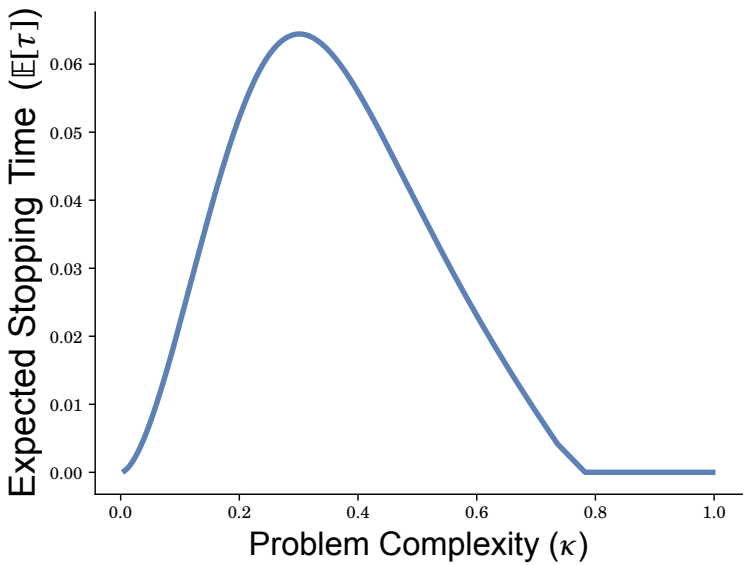
$\mathbb{P}(\alpha_\tau = \theta)$: boundary expansion + $p_0 = 1/2 \implies$ Blackwell-more information.

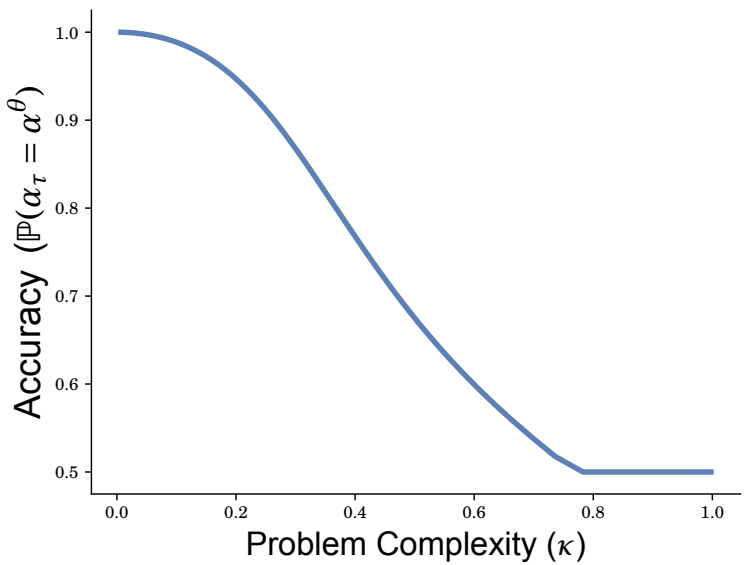
Under optimal stopping, stopping time becomes non-monotone.

Conclusion generalises to uncertain drift $M \sim \mu$: under conditions identified earlier for μ, comparative statics under prior $\mu \downarrow 0$ and $\mu \uparrow \infty$.

Speed-Accuracy Non-Monotonicity under Optimal Stopping







$\kappa = m^{-1}$.

Speed-Accuracy in Related Models

Uncertain Complexity.

DM's uncertain about problem complexity. Even if optimal given prior, stopping thresholds as if exogenous relative to *realised* complexity.

1. Results for *exogenous* stopping $\equiv$ Results wrt *realised complexity*.
   - *Unconditional* accuracy $\downarrow$ and stopping time $\uparrow$ in realised complexity.

Speed-Accuracy in Related Models

Uncertain Complexity.

DM's uncertain about problem complexity. Even if optimal given prior, stopping thresholds as if exogenous relative to *realised* complexity.

