MEASURE THEORETIC PROBABILITY III HW 6

TOMMENIX YU
ID: 12370130
STAT 38300
DUE THU MAY 4TH, 2023, 11AM

Discussed with classmates.

Exercise 1.

Proof.

(i)$\not\Rightarrow$ (ii):

Note that this is in general not true! This is because there's no requirement that $X_n \in L^1$.

So we can just pick X_1 to be unbounded, i.e. $\mathbb{E}[|X_1|] = \infty$ and let $X_{i\geq 2}$ to behave very well, then the family cannot be UI. This also does not violate the condition since weak convergence is a asymptotic property, and has nothing to do with the first few terms.

(i)$\Rightarrow$ (ii), assuming that $X_i \in L^1$:

This is fine since weak converge implies for all continuous bounded f we have

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$$

and so we define

$$f_n := \begin{cases} n & |x| \geq n \\ |x| & -n \leq x \leq n \end{cases}$$

in which case we have (since $\mathbb{E}[f_K(X)] \leq \mathbb{E}[|X|] = C < \infty$) that

$$\mathbb{E}[|X_n|] - \int_{|X_n|\geq K} X_n d\mathbb{P} = \mathbb{E}[f_K(X_n)] \to \mathbb{E}[f_K(X)] = \mathbb{E}[|X|] - \int_{|X|\geq K} X d\mathbb{P}$$

and since $\mathbb{E}[|X_n|] \to \mathbb{E}[|X|]$ is known, and $\int_{|X|\geq K} X d\mathbb{P} \to 0$ so the only term left also goes to 0, i.e.

$$\int_{|X_n|\geq K} X_n d\mathbb{P} < \varepsilon.$$

(ii)$\rightarrow$ (i):

The trick here is to argue for the existence of $\mathbb{E}[|X|]$, which is not given.

So we define f_k as the truncated absolute value function, then we define

$$a_{k,n} := \mathbb{E}[f_k(X_n)]$$

then by UI we know there's a uniform constant C in k, n such that $a_{k,n} \leq C$. But by weak converge we know for fixed k we have

$$\lim_{n\to\infty} a_{k,n} = \mathbb{E}[f_k(X)] =: \alpha_k$$

and since each term in the limit, as a real number, is bounded by C we know the limit is bounded by C. Thus, $\alpha_k \in [-C, C]$, a compact set in $\mathbb{R}$, so we know that there exists a subsequent $\phi(k)$ that $\alpha_{\phi(k)} \to \alpha$, where really we mean

$$\alpha := \lim_{k\to\infty} \mathbb{E}[f_{\phi(k)}(X)].$$

Now we use the diagonal argument, just take the diagonal subsequence $a_{\phi(n),\zeta(n)}$ then we know that

$$a_{\phi(n),\zeta(n)} \to \alpha \leq C$$

by a diagonal argument (pick each $\zeta(n)$ such that $||a_{\phi(n),\zeta(n)} - \alpha_{\phi(n)}|| \leq \varepsilon$). But it does not matter! We can peel off ϕ by noting that for any two convergent subsequence $\phi_1(n)$ and $\phi_2(n)$ we know they converges to the same limit, i.e. $\lim_{k\to\infty} a_{\phi_1(n),\zeta(n)} = \lim_{k\to\infty} a_{\phi_2(n),\zeta(n)}$ since, knowing what f_{ϕ_1} and f_{ϕ_2} are we have by UI, for the set $S = \left\{\max\{\phi_1(n), \phi_2(n)\} \geq |X_n| \geq \min\{\phi_1(n), \phi_2(n)\}\right\}$

$$\begin{aligned} a_{\phi_1(n),\zeta(n)} - a_{\phi_2(n),\zeta(n)} &= \int_S |X_n| - \min\{\phi_1(n), \phi_2(n)\} d\mathbb{P} + |\max - \min| \int_{X \geq \max} 1 d\mathbb{P} \\ &\geq \int_{|X_n| \geq \min\{\phi_1(n),\phi_2(n)\}} |X_n| d\mathbb{P} \to 0 \end{aligned}$$

thus the two limits are the same, and since we're in a compact set every subsequence away from α has a subsequence that goes to α, so we know $a_{n,\zeta(n)} \to \alpha$.

