PSTAT 213B Section Notes, winter 2020, Ming Min

Section 1. (7.1.1, 7.1.4, 7.2.4, 7.2.1)

7.1.1 Let $r \geq 1$, define $\|X\|_r = E[|X|^r]^{\frac{1}{r}}$. Show that:

(a) $\|c \cdot X\|_r = |c| \cdot \|X\|_r$ for $c \in \mathbb{R}$.

(b) $\|X+Y\|_r \leq \|X\|_r + \|Y\|_r$.

(c) $\|X\|_r = 0$ iff $P(X=0)=1$.

$\frac{1}{r}+\frac{1}{p}=1$

$\frac{1}{p}=1-\frac{1}{r}=\frac{r-1}{r}$

Proof: (a) $\|c \cdot X\|_r = E[|c \cdot X|^r]^{\frac{1}{r}} = |c| \cdot \|X\|_r$

(b) $\|X+Y\|_r = E[|X+Y|^r]^{\frac{1}{r}}$

$E[|X+Y| \cdot |X+Y|^{r-1}] \leq E[|X| \cdot |X+Y|^{r-1}] + E[|Y| \cdot |X+Y|^{r-1}]$ triangle

$\leq E[|X|^r]^{\frac{1}{r}} \cdot E[|X+Y|^r]^{\frac{r-1}{r}} + E[|Y|^r]^{\frac{1}{r}} E[|X+Y|^r]^{\frac{1}{r}}$ hölder.

$\Rightarrow$ desired result.

(c) "$\Rightarrow$"

if $\|X\|_r = E[|X|^r]^{\frac{1}{r}} = 0$, then $E[|X|^r]=0$

$\forall \varepsilon \in \mathbb{Q}, \varepsilon > 0$. $P(|X|>\varepsilon) = P(|X|^r > \varepsilon^r) < \frac{1}{\varepsilon^r} E[|X|^r] = 0$.

$P(|X|>0) = \lim_{\varepsilon \to 0} P(|X|>\varepsilon) = 0$.

"$\Leftarrow$" if $P(X=0)=1$,

then $|X|^r = 0$ a.s. $\Rightarrow E[|X|^r]^{\frac{1}{r}} = 0$

7.1.4 (Lévy metric) For two distribution functions F and G. let

$$d(F,G) = \inf\{\delta > 0 : F(x-\delta)-\delta \leq G(x) \leq F(x+\delta)+\delta \text{ for all } x \in \mathbb{R}\}$$

show d is a metric on the space of distribution function.

Proof: (i) $d(F,G) \geq 0$, — by definition of d.

(ii) $d(F,G)=0 \Leftrightarrow F=G$

(iii) $d(F,G) = d(G,F)$

(iv) $d(F,G) + d(G,H) \geq d(F,H)$.

(ii) $d(F,G)=0 \Rightarrow \forall \delta > 0$, $\begin{cases} G(x) \leq F(x+\delta)+\delta \\ F(x) \leq G(x+\delta)+\delta \end{cases}$ (shift)

let $\delta \downarrow 0$, we have $G(x) \leq F(x)$, $F(x) \leq G(x) \Rightarrow F=G$

if $F=G$, obvious.

(iii) denote $d(F,G) = \inf\left\{\delta : \begin{cases} G(x) \leq F(x+\delta)+\delta \\ F(x) \leq G(x+\delta)+\delta \end{cases}\right\}$

$d(F,G)$ admits the same form.

(iv) let $\delta_n \downarrow d(F,G)$, $\varepsilon_n \downarrow d(G,H)$. s.t.

$\forall x$, $G(x) \le F(x+\delta_n)+\delta_n$, and $F(x) \le G(x+\delta_n)+\delta_n$

$\forall x$, $G(x) \le H(x+\varepsilon_n)+\varepsilon_n$ and $H(x) \le G(x+\varepsilon_n)+\varepsilon_n$.

$\Rightarrow$ $F(x) \le G(x+\delta_n)+\delta_n \le H(x+\delta_n+\varepsilon_n)+(\delta_n+\varepsilon_n)$

and $H(x) \le F(x+\delta_n+\varepsilon_n)+\delta_n+\varepsilon_n$

$\Rightarrow$ $d(F,H) \le \lim_{n\to\infty}(\varepsilon_n+\delta_n) = d(F,G)+d(G,H)$.