1. Results for *exogenous* stopping $\equiv$ Results wrt *realised complexity*.
   – *Unconditional* accuracy $\downarrow$ and stopping time $\uparrow$ in realised complexity.

2. Results for *optimal* stopping $\equiv$ Results wrt *prior distribution*.
   – *Unconditional* accuracy $\downarrow$ and stopping time $\nearrow\searrow$ in complexity distribution.
   (Think: realised complexity = σ/μ + noise).

Reconciling Empirical Evidence

Speed-Complexity Monotonicity.

DM cannot adjust to problem complexity independently of learning the optimal solution: as if exogenous stopping.

Computational/numerical problems: identify larger numerical value of numbers/sums (Moyer Landauer 67; Buckley Gillman 74; Krajcsi et al. 16).

Perception problems: identify dominant colour/movement in set of items (Linde Paivio 79; Maanen et al. 11).

Smaller difference between alternatives makes it more complex; but indistinguishable from learning optimal solution.

More complex entails slower and worse.

Reconciling Empirical Evidence

Speed-Complexity Monotonicity.

DM *cannot* adjust to problem complexity independently of learning optimal solution: as if exogenous stopping.

Speed-Complexity Non-Monotonicity.

Problem complexity separable from determinants of optimal action.

DM *can* adjust to problem complexity independently of learning optimal solution: optimal stopping.

E.g., number summands or size of set considered (Gonçalves Nunnari Zarate-Pina WP).

Also: nature of T/F questions (Wright Ayton 88) and features of lottery choice problems and contingent reasoning problems (Agranov Schotter Trevino WP).

Overview

1. A Pioneer from Cognitive Sciences
2. Speed, Accuracy, and Complexity
3. Evidence and Beliefs
4. Exogenous Stopping and Problem Complexity
5. Optimal Stopping
6. Comparative Statics on Incentives and Revealing Complexity
7. Fast and Slow Errors

Properties of Optimal Stopping

Proposition (Gonçalves (WP) SAC, Propositions 7-8+)

Let $\sigma_\mu(t,p)$ satisfy (P1)-(P5). The following properties hold:

- $\overline{p}_t, \underline{p}_t$ are strictly increasing in $\tilde{p}$. (Comparative statics in payoffs.)
- $(\underline{p}_t, \overline{p}_t)$ is strictly decreasing in c wrt $\subseteq$ order. (Comparative statics in costs.)

Corollary

Let $\sigma_\mu(t,p)$ satisfy (P1)-(P5); $dp_t = \sigma_\mu(t,p)\,dB_t$; $v(p) := \max_{\alpha\in\{-1,1\}} u(\alpha,p)$. Suppose τ is an optimal stopping time for $V(p) := \sup_\tau \mathbb{E}_p[v(p_\tau) - c\tau]$.
Then, for any $t \geq 0$,
(1) $\mathbb{P}(\alpha = 1, \tau \leq t)$ increases and $\mathbb{P}(\alpha = -1, \tau \leq t)$ decreases in $\tilde{p}$; and
(2) $\mathbb{P}(\tau \leq t)$ increases and $\mathbb{P}(\alpha_\tau = \theta)$ decreases in c.

Hold in general learning problems of our class.

Results extend to discounting.

(Discrete-time version in Gonçalves (WP) *Sequential Sampling Equilibrium*.)

Inferring Complexity

Back to inferring complexity...

Complexity: cognition/info processing underlying several 'biases' in behaviour.

Inferring Complexity

Back to inferring complexity...

Complexity: cognition/info processing underlying several 'biases' in behaviour.

There is a correct answer (and analyst knows it):

Choices tend to identify complexity (even in heterog. populations).

(Avg) Accuracy decreasing in complexity.

Inferring Complexity

Back to inferring complexity...

Complexity: cognition/info processing underlying several 'biases' in behaviour.

There is a correct answer (and analyst knows it):

Choices tend to identify complexity (even in heterog. populations).

(Avg) Accuracy decreasing in complexity.

