Consistency of Bayesian Learning

- The standard result is that Bayesian learning is consistent— the limit beliefs concentrate around the true model.
- This requires that the prior is "correctly specified": Bayesians can't learn the true model if their prior gives its neighborhood probability 0.
- Will cover models of mispecified learners, but start out by reviewing the proof for the correctly-specified case.
- We consider the easy special case of finite-support priors; here "correctly specified" means that the prior assigns strictly positive probability to the true model.
- Also assume that the environment is exogenous i.i.d.

Environment

- Finite set of possible observations $Y \subset \mathbb{R}$.

- $(Y^\infty, \mathcal{B}(Y^\infty))$ is the measurable space of sequences in Y endowed with the product topology.

- The true model is i.i.d, and described by a time invariant probability measure $\hat{\theta} \in \Delta(Y)$ with full support.

- Denote the associated probability measure on infinite sequences by $\mathbb{P}_{\hat{\theta}}$, unique extension of the product measure over the finite sequences.

Beliefs

- Non-empty finite set $\{\theta_k\}_{k \in K} := \Theta \subset \Delta(Y)$ of models.
- Each $\theta_k \in \Theta$ induces a probability measure $\mathbb{P}_{\theta_k}$ over the infinite sequences of outcomes.
- Initial full support belief $\mu_0 \in \Delta(\Theta)$, with $\hat{\theta} \in \Theta$.
- Assume that for all y and k, $\hat{\theta}(y) > 0$ iff $\theta_k(y) > 0$: The truth and the agent's set of possible models are mutually absolutely continuous.
- Then Bayesian updating induces a well-defined stochastic process of beliefs:

$$\mu_t(y_1, ..., y_t)(\theta_k) = \frac{\mu_0(\theta_k) \prod_{\tau=1}^{t} \theta_k(y_\tau)}{\sum_{\theta_j \in \Theta} \mu_0(\theta_j) \prod_{\tau=1}^{t} \theta_j(y_\tau)}.$$

Consistency of correctly specified learning

Theorem

The posterior probability of the true model converges to 1*:*

$$\lim_{t \to \infty} \mu_t \left(\hat{\theta} \right) = 1 \quad \mathbb{P}_{\hat{\theta}} \text{ a.s.}$$

Reviewing the proof (Doob [1949]) will help set the stage for the analysis of incorrectly specified Bayesians.

Proof

- For $\theta_k \neq \hat{\theta}$, let $W_\tau^k (y_\tau) = \log \left(\frac{\theta_k (y_\tau)}{\hat{\theta} (y_\tau)} \right)$.

- The logarithm function is strictly concave, and $\theta_k \neq \hat{\theta}$ implies that the W_τ^k aren't constant.

- So by Jensen's inequality

$$\mathbb{E}_{\mathbb{P}_{\hat{\theta}}} \left(W_\tau^k \right) < \log \mathbb{E}_{\mathbb{P}_{\hat{\theta}}} \left(\exp W_\tau^k \right). \tag{1}$$

Bounding the expectation of the log-likelihood

- Also

$$\mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left(\exp W_{\tau}^{k}\right) = \mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left[\frac{\theta_k(\cdot)}{\hat{\theta}(\cdot)}\right] = \int_Y \frac{\theta_k(y)}{\hat{\theta}(y)}\hat{\theta}(y)\,dy = 1.$$

- Substituting into (1) yields

$$\mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left(W_1^k\right) < \log 1 = 0.$$

- And our absolute continuity assumption implies $\mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left(W_1^k\right) > -\infty.$

Applying the SLLN

- Since the y_τ are i.i.d, the W_τ^k are i.i.d as well.

- And since $-\infty < \mathbb{E}_{\mathbb{P}_{\hat{\theta}}}(W_1^k) < 0$, the strong law of large numbers implies that

$$\mathbb{P}_{\hat{\theta}}\left(\lim_{t\to\infty}\frac{\sum_{\tau=1}^{t} W_\tau^k\left(y_\tau\right)}{t} = \mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left(W_1^k\right)\right) = 1.$$

- So

$$\lim_{t\to\infty}\sum_{\tau=1}^{t} W_\tau^k\left(y_\tau\right) = \lim_{t\to\infty} t\cdot\mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left(W_1^k\right) = -\infty.$$

- Now define the random variable Z_t^k as the likelihood ratio between θ_k and $\hat{\theta}$ given observations until time t:

$$Z_t^k\left(y_1, ..., y_t\right) = \prod_{\tau=1}^{t} \frac{\theta_k\left(y_\tau\right)}{\hat{\theta}\left(y_\tau\right)} = \exp\left(\sum_{\tau=1}^{t} \log\left(\frac{\theta_k\left(y_\tau\right)}{\hat{\theta}\left(y_\tau\right)}\right)\right).$$