7.2.4. $d(F_n,F)\to 0$ (denote $d_n = d(F_n,F)$).

$\Leftrightarrow$ $F_n(x)\to F$ at all x s.t. $F(x)$ is continuous.

Proof. "$\Rightarrow$" $d(F_n,F)\to 0$. denote $C^F = \{$ continuous points of $F\}$.

so $\forall \varepsilon>0$, $\exists N$ s.t. $\forall n\ge N$. we have

$F(x-\varepsilon)-\varepsilon \le F_n(x) \le F(x+\varepsilon)+\varepsilon$ for all x.

let $n\to\infty$ and then $\varepsilon\to 0$

$\lim_{n\to\infty} F_n(x) \le F(x)$, $F(x^-) \le \lim F_n(x)$

, if F is continuous at x, then $F(x)=\lim_{n\to\infty} F_n(x)$.

"$\Leftarrow$" if $F_n(x)\to F(x)$ in C^F. let $\varepsilon>0$.

Let $\varepsilon>0$, find real numbers $a=x_1<x_2<\cdots<x_n=b$, each being points of continuity of F.

with (i) $F_m(a)<\varepsilon$ for all m, $F(b)>1-\varepsilon$.

(ii) $|x_{i+1}-x_i|<\varepsilon$ for $1\le i<n$.

This a can be found.
① find a' with $F(a')<\frac{1}{2}\varepsilon$. and a' continuity. so $\exists M$, s.t. $|F_n(a')-F(a')| \le \frac{\varepsilon}{2}$, $\Rightarrow F_n(a')<\varepsilon$ on $n\ge M$.
find a s.t. $F_n(a)<\varepsilon$ with $a<a'$, continu.

$\exists N$ s.t. $\forall n\ge N$, s.t. $|F_n(x_i)-F(x_i)|<\varepsilon$

for all i, (since finitely many x_i's).

Then for $x\in[x_i, x_{i+1})$,

$F_n(x) \le F_n(x_{i+1}) \le F(x_{i+1})+\varepsilon \le F(x+\varepsilon)+\varepsilon$

$F_n(x) \ge F_n(x_i) \ge F(x_i)-\varepsilon \ge F(x-\varepsilon)-\varepsilon$.

for $x\in(-\infty,a)$, $F_n(x)<\varepsilon \le F(x+\varepsilon)+\varepsilon$

and, $F(x)\le F(a)<\varepsilon$, $\Rightarrow$ $F_n(x)\ge 0 \ge F(x-\varepsilon)-\varepsilon$. large enough n.

for $x\in(b,\infty)$, $F_n(x)\le 1\le F(x+\varepsilon)+\varepsilon$.

$F_n(x)\ge 1-\varepsilon \ge F(x-\varepsilon)-\varepsilon$. large enough n.

7.2.1. (a) $X_n \xrightarrow{r} X$ when $r \geq 1$, show $E[|X_n|^r] \to E[|X|^r]$

(b) $X_n \xrightarrow{1} X$. Show $E[X_n] - E[X]$. converse is not true

(c) $X_n \xrightarrow{2} X$ Show $Var(X_n) \to Var(X)$

Proof: (a) $X_n \xrightarrow{r} X \Rightarrow E[|X_n - X|^r] \to 0$.

$$E[|X_n|^r]^{\frac{1}{r}} \le E[|X_n - X|^r]^{\frac{1}{r}} + E[|X|^r]^{\frac{1}{r}}$$

$\lim\limits_{n\to\infty} E[|X_n|^r]^{\frac{1}{r}} \le E[|X|^r]^{\frac{1}{r}}$, and vice versa.

(b) $E[|X_n - X|] \to 0$.

$$0 \le |E[X_n] - E[X]| \le E[|X_n - X|] \to 0.$$

(c) $E[(X_n - X)^2] \to 0. \Rightarrow E[X_n^2] \to E[X^2]$.

$$E[|X_n - X|] \le E[|X_n - X|^2]^{\frac{1}{2}} \to 0 \Rightarrow E[X_n] \to E[X].$$

so $var(X_n) = E[X_n^2] - E[X]^2 \to Var(X)$.

