Jan 29

- Weak Law of Large Numbers: (thm. 2.29)

$X_i \overset{iid}{\sim}$, with $E|X_i| < \infty$

Let $S_n = \sum_{i=1}^{n} X_i$, $EX_i = \mu$, then $S_n/n \xrightarrow{p} \mu$

ex: Monte Carlo integration

f defined on $[0,1]$, $I = \int_0^1 f(x)\,dx$

By WLLN:

If $\int_0^1 |f(x)|\,dx < \infty$, generate iid samples $u_i \sim [0,1]$

$$I_n = \frac{1}{n}\sum_{i=1}^{n} f(u_i) \quad \text{then } I_n \xrightarrow{p} I$$

Pf:/ $f(u_i) \overset{iid}{\sim}$, $E(f(u_i)) = \int_0^1 f(x) \cdot 1\,dx = I$

$$E|f(u_i)| = \int |f(x)|\,dx < \infty \quad \Rightarrow I_n \xrightarrow{p} I$$

Preliminaries:

- convergence in probability: $X_n \xrightarrow[n\to\infty]{P} X$

  if $\forall\, \varepsilon > 0$, $P(|X_n - X| > \varepsilon) \xrightarrow{n\to\infty} 0$

- uncorrelated RVs:

  X, Y are uncorrelated if $EX^2 < \infty$, $EY^2 < \infty$, and $E(XY) = E(X)E(Y)$

- Chebyshev's Inequality:

  $P(|X| > \varepsilon) \le EX^2/\varepsilon^2$. gives a bound for the tail probability in terms of the second moment.

  OR $P(|X| > \varepsilon) \le \dfrac{E|X|^p}{\varepsilon^p}$, $p > 0$

Thm 2.2.1: $X_1, \dots, X_n$ are uncorrelated. Then, $V(X_1 + \dots + X_n) = \sum_{i=1}^{n} V(X_i)$

Pf:/ $S_n = X_1 + \dots + X_n$

$V(S_n) = E(S_n^2) - (E(S_n))^2$. Let $EX_i = \mu_i$, $V(X_i) = \sigma_i^2$

$V(S_n) = E(\sum X_i)^2 - (\sum \mu_i)^2$

$= \sum EX_i^2 + \cancel{\sum_{i\neq j} X_i X_j} - \sum \mu_i^2 - \cancel{\sum_{i\neq j} \mu_i \mu_j}$

$$= \sum (EX_i^2 - \mu_i^2) = \sum V(X_i)$$

$\mathcal{L}^p$- convergence:

- $\mathcal{L}^p$- norm: $d(X,Y) = E|X-Y|^p$
- convergence in $\mathcal{L}^p$: $X_n \xrightarrow{d_p} X$ if $d(X_n, X) \xrightarrow{n\to\infty} 0$

Lemma 2.2: $\mathcal{L}^p$ convergence $\Rightarrow$ convergence in probability

Pf:/ (application of chebyshev's)

$$P(|X_n - X| > \varepsilon) \le \frac{1}{\varepsilon^p} E|X_n - X|^p \xrightarrow{n\to\infty} 0$$

Relationships between modes of convergence

Almost sure

conv. in prob.

weak conv / conv. in dist.

$\mathcal{L}^p$- conv.

Thm 2.2.3 (weak law under $\mathcal{L}^2$-bound):

- X_i's are uncorrelated. $EX_i = \mu$, $V(X_i) \le C \ \forall i$. Then, $S_n/n \to \mu$ in $\mathcal{L}^2$ and probability.

Pf:/

$$V\left(\frac{S_n}{n}\right) = \frac{1}{n^2} V(S_n) = \frac{1}{n^2} \sum V(X_i) \le \frac{1}{n^2} \sum C = \frac{Cn}{n^2} = \frac{C}{n} \xrightarrow{n\to\infty} 0$$

$$V\left(\frac{S_n}{n}\right) = E\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right)^2 = E\left(\frac{S_n}{n} - \mu\right)^2 \to 0$$

Thus $\frac{S_n}{n} \xrightarrow{\mathcal{L}^2} \mu$, $\mathcal{L}^2$ conv. $\Rightarrow$ conv. in prob:

$$\frac{S_n}{n} \xrightarrow{P} \mu.$$

Triangular Arrays of RVs

$$\begin{array}{lllll}
X_{1,1} & & & & \\
X_{2,1} & X_{2,2} & & & \\
X_{3,1} & X_{3,2} & X_{3,3} & & \\
 & \cdots & & & \\
X_{n,1} & \cdots & X_{n,k} & \cdots & X_{n,n}
\end{array}$$

Thm 2.2.4

$ES_n = \mu_n$, $V(S_n) = \sigma_n^2$. If $\frac{\sigma_n^2}{b_n^2} \to 0$, then

$$\frac{S_n - \mu_n}{b_n} \xrightarrow{P} 0$$

Pf:/ $E\left(\frac{S_n - \mu_n}{b_n}\right)^2 = \frac{1}{b_n^2} E(S_n - \mu_n)^2 = \frac{1}{b_n^2} \sigma_n^2 \to 0$

$$\Rightarrow \frac{S_n - \mu_n}{b_n} \xrightarrow{L^2} 0 \Rightarrow \frac{S_n - \mu_n}{b_n} \xrightarrow{P} 0$$

Thm. 2.2.6 (WLLN for triangular arrays):

$\{X_{n,k}\}$ is a triangular array with independence within rows

Let $\bar{X}_{n,k} = X_{n,k} I(|X_{n,k}| \le b_n)$ be a truncated version of $X_{n,k}$.

$b_n > 0$, $b_n \to \infty$. Suppose:

i) $\sum_{k=1}^{n} P(|X_{n,k}| > b_n) \xrightarrow{n\to\infty} 0$ (low probability of truncation)

ii) $\frac{1}{b_n^2} \sum E\bar{X}_{n,k}^2 \xrightarrow{n\to\infty} 0$ (constants dominate variances of truncated RVs)

($V(X) = EX^2 - (EX)^2 \le EX^2$

If the above conditions hold, then

$$\frac{S_n - a_n}{b_n} \xrightarrow{P} 0, \quad a_n = \sum_{k=1}^{n} \bar{X}_{n,k} \quad \text{(means of truncated variables)}$$

$$S_n = \sum_{k=1}^{n} X_{n,k} \quad \text{(sum of } n^{th} \text{ row)}$$

pf (idea):/

$$S_n = \sum_{k=1}^{n} X_{n,k}, \quad \bar{S}_n = \sum_{k=1}^{n} \bar{X}_{n,k}$$

Show $\frac{\bar{S}_n - a_n}{b_n} \to 0$ (as b_n dominates the variance) (thm. 2.2.4)

Show $S_n = \bar{S}_n$ with high probability (as probability of truncation is small)

Then $\frac{S_n - a_n}{b_n} \xrightarrow{P} 0$

Pf:/ i) $P(S_n \neq \bar{S}_n) \leq P\left(\bigcup_{k=1}^{n} \{|X_{n,k}| > b_n\}\right)$

$$\leq \sum_{k=1}^{n} P(|X_{n,k}| > b_n) \to 0 \quad (\text{by (i)})$$

truncated sum differs from untruncated sum by a very small probability

ii) $$P\left(\left|\frac{\bar{S}_n - a_n}{b_n}\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} E\left(\frac{\bar{S}_n - a_n}{b_n}\right)^2 = \frac{V(\bar{S}_n)}{\varepsilon^2 b_n^2}$$

within each row, RVs are independent so variance of sums = sum of variances

$$= \frac{1}{\varepsilon^2 b_n^2} \sum_{k=1}^{n} V(\bar{X}_{n,k}) \to 0 \quad (\text{by (ii)})$$

$$P\left(\left|\frac{S_n - a_n}{b_n}\right| > \varepsilon\right) \leq P\left(\{S_n \neq \bar{S}_n\} \cup \left\{\left|\frac{\bar{S}_n - a_n}{b_n}\right| > \varepsilon\right\}\right) \to 0$$

Jan 26

Preliminaries:

lemma 2.2.8: If $Y \geq 0$ and $p > 0 \Rightarrow EY^p = \int_0^\infty p y^{p-1} P(Y > y)\, dy$

special case: $EY = \int P(Y > y)\, dy$
$(p = 1)$

measure of prob. corresponding to Y on space where Ω is defined

pf:/ RHS: $\int_0^\infty p y^{p-1} \int I(Y > y)\, d\mu(Y)\, dy$

lesbegue measure on $\mathbb{R}$

$Y : \Omega \to \mathbb{R}$

$= \int \int_0^\infty p y^{p-1} I(y < Y)\, dy\, d\mu(y)$ ↓ by Fubini's thm b/c it's positive

$= \int \int_0^Y p y^{p-1}\, dy\, d\mu(y) = \int y^p\, d\mu(y) = E(Y^p)$

1.4.7 (v): X_1, X_2 RVs such that $X_1 = X_2$ a.s.

$E(|X_1|) < \infty$, $E(|X_2|) < \infty \Rightarrow EX_1 = EX_2$

e.g. $X_1 = 1$ everywhere

$X_2 = 1$ almost everywhere

$= \infty$ w/ probability 0

$$EX_1 = 1 = EX_2 = 1 \times P(X_2 = 1) + \infty \times P(X_2 = \infty)$$

$$= 1 + \underbrace{\infty \cdot 0}_{0}$$

$$= 1$$

Thm. 2.2.7: X_i's iid, $xP(|X_1| > x) \to 0$ as $x \to \infty$. $\mu_n = E(X_1 I(|X_1| \le n))$.

Then $\frac{S_n}{n} - \mu_n \xrightarrow{P} 0$

Pf (idea): / set up as triangular array.

X_1

$X_1 \quad X_2$

$X_1 \quad X_2 \quad X_3$

$\vdots$

$X_{n,k} = X_k$, $b_n = n$, $\bar{X}_{n,k} = X_k I(|X_k| \le n)$

$$a_n = E\left(\sum X_k I(|X_k| \le n)\right) = n\mu_n$$

(i) Show $\sum_{k=1}^{n} P(|X_{n,k}| > b_n) \to 0$ $\quad P(|X_k| > n)$

$= n P(|X_1| > n)$ because X_i's are iid

$\to 0$ (tail probability assumption)

(ii) Show $b_n^{-2} \sum_{k=1}^{n} E\bar{X}_{n,k}^2 \to 0$

$$E\bar{X}_{n,1}^2 = \int_0^\infty 2y\, P(|\bar{X}_{n,1}| > y)\, dy \quad \text{(lemma 2.2.8)}$$

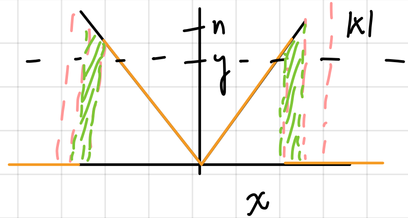
$$= \int_0^n 2y\, P(|\bar{X}_{n,1}| > y)\, dy \quad X_{n,1} = X_1 I(|X_1| \le n)$$

$$\le \int_0^n \underbrace{2y\, P(|X_1| > y)}_{g(y)}\, dy$$

$g(y) < 1$ as $n \to \infty$

$g(y) \xrightarrow{y \to \infty} 0$, show $\frac{1}{n}\int_0^n g(y)\, dy \to 0$

$$\frac{1}{n}\int_0^n g(y)\, dy = \frac{\int_0^n g(y)\, dy}{\int_0^n 1\, dy} \to 0$$



$z = |x| I(|x| \le n)$

$P(|\bar{X}_{n,1}| > y)$ $\{|x| > y\}$ $\quad \{|z| > y\} \le P(|X_1| > y)$

$$\frac{1}{n^2} \times n E\bar{X}_{n,1}^2 = \frac{E\bar{X}_{n,1}^2}{n} \le \frac{1}{n}\int_0^n g(y)\,dy \to 0$$

Since (i) and (ii) are satisfied, by WLLN for triangular arrays,

$$\frac{S_n - a_n}{b_n} \xrightarrow{P} 0 \quad , \quad \frac{S_n}{n} - \mu_n \xrightarrow{P} 0$$

<u>Proving the WLLN:</u>

Thm 2.2.9: X_i's iid, $E|X_i| < \infty$, $EX_1 = \mu$.

$$S_n = \sum_{i=1}^{n} X_i \ , \text{ then } \frac{S_n}{n} \xrightarrow{P} \mu$$

Pf:/ show $x P(|X_1| > x) \xrightarrow[]{x \to \infty} 0$

$$x P(|X_1| > x) \le \underbrace{E(|X_1| I(|X_1| > x))}_{f_x}$$

$$f_x \xrightarrow{x \to \infty} \infty I(|X_1| = \infty) \quad \text{(pointwise)}$$

$|f_x| \le |X_1|$ and $E|X_1| < \infty$

bounded by another RV

$$x P(|X_1| > x)$$

$$= x \cdot E(I(|X_1| > x))$$

$$= E(x \cdot I(|X_1| > x))$$

$$\le E(|X_1| I(|X_1| > x))$$

$$\le E(|X_1| I(|X_1| < x)) + \underbrace{E(|X_1| I(|X_1| \ge x))}_{\to 0}$$

$$E f_x \xrightarrow{x \to \infty} \infty \quad P(|X_1| = \infty) = 0$$

dominated convergence thm

so $x P(|X_1| > x) \to 0$

$$\frac{S_n}{n} - \mu_n \xrightarrow{P} 0 \Rightarrow P\left(\left|\frac{S_n}{n} - \mu_n\right| > \varepsilon/2\right) \to 0$$

truncated mean

$Y_n = X_1 \cdot I(|X_1| < n)$

$Y_n \to X_1$ as $n \to \infty$

Y_n is bounded by X_1

and $X_1 \leq X_1$.

Now need to show $\mu_n \xrightarrow{P} \mu$.

$$\mu_n = E X_1 I(|X_1| \leq n) \xrightarrow[DCT]{n \to \infty} E X_1 = \mu$$

$E X_1 I(|X_1| < n) \to$

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \leq \underbrace{P\left(\left|\frac{S_n}{n} - \mu_n\right| > \varepsilon\right)}_{\to 0} + \underbrace{I(|\mu_n - \mu| > \varepsilon/2)}_{\to 0}$$

$\mu_n \to E Y_n$?

$E X_1 = \mu$

$\mu_n \to \mu$.