- Then

$$\begin{aligned}\lim_{t\to\infty} Z_t^k\left(y_1, ..., y_t\right) &= \lim_{t\to\infty} \exp\left(\sum_{\tau=1}^{t} \log\left(\frac{\theta_k\left(y_\tau\right)}{\hat{\theta}\left(y_\tau\right)}\right)\right) \\ &= \exp\left(\lim_{t\to\infty} \sum_{\tau=1}^{t} W_\tau^k\left(y_\tau\right)\right).\end{aligned}$$

- Almost surely the likelihood ratio of the observed outcome path converges to 0.

Plug this into Bayes rule:

$$
\begin{aligned}
\mu_t(y_1, ..., y_t)\left(\hat{\theta}\right) &= \frac{\mu_0\left(\hat{\theta}\right)\prod_{\tau=1}^{t}\hat{\theta}(y_\tau)}{\sum_{\theta_k\in\Theta}\mu_0(\theta_k)\prod_{\tau=1}^{t}\theta_k(y_\tau)} \\
&= \mu_0\left(\hat{\theta}\right)\left(\sum_{\theta_k\in\Theta}\mu_0(\theta_k)\frac{\prod_{\tau=1}^{t}\theta_k(y_\tau)}{\prod_{\tau=1}^{t}\hat{\theta}(y_\tau)}\right)^{-1} \\
&= \mu_0\left(\hat{\theta}\right)\left(\mu_0\left(\hat{\theta}\right)+\sum_{\theta_k\neq\hat{\theta}}\mu_0(\theta_k)Z_t^k(y_1, ..., y_t)\right)^{-1}.
\end{aligned}
$$

So we have shown that

$$
\begin{aligned}
&\lim_{t\to\infty}\mu_t(y_1, ..., y_t)(\hat{\theta}) \\
= \;& \mu_0(\hat{\theta})\left(\mu_0(\hat{\theta})+\sum_{\theta_k\neq\hat{\theta}}\mu_0(\theta_k)\lim_{t\to\infty}Z_t^k(y_1, ..., y_t)\right)^{-1} = 1,
\end{aligned}
$$

where the last equality holds because the likelihood ratios converge to 0.

Identifiability

- So far we have defined a model to be a point in $\Delta(Y)$.
- Instead we could have an abstract family of models, with each model θ associated with a probability distribution over the observable outcomes through a map $\theta \mapsto p_\theta \in \Delta(Y)$.
- *Identifiability* requires that this map is 1-1.
- If $p_{\theta_1} = p_{\theta_2}$, the agent will be never able to distinguish the two models, since the likelihood of every observable event is the same under both.
- With endogenous signals, just what is identified and what is not will depend on the actions chosen by the agent.

Consistency in Hidden Markov Models

- The consistency of Bayesian updating is not limited to i.i.d. environments.
- A *hidden Markov model* is a stochastic process $(X_t, Y_t)_{t\in\mathbb{N}}$ where $(X_t)_{t\in\mathbb{N}}$ is an irreducible Markov chain with k states in $\mathcal{X}$.
- The agent only observes $(Y_t)_{t\in\mathbb{N}}$.
- Here a model $\theta = q \times M \in \Theta$ describes the initial distribution on the states q and the transition matrix $M : \mathcal{X} \to \Delta(\mathcal{X})$.
- The distribution of each Y_t depends only on X_t, i.e., there exists a (known) output kernel $K : \mathcal{X} \to \Delta(\mathcal{Y})$, and the $Y's$ are independent given the X's. (Note that the transition probabilities don't depend on Y.)

Theorem (De Gunst and Shcherbakova [2008])

Let $\hat{\theta} = \left(\hat{q}, \hat{M}\right) \in \operatorname{supp} \mu$ *where* $\hat{q}$ *has full support and* $\hat{M} \gg 0$*.*
Then, for every open set U *that contains* $\hat{\theta}$

$$\mathbb{P}_{\hat{\theta}}\left(\left\{\lim_{t \to \infty} \mu\left(U \mid Y_1, ..., Y_t\right) = 1\right\}\right) = 1.$$

- Having $\hat{M} \gg 0$ is stronger than necessary; can replace it with milder but more complicated conditions.

