Inequalities Lecture 11

Proposition 1 $\mathcal{L}_s \subset \mathcal{L}_r$, for any $0 < r < s$, if $\mu(\Omega) < \infty$.

Proof • Recall that $\mathcal{L}_p = \{X : \int |X|^p < \infty\}$.

[Notation: Recall also that we use the following notation; for any $(\Omega, \mathcal{A}, \mu)$, X and h measurable.

$$E\,h(X) = \int h(X)\,d\mu = \int h(X) = \int h(X(\omega))\,d\mu(\omega).$$

So, $\mathcal{L}_p = \{X : E|X|^p < \infty\}$, for $p \geq 0$]

Let $0 < r < s$.
Let $X \in \mathcal{L}_s \Rightarrow \int |X|^s d\mu < \infty$. To show that $X \in \mathcal{L}_r$, we need to show that $\int |X|^r d\mu < \infty$.

• Note that for any $a > 0$ and $0 < r < s$ we have $a^r < 1 + a^s$. (Consider two cases: $0 < a \leq 1$ & $a > 1$, each of which is obvious).

We apply this ~~for~~ $|x|^r$ and $|x|^s$. Hence:

$$\int |X|^r < \int 1 + |X|^s = \int 1 + \int |X|^s < \infty$$

qed.

Proposition 2 C_r - inequality

$E|X+Y|^r \leq C_r E|X|^r + C_r E|Y|^r$, where

a) $C_r = 2^{r-1}$, for $r \geq 1$.

b) $C_r = 1$, for $0 < r \leq 1$.

Proof: Page 47.

Proposition 3 (Hölder's inequality)

① $E|XY| \leq E^{1/r}|X|^r E^{1/s}|Y|^s$, for $r > 1$ and s such that $\frac{1}{r} + \frac{1}{s} = 1$.

Notation : $\|X\|_p \equiv \left(E|X|^p\right)^{1/p}$ = the p-th norm of X = the norm of X in $\mathcal{L}_p$ for $p > 0$

Thus ① $\Longleftrightarrow$ $\|XY\|_1 \leq \|X\|_r \cdot \|Y\|_s$.

Proof Suppose $E|X|^r, E|Y|^s \in (0, \infty)$, otherwise the proof is immediate.

We first prove Young's inequality:

$$|ab| \le \frac{|a|^r}{r} + \frac{|b|^s}{s} \quad \text{for all } a, b, \text{ and } \left.\begin{matrix} r>0 \\ s>0 \end{matrix}\right\} \text{ s.t. } \frac{1}{r} + \frac{1}{s} = 1.$$

Let $f(x) = e^x$. Since $f''(x) \ge 0$ for all x, then f is convex. Then:

$$f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y),$$

for all $0 \le \alpha \le 1$.

We let $\alpha = \frac{1}{r}$, $1-\alpha = \frac{1}{s}$, and

$x = r \log|a|$; $y = s \log|b|$.

<u>Then</u> $e^{(\log|a||b|)} \le \frac{1}{r} e^{\log|a|^r} + \frac{1}{s} e^{\log|a|^s} \Rightarrow$

$$\Rightarrow |ab| \le \frac{|a|^r}{r} + \frac{|a|^s}{s}.$$

- Take now $a = \frac{|X|}{\|X\|_r}$, $b = \frac{|Y|}{\|Y\|_s}$ and take expectations. Thus:

$$E\,\frac{|X|\,|Y|}{\|X\|_r\,\|Y\|_s} \le \frac{1}{r}\,E\,\frac{|X|^r}{\|X\|_r^r} + \frac{1}{s}\,E\,\frac{|Y|^s}{\|Y\|_s^s}$$

$$\Downarrow$$

$$\|XY\|_1 \le \frac{1}{r}\cdot\|X\|_r\cdot\|Y\|_s\cdot\cancel{\frac{E|X|^r}{E|X|^r}} + \frac{1}{s}\,\|X\|_r\cdot\|Y\|_s\cdot\cancel{\frac{E|Y|^s}{E|Y|^s}} \underset{\substack{\uparrow \\ \frac{1}{r}+\frac{1}{s}=1}}{=} \|X\|_r\,\|Y\|_s.$$

See page 47 for equality.

qed.

Cauchy-Schwarz inequality

$$\left(E|XY|\right)^2 \leq EX^2 \cdot EY^2$$

Or, if $\langle \, , \, \rangle$ denotes the inner product in some space H, and $x, y \in H$, we have:

$|\langle x, y \rangle| \leq \|x\| \, \|y\|$, where $\|\cdot\|$ is a norm.

Liapunov's inequality.

Let $\mu(\Omega) < \infty$. Then $\|X\|_r \equiv E^{1/r}|X|^r$ is increasing in r, for all $r > 0$.

