Lecture 22: Martingale convergence

Under mild conditions subMs converge a.e.

Doobs upcrossing inequality is the key to the proof.

Here is a motivation in terms of how to check a sequence of numbers $x_1, x_2, x_3, \ldots$ converges.

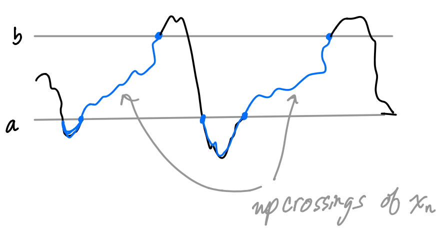
x_n does not converge

iff $\liminf_n x_n < \limsup_n x_n$

iff $\exists a, b \in \mathbb{Q}$ s.t. $a < b$

$x_n > b$ & $x_n < a$ infinitely often.

i.e.



iff $\exists a, b \in \mathbb{Q}$ s.t. x_n has infinitely many upcrossings of $[a,b]$

Doobs upcrossings

Let $(X_1, X_2, \ldots, X_n)$ be adapted to the filtration $(\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n)$.

Fix $-\infty < a < b < \infty$ and define

$$\left.\begin{aligned}\alpha_1 &= \min\left(\{k \geq 1 : X_k \leq a\} \cup \{n\}\right) \\ \beta_1 &= \min\left(\{k > \alpha_1 : X_k \geq b\} \cup \{n\}\right)\end{aligned}\right\} \text{upcrossing of } [a,b]$$

$$\left.\begin{aligned}\alpha_2 &= \min\left(\{k > \beta_1 : X_k \leq a\} \cup \{n\}\right) \\ \beta_2 &= \min\left(\{k > \alpha_2 : X_k \geq b\} \cup \{n\}\right)\end{aligned}\right\} \text{upcrossing of } [a,b]$$

$\vdots$

The number of upcrossings of $[a,b]$, denoted $U_{a,b}$, is defined as

$$U_{a,b} := \sum_{j=1}^{n} I_{\{X_{\alpha_j} \leq a,\, X_{\beta_j} \geq b\}}$$

can't be more than $n/2$ terms

Proposition:

$\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ are stopping times w.r.t $(\mathcal{F}_1, \ldots, \mathcal{F}_n)$ and U_{ab} is $\mathcal{F}_n$-measurable

Proof:

we've already shown that α_1 is a ST.

To show β_1 is a stopping time notice that

- If $1 \leq m < n$ then

$$\{\beta_1 = m\} = \{\alpha_1 < m,\ \beta_1 = m\}$$
$$= \bigcup_{j=1}^{m-1} \{\alpha_1 = j,\ \beta_1 = m\}$$
$$= \bigcup_{j=1}^{m-1} \{\alpha_1 = j,\ X_{j+1} < b, \ldots, X_{m-1} < b, X_m \geq b\}$$

$\in \mathcal{F}_j \subset \mathcal{F}_m$ in $\mathcal{F}_m$

$\in \mathcal{F}_m$

- If $m = n$ then

$$\{\beta_1 = n\} = \{\beta_1 < n\}^c = \left(\bigcup_{j=1}^{n-1} \{\beta_1 = j\}\right)^c \in \mathcal{F}_n$$

An induction argument shows $\alpha_2, \beta_2, \ldots, \alpha_n, \beta_n$ are ST.

$\therefore$ X_{α_j} is $\mathcal{F}_{\alpha_j}$-measurable &

X_{β_j} is $\mathcal{F}_{\beta_j}$-measurable

Since $\mathcal{F}_{\alpha_j}$ & $\mathcal{F}_{\beta_j}$ are sub σ-fields of $\mathcal{F}_n$

U_{ab} is $\mathcal{F}_n$-measurable. QED.

Theorem: (Doob's upcrossing ineq)

Suppose $(X_1, X_2, \dots, X_n)$ is a non-negative subM w.r.t filtration $(\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_n)$. Let $c>0$ & $\mathcal{U}_{0,c}$ denote the upcrossings of $[0,c]$. Then

$$E(\mathcal{U}_{0,c}) \le \frac{E(X_n) - E(X_1)}{c}.$$

Proof: The idea is to write $X_n - X_1$ as a telescoping sum:

$$\begin{aligned} X_n - X_1 = {} & (\overbrace{X_{\beta_n}}^{\text{must be } X_n} - X_{\alpha_n}) + (X_{\alpha_n} - X_{\beta_{n-1}}) \\ & + (X_{\beta_{n-1}} - X_{\alpha_{n-1}}) + (X_{\alpha_{n-1}} - X_{\beta_{n-2}}) \\ & + \\ & \vdots \\ & + (X_{\beta_2} - X_{\alpha_2}) + (X_{\alpha_2} - X_{\beta_1}) \\ & + (X_{\beta_1} - X_{\alpha_1}) + (X_{\alpha_1} - \underbrace{X_1}_{\text{set } \beta_0 = 1}) \end{aligned}$$

Notice that

$$\begin{aligned} E(\text{a red term}) &= E(X_{\alpha_j} - X_{\beta_{j-1}}) \\ &= E(E(X_{\alpha_j} - X_{\beta_{j-1}} \mid \mathcal{F}_{\beta_{j-1}})) \\ &= E(\underbrace{E(X_{\alpha_j} \mid \mathcal{F}_{\beta_{j-1}})}_{\overset{a.e.}{\geq} X_{\beta_{j-1}} \text{ by the Finite optimal Sampling Thm}} - X_{\beta_{j-1}}) \\ &\geq 0 \end{aligned}$$

Moreover

$$(X_{\beta_j} - X_{\alpha_j}) \begin{cases} > c & \text{if } (X_{\alpha_j}, X_{\beta_j}) \text{ an upcrossing} \\ 0 & \text{if } \beta_j = \alpha_j = n \\ > 0 & \text{if } \beta_j = n \text{ but } \alpha_j < n \\ & \text{since } X_{\alpha_j} = 0 \ \& \ X_{\beta_j} \geq 0 \text{ by non-negativity} \end{cases}$$

$\therefore$ The sum of all blue terms $\geq c\,\mathcal{U}_{0,c}$

$$\therefore \quad E(X_n - X_1) \geq E(\underbrace{\text{blues}}_{\geq c\,\mathcal{U}_{0,c}}) + E(\underbrace{\text{reds}}_{\geq 0}) \geq c\,E(\mathcal{U}_{0,c}).$$