Chp. 2.3 + 2.4: Strong Law of Large Numbers (SLLN)

- Limits of RVs:
  - $\limsup X_n = \overline{\lim}\, X_n$ is a RV X_{sup}

    such that $X_{sup}(\omega) = \overline{\lim}\,(X_n(\omega)) \ \forall\, \omega \in \Omega$

  - $\liminf X_n = \underline{\lim}\, X_n$ is a RV X_{inf}

    such that $X_{inf}(\omega) = \underline{\lim}\,(X_n(\omega)) \ \forall\, \omega \in \Omega$

Limit of X_n exists for some ω if $\overline{\lim}(X_n(\omega)) = \underline{\lim}\,(X_n(\omega))$

- $X_n \xrightarrow[a.e.]{a.s.} X$ if

  $\lim X_n(\omega)$ exists $\forall\, \omega \in \Omega_0$ s.t. $P(\Omega_0) = 1$

  $X(\omega) = \lim X_n(\omega)$, $\omega \in \Omega_0$

  $=$ anything, $\omega \notin \Omega_0$

$\cdot\ X_n \xrightarrow{P} X \underset{?}{\Longrightarrow} X_n \xrightarrow{a.s.} X$ ✓ typewriter sequence

$\Omega = (0,1] \qquad X_{n,k}(\omega) = I\left(\frac{k-1}{n} < \omega \le \frac{k}{n}\right), \ k = 1, \dots, n$

$$\begin{array}{lll} & \overset{Y_1}{X_{1,1}} & \\ Y_2 & X_{2,1} & \overset{Y_3}{X_{2,2}} \\ & \vdots & \end{array} \quad \Big\downarrow \quad Y_m$$

$Y_m(\omega)$ does not have a limit b/c

$I(\omega)$ is 1 at least once every row.

Every $Y_m = X_{n,k}$ for some n and k.

$$m \to \infty \iff n \to \infty$$

$$P(|Y_m| > \varepsilon) = P(|Y_m| \neq 0) = P(|X_{n,k}| \neq 0)$$
$$= \frac{1}{n} \xrightarrow[m\to\infty]{} 0$$

So $Y_m \xrightarrow{P} 0$ but $Y_m \overset{a.s.}{\not\to} 0$

Jan 31

Sequences of sets

· Limit of decreasing sets:

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$$

i.e.
$$\begin{array}{ccccccc} (& & & B_1 & & &) \\ & (& & B_2 & &) & \\ & & (& B_3 &) & & \\ & & & \vdots & & & \end{array}$$

$$\lim B_n = \{\omega : \omega \in B_m, \forall m\}$$

· Limit of increasing sets:

$$B_1 \subseteq B_2 \subseteq B_3 \cdots$$

$$\begin{array}{ccccccc} & & (& B_1 &) & & \\ & (& & B_2 & &) & \\ (& & & B_3 & & &) \\ & & & \vdots & & & \end{array}$$

for large enough

$$\lim B_n = \{\omega : \omega \in B_m, \forall m\}$$

(all but finitely many)

$A_1, A_2, \ldots$ sets $\subseteq \Omega$

$I(A_n)$ is a binary RV $\qquad I(\omega : \omega \in A_n)$, $\omega \in \Omega$

$\overline{\lim} \, I(A_n) = 1 \iff \exists$ at least one subsequence of 1's in $I(A_n)$

$\iff I(A_n) = 1$ infinitely often

$\iff \forall m \ \exists \ k > m$ s.t. $I(A_k) = 1$

$\iff I\left(\bigcup_{n=m}^{\infty} A_n\right) = 1 \ \forall m \qquad B_m = \bigcup_{n=m}^{\infty} A_n$

$\overline{\lim} A_n := \lim_{m \to \infty} \bigcup_{n=m}^{\infty} A_n$

$\iff I\left(\lim_{m \to \infty} B_m\right) = 1 \qquad B_m$ is decreasing as m increases (union of fewer sets)

$\iff I\left(\overline{\lim} A_n\right) = 1$

Thm: $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X$

pf:/

$$A_m = I(|X_m - X| > \varepsilon)$$

$$P(\overline{\lim} A_n) = P(|X_n - X| > \varepsilon, \text{ i.o.})$$

$$= 0$$ because $X_n \xrightarrow{a.s.} X$

$$P\left(\lim_{m\to\infty} \bigcup_{n=m}^{\infty} A_n\right) = 0 \Rightarrow \lim_{m\to\infty} P\left(\bigcup_{n=1}^{\infty} A_n\right) = 0$$

prob. is continuous under increasing/decreasing limits (MCT)

$$\lim_{m\to\infty} P(|X_m - X| > \varepsilon) = \lim_{m\to\infty} P(A_m) \le \lim_{m\to\infty} P\left(\bigcup_{n=m}^{\infty} A_n\right) = 0$$

$$\Rightarrow X_n \xrightarrow{P} X$$

$A_1, A_2, \dots$

$$\underline{\lim}(A_n) = 1 \iff I(A_m) = 1 \ \forall m$$

always true after a certain m.

$$\iff I\left(\bigcap_{n=m}^{\infty} A_n\right) = 1 \ \forall m$$

$B_m = \bigcap_{n=m}^{\infty} A_n$ is increasing

$$\iff I\left(\lim_{m\to\infty} \bigcap_{n=m}^{\infty} A_n\right) = 1$$

$$\underline{\lim}\, A_n := \lim_{m\to\infty} \bigcap_{n=m}^{\infty} A_n$$

Thm 2.3.1 Borel-Cantelli Lemma 1:

$$\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(\overline{\lim} A_n) = 0$$

Pf:/ $\overline{\lim} A_n = A_n$ i.o.

$$E \sum I(A_n) = \sum E I(A_n) \quad \text{by Fubini's Thm}$$

$$= \sum P(A_n)$$

$$< \infty$$

$X = \sum I(A_n)$ is finite a.s.

$P(A_n \text{ i.o.}) = 0$

Thm 2.3.2 : $X_n \xrightarrow{P} X \iff \forall$ subsequence n_m $\exists$ a further subsequence $n_{m(k)}$ s.t. $X_{n_{m(k)}} \xrightarrow{a.s.} X$

Pf:/ ($\Rightarrow$) $\varepsilon_k > 0$ s.t. $\varepsilon_k \xrightarrow{k\to\infty} 0$

$$X_n \xrightarrow{P} X \Rightarrow X_{n(m)} \xrightarrow{P} X$$

$$a_n = P(|X_n - X| > \varepsilon) \to 0 \Rightarrow a_{n(m)} \to 0$$

$\therefore \exists N_k$ s.t. $\forall m \geq N_k$, $P(|X_{n(m)} - X| > \varepsilon_k) \leq \frac{1}{2^k}$

Pick one such member of the sequence $\{X_{n(m)}\}$ and denote it by $X_{n_{m(k)}}$. Repeat this for $k = 1, 2, \ldots$

$\therefore$ $\forall k$, $P(|X_{n_{m(k)}} - X| > \varepsilon_k) \leq 1/2^k$

WLOG: $n_{m(k)} > n_{m(k-1)}$ (make it increasing)

$$\sum_k P(|X_{n_{m(k)}} - X| > \varepsilon_k) \leq \sum_k 1/2^k < \infty$$

$$P(|X_{n_{m(k)}} - X| > \varepsilon_k \text{ i.o.}) = 0$$

$$\therefore X_{n_{m(k)}} \xrightarrow{a.s.} X$$

$(\Leftarrow)$ $y_n = P(|X_n - X| > \varepsilon)$

$\therefore \forall n_m, \exists n_{m(k)}$ s.t. $X_{n_{m(k)}} \xrightarrow{a.s.} X$

$$\therefore X_{n_{m(k)}} \xrightarrow{P} X$$

$$\therefore y_{n_{m(k)}} = P(|X_{n_{m(k)}} - X| > \varepsilon) \to 0$$

$\therefore \ \forall n_{m_l}, \ \exists \ n_{m(k)}$ s.t. $y_{n_{m(k)}} \to 0$

$\Rightarrow \ y_n \to 0$ by properties of $\mathbb{R}$

Thm. 2.3.4 Continuous Mapping Theorem:

If f is a continuous function and $X_n \xrightarrow{P} X$, then $f(X_n) \xrightarrow{P} f(X)$

$X_n \xrightarrow{a.s.} X$, then $f(X_n) \xrightarrow{a.s.} f(X)$

pf:/ $Y_n = f(X_n)$

$\forall \ n_m \ \exists \ n_{m(k)}$ s.t. $X_{n_{m(k)}} \xrightarrow{a.s.} X$

$\therefore \ f(X_{n_{m(k)}}) \xrightarrow{a.s.} f(X)$

$\therefore \ Y_{n_{m(k)}} \xrightarrow{P} f(X)$

$$a_n = P(|f(X_n) - f(X)| > \varepsilon)$$

$$a_{n_{m_k}} = P(|Y_{n_{m(k)}} - f(X)| > \varepsilon) \to 0$$

?

$$\therefore \quad a_n \longrightarrow 0$$

Thm 2.3.6 Borel-Cantelli Lemma 2:

A_n's are independent events and $\sum P(A_n) = \infty \Rightarrow P(\overline{\lim} A_n) = 1$

pf:/ $\overline{\lim} A_n = \lim_{m\to\infty} \bigcup_{n=m}^{\infty} A_n = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$

( )

( )

( )

⋮

$$(\overline{\lim} A_n)^c = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n^c = \lim_{m\to\infty} \bigcap_{n=m}^{\infty} A_n^c$$

$$\lim_{m\to\infty} P\left(\bigcap_{n=m}^{\infty} A_n^c\right) = \lim_{m\to\infty} \prod_{n=m}^{\infty} P(A_n^c) = \lim_{m\to\infty} \prod_{n=m}^{\infty} (1 - P(A_n))$$

$$\le \exp\left(-\sum P(A_n)\right) = 0$$

Note: $1 - x \le e^{-x}$ if $0 < x < 1$

$$\lim_{m\to\infty} P\left(\bigcap_{n=m}^{\infty} A_n^c\right) = P\left(\lim_{m\to\infty} \bigcap_{n=m}^{\infty} A_n^c\right)$$

$$\Rightarrow P\left(\left(\overline{\lim} A_n\right)^c\right) = 0$$

WTS: $e^{x} \geq (1-x)$, $x \geq 0$.

$g(x) = e^{-x} - (1-x)$

$g'(x) = -e^{-x} + 1 \geq 0 \Rightarrow g(x) \geq g(0) = e^0 - (1-0) = 0.$

$\{a_n\}$ is a sequence of real numbers

for all subsequences a_{n_m},

$\exists\ a_{n_{m(k)}}$ s.t. $a_{n_{m(k)}} \to a$ (same a for all sub-subsequences)

then $a_n \to a$

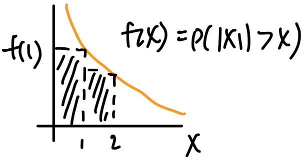
Feb 2

Thm. 2.4.1 SLLN:

Pf:/ $Y_k = X_k I(|X_k| \le k)$, $T_n = \sum_{i=1}^{n} Y_i$

Lemma 2.4.2: $\frac{S_n}{n} - \frac{T_n}{n} \xrightarrow{a.s.} 0$

Pf:/ $\sum P(|X_k| > k) = \sum_{k=1}^{\infty} \underbrace{P(|X_1| > k)}_{f(k)}$ X_i's are iid



$\le \int_0^{\infty} P(|X_1| > k)\,dx$ $\int_0^{\infty} P(|Y| > y)\,dy = \mathbb{E}|Y|$

summation = area of the bars $\le$ area under the curve

$= \mathbb{E}|X_i| < \infty$

for all k's $\Rightarrow P(|X_k| > k \text{ i.o.}) = 0$

$\Rightarrow P(X_k \ne Y_k \text{ i.o.}) = 0$ X_k only differs from Y_k if $|X_k| > k$

$S_n - T_n$ is finite w/ prob. 1

$\frac{(S_n - T_n)}{n} \xrightarrow{a.s.} 0$

$$\frac{T_{k(n)}}{k(n)} \xrightarrow{a.s.} \mu \text{ for a subsequence } k(n) = [\alpha^n], \alpha > 1$$

floor

by Chebyshev's

$$\sum_{1}^{\infty} P\left(\frac{T_{k(n)} - \mathbb{E}T_{k(n)}}{k(n)} > \varepsilon \right) \leq \frac{1}{\varepsilon^2} \sum_{1}^{\infty} V(T_{k(n)}) / k(n)^2$$

$$= \frac{1}{\varepsilon^2} \sum_{1}^{\infty} \frac{1}{k(n)^2} \sum_{m=1}^{k(n)} V(Y_m)$$

$$= \frac{1}{\varepsilon^2} \sum_{m=1}^{\infty} V(Y_m) \sum_{n: k(n) \geq m} \frac{1}{k(n)^2}$$ *

$$\leq \frac{1}{\varepsilon^2} \frac{1}{(1-\alpha^{-2})} \sum_{m=1}^{\infty} V(Y_m) / m^2$$

$$< \infty \quad (\text{lemma } 2.4.3)$$

$$\sum_{n: k(n) \geq m} \frac{1}{k(n)^2}$$

Note: $[x] \geq \frac{x}{2}$

$$= \sum_{n: [\alpha^n] \geq m} \frac{1}{[\alpha^n]^2}$$

$$\leq \sum_{n: \alpha^n \geq m} \frac{4}{(\alpha^n)^2} \leq \frac{4}{m^2} \sum_{n=0}^{\infty} \frac{1}{\alpha^{2n}}$$

$$= \frac{4}{m^2(1-\alpha^{-2})}$$

b/c $\alpha > 1$

$$\Rightarrow P\left(\frac{T_{k(n)} - \mathbb{E}T_{k(n)}}{k(n)} > \varepsilon \text{ i.o.} \right) = 0$$

$$\Rightarrow \frac{T_{k(n)} - \mathbb{E}T_{k(n)}}{k(n)} \xrightarrow{a.s.} 0$$

$$\frac{ET_n}{n} \to ?$$

$$EY_k = EX_1 I(|X_1| \le k) \xrightarrow[k\to\infty]{DCT} EX_1$$

$$a_k \to a \Rightarrow \underbrace{\frac{1}{k}\sum_{i=1}^{k} a_i}_{\text{cesaro sum}} \to a$$

$$\Rightarrow \frac{ET_n}{n} \to \mu \quad \text{b/c } \frac{ET_n}{n} \text{ is a}$$

cesaro sum of EY_k

So $\frac{T_{k(n)}}{k(n)} - \mu \xrightarrow{a.s.} 0 \qquad \frac{ET_{k(n)}}{k(n)} \to \mu$

What we have so far:

$$\frac{S_n}{n} - \frac{T_n}{n} \xrightarrow{a.s.} 0, \quad \frac{T_{k(n)}}{k(n)} \xrightarrow{a.s.} \mu, \qquad k(n) = [\alpha^n]$$

Assume all X_i's are ≥ 0, $k(n) \le m \le k(n+1)$

$$\frac{T_{k(n)}}{k(n+1)} \le \frac{T_m}{m} \le \frac{T_{k(n+1)}}{k(n)}$$

$$\underbrace{\frac{k(n)}{k(n+1)}}_{1/\alpha} \underbrace{\frac{T_{k(n)}}{k(n)}}_{\mu} \le \frac{T_m}{m} \le \underbrace{\frac{T_{k(n+1)}}{k(n)}}_{\mu} \underbrace{\frac{k(n+1)}{k(n)}}_{\alpha}$$

$$k(n) = [\alpha^n], \quad \lim_{n\to\infty} \frac{k(n+1)}{k(n)} = \alpha$$

$$\alpha^n - 1 \le [\alpha^n] \le \alpha^n \qquad \alpha > 1$$

So $\frac{1}{\alpha}\mu \le \underline{\lim} \frac{T_m}{m} \le \overline{\lim} \frac{T_m}{m} \le \mu\alpha$ a.s.