Thus: $\|X\|_1 \leq \|X\|_2 \leq \|X\|_3 \leq \ldots$

Proof: Use Hölder's inequality. Page 48.

Minkowski's inequality

$$E^{1/r}|X+Y|^r \le E^{1/r}|X|^r + E^{1/r}|Y|^r, \text{ for all } r \ge 1.$$

Proof: • For $r=1$, just note that $|X+Y| \le |X|+|Y|$.

• For $r>1$, let $s \equiv \frac{r}{r-1}$.

• Then:

$$E|X+Y|^r = E|X+Y|\,|X+Y|^{r-1} \le E(|X|+|Y|)\,|X+Y|^{r-1}$$

$$= E|X|\,|X+Y|^{r-1} + E|Y|\,|X+Y|^{r-1}$$

$$\le \|X\|_r \, \|\,|X+Y|^{r-1}\|_s + \|Y\|_r \, \|\,|X+Y|^{r-1}\|_s$$

↑ Hölder; with $\frac{1}{s} = \frac{r-1}{r}$

$$\le (\|X\|_r + \|Y\|_r)\left(E|X+Y|^{(r-1)\cdot s}\right)^{1/s} =$$

$$= (\|X\|_r + \|Y\|_r)\left(E|X+Y|^r\right)^{\frac{r-1}{r}} \quad \Big| : \left(E|X+Y|^r\right)^{\frac{r-1}{r}}$$

$$\Rightarrow \|X+Y\|_r \le \|X\|_r + \|Y\|_r$$

(So $\|\ \|_r$ is indeed a norm, since the triangle inequality is satisfied)

Basic inequality

Let $g \geq 0$ be $\uparrow$ on $[0,\infty)$ and even.
Then, for all measurable X we have..

$$\mu(|X| \geq \lambda) \leq \frac{E g(X)}{g(\lambda)}, \text{ for all } \lambda > 0.$$

Proof: $E g(X) = \int_{[|X| \geq \lambda]} g(X)\,d\mu + \int_{[|X| < \lambda]} g(X)\,d\mu \geq$

$$\geq \int_{[|X| \geq \lambda]} g(X)\,d\mu \underset{g \uparrow \text{ and even}}{\geq} g(\lambda) \int_{[|X| \geq \lambda]} d\mu$$

$$= g(\lambda)\,\mu(|X| \geq \lambda).$$

Corollaries

1. (Markov) $\mu(|X| \geq \lambda) \leq \frac{E|X|^r}{\lambda^r}$, for all $\lambda > 0$, $r \geq 1$..
2. (Chebyshev) $\mu(|X - \mu| \geq \lambda) \leq \frac{\operatorname{var} X}{\lambda^2}$, for all $\lambda > 0$

Jensen's inequality

- Suppose g is convex on (a,b).
- If $\mu(X \in (a,b)) = \mu(\Omega) = 1$ and if $a < E(X) < b$,

then:
$$g(EX) \le Eg(X).$$

Proof • Recall that if g is convex, then there exists a linear function l such that $g(x) \geq l(x)$, with equality at any pre-specified x_0 in the domain of g.

• In our case, let $x_0 \equiv E(X)$, and $g(x) \geq l(x)$.

Then:
$$Eg(X) \geq E\,l(X) = l(E(X))$$

Since l is linear and $\mu(\Omega)=1$, hence $\int$ constant = constant

$$= g(E(X)) \quad \text{qed}$$

Modes of Convergence   Lecture 12

Definition X_n converges in r-th mean to X ($X_n \longrightarrow_r X$ or $X_n \longrightarrow_{L_r} X$).

Theorem 1 (Completeness of $\mathcal{L}_r$).
Let X_n's be in any $\mathcal{L}_r$, $r>0$.

$$X_n \longrightarrow_r X \ (\text{in } \mathcal{L}_r) \iff X_n - X_m \longrightarrow_r 0$$

Proof See Homework

Notation If μ is a probability measure denoted P recall that we denote "X_n conv. in distr to X" by either $X_n \xrightarrow{d} X$ or $F_n \xrightarrow{d} F$.

Theorem 2 (Helly - Bray).
Let $(\Omega, \mathcal{A}, P)$ be a prob. space.
Suppose $F_n \to F$.

1) If g is bounded and continuous a.s. F, then

$$E\,g(X_n) = \int g\, dF_n \longrightarrow \int g\, dF = E g(X).$$

Conversely

2) If $Eg(X_n) \to Eg(X)$ for all bounded, continuous g, then $F_n \to_d F$.

<u>Theorem 3</u> (Mann-Wald).

If $X_n \to_d X$ and g is continuous a.s. F.

Then $g(X_n) \to_d g(X)$.