QED

Corollary:

If $(X_1, \dots, X_n)$ is a subM w.r.t. $(\mathcal{F}_1, \dots, \mathcal{F}_n)$ and $-\infty < a < b < \infty$ then

$$E(\mathcal{U}_{a,b}) \le \frac{E(X_n - a)^+ - E(X_1 - a)^+}{b-a} \le \frac{E(X_n^+) + a^-}{b-a}.$$

Proof:

For the first inequality:

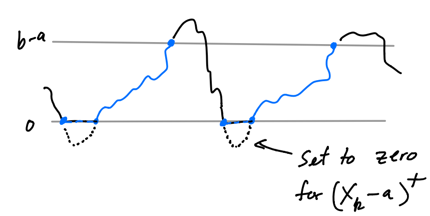
upcrossings of $[a,b]$ in $(X_1, \dots, X_n)$

$\parallel$

upcrossings of $[0, b-a]$ in $(X_1 - a, \dots, X_n - a)$

$\parallel$

upcrossings of $[0, b-a]$ in $((X_1 - a)^+, \dots, (X_n - a)^+)$



By Doobs result we therefore have

$$E(\mathcal{U}_{0,b-a}) \le \frac{E(X_n - a)^+ - E(X_1 - a)^+}{b-a}$$

For the second inequality just notice

$$E(X_n-a)^+ - E(X_1-a)^+ \leq E(X_n-a)^+$$
$$= E(X_n+(-a))^+$$
$$\leq E(X_n^+ + (-a)^+)$$

since $(\cdot)^+$ is convex

$$= E(X_n^+) + a^-$$

QED

a.e. convergence of martingales

For this section let

$$\mathcal{F}_\infty := \sigma\langle \mathcal{F}_1, \mathcal{F}_2, \dots \rangle$$

for a filtration $(\mathcal{F}_n)_{n\geq 1}$.

Theorem: ($E(X_n^+)$ bdd $\Rightarrow$ a.e. conv)

Let $(X_n)_{n\geq 1}$ be a subM w.r.t. filtration $(\mathcal{F}_n)_{n\geq 1}$. If $\sup_n E(X_n^+) < \infty$ then $\exists X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P)$ s.t.

$$X_n \xrightarrow{P\text{-a.e.}} X_\infty$$

Proof:

For $-\infty < a < b < \infty$ let $\mathcal{U}_{a,b}^n$ denote the number of upcrossings of $[a,b]$ from $(X_1, \dots, X_n)$, with $\mathcal{U}_{a,b}^\infty$ for $(X_n)_{n\geq 1}$.

We must have that

$$0 \leq \mathcal{U}_{a,b}^n \uparrow \mathcal{U}_{a,b}^\infty \quad \text{as } n \to \infty$$

$$\therefore \quad E(\mathcal{U}_{a,b}^\infty) \overset{\text{Big 3}}{=} \lim_n E(\mathcal{U}_{a,b}^n)$$
$$\leq \lim_n \frac{E(X_n^+) + a^-}{b-a}$$
$$= \frac{\sup_n E(X_n^+) + a^-}{b-a} < \infty$$

$$\therefore \quad \mathcal{U}_{a,b}^\infty < \infty \quad P\text{-a.e.}$$

Now

$$P(\liminf_n X_n < \limsup_n X_n)$$
$$\leq P\Big(\bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \{\mathcal{U}_{a,b}^\infty = \infty\}\Big)$$

must $\exists$ a, b with an infinite # of upcrossings.

$$= 0 \quad \text{since } \mathcal{U}_{a,b}^\infty < \infty \ P\text{-a.e.}$$

$$\therefore \quad X_n \xrightarrow{P\text{-a.e.}} \underbrace{\limsup_n X_n}_{\text{must be } \text{ⓜ } \mathcal{F}_\infty \text{ by closure Thm.}} =: X_\infty$$

To see why $X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P)$

since the limit exists

$$E|X_\infty| = E|\liminf_n X_n| = E(\liminf_n |X_n|)$$
$$\overset{\text{Fatou}}{\leq} \liminf_n E|X_n|$$
$$\leq \sup_n E|X_n|$$
$$= \sup_n E(2X_n^+ - X_n)$$

since $E(X_1) \leq E(X_n)$ by subM

$$\leq \sup_n \big(2E(X_n^+) - \underbrace{E(X_1)}_{<\infty}\big)$$
$$< \infty$$

QED

Remark: Note the analog to monotonic sequences when $\sup_n E(X_n^+) < \infty$ plays the roll of $\sup_n X_n^+ < \infty$

(7)

Another bdd type condition is that the X_n^+'s are UI.

Proposition: $\left((X_n^+)_{n\geq 1}\ UI \Rightarrow E(X_n^+)\ \text{bdd}\right)$

If $X_1, X_2, \ldots \in L_1(\Omega, \mathcal{F}, P)$ and $(X_n^+)_{n\geq 1}$ are UI then $\sup_n E(X_n^+) < \infty$.

Proof:

$$\sup_n E(X_n^+) \leq \underbrace{\underbrace{\sup_n E(X_n^+ I_{X_n^+ \geq c})}_{\xrightarrow{c\to\infty} 0 \text{ by definition of UI}} + \underbrace{\sup_n E(X_n^+ I_{X_n^+ < c})}_{< c}}_{\text{so for a large enough } c \text{ this is } < \infty.}$$

QED.

Here is an application of the above proposition which comes in handy for subsequent results.

Theorem: (Lévy's Smoothing Martingale)

If $X \in L_1(\Omega, \mathcal{F}, P)$ & $(\mathcal{F}_n)_{n\geq 1}$ is a filtration Then

$$E(X|\mathcal{F}_n) \xrightarrow{n\to\infty} E(X|\mathcal{F}_\infty) \quad \text{a.e. \& in } L_1$$

where $\mathcal{F}_\infty = \sigma\langle \mathcal{F}_1, \mathcal{F}_2, \ldots \rangle$.

Proof:

Let $X_n := E(X|\mathcal{F}_n)$.

We first show the X_n's are UI.

Since $|X_n| = |E(X|\mathcal{F}_n)| \leq E(|X|\,|\mathcal{F}_n)$ it will be sufficient to show the $E(|X|\,|\mathcal{F}_n)$'s are UI.

