Lecture 1/18

Thursday, December 28, 2017 9:20 PM

- Bonryuin (MCS 226)
- All materials on his personal webpage

Topics

- Law of Large Numbers
- General proof techniques.

Class Organization

- HW due every Thursday
- Midterm in March : 15
- Final evaluation: Presentations on current research topics

Topics (Cont)

- Limit theorems
- Error bounds
- Martingales / Iterative Methods

- Martingales / Iterative Methods
- Brownian motion

Laws of Large Numbers

Some intuition:

Suppose we have an experiment with two outcomes $\{S, F\}$ which we repeat several times.

Let X_n be the R.V. s.t.

$$X_n = \begin{cases} 1 & \text{if } S \\ 0 & \text{if } F \end{cases}$$

Q: What is relative freq of Success?

A: $$\frac{1}{n}\sum_{k=1}^{n} X_k$$

Intuitively we expect that

$\frac{1}{n}\sum_{k=1}^{n} X_k$ will stabalize.

$$\frac{1}{n}\sum_{k=1}^{n} X_k \quad \text{will stabalize.}$$

as $n \longmapsto \infty$.

We expect that $\frac{1}{n}\sum_{k=1}^{n} X_k \longmapsto \mu$

when $n \to \infty$.

Q: What does "$\longmapsto$" mean.

As we expect this stabalization to occur with prob 1 then "$\longmapsto$" should be a.s. Convergence.

a.s $\Longrightarrow$ We call this the Strong Law of Large Numbers.

For now we discuss Weak Laws of Large numbers.

This motivating example was just for Bernoulli.

Q: Can we generalize?

A: Yes

Preliminaries

Def: Two sequences X_n, Y_n of R.V. are said to be convergence equivalent when

$$\sum_{n=1}^{\infty} P(X_n \neq Y_n) < \infty$$

Prop: Let $r > 0$ and $\{X_n : n \geq 1\}$ be a sequence of iid RV s.t $E(|X|^r) < \infty$. Set $Y_n = X_n \mathbb{1}_{\{|X_n| \leq n^{1/r}\}}$ $\forall n$, (truncated sequence of Y_n)

then $\{X_n : n \geq 1\}$ and $\{Y_n : n \geq 1\}$ are convergence equivalent.

Pf: We have

$$P(X_n \neq Y_n) = P(X_n - Y_n \neq 0)$$

$$= P(X_n(1 - 1_{\{|X_n| \le n^{1/r}\}}) \neq 0)$$

$$= P(X_n 1_{\{|X_n| > n^{1/r}\}} \neq 0)$$

$$= P(|X_n| > n^{1/r})$$

We claim:

$$\sum P(|X_n| > n^{1/r}) \leq E(|X_1|^r) < \infty$$

Indeed

$$E(|X_1|^r) = \int_0^\infty x \, dF_{|X_1|^r}(x)$$

Typical Trick ✱

$$= \sum_{k=1}^{\infty} \int_{k-1}^{k} x \, dF_{|X_1|^r}(x)$$

$$\geq \sum_{k=1}^{\infty} (k-1) \int_{k-1}^{k} dF_{|X_1|^r}(x)$$

$$= \sum^{\infty} (k-1) P(k-1 < |X_1|^r \leq k)$$

$$= \sum_{k=1}^{\infty} (k-1)\, P(k-1 < |X_1|^r \le k)$$

sum with bad indice

$$= \sum_{k=1}^{\infty} \sum_{n=1}^{k-1} P(k-1 < |X_1|^r \le k)$$

$$= \sum_{n=1}^{\infty} \underbrace{\sum_{k=n+1}^{\infty} P(k-1 < |X_1|^r \le k)}_{P(|X_1|^r > n)}$$

$$= \sum_{n=1}^{\infty} P(|X_1|^r > n)$$

$$= \sum_{n=1}^{\infty} P(|X_1| > n^{1/r})$$

Hence $\displaystyle\sum_{n=1}^{\infty} P(Y_n \ne X_n) = \sum_{n=1}^{\infty} P(|X_1| > n^{1/r})$

$$\le E(|X_1|^r) < \infty .$$

Standard tricks used here.

- $\displaystyle\int_0^\infty x = \sum_{k=0}^{\infty} \int_k^{k+1}$

$n^{h(n)}$ for

- Truncate by $n^{h(n)}$ for which moments the sequence emits.

The Weak Law of Large Numbers.

<u>Thrm</u>: Let $\{X_n : n \geq 1\}$ be an iid seq. with finite mean μ.
Denote $S_n = \sum_{k=1}^{n} X_k$ to be the sequence of partial sums.
Then $\frac{1}{n} S_n \xrightarrow{P} \mu$ as $n \to \infty$.

<u>Pf</u>: Without loss of generality* we can assume $\mu = 0$.
We look to show $\frac{1}{n} S_n \xrightarrow{P} 0$
Let's take $\varepsilon > 0$ and define
$Y_k = X_k \mathbb{1}_{\{|X_k| \leq n\varepsilon^3\}}$ for

$Y_{k,n} = X_k \mathbb{1}_{\{|X_k| \le n\epsilon^3\}}$ for

$k = 1, \dots, n$ and $n \geq 1$.

Aux. Results (Truncate Chebyshev) (Gut 122)

- Let $\{X_n\}$ be iid. and $Y_n = X_n \mathbb{1}_{\{|X_n| \le c\}}$ for $c > 0$. Denote by S_n and S_n'

$S_n = \sum X_k \quad S_n' = \sum Y_k$. Then

$$P(|S_n - E(S_n')| > x) \leq \frac{n \operatorname{Var}(Y_1)}{x^2} + n P(|X_1| > c)$$

Denote by S_n, and S_n'

$\left(S_n' = \sum_{k=1}^{n} Y_{n,k}\right)$. Take $c = n\epsilon^3$, and

$x = n\epsilon$, then by the Truncated C.I.

$$P(|S_n - E(S_n')| > n\epsilon) \leq \frac{n \operatorname{Var}(Y_{1n})}{n^2 \epsilon^2} + n P(|X_1| > n\epsilon^3)$$

$$\leq \frac{1}{n\epsilon^2} E(Y_{1n}^2) + n P(|X_1| > n\epsilon^3)$$

$$= \frac{1}{n\epsilon^2} E\left(X_1^2 \mathbb{1}_{\{|X_1| \le n\epsilon^3\}}\right) + n P(|X_1| > n\epsilon^3)$$

$$\le \frac{n\epsilon^3}{n\epsilon^2} E\left(|X_1| \mathbb{1}_{\{|X_1| < n\epsilon^3\}}\right) + nP(|X_1| > n\epsilon^3)$$

$$\le \epsilon E(|X_1|) + nP(|X_1| > n\epsilon^3)$$

Focusing on the second term.

Observe that

$$E(|X_1|) = \int_0^\infty x \, dF_{|X_1|}(x) < \infty$$

Now, $n\epsilon^3 P(|X_1| > n\epsilon^3) = \int_{n\epsilon^3}^\infty n\epsilon^3 \, dF_{|X_1|}(x)$

$$\le \int_{n\epsilon^3}^\infty x \, dF_{|X_1|}(x)$$ ← Tail of a convergent integral.

So $\limsup P(|S_n - E(S_n')| > n\epsilon) \le \epsilon E(|X_1|)$

This implies

$$S_n - E(S_n') \quad P$$

$$\frac{S_n - E(S_n')}{n} \xrightarrow{P} 0$$

But we know that

$$|E(S_n')| = \left|E\left(\sum_{k=1}^{n} Y_{k,n}\right)\right| = |n E(Y_{1,n})|$$

$$= |n E(X_1 1_{\{|X_1| < n\varepsilon^3\}})|$$

Aux. Result: If $E(X) = 0$ then

then $E(X 1_{\{|X| \le c\}}) = -E(X 1_{\{|X| > c\}})$

$$= |-n E(X_1 1_{\{|X_1| > n\varepsilon^3\}})|$$

$$\le n E(|X_1| 1_{\{|X_1| > n\varepsilon^3\}})$$

So $\left|\frac{E(S_n')}{n}\right| \le E(|X_1| 1_{\{|X_1| > n\varepsilon^3\}}) \xrightarrow{*} 0$

This let's us write

$S_n \quad E(S_n') \qquad S_n \xrightarrow{P} 0$

$$\frac{S_n - E(S_n')}{n} = \frac{S_n}{n} \xrightarrow{P} 0$$

Thus WLLN holds

Lecture 1/23

Tuesday, January 23, 2018 12:24 PM

__Ex:__ Let $\{X_n : n \geq 1\}$ be iid Cauchy $(0,1)$ with density

$$f(z) = \frac{1}{\pi (1+x^2)} \text{ on } \mathbb{R}.$$

We have $\phi_{X_n}(t) = e^{-|t|}$

Note that

$$\phi_{\frac{S_n}{n}}(t) = \mathbb{E}\left(e^{it\frac{1}{n}\sum_{1}^{n} X_i}\right)$$

$$= \prod_{i=1}^{n} \mathbb{E}\left(e^{it/n X_i}\right)$$

$$= \prod_{i=1}^{n} \phi_{X_i}(t/n)$$

$$= \phi_{X_1}(t/n)^n$$

$$= \left[e^{-|t/n|}\right]^n$$

$$= \left[e^{-|t/n|} \right]$$

$$= e^{-|t|} = \phi_{X_1}(t)$$

- This shows that we get no increased precision as $n \to \infty$.
- WLLN does not apply as $\mathbb{E}(X_1) = +\infty$
- This shows finite mean is somewhat necessary.
- The following provides a generalization of the

ge

WLLN when we do not have finite means.

<u>Thrm:</u> (Marichiez-Zygmund WL)

Let $0<r<2$. Suppose $\{X_n : n \geq 1\}$ is a sequence of iid RV s.t. $E(|X_1|^r)<\infty$.

Let S_n be the sequence of partial sums of the X_i.

If $r \geq 0$ WLOG we assume $\mathbb{E}(X_1)=0$. Then $\dfrac{S_n}{n^{1/r}} \xrightarrow{P} 0$.

<u>Pf:</u> Case I: $r=1$ WLLN

modified Proof/

$1<r<2$

Case II: $1 < r < 2$ modified Proof/ HW

Case III: $0 < r < 1$

We will show that the convergence holds in $L^r \Rightarrow P$.

Aux. Results:

The C_r inequality. Suppose $r \geq 0$ and $X, Y \in L^r(\Omega)$.

Then

$$E(|X+Y|^r) \leq C^r \left[E(|X|^r) + E(|Y|^r) \right]$$

for $C^r = \begin{cases} 1 & r \leq 1 \\ 2^{r-1} & r \geq 1 \end{cases}$

for $c_r = \begin{cases} 1 & r \le 1 \\ 2^{r-1} & r \geq 1 \end{cases}$

Take $\varepsilon > 0$ and choose $m > 0$ s.t.

$$\mathbb{E}\left(|X_1|^r \mathbb{1}_{\{|X_1| > m\}}\right) < \epsilon .$$

We can always do this as $X_1 \in L^r(\Omega)$.

Define
$$Y_k = X_k \mathbb{1}_{\{|X_k| \le m\}}$$
$$Z_k = Z_k \mathbb{1}_{\{|X_k| > m\}} .$$

Then of course $X_k = Y_k + Z_k$.

Now

$$\mathbb{E}\left(\left|\frac{S_n}{n^{1/r}}\right|^r\right) = \frac{1}{n}\mathbb{E}\left(|S_n|^r\right)$$

$$= \frac{1}{n}\mathbb{E}\left[\left|\sum Y_k + \sum Z_k\right|^r\right)$$

$$\underset{CR}{\leq} \frac{1}{n}\left(\mathbb{E}\left(\left|\sum Y_k\right|^r\right) + \mathbb{E}\left(\left|\sum Z_k\right|^r\right)\right)$$

$$\underset{M+CR}{\leq} \frac{1}{n}\left[(nM)^r + \sum_k \mathbb{E}\left(|Z_k|^r\right)\right]$$

$$\underset{\text{how we chose } M}{\leq} \frac{1}{n}\left[(nM)^r + n\varepsilon\right]$$

$$= \frac{M^r}{n^{1-r}} + \varepsilon \longrightarrow \varepsilon$$

So $\limsup \mathbb{E}\left(\left|\frac{S_n}{n^{1/r}}\right|^r\right) \leq \epsilon$

n

So

$$\frac{S_n}{n^{1/r}} \xrightarrow{L^r} 0$$

$$\Longrightarrow$$

$$\frac{S_n}{n^{1/r}} \xrightarrow{P} 0$$

$\blacksquare$

Rmk: L^r convergence holds for all $r \in \mathbb{R}^+$.

Ex: Let's take $\{X_n : n \geq 1\}$ be a sequence of iid R.V. with symmetric stable dist. with index $\alpha \in (0,2)$. Then

$\phi_{r.v.}(t) = e^{-c|t|^\alpha}$ for $c > 0$,

$\phi_{X_1}(t) = e^{-c|t|^{\alpha}}$ for $c>0$, $-\infty < t < \infty$. Now let's look at $\frac{S_n}{n^{1/\alpha}}$.

$$\phi_{\frac{S_n}{n^{1/\alpha}}}(t) = \mathbb{E}\left(e^{it\frac{1}{n^{1/\alpha}}\sum X_k}\right)$$

$$= \prod_{j=1}^{n} \mathbb{E}\left(e^{it\,1/n^{1/\alpha}\,X_k}\right)$$

$$= \prod_{j=1}^{n} \phi_{X_1}\left(t/n^{1/\alpha}\right)$$

$$= \left[e^{-c\left|\frac{t}{n^{1/\alpha}}\right|^{\alpha}}\right]^{n} = e^{-c|t|^{\alpha}}$$

$$= \phi_{X_1}(t)$$

Again no increase in percision.

So the MZ-weak does not apply as

$$E(|X_1|^\alpha) = +\infty$$

So for Cauchy$(0,1)$ it would apply for $r < \alpha = 1$.

Since moments exist for $r < \alpha$.

Applications of Weak Laws

1. Empirical Dist. Funct.

1. Empirical Dist. Funct.

Let $\{X_1, \dots, X_n\}$ be iid obs. from unknown F. The EDF is given by

$$F_n(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{X_k \le x\}}$$

We have that

$\{\mathbb{1}_{\{X_k \le x\}} : k \ge 1\}$ is iid with

$$\mathbb{E}(\mathbb{1}_{\{X_1 \le x\}}) = P(X_1 \le x) = F(x).$$

By the WLLN we have that

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{X_k \le x\}} = F_n(x) \xrightarrow{P} F(x)$$

$$\frac{1}{n}\sum_{k=1}^{n} 1(x_k \le x) = F_n(x) \xrightarrow{J} F(x)$$

$\square$

2. Weirstrass Thrm. states that any cont. function on $[0,1]$. can be approximated by a polynomial $p(x)$ with uniform precision.

This is noneconstructive as we have no way to find $p(x)$.

In application, we need a Constructive approach.

Let u be continuous on $[0,1]$.

Let u be continuous on $[0,1]$.
and hence uniformly cont. and
bounded by M.

The approximating polynomial
of deg n is the n-th
Berstein polynomial.

$$u_n(x) = \sum_{k=0}^{n} u\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}$$

Pf of Approx & Uniform
precision:

Let $\{X_n : n \geq 1\}$ be a seq.
of $\mathrm{Bern}(x)$.

Let $Y_n = \frac{1}{n} \sum_{k=1} X_k$

We have

$$\mathbb{E}[u(Y_n)] = \sum_{k=0}^{n} u\left(\frac{k}{n}\right) P\left(Y_n = \frac{k}{n}\right)$$

$$= \sum_{k=0}^{n} u\left(\frac{k}{n}\right) P\left(\sum X_k = k\right)$$

$$= \sum_{k=0}^{n} u\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} = U_n(x).$$

So the connection is the polynomials in terms of a R.V.

We want to prove

$u_n(x) \simeq u(x)$ for large n.

e.g. $E(u(Y_n)) \simeq u(x)$.

The weak law tells us

$$Y_n \xrightarrow{P} E(Y_1) = \frac{1}{n} \sum E(X_k)$$
$$= x$$

So by continuous mapping theorem.

$$u(Y_n) \xrightarrow{P} u(x)$$

By uniform integrability we have

$$\mathbb{E}(u(Y_n)) \xrightarrow{P} u(x)$$

$$\Longleftrightarrow$$

$$u_n(x) \xrightarrow{P} u(x)$$

(Uniform Precision)

$$|u_n(x) - u(x)|$$

u is uniformly cont. $\forall \epsilon > 0$

$\exists \delta > 0$ s.t. $\forall x, y \in [0,1]$ w

$$|x-y| < \delta \Rightarrow |u(x) - u(y)| < \epsilon.$$

$$\left| \sum u(k/n) \binom{n}{k} x^k (1-x)^{n-k} - u(x) \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \right|$$

$$k=0$$

$$\le \sum \left|u\left(\tfrac{k}{n}\right) - u(x)\right| \left| \binom{n}{k} x^k (1-x)^{n-k} \right|$$

$$\le \sum_{k : \left|\frac{k}{n} - n\right| < \delta} \text{''} \; + \sum_{k : \left|\frac{k}{n} - n\right| \geq \delta} \text{''}$$

$$\le \epsilon \times 1 + 2M \, P(|Y_n - x| \geq \delta)$$

$$\le \epsilon + 2M \frac{1/n^2 \sum Var(Y_k)}{\delta^2}$$

$$\le \epsilon + \frac{M}{2n\delta^2}$$

does not depend on x.

Lecture 1/25

Thursday, January 25, 2018 12:31 PM

Convergence of Rand. Series

- It has its own importance and interest
- It will give us the strong law.
- So we will overview the general theorem

Kronecker Lemma: (Deterministic)

Let $\{X_n : n \geq 1\} \subset \mathbb{R}$ and $\{a_n : n \geq 1\}$ be a sequence of positive reals s.t.

of positive reals s.t.