Misspecified beliefs

- Now suppose that the true model $\hat{\theta}$ is not in the support of the prior.
- Many motivation have been offered for this. For example the agent might mistakenly believe outcomes are independent, as in e.g. Enke and Zimmermann [2019]; will see more examples next time in Esponda and Pouzo [2016].
- Berk [1966] shows that when the data is exogenous the posterior concentrates on the models that minimize the Kullback-Leibler divergence from the correct model $\hat{\theta}$.
- In class today we'll prove this for the simple case in which Θ is finite.
- An handout covers the more interesting but involved case of a compact set of models Θ.

Kullback-Leibler divergence

- The Kullback-Leibler (KL) divergence between model θ and the true model $\hat{\theta}$ is defined as

$$R\left(\hat{\theta}||\theta\right) = \int_Y \log\left(\frac{\hat{\theta}\left(y\right)}{\theta\left(y\right)}\right) d\hat{\theta}\left(y\right).$$

- Can view the KL divergence as a measure of the average inability of model θ to predict the realized state, where the expectation is taken w.r.t. $\hat{\theta}$.
- Note that it is convex in its second argument, and it is strictly convex on the set of probabilities that are absolutely continuous with respect to the first argument.
- If θ and θ' differ only on events that have probability 0 under $\hat{\theta}$, any convex combination of them has the same KL divergence with $\hat{\theta}$.

- Let $\Theta(\hat{\theta}) := \arg\min_{\theta \in \Theta} R\left(\hat{\theta} || \theta\right)$.

- If Θ is convex and $\hat{\theta}(y) > 0$ for all y (which rules out our simple case of finite Θ) the minimizer is unique but otherwise it need not be.

Theorem (Berk)

$$\lim_{t \to \infty} \mu_t(y_1, ..., y_t)(\Theta(\hat{\theta})) = 1 \quad \mathbb{P}_{\hat{\theta}} \text{ a.s.}$$

- With finite Θ the proof is a simple extension of the one for correctly specified Bayesians.

- In the more general case it follows by the SLLN and a result on belief concentration over the KL-minimizers with respect to the empirical distribution.

Proof of Berk's Theorem

- Fix a model $\bar{\theta} \in \Theta(\hat{\theta})$ that minimizes the KL divergence from the true model.
- For $\theta_k \notin \Theta(\hat{\theta})$, define the random variable Z_t^k as the likelihood ratio between model θ_k and $\bar{\theta}$ given the observations until time t:

$$Z_t^k\left(y_1, ..., y_t\right) = \prod_{\tau=1}^{t} \frac{\theta_k\left(y_\tau\right)}{\bar{\theta}\left(y_\tau\right)}.$$

- Z_t^k determines the ratio between the posterior probabilities of θ_k and $\bar{\theta}$.
- $Z_t^k\left(y_1, ..., y_t\right) = \exp\left(\sum_{\tau=1}^{t} \log\left(\frac{\theta_k\left(y_\tau\right)}{\bar{\theta}\left(y_\tau\right)}\right)\right).$

Bounding the expectation of the log-likelihood

- Recall that

$$W_\tau^k\left(y_\tau\right) = \log\left(\frac{\theta_k\left(y_\tau\right)}{\bar{\theta}\left(y_\tau\right)}\right) = \log\left(\frac{\theta_k\left(y_\tau\right)}{\hat{\theta}\left(y_\tau\right)}\right) - \log\left(\frac{\bar{\theta}\left(y_\tau\right)}{\hat{\theta}\left(y_\tau\right)}\right).$$

- Because θ_k is not a KL minimizer,

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left(W_\tau^k\right) &= \int_Y \left(\log\left(\frac{\theta_k\left(y_\tau\right)}{\hat{\theta}\left(y_\tau\right)}\right) - \log\left(\frac{\bar{\theta}\left(y_\tau\right)}{\hat{\theta}\left(y_\tau\right)}\right)\right) d\hat{\theta}(y) \\ &< 0.\end{aligned}$$

Applying the SLLN

- Since $\bar{\theta} \in \Theta(\hat{\theta})$,

$$-\infty < \int_Y \log\left(\frac{\hat{\theta}\,(y_\tau)}{\bar{\theta}\,(y_\tau)}\right)\hat{\theta}\,(y)\,dy$$

and therefore $0 > \mathbb{E}_{\mathbb{P}_{\hat{\theta}}}(W_1^k) > -\infty$.

- Since the y_τ are i.i.d, the W_τ^k are i.i.d as well.