<u>Proof</u>: We give a general proof based on the following strong result:

<u>Skorokhod Theorem</u>

If $X_n \to_d X$, there exists

on some prob. space $(\Omega, \mathcal{A}, P)$ Y_n and Y

such that: 1) $X_n \cong Y_n$ (Recall that "$\cong$" means "distributed as").

2) $X \cong Y$

$\boxed{3}$ $Y_n \to Y$ a.s. P_X.

(Thus, if we have variables <u>conv. in distribution</u>, we can always construct a probability space on which variables distributed like the ones we were originally interested in actually <u>converge almost surely</u>.)

dist^n induced by X_n i.e $P_{X_n}(\cdot)$

has dist^n funct^n of Y_n a.e

Proof of the Helly - Bray Theorem

We have $X_n \cong Y_n$ and
$X \cong Y$.

<u>Then</u>

$$E\,g(X_n) = E\,g(Y_n) = \int g(Y_n)\,dP$$
$$E\,g(X) = E\,g(Y) = \int g(Y)\,dP .$$

Thus, in order to prove that

$$E\,g(X_n) \xrightarrow[n\to\infty]{} E\,g(X)$$

we will prove that

$$\text{①}\; E\,g(Y_n) \xrightarrow[n\to\infty]{} E\,g(Y), \quad \text{where}$$

$$Y_n \xrightarrow[n\to\infty]{} Y \quad \text{a.s. } P.$$

- By the DCT, <u>if</u> a) $g(Y_n) \xrightarrow{a.s.} g(Y)$
and
b) $|g(Y_n)| \leq W$, for some W s.t. $\int |W| < \infty$.

<u>then</u> we would conclude the desired result ①, since we can apply Corollary 1, page 41 PfS.

It remains to verify a) and b).

- Since g is bounded $\Rightarrow \exists$ a constant $M \geq 0$ s.t.
  $|g(Y_n)| \leq M, \quad \forall n \geq 0.$

Then $\int |g(Y_n)| dP \leq \int M \cdot dP = M < \infty$, since P is a probability measure.

Thus, b) is verified.

- Observe now that showing a) means that we need to find a set A such that

  ② $g(Y_n(\omega)) \longrightarrow g(Y(\omega))$, for all $\omega \in A$ <u>and</u> $P(A) = 1$.

- In what follows we construct A such that $P(A) = 1$ and ② holds.

- First recall that $Y_n \longrightarrow Y$ a.s. P. ③

Let $A_1 \equiv \{\omega \mid Y_n(\omega) \xrightarrow[n \to \infty]{} Y(\omega)\}$.

Then ③ $\Rightarrow$ $\boxed{P(A_1) = 1.}$ ④

- Let $A_2 \equiv \{\omega \mid g$ is continuous at $Y(\omega)\}$.

$$P(A_2) = P_Y(\{y \mid g \text{ is continuous at } y\}) \quad ⑤$$

- Now, recall that $X \cong Y$. This means that:

$$P_X = P\circ X^{-1} = F = P\circ Y^{-1} = P_Y. \quad ⑥$$

- Now, ⑤ and ⑥ $\Rightarrow$

- $P(A_2) = P_X(\{y \mid g \text{ is cont. at } y\})$.
- By the hypothesis of the theorem, g is continuous a.s. F, that is, it is continuous a.s. with respect to P_X.

$\Rightarrow$

$\Rightarrow \boxed{P(A_2) = 1.}$ ⑤

- Notice now that on the set $A_1 \cap A_2$ we have:

$Y_n(\omega) \xrightarrow[n\to\infty]{} Y(\omega)$ and g is continuous.

Thus $\boxed{g(Y_n(\omega)) \to g(Y(\omega)), \text{ for all } \omega \in A_1 \cap A_2 \equiv A.}$

- Finally, $0 \le P(A^c) = P(A_1^c \cup A_2^c) \le P(A_1^c) + P(A_2^c) \underset{\uparrow}{=} 0 + 0$.

By ④ and 5.

Thus, $\boxed{P(A) = 1.}$

Q.E.D.

Proof of the Mann–Wald theorem

Consider again ~~a~~ sequence $\{Y_n\}_n$ and a r.v. Y such that

$$\text{(1)}\begin{cases} X_n \cong Y_n \\ X \cong Y \end{cases} \quad \text{and}$$

$$Y_n \longrightarrow Y \text{ a.s. } P.$$

Recall that, in the course of the proof of the previous theorem, we showed that for any a.e. $\exists$ continuous function g we have.

$$g(Y_n) \xrightarrow{a.s.} g(Y).$$

Since convergence a.s. implies convergence in distribution, we also have

$$g(Y_n) \xrightarrow{d} g(Y).$$

Now, we use (1) to deduce that

$$g(X_n) \xrightarrow{d} g(X).$$

q.e.d.