$a_n \nearrow \infty$ then

$$\sum_{k=1}^{\infty} \frac{x_k}{a_k} \text{ conv} \Rightarrow \frac{1}{a_n} \sum_{k=1}^{n} x_k \to 0$$

<u>Lemma</u>: (Random Kronecker)

Let $\{X_n : n \geq 1\}$ be a R.V. and let $\{a_n : n \geq 1\}$ be a seq. of positive reals s.t. $a_n \nearrow \infty$

Then

$$\sum_{k=1}^{\infty} \frac{X_k}{a_k} \text{ Conv. a.s} \Rightarrow \frac{1}{a_n} \sum_{k=1}^{n} X_k \xrightarrow{a.s.} 0$$

Pf: Let

$$A = \left\{ \omega \in \Omega : \sum_{k=1}^{\infty} \frac{X_k(\omega)}{a_k} \text{ conv.} \right\}$$

$$B = \left\{ \omega \in \Omega : \frac{1}{a_n} \sum_{k=1}^{n} X_k(\omega) \to 0 \right\}$$

By assumption $P(A) = 1$.

By the deterministic Kronecker Lemma for $\omega \in A$ we have $\omega \in B$ or $A \subseteq B$. Hence $P(B) = 1$.

Therefor almost all

ω, we see

$$\frac{1}{a_n}\sum_{k=1}^{n} X_k(\omega) \longrightarrow 0$$

or in other words

$$\frac{1}{a_n}\sum_{k=1}^{n} X_k \xrightarrow{a.s.} 0 \qquad \blacksquare$$

<u>Rmk</u>: This tells us that in order to prove

$$\frac{1}{n}\sum X_k = \frac{S_n}{n} \xrightarrow{a.s.} 0$$ it

is enough to show

$$\sum^{\infty} \frac{X_k}{k} \text{ converges a.s.}$$

$$\sum_{k=1} \frac{X_k}{k} \text{ converges a.s.}$$

Kolmogorov Convergence Criteria:

Thrm: Let $\{X_n : n \geq 1\}$ be a seq. of ind R.V. Then

$$\sum_{n=1}^{\infty} Var(X_n) < \infty \Rightarrow \sum_{k=1}^{\infty} [X_k - E(X_k)]$$

converges a.s.

In addition if

$\sum_{k=1}^{\infty} E(X_k)$ converges then

$\sum_{k=1} X_k$ Converges a.s.

<u>Aux. Result</u>: Let $\{X_1, \ldots, X_n\}$ be indep. R.V. with mean zero and $Var(X_k) < \infty$ for all k then for any $x > 0$

$$P\left(\max_{1 \le k \le n} |S_k| > x\right) \le \frac{\sum_{k=1}^{n} Var(X_k)}{x^2}$$

<u>Pf</u>: We will prove the sequence $\{S_n - E(S_n) : n \geq 1\}$ is a.s. Cauchy under the assumption $\sum_{1}^{\infty} Var(X_k) < \infty$.

the assumption $\sum_{k=1}^{\infty} Var(X_k) < \infty$.

Let $\varepsilon > 0$ and $n, m \in \mathbb{N}$ s.t. $n < m$. Then

$$P\left(\max_{n \le k \le m} \left| (S_k - E(S_k)) - (S_n - E(S_n)) \right| > \varepsilon \right)$$

$$= P\left(\max_{n \le k \le m} \left| \sum_{j=n+1}^{k} (X_j - E(X_j)) \right| > \varepsilon \right)$$

Kolmogorov Inequality

$$\le \frac{\sum_{j=n+1}^{m} Var(X_j)}{\varepsilon^2} \le \frac{\sum_{j=n+1}^{\infty} Var(X_j)}{\varepsilon^2}$$

Letting $m \to \infty$

$$P\left(\sup_{k \ge n} \left| (S_k - E(S_k)) - (S_n - E(S_n)) \right| > \epsilon \right)$$

$$\le \frac{\sum_{i=n+1}^{\infty} Var(X_j)}{\varepsilon^2} \longrightarrow 0 \text{ as } n \to \infty.$$

This shows the sequence $\{S_n - \mathbb{E}(S_n) ; n \geq 1\}$ is a.s. Cauchy. $\blacksquare$

Under the additional assumption that $\{X_n : n \geq 1\}$ is uniformly bounded then the converse holds.

Thrm: Let $\{X_n : n \geq 1\}$ be a sequence

Thrm: Let $\{X_n : n \geq 1\}$ be a sequence of uniformly bounded indep. random variables then

$$\sum \mathrm{Var}(X_j) < \infty \iff \sum (X_j - \mathbb{E}(X_j)) \text{ conv a.s.}$$

If in addition we have a uniform bound $\sup_n |X_n| \leq A$, $A > 0$ then

$$P\left(\max_{1 \leq k \leq n} |S_k| > x\right) \geq 1 - \frac{(x+A)^2}{\sum_{k=1}^{n} \mathrm{Var}(X_k)}$$

Pf: $(\Leftarrow)$ We will prove

the contrapositive.

Assume $\sum_{k=1}^{\infty} Var(X_k) = \infty$

We have that $\sup_n |X_n| \leq A$

with $A > 0$.

Then $\sup_n |X_n - E(X_n)| \leq 2A$ a.s.

So by the Kolmogorov ineq.

with $n, m \in \mathbb{N}$, $\varepsilon > 0$

$$P\left(\max_{n \leq k \leq m} \left| (S_k - E(S_k)) - (S_n - E(S_n)) \right| > \varepsilon \right)$$

$$\geq 1 - \frac{(\varepsilon + 2A)^2}{\sum_{k=n+1}^{m} Var(X_k)}.$$

$k=n+1$

Let $m \to \infty$ we arrive at:

$$P\left(\sup_{k\geq n} \left|(S_k - E(S_k)) - (S_n - E(S_n))\right| > \varepsilon\right) \geq 1.$$

This shows

$\{S_n - E(S_n) / n \geq 1\}$ does not

converge a.s. $\blacksquare$

Ex: Let $\{X_n : n \geq 1\}$ be a

sequence of iid symmetric

Then

Bernoullis. Then

$E(X_1) = 0$ and

$Var(X_1) = E(X_1^2) = 1 \cdot P(X_1^2 = 1) = 1.$

Consider the series

$$\sum_{k=1}^{\infty} \frac{X_k}{k} = \text{"} \sum_{k=1}^{\infty} \pm \frac{1}{k} \text{"}$$

We know $\sum \frac{1}{k} = \infty$

and $\sum \frac{(-1)^k}{k}$ conv.

What about the random series?

We have

$$\sum_{k=1}^{\infty} \mathrm{Var}\left(\frac{X_k}{k}\right) = \sum_{k=1}^{\infty} \frac{1}{k^2} \mathrm{Var}(X_k)$$

$$= \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$$

By the Kolmogorov Criterion

$\sum \frac{X_k}{k}$ Converges a.s.

What happens when we change p?

$P(X_1 = 1) = p \quad P(X_1 = -1) = 1-p$

$$\mathbb{E}(X_1) = 2p-1$$

$$\mathrm{Var}(X_1) = \mathbb{E}(X_1^2) - (2p-1)^2$$

$$= 1 - (2p-1)^2 < \infty$$

So we still have summable variances and by the K-criterion.

$$\sum_{k=1} \left(\frac{X_k}{k} - \mathbb{E}\left(\frac{X_k}{k}\right)\right)$$ Converges a.s.

So for $\sum \frac{X_k}{k}$ to converge we need to consider

$$\sum_{k=1}^{\infty} \mathbb{E}\left(\frac{X_k}{k}\right) = (2p-1)\sum_{k=1}^{\infty}\frac{1}{k} = \infty$$

for $p \neq \frac{1}{2}$.

which then shows that

$\sum_{k=1}^{\infty} \frac{X_k}{k}$ diverges a.s.

Possible issues

- Variances known?
- Uniformly Bounded?
- Characterization?

The Kolmogorov 3-series Thrm:

Thrm: Let $A > 0$. We assume $\{X_n : n \geq 1\}$ is a seq. of ind. R.V. For any $k \geq 1$ define

$$Y_k = X_k \mathbb{1}_{|X_k| < A}$$

then

$$\sum_{k=1}^{\infty} X_k \text{ converges a.s. iff}$$

(i) $\sum_{n=1}^{\infty} P(X_n \neq Y_n) = \sum_{n=1}^{\infty} P(|X_n| > A) < \infty$

(ii) $\sum_{k=1}^{\infty} \mathbb{E}(Y_k)$ conv.

(iii) $\sum^{\infty} \mathrm{Var}(Y_n) < \infty$.

(iii) $\sum_{k=1}^{\infty} Var(Y_k) < \infty$.

Lecture 1/30

Tuesday, January 30, 2018 12:27 PM

Recall the Kolmogorov Conv.

Thrm: Let $\{X_n, n \geq 1\}$ be a sequence of ind. R.V.

Then

$$\sum \mathrm{Var}(X_i) < \infty \implies \sum X_i - E(X) \text{ Conv. a.s.}$$

If additionally $\{X_n, n \geq 1\}$ is uniformly bounded

$$\sum \mathrm{Var}(X_i) < \infty \iff \sum X_n - \mathbb{E}(X_n) \text{ conv. a.s.}$$

Thrm: (3 Series) Let $A > 0$ and $\{X_n\}$ be a seq. of ind R.V. Set $Y_n = 1_{|X_n| < A}$

ind R.V. Set $Y_n = \mathbb{1}_{|X_n| \le A}$

than

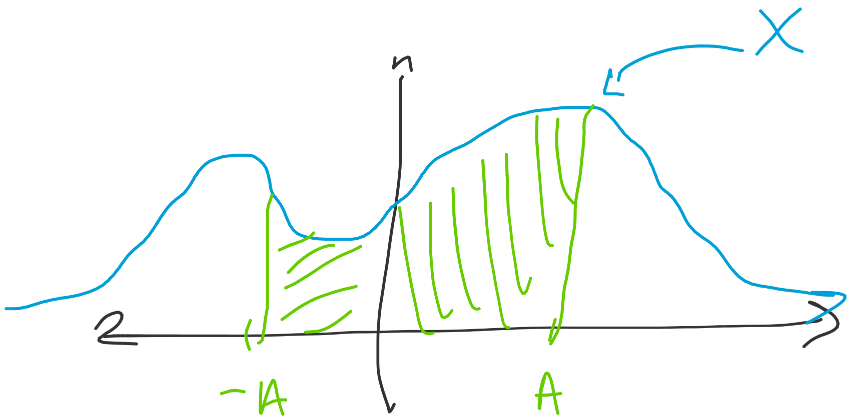
$\sum X_n$ conv. a.s. $\boxed{\text{iff}}$

(i) $\sum P(X_n \neq Y_n) = \sum P(|X_n| > A) < \infty$

(ii) $\sum \mathbb{E}(Y_n)$ conv.

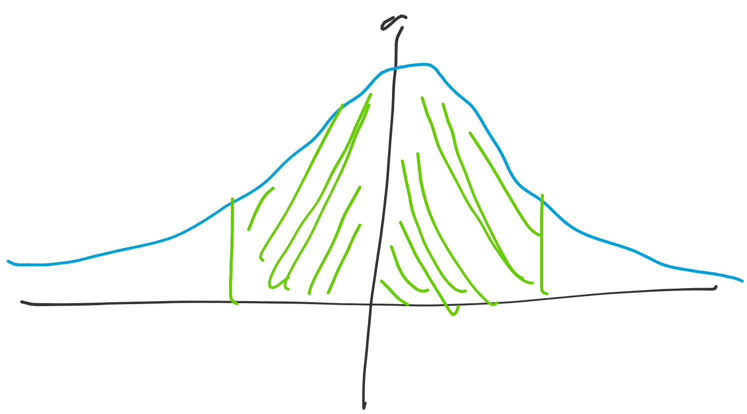
(iii) $\sum \operatorname{Var}(Y_n) < \infty$

Rmk: (i) and (iii) usually pretty easy to check. (ii) could be difficult. b/c even if $\mathbb{E}(X_n) = 0$, $\mathbb{E}(Y_n)$ are not necessarily centered.



may not be symmetric

If X symmetric however.



Pf: Recall that the symmetrization

X^s of a RV X is given by

$X^s = X - X'$ where $X' \stackrel{D}{=} X$

and $X \perp\!\!\!\perp X'$.

Hence $\mathbb{E}(X^s) = 0$ and

$\mathrm{Var}(X^s) = 2\,\mathrm{Var}(X)$

<u>Case I</u>: Assume that $\{X_n, n \geq 1\}$ is symmetric and uniformly bounded by $M > 0$. By choosing $A \geq M$, $X_n = Y_n \ \forall n \geq 1$.

As X_n are symmetric so are the Y_n.

Observe that

$$\sum P(|X_n|>A) = 0$$

$$\sum \mathbb{E}(Y_n) = 0$$

By the Bounded Kolmogor Criterion

$\sum Var(X_n) < \infty$ iff $\sum X_n$ convg. a.s.

<u>Case II</u>: Assume now that the $\{X_n\}$ are symmetric but not necessarily bounded.

Assume that the three series converge ($\Leftarrow$).

Observe that

$$\sum \mathbb{E}(Y_n) = 0.$$

(i) Tells us that X_n and Y_n are convergence equivalence.

(iii) Tells us by the Kolmogorov Convergence Criterion that $\sum Y_n$ conv. a.s.

Hence by convergence equivalence $\sum X_n$ conv. a.s.

($\Rightarrow$) We assume $\sum X_n$ conv. a.s. So $X_n \xrightarrow{a.s.} 0$

Then $\mathbb{P}(|X_n| > A \text{ i.o.}) = 0$.

So by Borel Cantelli we have

$$\sum P(|X_n| > A) < \infty.$$

which is (i). From this we deduce that X_n and Y_n are convergence equivalent. Hence

$\sum Y_n$ conv. a.s.

So by the Kolmogorov

Conv. Criterian

$\sum Var(Y_n) < \infty$.

<u>Case III</u>: (general case).

Assume that the

three series converge.

From (i) X_n and Y_n

are convergence equivalent.

From (iii) and the

Kolmogorov Convergence thrm,

$\sum [Y_n - \mathbb{E}(Y_n)]$ conv. a.s.

by from (ii) $\sum \mathbb{E}(Y_n)$

conv. a.s. So $\sum Y_n$ conv.

a.s. So by CE $\sum X_n$

conv. a.s.

($\Longrightarrow$) Assume $\sum X_n$ conv. a.s

By B.C. argument as above

$\sum P(|X_n| > A) < \infty$ and hence

X_n, Y_n are C.E.

By def. of symmetrization.

$\sum X_n^s$ conv. a.s. By

the same BC arguement.

X_n^s and Y_n^s are C.E.

So $\sum Y_n^s$ conv. a.s.

$\iff \sum Var(Y_n^s) < \infty$.

(

Bounded

Kolmogorov

Then $\sum Var(Y_n) = \frac{1}{2} \sum Var(Y_n^s) < \infty$

which is condition (iii).

which is condition (iii).

Finally, we have

$$\mathbb{E}(Y_n) = Y_n - (Y_n - \mathbb{E}(Y_n)) \quad (*)$$

We already know $\sum Y_n$ conv. a.s. But we just showed

$$\sum \mathrm{Var}(Y_n) < \infty \implies \underbrace{\sum (Y_n - \mathbb{E}(Y_n))}_{\text{conv. a.s}}$$

So

$$\sum \mathbb{E}(Y_n) = \underbrace{\sum Y_n}_{\text{conv. a.s.}} - \underbrace{\sum (Y_n - \mathbb{E}(Y_n))}_{\text{conv. a.s.}}$$

So by (*) $\sum \mathbb{E}(Y_n)$ conv. a.s.

$\blacksquare$

Rmk: We only used the assumption $\{X_n, n \geq 1\}$ are independent.

Other Results of Rand. Series

Thrm: (Levy's Thrm): Let X_n be a sequence of ind R.V.

Then

$\sum X_n$ conv. in dist

$\sum X_n$ conv. in dis.

$\Longleftrightarrow$

$\sum X_n$ conv in prob.

$\Longleftrightarrow$

$\sum X_n$ conv a.s.

Pf: HW #2.

The Strong Law of L.N.

We start with a preliminary version.

Thrm: Let X_n be a seq. of ind. R.V. with mean zero and finite mean zero and finite variance $Var(X_n) = \sigma_n^2$.

Then if

$$\sum \frac{\sigma_n^2}{n^2} < \infty \Rightarrow \frac{S_n}{n} \xrightarrow{a.s.} 0.$$

<u>Pf:</u> We have

$$\sum \frac{\sigma^2}{n^2} = \sum Var\left(\frac{X_n}{n}\right) < \infty$$

by Kol. Con. Criterion.

$\sum \frac{X_n}{n}$ conv. a.s.

which by Random Kronecker

Lemma

$$\frac{1}{n}\sum_{k=1}^{n} X_k = \frac{S_n}{n} \xrightarrow{a.s.} 0.$$

For the general strong law, need the following lemma.

<u>Lemma</u>: Let X_n be iid.

For $0<r<2$ let $Y_n = X_n \mathbb{1}_{|X_n| \le n^{1/r}}$

then if

$$\mathbb{E}\left[|X_1|^r\right] < \infty \Longrightarrow \sum \operatorname{Var}\left(\frac{Y_n}{n^{1/r}}\right) < \infty$$

Lecture 2/1

Thursday, February 1, 2018 12:26 PM

3 Series Thrm:

Recall at one point we said that

$\sum X_n^s$ conv. a.s $\Longrightarrow \sum Y_n^s$ cvg a.s.

$\hookrightarrow Y_n^s = X_n^s \mathbb{1}\{|X_n^s| \le A\}$

We used at one point

$$Var(Y_n) = \frac{1}{2} Var(Y_n^s)$$

where

$$Y_n^s = Y_n - Y_n' = X_n \mathbb{1}\{|X_n| \le A\}$$

$- X_n 1_{|X_k| \le A}$

So we need to be more careful for Y_n^s.

A more direct approach.

$\sum X_n$ conv. a.s.

$\longrightarrow \sum Y_n$ conv. a.s.