So we just define

$$Y := \lim_{n\to\infty} \mathbb{E}[f_n(X_{\zeta(n)})]$$

which exists by above argument. Then we know that $\zeta(n) \geq n$ is a growing sequence so we know that $\mathbb{E}[f_n(X_{\zeta(n)})]] \overset{as}{\to} Y$. Yet by UI we know almost sure convergence is L^1 convergence, so the convergence is L^1, in particular $\mathbb{E}[|X_n|] \to \mathbb{E}[|Y|] = \mathbb{E}[Y]$. Here ζ does not matter any more since if a subsequence converge then all converges, as they are uniformly bounded.

But UI + L^1 means convergence in probability, which mean weak convergence, so $X_n \overset{w*}{\rightharpoonup} Y$, yet by Hahn Banach we know that the weak* limit is unique since L^1 is a Hilbert space, so $X = Y$ and we get L^1 convergence, i.e. $\mathbb{E}[|X_n|] \to \mathbb{E}[|X|]$, as well as $\mathbb{E}[X] = \mathbb{E}[Y] = \alpha \leq C < \infty$. $\square$

Exercise 2.

Proof.

(i):

UI means for some K, $\int_{|X_n|\geq K} |X_n| d\mathbb{P} < \varepsilon$, where as tight means

$$\sup_n \mathbb{P}(|X_n| \geq K) \leq \varepsilon \iff \int_{|X_n|\geq K} 1 d\mathbb{P} < \varepsilon$$

so we just pick $K' = \max K, 1$ then we've shown tightness.

(ii):

Markov says

$$\mathbb{P}(|X_n| \geq K) \leq \frac{\mathbb{E}[|X_n|]}{K} \leq \frac{C}{K}$$

for some $C < \infty$ (this exists by condition). But then take sup on both sides we have

$$\sup \mathbb{P}(|X_n| \geq K) \leq \sup \frac{\mathbb{E}[|X_n|]}{K} \leq \frac{C}{K}$$

so we see that everything works through. Now we just pick large enough K such that $K \geq 1$ and $\frac{C}{K} < \varepsilon$ to get the result. $\square$

Exercise 3.

Proof.

We just write out

$$\{\tau \leq \sigma\} \cap \{\tau \vee \sigma \leq n\} = \bigcup_{k=0}^{n} \left(\{\sigma = k\} \cap \left(\bigcup_{i=0}^{k} \{\tau = i\} \right) \right) \in \mathcal{F}_n$$

since each term on the right is in $\mathcal{F}_n$. Thus we are done.

□

Exercise 4.

Proof.

(a) We require $R \ll Q \ll P$ such that the derivatives are defined. In this case, we note that $R \ll P$ is directly obtained from definition.

Now we have for all $F \in \mathcal{F}$ we have that $\frac{dR}{dQ} \in \mathcal{F}$ so we can approximate it with simple functions $\frac{dR}{dQ} = \lim \sum a_i \mathbb{1}_{A_i}$ for $A_i \in \mathcal{F}$, and for simple functions we have

$$\int_F \sum a_i \mathbb{1}_{A_i} \frac{dQ}{dP} dP = \sum \int_F a_i \frac{dQ}{dP} dP = \sum \int_F a_i dQ$$

and we now can use DCT to pass the limit by MCT since we can pick the simple function to be monotone (since positive) then we have

$$\int_F \frac{dR}{dQ}\frac{dQ}{dP} dP = \int_F \frac{dR}{dQ} dQ = R(F)$$

so

$$\frac{dR}{dQ}\frac{dQ}{dP} = \frac{dR}{dP}$$

since above holds for all $F \in \mathcal{F}$.

(b):

For any $F \in \mathcal{F}_n$ we have

$$\begin{aligned}\int_F Z_n dP &= \int_F \mathbb{E}_P\left[\frac{dQ}{dP}\middle|\mathcal{F}_n\right] \mathbb{E}_Q\left[\frac{dR}{dQ}\middle|\mathcal{F}_n\right] dP = \int_F \mathbb{E}_P\left[\frac{dQ}{dP}\mathbb{E}_Q\left[\frac{dR}{dQ}\middle|\mathcal{F}_n\right]\middle|\mathcal{F}_n\right] dP \\ &= \int_F \frac{dQ}{dP}\mathbb{E}_Q\left[\frac{dR}{dQ}\middle|\mathcal{F}_n\right] dP \overset{Y_n \in \mathcal{F}_n}{=} \int_F \mathbb{E}_Q\left[\frac{dR}{dQ}\middle|\mathcal{F}_n\right] dQ = \int_F \frac{dR}{dQ} dQ = R(F)\end{aligned}$$

and note that for Z_{n+1} everything is the same since $F \in \mathcal{F}_n \subset \mathcal{F}_{n+1}$ and $Y_{n+1} \in \mathcal{F}_{n+1}$, which justifies the take out what is known, and the simple function approximation part.