Theorem 4

If $X_n \xrightarrow{r} X$ then:

a) $E|X_n|^{r'} \longrightarrow E|X|^{r'}$

b) $EX_n^k \longrightarrow EX^k$

for all $0 \le k \le r' \le r$.

Proof See Th. 5.3 page .

Remark Theorem 4 implies that if

$E|X_n - X|^r \to 0$ then we also have:

$$E|X_n|^r \to E|X|^r$$

and

$$EX_n^r \to EX^r$$

Uniform integrability

<u>Definition 1</u> A collection of r.v.'s $\{X_t\}_{t\in T}$ is called <u>integrable</u> if

$$\sup_{t\in T} E|X_t| < \infty .$$

<u>Definition 2</u> A collection of r.v.'s $\{X_t\}_{t\in T}$ is called <u>uniformly integrable</u> if

$$\sup_{t\in T} E\left[|X_t| \times 1[|X_t| \geq \lambda] \right] \longrightarrow 0 ,$$

as $\lambda \to \infty$.

Remark

Suppose that $\{X_t\}_{t\in T}$ is such that

$|X_t| \leq Y$, $Y \in \mathcal{L}_1$, for all $t \in T$.

Recall that $Y \in \mathcal{L}_1$ means $\int |Y| < \infty \iff E|Y| < \infty$.

Ⓐ Then $\mu(|X_t| \geq \lambda) \leq \mu(|Y| \geq \lambda) \leq$

$$\leq \frac{E|Y|}{\lambda} \longrightarrow 0 \quad \text{as } \lambda \to \infty,$$

since $E|Y| < \infty$

Now: $E|X_t| \cdot 1_{[|X_t| \geq \lambda]} = \int |X_t| \cdot 1_{[|X_t| \geq \lambda]} d\mu =$

$$= \int_{[|X_t| \geq \lambda]} |X_t| d\mu \leq \int_{[|X_t| \geq \lambda]} Y d\mu \leq \int_{[|Y| \geq \lambda]} |Y| d\mu \longrightarrow 0,$$

by the absolute continuity of the integral (Th. 3.2.5) and Ⓐ

Thus, we have shown:

Lemma If $|X_t| \leq Y$, for all $t \in T$, for some $Y \in \mathcal{L}_1$, then X_t's are u.i.

Theorem 5 (Uniform absolute continuity of the integr
sume $\mu(\Omega) < \infty$.
$\{X_t\}_{t \in T}$ is uniformly integrable $\iff$

1) $\{X_t\}_{t \in T}$ is integrable ($\sup_t E|X_t| < \infty$).

2) $\mu(A) < \delta_\varepsilon \Rightarrow \sup_t \int_A |X_t| d\mu < \varepsilon$. (uniform abs. continui

Proof „$\Longrightarrow$" For every t we have

$$E|X_t| = \int_{[|X_t| < \lambda]} |X_t| d\mu + \int_{[|X_t| \geq \lambda]} |X_t| d\mu \leq$$

$$\leq \lambda \mu(\Omega) + E |X_t| \mathbb{1}_{[|X_t| \geq \lambda]} < \infty$$

For λ large enough

Thus $\sup_t E|X_t| < \infty$

Assume A is such that $\mu(A) < \delta$.

Then; as before:

$$\int_A |X_t|\,d\mu = \int_A |X_t| 1_{[|X_t|<\lambda]}\,d\mu +$$

$$\int_A |X_t| 1_{[|X_t|\geq\lambda]}\,d\mu$$

$$\leq \lambda\mu(A) + \int |X_t| 1_{[|X_t|\geq\lambda]}\,d\mu$$

$$\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$ since

$\{X_t\}_t$ is u.i., and for large λ with $\delta \leq \frac{\varepsilon}{2\lambda}$.

"$\Leftarrow$" Notice that by Markov's ineq.:

$$\mu(|X_t|\geq\lambda) \leq E|X_t|/\lambda \leq \sup_t E|X_t|/\lambda < \delta_\varepsilon,$$

Since $\sup_t E|X_t| < \infty$ and λ large.

We can use now (b) for $A = [|X_t|\geq\lambda]$:

$$\sup_t \int |X_t| \times 1_{[|X_t|\geq\lambda]}\,d\mu = \sup_t E|X_t| 1_{[|X_t|\geq\lambda]} \to 0$$

qed

Lecture 13

Theorem (Vitali)

Let $r > 0$. Suppose that X_n's $\in \mathcal{L}_r$ and that $X_n \xrightarrow{P} X$. Then, the following statements are equivalent:

1) $\{|X_n|^r : n \geq 1\}$ are u.i. rvs.

2) $X_n \to_r X$. ie $E|X_n - X|^r \to 0$ as $n \to \infty$

3) $E|X_n|^r \to E|X|^r$.