Section 2 (7.2.5(a), 7.2.7, 7.2.9, 7.3.1(a), 7.3.2)

7.2.5 (a) Suppose $X_n \xrightarrow{D} X$, $Y_n \xrightarrow{P} c$. Show $X_n Y_n \xrightarrow{D} cX$, $\frac{X_n}{Y_n} \xrightarrow{D} \frac{X}{c}$ if $c \neq 0$.

Proof: Let $r \in C^F$. $P(X_n \le r) \to P(X \le r)$

$$P(X_n Y_n \le r) = P(X_n Y_n \le r, |Y_n - c| \le \varepsilon) + P(X_n Y_n \le r, |Y_n - c| > \varepsilon)$$

$$\le P(X_n Y_n \le r, |Y_n - c| \le \varepsilon) + P(|Y_n - c| > \varepsilon)$$

On $\{|Y_n - c| \le \varepsilon\}$, $c - \varepsilon \le Y_n \le c + \varepsilon$, let $\varepsilon = \frac{1}{2}c$

let $0 < \varepsilon < c$

if $c > 0$, then $P(X_n Y_n \le r, |Y_n - c| \le \varepsilon)$

$r > 0$ $\quad \le P(X_n \le \frac{r}{c - \varepsilon})$

$\Rightarrow P(X_n Y_n \le r) \le P(X_n \le \frac{r}{c-\varepsilon}) + P(|Y_n - c| > \varepsilon)$

$P(X_n Y_n \le r) \ge P(X_n Y_n \le r, |Y_n - c| \le \varepsilon) \ge P(X_n \le \frac{r}{c+\varepsilon})$

$\Rightarrow$ let $n \to \infty$, $\varepsilon \to 0$,

$$\lim_{\varepsilon \to 0} \lim_{n\to\infty} P(X_n \le \tfrac{r}{c+\varepsilon}) \le \lim_{n\to\infty} P(X_n Y_n \le r) \le \lim_{\varepsilon\to 0} \lim_{n\to\infty} P(X_n \le \tfrac{r}{c-\varepsilon})$$

as long as $\frac{r}{c}$ is in C^X. (choose ε_r s.t. $[r, r+\varepsilon_r) \subseteq C^X$)

$$P(|X_nY_n|>r) = P(|X_nY_n|>r,\ |Y_n|>\varepsilon) + P(|X_nY_n|>r,\ |Y_n|\le\varepsilon)$$
$$\le P(|Y_n|>\varepsilon) + P(|X_n| > \tfrac{r}{\varepsilon}) \longrightarrow 0$$

7.2.7 Let $\{X_n\}$ be a sequence of RV, $\{C_n\}$ are Reals converging to c.

Show $X_n \to X$ entails $C_nX_n \to cX$ in (a.s., L^r, P, D)

Proof: If $X_n \to X$ a.s. $\Omega_0 = \{w\in\Omega : \lim_{n\to\infty} X_n(w) = X(w)\}$, $P(\Omega_0)=1$.

So for $\forall\ w\in\Omega_0$,

$$\lim_{n\to\infty} C_nX_n(w) = c\cdot X(w) \Rightarrow C_nX_n \to cX \text{ a.s.}$$

If $X_n \xrightarrow{L^r} X$, then $E[|X_n-X|^r]^{1/r} \to 0$.

$$E[|C_nX_n - cX|^r]^{1/r} \le E[|C_n|^r\cdot|X_n-X|^r]^{1/r} + E[|C_n-c|^r\cdot|X|^r]^{1/r}$$
$$= |C_n|\cdot E[|X_n-X|^r]^{1/r} + |C_n-c|\cdot E[|X|^r]^{1/r} \to 0$$

If $X_n \xrightarrow{P} X$, then $P(|X_n-X|>\varepsilon) \to 0$

$$P(|C_nX_n - cX|>\varepsilon) \le P(|C_n|\cdot|X_n-X|>\tfrac{\varepsilon}{2}) + P(|C_n-c||X|>\tfrac{\varepsilon}{2})$$
$$\le P(|X_n-X| > \tfrac{\varepsilon}{2(c+\varepsilon)}) + P(|X| > \tfrac{\varepsilon}{2|C_n-c|})$$
$$\to 0$$

7.2.9. do it by yourself.