There is no correct answer/Analyst doesn't know it:

Response time often used as proxy, but not always appropriate.

Non-monotone relationship creates inference problem.

Going to higher moments doesn't help.

Heterogeneous population: not even quasiconcave.

What to do?

Relative Incentives and Complexity

Subsidies and Relative Incentives: subsidise action $\mathbf{1}$ by s: $u(\mathbf{1}, \theta) + s$.

Affects relative incentives (indifference threshold),
not absolute incentives ($|u(\mathbf{1}, \theta) - u(\mathbf{1}, \hat{\theta})|$).

Makes DM more likely to choose $\mathbf{1}$ and do so faster
and less likely to choose $-\mathbf{1}$ and do so slower.

Pushes both stopping thresholds down.

Relative Incentives and Complexity

Subsidies and Relative Incentives: subsidise action $\mathbf{1}$ by s: $u(\mathbf{1}, \theta) + s$.

Affects relative incentives (indifference threshold),
not absolute incentives ($|u(\mathbf{1}, \theta) - u(\mathbf{1}, \hat{\theta})|$).

Makes DM more likely to choose $\mathbf{1}$ and do so faster
and less likely to choose $-\mathbf{1}$ and do so slower.

Pushes both stopping thresholds down.

Idea: look at how complexity and ability impact differential effect of relative incentives on choices.

Relative Incentives and Complexity

Subsidies and Relative Incentives: subsidise action 1 by s: $u(1, \theta) + s$.

Affects relative incentives (indifference threshold),
not absolute incentives ($|u(1, \theta) - u(1, \hat{\theta})|$).

Makes DM more likely to choose 1 and do so faster
and less likely to choose -1 and do so slower.

Pushes both stopping thresholds down.

Idea: look at how complexity and ability impact differential effect of relative incentives on choices.

Subsidies more effective if have greater effect on choices.

$\frac{\partial}{\partial s}\mathbb{P}(\alpha_\tau = 1)$, when $\mathbb{P}(\alpha_\tau = 1) \in (0, 1)$

Theorem 4

Subsidies are more effective in more complex problems.

Proof intuition

Subsidy s to action $\mathbf{1}$ lowers indifference belief $\tilde{p}$ but leaves stakes δ fixed.

Shifts both optimal thresholds upward.

Optional sampling gives: $\mathbb{P}(\text{Choose } \mathbf{1}) = \frac{p_0 - \underline{p}}{\overline{p} - \underline{p}}$.

Use equations characterising thresholds to show $\frac{\partial^2}{\partial \kappa \partial s} \mathbb{P}(\text{Choose } \mathbf{1}) > 0$ on $\tau > 0$.

Subsidy effectiveness decreases in m.

Result extends to case in which ability $\equiv$ effective learning rate.

Revealing Complexity and Ability

Theorem (Gonçalves (WP) SAC, Theorem 4)

Subsidies are more effective in more complex problems.

Intuition: 1 penny more for action **1** pushes stopping thresholds down.

If problem simple enough, then close to sure, very wide continuation region; won't affect choice prob. a lot.

If problem very hard, narrow continuation region; affects choice prob. much more.

Revealing Complexity and Ability

Theorem (Gonçalves (WP) SAC, Theorem 4)

Subsidies are more effective in more complex problems.

Intuition: 1 penny more for action **1** pushes stopping thresholds down.