$\longrightarrow \sum Y_n^s$ conv. a.s.

for $Y_n^s = X_n - Y_n'$

not $Y_n^s = X_n^s \mathbb{1}_{\{|X_n| < A}$

Then we can apply.

$$Var(Y_n) = \frac{1}{2} Var(Y_n^s)$$

Strong Law Preliminaries

Lemma: Let $0<r<2$ let $\{X_n, n\geq 1\}$ be iid R.V. Then writting

$$Y_n = X_n \mathbb{1}_{|X_n| < n^{1/r}}$$

Now, if $\mathbb{E}(|X_1|^r) < \infty$

we have $\sum_{n=1}^{\infty} Var\left(\frac{Y_n}{n^{1/r}}\right) < \infty$

Pf: $\sum_{n=1}^{\infty} Var\left(\frac{Y_n}{n^{1/r}}\right)$

$$\leq \sum_{n=1}^{\infty} \mathbb{E}\left[\left(\frac{Y_n}{n^{1/r}}\right)^2\right] = \sum_{n=1}^{\infty} \frac{1}{n^{2/r}} \mathbb{E}(Y_n^2)$$

$$= \sum_{n=1}^{\infty} \frac{1}{n^{2/r}} \mathbb{E}\left(X_n^2 \mathbb{1}_{|X_n| \leq n^{1/r}}\right)$$

$$= \sum_{n=1}^{\infty} \mathbb{E}\left[\frac{1}{n^{2/r}} X_n^2 \mathbb{1}_{|X_n| \leq n^{1/r}}\right]$$

as everything in here is positive by the

MCT we have

$$= \mathbb{E}\left[X_1^2 \sum_{n=1}^{\infty} \frac{\mathbb{1}_{|X_n| \le n^{1/r}}}{n^{2/r}} \right]$$

$$= \mathbb{E}\left[X_1^2 \sum_{n=|X_1|^r \vee 1} {}^{1}/_{n^{2/r}} \right]$$

start sum when $n \geq |X_1|^r$

$$= \mathbb{E}\left[X_1^2 \mathbb{1}_{\{|X_1| < 1\}} \sum_{n=1}^{\infty} {}^{1}/_{n^{2/r}} \right]$$

$$+ \mathbb{E}\left[X_1^2 \mathbb{1}_{1 \le |X_1| < 2^{1/r}} \sum_{n=|X_1|^r}^{\infty} {}^{1}/_{n^{2/r}} \right]$$

$$+ \mathbb{E}\left[X_1^2 \mathbb{1}_{|X_1| > 2^{1/r}} \sum_{n=|X_1|^r}^{\infty} {}^{1}/_{n^{2/r}} \right]$$

$$\leq \sum_{n=1}^{\infty} \frac{1}{n^{2/r}} + 2^{3/r} \sum_{n=1}^{\infty} \frac{1}{n^{2/r}}$$

$+$ see below

Calculus Trick $\forall \alpha > 0$

$$\sum_{n=k}^{\infty} \frac{1}{n^{\alpha+1}} = \sum_{n=k}^{\infty} \int_{n-1}^{n} \frac{1}{n^{\alpha+1}} dx$$

here $x < n$ so

$$\leq \sum_{n=k}^{\infty} \int_{n-1}^{n} \frac{1}{x^{\alpha+1}}$$

$$= \frac{1}{\alpha(k-1)^{\alpha}} = \frac{1}{\alpha k^{\alpha}} \cdot \frac{k^{\alpha}}{(k-1)^{\alpha}} \leq \frac{2^{\alpha}}{\alpha k^{\alpha}}$$

So taking $k = |X_1|^r$

$$\alpha + 1 = \frac{2}{r}$$

$$\alpha = \frac{2}{r} - 1$$

$$\leq \sum_{n=1}^{\infty} \frac{1}{n^{2/r}} + 2^{2/r} \sum_{n=1}^{\infty} \frac{1}{n^{2/r}} r$$

$$\mathbb{E}\left[X_1^2 \mathbb{1}_{|X_1| \geq 2^{1/r}} \frac{2^{2/r - 1}}{\left(\frac{2}{r} - 1\right)\left(|X_1|^r\right)^{\frac{2}{r} - 1}} \right]$$

$$\leq C_1 + C_2 + \frac{2^{2/r - 1}}{\left(\frac{2}{r} - 1\right)} \mathbb{E}\left[|X_1|^r\right] < \infty$$

Hence $\displaystyle\sum_{n=1}^{\infty} \operatorname{Var}\left(\frac{X_n}{n^{1/r}}\right) < \infty$. ∎

Strong Law of Large Numbers

Thm: (Kolmogorov Strong Law).

Let $\{X_n, n \geq 1\}$ be iid R.V.

(i) If $\mathbb{E}[|X_1|] < \infty$ with

$\mathbb{E}[X_1] = \mu$ then

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

(ii) If $\frac{S_n}{n} \xrightarrow{a.s.} c \in \mathbb{R}$ then

$\mathbb{E}[|X_1|] < \infty$ and $\mathbb{E}(X_1) = c$.

(iii) If $\mathbb{E}[|X_1|] = \infty$ then

$$\overline{\lim} \frac{|S_n|}{n} = \infty.$$

Pf: (i) Let $Y_n = X_n 1_{\{|X_n| < n\}}$

Then X_n and Y_n are conv. equivalent. As $\mathbb{E}[|X_1|] < \infty$ by the lemma

$$\sum Var\left(\frac{Y_n}{n}\right) < \infty$$

Now by the Kolmogorov Criterion.

$$\sum \left(\frac{Y_n}{n} - \frac{\mathbb{E}[Y_n]}{n}\right) \text{ cvg a.s.}$$

By the Kronecker Lemma

By the Kronecker Lemma

$$\frac{1}{n}\sum[Y_k - \mathbb{E}(Y_k)] \xrightarrow{a.s.} 0$$

So

$$\left(\frac{1}{n}\sum(Y_k - \mu) + \frac{1}{n}\sum[\mu - \mathbb{E}(Y_k)]\right) \xrightarrow{a.s.} 0$$

It suffices to show that

$$\frac{1}{n}\sum[\mu - \mathbb{E}(Y_k)] \longrightarrow 0$$

Well

$$\mathbb{E}[Y_k] = \mathbb{E}[X_n \mathbb{1}_{|X_n| \le n}] \longrightarrow \mathbb{E}(X_n) = \mu$$

So

$$a_k = \mu - \mathbb{E}(Y_k) \longrightarrow 0$$

$\Longleftrightarrow$

$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N}$ s.t $\forall n \geq n_0, |a_n| < \varepsilon.$

So

$$\left| \frac{1}{n} \sum_{k=1}^{n} a_k \right|$$

$$\leq \underbrace{\frac{1}{n} \sum_{k=1}^{n_0} |a_n|}_{\to 0} + \frac{1}{n} \sum_{k=n_0+1}^{n} |a_k|$$

$$< \frac{\quad\quad}{\quad\quad}\!\!/\!\!/ + \underbrace{\frac{n-n_0}{n}}_{\to 1} \cdot \underbrace{\frac{1}{n-n_0} \sum_{k=n_0+1}^{n} \varepsilon}_{\to \varepsilon}$$

Hence

$$\limsup \left(\left| \frac{1}{n} \sum_{k=1}^{n} a_k \right| \right) \le \epsilon$$

$$\Longrightarrow$$

Thus $\frac{1}{n} \sum \left(X_k - \mathbb{E}(X_k) \right) \xrightarrow{a.s.} 0$

(ii) Assume $\frac{S_n}{n} \xrightarrow{a.s.} c$. As

$$\frac{X_n}{n} = \frac{S_n - S_{n-1}}{n} = \frac{S_n}{n} - \frac{n-1}{n} \frac{S_{n-1}}{n-1}$$

$\xrightarrow{a.s.} c - 1 \times c = 0$. So by HW#1

$\mathbb{E}[|X_1|] < \infty$.

Now by part (i)

$$\frac{S_n}{n} \xrightarrow{a.s.} \mathbb{E}(X_1)$$

But by uniqueness of limits

$\mathbb{E}(X_1) = \mu = c$.

(ii) Assume $\mathbb{E}[|X_1|] = \infty$.

In HW we showed

that

$$\mathbb{E}[|X_1|] < \infty \iff \sum_{n=1}^{\infty} P(|X_n| > n\varepsilon) < \infty$$

So with our assumption

$$\sum P(|X_n| > n\varepsilon) = +\infty$$

and by BC II

$$P(|X_n| > n\varepsilon, \text{ i.o.}) = 1.$$

Now,

$$|X_n| = |S_n - S_{n-1}| \leq |S_n| + |S_{n-1}|$$

$$= \begin{cases} 2|S_n| & \text{if } |S_n| \geq |S_{n-1}| \\ 2|S_{n-1}| & \text{if } |S_{n-1}| \geq |S_n| \end{cases}$$

So either

$$P(2|S_n| > n\varepsilon \text{ i.o.}) = 1$$

or

$$P(2|S_{n-1}| > n\varepsilon \text{ i.o.})$$

$$\Longrightarrow$$

$$P(2|S_{n-1}| > (n-1)\varepsilon)$$

So $\forall$ $n \geq 1$

$$P\left(\frac{|S_n|}{n} > \frac{\varepsilon}{2} \text{ i.o.}\right) = 1$$

$$\iff$$

$$\limsup \frac{|S_n|}{n} = +\infty$$

$\underline{\text{Thm:}}$ Let $0 < r < 2$ and $\{X_n : n \geq 1\}$ be iid R.V. s.t. $\mathbb{E}(|X_1|^r) < \infty$ and $\mathbb{E}(X_1) = 0$ if $r \geq 1$.

Then

$$\frac{S_n}{n} \xrightarrow{a.s.} 0$$

$n^{1/r}$

Conversely, $\frac{S_n}{n^{1/r}} \xrightarrow{a.s.} 0$ then

$\mathbb{E}(|X_1|^r) < \infty$ and $\mathbb{E}(X_1) = 0$

if $r \geq 1$.

Proof: Exercise / HW.

Applications of SLLN

Normal Numbers

Def: A number is normal in base 10 if the relative numb. of the coefficients in its decimal expansion is $\frac{1}{10}$.

decimal expansion is $\frac{1}{10}$.

Q: How many normal in $[0,1]$ are there?

<u>Thrm</u>: Almost all numbers in $[0,1]$ are normal.

Good example of uncountable set of measure zero.

<u>Ex</u>: $0.0123456789101112131 4 \ldots$

<u>Pf</u>: Pick a number in $[0,1]$ accord. to $\mathrm{Unif}(0,1)$.

If x_i is the ith decimal

$P(X_i = j) = \frac{1}{10}$ and mcnarr

$X_i \perp\!\!\!\perp X_j$ By the SLLN.

$$\underbrace{\frac{1}{n}\sum_{k=1}^{n} 1_{\{X_k = j\}}}_{\text{Relative frequency of } j} \xrightarrow{a.s.} \mathbb{E}(1_{\{X_k = j\}}) = \frac{1}{10}$$

Lecture 2/6

Tuesday, February 6, 2018 12:26 PM

Central Limit Theorems

The LLN tells us that if $\{X_n; n \geq 1\}$ are iid with finite mean μ then

$$\frac{S_n - n\mu}{n} \xrightarrow{a.s.} 0$$

In asymptotic analysis we found write

$$f(n) = a_1 \phi_1(n) + a_2 \phi_2(n) + O(\phi_3(n))$$

then

$$\frac{f(n)}{\phi_1(n)} \longrightarrow a_1$$

This shows that

$f(n) \simeq a_1 \phi_1(n)$. asymptotically

We could go one step further

$$\frac{f(n) - a_1 \phi_1(n)}{\phi_2(n)} \longrightarrow a_2$$

this shows that

$f(n) \simeq a_1 \phi_1(n) + a_2 \phi_2(n)$.

at the rate of $\phi_2(n)$.

But what happens when

But what happens when we don't have this decomposition.

Going back to the LLN. We know

$$S_n \simeq n\mu.$$

We would like to know the rate at which this equivalence holds.

<u>Q</u>: $S_n = n\mu + \phi_2(n) \}$?

Going back SLLN

$$S_n - n\mu \quad \text{a.s.}$$

$$\underbrace{\frac{S_n - n\mu}{n}}_{\text{R.V.}} \xrightarrow{a.s.} \underbrace{0}_{\text{degenerate}}$$

n too strong/fast to give anything interesting.

The CLT will tell us ($\mu = 0$, $\sigma^2 = 1$)

$$\frac{S_n}{\sqrt{n}} = \left(\frac{S_n}{\mu}\right)\sqrt{n} \xrightarrow{D} \xi$$

with $\xi \sim N(0,1)$.

So to answer our

question

$$\frac{S_n - n\mu}{\sqrt{n}} \longrightarrow \xi$$

$\Rightarrow$

$$S_n \simeq n\mu + \sqrt{n}\,\xi$$

"$f(n)$" "$\phi_1(n)a_1$" + "$\phi_2(n)a_2$"

only difference is that a_2 in our case is a R.V.

<u>Analogy</u>: LLN just need

$\mathbb{E}|X_1| < \infty \longrightarrow$ 1 Taylor expansion

expansion

In CLT we need

$E|X_1|^2 < \infty \longrightarrow 2$ Taylor expansion

So LLN looked at asymptotic equivalence of S_n now we need rates or second order equivalence and will get distributional terms.

CLT IID Case

Thrm: Take $\{X_n, n \geq 1\}$ iid seq. of R.V. with finite mean μ and variance $\infty > \sigma^2 > 0$. Then

$$\boxed{\frac{S_n - n\mu}{\sqrt{n}\,\sigma} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_k - \mu}{\sigma} \xrightarrow{D} N(0,1)}$$

Pf: As $\dfrac{S_n - n\mu}{\sqrt{n}\,\sigma} = \dfrac{1}{\sqrt{n}} \displaystyle\sum_{k=1}^{n} \frac{X_k - \mu}{\sigma}$

then $\dfrac{X_k - \mu}{\sigma}$ has mean 0 and variance 1.

WLOG assume $\mu = 0$ $\sigma = 1$

We need to show

$-t^2/2$

$$\phi_{\frac{S_n}{\sqrt{n}}}(t) \longrightarrow e^{-t^2/2} = \phi_{N(0,1)}(t)$$

First observe that

$$\phi_{\frac{S_n}{\sqrt{n}}}(t) = \mathbb{E}\left[e^{i\frac{t}{\sqrt{n}}\sum X_k}\right]$$

$$= \prod_{i=1}^{n} \mathbb{E}\left[e^{it/\sqrt{n}\,X_k}\right] = \phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n$$

Recall that if X is s.t. $E(|X|^n) < \infty$ $n \geq 1$ then

$$\phi_X(t) = \mathbb{E}\left[\sum_{k=0}^{n} \frac{\phi_X^{(k)}(0)}{k!}\right] + o(|t|^n)$$

$$= \mathbb{E}\left[\sum_{k=0}^{n} \frac{(iX)^k}{k!} t^k\right] + o(|t|^n)$$

For us this true for $n=2$. So we have

$$\phi_{X_1}(t/\sqrt{n}) = 1 + i\, t/\sqrt{n}\, \mathbb{E}(X_1) - \frac{t^2}{2n} \mathbb{E}(X_1^2) + o(t^2/n)$$

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

So

$$\lim_n \phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)^n = \lim_n \left(1 - \frac{t^2}{2n}\right)^n$$

Observe that

$$\log\left(1-\frac{t^2}{2n}\right)^n = \frac{\ln\left(1-\frac{t^2}{2n}\right)}{1/n}$$

So

$$\lim_n \frac{\ln\left(1-\frac{t^2}{2n}\right)}{1/n} = \lim_n \frac{\frac{t^2}{n^2}\times\frac{1}{1-\frac{t^2}{2n}}}{-\frac{1}{n^2}}$$

$$= -\frac{t^2}{2}$$

This gives

$$\lim_n \phi_{X_1}\left(t/\sqrt{n}\right)^n = \lim\left(1-\frac{t^2}{2n}\right)^n$$

$$= e^{-t^2/2}$$

[illegible] relax some of

Q: Can we relax some of the assumption?

Lindeberg-Lévy-Feller CLT

- We are going to relax the assumption of iid.
- Replace them with the Lindeberg Conditions

Let $\{X_n, n \geq 1\}$ be independent with finite means and

with finite mean-
finite variances (not all zero).
with

$$\mathbb{E}(X_k) = \mu_k \qquad S_n = \sum_{k=1}^{n} X_k$$

$$\mathrm{Var}(X_k) = \sigma_k^2 \qquad s_n^2 = \sum_{k=1}^{n} \sigma_k^2$$

The Lindeberg Conditions

(1) $$L_1(n) = \max_{1 \le k \le n} \frac{\sigma_k^2}{s_n^2} \longrightarrow 0$$

(2) $$L_2(n) = \frac{1}{s_n^2} \sum_{k=1}^{n} \mathbb{E}\left[(X_k - \mu_k)^2 \mathbb{1}_{(|X_k - \mu_k| > \varepsilon s_n)}\right] \longrightarrow 0$$

Thm: We take $(X_n)_{n=1}^{\infty}$ as

Thrm: We take $(X_n)_{n=1}$ as before. If L_2 holds then L_1 holds and

$$\frac{1}{s_n}\sum_{k=1}^{n}(X_k-\mu_k) \xrightarrow{D} N(0,1). \quad (CLT)$$

(ii) If L_1 and CLT hold then L_2 holds.

Rmk: If $(X_n)_{n=1}^{\infty}$ is s.t. if L_1 holds then $L_2 \iff CLT$ which implies that L_2 is almost necessary.

Interpretation of L_1 and L_2.

- They mean that the contribution of any individual R.V. to the sum of variances can be made arb. small.
- No individual R.V. dominates

Pf: (Lindeberg CLT).

$L_1 \Rightarrow \Delta_n^2 \longrightarrow \infty$. As not all variances are zero then $\exists m \in \mathbb{N}$ s.t. $\sigma_m^2 > 0$.

Then $n > m$, $\frac{\sigma_m^2}{s_n^2} \le L_1(n) \longrightarrow 0$

and $n \longrightarrow \infty$. Hence $s_n^2 \longrightarrow \infty$.

$L_2 \Rightarrow L_1$ We have that

$$\sigma_k^2 = Var(X_k) \le \mathbb{E}(X_k^2).$$

$$= \mathbb{E}[X_k^2 1\{|X_k| \le \varepsilon s_n\}] + \mathbb{E}[X_k 1_{(|X_k| > \varepsilon s_n)}]$$

WLOG take $E(X_k) = 0$.

$$L_1(n) = \max_{1 \le k \le n} \frac{\sigma_k^2}{s_n^2}$$

$$\le \max_{1 \le k \le n} \frac{1}{s_n^2} \mathbb{E}(X_k^2 1_{(|X_k| \le \varepsilon s_n)})$$

$+$

$$\max_{1 \le k \le n} \frac{1}{s_n^2} \mathbb{E}\left(X_k^2 \mathbb{1}(|X_k| \ge \varepsilon s_n) \right)$$

$$\le \varepsilon^2 + L_2(n)$$

Hence

$$\limsup L_1(n) \le \varepsilon^2.$$

Therefore

$$L_2(n) \longrightarrow 0 \implies L_1(n) \longrightarrow 0.$$

Lecture 2/8

Thursday, February 8, 2018 12:26 PM

Suppose we have $\{X_n, n \geq 1\}$ with finite means μ_k and finite variances σ_k^2.