7.3.1 (a) Suppose $X_n \xrightarrow{P} X$. Show $\{X_n\}$ is Cauchy convergent in P, i.e. $\forall\varepsilon>0$, $P(|X_m-X_n|>\varepsilon) \to 0$ as $m,n\to\infty$. In what sense is the converse true?

Proof: If $X_n \xrightarrow{P} X$. then

$$P(|X_m-X_n|>\varepsilon) \le P(|X_m-X|>\tfrac{1}{2}\varepsilon) + P(|X_n-X|>\tfrac{1}{2}\varepsilon)$$
$$\to 0$$

If Cauchy convergent in P.

$\forall k$, $\exists n_k$ s.t. $P(|X_m-X_n|>\frac{1}{k^2}) \le \frac{1}{k^2}$ for $\forall m,n\ge n_k$

Define $N_1 = n_1$, $N_k = \max(N_{k-1}+1, N_k)$

So we have $\sum_{k=1}^{\infty} P(|X_{N_{k+1}} - X_{N_k}| > \frac{1}{k^2}) < \infty \Rightarrow$

$A_k = \{|X_{N_{k+1}} - X_{N_k}| > \frac{1}{k^2}\}$ happens finitely often

So $X_{N_1} + \sum_{k=1}^{\infty}(X_{N_{k+1}} - X_{N_k})$ converges, denote it by X.

$$X_{N_m} = X_{N_1} + \sum_{k=1}^{m-1}(X_{N_{k+1}} - X_{N_k}) \Rightarrow X - X_{N_m} = \sum_{k=m}^{\infty}(X_{N_{k+1}} - X_{N_k})$$

$\rightarrow$ $X_{N_m} \to X$ a.s. as $m \to \infty$.

$$P(|X_n - X| > \varepsilon) \le P(|X_n - X_{N_m}| > \frac{\varepsilon}{2}) + P(|X_{N_m} - X| > \varepsilon)$$

for $\forall m$, let $m \to \infty$, $n \to \infty$.

7.3.2 Show that the prob. that infinitely many of events $\{A_n : n \geq 1\}$ occur satisfies $P(A_n \text{ i.o.}) \geq \limsup_{n\to\infty} P(A_n)$

Proof: $P(A_n \text{ i.o.}) = P(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m) = \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty} A_m)$

$P(\bigcup_{m=n}^{\infty} A_m) \geq P(A_m)$ for $\forall m \geq n$

$\Rightarrow$ $P(\bigcup_{m=n}^{\infty} A_m) \geq \sup_{m \geq n} P(A_m)$

$\lim_{n\to\infty} P(\cup A_m) \geq \limsup P(A_m)$

Section 3. (7.2.10, 7.3.9, 7.7.1, 7.7.3, 7.3.13)

7.2.10. $\{X_r, r \geq 1\}$ independent Poisson variables with parameters $\{\lambda_r : r \geq 1\}$

Show that $\sum X_r$ converges or diverges a.s. according as $\sum_{r=1}^{\infty} \lambda_r$ converges or diverges.

Proof: Denote $S_n = \sum_{r=1}^{n} X_r$. let $m > n$, so $S_m - S_n \sim \text{Poisson}(\sum_{r=n}^{m} \lambda_r)$

$$P(\sum_{r=n+1}^{\infty} X_r > 0) = 1 - P(\sum_{r=n+1}^{\infty} X_r = 0) = 1 - \prod_{r=n+1}^{\infty} P(X_r = 0) = 1 - e^{-\sum_{r=n+1}^{\infty} \lambda_r}$$

$$\Rightarrow P(\lim_{n\to\infty} |S_n - \sum_{r=1}^{\infty} X_r| = 0) = P(\lim_{n\to\infty} \sum_{r=n+1}^{\infty} X_r = 0) = \lim_{n\to\infty} e^{-\sum_{r=n+1}^{\infty} \lambda_r}$$

$$\left(= P(\bigcup_{n=1}^{\infty} \bigcap_{r=n+1}^{\infty} \{X_r = 0\}) = \lim_{n\to\infty} P(\bigcap_{r=n+1}^{\infty} \{X_r = 0\}) \right)$$

7.3.9 $\{X_n ; n \geq 1\}$ iid $\exp(1)$. Show $P(\limsup_{n\to\infty} \frac{X_n}{\log n} = 1) = 1$

$$P(\frac{X_n}{\log n} \geq 1+\varepsilon) = \int_{(1+\varepsilon)\log n}^{\infty} e^{-x}\,dx = n^{-(1+\varepsilon)} \Rightarrow \sum_{n=1}^{\infty} P(\frac{X_n}{\log n} \geq 1+\varepsilon) < \infty$$

By Borel-Cantelli lemma 1, $\{\frac{X_n}{\log n} \geq 1+\varepsilon\}$ finitely often $\Rightarrow$ $\limsup \frac{X_n}{\log n} \leq 1$. a.s.