(8)

i.e. show

$$\lim_{c\to\infty} \sup_n \underbrace{\int_{E(|X||\mathcal{F}_n)\geq c} E(|X|\,|\mathcal{F}_n)\,dP}_{=\int_{E(|X||\mathcal{F}_n)\geq c} |X|\,dP} = 0$$

since $\{E(|X||\mathcal{F}_n)\geq c\}$ is a $\mathcal{F}_n$-set

Let $A_{n,c} := \{E(|X||\mathcal{F}_n)\geq c\}$

we then want to bound $\int_{A_n} |X|\,dP$

Note $\int_\cdot |X|\,dP \ll P(\cdot)$ on $\mathcal{F}_n$

$\therefore$ by HWK 3 in 235A

$$\lim_{\delta\downarrow 0} \sup\left\{\int_F |X|\,dP : F\in\mathcal{F}_n,\ P(F)\leq\delta\right\} = 0$$

i.e. $\forall \varepsilon>0, \exists \delta_\varepsilon > 0$ s.t.

$$\underset{\in\mathcal{F}_n}{P(F)} \leq \delta_\varepsilon \Rightarrow \int_F |X|\,dP \leq \varepsilon$$

Now since

$$P(A_{n,c}) = P(E(|X||\mathcal{F}_n)\geq c)$$
$$\leq \frac{E(E(|X||\mathcal{F}_n))}{c} = \frac{E|X|}{c} \quad (*)$$

we have $\forall \varepsilon>0, \exists C>0$ s.t.

$$c \geq C \Rightarrow P(A_{n,c}) \leq \delta_\varepsilon \quad \text{by } (*)$$

$$\Rightarrow \int_{A_{n,c}} |X|\,dP \leq \varepsilon$$

$$\Rightarrow \sup_n \int_{A_{n,c}} |X|\,dP \leq \varepsilon$$

$\therefore$ the X_n's are indeed UI

$\therefore$ the X_n^+'s are UI so the $E(X_n^+)$'s bdd

$\therefore$ $\exists X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P)$ s.t.

$X_n \xrightarrow{a.e.} X_\infty$ by subM $\xrightarrow{a.e.}$ Thm

$\therefore$ $X_n \xrightarrow{L_1} X_\infty$ by the L_p convergence Thm since the X_n's are UI

To finish we show $X_\infty \stackrel{a.e.}{=} E(X \mid \mathcal{F}_\infty)$.

In particular show:

(i) X_∞ is $\mathcal{F}_\infty$-measurable ✓

(ii) $X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P)$ ✓

(iii) Integrates like X over $\mathcal{F}_\infty$-sets.

i.e. $E(XI_A) = E(X_\infty I_A)\ \forall A \in \mathcal{F}_\infty$

For (iii) notice that $\forall A \in \bigcup_{n=1}^\infty \mathcal{F}_n$

$$E(XI_A) \stackrel{(*)}{=} E(X_n I_A) \quad \forall \text{ large } n$$
since $X_n = E(X \mid \mathcal{F}_n)$

$\therefore$ $\forall A \in \bigcup_{n=1}^\infty \mathcal{F}_n$ ← a field generating $\mathcal{F}_\infty$

$$E(X_\infty I_A) = E(\lim_n X_n I_A) \quad \text{since } X_n \xrightarrow{a.e.} X_\infty$$
$$= \lim_n E(X_n I_A) \quad \text{by UI cond. for } \lim \int = \int \lim$$
$$\stackrel{(*)}{=} E(XI_A) \quad \text{by } (*)$$

To show $E(X_\infty I_A) = E(XI_A)\ \forall A \in \mathcal{F}_\infty$

Let
$$\mathcal{L} = \{A \in \mathcal{F} : E(X_\infty I_A) = E(XI_A)\}$$

By $(*)$ we have

field & $\therefore$ a π-sys → $\bigcup_{n=1}^\infty \mathcal{F}_n \subset \mathcal{L}$ ← a λ-system

$\therefore$ $\lambda\langle \bigcup_{n=1}^\infty \mathcal{F}_n \rangle \subset \mathcal{L}$

$\parallel$ Dynkin's π-λ

$\sigma\langle \bigcup_{n=1}^\infty \mathcal{F}_n \rangle$

$\therefore$ $E(X_\infty I_A) = E(XI_A)\ \forall A \in \mathcal{F}_\infty$

by "Good sets".

QED

Example

Lets use Lévy Thm to deduce the Kolmogorov 0-1 law.

Let $\mathcal{C}_1, \mathcal{C}_2, \dots$ be indep sub-σ-fields of $\mathcal{F}$.

The tail σ-field is defined as

$$\mathcal{T} = \bigcap_{n=1}^\infty \sigma\langle \mathcal{C}_1, \mathcal{C}_2, \dots \rangle$$

We want to show

$$P(A) \in \{0,1\} \quad \forall A \in \mathcal{T}.$$

Set $\mathcal{F}_n = \sigma\langle \mathcal{C}_1, \dots, \mathcal{C}_n \rangle$ and notice that for $A \in \mathcal{T}$,

$$\underbrace{E(I_A \mid \mathcal{F}_n)} \xrightarrow{a.e.} E(I_A \mid \mathcal{F}_\infty)$$
$= E(I_A)$ since $A \in \sigma\langle \mathcal{C}_{n+1}, \dots \rangle \perp \mathcal{F}_n$ by ANOVA

$= P(A)$

$\therefore$ $P(A) \stackrel{a.e.}{=} E(I_A \mid \mathcal{F}_\infty) \stackrel{a.e.}{=} I_A$

since $A \in \mathcal{T} \subset \sigma\langle \mathcal{C}_1, \mathcal{C}_2, \dots \rangle = \mathcal{F}_\infty$

$\therefore$ $P(A) = 0$ or 1.

Another way to quantify "bddness" of subMs is with "closers".