Denote $S_n = \sum_{k=1}^{n} X_k$

$s_n^2 = \sum_{k=1}^{n} \sigma_k^2$

Recall the Lindeberg conditions

$$L_1(n) = \max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} \longrightarrow 0 \quad (L_1)$$

$$L_2(n) = \frac{1}{s_n^2} \sum_{k=1}^{n} \mathbb{E}\left[(X_k - \mu_k)^2 \mathbb{1}_{|X_k - \mu_k| > \varepsilon s_n} \right]$$

$$\longrightarrow 0 \quad \forall \varepsilon > 0 \quad (L_2).$$

Thm: For $\{X_n, n \geq 1\}$

(i) If L_2 holds then

L_1 holds and

$$\frac{1}{s_n} \sum_{k=1}^{n} (X_k - \mu_k) \longrightarrow N(0,1)$$

(ii) If L_1 and CLT hold

then L_2 holds.

Pf: We saw that $L_2 \Rightarrow$

$s_n^2 \longrightarrow \infty$ and

$L_2 \Rightarrow L_1$

(Pf of sufficiency) ($L_2 \Rightarrow$ CLT)

[illegible]

We assume L_2 holds which implies L_1 holds.

Aux Result. For $\{z_k : 1 \leq k \leq n\}$ and $\{w_k ; 1 \leq k \leq n\}$ be sequences in $\mathbb{C}$ with $|z_k| \leq 1$, $|w_k| \leq 1$ then

$$\left| \prod_{k=1}^{n} z_k - \prod_{k=1}^{n} w_k \right| \leq \sum_{k=1}^{n} |z_k - w_k| .$$

Pf: Induction. Let

$$z^* = \prod^{n} z_k \qquad w_n^* = \prod^{n} w_k$$

$$z_n^* = \prod_k z_k \qquad w_n^* = \prod_{k=1} w_k$$

$$z_{n+1}^* - w_{n+1}^*$$

$$= z_n^* (z_{n+1} - w_{n+1})$$
$$+ w_{n+1} (z_n^* - w_n^*)$$

So

$$|z_{n+1}^* - w_{n+1}^*| \leq$$

$$|z_{n+1} - w_{n+1}| + |z_n^* - w_n^*|$$

$$\leq |z_{n+1} - w_{n+1}| + \sum_{k=1} |z_k - w_k|$$

Aux Result: For $n \geq 0$

$$\left| e^{iy} - \sum^{n} \frac{(iy)^k}{!!} \right|$$

$$\left| e^{iy} - \sum_{k=0}^{n} \frac{(iy)^k}{k!} \right|$$

$$\leq \min \left\{ \frac{2|y|^n}{n!}, \frac{|y|^{n+1}}{(n+1)!} \right\}$$

We want to prove that

$$\left| \phi_{\frac{S_n}{s_n}}(t) - e^{t^2/2} \right| \longrightarrow 0$$

Recall that we assumed WLOG that the X_k are centered.

$$\phi_{\frac{S_n}{s_n}}(t) = \mathbb{E}\left(e^{it \frac{1}{s_n} \sum X_k} \right)$$

$$= \prod_{k=1}^{n} \phi_{X_k}\left(\frac{t}{s_n}\right)$$

Also $e^{-t^2/2} = e^{-\frac{t^2}{2} \cdot \frac{s_n^2}{s_n^2}}$

$$= e^{-\frac{t^2}{2} \frac{\sum \sigma_k^2}{s_n^2}}$$

$$= \prod_{i=1}^{n} e^{-\frac{t^2 \sigma_k^2}{2 s_n^2}}$$

Looking at the difference

$$\left| \phi_{\frac{S_n}{s_n}}(t) - e^{-\frac{t^2}{2}} \right| = \left| \prod_{k=1}^{n} \phi_{X_k}\left(\frac{t}{s_n}\right) - \prod_{k=1}^{n} e^{-\frac{t^2 \sigma_k^2}{2 s_n^2}} \right|$$

Seeing that both of these have $|\cdot| \leq 1$ we can use the Aux. Result I. to write

$$\left| \phi_{\frac{S_n}{s_n}}(t) - e^{-\frac{t^2}{2}} \right| \leq \sum_{k=1}^{n} \left| \phi_{X_k}\left(\frac{t}{s_n}\right) - e^{-\frac{t^2 \sigma_k^2}{2 s_n}} \right|$$

Add/Sub. 2 order Taylor approximation.

$$\leq \sum_{k=1}^{n} \left| \phi_{X_k}\left(\frac{t}{s_n}\right) - \left(1 - \frac{t^2 \sigma_k^2}{2 s_n^2}\right) \right|$$

$$+ \sum_{k=1}^{n} \left| e^{-\frac{t^2 \sigma_k^2}{2 s_n^2}} - \left(1 - \frac{t^2 \sigma_k^2}{2 s_n^2}\right) \right|$$

observe that the second term is a particular case when $X_k \sim N(0, \sigma_k^2)$.

case when $X_k \sim N(0, \sigma_k^2)$.

So if we can handle the general case we can handle the normal case.

Now, we have

$$\sum_{k=1}^{n} \left| \mathbb{E}\left[e^{i\frac{t}{s_n} X_k} - \left(1 + i\frac{t}{s_n} X_k - \frac{t^2 X_k^2}{2 s_n}\right)\right]\right|$$

$$\leq \sum_{k=1}^{n} \mathbb{E}\left| e^{i t/s_n X_k} - \left(1 + i\frac{t}{s_n} X_k - \frac{t^2 X_k^2}{2 s_n^2}\right)\right|$$

Jensens

Let $n = 2$ and $y = \frac{t}{s_n} X_k$

then we get

$$\leq \sum^{n} \mathbb{E}\left[\min\left\{ \frac{t^2 X_k^2}{\ldots}, \frac{|t|^3 |X_k|^3}{\ldots} \right\}\right]$$

$$\leq \sum_{k=1}^{n} \mathbb{E}\left[\min\left\{ \frac{t^2 X_k^2}{\Delta_n^2}, \frac{|t|^3 |X_k|^3}{6\Delta_n^3} \right\} \right]$$

$$= \sum_{k=1}^{n} \mathbb{E}\left[\min\{\ \} \mathbb{1}_{|X_k| \leq \varepsilon \Delta_n} \right]$$

$$+$$

$$\sum_{k=1}^{n} \mathbb{E}\left[\min\{\ \} \mathbb{1}_{|X_k| > \varepsilon \Delta_n} \right]$$

$$\leq \sum_{k=1}^{n} \mathbb{E}\left[\frac{|t|^3 |X_k|^3}{6\Delta_n^3} \mathbb{1}_{|X_k| \leq \varepsilon \Delta_n} \right]$$

$$+$$

$$\sum_{k=1}^{n} \mathbb{E}\left[\frac{t^2 X_k^2}{\Delta_n^2} \mathbb{1}_{|X_k| > \varepsilon \Delta_n} \right]$$

$$\leq \frac{|t|^3 \varepsilon \Delta_n}{6\Delta_n^3} \sum_{k=1}^{n} \mathbb{E}\left(X_k^2 \mathbb{1}_{|X_k| \leq \varepsilon \Delta_n} \right)$$

$$+ t^2 L_2(n)$$

$$\leq \frac{|t|^3 \varepsilon \Delta_n}{6 \Delta_n^3} \sum_{k=1}^{n} \sigma_k^2 + t^2 L_2(n)$$

$$= \frac{|t|^3}{6} \varepsilon + t^2 L_2(n)$$

Hence

$$\limsup_n \left| \phi_{\frac{S_n}{\Delta_n}}(t) - e^{-t^2/2} \right|$$

$$\leq \frac{|t|^3}{3} \varepsilon \quad (\text{as } L_2(n) \to 0)$$

Therefore

$$\phi_{\frac{S_n}{\Delta_n}}(t) \xrightarrow{D} e^{-t^2/2}$$

Pf of necessity: (L_1 + CLT $\Rightarrow$ L_2)

Overview:

Assume L_1 + CLT. Well

CLT $\Rightarrow$ $\left| \phi_{\frac{S_n}{s_n}}(t) - e^{-\frac{t^2}{2}} \right| \to 0$

① Prove

$$\lim_n \phi_{X_n}\left(\frac{t}{s_n}\right) = \lim_n e^{\phi_{X_n}\left(\frac{t}{s_n}\right) - 1}$$

② Deduce

$$e^{t^2/2} \prod_{k=1}^{n} e^{\phi_{X_n}(t/s_n) - 1} \longrightarrow 1.$$

Aux. Result if $L_1(n) \to \infty$

$$\left|\phi_{X_k}\left(\frac{t}{s_n}\right)-1\right| \le \mathbb{E}\left(\left|e^{\frac{itX_k}{s_n}}-1\right|\right)$$

$$\le \mathbb{E}\left[e^{\frac{itX_k}{s_n}}-\left(1+\frac{it}{s_n}X_k\right)\right]$$

So by Lemma 2.

$$\le \mathbb{E}\left[\min\left\{\frac{2(\varepsilon(|X_k|}{s_n}, \frac{t^2X_k^2}{2s_n^2}\right\}\right]$$

So $\displaystyle\sum_{k=1}^{n}\left|\phi_{X_k}\left(\frac{t}{s_n}\right)-1\right|^2$

$$\le \max_{1\le k\le n}\left|\phi_{X_k}\left(\frac{t}{s_n}\right)-1\right|\sum_{k=1}^{n}\left|\phi_{X_k}\left(\frac{t}{s_n}\right)-1\right|$$

$$\le \max \frac{t\sigma_k^2}{2} \cdot \sum_{k=1}^{n}\frac{t^2\sigma_k^2}{2s_n^2}$$

$$\leq \max_{1\leq k\leq n} \frac{t^2\sigma_k^2}{2s_n^2} \cdot \sum_{k=1}^{n} \frac{t^2\sigma_k^2}{2s_n^2}$$

$$= \frac{t^4}{4} L_1(n) \longrightarrow 0$$

Pf of ①.

So now we want to show

$$\left| \prod_{k=1}^{n} \phi_{X_k}\left(\frac{t}{s_n}\right) - \prod_{k=1}^{n} e^{\phi_{X_k}\left(\frac{t}{s_n}\right)-1} \right|$$

$$\leq \sum_{k=1}^{n} \left| e^{\phi_{X_k}(t/s_n)-1} - \phi_{X_k}\left(\frac{t}{s_n}\right) \right|$$

$$= \sum_{k=1}^{n} \Bigl| \underbrace{e^{\phi_{X_k}(t/s_n)-1} - 1 - \left(\phi_{X_k}(t/s_n)-1\right)}_{e^z-1-z} \Bigr|$$

For $|z| \le \frac{1}{2}$ then

$$\left|e^z - 1 - z\right| = \left|\sum_{k=2}^{\infty} \frac{z^k}{k!}\right|$$

$$\le \sum_{k=2}^{\infty} \frac{|z|^k}{k!} \le \frac{1}{2} \sum_{k=2}^{\infty} |z|^k$$

$$= \frac{|z|^2}{2} \sum_{k=2}^{\infty} |z|^{k-2}$$

$$= \frac{|z|^2}{2} \sum_{k=0}^{\infty} |z|^k = \frac{|z|^2}{2} \frac{1}{1-|z|} \le |z|^2$$

Applying this to our situation for sufficiently large n $\quad t_n \longrightarrow \infty$

$$\left| \phi_{X_k}(t/s_n) - 1 \right| \leq 1/2$$

So our result is given by

$$\sum \left| e^{\phi_{X_k}(t/s_n) - 1} - 1 - (\phi_{X_k}(t/s_n) - 1) \right|$$

$$\leq \sum_{k=1}^{n} \left| \phi_{X_k}(t/s_n) - 1 \right|^2 \to 0$$

So

$$e^{t^2/2} \prod_{k=1}^{n} e^{\phi_{X_k}(t/s_n) - 1} \longrightarrow 1.$$

Hence

$$\sum \left(\phi_{X_k}(t/s_n) - 1 + \frac{t^2 \sigma_k^2}{\cancel{8} s_n^2} \right)$$

$e \longrightarrow 1$

So

$\mathrm{Re}(LHS) \longrightarrow 1$

So

$$\sum \mathbb{E}\left(\cos\left(\frac{t}{s_n}X_k\right) - 1 + \frac{t X_k^2}{2 s_n^2}\right) \longrightarrow 0$$

The function

$$\cos(x) - 1 - \frac{x^2}{2} \geq 0$$

So we can write

$$\sum_{k=1}^{n} \mathbb{E}\left[\cancel{+}\ 1_{|X_k| > \varepsilon s_n}\right] \longrightarrow 0$$

Lecture 2/13

Tuesday, February 13, 2018 12:33 PM

[Necessity (L_1 + CLT $\Rightarrow L_2$)]

Take $\{X_n, n \geq 1\}$ to be a sequence as before

Thrm:

(i) $L_2 \Rightarrow L_1$ + $L_2 \Rightarrow$ CLT

(ii) L_1 + CLT $\Rightarrow L_2$

From before,

$$e^{t^2/2} \prod_{k=1}^{n} e^{\phi_{X_k}(t/s_n) - 1} \longrightarrow 1$$

because

$$\phi_{\frac{S_n}{s_n}}(t) = \prod_{k=1}^{n} \phi_{X_k}\left(\frac{t}{s_n}\right) \to e^{-t^2/2}$$

Now this implies

$$\exp\left\{ \sum_{k=1}^{n} \phi_{X_k}(t/s_n) - 1 + \frac{t^2 \sigma_k^2}{2 s_n^2} \right\} \xrightarrow[n\to\infty]{} 1$$

So we can conclude that

$$\mathrm{Re}\left\{ \sum_{k=1}^{n} \phi_{X_k}\left(\frac{t}{s_n}\right) - 1 + \frac{t^2 \sigma_k^2}{2 s_n^2} \right\} \longrightarrow 0$$

or

$$\sum^{n} E\left[\quad \left(\frac{t}{\quad}\right) \quad \frac{t^2 X_k^2}{\quad} \right]$$

$$\sum_{k=1}^{n} \mathbb{E}\left[\cos\left(\frac{t}{s_n} X_k\right) - 1 + \frac{t^2 X_k^2}{2 s_n^2}\right] \to 0$$

Observe that

$$\cos(x) - 1 + \frac{x^2}{2} \geq 0$$

So this gives

$$\sum_{k=1}^{n} \mathbb{E}\left[\cos\left(\frac{t X_k}{s_n}\right) - 1 + \frac{t^2 X_k^2}{2 s_n^2}\right] 1_{\{|X_k| > \varepsilon s_n\}}$$

$$\xrightarrow[n \to \infty]{} 0$$

So we have the fact

that

$$\frac{t^2}{2}L_2(n) - \sum \mathbb{E}\left[\left(1-\cos\left(\frac{t}{s_n}X_k\right)\right)\mathbb{1}_{|X_k|>\varepsilon s_n}\right]$$

$$\longrightarrow 0$$

So

$$\limsup \frac{t^2}{2}L_2(n)$$

$$= \limsup \sum \mathbb{E}\left[\left(1-\cos\left(\frac{t}{s_n}X_k\right)\right)\mathbb{1}_{|X_k|>\varepsilon s_n}\right]$$

So focusing on the bounding term.

$$\left|\frac{2}{t^2}\sum \mathbb{E}\left([\;]\mathbb{1}_{|X_k|>\varepsilon s_n}\right)\right|$$

$$\leq \frac{4}{} \sum^{n} \mathbb{P}\left(|X_k|>\varepsilon s_n\right)$$

$$\le \frac{1}{t^2} \sum_{k=1}^{n} \mathbb{P}(|X_k| > \varepsilon s_n)$$

$$\le \frac{4}{t^2} \sum \frac{\mathbb{E}(X_k^2)}{\varepsilon^2 s_n^2} = \frac{4}{t^2 \varepsilon^2}$$

Hence

$$\limsup L_2(\omega) \le \frac{4}{(t\varepsilon)^2}$$

As t is arbitrary, we can make this quantity arb. small.

<u>Rmk</u>: Assume that

$s_n^2 \longrightarrow s^2 < \infty$. Then for

$n \geq m$, $\mathbb{E}[\,|S_n - S_m|^2\,]$

$= \sum_{k=m+1}^{n} \sigma_k^2$ by ind/centering

$\longrightarrow 0$ and $n, m \to \infty$

So $\{S_n, n \geq 1\}$ Cauchy/Converges

in L^2 to S $\Rightarrow$

$S_n \xrightarrow{D} S$. (No normalizing here).

<u>Thrm</u>: (Cramér's Thrm).

If $X \perp\!\!\!\perp Y$ and nondegenerate

R.V. s.t. $X+Y \sim N$ then

$X, Y \sim N$.

Assuming that the CLT holds i.e. $S \sim N(0,1)$.

Then by Cramérs Thrm.

$X_k \sim N$. from the start.

Which isn't exactly exciting... So $L_1(n)$ only excludes trivial cases.

- So the conditions are sharp in the sense that the conditions extend past trivial cases.

<u>Rmk:</u> If $X_k \sim N(0, \sigma_k^2)$ and $\sum \sigma_k^2 = s^2 < \infty$

then $\frac{S_n}{s_n} \sim N(0,1)$ for all n and $\frac{S_n}{s_n} \xrightarrow{D} N(0,1)$

But

$L_n(\eta) = \frac{1}{\quad} \sum \mathbb{E}[X_k^2 \mathbb{1}_{|X_k| \ldots}]$

$$L_2(n) = \frac{1}{s_n^2} \sum \mathbb{E}\left[X_k^2 1_{|X_k| > \varepsilon s_n}\right]$$

$$\geq \frac{1}{s^2} \mathbb{E}\left[X_k^2 1_{|X_k| > \varepsilon s_n}\right]$$

$$\geq \frac{1}{s^2} \mathbb{E}\left[X_1^2 1_{|X_1| > \varepsilon s}\right] > 0$$

So in any case

$$\lim_n L_2(n) > 0.$$

Taking $\sigma_k^2 = \frac{1}{2^k}$ then

$$s_n^2 = \sum_{k=1}^{n} \frac{1}{2^k} = 1 - \frac{1}{2^n}$$

and

$$L_1(n) = \max_{1 \le k \le n} \frac{\sigma_k^2}{s_n^2} = \frac{1/2}{1 - 1/2^n}$$

$$\longrightarrow 1/2 > 0.$$

So these summabilities imply neither L_1 + L_2 apply.