$P(\frac{X_n}{\log n} \geq 1) = n^{-1}$ $\Rightarrow$ By Borel-Cantelli lemma 2, $\{\frac{X_n}{\log n} \geq 1\}$ i.o. $\Rightarrow$ $\limsup \frac{X_n}{\log n} \geq 1$ a.s.

7.7.1 $X_1, X_2, \cdots$ r.v.s s.t. $S_n = X_1 + \cdots + X_n$ determine a martingale.
Show $E[X_i X_j] = 0$ if $i \neq j$.

Proof: $E[S_j \mid \mathcal{F}_{j-1}] = S_{j-1} \Rightarrow E[X_j \mid \mathcal{F}_{j-1}] = 0$.

suppose $j \geq i+1 > i$.

$\Rightarrow E[X_i \cdot X_j] = E[X_i \cdot E[X_j \mid \mathcal{F}_{j-1}]] = 0$.

7.7.3. $X_0, X_1, X_2, \cdots$ have finite means and $E[X_{n+1} \mid X_0, X_1, \cdots X_n] = aX_n + bX_{n-1}$. for $n \geq 1$.
where $0 < a, b < 1$. and $a+b=1$. Find α for $S_n = \alpha X_n + X_{n-1}$, $n \geq 1$, s.t.
S_n is a martingale w.r.t. sequence X.

$$E[S_n \mid \mathcal{F}_{n-1}] = E[\alpha X_n + X_{n-1} \mid \mathcal{F}_{n-1}]$$
$$= \alpha \cdot E[X_n \mid \mathcal{F}_{n-1}] + X_{n-1} \quad (= \alpha X_{n-1} + X_{n-2})$$
$$= \alpha(aX_{n-1} + bX_{n-2}) + X_{n-1}$$
$$= (\alpha \cdot a + 1) X_{n-1} + \alpha \cdot b\, X_{n-2}$$
$$\Rightarrow \alpha b = 1, \ \alpha = \tfrac{1}{b}. \text{ hence } \tfrac{1}{b} \cdot a + 1 = \tfrac{1}{b} = \alpha.$$

7.3.13. $\{X_r : 1 \leq r \leq n\}$ iid with mean μ and finite variance σ^2.
Let $\bar{X}_n = \frac{1}{n} \cdot \sum_{r=1}^{n} X_r$. Show that:

$$\sum_{r=1}^{n} (X_r - \mu) \Big/ \sqrt{\sum_{r=1}^{n} (X_r - \bar{X})^2} \quad \text{converges in distribution to } N(0,1)$$

Proof: $\dfrac{\sum_{r=1}^{n} (X_r - \mu)}{\sqrt{n} \cdot \sigma} \xrightarrow{D} N(0,1)$ by CLT.

and $\sum_{r=1}^{n} (X_r - \bar{X})^2 = \sum_{r=1}^{n} (X_r - \mu)^2 +$

$$\sum_{r=1}^{n} (X_r - \mu + \mu - \bar{X})^2 = \sum (X_r - \mu)^2 + n \cdot (\mu - \bar{X})^2 + \sum_{r=1}^{n} 2(X_r - \mu)(\mu - \bar{X})$$
$$= \sum (X_r - \mu)^2 + n(\mu - \bar{X})^2 + 2(\mu - \bar{X})(\sum X_r - n\mu)$$
$$= \sum (X_r - \mu)^2 - n(\mu - \bar{X})^2$$
$$\Rightarrow \frac{1}{n} \cdot \sum (X_r - \mu)^2 - (\mu - \bar{X})^2 \longrightarrow \sigma^2. \text{ a.s. by LLN.}$$

so the result holds.