Definition: A pair $(X_\bullet, \mathcal{F}_\bullet)$ closes a subM $(X_n)_{n\geq 1}$ on the right if

r.v. ; sub σ-field of $\mathcal{F}$

$$X_1, X_2, \dots, X_n, \dots X_\bullet$$

is a subM w.r.t. $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_n, \dots, \mathcal{F}_\bullet$, i.e if

$$E(X_\bullet | \mathcal{F}_n) \overset{P\text{-a.e.}}{\geq} X_n \quad \forall n \in \mathbb{N}.$$

Definition: $(X_0, \mathcal{F}_0)$ is a nearest closer of $(X_n)_{n\geq 1}$ if

$$X_1, X_2, \dots X_0, X_\bullet$$

is a subM w.r.t $\mathcal{F}_1, \mathcal{F}_2, \dots \mathcal{F}_0, \mathcal{F}_\bullet$ for every closer $(X_\bullet, \mathcal{F}_\bullet)$ of $(X_n)_{n\geq 1}$.

The existance of a closer is equivalent to the UI condition:

Theorem: ($\exists$ closer $\Leftrightarrow$ X_n^+'s are UI)

If $(X_n)_{n\geq 1}$ is a subM w.r.t filt $(\mathcal{F}_n)_{n\geq 1}$ then

$$\exists \text{ a closer for } (X_n)_{n\geq 1} \Leftrightarrow (X_n^+)_{n\geq 1} \text{ are UI}$$

Proof:

($\Rightarrow$): Suppose $(X_\bullet, \mathcal{F}_\bullet)$ closes $(X_n)_{n\geq 1}$.

$\therefore$ $(X_\bullet^+, \mathcal{F}_\bullet)$ closes the subM $(X_n^+)_{n\geq 1}$ by Trans of Ms.

$\therefore$ $\underbrace{E(X_\bullet^+ | \mathcal{F}_n)}_{\text{These are UI by proof of Lévy's thm}} \overset{a.e.}{\geq} \underbrace{X_n^+}_{\therefore \text{ these are too.}} \quad \forall n \in \mathbb{N}$

($\Leftarrow$) Suppose $(X_n^+)_{n\geq 1}$ are UI.

$\therefore$ $\exists X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P)$ s.t.

$$X_n \xrightarrow{a.e.} X_\infty$$

we show that $(X_\infty, \mathcal{F}_\infty)$ is a closer.

Case 1: $X_n \geq c > -\infty \quad \forall n \in \mathbb{N}$.

$\therefore$ The X_n^-'s are UI & hence the X_n's are UI.

$\therefore$ $X_n \xrightarrow{L_1} X_\infty$ by the L_p-convergence Thm.

Now to show $E(X_\infty | \mathcal{F}_n) \overset{a.e.}{\geq} X_n$. Let $A \in \mathcal{F}_n$ so that

$$\int_A X_n \, dP \overset{\text{subM}}{\leq} \int_A E(X_{n+m} | \mathcal{F}_n) \, dP$$

$$= \int_A X_{n+m} \, dP$$

$$\xrightarrow{m\to\infty} \int_A X_\infty \, dP \quad \text{since } X_{n+m} I_A \xrightarrow{L_1} X_\infty I_A$$

$$= \int_A E(X_\infty | \mathcal{F}_n) \, dP$$

$\therefore$ $X_n \overset{a.e.}{\leq} E(X_\infty | \mathcal{F}_n)$ by our results on indefinite integrals.

$\therefore$ $(X_\infty, \mathcal{F}_\infty)$ closes $(X_n)_{n\geq 1}$.

Case 2: $X_n \in L_1(\Omega, \mathcal{F}, P)$.

Now case 1 applies to $\underbrace{X_n \vee c}_{\text{a subM since } \max(X_n, c) \text{ is a subM, also UI and bdd below.}}$ where $c > -\infty$.

Also $X_n \vee c \xrightarrow{L_1} X_\infty \vee c$

$\therefore$ $E(X_\infty \vee c | \mathcal{F}_n) \overset{a.e.}{\geq} X_n \vee c$

$\therefore$ $E(X_\infty | \mathcal{F}_n) \overset{a.e.}{\geq} X_n$ by taking limits as $c \to -\infty$ by MCT for $E(\cdot | \mathcal{F}_n)$

QED

Now we get the "closer" condition for a.e. convergence almost as a corollary.

Theorem: ($\exists$ closer $\Rightarrow$ a.e. convergence)

If the subM $(X_n)_{n\geq 1}$ has a closer then $\exists X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P)$ s.t.

$\mathcal{F}_\infty = \sigma\langle \mathcal{F}_1, \mathcal{F}_2, \dots \rangle$

$$X_n \xrightarrow{a.e.} X_\infty$$

and $(X_\infty, \mathcal{F}_\infty)$ is the nearest closer.

Proof:

By the previous result

$$\exists \text{ closer} \Leftrightarrow (X_n^+)'s \text{ are } uI$$
$$\Rightarrow E(X_n^+)'s \text{ are bdd}$$
$$\Rightarrow \exists X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P) \text{ s.t. } X_n \xrightarrow{a.e.} X_\infty.$$

The proof of the previous Thm also establishes that $(X_\infty, \mathcal{F}_\infty)$ closes $(X_n)_{n\geq 1}$.

Now we just show $(X_\infty, \mathcal{F}_\infty)$ is the nearest closer.

Let $(X_\bullet, \mathcal{F}_\bullet)$ be some closer.

We need to show

$$X_1, X_2, \dots X_\infty, X_\bullet \quad \text{is a subM}$$

w.r.t filtration $\mathcal{F}_1, \mathcal{F}_2, \dots \mathcal{F}_\infty, \mathcal{F}_\bullet$

it's a filtration since $\bigcup_{i=1}^\infty \mathcal{F}_i \subset \mathcal{F}_\bullet$, hence $\underbrace{\sigma\langle \bigcup_{i=1}^\infty \mathcal{F}_i \rangle}_{=\mathcal{F}_\infty} \subset \mathcal{F}_\bullet$

Notice it is sufficient to show

$$(*) \qquad E(X_\bullet | \mathcal{F}_\infty) \overset{P\text{-a.e.}}{\geq} X_\infty$$

To show $(*)$ use Lévy's smoothing result as follows

$$\begin{array}{ccc} E(X_\bullet|\mathcal{F}_n) & \overset{a.e.}{\geq} & X_n \\ \downarrow{\scriptstyle a.e.} & & \downarrow{\scriptstyle a.e.} \\ E(X_\bullet|\mathcal{F}_\infty) & \overset{a.e.}{\geq} & X_\infty \end{array}$$

since $(X_\bullet, \mathcal{F}_\bullet)$ is a closer

as way to be shown.