(c):

We know off hand that $\frac{dR}{dP}$ exist, so we define $Z := \frac{dR}{dP}$ then if we can show $Z_n = \mathbb{E}[Z|\mathcal{F}_n]$ then since Z_n are the conditional expectation of some random variable that is L^1 bounded (since probability measures) then we know they are UI (14.2).

But that's very easy to show because obviously for all n, and for all $F \in \mathcal{F}_n$

$$\int_F Z dP = R(F)$$

by what it is, then we are done. □

Exercise 5.

Proof.

(a): Assume that the R.N. derivatives are bounded:

We have by Chain rule (and mutually absolutely continuous means all terms below are well defined) that

$$\dot{l}_n(\theta) = \frac{dP_0}{dP_\theta}\frac{d}{d\theta}\frac{dP_\theta}{dP_0}$$

where for convenience we denote $Z_n(\theta) := Z_n := \frac{dP_\theta}{dP_0}$. Thus we write out:

$$\int_{F\in\mathcal{F}_n} \dot{l}_n(\theta)dP_\theta = \int_F \frac{dP_0}{dP_\theta}\frac{d}{d\theta}\frac{dP_\theta}{dP_0}dP_\theta = \int_F \lim_{h\to 0}\frac{\frac{dP_{\theta+h}}{dP_0} - \frac{dP_\theta}{dP_0}}{h}dP_0$$

as everything in the quotient is F_n measurable, and taking limit does not change that. Thus if we assume that we can pass the limit, i.e. we assume that the derivatives are bounded in θ, then we can pass the limit to get

$$\begin{aligned}\int_F \dot{l}_n(\theta)dP_\theta &\overset{DCT}{=} \lim_{h\to 0}\frac{1}{h}\left[\int_F \left.\frac{dP_{\theta+h}}{dP_0}\right|_n dP_0 - \int_F \left.\frac{dP_\theta}{dP_0}\right|_n dP_0\right] \\ &= \frac{d}{d\theta}\int_F \left.\frac{dP_\theta}{dP_0}\right|_n dP_0 \overset{F\in\mathcal{F}_n}{=} \frac{d}{d\theta}P_\theta(F)\end{aligned}$$

Now we see that there's nothing stopping us from doing the same for $\dot{l}_{n+1}(\theta)$ since all is $\mathcal{F}_{n+1}$ measurable. Thus $\dot{l}_n(\theta)$ is a Martingale.

(b): We assume that both the R.N. derivatives and their derivatives in θ are bounded.

Now we show both terms are Martingale. For the first we have

$$\int_{F\in\mathcal{F}_n} \ddot{l}_n(\theta)dP_\theta = \int_F \frac{d}{d\theta}\left.\frac{dP_0}{dP_\theta}\right|_n \frac{d}{d\theta}\left.\frac{dP_\theta}{dP_0}\right|_n + \left.\frac{dP_0}{dP_\theta}\right|_n \frac{d^2}{d\theta^2}\left.\frac{dP_\theta}{dP_0}\right|_n dP_\theta$$

and we see clearly that the second term is nothing but passing limit as we did for (a). So the second term is $\frac{d}{d\theta}P_\theta(F)$. As for the first term we expand to get (we use DCT multiple times by our assumption that all are bounded)

$$\begin{aligned}\int_F \frac{d}{d\theta}\left.\frac{dP_0}{dP_\theta}\right|_n \frac{d}{d\theta}\left.\frac{dP_\theta}{dP_0}\right|_n dP_\theta &= \int_F \lim_{s\to 0}\left.\frac{\frac{dP_0}{dP_{\theta+s}} - \frac{dP_0}{dP_\theta}}{s}\right|_n \lim_{h\to 0}\left.\frac{\frac{dP_{\theta+h}}{dP_0} - \frac{dP_\theta}{dP_0}}{h}\right|_n dP_\theta \\ = \lim_{s\to 0}\frac{1}{s}\Bigg[\int_F \left.\frac{dP_0}{dP_{\theta+s}}\right|_n \lim_{h\to 0}\left.\frac{\frac{dP_{\theta+h}}{dP_0} - \frac{dP_\theta}{dP_0}}{h}\right|_n dP_\theta &- \int_F \left.\frac{dP_0}{dP_\theta}\right|_n \lim_{h\to 0}\left.\frac{\frac{dP_{\theta+h}}{dP_0} - \frac{dP_\theta}{dP_0}}{h}\right|_n dP_\theta\Bigg]\end{aligned}$$