<u>Ex</u>: Let $X_k \sim \text{Pois}(1/2^k)$ be iid then $S_n \to \text{Pois}(1)$

$\not\to N(0,1)$.

<u>Thrm</u>: (Lyapounov) Let $\{X_k\}$ be as in Lindeberg.

Assume $\mathbb{E}[|X_k|^r] < \infty$

for $r > 0$. Then if

$$\beta(n,r) = \frac{1}{s_n^r} \sum_{k=1}^{n} \mathbb{E}(|X_k - \mu_k|^r) \to 0$$

then the CLT hold.

<u>Application</u>: Stirling Formula

$$\lim_n \frac{\left(\frac{n}{e}\right)^n \sqrt{2n\pi}}{n!} = \underline{1}$$

$$\left(\frac{n}{e}\right)^n \sqrt{2n\pi} \sim n!$$

<u>Pf</u>: Let $\{X_n\}$ be a

Seq. of $\perp\!\!\!\perp$ standard Exponentials. $X_k \sim \text{Exp}(1)$.

So $\mathbb{E}(S_n) = n$. Then the CLT says

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

As $\mathbb{E}\left[\left(\frac{S_n - n}{\sqrt{n}}\right)^2\right] = 1 \ \forall n \geq 1$.

Uniform Integrability then implies that

$$\mathbb{E}\left(\left|\frac{S_n - n}{\sqrt{n}}\right|\right) \to \mathbb{E}(|N(0,1)|)$$
$$= \sqrt{\tfrac{2}{\pi}}$$

$$= \sqrt{\frac{2}{\pi}}$$

But as $\mathrm{Exp}(1) \sim \mathrm{Gamma}(1,1)$

$S_n \sim \mathrm{Gamma}(n,1)$.

So

$$\mathbb{E}\left(\left|\frac{S_n - n}{\sqrt{n}}\right|\right)$$

$$= \int_0^\infty \left|\frac{x-n}{\sqrt{n}}\right| \frac{1}{\Gamma(n)} x^{n-1} e^{-x}\, dx \longrightarrow \sqrt{\frac{2}{\pi}}$$

which gives Sterling Formula.

Lecture 2/15

Thursday, February 15, 2018 12:26 PM

- CLT $\Longrightarrow$ rate for LLN
- Can we find rates for C.L.T.?
- We will need metrics on prob. measures to answer this.

Stien's Method:

Quantifying the C.L.T.

I – Metrics on the set of prob measures on $\mathbb{R}$.

Def: Take $\mathcal{H}$ to be a collection of Borel measurable functions on $\mathbb{R}$. Let

Th...

F, G be s.t. $h(F), h(G) \in L^1(\Omega)$ for all $h \in H$. We say H is seperating if for any F, G, R.V. s.t. $h(F), h(G) \in L^1$

$\mathbb{E}[h(F)] = \mathbb{E}[h(G)] \quad \forall h \in H$

$\Rightarrow$ F and G have the same law.

<u>Ex:</u> $H = \{ \mathbb{1}_{(-\infty, z)}(\cdot), z \in \mathbb{R} \}$

then $\mathbb{E}[h(F)] = \mathbb{E}[h(G)]$

$\Leftrightarrow \mathbb{P}(F \leq z) = \mathbb{P}(G \leq z)$

$H = \{ \mathbb{1}_B, B \in \mathcal{B}(\mathbb{R}) \}$

$$\mathcal{H} = \{ \mathbb{1}_B , B \in \mathcal{B}(\mathbb{R}) \}$$

$\mathcal{H} : \{ h : h$ is Borel measurable & Bouned $\}$

Def: Let $\mathcal{H}$ be a seperating class and F, G R.V. s.t. $h(F)\ h(G) \in L^1$. the distance between the laws of F, G indeced $\mathcal{H}$ is given

$$d_{\mathcal{H}} \langle F, G \rangle = \sup_{h \in \mathcal{H}} \left| \mathbb{E}(h(F)) - \mathbb{E}(h(G)) \right|$$

Prop: $d_{\mathcal{H}}$ is a metric on a subset of prob. measures on $\mathbb{R}$.

<u>Pf</u>: Exercise.

<u>Some Special Cases</u>

<u>Def</u>: (Kolmogorov metric) the K-M. is induced by the

$$H_{kol} = \{ 1_{(-\infty, z]} : z \in \mathbb{R} \}.$$

Then

$$d_{Kol}(F,G) = \sup_{h \in H_{kol}} \left| \mathbb{E}(h(F)) - \mathbb{E}[h(G)] \right|$$

$$= \sup_{z \in \mathbb{R}} \left| \mathbb{P}(F \le z) - \mathbb{P}(G \le z) \right|$$

<u>Def</u>: (Total Variation metric) The TV metric is induced by $H_{TV} = \{ 1_B : B \in \mathcal{B}(\mathbb{R}) \}$

by $\mathcal{H}_{TV} = \{ 1_B : B \in \mathcal{B}(\mathbb{R}) \}$

We write

$$d_{TV}(F,G) = \sup_{h \in \mathcal{H}_{TV}} \left| \mathbb{E}(h(F)) - \mathbb{E}(h(G)) \right|$$

$$= \sup_{B \in \mathcal{B}(\mathbb{R})} \left| \mathbb{P}(F \in B) - \mathbb{P}(G \in B) \right|$$

Def: (Wasserstein-metric) The WM induced by $\mathcal{H}_W = \{ h : \mathbb{R} \to \mathbb{R} : \|h\|_L \leq 1 \}$

where $\|h\|_L = \sup_{\substack{x \neq y \\ x,y \in \mathbb{R}}} \frac{|h(x) - h(y)|}{|x-y|}$

$$d_W(F,G) = \sup_{h \in \mathcal{H}_W} \left| \mathbb{E}(h(F)) - \mathbb{E}(h(G)) \right|$$

Proposition $d_{TV} \geq d_{KoL}$ and

$$d_{Kol}(F, N(0,1)) \leq 2\sqrt{d_W(F, N(0,1))}$$

$d_{Kol}(F, N(0,1)) \le 2\sqrt{d_W(F, N(0,1))}$

<u>Proposition</u>: d_{TV}, d_{Kol}, d_W induce topologies that are <u>strictly stronger</u> than the topology of convergence in law.

<u>Pf</u>: HW 4.

<u>Prop</u>: Let $\{F, F_n; n \geq 1\}$ be a seq. of R.V. s.t. the map $x \longmapsto \mathbb{P}(F_n \le x)$ is cont. $\forall x \in \mathbb{R}$. Then

$$d_{Kol}(F, F_n) \longrightarrow 0 \iff F_n \xrightarrow{D} F$$

In this case, d_{Kol} metrizes convergence in law.

Recall convergence in law
is characterized by
$\mathbb{E}(h(F_n)) \longrightarrow \mathbb{E}(h(F))$
$\forall h$ cont and bounded

Rmk: d_{TV}, d_W never metrizes
Convergence in distribution.

II - Stien's method

We begin by characterizing
$N(0,1)$

Take $\gamma(dx) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$

Then

$$\int_{\mathbb{R}} f'(x)\,\gamma(dx) = \int_{\mathbb{R}} x f(x)\,\gamma(dx)$$

Stien's Lemma: A real valued R.V. $N \sim (0,1)$ dist iff $\forall h: \mathbb{R} \to \mathbb{R}$ that are differentiable it holds that if $\mathbb{E}(h'(N)) < \infty$ and $\mathbb{E}[N h(N)] < \infty$ then

$$\mathbb{E}[h'(N) - N h(N)] = 0$$

Pf: HW#4.

Stein's Hueristic: If $\mathbb{E}[h'(a) - a h(a)]$ is close to zero for all h s.t.

to zero for all h s.t.

$\mathbb{E}[h'(N)] < \infty$ $\mathbb{E}[a h(a)] < \infty$

Could we say that the law is close to $N(0,1)$?

If so, how close?

Stein's Equation

<u>Def</u>: Let $h : \mathbb{R} \to \mathbb{R}$ be a Borel measurable s.t. $\mathbb{E}(|h(N)|) < \infty$ where $N \sim N(0,1)$.

The Stein equation associated with the function h is the ODE.

$$f'(x) - x f(x) = h(x) - \mathbb{E}[h(N)].$$

A solution to the above

A solution to the above

is an abs. cont. function

f s.t. f' satisfies above.

<u>Prop</u>: Every solution to the Stein equation has the form

$$f(x) = c e^{x^2/2} + e^{x^2/2} \int_{-\infty}^{x} \{h(y) - \mathbb{E}(h(N))\} e^{-y^2/2} \, dy$$

In particular

$$f_h(x) = e^{x^2/2} \int_{-\infty}^{x} \{h(y) - \mathbb{E}(h(N))\} e^{-y^2/2} \, dy$$

then if f_h is the unique solution to Stein's equation s.t.

$$\lim e^{-x^2/2} f_h(x) = 0$$

$$\lim_{x \mapsto \pm\infty} e^{-x^2/2} f_n(x) = 0$$

Pf: HW 4

- We are now going to combine Stein's Lemma with metrics

Let F be a R.V. that we want to compare its law to $N \sim N(0,1)$. Let $h \in H$. when H is seperable. Then by Stien's eq.

$$f_n'(x) - x f_n(x) = h(x) - \mathbb{E}(h(N))$$

Take $x = F$ and taking

expectations

$$\underbrace{\mathbb{E}\left(f_n'(F) - F f_n(F)\right)}_{\text{Stein's Lemma}} = \underbrace{\mathbb{E}(h(F)) - \mathbb{E}(h(N))}_{\text{quant. of metric.}}$$

Then taking $\sup_{h \in \mathcal{H}}$ we get.

$$\sup_{h \in \mathcal{H}} \left| \mathbb{E}\left(f_h'(F) - F f_h(F)\right) \right| = d_{\mathcal{H}}(F, N).$$

Distance independent of N.

Need only consider

$$\sup_h \mathbb{E} \left| F - F f_h(F) \right|$$

to measure distance from Z

So we have successfully quantified the distance between laws in a proper metric space.

Lecture 2/27

Tuesday, February 27, 2018 12:33 PM

Recall from HW 1 that

Convergence in d_{Kol}

is strictly stronger than

$d_{Distr.}$

<u>Comments on HW 5</u>

$$d_{\mathcal{H}}(F,N) = \sup_{h \in \mathcal{H}} \left| \mathbb{E}(f_h'(F)) - Ff_h(F) \right|$$

By Stein equation

$\|f_h'\|_\infty \leq c_1$ $\|f_h\|_\infty \leq c_2$

Then we can enlarge $\mathcal{H}$

to the class

$$F_{\mathcal{H}} = \left\{ f : \|f\|_\infty \leq c_1 \ \ \|f'\|_\infty \leq c_2 \right\}$$

So

$$d_{\mathcal{H}}(F,N) \leq \sup_{f \in F_{\mathcal{H}}} \mathbb{E}\left| f'(F) - Ff(F) \right|$$

In exercise 3 we want to compare $N_1 \sim N(0, \sigma_1^2)$
$N_2 \sim N(0, \sigma_2^2)$

Berry-Essen Thrm

Thrm: Let $\{X_n, n \geq 1\}$ be iid seq. of R.V. s.t. $\mathbb{E}(X_1) = 0$, $\mathbb{E}(X_1^2) = 1$ then letting $V_n = \frac{S_n}{\sqrt{n}}$ we have

$$V_n \xrightarrow{D} N(0,1)$$

by the C.L.T. Moreover

$$d_{Kol}(V_n, N) \leq \frac{C \, \mathbb{E}(|X_1|^3)}{\sqrt{n}}$$

Pf: Assume that $\mathbb{E}(|X_1|^3) < \infty$ and define c_n be the best constant s.t.

$$d_{Kol}(V_n, N) = \frac{c_n \, \mathbb{E}(|X_1|^3)}{}$$

s.t.

$$d_{Kol}(V_n, N) \leq \frac{c_n \, \mathbb{E}(|X_1|^3)}{\sqrt{n}}$$

Take $c_n \leq \frac{\sqrt{n}}{\mathbb{E}(|X_1|^3)}$

then

$$d_{Kol}(V_n, N) \leq 1$$

So c_n is a well defined.

Goal is to show $\sup_n c_n < \infty$

By Hölder the Cauchy-Swartz

$$\mathbb{E}(X_1^2) \leq \left(\mathbb{E}|X_1|^3\right)^{1/3}\left(\mathbb{E}|X_1|^{3/2}\right)^{2/3}$$

$$\leq \left(\mathbb{E}|X_1|^3\right)^{1/3}\left(\mathbb{E}|X_1|^3\right)^{1/3}$$

which gives

$$1 = \mathbb{E}(X_1^2) \leq \mathbb{E}(|X_1|^3)^{2/3}$$

Then

$$\frac{1}{\mathbb{E}(|X_1|^3)} \leq 1 \Leftrightarrow \frac{\sqrt{n}}{\mathbb{E}(|X_1|^3)} \leq \sqrt{n}$$

$$\iff \quad c_n \le \frac{\sqrt{n}}{\mathbb{E}(|X_1|^3)} \le \sqrt{n}$$

So $\quad c_n \le \sqrt{n}$

We need to prove

$$\sup_n c_n < \infty.$$

In fact $\quad \sup_n c_n < 33$

Returning to $d_{Kol}(V_n, N)$.

$$|\mathbb{P}(V_n \le z) - \mathbb{P}(N \le z)|$$

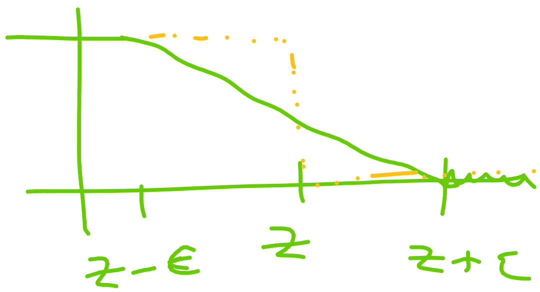
$$= \left| \mathbb{E}\left(1_{(-\infty, z]}(V_n) - 1_{[-\infty, z)}(N) \right) \right|$$

Would be really helpful to have regularizing cont. functions.

For any $z \in \mathbb{R}$ $\epsilon > 0$ let our approximating function be given by

$$h_{z,\epsilon}(x) = \begin{cases} 1 & x \le z - \epsilon \\ \text{linear} & x \in (z-\epsilon, z+\epsilon) \end{cases}$$

$$h_{z,\varepsilon}(x) = \begin{cases} \text{linear} & x \in (z-\varepsilon, z+\varepsilon) \\ 0 & x \geq z-\varepsilon \end{cases}$$

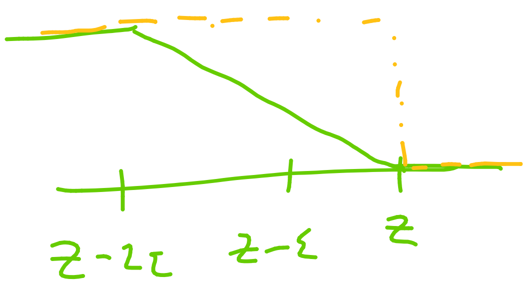


then

$$\lim_{\varepsilon \to 0} h_{z,\varepsilon}(x) = 1_{(-\infty,z]}(x)$$

Observe that

$$h_{z-\varepsilon,\varepsilon}(x) \leq 1_{(-\infty,z)}(x)$$



$$\mathbb{E}\left(h_{z-\varepsilon,\varepsilon}(V_n)\right) \leq \mathbb{P}(V_n \leq z)$$

Similarly

$$1_{(-\infty,z)}(x) \leq h_{z+\varepsilon,\varepsilon}(x) \quad \forall x$$

$$1_{(-\infty,z)}(x) \le h_{z+\varepsilon,\varepsilon}(x) \quad \forall x$$

and hence

$$\mathbb{P}(V_n \le z) \le \mathbb{E}\left(h_{z+\varepsilon,\varepsilon}(V_n)\right)$$

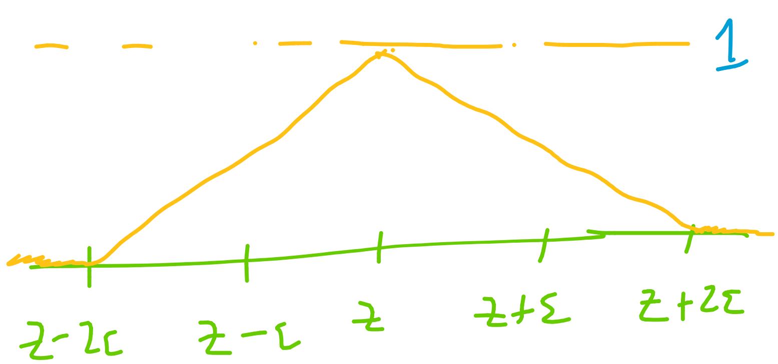
So

$$\mathbb{E}\left(h_{z-\varepsilon,\varepsilon}(V_n)\right) \le \mathbb{P}(V_n \le z) \le \mathbb{E}\left(h_{z+\varepsilon,\varepsilon}(V_N)\right)$$

Observe that

$$h_{z+\varepsilon,\varepsilon}(x) - h_{z-\varepsilon,\varepsilon}(x) = 0$$

for $x \notin [z-2\varepsilon, z+2\varepsilon]$



So

$$\mathbb{E}\left(h_{z+\varepsilon}(N)\right) - \mathbb{E}\left(h_{z-\varepsilon,\varepsilon}(N)\right)$$

$$\mathbb{E}(h_{z+\varepsilon}(N)) - \mathbb{E}(h_{z-\varepsilon,\varepsilon}(N))$$

$$= \int_{z-2\varepsilon}^{z+2\varepsilon} \underbrace{\{h_{z+\varepsilon}(x) - h_{z-\varepsilon,\varepsilon}(x)\}}_{\leq 1} dN(x)$$

$$\leq \frac{1}{\sqrt{2\pi}} [x]_{z-2\varepsilon}^{z+2\varepsilon} = \frac{4\varepsilon}{\sqrt{2\pi}}$$