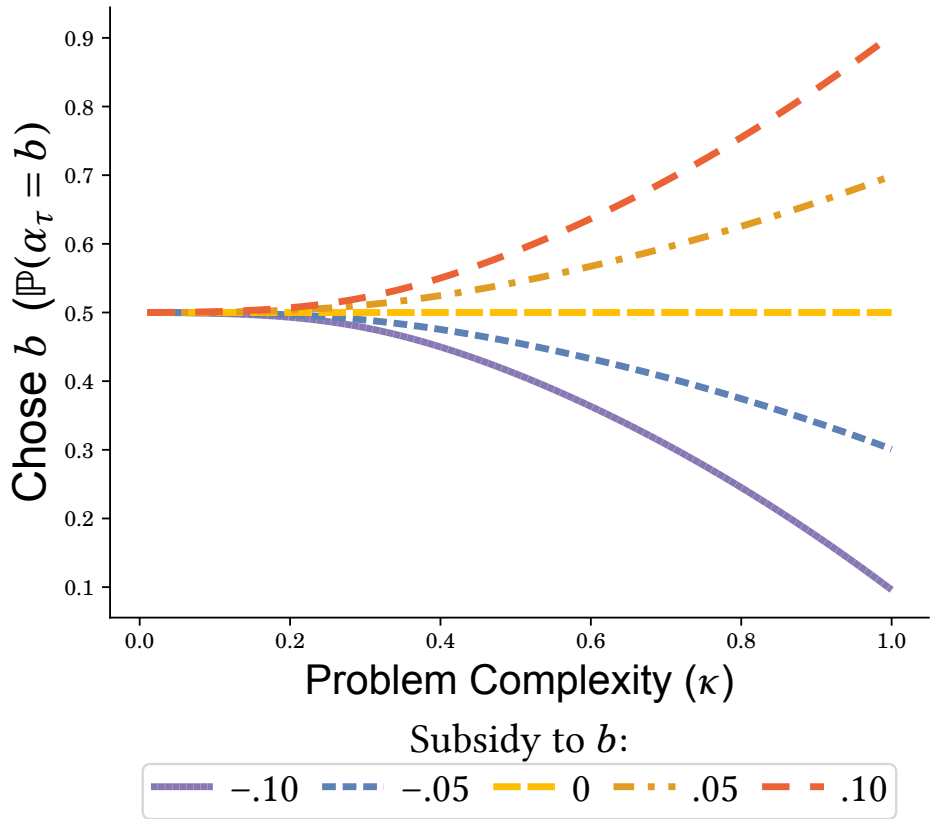
If problem simple enough, then close to sure, very wide continuation region; won't affect choice prob. a lot.

If problem very hard, narrow continuation region; affects choice prob. much more.

Limitations: hold fixed (distribution of) absolute incentive levels / stakes fixed.

Holds beyond DDM framework — e.g., Shannon costs — but not in general UPS costly info acquisition.

Revealing Complexity and Ability



Overview

1. A Pioneer from Cognitive Sciences
2. Speed, Accuracy, and Complexity
3. Evidence and Beliefs
4. Exogenous Stopping and Problem Complexity
5. Optimal Stopping
6. Comparative Statics on Incentives and Revealing Complexity
7. Fast and Slow Errors

Fast and Slow Errors

Comparative Statics Across Problems.

Results thus far: Changing drift rate, changing prior distribution, etc.

Nuanced conclusions: Across problems, time is somewhat informative of problem complexity and accuracy, but not when DM freely adjusts to it.

Fast and Slow Errors

Comparative Statics Across Problems.

Results thus far: Changing drift rate, changing prior distribution, etc.

Nuanced conclusions: Across problems, time is somewhat informative of problem complexity and accuracy, but not when DM freely adjusts to it.

Comparative Statics Within Problem.

Holding problem fixed, is time be informative of choice quality? Are errors fast or slow?

Fast and Slow Errors

Comparative Statics Across Problems.

Results thus far: Changing drift rate, changing prior distribution, etc.

Nuanced conclusions: Across problems, time is somewhat informative of problem complexity and accuracy, but not when DM freely adjusts to it.

Comparative Statics Within Problem.

Holding problem fixed, is time be informative of choice quality? Are errors fast or slow?

Focus: Fudenberg, Strack, and Strzalecki (2018 AER) and Cripps and Gonçalves (WP).

Fast and Slow Errors

Comparative Statics Within Problem.

Holding problem fixed, is time be informative of choice quality? Are errors fast or slow?