- So the strong law of large numbers implies that

$$\mathbb{P}_{\hat{\theta}}\left(\lim_{t\to\infty}\frac{\sum_{\tau=1}^{t} W_\tau^k\,(y_\tau)}{t} = \mathbb{E}_{p_{\hat{\theta}}}\left(W_1^k\right)\right) = 1.$$

Therefore, $\mathbb{P}_{\hat{\theta}}$ a.s. we have

$$\lim_{t\to\infty}\sum_{\tau=1}^{t} W_{\tau}^{k}\left(y_{\tau}\right)=\lim_{t\to\infty} t\cdot\mathbb{E}_{\mathbb{P}_{\hat{\theta}}}\left(W_{1}^{k}\right)=-\infty.$$

$$\Longrightarrow$$

$$\begin{aligned}\lim_{t\to\infty} Z_{t}^{k}\left(y_{1},...,y_{t}\right) &= \lim_{t\to\infty}\exp\left(\sum_{\tau=1}^{t}\log\left(\frac{\theta_{k}\left(y_{\tau}\right)}{\bar{\theta}\left(y_{\tau}\right)}\right)\right)\\ &= \exp\left(\lim_{t\to\infty}\sum_{\tau=1}^{t} W_{\tau}^{k}\left(y_{\tau}\right)\right)\\ &= 0.\end{aligned}$$

Almost surely the likelihood ratio converges to 0.

Plug this into Bayes rule:

$$\begin{aligned}\mu_t\left(y_1,...,y_t\right)\left(\theta_k\right) &= \frac{\mu_0\left(\theta_k\right)\prod_{\tau=1}^{t}\theta_k\left(y_\tau\right)}{\sum_{\theta\in\Theta}\mu_0\left(\theta\right)\prod_{\tau=1}^{t}\theta\left(y_\tau\right)} \\ &= \mu_0\left(\theta_k\right)\left(\sum_{\theta\in\Theta}\mu_0\left(\theta\right)\frac{\prod_{\tau=1}^{t}\theta\left(y_\tau\right)}{\prod_{\tau=1}^{t}\theta_k\left(y_\tau\right)}\right)^{-1} \\ &\leq \mu_0\left(\theta_k\right)\left(\mu_0\left(\bar{\theta}\right)\frac{\prod_{\tau=1}^{t}\bar{\theta}\left(y_\tau\right)}{\prod_{\tau=1}^{t}\theta_k\left(y_\tau\right)}\right)^{-1}.\end{aligned}$$

Taking the limit yields:

$$\begin{aligned}&\lim_{t\to\infty}\mu_t\left(y_1,...,y_t\right)\left(\theta_k\right) \\ \leq\ &\mu_0\left(\theta_k\right)\left(\frac{\mu_0(\bar{\theta})}{\lim_{t\to\infty}Z_t^k\left(y_1,...,y_1\right)}\right)^{-1}=0.\end{aligned}$$

This holds for every $\theta_k \notin \Theta(\hat{\theta})$, which proves the theorem (and explains why we'll be seeing a lot of the KL divergence and nothing about other divergence notions.)

Uniform Exponential Concentration

- When Θ is finite the proof above also gives a uniform rate of convergence.
- Berk's more general proof doesn't provide that.
- Fudenberg, Lanzani, and Strack [forthcoming] provides one.
- It first shows that under an additional "ϕ positivity" assumption on the prior, beliefs concentrate around the parameters that almost best fit the observed distribution, with a rate that is exponential in t, and uniform over the empirical frequencies.
- This lets us give a rate of convergence for Berk's theorem.

Idea of Proof:

- With probability that goes to 1 at an exponential rate the empirical distribution is very close to the true data generating process. (This is Sanov's theorem, a "concentration inequality.")

- And when this concentration happens, the belief assigned to every θ outside a ball around the KL-minimizer drops to 0 exponentially fast.

- The full support condition can be relaxed.

Oscillation in beliefs

- The fact that the limit probability assigned to $\Theta\left(\hat{\theta}\right)$ converges to 1 doesn't imply beliefs converge.

- Simple example: $Y = \{0, 1\}$, and the true data generating process is a fair coin.