Now by the same argument as before.

$$\mathbb{E}(h_{z-\varepsilon,\varepsilon}(N)) \leq \mathbb{P}(N \leq g) \leq \mathbb{E}(h_{z+\varepsilon}(N))$$

$$\leq \mathbb{E}(h_{z-\varepsilon,\varepsilon}(N)) + \frac{4\varepsilon}{\sqrt{2\pi}}$$

Building the lower bound

$$\mathbb{E}(h_{z+\varepsilon,\varepsilon}(N)) - \frac{4\varepsilon}{\sqrt{2\pi}} \leq \mathbb{P}(N \leq z).$$

Now take

$$\mathbb{P}(V_n \leq z) - \mathbb{P}(N \leq z)$$

$$\leq \mathbb{E}(h_{z+\varepsilon,\varepsilon}(N)) - \mathbb{E}(h_{z+\varepsilon,\varepsilon}(N)) + \frac{4\varepsilon}{2\pi}$$

and also

$$\mathbb{P}(V_n \le z) - \mathbb{P}(N \le z)$$

$$\ge \mathbb{E}(h_{z-\varepsilon,\varepsilon}(V_n)) - \mathbb{E}(h_{z-\varepsilon,\varepsilon}(N)) - \frac{4\varepsilon}{2\pi}.$$

So

$$|\mathbb{P}(V_n \le z) - \mathbb{P}(N \le z)|$$

$$\le \max\left\{ \left|\mathbb{E}(h_{z-\varepsilon,\varepsilon}(V_n) - h_{z-\varepsilon,\varepsilon}(N) - \frac{4\varepsilon}{2\pi}\right|, \left|\mathbb{E}(h_{z+\varepsilon,\varepsilon}(V_n) - h_{z.}(N) + \frac{4\varepsilon}{2\pi}\right| \right.$$

Then

$$\sup_{z\in\mathbb{R}} |\mathbb{P}(V_n \le z) - \mathbb{P}(N \le z)|$$

$$\le \sup_z |\mathbb{E}(h_{z,\varepsilon}(V_n) - h_{z,\varepsilon}(N))| + \frac{4\varepsilon}{2\pi}$$

as $\sup_{z\in\mathbb{R}} |\mathbb{E}(h_{z+\varepsilon,\varepsilon}(V_n) - \dots|$

$$= \sup_z |\mathbb{E}(h_{z,\varepsilon}(V_n) - \mathbb{E}(h_{z,\varepsilon}(N))|$$

Therefore

$$d_{kol}(V_n, N) \le \sup_{z\in\mathbb{R}} |\mathbb{E}(h_{z,\varepsilon}(V_n) - \mathbb{E}(h_{z,\varepsilon}(N))| + \frac{4\varepsilon}{2\pi}$$

So here is where
will use Stein's method.

Lecture 3/20

Tuesday, March 20, 2018 12:29 PM

<u>Thrm</u>: Take $\{X_n\}$ be iid

s.t. $\mathbb{E}(X_i) = 0$ $\mathbb{E}(X_i^2) = 1$.

Then setting $V_n = \frac{1}{\sqrt{n}} \sum X_i$

then

$V_n \xrightarrow{D} N(0,1)$ (CLT)

$d_{Kol}(V_n, N) \leq \dfrac{C\,\mathbb{E}(|X_1|^3)}{\sqrt{n}}$ (Berry-Essen)

for $c > 0$ ind. of $\{X_i\}$

<u>Pf</u>: Recall we let

c_n be the best const. s.t.

$d_{Kol}(V_n, N) \leq \dfrac{c_n\,\mathbb{E}(|X_1|^3)}{\sqrt{n}}$

and found $c_n \leq \sqrt{n}$

Remains to show

$$\sup_n c_n = C.$$

We went onto define

$$h_{z,\varepsilon}(x) = \begin{cases} 1 & x < z-\varepsilon \\ \text{linear} & z-\varepsilon \le x \le z+\varepsilon \\ 0 & x > z+\varepsilon \end{cases}$$

linear slope is $-\dfrac{1}{2\varepsilon}$

$$d_{Kol}(V_n, N) \le \sup_{z\in\mathbb{R}} \left| \mathbb{E}(h_{z,\varepsilon}(V_n)) - \mathbb{E}(h_{z,\varepsilon}(N)) \right| + \frac{4\varepsilon}{\sqrt{2\pi}}$$

All done last lecture $\longrightarrow\!\!/\!\!/$

Assume we could prove

(*) $$\sup_z \left| \mathbb{E}(h_{z,\varepsilon}(V_n)) - \mathbb{E}(h_{z,\varepsilon}(N)) \right| \le \frac{6\,\mathbb{E}(|X_1|^3)}{\sqrt{n}} + \frac{3C_{n-1}\,\mathbb{E}(|X_1|^3)^2}{n\varepsilon}$$

For $\varepsilon = \frac{\mathbb{E}(|X_1|^3)\sqrt{c_{n-1}}}{\sqrt{n}}$ we have

$$d_{Kol}(V_n, N) \leq \frac{6\,\mathbb{E}(|X_1|^3)}{\sqrt{n}} +$$

$$\sqrt{c_{n-1}}\,\frac{3c_{n-1}\mathbb{E}(|X_1|^3)}{\sqrt{n}} + \frac{4\,\mathbb{E}(|X_1|^3)\sqrt{c_{n-1}}}{\sqrt{2\pi}\sqrt{n}}$$

$$= \frac{\mathbb{E}(|X_1|^3)}{\sqrt{n}}\left\{6 + \sqrt{c_{n-1}}\left(3 + \frac{4}{\sqrt{2\pi}}\right)\right\}$$

As c_n is the best constant we see

$$c_n \leq 6 + \sqrt{c_{n-1}}\left(3 + \frac{4}{\sqrt{2\pi}}\right)$$

Moreover $c_n \leq \sqrt{n}$ so

$$c_1 \leq 1 \leq 33$$

So $c_n \leq 33$ for $n=1$

Assume it holds at $n-1$

Then

$$c_n \leq 6 + \sqrt{33}\left(3 + \frac{4}{\sqrt{2\pi}}\right) \leq 33$$

So $\sup_n c_n \leq 33$

which then proves Berry-Essen. If we can show (*).

We will use Steins eq. associated with $h_{z,\varepsilon}$.

Let $f = f_{z,\varepsilon}$ denote the sol to Stein associated with $h_{z,\varepsilon}$.

As $h_{z,\varepsilon}$ is cont $f \in C^1$.

Also

$$\|f\|_\infty \leq \sqrt{\frac{\pi}{2}} \qquad \|f'\|_\infty \leq 2$$

Aux Result: Denote $\tilde{f}(x) = xf(x)$

then $|\tilde{f}(x) - \tilde{f}(y)|$

$$= |f(x)(x-y) + (f(x) - f(y))y|$$

$$\overset{MVT}{=} |f(x)(x-y) + f'(c)(x-y)y|$$

for $c \in (x,y)$

$$\leq \sqrt{\frac{\pi}{2}}|x-y| + 2|y||x-y|$$

$$= \left(\sqrt{\pi/2} + 2|y|\right)|x-y|$$

Now, let $V_n^i = V_n - \frac{X_i}{\sqrt{n}}$

and note $V_n^i \perp\!\!\!\perp X_i$.

$$\mathbb{E}(h(V_n)) - \mathbb{E}(h(N))$$

$$= \mathbb{E}(f'(V_n) - V_n f(V_n)) \text{ by Stein}$$

$$= \sum_{i=1}^{n} \mathbb{E}\left\{\frac{1}{n} f'(V_n) - f(V_n)\frac{X_i}{\sqrt{n}}\right\}$$

$$= \sum_{i=1}^{n} \mathbb{E}\left\{ \frac{1}{n} f'(V_n) - \left(f(V_n) - f(V_n^i)\right)\frac{X_i}{\sqrt{n}} \right\}$$

$$\mathbb{E}\left(f(V_n^i)\frac{X_i}{\sqrt{n}}\right)$$

$$= \frac{\mathbb{E}(f(V_n^i))\,\mathbb{E}(X_i)}{\sqrt{n}}$$

$$= 0$$

Focusing on this second term

$$\mathbb{E}\left(\frac{X_i}{\sqrt{n}}\left(f(V_n) - f(V_n^i)\right)\right)$$

$$= \mathbb{E}\left(\frac{X_i}{\sqrt{n}} \int_{V_n^i}^{V_n} f'(x)\,dx\right)$$

$$= \mathbb{E}\left(\frac{X_i}{\sqrt{n}} \int_0^1 (V_n - V_n^i)\, f'(V_n^i + y(V_n - V_n^i))\right)$$

$$dx = (V_n - V_n^i)\,dy \qquad y = \frac{x - V_n^i}{V_n - V_n^i}$$

Now notice

$$\int_0^1 f(x)\,dx = \mathbb{E}(f(\Theta)) \quad \Theta \sim \mathrm{Unif}(0,1).$$

$$\int_0^1 f(x)\,dx = \mathbb{E}(f(\Theta)) \quad \Theta \sim \mathrm{Unif}(0,1)$$

So we have

$$= \mathbb{E}\left(\mathbb{E}\left(\frac{X_i}{\sqrt{n}}(V_n - V_n^i)\int_0^1 f'(V_n^i + y(V_n - V_n^i))\,dy \,\middle|\, X_1,\dots,X_n\right)\right)$$

$$= \mathbb{E}\left[\frac{X_i^2}{n}\,\mathbb{E}\left(\int_0^1 \text{---}''\text{---}\,dy \,\middle|\, X_1,\dots,X_n\right)\right]$$

$$= \mathbb{E}\left(\frac{X_i^2}{n}\,\mathbb{E}\left(f'\left(V_n^i + \Theta\frac{X_i}{\sqrt{n}}\right) \,\middle|\, X_1,\dots,X_n\right)\right)$$

$$= \mathbb{E}\left(\frac{X_i^2}{n} f'\left(V_n^i + \Theta\frac{X_i}{\sqrt{n}}\right)\right)$$

for $\Theta \sim \mathrm{Unif}(0,1)$

So

$$\mathbb{E}(h(V_n)) - \mathbb{E}(h(N))$$

$$= \sum_{i=1}^{n} \mathbb{E}\left(\frac{1}{n} f'(V_n) - \frac{X_i^2}{n} f'\left(V_n^i + \Theta\frac{X_i}{\sqrt{n}}\right)\right)$$

Observe that $\Theta \perp\!\!\!\perp X_i$

and Recall that f is

sol to Stein so

$$f' = xf + h - \mathbb{E}(h(N))$$

$$= \tilde{f} + h(x) - \mathbb{E}(h(N))$$

So we can reform to get

$$= \sum_{i=1}^{n} \mathbb{E}\Big\{ \frac{1}{n}\tilde{f}(V_n) + \frac{1}{n}h(V_n) - \frac{1}{n}\mathbb{E}\cancel{(h(N))}$$

$$- \frac{X_i^2}{n}\tilde{f}\Big(V_n^i + \frac{\sigma X_i}{\sqrt{n}}\Big) - \frac{X_i^2}{n}h\Big(V_n^i + \sigma\frac{X_i}{\sqrt{n}}\Big)$$

$$+ \frac{X_i^2}{n}\cancel{\mathbb{E}(h(N))}\Big\} \quad \mathbb{E}(X_i^2) = 0$$

$$= \sum_{i=1}^{n} \mathbb{E}\Big\{ \frac{1}{n}\tilde{f}(V_n) \underbrace{- \frac{1}{n}\tilde{f}(V_n^i) + \frac{X_i^2}{n}\cancel{\tilde{f}(V_n^i)}}_{\text{I same}}$$

$$+ \frac{1}{n}h(V_n) \underbrace{- \frac{1}{n}h(V_i^i) + \cancel{\frac{X_i^2}{n}h(V_n^i)}}_{\text{I same}}$$

$$- \frac{X_i^2}{n}\tilde{f}\Big(V_n^i + \sigma\frac{X_i}{\sqrt{n}}\Big) \underbrace{+ \frac{X_i^2}{n}\tilde{f}(V_n^i)}$$

$$\underbrace{- \frac{X_i^2}{n}\cancel{\tilde{f}(V_n^i)}}_{\text{I}}$$

$$-\frac{X_i^2}{n} h\left(U_n^i+\theta\frac{X_i}{\sqrt{n}}\right)+\underbrace{\frac{X_i^2}{n} h(U_n^i)}_{\pm}$$

$$\left.-\underbrace{\cancel{\frac{X_i^2}{n} h(U_n^i)}}\right\}$$

$$=\sum_{i=1}^{n}\left\{a_i(\tilde{f})-b_i(\tilde{f})+a_i(h)-b_i(h)\right\}$$

$$a_i(g)=\frac{1}{n}\mathbb{E}\left(g(U_n)-g(U_n^i)\right)$$

$$b_i(g)=\mathbb{E}\left(\frac{X_i^2}{2}\left(g\left(U_n^i+\theta\frac{X_i}{\sqrt{n}}\right)-g(U_n^i)\right)\right)$$

All we need is bounds on these four terms.

<u>First term:</u> $a_i(\tilde{f})\leq\frac{1}{n}\mathbb{E}\left|\tilde{f}(U_n)-\tilde{f}(U_n^i)\right|$

$$\leq\frac{1}{n}\mathbb{E}\left(\left(\sqrt{\frac{\pi}{2}}+2|U_n^i|\right)\left|\frac{X_i}{\sqrt{n}}\right|\right)$$

$$= \frac{1}{n\sqrt{n}} \mathbb{E}\left(|X_i| \sqrt{\pi/2} + 2 |X_i| \, |U_n^i| \right)$$

$$= \frac{1}{n\sqrt{n}} \left(\sqrt{\frac{\pi}{2}} \mathbb{E}|X_i| + 2 \, \mathbb{E}|X_i| \, \mathbb{E}|U_n^i| \right)$$

But $\mathbb{E}(|X_i|) \leq \mathbb{E}(|X_i|^2)^{1/2} = 1$

$$\mathbb{E}(|U_n^i|) \leq \left(\mathbb{E}(U_n^i)^2\right)^{1/2}$$

$$= \mathbb{E}\left(\frac{1}{n}\Big(\sum_{i \neq j} X_j\Big)^2\right)^{1/2} = \mathbb{E}\left(\frac{n-1}{n}\right) \leq 1$$

So $|a_i(\tilde{f})| \leq \frac{1}{n\sqrt{n}} \left(\sqrt{\frac{\pi}{2}} + 2 \right)$

<u>Second Term</u>:

$$|b_i(\tilde{f})| \leq \mathbb{E}\left(\left| \frac{X_i^2}{n} \tilde{f}(U_n^i) \right| + \left| \frac{G X_i}{\sqrt{n}} \tilde{f}(U_n^i) \right| \right)$$

$$= \mathbb{E}\left(\frac{X_i^2}{n} |G| \Big/ \frac{X_i}{\sqrt{n}} \right) \left(\sqrt{\frac{\pi}{2}} + 2 |U_n^i| \right)$$

$$= \mathbb{E}\left(\frac{X_i}{n} |\theta| \left|\frac{X_i}{\sqrt{n}}\right|\right)\left(\sqrt{\frac{\pi}{2}} + 2|V_n'|\right)$$

$$\leq \frac{1}{n\sqrt{n}} \mathbb{E}\left(|X_i|^3 |\theta|\right)\sqrt{\frac{\pi}{2}} + \frac{1}{n\sqrt{n}} \mathbb{E}\left(2|X_i|^3|\theta||V_n'|\right)$$

$$= \frac{\mathbb{E}(|X_i|^3)}{n\sqrt{n}}\left[\frac{1}{2}\sqrt{\pi/2} + 1\right]$$

Third Term:

$$h(x) - h(y) = \int_y^x h'(u)\,du$$

$$= \int_0^1 (x-y)\underbrace{h'(y + t(x-y))}_{-1/2\varepsilon}\,dt$$

$$= -\frac{(x-y)}{2\varepsilon}\int_0^1 \mathbb{1}_{(z-\varepsilon, z+\varepsilon)}(y + t(x-y))\,dt$$

$$= -\frac{(x-y)}{} \mathbb{E}\left(\mathbb{1}_{(z-\varepsilon, z+\varepsilon)}(y + \hat{\theta}(x-y))\right)$$

$$\cdots \frac{\cdots}{2\varepsilon}\,\mathbb{E}\left(\mathbb{1}_{(z-\varepsilon,\,z+\varepsilon)}(\cdots)\right)$$

$$|a_i(n)| \le \frac{1}{n}\,\mathbb{E}\left(|h(U_n)-h(U_n^i)|\right)$$

$$\le \frac{1}{n}\,\mathbb{E}\left(\frac{|U_n-U_n^i|}{2\varepsilon}\,\mathbb{1}_{(z-\varepsilon,\,z+\varepsilon)}\left(U_n^i+\hat{\theta}\left(X_i/\sqrt{n}\right)\right)\right)$$

$$= \frac{1}{n}\,\mathbb{E}\left(\frac{|X_i|}{2\varepsilon\sqrt{n}}\,\mathbb{P}\left(z-\varepsilon \le U_n^i + t\frac{x}{\sqrt{n}} \le z+\varepsilon \;\middle|\; t=\hat{\theta},\, x=X_i\right)\right)$$

$$\le \frac{1}{2\varepsilon n\sqrt{n}}\,\mathbb{E}\left(|X_i|\right)\sup_{t\in[0,1]}\sup_{x\in\mathbb{R}}\mathbb{P}\left(z-\varepsilon-\frac{tx}{\sqrt{n}} \le U_n^i \le z+\varepsilon-\frac{tx}{\sqrt{n}}\right)$$

$$\le \frac{1}{2\varepsilon n\sqrt{n}}\,\sup\sup\,\mathbb{P}\left(\text{——}\,''\,\text{——}\right)$$

Lecture 3/22

Thursday, March 22, 2018 12:33 PM

Recall we obtained

$$\sum_{i=1}^{n} \left\{ \underline{a_i(\tilde{f})} - \underline{b_i(\tilde{f})} + a_i(h) - b_i(h) \right\}$$

$\checkmark \qquad \checkmark$

$$a_i(g) = \frac{1}{n} \mathbb{E}\left(g(V_n) - g(V_n^i) \right)$$

$$b_i(g) = \mathbb{E}\left(\frac{X_i^2}{n} \left(g\left(V_n^i + \Theta \frac{X_i}{\sqrt{n}}\right) - g(V_n^i) \right) \right)$$

$$h(x) - h(y) = -\frac{x-y}{2\varepsilon} \mathbb{E}\left(\mathbb{1}_{(z-\varepsilon, z+\varepsilon)}(y + \hat{\Theta}(x-y)) \right)$$

$\Downarrow$

$$|a_i(h)| \leq \frac{1}{n} \mathbb{E}\left(\frac{|X_i|}{2\varepsilon\sqrt{n}} \mathbb{1}_{(z-\varepsilon, z+\varepsilon)}\left(V_n^i + \Theta \frac{X_i}{\sqrt{n}}\right) \right)$$

$$= \frac{1}{2\varepsilon n\sqrt{n}} \mathbb{E}\left(|X_i| \, \mathbb{P}\left(z - \varepsilon - \frac{tx}{\sqrt{n}} \leq V_n^i \leq z + \varepsilon - \frac{t}{\sqrt{n}} \right) \right)$$

$$\Big|_{\substack{t = \hat{\varepsilon} \\ x = x_i}}$$

$$\le \frac{\mathbb{E}|X_i|}{2\varepsilon u\sqrt{n}} \sup_t \sup_x \mathbb{P}\left(z-\varepsilon - \frac{tx}{\sqrt{n}} \le V_n^i \le \#\right)$$

$\mathbb{E}|X_i| \le 1$ by C.S.