4) $\overline{\lim} \, E|X_n|^r \leq E|X|^r < \infty$

Corollary ($\mathcal{L}_1$ convergence).

$$X_n \to_1 X \iff \begin{cases} a) \; X_n \xrightarrow{ui} X \\ b) \; \{|X_n| : n \geq 1\} \text{ are uniformly integrable.} \end{cases}$$

- Proof of Corollary.

- Notice that a) + b) $\Rightarrow$ $X_n \xrightarrow{L_1} X$ by Vitali's theorem, with $r = 1$.

- For "$\Leftarrow$" ~~No~~tice that
$\mu(|X_n - X| > \varepsilon) \leq \frac{E|X_n - X|}{\varepsilon} \xrightarrow[n \to \infty]{} 0$, if $X_n \xrightarrow{L_1} X$.
So, $X_n \xrightarrow{\mu} X$ and $X_n \xrightarrow{L_1} X$. Then, by Vitali's theorem again, $2 \Rightarrow 1$, we have the desired result.

- Proof of Vitali's theorem: Let $X_n \xrightarrow{P} X$ ⊛.
We show that 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 1).
- Show that 1) $\Rightarrow$ 2: By ⊛ and Th. 2.3.1 $\Rightarrow$
$\exists n'$ s.t. $X_{n'} \xrightarrow{a.s.} X \Rightarrow |X_{n'}|^r \xrightarrow{a.s.} |X|^r$.
- Thus $E|X|^r \underset{(101)}{=} E \lim |X_{n'}|^r \underset{\text{Fatou's Lemma}}{\leq} \underline{\lim}\, E|X_{n'}|^r \underset{|X_{n'}|^r \text{ are u.i., hence integrab}}{<} \infty \Rightarrow X \in L_r$

- Since $\{|X_n|^r\}_n$ are u.i. and $X \in L_r$ $\underset{(102)}{\Longrightarrow}$ $\{|X_n - X|^r\}_n$ are u.i.

Exercise: Justify (101) and (102).

$|X_{n'}|^r \to |X|^r \Rightarrow \underline{\lim}\, |X_{n'}|^r = \lim |X_{n'}|^r = |X|^r$
$\therefore E(\underline{\lim}\, |X_{n'}|^r) = E(|X|^r)$

Now

$$E|X_n - X|^r = E|X_n-X|^r \cdot 1_{[|X_n-X|>\varepsilon]} + E|X_n-X|^r \cdot 1_{[|X_n-X|\le \varepsilon]}. \qquad (100)$$

By (**) and Theorem 3.5.4, we have that

$E|X_n-X|^r \cdot 1_{[|X_n-X|>\varepsilon]} < \varepsilon$ since

$P(|X_n-X|\geq \varepsilon) \to 0$ by (*).

$\therefore$ $E|X_n-X|^r \leq \varepsilon + \varepsilon^r$ and so $X_n \xrightarrow{r} X$.

Thus, we showed that 1) $\Rightarrow$ 2).

Also, notice that 2) $\Rightarrow$ 3) by Th. 3.5.3, and 3) $\Rightarrow$ 4).

It remains to show that 4) $\Rightarrow$ 1).

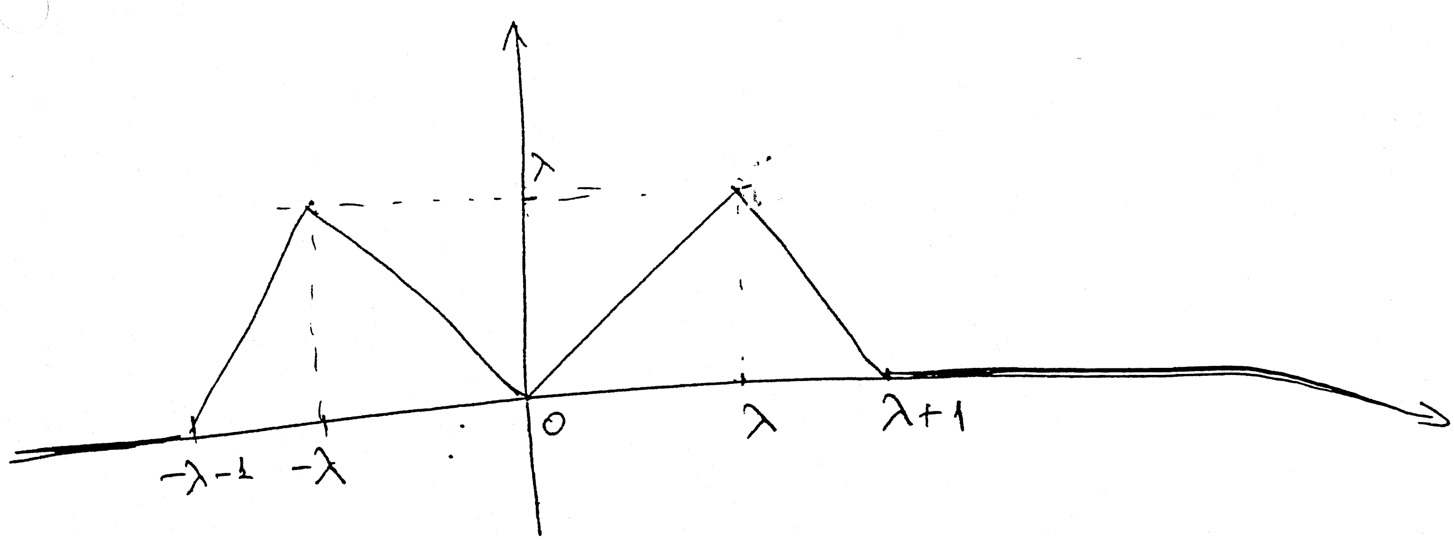
Remark: Notice that in (100) we used $1_{[|X_n-X|>\varepsilon]}$, NOT $1_{[|X_n-X|^r>\varepsilon]}$. If $\{|X_n-X|^r\}_n$ are u.i. we know that $\sup_n E|X_n-X|^r 1_{[|X_n-X|^r>\varepsilon]} \xrightarrow[\varepsilon\to\infty]{} 0$. HOWEVER: this is not useful here, as we are interested in small values of ε. That's why we need Th. 3.5.4 ($\to$ We can then use $1_{[|X_n-X|^r>\varepsilon]}$ if desired).