- Three biased coins: $\Theta = \{p, q, r\}$ with

$$p\left(y\right) = \begin{cases} \frac{3}{4} & y = 1 \\ \frac{1}{4} & y = 0 \end{cases} \quad q\left(y\right) = \begin{cases} \frac{1}{4} & y = 1 \\ \frac{3}{4} & y = 0 \end{cases} \quad r\left(y\right) = \begin{cases} \frac{1}{10} & y = 1 \\ \frac{9}{10} & y = 0 \end{cases}$$

- The closest model to the truth are p and q, so Berk's theorem guarantees that

$$\lim_{n\to\infty} \mu\left(\{p,q\} \,|\, (Y_1,...,Y_n)\right) = 1.$$

- However, Berk showed that regardless of the prior μ, the beliefs will oscillate between concentrating around p and q:

$$\lim \sup_{n\to\infty} \mu\left(p|\,(Y_1,...,Y_n)\right) = \lim \sup_{n\to\infty} \mu\left(q|\,(Y_1,...,Y_n)\right) = 1.$$

- Fudenberg, Lanzani, and Strack [2021] generalizes this: $\limsup_{n\to\infty} \mu\left(B_\varepsilon(p)|\,(Y_1,...,Y_n)\right) = 1$ for every $p \in \Theta(\hat{\theta})$ and $\varepsilon > 0$.

- In active learning problems, this implies that behavior can only converge to actions that are best replies to *all* of the models that minimize the KL divergence.

Oscillation in location experiments

- Diaconis and Freedman [1986] give an example with an infinite-dimensional prior where oscillations occur even with a correctly specified agent.
- In the example the data generating process is a real-valued location experiment
$$Y_t = \nu + \varepsilon_t$$
where the ε's are independent with unknown cumulative distribution function F.
- They consider priors that are common in Bayesian statistics: normal distribution on ν and a Beta distribution for the errors, with a Cauchy measure on the Beta parameter.
- Trade-off between better behavior of beliefs in finite dimensional models and the increased risk of being misspecified.

Martingales and upcrossings

- Martingales are fundamental in learning theory and the Martingale Convergence Theorem is a key result.

- The proof of the theorem uses the idea of "upcrossings."

- The properties of supermartingale upcrossings can be are also used directly in e.g. in the reputation literature (Fudenberg and Levine [1992a] or Pei [2020]).

- Consider a filtered probability space $\left(\Omega, \mathcal{F}, \left(\mathcal{F}_n\right)_{n\in\mathbb{N}}, \mathbb{P}\right)$.
- Each $\mathcal{F}_n$ is a sigma-algebra; x_n is what has been observed by stage n. (Filtration means "no forgetting": any set that is measurable given period n information is measurable at $n+1$.)
- $\mathcal{F}$ is the sigma-algebra generated by the union of the $\mathcal{F}_n$.

Definition

- $\left(X_n\right)_{n\in\mathbb{N}}$ is adapted to the filtration if X_n is $\mathcal{F}_n$-measurable $\forall n \in \mathbb{N}$.
- A stopping time is a random variable τ with values in $\mathbb{N} \cup \infty$ that is adapted to the filtration, that is $\{\tau = n\} \in \mathcal{F}_n$ for all n: the decision to stop at n can only depend on information that is available then.
- A supermartingale is an adapted stochastic process such that $|\mathbb{E}[X_n]| \leq \infty$ and $\mathbb{E}[X_{n+1}|X_1, ..., X_n] \leq X_n$ for all $n \in \mathbb{N}$.

Upcrossing Numbers

- For an interval $[a, b]$ of $\mathbb{R}$, the upcrossing number $U_N[a, b](\omega)$ of upcrossings made in state ω by time N is the largest $k \in \mathbb{N}$ such that

$$0 \leq s_1 < t_1 < s_2 < t_2 < ... < s_k < t_k \leq N$$

with

$$X_{s_i} < a \text{ and } X_{t_i} > b.$$

- In words, $U_N[a, b](\omega)$ counts how many times the stochastic process rises from below a to above b before time N.

- $U_\infty[a, b](\omega)$ is the total number of upcrossings on the whole sample path.

Theorem (Dubins' Upcrossing Inequality)

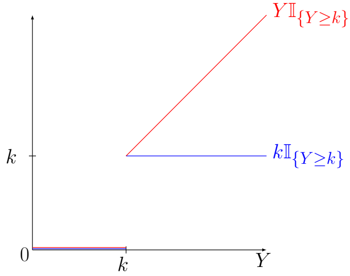
If $(X_t)_{t\in\mathbb{N}}$ *is a positive supermartingale, and* $0 < a < b < \infty$*, then*

$$\mathbb{P}\left(U_\infty\left[a,b\right] \geq k\right) \leq \left(\frac{a}{b}\right)^k \left(\min\left\{\frac{X_0}{a}, 1\right\}\right).$$

- Dubins' inequality is a generalization of Markov's inequality for non-negative random variables, which says

$$\mathbb{P}(Y \geq k) \leq \frac{\mathbb{E}\left[Y\mathbb{I}_{\{Y\geq k\}}\right]}{k} \leq \frac{\mathbb{E}[Y]}{k}.$$

- This follows from $k\mathbb{I}_{\{Y\geq k\}} \leq Y\mathbb{I}_{\{Y\geq k\}}$ and taking expectations.