Let $a, b \in \mathbb{R}$ we need to bound $\mathbb{P}(a \le V_n^i \le b)$

Set

$$\tilde{V}_n^i = \frac{1}{\sqrt{n-1}} \sum_{j \ne i}^{n} X_j$$

$$V_n^i = \frac{1}{\sqrt{n}} \sum_{j \ne i}^{n} X_j = \frac{\sqrt{n-1}}{\sqrt{n}} \tilde{V}_n^i$$

$$\mathbb{P}(a \le V_n^i \le b)$$

$$= \mathbb{P}\left(\frac{a}{\sqrt{\frac{n-1}{n}}} \le \tilde{V}_n^i \le \frac{b}{\sqrt{\frac{n-1}{n}}}\right)$$

$$= \mathbb{P}\left(\frac{a}{\sqrt{n-1}} \le N \le \frac{b}{\sqrt{n-1/n}}\right)$$

$$= \mathbb{P}\left(\frac{a}{\sqrt{n-1/n}} \le N \le \frac{\quad}{\sqrt{n-1/n}} \right)$$

$$\left.\begin{array}{l} - \mathbb{P}\left(\frac{a}{\sqrt{n-1/n}} \le N \le \frac{b}{\sqrt{n-1/n}} \right) \\ + \mathbb{P}\left(\frac{a}{\sqrt{\frac{n-1}{n}}} \le \hat{V}_n^i \le \frac{b}{\sqrt{n-1/n}} \right) \end{array}\right\} \text{almost Kol Distance}$$

$$\le \mathbb{P}\left(\frac{a}{\sqrt{n-1/n}} \le N \le \frac{b}{\sqrt{n-1/n}} \right)$$

$$+ 2\, d_{KoL}(\hat{V}_n^i, N)$$

$$\le \frac{b-a}{\sqrt{2\pi \frac{n-1}{n}}} + 2\, d_{KoL}(\hat{V}_n^i, N)$$

Using Berry-Essen bound with c_{n-1}

$$\le \frac{b-a}{\sqrt{2\pi \frac{n-1}{n}}} + \frac{2 c_{n-1} \mathbb{E}(|X_1|^3)}{\sqrt{n-1}}$$

Hence we can write

$$|a_i(h)| \le$$

$$\frac{1}{2\varepsilon n\sqrt{n}}\left[\frac{2\varepsilon}{\sqrt{2\pi}\sqrt{n-1/n}} + \frac{2C_{n-1}\mathbb{E}(|X_1|^3)}{\sqrt{n-1}}\right]$$

$$\le \frac{1}{\sqrt{2\pi}\, n\sqrt{n-1}} + \frac{C_{n-1}\,\mathbb{E}(|X_1|^3)}{\varepsilon\, n\sqrt{n}\sqrt{n-1}}$$

Fourth Term:

$$|b_i(h)| \le \mathbb{E}\left(\frac{X_i^2}{n}\frac{|\theta||X_i|}{2\varepsilon\sqrt{n}}\mathbb{1}_{(z-\varepsilon,z+\varepsilon)}\left(U_n^i + \hat{\theta}\sigma\frac{X_i}{\sqrt{n}}\right)\right)$$

$$= \frac{1}{2\varepsilon n\sqrt{n}}\,\mathbb{E}(|X_i|^3|\theta|\,\mathbb{P}\left(U_n^i\in\left(z\pm\varepsilon-\frac{tx}{n}\right)\right)\Bigg|_{\substack{t=\hat{\theta}\sigma\\ x=X_i}}$$

$$\le \frac{\mathbb{E}|X_i|^3}{4\varepsilon n\sqrt{n}}\sup_{t,x}\mathbb{P}\left(U_n^i\in\left(z\pm\varepsilon-\frac{tx}{n}\right)\right)$$

$$\le \frac{\mathbb{E}(X_i|^3}{4\varepsilon n\sqrt{n}}\left[\frac{2\varepsilon}{\sqrt{2\pi}\sqrt{\frac{n-1}{n}}} + \frac{2C_{n-1}\mathbb{E}(X_1^3)}{\sqrt{n-1}}\right.$$

$$= \frac{\mathbb{E}(|X_1|^3)}{2\sqrt{2\pi}\, n\sqrt{n-1}} + \frac{c_{n-1}\left(\mathbb{E}|X_1|^3\right)^2}{2\varepsilon\, n\sqrt{n}\,\sqrt{n-1}}$$

Taking all of these bounds together and summing over n we get (*). ∎

Law of Iterated Logarithm

Let $\{X_n : n \geq 1\}$ be a seq of iid R.V. with $\mathbb{E}(X_n) = 0$ and $\mathbb{E}(X_1^2) = 1$

Then

$$\frac{S_n}{n} \xrightarrow{a.s.} 0$$

n

$$\frac{S_n}{\sqrt{n}} \xrightarrow{D} Z$$

Then for any $c>0$

by Kol. 0-1 law

$$\mathbb{P}\left(\limsup \frac{S_n}{\sqrt{n}} > c\right) = 0 \text{ or } 1$$

$$\mathbb{P}\left(\limsup \frac{S_n}{\sqrt{n}} > c\right)$$

$$\geq \limsup \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > c\right) = \mathbb{P}(N>c) > 0$$

So $\mathbb{P}\left(\limsup \frac{S_n}{\sqrt{n}} > c\right) = 1$

$$\Rightarrow \limsup \frac{S_n}{\sqrt{n}} = +\infty \text{ (a.s.)}$$

Similarly

$$\liminf \frac{S_n}{\sqrt{n}} = -\infty \quad (a.s.)$$

Note: $\frac{1}{n}$ erases fluctuation

$\frac{1}{\sqrt{n}}$ doesn't erase any fluctuation

Both uninformative. So we seek $f(n)$

$$\limsup \frac{S_n}{f(n)} = 1$$

$$\liminf \frac{S_n}{f(n)} = -1$$

$$\boxed{\text{What function flucuates like } S_n\text{?}}$$

$$\mathbb{P}\left(\left|\frac{S_n}{\sqrt{2n\log\log n}}\right| > \varepsilon\right)$$

$$\leq \frac{n}{\varepsilon^2 2n\log\log n} = \frac{1}{2\varepsilon^2 \log\log n} \mapsto 0$$

So with high probability

we can't get crazy

fluctuations.

<u>Thrm:</u> (Hartman - Wintner LIL)

Let $\{X_n : n \geq 1\}$ be an iid R.V.

with $\mathbb{E}(X_1) = 0$ $\mathbb{E}(X_1^2) = 1$

Then

$$\limsup \frac{S_n}{\sqrt{2n\log\log n}} = 1 \quad \text{a.s.}$$

(*)

$$\liminf \frac{S_n}{\quad} = -1 \quad \text{a.s.}$$

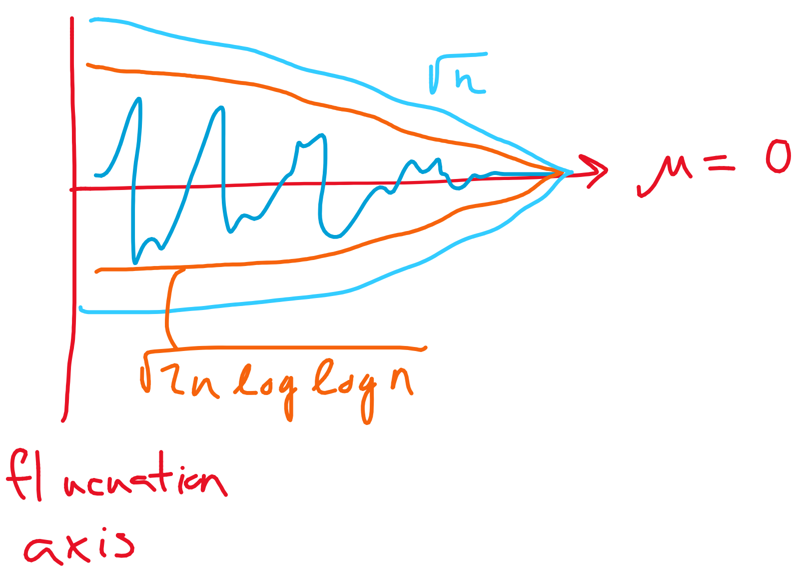
$$\liminf \frac{S_n}{\sqrt{2n\log\log n}} = -1 \quad \text{a.s.}$$

Conversly if

$$P\left(\limsup \frac{S_n}{\sqrt{2n\log\log n}} < \infty\right) > 0$$

then $\mathbb{E}(X_1) = 0$ $\mathbb{E}(X_1^2) = 1$

and (*) holds



Lecture 3/27

Tuesday, March 27, 2018 12:30 PM

Martingales

- Central concept in probability Theory
- Generalizations of centered random walks

Review of conditional

Exp.

Consider the triple $(\Omega, \mathcal{F}, \mathbb{P})$ with sub σ-algebra $\mathcal{G} \subseteq \mathcal{F}$

Def: Let $X \in L^1(\Omega)$. Then the conditional exp. of X w.r.t. $\mathcal{G}$ is any R.V.

X w.r.t. $\mathcal{G}$ is any r.v.

$Z \in \mathcal{G}$ s.t. $Z \in L^1(\Omega)$ s.t.

$$\int_A Z \, d\mathbb{P} = \int_A X \, d\mathbb{P} \quad \forall A \in \mathcal{G}$$

In which case

$$Z = \mathbb{E}(X \mid \mathcal{G})$$

Rmk: Z is $\mathcal{G}$ measurable as X is $\mathcal{F}$ measurable

Rmk: We could rewrite the equality as

$$\mathbb{E}(Z \mathbb{1}_A) = \mathbb{E}(X \mathbb{1}_A)$$

Prop: The cond. exp. is a.s. unique.

a.s. unique.

<u>Pf</u>: If Y, Z are cond. exp. of X wrt $\mathcal{G}$ then

$$\int_A (Y-Z)\,d\mathbb{P} = 0 \quad \forall A \in \mathcal{G}$$

In particular take

$$A^* = \{\omega \in \Omega : Y(\omega) > Z(\omega)\}$$

Hence

$$\int_{A^*} \underbrace{(Y-Z)}_{>0}\,d\mathbb{P} = 0$$

$$\Rightarrow \mathbb{P}(A^*) = 0$$

By a symmetric argument

$$\mathbb{P}(\{\omega\in\Omega : Y(\omega)<Z(\omega)\})=0$$

So $Y \overset{a.s.}{=} Z$. $\blacksquare$

<u>Rmk</u>: We know they are unique but do they exist?

For existance we need def and a thrm.

<u>Def</u>: Let $\mathbb{Q}, \mathbb{P}$ be measures. We say $\mathbb{Q}$ is <u>abs. cont</u> wrt $\mathbb{P}$ iff

$$\mathbb{P}(A)=0 \Rightarrow \mathbb{Q}(A)=0 \quad \forall A\in\mathcal{F}$$

and we write $\mathbb{Q}\ll\mathbb{P}$

Thrm: (Radon - Nikodym)

Let $\mathbb{Q}$ be a finite measure with $\mathbb{Q} \ll \mathbb{P}$ then there exists a R.V. $Y \in L^1(\Omega)$ s.t.

$$\mathbb{Q}(A) = \int_A Y \, d\mathbb{P} \quad \forall A \in \mathcal{F}$$

We denote $Y = \frac{d\mathbb{Q}}{d\mathbb{P}}$

and call it the Radon-Nikodym derivative.

Thrm: (existance of cond. exp)

Let $X \in L^1(\Omega)$ and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ-alg.

Then there exists a a.s. unique r.v. $Z \in \mathcal{G}$ s.t.

unique r.v. $Z \in \mathcal{G}$ s.t.

$$\int_A Z \, d\mathbb{P} = \int_A X \, d\mathbb{P} \quad \forall A \in \mathcal{G}.$$

Pf: (we already showed a.s. uniqueness). Let $\mathbb{Q}$ be a measure on $\mathcal{G}$ defined by

$$\mathbb{Q}(A) = \int_A X \, d\mathbb{P} \quad \forall A \in \mathcal{G}$$

$\mathbb{Q}(A)$ is finite as $\mathbb{P}$ is a prob measure and $X \in L^1(\Omega)$.

To use Radon-Nikodym

thrm we see that

$$\mathbb{Q} \ll \mathbb{P}|_{\mathcal{G}}$$

So we can say that

there exists a R.V.

$Y \in \mathcal{G}$, $Y \in L^1(\Omega)$ s.t.

$$\mathbb{Q}(A) = \int_A Y \, d\mathbb{P}|_{\mathcal{G}} \quad \forall A \in \mathcal{G}$$

But recall by def.

of $\mathbb{Q}$

$$\mathbb{Q}(A) = \int_A X \, d\mathbb{P} = \int_A Y \, d\mathbb{P} \quad \forall A \in \mathcal{G}$$

which is the defining

relation of the conditional

expectation.

$$Y = \mathbb{E}(X|\mathcal{G}) \left(= \frac{d\mathbb{Q}}{d\mathbb{P}|_{\mathcal{G}}} \right)$$

$\blacksquare$

<u>Rmk</u>: Observe that

$$\int_A \underbrace{\left(X - \mathbb{E}(X|\mathcal{G})\right)}_{=0 \text{ on } \mathcal{G}} Z \, d\mathbb{P} = 0$$

$\forall Z \in \mathcal{G}$ bounded $\forall A \in \mathcal{G}$

For $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$

$Z \in L^2(\Omega, \mathcal{G}, \mathbb{P})$

We get that

$X - \mathbb{E}(X|\mathcal{G})$ is orthogonal

to Z in $L^2(\Omega, \mathcal{G}, \mathbb{P})$

So $X - \mathbb{E}(X|\mathcal{G}) \in L^2(\Omega, \mathcal{G}, \mathbb{P})^{\perp}$

orthogonal compliment

As

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) = L^2(\Omega, \mathcal{G}, \mathbb{P}) \oplus L^2(\Omega, \mathcal{G}, \mathbb{P})^{\perp}$$

So we can write X uniquely as

$$X = \mathbb{E}(X|\mathcal{G}) + (X - \mathbb{E}(X|\mathcal{G}))$$

orthongonal projection of X onto $L^2(\Omega, \mathcal{G}, \mathbb{P})$

Main Properties of Cond Exp

- $\mathbb{E}(X|\mathcal{G}) \overset{a.s.}{=} X \quad \forall \quad X \in \mathcal{G}$
- $\mathbb{E}(XY|\mathcal{G}) \overset{a.s.}{=} Y\,\mathbb{E}(X|\mathcal{G}) \quad \forall Y \in \mathcal{G}$

  s.t. $XY \in L^1(\Omega)$
- $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X) \quad \forall \quad X \perp\!\!\!\perp \mathcal{G}$
- $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$

If $\mathcal{G}_1 \subseteq \mathcal{G}_2$ then

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G}_2)|\mathcal{G}_1) = \mathbb{E}(X|\mathcal{G}_1)$$

$$= \mathbb{E}(\mathbb{E}(X|\mathcal{G}_1)|\mathcal{G}_2)$$

think projections here.

Martingales

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a prob. space. and consider the increasing seq. of sub σ-alg. of $\mathcal{F}$

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$$

$\{\mathcal{F}_n\}_{n=0}^{\infty}$ is a <u>fitration</u> of $\mathcal{F}$.

<u>Rmk</u>: n can be continuous index.

<u>Def</u>: A seq. of R.V. $\{X_n\}$ is <u>adapted to the fitration</u> $\{\mathcal{F}_n\}$ iff $X_n \in \mathcal{F}_n \ \forall n$

In particular if

$\mathcal{F}_n = \sigma\{X_0, \dots, X_n\}$ $\forall n$

then X_n is automatically adated to $\mathcal{F}_n$.

We call $\mathcal{F}_n$ the natural filration of X_n

Def: A sequence X_n is called $\mathcal{F}_n$-predictable if

$X_n \in \mathcal{F}_{n-1}$ $\forall n \geq 1$

Def: An increasing process $\{A_n; n \geq 1\}$ is a seq. of R.V. s.t.