Let $f_\lambda : [0,\infty) \to \mathbb{R}$, continuous, ~~such~~ that

$$f_\lambda(x) = \begin{cases} |x|^r, & \text{if } |x|^r \le \lambda \\ 0, & \text{if } |x|^r \ge \lambda + 1 \\ \text{linear}, & \text{if } \lambda \le |x|^r \le \lambda+1. \end{cases}$$

We graph it for $r=1$.

$$f_\lambda(x) = \begin{cases} |x|, & \text{if } x \in [-\lambda, \lambda]. \\ 0, & \text{if } x \in (-\infty, -\lambda-1) \cup (\lambda+1, \\ \text{linear}, & \text{if } x \in (-\lambda-1, -\lambda) \cup (\lambda, \lambda+ \end{cases}$$



Then

$$\overline{\lim} \int_{[|X_n|^r > \lambda+1]} |X_n|^r \, d\mu = \overline{\lim} \left\{ \int |X_n|^r \, d\mu - \int_{[|X_n|^r \le \lambda+1]} |X_n|^r \, d\mu \right\}.$$

$$\underset{4)}{\le} E|X|^r - \underline{\lim} \int_{[|X_n|^r \le \lambda+1]} |X_n|^r \, d\mu$$

$$\le E|X|^r - \underline{\lim}\, E f_\lambda (X_n).$$

By the definition of f_λ

$$= E|X|^r - E f_\lambda (X)$$

By Helly-Bray Theorem, since f_λ is bounded and cont

$$\le \int_{[|X| \ge \lambda]} |X|^r \, d\mu \longrightarrow 0, \text{ as } \lambda \to \infty, \text{ since } X \in L_r.$$

Note that this follows by th. 2.5 page 42 combined with $P(|X| > \lambda) \le \frac{E|X|}{\lambda} \xrightarrow[\lambda \to \infty]{} 0$, since $E|X| < \infty$ ($X \in L_r$, $r \ge 1 \Rightarrow X \in L_1$).

q.e.d.

Theorem of characterization of uniform integrab
(de la Vallée Poussin).

Let $\mu(\Omega) < \infty$.

A family $\{X_t\}_t$ of $\mathcal{L}_1$ integrable functions is uniformly integrable $\iff \exists$ a convex function G on $[0,\infty)$ s.t.

$$\begin{cases} 1)\ G(0) = 0 \\ 2)\ \dfrac{G(x)}{x} \xrightarrow[x \to \infty]{} \infty \\ 3)\ \sup_t E\, G(|X_t|) < \end{cases}$$

Proof "$\Leftarrow$" Assume we have a G satisfying 1).

Then, re-write 2) as: For all $x \geq \lambda$ we have $\frac{G(x)}{x}$

Thus: $\displaystyle\int_{[|X_t| \geq \lambda]} |X_t|\, d\mu \leq \frac{1}{c} \int_{[|X_t| \geq \lambda]} G(|X_t|)\, d\mu \leq \frac{1}{c} \int G(|X_t$

$$\leq \frac{1}{c} \sup_t E\, G(|X_t|) < \varepsilon .$$

By 3) and for c lar

"$\Longrightarrow$" We show that u.i. implies [redacted]

- Let $\{b_n\}_n$ be a sequence of [redacted]
that $b_n \uparrow \infty$ and $b_0 = 0$. [redacted]

- Define $g(x) = \begin{cases} b_n, & \text{if } n \le x < n \ldots \\ \text{~~otherwise~~ [redacted]} \end{cases}$

- Define $G(x) = \displaystyle\int_0^x g(y)\,dy$.