Theorem (Dubins' Upcrossing Inequality)

If $(X_t)_{t\in\mathbb{N}}$ *is a positive supermartingale, and* $0 < a < b < \infty$*, then*

$$\mathbb{P}\left(U_\infty[a,b] \geq k\right) \leq \left(\frac{a}{b}\right)^k \left(\min\left\{\frac{X_0}{a}, 1\right\}\right).$$

- *Idea*: Because X is a supermartingale, the probability of moving from below a to above b is at most a/b, so the probability $k+1$ upcrossings of $[a,b]$ is at most a/b times the probability of k upcrossings.

The proof of Dubin's inequality uses the following lemma:

Lemma

If $\left(X_t^1\right)_{t\in\mathbb{N}}$ *and* $\left(X_t^2\right)_{t\in\mathbb{N}}$ *are positive supermartingales and* v *is a stopping time such that* $X_{v(\omega)}^1(\omega) \geq X_{v(\omega)}^2(\omega)$,

$$Z_n(\omega) = \begin{cases} X_n^1(\omega) & 0 \leq n < v(\omega) \\ X_n^2(\omega) & v(\omega) \leq n \end{cases}$$

is a positive supermartingale.

- *Intuition*: If we start with a process that decreases on average, and replace it (at a random stopping time) with something weakly lower, the new process is also decreasing.

Proof of Dubins' upcrossing inequality

Define stopping times by when X_n crosses below a or above b:

$$\begin{aligned} v_1(\omega) &= \min\{n : X_n(\omega) < a\} \\ v_2(\omega) &= \min\{n \geq v_1 : X_n(\omega) > b\} \end{aligned}$$

$$v_3(\omega), v_4(\omega), \ldots$$

Now fix a k and define an ancillary process that caps X_n at $(b/a)^i$ in the i-th excursion above b, $1 \leq i \leq k$

$$Y_n(\omega) = \begin{cases} 1 & 0 \leq n < v_1(\omega) \\ \frac{X_n(\omega)}{a} & v_1(\omega) \leq n < v_2(\omega) \\ \frac{b}{a} \cdot 1 & v_2(\omega) \leq n < v_3(\omega) \\ \frac{b}{a} \cdot \frac{X_n(\omega)}{a} & v_3(\omega) \leq n < v_4(\omega) \\ \ldots & \ldots \\ \left(\frac{b}{a}\right)^{k-1} \cdot \frac{X_n(\omega)}{a} & v_{2k-1}(\omega) \leq n < v_{2k}(\omega) \\ \left(\frac{b}{a}\right)^{k} & v_{2k}(\omega) \leq n. \end{cases}$$

- Let Z_1 be identically 1, and Z_2 be X_n/a. Then switching from Z_1 to Z_2 at $v_1(\omega)$ is switching to a lower supermartingale.

- Then let $Z_3 = b/a$; switching to Z_3 at $v_2(\omega)$ is switching to a lower supermartingale.

- And iteration of the switching lemma shows that $(Y_n)_{n\in\mathbb{N}}$ is a supermartingale

- If $X_0/a > 1$ then $Y_0 = 1$; if $X_0/a \leq 1$ then and $Y_0 = X_0/a$.
- So $Y_0 = \min\left\{1, \frac{X_0}{a}\right\} = \mathbb{E}\left(Y_0\right)$.
- Since $\left(Y_n\right)_{n\in\mathbb{N}}$ is a supermartingale,

$$\min\left\{1, \frac{X_0}{a}\right\} = \mathbb{E}\left(Y_0\right) \geq \mathbb{E}\left(Y_n\right).$$

- And $Y_n \geq \left(\frac{b}{a}\right)^k 1_{\{v_{2k}\leq n\}}$.
- So

$$\left(\frac{a}{b}\right)^k \min\left\{1, \frac{X_0}{a}\right\} \geq \mathbb{P}\left(\{v_{2k} \leq n\}\right).$$

- Because

$$\{v_{2k} < \infty\} = \{U_\infty\left[a, b\right] \geq k\},$$

the theorem follows.

More Upcrossings

- An upcrossing argument due to Doob is also used to prove the Martingale Convergence Theorem.

- Dubins' result bounds the *probability* of more than k upcrossing in the whole sample path; Doob's result bounds the *expected value* of the upcrossing in a finite time.