As $\alpha \to 1$, $\underline{\lim} \frac{T_m}{m} = \overline{\lim} \frac{T_m}{m}$

making subsequence denser and denser

$$\Rightarrow \mu \le \lim \frac{T_m}{m} \le \mu$$

So $\frac{S_n}{n} \xrightarrow{a.s.} \mu$

If X_i's not ≥ 0:

$X_i = X_i^+ - X_i^-$ $\qquad \mu^+ = E X_i^+$

$\mu = \mu^+ - \mu^-$ $\qquad \mu^- = E X_i^-$

$$\frac{S_n}{n} = \frac{\sum X_i^+}{n} - \frac{\sum X_i^-}{n}$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\mu \qquad \mu^+ \qquad \mu^- \qquad \blacksquare$$

Example: $X_i \overset{iid}{\sim}$, $EX_i = \mu$, $EX_i^2 < \infty$

$\frac{S_n}{n} \xrightarrow{a.s.} \mu$, sample variance $= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

Notation:

$EX_i^2 = \mu_2$, $V(X_i) = \sigma^2$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$$

$$\downarrow$$

$$EX_1^2$$

$= \mu_2 - \mu^2$ by CMT

a.s.

$= \sigma^2$ a.s.

Example : Empirical CDF

$$X_i \overset{iid}{\sim} F. \qquad F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \underbrace{I(X_i \le x)}_{\text{Finite expectation}}$$

$$\text{For a fixed } x, \quad F_n(x) \xrightarrow{a.s.} E\,I(X_i \le x)$$

$$= P(X_1 \le x)$$

$$= F(x) \quad \text{by SLLN}$$

Thm. 2.4.7 Glivenko-Cantelli:

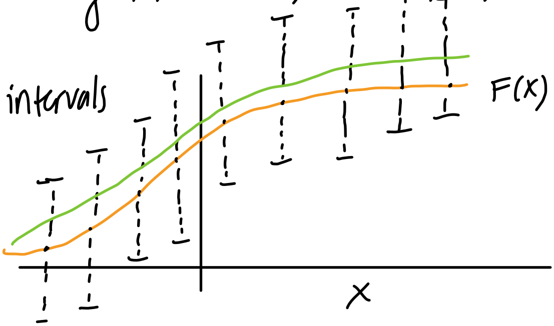
The convergence of $F_n(x)$ is uniform.

$$Z_n = \sup_x \left| F_n(x) - F(x) \right| \xrightarrow{a.s.} 0$$

Pf. (idea) / F_n and F are non-decreasing.

For any x, $F_n(x) \xrightarrow{a.s.} F(x)$

tighten intervals



for each point, can pick n large enough such that $|F_n(x) - F(x)|$ is small.

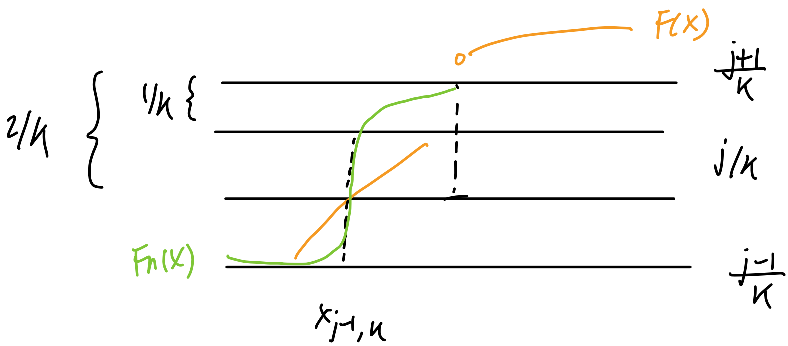
Pf:/ $F_n(x) \xrightarrow{a.s.} F(x)$, $F_n(x^-) = \frac{1}{n}\sum I(X_i < x) \xrightarrow{a.s.} F(x^-)$

$$x_{j,k} = \inf\{y : F(y) \geq j/k\}, \quad j = 1, \dots, k-1$$

$$x_{0,k} = -\infty, \quad x_{k,k} = \infty$$

$\therefore\ \exists\ n_k$ s.t. $\forall\ n \geq n_k$, $|F_n(x_{j,k}) - F(x_{j,k})| \leq 1/k$

$|F_n(x_{j,k}^-) - F(x_{j,k}^-)| \leq 1/k \quad \forall j$



no curve will have a worse difference than $2/k$.

$$n \geq n_k, \quad \sup_x |F_n(x) - F(x)| \leq \frac{2}{k} \xrightarrow{k\to\infty} 0$$

Feb 7

- Convergence in Dist. / weak convergence

$$X_n \xrightarrow{d} X \sim F \text{ if } \underbrace{F_n(x)}_{\text{CDF of } X_n} \xrightarrow{n\to\infty} F(x)$$

for all continuity points x of F

Ex: $X_1, \dots, X_n \overset{iid}{\sim} F$

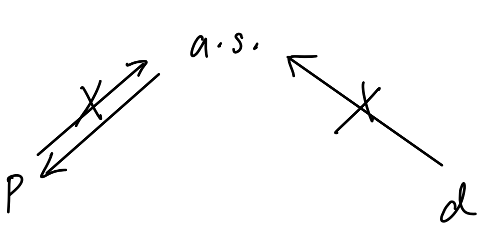
Z_n = randomly selected one of $X_1, \dots, X_n$.

$$F_{Z_n}(x) = P(Z_n \le x)$$

$$= \sum_{i=1}^{n} P(Z_n \le x \mid Z_n = i\text{th data point})\, \underbrace{P(Z_n = i\text{th data point})}_{1/n}$$

$$= \frac{1}{n}\sum_{i=1}^{n} P(X_i \le x)$$

$$= F(x)$$



Ex: $X \sim N(0,1)$

$X_1 = X, \ X_2 = -X, \ X_3 = X, \ X_4 = -X.$

$X_1 \sim N(0,1), \quad X_2 \sim N(0,1) \ldots$

$F_n = \text{cdf}(X_n) = F$

$X_n \xrightarrow{d} X$ but $X_n \not\xrightarrow{a.s.} X$

Thm 3.2.2 Skorohod's Thm.

$X_n \sim F_n, \ X_n \xrightarrow{d} X \sim F$. Then, $\exists \ Y_n \sim F_n$

and $Y \sim F$ s.t. $Y_n \xrightarrow{a.s.} Y$

Pf:/ $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}((0,1))$, $P =$ Lesbesgue

$U(\omega) = \omega$ (uniform 0-1)

$P(U(\omega) \leq x) = P(\omega \leq x) = x$

$Y_n(\omega) = F_n^{-1}(U(\omega)) = F_n^{-1}(\omega)$

$= \inf\{x : F_n(x) \geq \omega\}$

$Y_n \xrightarrow{d} F_n$ (Durrett Thm 1.2.2)

$Y(\omega) = F^{-1}(U(\omega)) = \inf\{x : F(x) \geq \omega\}$

$Y \xrightarrow{d} F$

1) Inverses of non-decreasing fn's are non-decreasing
2) Non-decreasing fns have at most countably many discontinuities

Take $\omega \in (0,1)$, $\varepsilon > 0$ and x s.t. $Y(\omega) - \varepsilon < x < Y(\omega)$ and x is a continuity point.

If $F(x) \geq \omega$ then $\underbrace{Y(\omega) \leq x}$ (contradiction)

$\therefore F(x) < \omega$ b/c $Y(\omega)$ is the infimum

$\therefore F_n(x) < \omega$ for sufficiently large n.

If $x \geq Y_n(\omega) \Rightarrow F_n(x) \geq F_n(Y_n(\omega))$

$\geq \omega$ (contradiction)

$x_m \downarrow Y_n(\omega)$ $F_n(x_m) \geq \omega$, $F_n(Y_n(\omega)) \geq \omega$

$\therefore x < Y_n(\omega)$ for sufficiently large n.

$\Rightarrow x \leq \underline{\lim} Y_n(\omega)$

$\underline{\lim} Y_n(\omega) > Y(\omega) - \varepsilon$

$\therefore \underline{\lim} Y_n(\omega) \geq Y(\omega)$

Let $\omega' > \omega$. Choose y s.t. $Y(\omega') < y < Y(\omega') + \varepsilon$ and F is continuous at y.

$$\omega < \omega' \leq F(Y(\omega')) \leq F(y)$$

$\therefore$ $F(y) > \omega \Rightarrow F_n(y) > \omega$ for sufficiently large n.

$\therefore$ $\omega \leq F_n(y)$ for large enough n

$$\Rightarrow Y_n(\omega) \leq y < Y(\omega') + \varepsilon \quad \text{for large enough } n$$

$\therefore$ $Y_n(\omega) \leq y < Y(\omega')$

If Y is continuous at ω, $\lim Y_n(\omega) \leq Y(\omega)$

$\therefore$ If Y is continuous at ω, $Y_n(\omega) \to Y(\omega)$

$$Y_n \xrightarrow{a.s.} Y$$

b/c Y has countably many discontinuities.

Portmanteau Lemma: The following are equivalent:

i) $X_n \xRightarrow{d} X$

ii) $\mathbb{E} f(X_n) \to \mathbb{E} f(X)$ $\forall$ bounded, cont. f

iii) $\mathbb{E} f(X_n) \to \mathbb{E} f(X)$ $\forall$ bounded, Lipschitz f

((i) $\Rightarrow$ (ii)): $X_n \xRightarrow{d} X$

By Skorohod, $Y_n \xRightarrow{d} Y$, $Y \overset{d}{\Rightarrow} X$, $Y_n \xrightarrow{a.s.} Y$

$\therefore f(Y_n) \xrightarrow{a.s.} f(Y)$ (f is cont.)

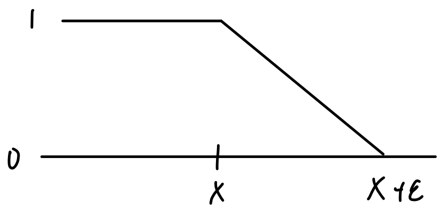
$$\begin{array}{ccc} \mathbb{E} f(Y_n) & \xrightarrow{DCT} & \mathbb{E} f(Y) \\ \downarrow & & \downarrow \\ \mathbb{E} f(X_n) & \to & \mathbb{E} f(X) \end{array}$$

b/c X_n and Y_n have the same dist.

((ii) $\Rightarrow$ (iii)) f is Lipschitz if $|f(x) - f(y)| \leq \gamma |x-y|$ for some $\gamma > 0$ and all x, y

Lipschitz $\Rightarrow$ continuous

((iii) $\Rightarrow$ (i))



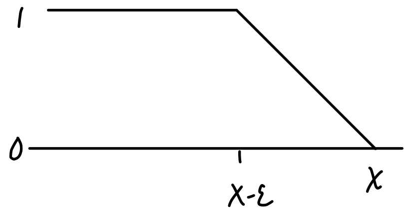
$$g_{\varepsilon,x}(u) = 1 \text{ if } u \leq x$$
$$= 0 \text{ if } u \geq x+\varepsilon$$
$$= \text{linear in between}$$

$\mathbb{E}\, g_{\varepsilon,x}(X_n) \to \mathbb{E}\, g_{\varepsilon,x}(X)$ as $g_{\varepsilon,x}$ is Lipschitz

$$P(X_n \leq x) = \mathbb{E}\, I(X_n \leq x) \leq \mathbb{E}\, g_{\varepsilon,x}(X_n)$$

$$\overline{\lim}\, P(X_n \leq x) \leq \mathbb{E}\, g_{,\varepsilon}(X) \leq P(X \leq x+\varepsilon)$$

$\varepsilon \to 0 \Rightarrow \overline{\lim}\, P(X_n \le x) \le P(X \le x)$



$$g_{\varepsilon, x-\varepsilon}(u) = 1 \text{ if } u \le x-\varepsilon$$
$$= 0 \text{ if } u \geq x$$
$$= \text{linear in between}$$

$$P(X_n \le x) \geq \mathbb{E} g_{\varepsilon, x-\varepsilon}(X_n)$$

$$\therefore \underline{\lim}\, P(X_n \le x) \geq \mathbb{E}_{g, x-\varepsilon}(X) \geq P(X \le x - \varepsilon)$$

$$\therefore \underline{\lim}\, P(X_n \le x) \geq P(X \le x - \varepsilon)$$

If x is a continuity point,

$$\underline{\lim}\, P(X_n \le x) \geq P(X \le x)$$

$$\Rightarrow \lim F_n(x) = F(x)$$

(iv) $\underline{\lim}\, \mathbb{E} f(X_n) \geq \mathbb{E} f(X)$ $\forall$ cont. non-negative f

(v) $\underline{\lim}\, P(X_n \in G) \geq P(X \in G)$ $\forall$ open sets G

(vi) $\overline{\lim}\, P(X_n \in F) \le P(X \in F)$ $\forall$ closed sets F

(vii) $P(X_n \in B) \to P(X \in B)$ $\forall$ Borel sets B w/ $P(X \in \partial B) = 0$ where $\partial B =$ boundary of B

Feb 9

Application of SLLN:

$$(Y_i, X_i) \overset{iid}{\sim} \text{ s.t. } E(X_i^2) < \infty, \ E(Y_i^2) < \infty$$

$$\text{Cov}(X_i, Y_i) = E(X_iY_i) - E(X_i)E(Y_i)$$

$$\text{Corr}(X_i, Y_i) = \rho = \frac{\text{Cov}(Y_i, X_i)}{\sqrt{V(X_i)\,V(Y_i)}}$$

$$\hat{\rho} = \frac{\frac{1}{n}\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n}\sum(X_i-\bar{X})^2 \cdot \frac{1}{n}\sum(Y_i-\bar{Y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

$$= \frac{\underbrace{\frac{1}{n}\sum Y_iX_i}_{\xrightarrow{a.s.} E(X_iY_i)} - \underbrace{\bar{Y}}_{\xrightarrow{a.s.} \mu_Y}\,\underbrace{\bar{X}}_{\xrightarrow{a.s.} \mu_X}}{\sqrt{\underbrace{S_{xx}}_{\xrightarrow{a.s.} \sigma_X^2}\ \underbrace{S_{yy}}_{\xrightarrow{a.s.} \sigma_Y^2}}}$$

$$\Rightarrow \quad \hat{\rho} \xrightarrow{a.s.} \frac{E(X_1Y_1) - E(X_1)E(Y_1)}{\sqrt{V(X_1)\,V(Y_1)}} = \rho$$