- Define $a_n(t) = \mu(|X_t| \geq$ [redacted] $n)$.

Notice that this G is convex [redacted] and sat…
It remains to show that $\sup_t E$ [redacted] $(|X_t|)$

- Note that

$$EG(|X_t|) = \int \left(\int_0^{|X_t|} g(y)\,dy \right) d \text{[redacted]} =$$

$$= \int_{[0 \le |X_t| < 1]} \left(\int_0^{|X_t|} g(y)\,dy \right) d\mu + \ldots + \int_{[n \le |\text{[redacted]}_t| < n+1]} \left(\int_0^{|X_t|} g(y) \ldots \right.$$

Recall $g(x) = \begin{cases} b_0, & \text{if } 0 \le x < 1 \\ b_1, & \text{if } 1 \le x < 2 \\ b_1, & \text{if } 2 \le \cancel{x} < 3 \\ \vdots & \\ b_n, & \text{if } n \le \cancel{x} < n+1 \\ \vdots & \end{cases}$

[So $G(\frac{3}{2}) = \int_0^{\frac{3}{2}} g(y)\,dy = \int_0^1 g(y)\,dy + \int_1^{\frac{3}{2}} g(y)\,dy \le$

$\le b_0 + b_1$.]

Thus :

$$E\,G(|X_t|) \le b_1\,\mu(1 \le |X_t| < 2) + (b_1 + b_2)\,\mu(2 \le |X_t| <$$

$$= \sum_{n=1}^{\infty} b_n a_n(t).$$

To show that 3) holds it is enough to find $b_n \nearrow \infty$ such that

$$\sup_t \sum_{n=1}^{\infty} b_n a_n(t) < \infty.$$

Recall that $\{X_t\}_t$ are u. i. $\Longrightarrow$

$$\sup_t \int_{[|X_t| \geq c_n]} |X_t| \, d\mu \leq \frac{1}{2^n}, \quad \text{for } c_n \nearrow$$

<u>Then</u>:

since $i \leq |X_t| < i+1$

$$\frac{1}{2^n} \geq \int_{[|X_t| \geq c_n]} |X_t| \, d\mu \geq \sum_{i=c_n}^{\infty} i \, \mu(i \leq |X_t| < i+1)$$

reiting $i = \sum_{j=1}^{i} 1$

$$= \sum_{i=c_n}^{\infty} \sum_{j=1}^{i} \mu(i \leq |X_t| < i+1)$$

$$\geq \sum_{j=c_n}^{\infty} \sum_{i=j}^{\infty} \mu(i \leq |X_t| < i+1).$$

The sum of the first terms of a conv series $\geq$ the rest of the series (for a conv. series)

$$= \sum_{j=c_n}^{\infty} \mu(|X_t| \geq j).$$

$$= \sum_{j=c_n}^{\infty} a_j(t).$$

So, we showed that:

$$\sum_{j=c_n}^{\infty} a_j(t) \le \frac{1}{2^n} .$$ Thus:

$$1 = \sum_{n=1}^{\infty} \frac{1}{2^n} \geq \sup_t \sum_{n=1}^{\infty} \sum_{j=c_n}^{\infty} a_j(t) =$$

$$= \sup_t \sum_{j=1}^{\infty} b_j a_j(t) ,$$ where

we recall that

$a_j(t) \equiv \mu(|X_t| \geq j)$ and we define

$b_j \equiv$ "the number of integers n such that c_n is $\le j$". (Check!)

Hence, for this choice of b_j we have:

$$\sup_t E\, G(|X_t|) \le \sup_t \sum_{n=1}^{\infty} b_n a_n(t) \le 1 < \infty \;!$$

q.e.d.

Remark: In nonparametric statistics it is customary to look at the L_1 or L_2 "distances" between the parameter of interest (typically a function f) and its estimate. Since good estimation means "small" distance between the target and its estimator, developing good estimates requires (most of the time) the study of the L_2 or L_1 convergence of some sequences.

Important reading exercise:
Theorem 5.7 page 57 and the diagram at page 58. **

Theorem (See Theorem 5.5.8, page 59).

Let $S_c = \left\{ \sum_{j=1}^{m} y_j 1_{I_j} \mid I_j \text{ disjoint finite intervals} \right\}$.

Then, $\forall\, \varepsilon > 0 \;\exists\, Y_\varepsilon \in S_c$ such

that $\int |X - Y_\varepsilon| \, d\mu_F < \varepsilon$, for any $X \in \mathcal{L}_1$.