Doob's Upcrossing Lemma

Let X be a supermartingale. Then

$$(b-a)\,\mathbb{E}\left(U_N\left[a,b\right]\right) \leq \mathbb{E}\left(\max\{a-X_N,0\}\right).$$

As a consequence, if $\sup_n \mathbb{E}\left(|X_n|\right) < \infty$,

$$P\left(U_\infty\left[a,b\right] \text{ is finite}\right) = 1.$$

The Martingale Convergence Theorem

Theorem

Let X be a supermartingale with $\sup_n \mathbb{E}\left(|X_n|\right) < \infty$. Then, a.s., $X_\infty := \lim X_n$ exists and is finite.

- Let

$$\Lambda_{a,b} := \left\{\omega : \liminf X_n\left(\omega\right) < a < b < \limsup X_n\left(\omega\right)\right\}.$$

- Then,

$$\begin{aligned}
\Lambda : \quad &= \quad \left\{\omega : X_n\left(\omega\right) \text{ does not converge to a limit in } [-\infty, +\infty]\right\} \\
&= \quad \left\{\omega : \liminf X_n\left(\omega\right) < \limsup X_n\left(\omega\right)\right\} \\
&= \quad \bigcup_{\{a,b \in \mathbb{Q} : a < b\}} \left\{\omega : \liminf X_n\left(\omega\right) < a < b < \limsup X_n\left(\omega\right)\right\}.
\end{aligned}$$

- But

$$\Lambda_{a,b} \subseteq \{\omega : U_\infty [a,b] (\omega) = \infty\}$$

and by Doob's Upcrossing Lemma, $P(\Lambda_{a,b}) = 0$, and so $P(\Lambda) = 0$. Therefore

$$X_\infty := \lim X_n \text{ exists a.s. in } [-\infty, \infty].$$

- Finally

$$\mathbb{E}(|X_\infty|) = \mathbb{E}(\liminf |X_n|) \leq \liminf \mathbb{E}(|X_n|) \leq \sup_n \mathbb{E}(|X_n|) < \infty$$

where the equality follows from the previous argument, the first inequality by Fatou's Lemma, and the strict inequality is assumed by the theorem. But this implies that

$$P(X_\infty \text{ is finite}) = 1.$$

Corollary

If X is a positive supermartingale then $\mathbb{E}(|X_n|) = \mathbb{E}(X_n) \leq \mathbb{E}(X_0)$ so a.s., $X_\infty := \lim X_n$ exists and is finite.

Reputation Effects

- LR player 1, infinite sequence of SR player 2's, time periods $n = 0, 1, 2, \ldots.$
- Each period n players 1 and 2 simultaneously choose actions a_{1n}, a_{2n} respectively (or mixed actions $\alpha_i \in \Delta(A_i)$)
- Simplify by assuming that at end of each period, players observe actions played. (Results extend to case of signals that needn't fully identify the actions.)
- Each player 2 picks a_{2n} to maximize expected value of $u_2(a_{1n}, a_{2n})$.

- The "rational type" θ^* of player 1 has time stationary preferences: maximizes expected discounted value of utility u_1 with discount factor δ.
- $B(\alpha_1) = \{\alpha_2 \in \operatorname{argmax} u_2(\alpha_1, \alpha_2)\}$
- $B_\varepsilon(\alpha_1) = \{\alpha_2 \in B(\alpha_1'), ||\alpha_1' - \alpha_1|| < \varepsilon\}$
- The Stackelberg payoff is $\max_{\alpha_1} \min_{\alpha_2 \in B(\alpha_1)} u_1(\alpha_1, \alpha_2)$.
- Intuition behind this literature is that when player 1 is patient they should be able to do about as well as their Stackleberg payoff.
- Let α_1^* be a Stackelberg action—an element of the argmax.
- To model the possibility of building a reputation for playing Stackelberg, suppose there is positive prior probability that player 1 is a "Stackleberg type" ω^* that always plays α_1^*.
- This type is private information, not known to the player 2's.