· Limar regression

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad X_i \overset{iid}{\sim}, \quad EX_i = \mu_X, \quad EX_i^2 < \infty$$

$$\varepsilon_i \overset{iid}{\sim}, \quad E\varepsilon_i = 0, \quad E\varepsilon_i^2 < \infty,$$

$$\varepsilon\text{'s} \perp X\text{'s}$$

$$\hat{\beta} = \frac{S_{XY}}{S_{XX}} \xrightarrow{a.s.} \frac{E(X_1 Y_1) - E(X_1)E(Y_1)}{V(X_1)}$$

$$E(X_1 Y_1) = E(X_1(\alpha + \beta X_1)) = \alpha \mu_X + \beta E X_1^2$$

$$E(X_1)E(Y_1) = \mu_X(\alpha + \beta \mu_X)$$

$$\mathrm{Cov}(X_1, Y_1) = \beta\,(EX_1^2 - \mu_X^2) = \beta V(X_1)$$

$$\Rightarrow \hat{\beta} \xrightarrow{a.s.} \beta$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \xrightarrow{a.s.} (\alpha + \beta \mu_X) - \beta \mu_X = \alpha$$

Convergence of Multivariate RVs

$$X_n(\omega) \in \mathbb{R}^J, \quad X_n = \left(X_n^{(1)}, \ldots, X_n^{(J)}\right)$$

$$X_n \xrightarrow{a.s.} X : \quad X_n^{(j)} \xrightarrow{a.s.} X^{(j)}, \quad j = 1, \ldots, J$$

$$X_n \xrightarrow{P} X : \quad P(\|X_n - X\|_2 > \varepsilon) \xrightarrow{n\to\infty} 0 \quad \forall \varepsilon$$

$$P(d(X_n, X) > \varepsilon) \longrightarrow 0$$

$$\|X_n - X\|^2 = \sum_{j=1}^{J} \left(X_n^{(j)} - X^{(j)}\right)^2$$

$$\therefore \quad X_n \xrightarrow{P} X \Rightarrow X_n^{(j)} \xrightarrow{P} X^{(j)} \ \forall j$$

$$\underset{\sim}{x} \in \mathbb{R}^J$$

$$X_n \overset{d}{\Rightarrow} X : \quad F_n(\underset{\sim}{x}) = P\left(X_n^{(1)} \le \underset{\sim}{x}_1, \ldots, X_n^{(J)} \le \underset{\sim}{x}^{(J)}\right)$$

$$F_n(\underset{\sim}{x}) \to F(\underset{\sim}{x}) \quad \forall \underset{\sim}{x} \text{ where } F \text{ is cont.}$$

VDW Thm. 2.7:

iv) $X_n \overset{d}{\Rightarrow} X$ and $d(X_n, Y_n) \xrightarrow{P} 0$ then

$$Y_n \overset{d}{\Rightarrow} X$$

Pf:/ f denotes any bounded, Lipschitz function.

Enough to show that $Ef(Y_n) \to Ef(X)$ $\forall$ such f (portmanteau lemma)

$$\underbrace{|Ef(Y_n) - Ef(X)|}_{t} \le |Ef(Y_n) - Ef(X_n)| + \underbrace{|Ef(X_n) - Ef(X)|}_{\downarrow 0 \text{ by portmanteau } (X_n \xrightarrow{d} X)}$$

$$t_1 \Big[\quad \le E|f(Y_n) - f(X_n)|\,\mathbb{1}(d(X_n, Y_n) \le \varepsilon)$$

$$t_2 \Big[\quad + E|f(Y_n) - f(X_n)|\,\mathbb{1}(d(X_n, Y_n) > \varepsilon)$$

$$t_1 \le \gamma E\, d(X_n, Y_n)\,\mathbb{1}(d(X_n, Y_n) \le \varepsilon) \le \gamma \varepsilon$$

$$t_2 \le \underbrace{2M}_{f \text{ is bounded}} \underbrace{P(\overbrace{d(X_n, Y_n)}^{\to 0} > \varepsilon)}_{E \text{ of indicator is the probability}} \to 0$$

$t \le \gamma \varepsilon$ for sufficiently large n.

$$\therefore |Ef(X_n) - Ef(X)| \xrightarrow[\varepsilon \to 0]{} 0 \qquad \therefore Y_n \xrightarrow{d} X$$

i) $X_n \xrightarrow{a.s.} X$ then $X_n \xRightarrow{d} X$

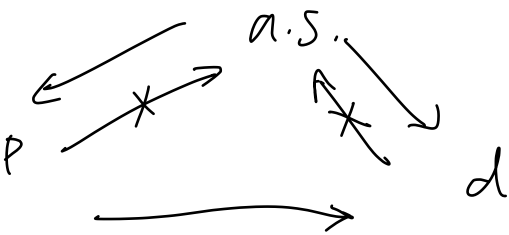
ii) $X_n \xrightarrow{p} X$ then $X_n \xRightarrow{d} X$

Pf of (ii): $Z_n = X$, $Y_n = X_n$

$Z_n \xRightarrow{d} X$, $d(Z_n, Y_n) = d(X, X_n) \xrightarrow{p} 0$

$\therefore Y_n \xRightarrow{d} X$ by part iv

$\therefore X_n \xRightarrow{d} X$



Continuous Mapping Thm:

$X_n \xRightarrow{d} X$, f is continuous, then

$$f(X_n) \xRightarrow{d} f(X)$$

Pf:/ $X_n \xrightarrow{d} X$, then $\exists\ Y_n \xrightarrow{d} X_n$, $Y \xrightarrow{d} X$

s.t. $Y_n \xrightarrow{a.s.} Y$

$$f(Y_n) \xrightarrow{a.s.} f(Y), \quad f(Y_n) \xrightarrow{d} f(Y)$$

$$f(X_n) \xrightarrow{d} f(X)$$

if two RVs have same dist, any function of them will too.

iii) $X_n \xrightarrow{d} c$, c is a constant. Then $X_n \xrightarrow{P} c$.

Pf:/ $\overline{\lim}\, P(|X_n - c| \geq \varepsilon)$

$= \overline{\lim}\, P(\{X_n \geq c+\varepsilon\} \cup \{X_n \leq c-\varepsilon\})$

Z is the degenerate RV at c.

F is a closed set. Then $\overline{\lim}\, P(X_n \in F) \leq P(Z \in F)$

b/c $Z=c$ when $X_n \xrightarrow{d} Z$.

$\leq P(\{Z \geq c+\varepsilon\} \cup \{Z \leq c-\varepsilon\}) = 0$

v) $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$ then $(X_n, Y_n) \xrightarrow{d} (X, c)$

pf:/ f is a bounded, cont. function on $\mathbb{R}^2$.

Enough to show $Ef(X_n, Y_n) \to Ef(X, c)$

$\forall$ such f.

$g(x) = f(x, c)$. g is cont., bounded on $\mathbb{R}$.

$$E f(X_n, c) = E g(X_n) \to E g(X) \quad \text{by Portmanteau}$$

$$= Ef(X, c)$$

$$(X_n, c) \xrightarrow{d} (X, c).$$

$$d((X_n, c), (X_n, Y_n)) = d(Y_n, c) \xrightarrow{P} 0$$

$$\therefore \; X_n, Y_n \xrightarrow{d} X, c$$

Slutsky's Thm: $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$, then

i) $X_n + Y_n \xrightarrow{d} X + c$

ii) $X_n Y_n \xrightarrow{d} cX$

iii) $X_n / Y_n \xrightarrow{d} X/c$ if $c \neq 0$

Pf:/ $(X_n, Y_n) \xrightarrow{d} (X, c)$ (proved above)

i) $g(x,y) = x + y$

ii) $g(x,y) = xy$

iii) $1/Y_n \xrightarrow{P} 1/c$ (HW1),

$(X_n, 1/Y_n) \xrightarrow{d} (X, 1/c)$

$\therefore \dfrac{X_n}{Y_n} \xrightarrow{d} \dfrac{X}{c}$ by part (ii)

Feb 14

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad x_i \overset{iid}{\sim}, \quad E x_i = \mu_x, \quad E x_i^2 < \infty$$

$$\varepsilon_i \overset{iid}{\sim}, \quad E\varepsilon_i = 0, \quad V(\varepsilon_i) = \sigma^2, \quad \varepsilon's \perp x's$$

$$\hat{\beta} \xrightarrow{p} \beta, \quad \hat{\alpha} \xrightarrow{p} \alpha, \quad \hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum (y_i - \hat{y}_i)^2$$

$$= \frac{1}{n}\sum (\alpha + \beta x_i + \varepsilon_i - \hat{\alpha} - \hat{\beta} x_i)^2$$

$$= (\alpha - \hat{\alpha})^2 + (\beta - \hat{\beta})^2 \frac{1}{n}\sum x_i^2 + \frac{1}{n}\sum \varepsilon_i^2$$

0 $\qquad E(\varepsilon_i^2) = \sigma^2$

$$+ 2(\alpha - \hat{\alpha})(\beta - \hat{\beta})\frac{1}{n}\sum x_i \to 0$$

$$+ 2(\alpha - \hat{\alpha})\frac{1}{n}\sum \varepsilon_i \to 0$$

$$+ 2(\beta - \hat{\beta})\frac{1}{n}\sum x_i\varepsilon_i \to 0 \quad \text{b/c}$$

$$\frac{1}{n}\sum x_i\varepsilon_i \xrightarrow{SLLN} E x_i\varepsilon_i$$

$\hat{\sigma}_2 \xrightarrow{p} \sigma^2$, so $\hat{\sigma}^2$ is a consistent estimator.

Portmanteau Thm

iv) $\underline{\lim} \, E f(X_n) \geq E f(X)$ for all cont. $f \geq 0$

v) $\underline{\lim} \, P(X_n \in G) \geq P(X \in G)$ $\forall$ open set G

vi) $\overline{\lim} \, P(X_n \in F) \leq P(X \in F)$ $\forall$ closed sets F

vii) $P(X_n \in B) \to P(X \in B)$ for all Borel sets B with $P(X \in \partial B) = 0$ where $\partial B =$ boundary of B

pf:/
$(i) \Rightarrow (iv)$: $X_n \xRightarrow{d} X$

$\Rightarrow \underline{\lim} \, Ef(X_n) \geq Ef(X)$ $\forall$ cont. $f \geq 0$

To go from dist. to a.s., use Skorokhod's Thm: $Y_n \xRightarrow{d} X$, $Y \xRightarrow{d} X$, $Y_n \xrightarrow{a.s.} Y$

$$f(Y_n) \xrightarrow{a.s.} f(Y) \quad (CMT)$$

$$\begin{array}{ccc} f(Y_n) & & f(Y) \\ \| & & \| \\ Z_n & & Z \end{array}$$

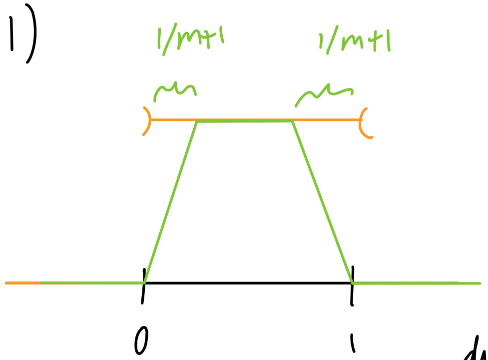
$$\underline{\lim}\, E Z_n \geq E Z = E \lim Z_n \quad \text{Fatou's lemma}$$

$$\underline{\lim}\, E f(X_n) \geq E f(X) \qquad \square$$

(iv) $\Rightarrow$ (v) : $\underline{\lim}\, P(X_n \in G) \geq \underbrace{P(X \in G)}_{E I(X \in G)}$ $\forall$ open G

$f(x) = I(x \in G)$

$G = (0,1)$



$\underbrace{f_m(x)}$ can be any decreasing function of m

$f_m(x) \geq 0$, $f_m(x) \uparrow I(x \in G)$ $\forall$ x

If G is closed, $f(x)$ is always 1 at 0 and 1, but $f_m(x)$ is 0.

f_m's are continuous and Lipschitz for fixed m

By (iv):

$$\underline{\lim}_n Ef(X_n) \geq \underline{\lim}_n f_m(X_n) \geq Ef_m(X)$$

$$\underline{\lim}\, P(X_n \in G) \geq Ef_m(X)$$

$$\downarrow m \to \infty$$

$$Ef(X) \quad \text{by MCT}$$

$$= P(X \in G) \quad \square$$

(v) $\Rightarrow$ (vi) : $\overline{\lim}\, P(X_n \in F) \leq P(X \in F)$, F closed

$F^c = G$, open.

$$P(X_n \in G) \geq P(X \in G)$$

$$\therefore \overline{\lim}\, P(X_n \in F) \leq P(X \in F) \quad \square$$

(vi) $\Rightarrow$ (v): Reverse the above.

Thus (v) $\Leftrightarrow$ (vi)

$\square$

(vi) $\Rightarrow$ (vii) : (vi) implies (v)

B is a set such that $P(X \in \partial B) = 0$,

$\partial B = \bar{B} \setminus B^{\circ}$. $B^{\circ} = (0,1)$, $\bar{B} = [0,1]$, $\partial B = \{0,1\}$

$$P(X \in \bar{B}) = P(\{X \in B^{\circ}\} \cup \{X \in \partial B\})$$

$$= P(\{X \in B^{\circ}\}) + P(\{X \in \partial B\})$$ b/c disjoint

$$= P(X \in B^{\circ})$$

$$\underline{\lim}\, P(X_n \in B) \geq \underbrace{\underline{\lim}\, P(X_n \in B^{\circ}) \geq P(X \in B^{\circ})}_{(v)}$$

$$\overline{\lim}\, P(X_n \in B) \leq \underbrace{\overline{\lim}\, P(X_n \in \bar{B}) \leq P(X \in \bar{B})}_{(vi)}$$

$$\therefore \lim P(X_n \in B) = P(X \in B^{\circ}) = P(X \in B)$$

$\square$

(vii) $\Rightarrow$ (i) : $F_n(X) \to F(X)$, X is a cont. point of F.

$$B = (-\infty, x] . \quad \partial B = \{x\}$$

$$P(X \in \partial B) = P(X = x) = 0$$ b/c F cont. at X (no mass at x)

$$\| \qquad\qquad \|$$

$$F_n(X) \longrightarrow F(X)$$

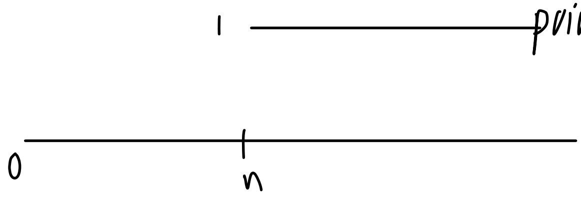
$\square$

Thm 3.2.6 Helly's Selection Thm:/

F_n is a sequence of cdfs. Then there exists a subsequence F_{n_k} such that $F_{n_k}(y) \to F(y)$ for a non-decreasing, right-continuous F and for all y such that F is continuous at y.