and this might seem problematic because of the term $\frac{dP_0}{dP_{\theta+s}}dP_\theta$ does not cancel. But that's no matter since we have

$$\int_F \frac{dP_0}{dP_{\theta+s}}dP_\theta - \int_F dP_0 \to 0$$

thus we can still validly pass that limit. So we get

$$\int_F \frac{d}{d\theta}\frac{dP_0}{dP_\theta}\bigg|_n \frac{d}{d\theta}\frac{dP_\theta}{dP_0}\bigg|_n dP_\theta = \frac{d}{d\theta}\int_F \frac{dP_0}{dP_\theta}\bigg|_n \lim_{h\to 0}\frac{\frac{dP_{\theta+h}}{dP_0} - \frac{dP_\theta}{dP_0}}{h}\bigg|_n dP_\theta = \frac{d^2}{d\theta^2}P_\theta(F)$$

and thus

$$\int_{F\in\mathcal{F}_n} \ddot{l}_n(\theta)dP_\theta = 2\frac{d^2}{d\theta^2}P_\theta(F)$$

which is independent of n, so this term is a martingale.

As for the other term, exactly the same passing in the limit yields

$$\int_F [\dot{l}(\theta),\dot{l}(\theta)]_{n+1} - [\dot{l}(\theta),\dot{l}(\theta)]_n dP_\theta = \int_F \left(\frac{dP_0}{dP_\theta}\frac{d}{d\theta}\frac{dP_\theta}{dP_0}\bigg|_n - \frac{dP_0}{dP_\theta}\frac{d}{d\theta}\frac{dP_\theta}{dP_0}\bigg|_{n-1}\right)^2 dP_\theta$$

$$= \int_F \frac{dP_0}{dP_\theta}\bigg|_n \left(\frac{d}{d\theta}\frac{dP_\theta}{dP_0}\bigg|_n\right)^2 dP_0 + \int_F \frac{dP_0}{dP_\theta}\bigg|_{n-1}\left(\frac{d}{d\theta}\frac{dP_\theta}{dP_0}\bigg|_{n-1}\right)^2 dP_0$$

$$-2\int_F \frac{dP_0}{dP_\theta}\bigg|_n \frac{dP_0}{dP_\theta}\bigg|_{n-1}\left(\frac{d}{d\theta}\frac{dP_\theta}{dP_0}\bigg|_n\right)\left(\frac{d}{d\theta}\frac{dP_\theta}{dP_0}\bigg|_{n-1}\right)dP_\theta$$

but we see the first term is equal to one of the two minused terms by putting the condition on $n-1$ term inside d. The other two is dealt with similarly so the above is 0, which means it is a Martingale.

(c): Use Taylor on $l'(\hat{\theta}_n)$ to get

$$0 = l'(\hat{\theta}_n) = l'(\theta) + l''(\tilde{\theta})(\hat{\theta}_n - \theta)$$

and shifting term we get

$$\frac{\hat{\theta}_n - \theta}{\sqrt{n}} = \frac{l'(\theta)/\sqrt{n}}{-l''(\tilde{\theta})/n}$$

and if we can prove two things:

(1) The nominator goes weakly to $N(0,\sigma^2)$;
(2) The denominator goes in probability to a constant σ^2.

then dividing and we get the result. For the first result we use Martingale CLT, where the martingale is shown in part (a), and since we need to assume bounded in order to even show the first two parts we automatically get

$$\frac{1}{\sqrt{n}}\mathbb{E}[\max_{1\le n\le N}|Z_n|] \to 0$$

where $Z_n = \left.\frac{dP_\theta}{dP_0}\right|_n$. Thus we only need to show the other condition, but that is direclty given to us that

$$\frac{[\dot{l}(\theta), \dot{l}(\theta)]_n}{n} \xrightarrow{p} \sigma^2$$

and so we use MG CLT to get that the nominator goes weakly to $N(0, \sigma^2)$.

Now, as for the denominator, we have by part (b) that

$$\frac{[\dot{l}(\theta), \dot{l}(\theta)]_n}{n} - \frac{l''(\theta)}{n} \xrightarrow{p} 0$$

so as long as we can show $\hat{\theta}_n \to \theta$ in probability then we are done, since we can squeeze $\tilde{\theta}$ between $\hat{\theta}_n$ and θ.

So we just assume that $\hat{\theta}_n \to \theta$ in probability and by above argument we automatically get the second result.

$\square$