QED

Likelihood ratio Example from last Lecture

Two models P & Q for a random ω genating an infinite sequence of r.v.s

$$X = (X_1, X_2, \dots)$$

$$(\Omega, \mathcal{F}, P \text{ or } Q) \xrightarrow{(X_1, X_2, \dots)(\omega)} (\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$$

where

P & Q are distinguishable (a.e.) from one sample of X $\Leftrightarrow PX^{-1} \perp QX^{-1}$

$$\Leftrightarrow Q_\infty \perp P_\infty$$

where $\mathcal{F}_\infty := \sigma\langle X_i : i \geq 1\rangle$, $Q_\infty := Q|_{\mathcal{F}_\infty}$

$$\Leftrightarrow \frac{dQ_\infty^a}{dP_\infty} = 0$$

$$\overset{?}{\Leftrightarrow} \frac{dQ_n^a}{dP_n} \xrightarrow{a.e.} 0$$

where $\mathcal{F}_n := \sigma\langle X_1, \dots, X_n\rangle$, $Q_n := Q|_{\mathcal{F}_n}$

where $\frac{dQ_n^a}{dP_n}$ & $\frac{dQ_\infty^a}{dP_\infty}$ represents the finite & infinite data likelihood ratio

15

Recall that we showed $\frac{dQ_n^a}{dP_n}$ is a supM wrt $(\mathcal{F}_n)_{n\geq 1}$ under P. We will use this to show $\frac{dQ_n^a}{dP_n}$ converge

Theorem: $\frac{dQ_n^a}{dP_n} \xrightarrow{P\text{-a.e.}} \frac{dQ_\infty^a}{dP_\infty}$ as $n\to\infty$.

Proof:

$\frac{dQ_n^a}{dP_n}$ is a supM and non-neg by construction

$\therefore \left(-\frac{dQ_n^a}{dP_n}\right)_{n\geq 1}$ is a non-positive subM

$\therefore (0, \mathcal{F}_\infty)$ closes $\left(-\frac{dQ_n^a}{dP_n}\right)_{n\geq 1}$ on the right:

i.e. $-\frac{dQ_1^a}{dP_1}, -\frac{dQ_2^a}{dP_2}, \dots, 0$

is a subM w.r.t. $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_\infty$ because

$$0 = E(0 \mid \mathcal{F}_n) \geq -\frac{dQ_n^a}{dP_n} \quad \forall n\in\mathbb{N}.$$

By the closer Thm $\exists X_\infty \in L_1(\Omega, \mathcal{F}_\infty, P)$ s.t.

$$\frac{dQ_n^a}{dP_n} \xrightarrow{a.e.} \underbrace{X_\infty}_{\text{note } X_\infty \overset{a.e.}{\geq} 0} \quad \& \quad (-X_\infty, \mathcal{F}_\infty) \text{ is the nearest closer of } \left(-\frac{dQ_n^a}{dP_n}\right)_{n\geq 1}.$$

We show $X_\infty \overset{a.e.}{=} \frac{dQ_\infty^a}{dP_\infty}$

($\underline{\text{show } X_\infty \overset{a.e.}{\leq} \frac{dQ_\infty^a}{dP_\infty}}$)

Use that $-X_\infty$ closes $\left(-\frac{dQ_n^a}{dP_n}\right)_{n\geq 1}$ on the right.

Since X_∞ closes we have

$$E(X_\infty \mid \mathcal{F}_n) \overset{P\text{-a.e.}}{\leq} \frac{dQ_n^a}{dP_n}$$

16

$$\therefore \int_A X_\infty \, dP \leq \int_A \frac{dQ_n^a}{dP_n} dP \leq Q_\infty(A) \quad \forall A \in \bigcup_{k=1}^\infty \mathcal{F}_k$$

Since $\bigcup_{k=1}^\infty \mathcal{F}_k$ is a field generating $\mathcal{F}_\infty$ & $X_\infty \overset{a.e.}{\geq} 0$ we can apply "Good sets" to show

$$\underbrace{\int_A X_\infty \, dP}_{\text{a finite measure}} \leq Q_\infty(A) \quad \forall A \in \mathcal{F}_\infty$$

but $\frac{dQ_\infty^a}{dP_\infty}$ is the P-largest (by Leb Decomp) such, so we have

$$X_\infty \overset{P\text{-a.e.}}{\leq} \frac{dQ_\infty^a}{dP_\infty}$$

($\underline{\text{show } X_\infty \overset{a.e.}{\geq} \frac{dQ_\infty^a}{dP_\infty}}$)

Use that $-\frac{dQ_\infty^a}{dP_\infty}$ closes $\left(-\frac{dQ_n^a}{dP_n}\right)_{n\geq 1}$ on the right.

Indeed this follows from the same method used to show $\left(\frac{dQ_n^a}{dP_n}\right)_{n\geq 1}$ is a supM.

i.e. $-\frac{dQ_1^a}{dP_1}, -\frac{dQ_2^a}{dP_2}, \dots, \overset{\text{nearest closer}}{-X_\infty}, -\frac{dQ_\infty^a}{dP_\infty}, 0$

is a subM w.r.t. $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_\infty, \mathcal{F}_\infty, \mathcal{F}_\infty$

$$\therefore E\left(-\frac{dQ_\infty^a}{dP_\infty} \,\Big|\, \mathcal{F}_\infty\right) \overset{a.e.}{\geq} -X_\infty$$

$$\overset{a.e.}{\|}$$

$$-\frac{dQ_\infty^a}{dP_\infty}$$

QED

L_p convergence

To summarize what we know so far for subMs:

$$\sup_n E(X_n^+) < \infty \implies X_n \xrightarrow{a.e.} X_\infty \in L_1(\Omega, \mathcal{F}, P)$$
$$\Uparrow$$
$$X_n^+ \text{ are UI} \iff X_n \text{ is closable}$$

X_n have a closer $\implies X_n \xrightarrow{a.e.} X_\infty \in L_1$ & X_∞ is the nearest closer.

The next result shows that for $p \geq 1$

$$|X_n|^p \text{ UI} \implies X_n \xrightarrow{L_p} X_\infty \in L_p(\Omega, \mathcal{F}, P)$$

where X_∞ is the nearest closer.

Theorem: (subM L_p convergence Thm for $1 \leq p < \infty$)

Suppose $1 \leq p < \infty$ and $(X_n)_{n\geq 1}$ is a subM wrt filtration $(\mathcal{F}_n)_{n \geq 1}$.