Note: F is not necessarily a CDF

e.g. $F_n(X)\ I(X > n)$, X is degenerate w/ point mass at n.



$F(X) = \lim F_n(X) = 0 \quad \forall X$

Notation: $F_{n_k} \overset{v}{\Rightarrow} F$

Pf. (Diagonal subsequence argument):/

$\mathbb{Q} = \{q_1, q_2, \dots\}$

$F_n(q_1) \in [0,1] \Rightarrow \exists$ a subsequence $m_1(i)$ s.t.

$$F_{m_1(i)}(q_1) \xrightarrow{i \to \infty} \text{some limit } g(q_1)$$

$$F_{m_1(i)}(q_2) \xrightarrow{i \to \infty} \text{some limit } g(q_2)$$

$F_{m_1(i)}(q_2) \in [0,1] \Rightarrow$ a subsequence $m_2(i)$

s.t. $F_{m_2(i)}(q_2) \xrightarrow{i \to \infty}$ some limit $= G(q_2)$

$F_{m_1}(1) \quad F_{m_1}(2) \quad F_{m_1}(3) \quad \cdots \to G(q_1)$ at q_1

$F_{m_2}(1) \quad F_{m_2}(2) \quad F_{m_2}(3) \quad \cdots \to G(q_2)$ at q_2

$F_{m_3}(1) \quad F_{m_3}(2) \quad F_{m_3}(3) \quad \cdots \to G(q_3)$ at q_3

entire diagonal is part of the first row $\Rightarrow$ converges at q_1

entire diagonal is part of the j^{th} row except for the first j elements $\Rightarrow$ converges at q_j

$n_k = m_k(k)$

So $F_{n_k}(q) \to G(q) \ \forall q \in \mathbb{Q}$

Still need to show g is non-decreasing, right-continuous, and define it outside of $\mathbb{Q}$. Call it F instead:

$$F(x) = \inf\{G(q) : q \in \mathbb{Q}, q > x\}$$

Some notation:

$\{X_n\}$ is $o(a_n)$ if $\frac{X_n}{a_n} \to 0$

e.g. $X_n = n$, $a_n = n^2$

$X_n = \log n$, $a_n = n$

$\{X_n\}$ is $O(a_n)$ if $\frac{X_n}{a_n}$ is bounded, i.e.

$\exists M > 0$ s.t. $\left|\frac{X_n}{a_n}\right| \leq M \ \forall n$.

e.g. $X_n = an^2 + bn + c$

$$X_n = O(n^2)$$

Probabilistic version:

$\{X_n\}$ is $o_p(Y_n)$ if $\frac{X_n}{Y_n} \xrightarrow{p} 0$

e.g. $Z_i \overset{iid}{\sim}$, $X_n = \sum_i^n Z_i$, $EZ_i = 0$

$$X_n = o_p(n)$$

if $Y_n = 1$, $X_n = o_p(1)$

$\{X_n\}$ is $O_p(1)$ = bounded in probability

$\forall \varepsilon > 0, \exists M(\varepsilon)$ s.t. $P(|X_n| > M(\varepsilon)) \leq \varepsilon$ $\forall n.$

Feb 16

Linear Regression

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

same assumptions as before

$$Y_{new} = \alpha + \beta X_{new} + \varepsilon_{new}$$

Predicted Y_{new}:

$$\hat{Y}_{new} = \hat{\alpha} + \hat{\beta} X_{new}$$

$$\hat{Y}_{new} - E(Y_{new}) = \hat{\alpha} + \hat{\beta} X_{new} - \alpha - \beta X_{new}$$

$$= \underbrace{(\hat{\alpha} - \alpha)}_{o_p(1)} + \underbrace{\underbrace{(\hat{\beta} - \beta)}_{o_p(1)} \underbrace{X_{new}}_{O_p(1)}}_{o_p(1)}$$

$$= o_p(1)$$

$$\xrightarrow{P} 0$$

$$Y_{new} - \hat{Y}_{new} = \underbrace{E(Y_{new}) - \hat{Y}_{new}}_{o_p(1)} + \underbrace{\varepsilon_{new}}_{\text{constant RV, doesn't change w/ } n} \leftarrow \text{residuals}$$

$$\xrightarrow{d} \varepsilon_{new} \qquad \text{Slutsky's}$$

Tightness

A sequence of CDF's F_n is tight if the sequence RVs $X_n \sim F_n$ is $O_p(1)$.

For every $\varepsilon > 0$, $\exists$ $M(\varepsilon)$ s.t.

$$F_n(-M) + 1 - F_n(M) \leq \varepsilon \quad \forall n$$

Thm 3.2.7. : A sequence of CDFs F_n is tight $\Leftrightarrow$ $\forall$ subsequences n_k s.t. F_{n_k} has a limit F, then F is a CDF.

Pf:/ ($\Rightarrow$) F_n is tight. Suppose $F_{n_k} \xrightarrow{v} F$.

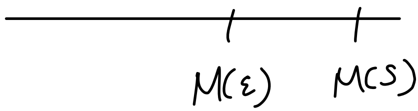
For $\varepsilon > 0$, $\exists$ $M(\varepsilon)$ such that

$$1 - F_n(M(\varepsilon)) + F_n(-M(\varepsilon)) \le \varepsilon \quad \forall n.$$

Pick r, s such that $s > M(\varepsilon)$, $r < -M(\varepsilon)$, r, s, are continuity points of F.

$$1 - F_n(s) + F(r) = 1 - \lim F_{n_k}(s) + \lim F_{n_k}(r)$$

$$= \lim\left(1 - \underbrace{F_{n_k}(s)}_{\text{mass beyond } s} + F_{n_k}(r)\right)$$



b/c we don't know if the limit exists

$$\le \overline{\lim}\left(1 - F_{n_k}(\underbrace{M(\varepsilon)) + F_{n_k}(-M(\varepsilon)}_{\text{might not be continuity points}})\right)$$

$$\le \varepsilon$$

$$\lim_{x \to \infty} 1 - F(x) + F(-x) = 0, \quad \lim_{n \to \infty} F(-x) = 0$$

$$\Rightarrow \lim_{n \to \infty} F(x) = 1,$$ F is a CDF

($\Leftarrow$) Suppose for a contradiction the sequence is not tight. $\exists\, \varepsilon > 0$ s.t. $\forall M \,\exists\, n$ for which $1 - F_n(M) + F_n(-M) > \varepsilon$.

$M = 1$: $1 - F_{n_1}(1) + F_{n_1}(-1) > \varepsilon$

$M = 2$: $1 - F_{n_2}(2) + F_{n_2}(-2) > \varepsilon$

$\vdots$

$n_k > n_{k-1} > \cdots$

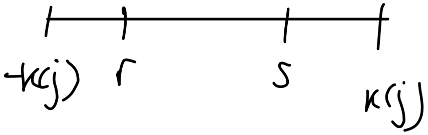
$1 - F_{n_k}(k) + F_{n_k}(-k) > \varepsilon \quad \forall k$

By Helly's thm:

For each n_k, $\exists\, n_{k(j)}$ s.t. $F_{n_{k(j)}} \xrightarrow[j\to\infty]{v} F$,

F is a CDF. Let $r < 0 < s$ be continuity points of F.

$1 - F(s) + F(r) = \lim_j \left(1 - F_{n_{k(j)}}(s) + F_{n_{k(j)}}(r)\right)$



don't know if the limit exists

$$\geq \underline{\lim}_j \left(1 - F_{n_{k(j)}}k(j) + F_{n_{k(j)}}(-k(j))\right)$$

$$\geq \varepsilon$$

$$\lim_{\substack{r \to -\infty \\ s \to \infty}} 1 - F(s) + F(r) = 0$$

$\therefore$ $0 \geq \varepsilon$, which is a contradiction. So, the sequence must be tight.

Characteristic Functions

$$z = a + ib, \quad i^2 = -1, \quad a, b \in \mathbb{R}$$

$$\varphi_X(t) = E(\cos tX + i \sin tX)$$

$$= E\cos tX + iE\sin tX$$

$$\bar{z} = a - ib, \quad |z| = \sqrt{a^2 + b^2} = \sqrt{z\bar{z}}$$

Thm. 3.3.1:

i) $\varphi_X(0) = 1$

Pf:/ $\varphi_X(0) = E\cos 0X + iE\sin 0X$

$= 1$

ii) $\varphi_X(-t) = \overline{\varphi_X(t)}$

Pf:/ $\varphi_X(-t) = E\cos(-t)X + iE\sin(-t)X$

$= E\cos tX - iE\sin tX$

$= \overline{\varphi_X(t)}$

iii) $|\varphi(t)| \leq 1$

Pf:/ $|\varphi(t)| = |E\cos tX + iE\sin tX|$

$= |E(\cos tX + i\sin tX)|$

$\leq E\underbrace{|\cos tX + i\sin tX|}_{\cos^2 + \sin^2 = 1}$ by Jensen's inequality

$= 1$

iv) $|\varphi_X(t+h)-\varphi_X(t)| \le E|e^{ihX}-1|$

Euler's formula: $e^{it}=\cos t + i\sin t$

Pf:/ $|zw| = |z||w|$

$$|\varphi_X(t+h)-\varphi_X(t)| = |Ee^{i(t+h)X} - Ee^{itX}|$$

$$= |Ee^{itX}e^{ihX} - Ee^{itX}|$$

$$\le E|e^{itX}e^{ihX} - e^{itX}| \quad \text{Jensen's}$$

$$= E|e^{itX}||e^{ihX}-1|$$

$$= E|e^{ihX}-1|$$

$\varphi_X(t)$ is uniformly continuous.

$$e^{ihX}-1 = (\cos hX - 1) + i\sin hX$$

$$E|e^{ihX}-1| \xrightarrow[h\to 0]{DCT} 0$$

v) $Ee^{it(aX+b)} = Ee^{itaX} e^{itb}$

$$= e^{itb} Ee^{itaX}$$

$$= e^{itb} \varphi_X(at)$$

ex: X is a random coin flip.

$X = +1$ w.p. $1/2$

$X = -1$ w.p. $1/2$

$$\varphi_X(t) = Ee^{itX} = (e^{it})1/2 + (e^{-it})1/2$$

$$= (\cos t + i\sin t)/2$$

$$+ (\cos t - i\sin t)/2$$

$$= \cos t$$

Feb 21

$$\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$$

$$e_i = \hat{y}_i - y_i = (\hat{\alpha} - \alpha) + (\hat{\beta} - \beta) x_i - \varepsilon_i$$

Properties: $\sum e_i = 0$, $\sum x_i e_i = 0$

$$S_{YE} = \frac{1}{n} \sum_i e_i y_i$$

$$= \frac{1}{n} \sum_i e_i (\alpha + \beta x_i + \varepsilon_i)$$

$$= \frac{1}{n} \sum e_i \varepsilon_i$$

$$= (\alpha - \hat{\alpha}) \frac{1}{n} \sum \varepsilon_i + (\hat{\beta} - \beta) \frac{1}{n} \sum x_i \varepsilon_i - \frac{1}{n} \sum \varepsilon_i^2$$

$$\xrightarrow{P} -\sigma^2$$

ex: coin flip.

$X \sim Bin(n, p)$

$$\varphi_X(t) = E e^{iXt} = \sum \binom{n}{k} p^k (1-p)^{n-k} e^{ikt}$$

$$= \sum \binom{n}{k} (pe^{it})^k (1-p)^{n-k}$$

$$= (pe^{it} + 1 - p)^n$$

binomial expansion

$n=1$, $\varphi_X(t) = pe^{it} + 1 - p$ (Bernoulli)

ex: $X \sim Poi(\lambda)$.

$$\varphi_X(t) = E(e^{itX}) = \sum_k \frac{e^{-\lambda}\lambda^k}{k!} e^{itk}$$

$$= e^{-\lambda} \sum_k \frac{(\lambda e^{it})^k}{k!}$$

$$= e^{-\lambda} e^{\lambda e^{it}}$$

ex: $X \sim N(0,1)$

$$\varphi_X(t) = Ee^{itX} = \int_{-\infty}^{\infty} \cos tx \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx + \underbrace{i\int_{-\infty}^{\infty} \sin tx \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx}_{0}$$

sin is odd

$$= \int_{-\infty}^{\infty} \cos tx \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$

$$\varphi'_X(t) = \int_{-\infty}^{\infty} \frac{-x \sin tx \, e^{-x^2/2}}{\sqrt{2\pi}} dx$$

Note: integration by parts

$$\int u\,dv = uv - \int v\,du$$

$$\varphi_X'(t) = \int_{-\infty}^{\infty} \sin tx \, d\left(\frac{e^{-x^2/2}}{\sqrt{2\pi}}\right)$$

$$= \left[\frac{\sin tx \, e^{-x^2/2}}{\sqrt{2\pi}}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{t\cos tx \, e^{-x^2/2}}{\sqrt{2\pi}} \, dx$$

$$= -t\,\varphi_X(t)$$

$\varphi_X'(t) + t\,\varphi_X(t) = 0$ for standard normal

$$e^{t^2/2}\left(\varphi'(x) + \varphi_X(t)\right) = 0$$

$$\frac{d}{dt}\left(e^{t^2/2}\varphi_X(t)\right) = 0$$

$\Rightarrow e^{t^2/2}\varphi_X(t) = c$ (constant)

$\varphi_X(0) = 1$, $e^0\varphi_X(0) = c$ $\therefore$ $c = 1$

$$\therefore \varphi_X(t) = e^{-t^2/2}$$

ex: $X \sim N(\mu, \sigma^2)$

$$X = \mu + \sigma Z, \quad Z \sim N(0,1)$$

$$\varphi_X(t) = e^{it\mu} \varphi_Z(\sigma t)$$

$$= e^{it\mu - \sigma^2 t^2/2}$$

Thm 3.3.4: Inversion formula

Let $\varphi_X(t)$ be the c.f. of X. Then,

$$\lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{ita} - e^{itb}}{it} \varphi_X(t)\, dt$$

$$= P(a < X < b) + \frac{1}{2} P(X \in \{a, b\})$$

Corollary: c.f. uniquely determines the CDF.