Focus: Fudenberg, Strack, and Strzalecki (2018 AER) and Cripps and Gonçalves (WP).

No time this year...

Fast and Slow Errors

Comparative Statics Within Problem.

Holding problem fixed, is time be informative of choice quality? Are errors fast or slow?

Focus: Fudenberg, Strack, and Strzalecki (2018 AER) and Cripps and Gonçalves (WP).

No time this year...

Highlights:

FFS (1): $\theta = \pm 1$, $p_0 = \tilde{p} = 1/2$, fixed m. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ constant in t.

(2): $\theta = \pm 1$, $p_0 = \tilde{p} = 1/2$, $m = |\varepsilon|$, $\varepsilon \sim N(0,1)$. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ decreasing in t.

Fast and Slow Errors

Comparative Statics Within Problem.

Holding problem fixed, is time be informative of choice quality? Are errors fast or slow?

Focus: Fudenberg, Strack, and Strzalecki (2018 AER) and Cripps and Gonçalves (WP).

No time this year...

Highlights:

FFS (1): $\theta = \pm 1$, $p_0 = \tilde{p} = 1/2$, fixed m. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ constant in t.

(2): $\theta = \pm 1$, $p_0 = \tilde{p} = 1/2$, $m = |\varepsilon|$, $\varepsilon \sim N(0,1)$. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ decreasing in t.

CG: revisit $\theta = \pm 1$, $p_0, \tilde{p}, \in (0,1)$, fixed m. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ strictly monotone in t, except at knife-edge symmetry conditions; more: may be increasing or decreasing!

Implications: taking longer has clear implications on choice quality!

E.g. embed into pricing problem where price shifts $\tilde{p}$ (qlinearity).

Fast and Slow Errors

Comparative Statics Within Problem.

Holding problem fixed, is time be informative of choice quality? Are errors fast or slow?

Focus: Fudenberg, Strack, and Strzalecki (2018 AER) and Cripps and Gonçalves (WP).

No time this year...

Highlights:

FFS (1): $\theta = \pm 1$, $p_0 = \tilde{p} = 1/2$, fixed m. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ constant in t.

(2): $\theta = \pm 1$, $p_0 = \tilde{p} = 1/2$, $m = |\varepsilon|$, $\varepsilon \sim N(0,1)$. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ decreasing in t.

CG: revisit $\theta = \pm 1$, $p_0, \tilde{p}, \in (0,1)$, fixed m. Get $\mathbb{P}(\alpha_\tau = \theta \mid \tau = t)$ strictly monotone in t, except at knife-edge symmetry conditions; more: may be increasing or decreasing!

Implications: taking longer has clear implications on choice quality!

E.g. embed into pricing problem where price shifts $\tilde{p}$ (qlinearity).

Cool Tools:

FFS use Brownian bridge argument for their result.

CG use total positivity of killed Brownian density + parabolic comparison principle.

Learning in Continuous Time

Duarte Gonçalves

University College London

Topics in Economic Theory

References

Pham (2009), *Continuous-time Stochastic Control and Optimization with Financial Applications*. Ch. 1, 3, 5.

Øksendal (2003), *Stochastic Differential Equations*, Ch. 10-11.

Karatzas and Shreve (1998), *Brownian Motion and Stochastic Calculus*. Ch. 1, 3, 5.

Fleming and Soner (2006), *Controlled Markov Processes and Viscosity Solutions*. Ch. IV-V.

Liptser and Shiryaev (2001), *Statistics of Random Processes I*. Ch. 8-10.

Peskir and Shiryaev (2006), *Optimal Stopping and Free-Boundary Problems*. Ch. 2, 9.

Teschl (2012), *Ordinary Differential Equations and Dynamical Systems*. Ch. 1.

Gonçalves (WP), Speed, Accuracy, and Complexity.

Fudenberg, Strack, and Strzalecki (2018 AER), Speed, Accuracy, and the Optimal Timing of Choices.

Cripps and Gonçalves (WP), Fast and Slow Errors.