If $|X_n|^p$ is UI then $\exists X_\infty \in L_p(\Omega, \mathcal{F}_\infty, P)$ s.t.

$$X_n \longrightarrow X_\infty \quad P\text{-a.e. \& in } L_p$$

where $(X_\infty, \underbrace{\mathcal{F}_\infty}_{=\sigma\langle \mathcal{F}_1, \mathcal{F}_2, \ldots \rangle})$ is the nearest closer of $(X_n)_{n\geq 1}$

Proof:

Clearly $|X_n|$'s are UI by the dilation criterion.

$\therefore (X_n^+)_{n\geq 1}$ are UI since $X_n^+ \leq |X_n|$

$\therefore X_n \xrightarrow{a.e.} X_\infty \in L_1(\Omega, \mathcal{F}, P)$ with X_∞ being the nearest closer.

Now all we need is to show

$$X_n \xrightarrow{L_p} X_\infty$$

but this follows immediately by the L_p convergence Thm, since $(|X_n|^p)_{n\geq 1}$ are UI.

QED.

Theorem: (Checking $|X_n|^p$ UI for $X_n \geq 0$ subM)

If $(X_n)_{n\geq 1}$ is a subM, $X_n \geq 0$ $\forall n$ & $p > 1$ then

$$X_n^p \text{ are UI} \iff \sup_n E(X_n^p) < \infty$$
$$\iff E(\sup_n X_n^p) < \infty$$

Proof:

$(X_n^p \text{ UI} \Rightarrow \sup_n E(X_n^p) < \infty)$

Follows by UI properties

$(\sup_n E(X_n^p) < \infty \Rightarrow E(\sup_n X_n^p) < \infty)$

Let $M_n = \max_{1 \leq k \leq n} X_k$ & $q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$

Now

$$E M_n^p = \int_0^\infty P(M_n^p \geq t)\,dt$$

$$= \int_0^\infty P(M_n \geq t^{1/p})\,dt$$

$w = t^{1/p}$, $\therefore dw = \frac{t^{\frac{1}{p}-1}}{p} dt$, $\therefore dw = \frac{w^{1-p}}{p} dt$

$$= \int_0^\infty P(M_n \geq w)\, p w^{p-1}\,dw$$

$$\stackrel{\text{Kolm}}{\leq} \int_0^\infty w^{-1} E(X_n I_{\{M_n \geq w\}})\, p w^{p-1}\,dw$$

$$\stackrel{\text{Fubini}}{=} p E\left[\int_0^\infty X_n I_{\{M_n \geq w\}} w^{p-2}\,dw\right]$$

$$= p E\left[X_n \left[\frac{w^{p-1}}{p-1}\right]_0^{M_n}\right]$$

$$\therefore\ E(M_n^p) = \frac{p}{p-1} E(X_n M_n^{p-1})$$

$\frac{1}{p}+\frac{1}{q}=1$ $\therefore\ q+p=pq$

$$\overset{\text{Hölder}}{\leq} \frac{p}{p-1} \|X_n\|_p \|M_n^{p-1}\|_q$$

$$= \frac{1}{1-\frac{1}{p}} [E(X_n^p)]^{1/p} [E(M_n^{\overbrace{qp-q}^{p}})]^{1/q}$$

$$= \frac{1}{q} [E(X_n^p)]^{1/p} [E(M_n^p)]^{1/q}$$

$$\therefore\ [E(M_n^p)]^{1-\frac{1}{q}} \leq q [E(X_n^p)]^{1/p}$$

since $E(M_n^p) < \infty$ because $M_n \leq \underbrace{X_1+\cdots+X_n}_{\in L_p}$

$$\therefore\ E(M_n^p) \leq q^p E(X_n^p) \leq q^p \sup_k E(X_k^p) \quad (*)$$

Now since $M_n^p = \max_{1\leq k\leq n} X_k^p \uparrow \sup_n X_n^p$

we can take a limit in n in $(*)$ to obtain

$$E(\sup_n X_n^p) \leq q^p \sup_k E(X_k^p).$$

($\underline{E(\sup_n X_n^p) < \infty \implies X_n^p \text{ are UI}}$)

$$X_n^p \leq \sup_n X_n^p =: Y$$

Since $Y \in L_1$ the single Y is UI.

$\therefore$ X_n^p's are UI.

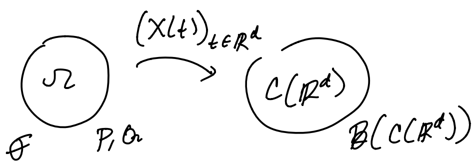
QED

Application of $Q \perp P$ & $Q \ll P$ for Gaussian random fields (GRF)

Let $(\Omega, \mathcal{F})$ be a measure space & P, Q be two probability measures on $(\Omega, \mathcal{F})$.

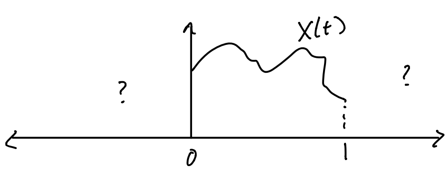
Suppose $(X(t))_{t \in \mathbb{R}^d}$ is a collection of r.v.s defined on $(\Omega, \mathcal{F})$ s.t. under both P & Q $(X(t))_{t \in \mathbb{R}^d}$ induces a measure on $C(\mathbb{R}^d)$ s.t. $\forall t_1, \dots, t_n \in \mathbb{R}^d$

$(X(t_1), \dots, X(t_n))$ is multivariate Gaussian.



f.d.d of PX^{-1} & QX^{-1} are Gaussian.

Suppose I sample $w \in \Omega$ according to P or Q & show you $X(t)(w)$ $\forall t \in [0,1]^d$



You need to determine if I used P or Q to sample $w \in \Omega$.

The information available to you is encoded in $\sigma\langle X(t) : t \in [0,1]^d \rangle$.