Pf:/ $I(T) = \displaystyle\int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt$

$$g(t) = \frac{e^{-ita} - e^{-itb}}{it} = \int_a^b e^{-itu}\, du$$

$$|g(t)| \le \int_a^b |e^{-itu}|\,du \le (b-a)$$

$$\Rightarrow I(T) = \int_{-T}^{T} g(t) \int e^{itx} P(dx)\,dt, \quad |g(t)| \le b-a$$

$$= \iint_{-T}^{T} g(t)\, e^{itx}\, dt\, P(dx) \qquad \text{Fubini's Thm}$$

$$= \iint_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{it}\, dt\, P(dx)$$

Note: $$\int_{-T}^{T} \frac{e^{it\theta}}{it}\, dt = \underbrace{\int \frac{\cos t\theta}{t}\, dt}_{0} + \int \frac{\sin t\theta}{t}\, dt$$

$$= \iint_{-T}^{T} \frac{\sin t(x-a) - \sin t(x-b)}{t}\, dt\, P(dx)$$

$$= 2\iint_{0}^{T} \frac{\sin t(x-a) - \sin t(x-b)}{t}\, dt\, P(dx)$$

$$I(t) = \int \left(R(x-a,T) - R(x-b,T) \right) P(dx)$$

$$R(\theta,T) = 2\int_0^T \frac{\sin\theta t}{t}\,dt = 2\int_0^{T\theta} \frac{\sin u}{u}\,du$$

$$= 2S(T\theta)$$

$$S(T) = \int_0^T \frac{\sin u}{u}\,du \longrightarrow \frac{\pi}{2} \text{ as } T \to \infty$$

(Dirichlet integral)

$$R(\theta,T) \xrightarrow{T\to\infty} \pi \operatorname{sgn}(\theta)$$

$$\operatorname{sgn}(\theta) = \begin{cases} 1 & \text{if } \theta > 0 \\ -1 & \text{if } \theta < 0 \\ 0 & \text{if } \theta = 0 \end{cases}$$

$$\lim_{T\to\infty} \frac{1}{2\pi} I(T)$$

$$= \lim_{T\to\infty} \frac{1}{2\pi} \int \left(R(x-a,T) - R(x-b,T) \right) P(dx)$$

want to apply DCT

finding a pointwise limit: $\quad a<b$

$$R(x-a,T) - R(x-b,T) \xrightarrow{T\to\infty} 0 \quad \text{if } x>b \text{ or } x<a$$

$$R(x-a,T) - R(x-b,T) \xrightarrow{T\to\infty} \pi \quad \text{if } x=a \text{ or } x=b$$

$$R(x-a,T) - R(x-b,T) \xrightarrow{T\to\infty} 2\pi \quad \text{if } a<x<b$$

$$\underbrace{|R(x-a,T) - R(x-b,T)| \leq 4 \sup_y S(y)}_{\text{by triangle inequality}} < \infty$$

Now we can use DCT:

$$LHS = P(a<X<b) + \underbrace{\frac{1}{2}P(X\in\{a,b\})}_{\text{vanishes if } a,b \text{ are continuity points}}$$

Thm 3.3.2: $X_1 \perp X_2 \Rightarrow \varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\,\varphi_{X_2}(t)$

Pf:/ $\varphi_{X_1+X_2}(t) = E\, e^{it(X_1+X_2)}$

$= E\, e^{itX_1} e^{itX_2}$

$= E\, e^{itX_1}\, E e^{itX_2}$ $X_1 \perp X_2$

$= \varphi_{X_1}(t)\,\varphi_{X_2}(t)$

$X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, $X_1 \perp X_2$

$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\,\varphi_{X_2}(t)$$

$$= e^{it\mu_1 - it\sigma_1^2/2}\, e^{it\mu_2 - it\sigma_2^2/2}$$

$$= e^{it(\mu_1+\mu_2) - t^2(\sigma_1^2+\sigma_2^2)/2}$$

$$= \text{c.f.}\left(N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)\right)$$

$\therefore\ X_1 + X_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

Feb 23

Multiple Linear Regression

$y_i = X_i'\beta + \varepsilon_i$, $X_i \in \mathbb{R}^p$, $X_i \perp \varepsilon_i$

$X_i \overset{iid}{\sim}$, $E(X_i) = \mu \in \mathbb{R}^p$, $E X_i X_i^T = C \in \mathbb{R}^{p \times p}$

$Y = (y_1, \ldots, y_n)'$, $X = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)'$

$$\hat{\beta} = (X'X)^{-1}X'Y = (X'X)^{-1}X'(X\beta + \varepsilon)$$

$$= \beta + \underbrace{(X'X)^{-1}X'\varepsilon}_{\text{bias term}}$$

$$\hat{\sigma}^2 = \frac{1}{n} Y'(I - P_X)Y$$

$$= \frac{1}{n}\varepsilon'(I - P_X)\varepsilon$$

Matrix Norms:

Let M be a matrix $\in \mathbb{R}^{n \times p}$

l_2-norm:

$$\|M\|_2 = \sup_x \left\{ \frac{\|Mx\|_2}{\|x\|_2}, \ x \in \mathbb{R}^p \right\}$$

$$= \sup_{x : \|x\| = 1} \left\{ \sqrt{x' M' M x} \right\}$$

= largest singular value of M

= largest eigenvalue if M is symmetric psd

Multiplicative property:

$$\|Mv\|_2 = \frac{\|Mv\|_2}{\|v\|_2} \|v_2\| \leq \|M\|_2 \|v\|_2$$

Frobenius Norm:

$$\|M\|_F = \left(\sum_{i,j} M_{ij}^2 \right)^{1/2}$$

$$\|M\|_2 \leq \|M\|_F \leq r \|M_2\| \quad (r = \text{rank}(M))$$

Defn: $M_n \xrightarrow{P} M$ if $\| M_n - M \|_2 \xrightarrow{P} 0$

$$\Leftrightarrow \|M_n - M\|_F \xrightarrow{P} 0$$

$$\Leftrightarrow M_{n_{i,j}} \xrightarrow{P} M_{i,j} \quad \forall i,j$$

M is a square matrix:

$\det(M)$ is a continuous function.

$$\text{inv}(M) = M^{-1} = \frac{\text{adj}(M)}{\det(M)}$$ is continuous at any non-singular M (i.e. $\det(M) \neq 0$)

$$\hat{\beta} = \beta + \left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'\varepsilon\right)$$

$$\frac{1}{n}X'X = \frac{1}{n}\sum x_i x_i' \xrightarrow{P} E x_i x_i' = C > 0$$

$$\frac{1}{n}X'\varepsilon = \frac{1}{n}\sum \varepsilon x_i' \xrightarrow{P} \vec{0} \quad \text{by SLLN}$$

$$\|\hat{\beta} - \beta\|_2 \leq \left\|\left(\frac{1}{n}X'X\right)^{-1}\right\|_2 \underbrace{\left\|\left(\frac{1}{n}X'\varepsilon\right)\right\|_2}_{op(1)}$$

Let $M_n = \frac{1}{n}(X'X)^{-1}$, $(M_n^{-1})_{i,j} = \dfrac{\text{adj}(M_n)_{i,j}}{\det(M_n)}$

$$\longrightarrow \frac{\text{adj}(C)_{i,j}}{\det(C)}$$

$$(M_n^{-1})_{i,j} \xrightarrow{P} (C^{-1})_{i,j} \Rightarrow M_n^{-1} \xrightarrow{P} C^{-1}$$

$$\|M_n^{-1}\|_2 \le \|M_n^{-1} - C^{-1}\|_2 + \|C^{-1}\|_2$$

$$= o_p(1) + O(1) = O_p(1)$$

So $\|\hat{\beta} - \beta\|_2$ is $O_p(1) \times o_p(1) = o_p(1)$

and $\hat{\beta}$ is consistent.

$$\hat{\sigma}^2 = \frac{1}{n}\varepsilon'(I - P_X)\varepsilon$$

$$= \underbrace{\frac{1}{n}\varepsilon'\varepsilon}_{\to \sigma^2} - \underbrace{\frac{1}{n}\varepsilon' X}_{o_p(1)} \underbrace{\left(\frac{X'X}{n}\right)^{-1}}_{O_p(1)} \underbrace{\frac{X'\varepsilon}{n}}_{o_p(1)}$$

$$\Rightarrow \hat{\sigma}^2 \xrightarrow{p} \sigma^2$$

Thm 3.3.6: continuity Thm

Let X_n's denote RVs with c.f. φ_n.

i) $X_n \overset{d}{\Rightarrow} X \Rightarrow \varphi_n(t) \to \varphi_X(t) \ \forall t$

ii) $\varphi_n(t) \longrightarrow \varphi(t) \ \forall t$ and φ is cont. at 0

then X_n is tight, $X_n \overset{d}{\Rightarrow} X$. $\varphi_X(t) = \varphi(t) \ \forall t$

Pf:/ i) $\varphi_n(t) = E e^{itX_n}$

$$= E\cos tX_n + iE\sin tX_n$$

$$\to E\cos tX + iE\sin tX = \varphi_X(t)$$

ii) $\frac{1}{u}\int_{-u}^{u}(1-e^{itx})dt = \frac{1}{u}\int_{-u}^{u}(1-\cos tx)dt$

$$= 2 - \frac{2\sin ux}{ux}$$

$$\int \frac{1}{u}\int_{-u}^{u}(1-e^{itx})dt\, \mu_n(dx) = \int\left(2-\frac{2\sin ux}{ux}\right)\mu_n(dx)$$

$$\frac{1}{u}\int_{-u}^{u}(1-\varphi_n(t))dt \geq 2\int_{\{|x|\geq \frac{2}{u}\}}\left(1-\frac{\sin ux}{ux}\right)\mu_n(dx)$$

Note: if $x > 0$,

$$\sin x = \int_0^x \cos u\,du \leq \int_0^x 1\,du = x$$

$|x| > 2/u \Rightarrow |ux| \geq 2$, $\frac{\sin ux}{ux} \leq \frac{1}{2}$

$$2\left(1-\frac{\sin ux}{ux}\right) \geq 1$$

So RHS $\geq P(|X_n| \geq \frac{2}{u})$

$$P(|X_n| \geq 2/u) \leq \frac{1}{u}\int_{-u}^{u}(1-\varphi_n(t))\,dt$$

$$\varphi(t) \xrightarrow{t\to 0} \varphi(0) = \lim_n \varphi_n(0) = 1$$

DCT

$$\leq \frac{1}{u}\int_{-u}^{u}(1-\varphi(t))\,dt + \varepsilon/2 \quad \forall n$$

Choose u such that $|t| \leq u$, $1-\varphi(t) \leq \varepsilon/u$

$$\leq \frac{\varepsilon}{u} \cdot \frac{1}{u} \cdot 2u + \varepsilon/2$$

$$\leq \varepsilon \quad \forall n.$$

$\therefore$ X_n is tight.

Let X_{n_k} be a subsequence.

By Helly's Thm, $\exists$ $X_{n_{k(j)}}$ s.t. $F_{n_{k(j)}} \overset{v}{\Longrightarrow} F$

By tightness, F is a cdf of some RV X.

$$X_{n_{k(j)}} \xRightarrow{d} X$$

$$\therefore \varphi_{n_{k(j)}}(t) \to \varphi_X(t).$$

$$\hookrightarrow \varphi(t)$$

$$\therefore \varphi_X(t) = \varphi(t).$$

f is bounded, cont.

$$Ef(X_{n_{k(j)}}) \to Ef(X)$$

$$\therefore Ef(X_n) \to Ef(X)$$

$$\therefore X_n \xRightarrow{d} X$$

Thm 3.4.1 Central Limit Thm:

$X_1, \dots, X_n \overset{iid}{\sim}$ $EX_i = \mu$, $V(X_i) = \sigma^2$

Then $\dfrac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xRightarrow{d} N(0,1)$.

Ex: $X_i \overset{iid}{\sim} Ber(p)$

$$E(X_i) = p, \quad var(X_i) = p(1-p)$$

$$\frac{\sqrt{n}(\bar{X}-p)}{\sqrt{p(1-p)}} \xrightarrow{d} N(0,1)$$

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$$

$$= \frac{Bin(n,p) - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$$

Ex: $X_i \overset{iid}{\sim} Poi(\lambda)$

$$\frac{\sqrt{n}(\bar{X}-\lambda)}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$$

$$\frac{S_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{d} N(0,1)$$

$$\lambda = 1 \Rightarrow \frac{S_n - n}{\sqrt{n}} = \frac{Poi(n) - n}{\sqrt{n}} \underset{n\to\infty}{\overset{d}{\Rightarrow}} N(0,1)$$

$$\frac{Poi(\lambda) - \lambda}{\sqrt{\lambda}} \underset{\lambda\to\infty}{\Rightarrow} N(0,1)$$

Feb 28

Thm. 3.3.8:

$EX^2 < \infty$, then $\varphi_X(t) = 1 + itEX - \frac{t^2}{2}EX^2 + o(t^2)$

Pf:/ $e^{itX} = \sum_{m=0}^{\infty} \frac{(itX)^m}{m!}$

lemma 3.3.7:

$$\left| e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!} \right| \le \min\left\{ \frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right\}$$

$n=2 \Rightarrow$

$$\left| e^{itx} - \left(1 + itx - \frac{t^2x^2}{2}\right) \right| \le \min\left\{ \frac{|t|^3|x|^3}{6}, t^2x^2 \right\}$$

$$\therefore \left| Ee^{itX} - \left(1 + itEX - \frac{t^2}{2}EX^2\right) \right|$$

$$\le E\left| e^{itX} - \left(1 + itX - \frac{t^2}{2}X^2\right) \right|$$ Jensen's inequality

$$\le t^2 E \min\left\{ \frac{|t||X|^3}{6}, X^2 \right\}$$

$\min\left\{ \frac{|t||X|^3}{6}, X^2 \right\} \xrightarrow{t \to 0} 0$, X^2 is dominator, $EX^2 < \infty$

$$E \min\left\{ \frac{|t||X|^3}{6}, X^2 \right\} \xrightarrow{DCT} 0$$

$\therefore$ RHS $= o(t^2)$

Thm. 3.4.1: Central Limit Theorem:

$X_1, \dots, X_n \overset{iid}{\sim}$, $EX_i = \mu$, $V(X_i) = \sigma^2$

Then $\dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \overset{D}{\Rightarrow} N(0,1)$, $\bar{X} = \dfrac{X_1 + \cdots + X_n}{n}$

Pf:/ $X_i^* = \dfrac{X_i - \mu}{\sigma}$, $EX_i^* = 0$, $V(X_i^*) = 1$, $X_i^* \overset{iid}{\sim}$

$$Z_n = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} = \sqrt{n}\bar{X}^*, \quad \bar{X}^* = \frac{X_1^* + \cdots + X_n^*}{n}$$

$$\varphi_{Z_n}(t) = Ee^{it\sqrt{n}\bar{X}^*} = Ee^{\frac{it}{\sqrt{n}}(X_1^* + \cdots + X_n^*)}$$

$$= \left(\varphi_{X_1^*}\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

$$= \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n$$

$$(1+x)^{1/x} \xrightarrow{x \to 0} e, \quad \left(1+\frac{x}{n}\right)^n = \left[\left(1+\frac{x}{n}\right)^{1/\frac{x}{n}}\right]^x \xrightarrow{n \to \infty} e^x$$