- To simplify we suppose that the support of the prior is $\Theta^* = \{\theta^*, \omega^*\}$.
- A Nash equilibrium exists in this game.
- Let $W(\delta)$ be the infimum of rational type's payoff over all of the Nash equilibria when the discount factor is δ.
- Let $\underline{u_1}$ be 1's lowest possible payoff, and

$$u_1(\varepsilon) = \inf_{\alpha_2 \in B\left(\alpha_1^*\right)} u_1\left(\alpha_1^*, \alpha_2\right) - \varepsilon.$$

Theorem (Fudenberg and Levine [1992b])

$\forall \varepsilon > 0$, $\exists k$ s.t. $\forall \delta$:

$$(1-\varepsilon)\delta^k u_1(\varepsilon) + \left[1 - (1-\varepsilon)\delta^k\right] \underline{u_1} \leq W(\delta).$$

Outline of proof

- Fix a Nash equilibrium (NE); the NE strategies and the prior determine a joint probability distribution over types and histories.
- In equilibrium, SR players use Bayesian updates from this distribution to form their posterior beliefs at any history that has positive probability.
- Suppose that every period LR plays α_1^*. (This needn't be the optimal play, but it is a feasible one.)
- Because the Stackelberg type has positive probability, SR beliefs should come to expect this play.
- Note this doesn't say that the SR learn the LR is the commitment type—there may be a "pooling equilibrium" where the rational type also always plays α_1^*.

- In a pooling equilibrium beliefs converge from the start- and LR gets at least the Stackelberg payoff associated with α_1^*.
- Fix an ε and $\delta > 0$ and say a period is "bad" if SR play an $\alpha_2 \notin B_\varepsilon(\alpha_1^*)$.
- In the pooling equilibrium there aren't any bad periods.
- Now we need to bound how many bad periods there are.
- Study the evolution of the SR beliefs *in the bad periods only*.
- Since best response correspondence has closed graph, the equilibrium play of the rational type in the bad periods is uniformly bounded away from α_1^*.
- *Idea*: In good periods the LR player gets a high payoff; in bad periods the SR players are "surprised" and increase the probability they assign to ω^*.

- Want to show that there can't be too many such surprises, as once the probability of ω^* is high enough all subsequent periods will be good.
- Here the play of the rational type is history dependent and not i.i.d.
- *Easy lemma* There is a γ s.t. for all bad periods t, $||\alpha_1^* - \theta^*|| > \gamma$.
- Under strategy α_1^* the process $\frac{1-\mu_t(\omega^*)}{\mu_t(\omega^*)}$ is a supermartingale.
- And one can use Dubins' inequality to show the following:

Claim

For every $\underline{L} > 0$ *and* $\epsilon \in (0,1)$*, there is* $T < \infty$ *s.t.* $\mathbb{P}_{\alpha_1^*}[\sup_{t \geq T} \frac{1-\mu_t(\omega^*)}{\mu_t(\omega^*)} \leq \underline{L}] \geq (1-\epsilon)\,.$

So it is unlikely there will be many bad periods, which proves the theorem.

Fatou's Lemma

Lemma

Let $(f_n)_{n\in\mathbb{N}}$ *be a sequence of measurable functions and*

$$f(x) = \liminf_{n\in\mathbb{N}} f_n(x).$$

Then f *is measurable and*

$$\int f d\mu \leq \liminf_{n\in\mathbb{N}} \int f_n d\mu$$

Back to slides

References I

[1] Robert H Berk. "Limiting behavior of posterior distributions when the model is incorrect." *The Annals of Mathematical Statistics* (1966).

[2] MCM De Gunst and O Shcherbakova. "Asymptotic behavior of Bayes estimators for hidden Markov models with application to ion channels." *Mathematical Methods of Statistics* (2008).

[3] Persi Diaconis and David Freedman. "On the consistency of Bayes estimates." *The Annals of Statistics* (1986).

[4] Joseph L Doob. "Application of the theory of martingales." *Le calcul des probabilites et ses applications* (1949).

[5] Benjamin Enke and Florian Zimmermann. "Correlation neglect in belief formation." *The Review of Economic Studies* (2019).

References II

[6] Ignacio Esponda and Demian Pouzo. “Berk–Nash equilibrium: A framework for modeling agents with misspecified models.” *Econometrica* (2016).

[7] Drew Fudenberg, Giacomo Lanzani, and Philipp Strack. “Pathwise concentration bounds for Bayesian beliefs.” *Theoretical Economics* (forthcoming).

[8] Drew Fudenberg, Giacomo Lanzani, and Philipp Strack. “Limit Points of Endogenous Misspecified Learning.” *Econometrica* (2021).

[9] Drew Fudenberg and David K Levine. “Maintaining a Reputation When Strategies are Imperfectly Observed.” *Review of Economic Studies* (1992).

[10] Drew Fudenberg and David K Levine. “Maintaining a Reputation when Strategies are Imperfectly Observed..” *Review of Economic Studies* (1992).

References III

[11] Harry Pei. "Reputation effects under interdependent values." *Econometrica* (2020).