If we let P, Q denote the restriction to $\sigma\langle X(t) : t \in [0,1]^d \rangle$ then

$P \perp Q \iff$ Perfect discrimination possible (w.p. 1)

$P \equiv Q \iff$ Impossible to distinguish P from Q

Example: GRF's are characterized by $m(t) := E(X(t))$ &

$$K(t,s) := \operatorname{cov}(X(t), X(s)).$$

Suppose that $X(t)$ is defined on $\mathbb{R}$ &

Under P: $m(t) = 0$, $K(t,s) = e^{-|t-s|}$

Under Q: $m(t) = 0$, $K(t,s) = (1 + |t-s|)e^{-|t-s|}$

Now let

$$X_n = n\left(X(\tfrac{1}{n}) - X(0)\right)$$

$$\therefore \operatorname{var}_P(X_n) = n^2\left(K(\tfrac{1}{n}, \tfrac{1}{n}) - 2K(\tfrac{1}{n}, 0) + K(0,0)\right)$$

$$= 2n^2\left(1 - e^{-1/n}\right) \sim n \to \infty$$

$$\operatorname{var}_Q(X_n) = 2n^2\left(1 - e^{-1/n}(1 + \tfrac{1}{n})\right)$$

$$= 2n^2\left(1 - \left(1 - \tfrac{1}{n} + \tfrac{1}{2n^2} + O(\tfrac{1}{n^3})\right)(1 + \tfrac{1}{n})\right)$$

$$= 2n^2\Big(\cancel{1} - \big(\cancel{1} - \cancel{\tfrac{1}{n}} + \tfrac{1}{2n^2} + O(\tfrac{1}{n^3})\big) - \big(\cancel{\tfrac{1}{n}} - \tfrac{1}{n^2} + O(\tfrac{1}{n^3})\big)\Big)$$

$$\to 1$$

$\therefore$ under P: $X_n \sim N(0, \theta(n))$

under Q: $X_n \sim N(0, \theta(1))$

$\therefore$ one can find $F_n \in \mathcal{F}$ s.t.

$P(F_n) \xrightarrow{n\to\infty} 0$

$Q(F_n) \xrightarrow{n\to\infty} 1$

scale X_n so $\mathrm{var}(c_n X_n) \to \infty$ under P and $\to 0$ under Q

$\therefore$ $P \perp Q$ & perfect discrimination is possible from "one sample" of $X(t)$ on $t \in [0,1]$.

Now choose an infinite sequence of r.v.s $X_1, X_2, \ldots \in$ finit linear span $\{X(t) : t \in [0,1]^d\}$ which are linearly independent under both P & Q (if no such sequence exists then $P \perp Q$).

Let $\mathcal{F}_n = \sigma\langle X_1, \ldots, X_n\rangle$

$\mathcal{F}_\infty = \sigma\langle X_1, X_2, \ldots\rangle$

$P_n = P|_{\mathcal{F}_n}$, $Q_n = Q|_{\mathcal{F}_n}$ $\forall n \in \mathbb{N} \cup \{\infty\}$.

Note: Since the sample paths of X are continuous we can choose X_n's so that

$P_\infty \equiv Q_\infty \Leftrightarrow P \equiv Q$ &

$P_\infty \perp Q_\infty \Leftrightarrow P \perp Q$ remember these are defined on $\sigma\langle X(t) : t \in [0,1]^d\rangle$

Let p_n & q_n be the density of $P(X_1, \ldots, X_n)^{-1}$ & $Q(X_1, \ldots, X_n)^{-1}$ w.r.t Lebesgue measure on $\mathbb{R}^n$.

Let

$$KL(P_n, Q_n) = E_P \log \frac{dP_n}{dQ_n} - E_Q \log \frac{dP_n}{dQ_n}$$

$$= E_P \log \frac{p_n(X_1, \ldots, X_n)}{q_n(X_1, \ldots, X_n)} - E_Q \log \frac{p_n(X_1, \ldots, X_n)}{q_n(X_1, \ldots, X_n)}$$

Claim: For GRF models P & Q

(i) $KL(P_n, Q_n)$ is monotonic

(ii) $KL(P_n, Q_n) \to \infty \Rightarrow P_\infty \perp Q_\infty$

(iii) $KL(P_n, Q_n) \to c < \infty \Rightarrow P_\infty \equiv Q_\infty$

Proof:

(i) Follows since

$$KL(P_n, Q_n) = E_P\Big(\underbrace{-\log \overbrace{\frac{dQ_n}{dP_n}}^{\text{a martingale under } P \text{ w.r.t. } (\mathcal{F}_n)_{n\geq 1}}}_{\text{a subM under } P}\Big) + E_Q\Big(\underbrace{-\log \overbrace{\frac{dP_n}{dQ_n}}^{\text{a martingale under } Q \text{ w.r.t } (\mathcal{F}_n)_{n\geq 1}}}_{\text{a subM under } Q}\Big)$$

$\therefore$ a subM under P since $-\log(x)$ is convex.

For (ii) Suppose $KL(P_n, Q_n) \to \infty$.

For each n find a linear transformation of $(X_1, \ldots, X_n)$ to $(Z_{1,n}, \ldots, Z_{n,n})$ s.t.

under Q: $(Z_{1,n}, \ldots, Z_{n,n}) \sim N(0, I)$

under P: $(Z_{1,n}, \ldots, Z_{n,n}) \sim N\left(\begin{pmatrix} m_{1,n} \\ \vdots \\ m_{n,n}\end{pmatrix}, \begin{pmatrix} \lambda_{1,n}^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_{n,n}^2\end{pmatrix}\right)$

24

Now

$$\frac{dP_n}{dQ_n} = \frac{p_n(X_1,\dots,X_n)}{q_n(X_1,\dots,X_n)}$$

$$= \frac{\prod_{k=1}^{n} \lambda_{kn}^{-1} \exp\left(-\frac{1}{2}\frac{(z_{kn}-m_{kn})^2}{\lambda_{kn}^2}\right)}{\prod_{k=1}^{n}\exp\left(-\frac{1}{2}z_{kn}^2\right)}$$

so that

$$E_P\left(\log\frac{dP_n}{dQ_n}\right) = E_P\left(-\frac{1}{2}\sum_{k=1}^{n}\left[\frac{(z_{kn}-m_{kn})^2}{\lambda_{kn}^2} - z_{kn}^2\right]\right) - \sum_{k=1}^{n}\log\lambda_{kn}$$

$$= -\frac{1}{2}\sum_{k=1}^{n}\left[1 - m_{kn}^2 - \lambda_{kn}^2\right] - \sum_{k=1}^{n}\log\lambda_{kn}$$

$$E_Q\left(\log\frac{dP_n}{dQ_n}\right)$$

$$= -\frac{1}{2}\sum_{k=1}^{n}\left[\frac{1+m_{kn}^2}{\lambda_{kn}^2} - 1\right] - \sum_{k=1}^{n}\log\lambda_{kn}$$