Thm 3.4.2: $c_n \in \mathbb{C}$, $c_n \to c$

$$\Rightarrow \left(1 + \frac{c_n}{n}\right)^n \to e^c$$

$$\left(1 + \frac{x_n}{n}\right)^n \xrightarrow{n \to \infty} e^x, \quad x_n \to x$$

$$\varphi_{Z_n}(t) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n$$

$$= \left(1 + \frac{c_n}{n}\right)^n$$

$$c_n = \frac{-t^2}{2} + o(t^2)$$

$$= \frac{-t^2}{2} + o(1)$$ b/c t is fixed

$$\xrightarrow{n \to \infty} -t^2/2$$

$$\varphi_{Z_n}(t) \to e^{-t^2/2} = \varphi_{N(0,1)}(t)$$

e.g. confidence intervals:

$X_i \overset{iid}{\sim}$, $EX_i = \mu$, $V(X_i) = \sigma^2$

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \overset{d}{\Rightarrow} N(0,1)$$

variance is unknown, so we can estimate it:

$$S_n^2 = \frac{1}{n}\sum (X_i - \bar{X})^2 \to \sigma^2$$

$$\frac{\sqrt{n}(X - \mu)}{S_n} \overset{d}{\Rightarrow} N(0,1)$$ slutsky's thm

$$P\left(\frac{\sqrt{n}(\bar{X} - \mu)}{S_n} \le x\right) \longrightarrow P(N(0,1) \le x) = \Phi(x)$$

$$P\left(-x \le \frac{\sqrt{n}(\bar{X}-\mu)}{S_n} \le x\right) \to \Phi(x) - \Phi(-x) \qquad \alpha = 2\Phi(-x)$$

$$= 1 - 2\Phi(-x) = 1 - \alpha$$

$$\Phi(-x) = \frac{\alpha}{2}, \quad -x = \Phi^{-1}(\alpha/2)$$

$$x = -\Phi^{-1}(\alpha/2)$$

$$= \Phi^{-1}(1-\alpha/2) = z_{1-\alpha/2}$$

$$P\left(-z_{1-\alpha/2} \le \frac{\sqrt{n}(\bar{X}-\mu)}{S_n} \le z_{1-\alpha/2}\right) \to 1-\alpha$$

$$P\left(\bar{X} - z_{1-\alpha/2}\frac{S_n}{\sqrt{n}} \le \mu \le \bar{X} + z_{1-\alpha/2}\frac{S_n}{\sqrt{n}}\right) \to 1-\alpha$$

Poisson approx. to Normal:

$$\frac{Poi(n) - n}{\sqrt{n}} \xRightarrow{d} N(0,1), \quad n \to \infty$$

$$\frac{Poi(\lambda) - \lambda}{\sqrt{\lambda}} \xRightarrow{d} N(0,1), \quad \lambda \to \infty$$

$$[\lambda] \le \lambda < [\lambda] + 1$$

$$z_1 \sim Poi([\lambda]), \quad z_2 \sim Poi(\lambda - [\lambda]), \quad z_3 \sim Poi([\lambda] + 1 - \lambda)$$

$$z_1 \perp z_2 \perp z_3$$

$$\frac{z_1+z_2+z_3-\lambda}{\sqrt{\lambda}} = \left(\underbrace{\frac{poi([\lambda]+1)-([\lambda]+1)}{\sqrt{[\lambda]+1}}}_{\xrightarrow{D} N(0,1)} + \underbrace{\frac{([\lambda]+1-\lambda)}{\sqrt{[\lambda]+1}}}_{o(1)}\right)\underbrace{\frac{\sqrt{[\lambda]+1}}{\sqrt{\lambda}}}_{1}$$

$$\xRightarrow{D} N(0,1)$$

$$z_1 \sim poi([\lambda]), \quad \frac{z_1-\lambda}{\sqrt{\lambda}} \xRightarrow{d} N(0,1)$$

$$P\left(\frac{\overbrace{z_1+z_2+z_3}-\lambda}{\sqrt{\lambda}} \le x\right) \le P\left(\frac{z_1+z_2-\lambda}{\sqrt{\lambda}} \le x\right) \le P\left(\frac{z_1-\lambda}{\sqrt{\lambda}} \le x\right)$$

Multivariate c.f. :

$$X \in \mathbb{R}^d, \ c.f.(X) = \varphi_X(t) = E e^{it'X}, \quad t \in \mathbb{R}^d$$

e.g. $X \sim N(0, I)$, $cf(X) = E e^{it'X}$

$$= E e^{i\sum_{i=1}^{d} t_i X_i}$$

$$= \prod \varphi_{X_i}(t_i) = e^{-\sum t_i^2/2}$$

$$= e^{-t't/2}$$

e.g. $X \sim N(\mu, \Sigma)$, $X = \mu + \Sigma^{1/2} Z$, $Z \sim N(0, I)$

$$\varphi_X(t) = E e^{i(t'\mu + t'\Sigma^{1/2} Z)}$$

$$= e^{it'\mu} \varphi_Z(\Sigma^{1/2} t)$$

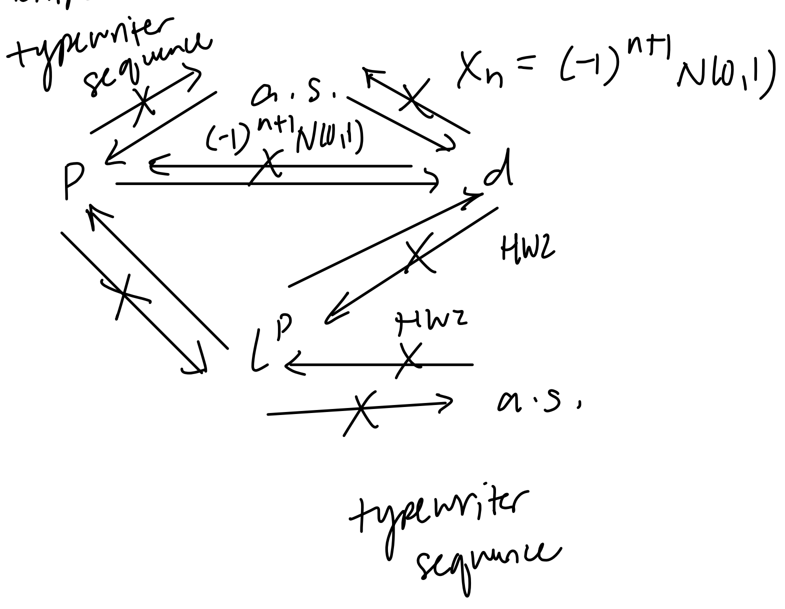
$$= e^{it'\mu + t'\Sigma t/2}$$

Thm. 3.9.6: Multivariate CLT:

$X_i \overset{iid}{\sim}$, $EX_i = \mu$, $V(X_i) = \Sigma$, $X_i \in \mathbb{R}^d$

$$\sqrt{n}(\bar{X} - \mu) \overset{d}{\Rightarrow} N(0, \Sigma)$$

Relationships between modes of convergence:



March 2

Conditional Probabilities

$(\Omega, F_0, P) =$ probability space.

X is an RV (measurable). $E|X| < \infty$ (integrable).

$F \subseteq F_0$ is a σ-algebra.

Y is the conditional expectation of X given F if:

i) Y is F-measurable

ii) $\int_A Y dP = \int_A X dP \quad \forall A \in F$.

Y is denoted $E(Y|F)$

Uniqueness:

Lemma 5.1.1 : $E|E(X|F)| < \infty$

Pf:/ $Y = E(X|F)$, $A = \{Y \geq 0\}$

$A \in F \therefore Y$ is F-measurable

$$\int_A Y\,dP = \int_A X\,dP \le \int_A |X|\,dP$$

$$-\int_{A^c} Y\,dP = -\int_{A^c} X\,dP \le \int_{A^c} |X|\,dP$$

$$E|Y| = \int_\Omega |Y|\,dP = \int_A Y\,dP - \int_{A^c} Y\,dP$$

$$\le \int_\Omega |X|\,dP < \infty$$

Let Y, Y' satisfy i) and ii).

$$A = \{Y - Y' > \varepsilon\} \in F.$$

$$\int_A Y\,dP = \int_A X\,dP = \int_A Y'\,dP$$

lemma 5.1.1

$$0 = \int (Y - Y')\,dP \le \varepsilon P(A)$$

$$\therefore P(A) = 0$$

$Y \le Y'$ a.s. Similarly, $Y' \le Y$ a.s.

So $Y = Y'$ a.s.

Existence:

A measure γ is absolutely continuous wrt a measure μ if $\mu(A) = 0 \Rightarrow \gamma(A) = 0$.

Denoted by $\gamma \ll \mu$.

Radon-Nikodym Thm: γ, μ are σ-finite on (Ω, F) and $\gamma \ll \mu$. Then $\exists f$ (F-measurable) s.t.

$$\gamma(A) = \int_A f \, d\mu.$$

R.N. derivative of γ wrt μ

Assume $X \geq 0$. Let $\gamma(A) = \int_A X \, dP$, $A \in F$.

Clearly, $\gamma(A) \geq 0 \ \forall A$.

$E|X| < \infty \Rightarrow \gamma$ is finite.

A_i's disjoint, $\gamma(\cup A_i) = \sum \gamma(A_i)$

$$\gamma(\cup A_i) = \int_{\cup A_i} X \, dP = \int_{\Omega} X(\sum I(A)) \, dP$$

$$= \sum_i \int_{\Omega} X I(A_i) \, dP$$

$$= \sum_i \gamma(A_i)$$

$$P(A) = 0 \Rightarrow \gamma(A) = 0, \quad \gamma << P$$

By Radon - Nikodym, $\exists\ f$ that is F-measurable

s.t. $\gamma(A) = \int_A f\,dP$, f is F-measurable

$$\therefore f = E(X|F)$$

$$\int_A X\,dP = \int_A X^+ dP - \int_A X^- dP$$

$$= \int_A f_1 dP - \int_A f_2 dP$$

$f_1 = E(X^+|F)$, $f_2 = E(X^-|F)$, f_1, f_2 are

F-measurable.

$$= \int_A (f_1 - f_2) dP$$

$f_1 - f_2$ is F-measurable

$\Rightarrow f_1 - f_2 = E(X|F)$.

Ex:

5.1.1. X is F measurable. $E(X|F) = X$

5.1.2. $X \perp F$. $E(X|F) = E(X)$

Pf:/

$$\int_A E(X)\,dP = EX \int_A dP = EX\,EI_A$$

$$= EXI_A \quad \big\downarrow X \perp F$$

$A \in F$.

$$= \int_A X\,dP$$

5.1.3 conditional Probability

Let $\Omega_1, \Omega_2, \dots$ be disjoint sets. $\cup \Omega_i = \Omega$

and $P(\Omega_i) > 0 \ \forall i$. $S = \{\Omega_1, \Omega_2, \dots\}$

$F = \sigma(S)$. Find $E(X|F)$.

Claim: If $\exists\, Y$ s.t. $\int_{\Omega_i} Y\,dP = \int_{\Omega_i} X\,dP \ \forall i$, then $Y = E(X|F)$.

$$Y(\omega) = \sum_i b_i I(\omega \in \Omega_i)$$

$$\int_{\Omega_i} Y dP = b_i \int_{\Omega_i} dP = b_i P(\Omega_i)$$

$$b_i P(\Omega_i) = E(X I(\Omega_i))$$

$$\therefore b_i = \frac{E(X I(\Omega_i))}{P(\Omega_i)}$$

$X = I(G)$,

$$b_i = \frac{P(G \cap \Omega_i)}{P(\Omega_i)}$$

$$E(I(G)|F)(\omega) = \sum_i \underbrace{\frac{P(G \cap \Omega_i)}{P(\Omega_i)}}_{\text{conditional probability of } G|\Omega_i = P(G|\Omega_i).} I(\omega \in \Omega_i)$$

March 7

Claim: If $\exists\, Y$ s.t. $\int_A X\,dP = \int_A Y\,dP \quad \forall A \in S = \{\Omega_i\}$

disjoint

then $Y = E(X \mid \sigma(S))$

Pf:/ Let $S^* = S \cup \emptyset$. $U, V \in S^*$, then $\underbrace{U \cap V}_{\text{disjoint}} = \emptyset \in S^*$

So S^* is closed under finite intersection (π-system)

λ-system: A collection of sets L is a λ-system if

i) $\Omega \in L$

ii) $A, B \in L$, $A \subseteq B \Rightarrow B \setminus A \in L$

iii) $A_n \uparrow$, $A_n \in L \Rightarrow \lim A_n \in L$

(closed under increasing limits)

π-λ Thm: P is a π-system, L is a λ-system.

$P \subseteq L \Rightarrow \sigma(P) \subseteq L$

S^* is a π-system. Define $\mathcal{L} = \{A \mid \int_A X dP = \int_A Y dP\}$

$$\int_A Y dP = \int_A X dP \quad \forall A \in S$$

So $S^* \subseteq \mathcal{L}$. Now need to show $\mathcal{L}$ is a λ-system.

i) $$\int_\Omega X dP = \int X \sum_{i=1}^{\infty} I(\Omega_i) dP$$

$$= \int X \lim_{n\to\infty} \sum_{i=1}^{n} I(\Omega_i) dP$$

$$= \lim_{n\to\infty} \sum \int_{\Omega_i} X dP$$

by DCT

$$= \sum_{i=1}^{n} \int_{\Omega_i} X dP$$

$$= \sum_i \int_{\Omega_i} Y dP \overset{DCT}{=} \int_\Omega Y dP$$

ii) $A, B \in \mathcal{L}$, $A \subseteq B$

$$\int_{B \setminus A} X dP = \int_B X dP - \int_A X dP$$

Assumes X is integrable, won't have $\infty - \infty$

$$= \int_B Y dP - \int_A Y dP$$

$$= \int_{B \setminus A} Y dP$$

iii) $A_n \uparrow A$, $A_n \in \mathcal{L}$.

$$\int_{A_n} X dP = \int X I(A_n) dP$$

$$\int_A X dP = \lim_n \int_{A_n} X dP$$

$$= \lim_n \int_{A_n} Y dP$$

$$\overset{DCT}{=} \int_A Y dP$$

$$E(E(X|F)) = EX$$

$$\int_A E(X|F) dP = \int_A X dP, \quad \forall A \in F$$

and $\Omega \in F$ so let $A = \Omega$.

Ex: 5.1.4 conditional density

marginal density of y

$(X,Y) \sim$ joint density $f(x,y)$ s.t. $\int f(x,y)\,dx > 0 \ \forall y$.

want to find $E(g(X) \mid \sigma(Y))$, $E|g(X)| < \infty$

$$E(X|Y) = \int x f_{X|Y}(x)\,dx$$

$$E(g(x)|Y) = \int g(x) f_{X|Y}(x)\,dx$$

$$= \frac{\int g(x) f(x,y)\,dx}{\int f(x,y)\,dx}$$

Claim: $h(y) = \dfrac{\int g(x) f(x,y)\,dx}{\int f(x,y)\,dx}$,

$$\int_A h(Y)\,dP = \int_A g(X)\,dP = \quad \forall A \in \sigma(Y)$$

want to show that $h(Y)$ is $\sigma(Y)$ measurable.