That is, S_c is ε dense in $\mathcal{L}_1$,
where we recall that

$\mathcal{L}_1 = \left\{ \int |X| \, d\mu < \infty \right\}$, where

$X : (\Omega, \mathcal{A}, \mu) \longrightarrow (\mathbb{R}, \mathcal{B})$ in general.

In this theorem we consider r.v.'s

$X : (\mathbb{R}, \mathcal{B}, \mu_F) \longrightarrow (\mathbb{R}, \mathcal{B})$,

where μ_F is a Lebesgue-Stieltjes measure associated with a generalized d.f. F.

Proof. Enough to show for $X \geq 0$. In what follows we assume that this holds.

- Recall that a simple function is of the form

$$\cancel{Z}\, Z = \sum_{i=1}^{n} z_i 1_{A_i}, \text{ where } \sum_{i=1}^{n} A_i = B,$$

and $z_i \in \mathbb{R}$.

By proposition 2.2.3 page 25, for any $\varepsilon > 0$, $\exists\, Z_\varepsilon$, a simple function, such that

$$\text{(1)} \quad |X - Z_\varepsilon| < \varepsilon.$$

$\boxed{\text{Recall that } Z_\varepsilon \text{ depends on } X\,!!}$

(2) • Notice first that $\mu(A_i) < \infty$, for all $i = 1, \dots, n$.

To see that, assume that $\exists\, i_0 \in \{1, \dots, n\}$ such that $\mu(A_{i_0}) = \infty$. Then

$$\int |X| \, d\mu_F \leq \int |X - Z_\varepsilon| \, d\mu_F + \int |Z_\varepsilon| \, d\mu_F$$

$$\leq \int |X - Z_\varepsilon| \, d\mu_F + \sum_{i=1}^{n} \int |z_i| \, 1_{A_i} \, d\mu_F.$$

$= \infty$, contradiction with $\int |X| \, d\mu_F < \infty$.

- Choose now Z_ε satisfying ① (and implicitely ②)
  such that $\int |X - Z_\varepsilon| \, d\mu_F < \frac{\varepsilon}{2}$. ③
  (Notice that $\int |X - Z_\varepsilon| d\mu_F < \int |X| d\mu_F + \int |Z_\varepsilon| d\mu_F < \infty$
  So ③ is always possible).
- Recall now that $\mathcal{B} = \sigma[\mathcal{C}_F]$,
  where $\mathcal{C}_F \supset \{ \sum_n I_n$, with each I_n of the type $(a,b] \}$
- Recall also that for two arbitrary sets A and C
  $A \Delta C \equiv A^c C \cup A C^c$
- We apply now the Approximation Lemma of Ex. 1.2. page 16 to each of the sets $A_1, \dots, A_n \in \mathcal{B}$ (~~since~~
  We can apply it since we made sure that $\mu(A_i) < \infty$ for all i). Then, $\exists\ B_1, \dots, B_n \in \mathcal{C}_F$
  such that for all $\varepsilon > 0$,
  $\mu(A_i \Delta B_i) < \frac{\varepsilon}{2n|z_i|}$.

Define now $\cancel{Y}_\varepsilon \equiv \sum_{i=1}^{n} z_i 1_{B_i}$.

Then $\int |Z_\varepsilon - Y_\varepsilon| d\mu_F = \int \left| \sum_{i=1}^{n} z_i (1_{B_i} - 1_{A_i}) \right| d\mu_F$

$\leq \sum_{i=1}^{n} |z_i| \int |1_{B_i} - 1_{A_i}| d\mu_F$;

$$|1_{B_i} - 1_{A_i}| = \begin{cases} 1, & \text{if } \omega \in B_i \setminus A_i = B_i A_i^c \quad \underline{\underline{\text{or}}} \\ 1, & \text{if } \omega \in A_i \setminus B_i = A_i B_i^c \\ 0 & \text{otherwise}. \end{cases}$$

<u>Thus</u> $\int |Z_\varepsilon - Y_\varepsilon| d\mu_F \leq \sum_{i=1}^{n} |z_i| \mu(A_i \Delta B_i)$

$\leq \sum_{i=1}^{n} \cancel{|z_i|} \cdot \frac{\varepsilon}{2n\cancel{|z_i|}} = \frac{\varepsilon}{2}$.

<u>Hence</u> $\int |X - Y_\varepsilon| d\mu_F \leq \int |X - Z_\varepsilon| d\mu_F + \int |Y_\varepsilon - Z_\varepsilon| d\mu_F$

$\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

To complete the proof, re-write

$Y_\varepsilon = \sum_{j=1}^{m} y_j 1_{I_j}$, with I_j ~~that~~ <u>disjoint</u> finite intervals.

q.ed.