$\therefore$

$$KL(P_n,Q_n) = E_P\left(\log\frac{dP_n}{dQ_n}\right) - E_Q\left(\log\frac{dP_n}{dQ_n}\right)$$

$$= -\frac{1}{2}\sum_{k=1}^{n}\left[2 - m_{kn}^2 - \lambda_{kn}^2 - \frac{1+m_{kn}^2}{\lambda_{kn}^2}\right]$$

$$= \frac{1}{2}\sum_{k=1}^{n}\left[\frac{(1-\lambda_{kn}^2)^2}{\lambda_{kn}^2} + m_{kn}^2 + \frac{m_{kn}^2}{\lambda_{kn}^2}\right]$$

25

Also

$$\mathrm{var}_Q\left(\log\frac{dP_n}{dQ_n}\right) = \frac{1}{2}\sum_{k=1}^{n}\left[\frac{(1-\lambda_{kn}^2)^2 + 2m_{kn}^2}{\lambda_{kn}^4}\right]$$

$$\mathrm{var}_P\left(\log\frac{dP_n}{dQ_n}\right) = \frac{1}{2}\sum_{k=1}^{n}\left[(1-\lambda_{kn}^2)^2 + 2m_{kn}^2\lambda_{kn}^2\right]$$

If $\inf_{k,n}\lambda_{k,n}^2 = 0$, $\exists\, k_n$ s.t. $\lambda_{k_n,n}^2 \to 0$ then

$\mathrm{var}_Q(Z_{k_n,n}) = 1$ but $\mathrm{var}_P(Z_{k_n,n}) = \lambda_{k_n,n}^2 \to 0$

which implies $P \perp Q$. Similarly $P \perp Q$

if $\sup_{k,n}\lambda_{k,n}^2 = \infty$.

So suppose

$$0 < \inf_{k,n}\lambda_{k,n}^2 \le \sup_{k,n}\lambda_{k,n}^2 < \infty$$

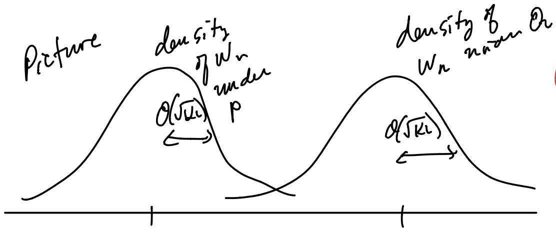
In this case

$$KL(P_n,Q_n) \asymp \mathrm{var}_P\left(\log\frac{dP_n}{dQ_n}\right) \asymp \mathrm{var}_Q\left(\log\frac{dP_n}{dQ_n}\right)$$

Now setting $W_n := \log\dfrac{dP_n}{dQ_n}$

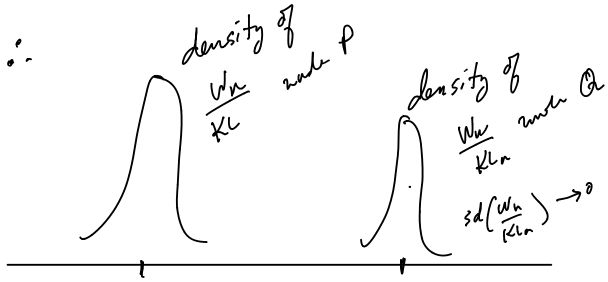
$$KL(P_n,Q_n) = \underbrace{E_P(W_n) - E_Q(W_n)}_{\substack{\text{same order as} \\ \mathrm{var}_P(W_n)\ \&\ \mathrm{var}_Q(W_n)}} \to \infty$$

Picture



$KL(P_n, Q_n) \longrightarrow \infty$

$\therefore$



$O(1)$

$\therefore P_\infty \perp Q_\infty$

For (iii) Suppose $KL(P_n, Q_n) \longrightarrow c < \infty$

and show $P_\infty \equiv Q_\infty$.

Suppose $P_\infty \not\equiv Q_\infty$ so $\exists A \in \mathcal{F}_\infty$ s.t.

$$P(A) > 0 \quad \text{but} \quad Q(A) = 0$$

Approximate A with $A_n \in \bigcup_{k=1}^{\infty} \mathcal{F}_k$ ← Field generators of $\mathcal{F}_\infty$

such that

$$P(A_n) \longrightarrow P(A) > 0$$

$$Q(A_n) \longrightarrow Q(A) = 0$$

Let $\mathcal{Q}_n := \{\emptyset, \Omega, A_n, A_n^c\}$

Now

$$KL(P_n, Q_n) = E_P \log \frac{dP_n}{dQ_n} - E_Q \log \frac{dP_n}{dQ_n}$$

since

$$-E_Q \log \frac{dP_n}{dQ_n} \overset{\text{jensen}}{\geq} -\log E_Q \frac{dP_n}{dQ_n} = -\log \int_\Omega \frac{dP_n}{dQ_n} dQ = 0$$

we have

$$KL(P_n, Q_n) \geq E_P \log \frac{dP_n}{dQ_n}$$

$$= -E_P \log \frac{dQ_n}{dP_n}$$

$$= E_P\left(E_P\left(-\log \frac{dQ_n}{dP_n} \,\middle|\, \mathcal{Q}_n\right)\right)$$

$$\overset{\text{jensen}}{\geq} E_P\left(-\log \underbrace{E_P\left(\frac{dQ_n}{dP_n} \,\middle|\, \mathcal{Q}_n\right)}_{= \begin{cases} Q(A_n)/P(A_n) & \text{on } A_n \\ Q(A_n^c)/P(A_n^c) & \text{on } A_n^c \end{cases}}\right)$$

$$= -E_P\left(\log\left(\frac{Q(A_n)}{P(A_n)}\right) I_{A_n} + \log\left(\frac{Q(A_n^c)}{P(A_n^c)}\right) I_{A_n^c}\right)$$

$$= -P(A_n)\log\left(\frac{Q(A_n)}{P(A_n)}\right) - P(A_n^c)\log\left(\frac{Q(A_n^c)}{P(A_n^c)}\right)$$

$$\longrightarrow \infty$$

Contradiction.

$\therefore P_\infty \equiv Q_\infty$.

QED