$A \in \sigma(Y)$, $A = \{\omega \mid Y(\omega) \in B\}$, $B \in \mathcal{B}(\mathbb{R})$.

$$\int_A h(Y)\,dP = \iint h(y) I(y \in B) f(x,y)\,dx\,dy$$

$$= \iint \frac{\int g(u) f(u,y)\,du}{\int f(u,y)\,du} I(y \in B) f(x,y)\,dx\,dy$$

$$= \int \frac{\int g(u) f(u,y)\,du}{\cancel{\int f(u,y)\,du}} I(y \in B) \cancel{\int f(x,y)\,dx}\,dy$$

$$= \iint g(x) f(x,y) I(y \in B)\,dx\,dy$$

$$= E(g(x) I(y \in B))$$

$$= \int_A g(x)\,dP$$

$$g(x) = I(x \le u)$$

$$E(g(x) \mid Y) = \int_{-\infty}^{u} f(x,y)\,dx \Big/ \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$P(X \le u \mid Y) = \int_{-\infty}^{u} f(x,y)\,dx \Big/ \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$\text{conditional density} = \frac{f(x,y)}{\int f(x,y)\,dx}$$

Properties of Conditional Expectations

Thm. 5.1.2

i) $E(aX + Y \mid F) = aE(X \mid F) + E(Y \mid F)$

Pf:/ $\int_A [aE(X \mid F) + E(Y \mid F)]\,dP$

$= a\int_A E(X \mid F)\,dP + \int_A E(Y \mid F)\,dP$

$= a\int_A X\,dP + \int_A Y\,dP$

$= \int (aX + Y)\,dP$

$\therefore aE(X \mid F) + E(Y \mid F) = E(aX + Y \mid F)$

ii) $X \leq Y \Rightarrow E(X|F) \leq E(Y|F)$

$$\int_A E(X|F)dP + \varepsilon P(A) \leq \int_A X dP \leq \int_A Y dP$$

$$= \int_A E(Y|F)dP \quad \forall A \in F.$$

Let $A = \{\omega \mid \underbrace{E(X|F) \geq E(Y|F) + \varepsilon}_{F\text{-measurable}}\} \in F$

so $P(A) = 0$.

$\varepsilon \to 0 \Rightarrow E(X|F) \leq E(Y|F)$.

iii) $X_n \geq 0$, $X_n \uparrow X$, $EX < \infty$ then

$E(X_n|F) \uparrow E(X|F)$

Pf:/ $E(X_n|F) \geq 0$ by ii)

$E(X_n|F) \uparrow Y$ by ii)

$$\int_A Y\,dP \overset{\uparrow}{\underset{MCT}{}} \int_A E(X_n|F)\,dP = \int_A X_n\,dP \underset{MCT}{\uparrow} \int X\,dP$$

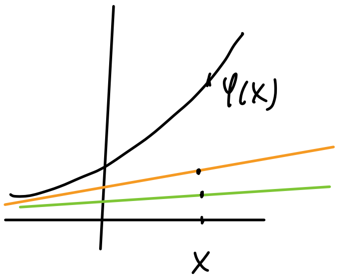
$$\therefore\ Y = \lim E(X_n|F) = E(X|F)$$

Thm. 5.1.3 : Jensen's :

φ is convex, measurable and $E|X|<\infty$, $E|\varphi(X)|<\infty$.

Then $\varphi(E(X|F)) \le E(\varphi(X)|F)$

Pf:/



$$\varphi(x) = \sup_{(a,b)} \{a+b \mid \underbrace{a+bx \le \varphi(x)\ \forall x,\ a,b \in \mathbb{Q}}_{\text{countable}}\}$$

$$\varphi(x) \ge a+bx$$

$$E(\varphi(x)|F) \ge a + bE(X|F)$$

$$\ge \sup\{a+bE(X|F) \mid a,b\in\mathbb{Q},\ a+bx\le\varphi(x)\ \forall x\}$$

$$= \varphi(E(X|F)).$$

March 9

$Y_i \overset{iid}{\sim}$, $E|Y_i| < \infty$, $EY_i = \mu$, N is RV, $N \perp \{Y_i\}$

N is a positive integer

$$E\left(\sum_{i=1}^{N} Y_i\right) = E\left(E\left(\sum_{i=1}^{n} Y_i \mid N=n\right)\right)$$

by independence $= E(N\mu) = \mu E(N)$

$$\underbrace{Y = \{Y_i\}}_{\text{infinite dimensional (entire sequence)}}, N$$

$f(Y,N) = \sum_{i=1}^{N} Y_i$, is this measurable?

(preimage needs to be measurable

$$\{\omega : f(Y,N)(\omega) \in B\}$$

$$= \bigcup_{n=1}^{\infty} \left\{\omega : \sum_{i=1}^{N(\omega) \to n} Y_i(\omega) \in B, N(\omega) = n\right\}$$

$$= \bigcup_{n=1}^{\infty} \left(\overbrace{\left\{\sum_{i=1}^{n} Y_i \in B\right\}}^{\in F} \cap \overbrace{\{N = n\}}^{\in F}\right) \in F$$

So $f(Y,N)$ is measurable. Is it integrable?

$$E\left|\sum_{i=1}^{N} y_i\right| \le E\left(\sum_{i=1}^{N} |y_i|\right)$$

$$\left|E\sum_{i}^{N} y_i\right| \le E\left|\sum_{i}^{N} Y_i\right| \le E\left(\sum_{i}^{N} |Y_i|\right)$$

$$X_n = \left(\sum_{i}^{N} |Y_i|\right) I(N=n) = \left(\sum_{i}^{n} |Y_i|\right) \underbrace{I(N=n)}_{\text{indep.}}$$

$$EX_n = n\, E|Y_1|\, P(N=n)$$

$$f_m = \sum_{n=1}^{m} X_n \uparrow \sum_{n=1}^{\infty} X_n = \sum_{i=1}^{N} |Y_i| \quad \sum_{n=1}^{\infty} I(N=n) = 1$$

$$E\sum_{i=1}^{N} |Y_i| = \lim_{m\to\infty} E f_m = \underbrace{\lim_{m\to\infty} \sum_{n=1}^{m} n\, P(N=n)}_{E(N)} E|Y_1|$$

$$= EN\, E|Y_1| < \infty$$

So the function is integrable.

Then $E\left(\sum_{i=1}^{N} Y_i\right) = EN\, EY_1 = \mu EN$

Cannot use cond. exp. before you know it's integrable.

means σ-algebra generated by N

$$E\left(\sum_{i=1}^{N} Y_i \,\middle|\, N\right) = E\left(\underbrace{\sum_{i=1}^{N} Y_i}_{X} \,\middle|\, \underbrace{\sigma(N)}_{\sigma\{\overset{\Omega_1}{\{N=1\}},\ \overset{\Omega_2}{\{N=2\}},\ \dots\}}\right)$$

$$= \sum_n \frac{E X I(\Omega_n)}{P(\Omega_n)} I(\Omega_n)$$

$$E X I(\Omega_n) = E\left(\sum_{i=1}^{n} Y_i\right) I(N=n) = n\mu P(N=n)$$

$$= \sum_n \frac{n\mu P(N=n)}{P(N=n)} I(\Omega_n)$$

$$= \mu \sum_n n I(N=n)$$

$$= \mu N$$

Example 5.1.5:

$X \perp Y$, φ is measureable s.t. $E|\varphi(X,Y)| < \infty$

Then $E(\varphi(X,Y) \mid X) = g(X)$ where $g(x) = E\varphi(x,Y)$ fixed value

e.g. $g(n) = n\mu$,

$g(N) = N\mu$

Pf:/ $g(x) = \int \varphi(x,y)\,\nu(dy)$ is $\sigma(X)$ measureable (Fubini's thm)

Now just need to show that it's a conditional expectation.

$A \in \sigma(X) \therefore A = \{\omega : X(\omega) \in C\}$

$$\int_A g(X)\,dP = \int_C g(x)\,\mu\,dx$$

C ← Borel set whose preimage is A

$$= \int_C \int \varphi(x,y)\,\nu(dy)\,\mu(dx)$$

$$= \iint \varphi(x,y)\,I(x \in C)\,\nu(dy)\,\mu(dx)$$

$$= E\big(I(X \in C)\,\varphi(X,Y)\big) = E\big(I(A)\,\varphi(X,Y)\big)$$

$$= \int_A \varphi(X,Y)\,dP$$

Thm. 5.1.6:

$F_1 \subseteq F_2$ then i) $E(E(X|F_1)|F_2) = E(X|F_1)$

ii) $E(E(X|F_2)|F_1) = E(X|F_1)$

Pf:/ i) $Y = E(X|F_1)$ is F_1-measureable $\Rightarrow$ F_2-measureable b/c $F_1 \subseteq F_2$

$\therefore E(Y|F_2) = Y$ conditional exp. is itself if it's measureable wrt the σ-algebra

ii) Let $A \in F_1$, $\therefore A \in F_2$, $Y = E(X|F_1)$, $Z = E(X|F_2)$

$$\int_A Y\,dP = \underbrace{\int_A X\,dP = \int_A Z\,dP}_{\text{b/c } A \in F_2}$$

$\therefore Y = E(Z|F_1)$

Thm. 5.1.7: $X \in F$, $E|Y| < \infty$, $E|XY| < \infty$, then $E(XY|F) = XE(Y|F)$

Pf:/ RHS is $\mathcal{F}$-measureable. Let $A \in \mathcal{F}$ and

$X = I(B)$, $B \in \mathcal{F}$

$$\int_A XY\,dP = \int_A I(B)Y\,dP = \int_{\substack{A\cap B \\ \in \mathcal{F}}} Y\,dP = \int_{A\cap B} E(Y|\mathcal{F})\,dP$$

$$= \int_A I(B)\,E(Y|\mathcal{F})\,dP$$

$$= \int_A X\,E(Y|\mathcal{F})\,dP$$

$\therefore$ true for $X = I(B)$

$\therefore$ true for $X = \sum_{i=1}^{k} l_i I(B_i)$ (simple functions)

$\therefore$ true for $X \geq 0$, by DCT

$\therefore$ write $X = X^+ - X^-$ and use the previous result for X^+ and X^-

Thm 5.1.8: suppose $EX^2 < \infty$.

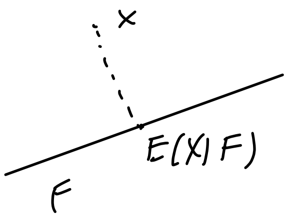
i) Residual from conditional expectation is uncorrelated with all $Z \in L_2(F)$ (Z is F-measureable, $EZ^2 < \infty$)

$Y = X\beta + \varepsilon$ $\qquad$ $Y = f(X) + \varepsilon$

$E(Y|X) = X\beta$ $\qquad$ $E(Y|X) = f(x)$

$\varepsilon = Y - X\beta \perp \sigma(x)$ $\qquad$ $Y - f(x) \perp \sigma(x)$

ii) $E(X|F)$ is the variable Y which is F-measureable and minimizes the "mean square error" $E(X-Y)^2$



Pf:/ i.) $Z \in L_2(F)$ then $E(ZX|F) = ZE(X|F)$

$\therefore E(ZX) = E(E(ZX|F)) = E(ZE(X|F))$

$\therefore E(Z(\underbrace{X - E(X|F)}_{E\text{ residual}=0})) = 0$

ii) $E(X-Y)^2 = E(X - E(X|F) + E(X|F) - Y)^2$

$$= E(X - E(X|F))^2 + E(E(X|F) - Y)^2$$

$$+ 2E\big((E(X|F) - Y)(X - E(X|F))\big)$$

$$= E(X - E(X|F))^2 + EZ^2 + \underbrace{2E(Z(X - E(X|F)))}_{=0}$$

$Z = E(X|F) - Y \in L_2(F)$

Minimized when $Y = E(X|F)$

March 14

Conditional variances

Let $EX^2 < \infty$.

Show that $E(X|F) \in \underbrace{L_2(F)}_{}$

F-measurable + square integrable

Pf:/ F-measurable is obvious.

Only need to show $\int E(X|F)^2 dP < \infty$

$$= E\left(E(X|F)^2\right)$$

$$\leq EE(X^2|F) \quad \text{(Jensen's)}$$

$$= EX^2 < \infty$$

So $E(X|F)$ is square integrable and

$E(X|F) \in L_2(F)$.

Exercise 5.1.8: $G \subseteq F$, $EX^2 < \infty$ then

$$E(X-E(X|G))^2 = E(X-E(X|F))^2 + E\big(E(X|F)-E(X|G)\big)^2$$

Pf:/ $LHS = E\big(X-E(X|F)+E(X|F)-E(X|G)\big)^2$

$$= RHS + 2E\underbrace{(X-E(X|F))}_{\text{residual, uncorrelated w/ } L_2(F).}(\overset{\in L_2(F)}{E(X|F)} - \overset{\in L_2(G) \in L_2(F)}{E(X|G)})$$

$= RHS$

Let $EX^2 < \infty$, define $Var(X|F) = E(X^2|F) - E(X|F)^2$

≥ 0 (Jensen's)

$$var(X|Z) := var(X|\sigma(Z))$$

Exercise 5.1.9:

$$var(X) = var(E(X|F)) + E(var(X|F))$$

Pf:/ $G = \{\phi, \Omega\}$

$$P(X \in B, A) = P(X \in B)\, P(A) \quad \forall\ A \in G$$

Show this to show independence

Can only pick ϕ or Ω for A, and for both choices the identity holds.

Then,

$$E(X - E(X|G))^2 = E(X - E(X|F))^2 + E(E(X|F) - E(X|G))^2$$

$$LHS = E(X - E(X))^2 = var(X)$$

2^{nd} term of RHS $= E(E(X|F) - E(X))^2$

$$= E(E(X|F) - E(E(X|F)))^2$$

$$= E(Y - E(Y))^2, \quad Y = E(X|F)$$

$$= var(Y)$$

$$= var(E(X|F))$$

1st term of RHS: $E(X-E(X|F))^2$

$$= EX^2 - 2E(\underbrace{XE(X|F)}_{Z}) + E(E(X|F)^2)$$

$$= EX^2 - 2E(E(Z|F)) + E(E(X|F))$$

$$E(Z|F) = E(XE(X|F)|F)$$

belongs to F so we can take it out

$$= E(X|F)E(X|F)$$

$$= E(X|F)^2$$

$$= E(E(X^2|F)) - E(E(X|F)^2)$$

$$= E(Var(X|F))$$