(i) $A_0 = 0$

(ii) A_n increasing

(ii) A_n increasing

(iii) $\mathcal{F}_n$ - predictable

<u>Def</u>: An integrable $\mathcal{F}_n$ adapted seq. $\{X_n\}$ of R.V. is a

(i) <u>martingale</u> if

$$\mathbb{E}(X_{n+1} / \mathcal{F}_n) = X_n \ \forall n$$

(ii) <u>submartingale</u> if

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) \geq X_n \quad \forall n$$

(iii) <u>supermartingale</u>. if

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) \leq X_n$$

Lecture 3/29

Thursday, March 29, 2018 12:30 PM

Martingales

Def: An itegrable $\{\mathcal{F}_n\}$-adapted sequence $\{X_n; n \geq 1\}$ is a martingale w.r.t. $\{\mathcal{F}_n : n \geq 0\}$ if

$$\mathbb{E}(X_{n+1} / \mathcal{F}_n) = X_n \quad a.s.$$

$\{X_n\}$ submartingale

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) \geq X_n \quad a.s.$$

supermartingale

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) \leq X_n \quad a.s.$$

Def: $\{X_n\}$ is an L^P-martingale if $\mathbb{E}(|X_n|^P) < \infty \quad \forall n \geq 0$

if $\mathbb{E}(|X_n|^p) < \infty$ $\forall n \geq 0$

Moreover, $\{X_n\}$ is <u>L^p-bounded</u>

if $\sup_{n \geq 0} \mathbb{E}(|X_n|^p) < \infty$

<u>Rmk</u>: All equalities for martingales are in the almost sure sense based on def of $\mathbb{E}(X_{n+1} | \mathcal{F}_n)$

<u>Rmk</u>: Martingale's must be a sequence. Must define both the fitration and the sequence.

$\{\mathcal{F}_n\}$ $\{X_n\}$

<u>Def</u>: If $\{U_n : n \geq 0\}$ is $\{\mathcal{F}_n\}$-adapted

is $\{\mathcal{F}_n\}$ -adapted

Seq. with

$\mathbb{E}(U_{n+1} | \mathcal{F}_n) = 0$

then $\{U_n\}$ is a

martingale difference

Rmk: Same "sub" "super"

definitions

Rmk: Think of this as

incriments

Ex: (Canonical) Let

$\{Y_n : n \geq 1\}$ be integrable,

centered random varibles.

Define $X_n = \sum^n Y_k \quad n \geq 0$

Define $X_n = \sum_{k=1}^{n} Y_k \quad n \geq 0$

Then $\{X_n\}$ is a martingale with respect to the natural filtration.

Pf: Y_n integrable $\Rightarrow$ X_n integrable. Furthermore

$$\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) = \mathbb{E}\left(\sum_{k=1}^{n+1} Y_k \mid \mathcal{F}_n\right)$$

As $Y_{n+1} \perp\!\!\!\perp \mathcal{F}_n$ we have

$$= \mathbb{E}\left(Y_{n+1} \mid \mathcal{F}_n\right) + \mathbb{E}\Big(\underbrace{\sum_{k=1}^{n} Y_k}_{\in \mathcal{F}_n} \mid \mathcal{F}_n\Big)$$

Think Projection

$$= \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) + X_n$$

$$= \mathbb{E}(Y_{n+1}) + X_n$$

$$= 0 + X_n = X_n$$

So we see that symmetric R.W. are in fact Martingales.

Characterization of Martingales

An integrable $\{\mathcal{F}_n\}$-adapted seq. $\{X_n\}$ is a martingale w.r.t. $\{\mathcal{F}_n\}$ iff $\mathbb{E}(X_n|\mathcal{F}_m) = X_m$ (i)

$\forall\, 0 \leq m \leq n$

Furthermore

$$\mathbb{E}(\mathbb{E}(X_n|\mathcal{F}_m)) = \mathbb{E}(X_m)$$

$$\mathbb{E}(X_n) = \mathbb{E}(X_m) \quad \text{(ii)}$$

Constant mean

Pf: HW #8

Rmk: From a game perspective,

(ii) $\Rightarrow$ fairness of a game

(i) $\Rightarrow$ future expcted State of winings is the current state

Thm: Let $\{U_n\}$ be an int. $\{\mathcal{F}_n\}$ adapted seq. and let

$$X_n = \sum_{1}^{n} U_k$$

$$X_n = \sum_{k=0}^{n} U_k$$

(i) $\{X_n\}$ is a martingale iff $\{U_n\}$ is a martingale diff seq.

(ii) A martingale diff seq. has constant expectation $(=0)$

<u>Pf</u>: HW #8

<u>Ex</u>:

1. (Compensated Square of Mart.)

Let $\{Y_n\}$ be square-int. $\{\mathcal{F}_n\}$ adapted seq. with mean zero. We saw that mean zero. We saw that

$$\left\{ X_n = \sum_{k=1}^{n} Y_k \right\} \text{ is } \{\mathcal{F}_n\}\text{-mart.}$$

Let $\sigma_n^2 = \mathrm{Var}(Y_n)$

$s_n^2 = \sum_{k=1}^{n} \sigma_k^2$ then

$\{X_n^2 - s_n^2 : n \geq 0\}$ is

an $\{\mathcal{F}_n\}$-mart.

2. (Gambler's Martingale)

Let $X_0 = 1$. and

$$X_{n+1} = \begin{cases} 2X_n & p = 1/2 \\ 0 & p = 1/2 \end{cases}$$

$$\mathbb{P}(X_n = 2^n) = p^n = (1/2)^n$$

$$\mathbb{P}(X_n = 0) = 1 - p^n = 1 - \frac{1}{2^n}$$

So we can write

$$X_n \overset{D}{=} \prod_{k=1}^{n} Y_k , \quad Y_k = \begin{cases} 2 & p = 1/2 \\ 0 & p = 1/2 \end{cases}$$

X_n int b/c Y_k bounded

natural filtration

$$\mathbb{E}(X_n \mid \mathcal{F}_m) \qquad m \leq n$$

$$= \mathbb{E}\left(\prod_{k=1}^{n} Y_k \mid \mathcal{F}_m \right)$$

$$= \mathbb{E}\left(\underbrace{\prod_{k=1}^{m} Y_k}_{\in \mathcal{F}_m} \underbrace{\prod_{j=m+1}^{n} Y_k}_{\perp\!\!\!\perp \mathcal{F}_m} \mid \mathcal{F}_m \right)$$

$$= X_m \, \mathbb{E}\left(\prod_{k=m+1}^{n} Y_k \right)$$

$$= X_m \prod_{k=m+1}^{n} \mathbb{E}(Y_k)$$

$= X_m$

So this game is fair.

3) Martingales as Cond. exp.

Let $Z \in L^1(\Omega)$ For each $n \geq 0$ define

$X_n = \mathbb{E}(Z / \mathcal{F}_n)$ for a given filtration.

$$\mathbb{E}(|X_n|) = \mathbb{E}(|\mathbb{E}(Z \mid \mathcal{F}_n)|)$$

$$\leq \mathbb{E}(\mathbb{E}(|Z| \mid \mathcal{F}_n))$$

$$= \mathbb{E}(|Z|) < \infty \quad \forall n \geq 0$$

So it is integrable.

For $m \leq n$

$$\mathbb{E}(X_n \mid \mathcal{F}_m)$$

$$= \mathbb{E}(\mathbb{E}(Z \mid \mathcal{F}_n) \mid \mathcal{F}_m)$$

$$= \mathbb{E}(Z \mid \mathcal{F}_m) = X_m$$

Lecture 4/3

Tuesday, April 3, 2018 12:33 PM

Orthogonality

Thrm: Let $\{X_n, n \geq 1\}$ be a martingale wrt $\{\mathcal{F}_n\}$ and square integrable. Let $\{U_n, n \geq 1\}$ be its difference seq. Then

(i) $\mathbb{E}(U_n U_m) = \mathbb{1}_{\{n=m\}} \mathbb{E}(X_n^2)$

orthogonal in $L^2(\Omega)$

(ii) For $m < n$

$$\mathbb{E}[U_n X_m] = 0$$

and

$$\mathbb{E}(X_n X_m) = \mathbb{E}(X_m^2)$$

(iii) $\mathbb{E}\left[(X_n - X_m)^2\right] = \mathbb{E}(X_n^2) - \mathbb{E}(X_m^2)$

Pf: HW #8

Decompositions

Thrm: (Doob Decomp.) Any $\{\mathcal{F}_n\}$-submartingal $\{X_n, n \geq 1\}$ can be decomposed as the sum of a martingale $\{M_n, n \geq 1\}$ and an increasing process $\{A_n, n \geq 1\}$. That is

$$X_n = M_n + A_n \quad \forall n \in \mathbb{N}.$$

unique

unique
as well.

Pf: Let $M_0 = X_0$ so by necessity $A_0 = 0$. Now let

$$M_n = \sum_{k=1}^{n} \left[X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) \right]$$

and

$$A_n = X_n - M_n .$$

First step: $\{M_n, n \geq 1\}$ is an $\{\mathcal{F}_n, n \geq 1\}$ martingale.

(a) M_n is integrable as X_n is integrable

(b) $\mathbb{E}(M_{n+1} \mid \mathcal{F}_n)$

$$= \mathbb{E}\left[\sum_{k=1}^{n+1} \left[X_k - \mathbb{E}(X_k / \mathcal{F}_{k-1}) \right] \Big| \mathcal{F}_n \right]$$

$$= \mathbb{E}\left(\sum_{k=1} [X_k - \mathbb{E}(X_k / \mathcal{F}_{k-1})] \Big| \mathcal{F}_n \right]$$

$$= \mathbb{E}\Big(X_{n+1} - \underbrace{\mathbb{E}(X_{n+1} | \mathcal{F}_n)}_{\in \mathcal{F}_n} \Big/ \mathcal{F}_n \Big)$$

$$+ \sum_{k=1}^{n} [X_k - \mathbb{E}(X_k | \mathcal{F}_{k-1})]$$

$$= \mathbb{E}(X_{n+1} / \mathcal{F}_n) - \mathbb{E}(X_{n+1} | \mathcal{F}_n)$$

$$+ M_n$$

$$= M_n$$

Step 2: $\{A_n, n \geq 1\}$ increasing process.

(a) $A_0 = 0$ ✓

(b) $A_n \nearrow$

(c) A_n $\mathcal{F}_n$-predictable

Starting with predictability

$$A_n = X_n - M_n$$

$$= X_n - \sum_{k=1}^{n} [X_k - \mathbb{E}(X_k / \mathcal{F}_{k-1})]$$

$$= \underbrace{\mathbb{E}(X_n \mid \mathcal{F}_{n-1})}_{\in \mathcal{F}_{n-1}} - \underbrace{\sum_{k=1}^{n-1} [X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1})]}_{\in \mathcal{F}_{n-1}}$$

Hence $A_n \in \mathcal{F}_{n-1}$ so the process is indeed predictable.

For the increasing argument:

$$A_{n+1} - A_n \overset{\text{predictability}}{=} \mathbb{E}(A_{n+1} - A_1 \mid \mathcal{F}_n)$$

$$= \mathbb{E}[(X_{n+1} - X_n) - (M_{n+1} - M_n) \mid \mathcal{F}_n]$$

$$= \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) - \underbrace{\mathbb{E}(M_{n+1} \mid \mathcal{F}_n)}_{\text{Martingale}}$$
$$- X_n + M_n$$

$$= \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) - X_n$$
$$- M_n + M_n$$

$$= \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) - X_n$$

≥ 0

X_n submartingale.

Therefore there exists a decomposition.

Uniqueness:

Assume $X_n = M_n' + A_n'$ is another decomposition

Then

$$A_{n+1}' - A_n' = \mathbb{E}(A_{n+1}' - A_n' \mid \mathcal{F}_n)$$

$$= \mathbb{E}(X_{n+1} - M_{n+1}' - X_n + M_n' \mid \mathcal{F}_n)$$

$$= \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) - X_n$$

$$= A_{n+1} - A_n$$

Taking $n=0$ with $A_0' = A_0 = 0$

$A_1' = A_1$ then applying

$A_1' = A_1$ then applying this recursively we see

$$A_n = A_n' \quad \forall n \geq 0$$

Then

$$M_n' = X_n - A_n' = X_n - A_n = M_n .$$

$\blacksquare$

<u>Thrm</u>: Let $\{X_n, n \geq 0\}$ be a $\mathcal{F}_n$ mart. s.t.

$$\sup_{n \geq 0} \mathbb{E}[X_n^+] < \infty$$

then there exists two non-negative mart. $\{M_n^{(1)}\}$ $\{M_n^{(2)}\}$ s.t.

$$X_n = M_n^{(1)} - M_n^{(2)}, \quad n \geq 0$$

Stopping Times

<u>Def:</u> A positive, integer valued, possibly infinite R.V. τ is called a <u>Stopping time</u> w.r.t. $\{\mathcal{F}_n\}_{n=1}^{\infty}$ if

$$\{\tau = n\} \in \mathcal{F}_n \quad \forall n$$

<u>Rmk:</u> As $\{\tau \leq n\} = \bigcup_{k=0}^{n} \{\underbrace{\tau = k}_{\substack{\in \mathcal{F}_k \\ \subseteq \mathcal{F}_n}}\}$

So $\{\tau \leq n\} \in \mathcal{F}_n$

$$\{\tau = n\} = \underbrace{\{\tau \leq n\}}_{\in \mathcal{F}_n} \setminus \underbrace{\{\tau \leq n-1\}}_{\in \mathcal{F}_{n-1}}$$

So we have a two way equivalence

equivalence

$$\{\tau > n\} = \left(\bigcup_{k=0}^{n} \{\tau = k\}\right)^c \in \mathcal{F}_n$$

- Decide at each time with <u>certainty</u> if the stopping time is now or in the future.

<u>Ex:</u> R.W. with first entrance time $\tau = \min\{n \geq 0 : X_n = 5\}$

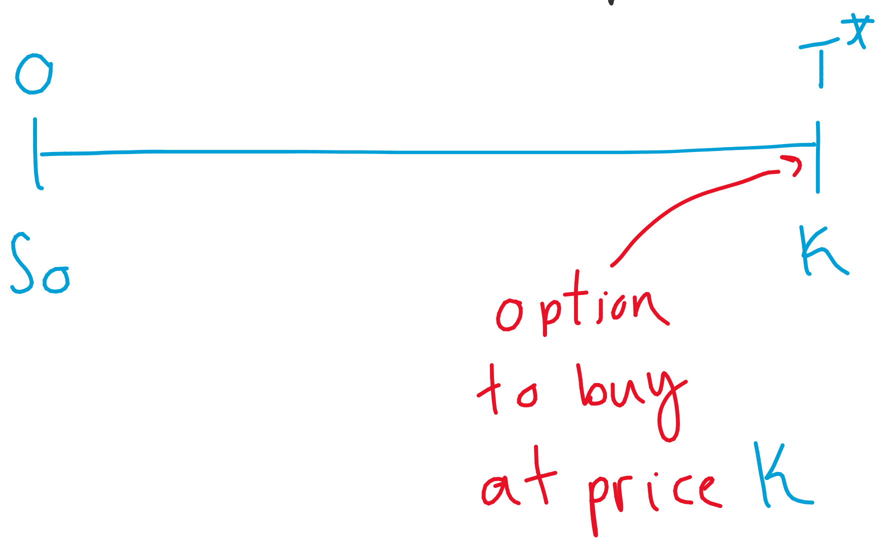
This is a stopping time.

<u>Ex:</u> The last exist time

$$\tau = \max\{n \geq 0 : X_n = 0\}$$

Not a stopping time.

Ex: (Put & Call Option)



At time T^* you recieve

$(S_{T^*} - K)^+ \longleftarrow$ call option payoff

$(K - S_{T^*})^+ \longleftarrow$ put optim payoff.

So when first selling put/call option you calculate the adjusted

$(S_{T^*} - K)^+ \quad (K - S_{T^*})^+$

price - discounted back

today.

These discounted prices

follow martingale theory.

April 4/5

Thursday, April 5, 2018 12:25 PM

Perpetual Put Option

At the time τ the payoff would be

$$(K - S_\tau)^+$$

So the cost today is given

$$v^*(x) = \max_{\tau \in T} \mathbb{E}\left(e^{-r\tau} (K - S_\tau)^+ 1_{\{\tau < \infty\}} \mid S_0 = x \right)$$

optimal time discount payoff exercise eventual

Let $\tau_L = \min\{t \geq 0 : S_t = L\}$

then τ_L is a stopping/hitting time. Then

$$v_L(x) = \mathbb{E}\left(e^{-r\tau_L}(K - S_{\tau_L})^+ 1_{\{\tau_L < \infty\}} \,\middle|\, S_0 = x\right)$$

$$= \mathbb{E}\left(e^{-r\tau_L}(K - L) 1_{\{\tau_L < \infty\}} \,\middle|\, S_0 = x\right)$$

So this is only random in τ_L. So we can maximize over L. Now

$$v^*(x) = \max_{L \in (0,\infty)} v_L(x)$$

provides a maximal expected return.

<u>Def</u>: The σ-algebra of the stopping time τ is defined as

$$\mathcal{F}_\tau = \{ A \in \mathcal{F}_\infty : A \cap \{\tau = n\} \in \mathcal{F}_n, \forall n \}$$

$$= \{ A \in \mathcal{F}_\infty : A = \bigcup_{n=0}^{\infty} (A_n \cap \{\tau = n\}), \; A_n \in \mathcal{F}_n, \forall n \geq 0 \}$$

Exercise: $\mathcal{F}_\tau$ a sigma-alg.?

<u>Thrm</u>: (Basic Properties)

(i) Every positive integer is a stopping time

(ii) If $\tau \equiv k$ then $\mathcal{F}_\tau = \mathcal{F}_k$

(iii) $\mathcal{F}_\tau \subseteq \mathcal{F}_\infty$

(iv) $\tau \in \mathcal{F}_\tau$, $\tau \in \mathcal{F}_\infty$

(v) $\{\tau = \infty\} \in \mathcal{F}_\infty$

<u>Pf</u>:

(i) $\tau \equiv k$ $\quad \{\tau = n\} = \begin{cases} \Omega & n = k \\ \emptyset & n \neq k \end{cases}$

1 †.

(i) $\tau \equiv k$ $\quad \{\tau = n\} = \begin{cases} \Omega & n = k \\ \emptyset & n \neq k \end{cases}$

So by def of sig-alg.

$\Omega, \emptyset \in \mathcal{F}_n$. In either case

$\{\tau = n\} \in \mathcal{F}_n \quad \forall n$

(ii) $\mathcal{F}_\tau = \{A : A \cap \{\tau = n\} \in \mathcal{F}_n \quad \forall n\}$

$= \{A \in \mathcal{F}_\infty : A \in \mathcal{F}_k\} = \mathcal{F}_k$

(iii) By def of $\mathcal{F}_\tau$

(iv) observe

$\{\tau = m\} \cap \{\tau = n\} = \begin{cases} \{\tau = n\} & m = n \\ \emptyset & n \neq m \end{cases}$

So in either case $\in \mathcal{F}_n$

Therefore $\tau \in \mathcal{F}_\tau \subseteq \mathcal{F}_\infty$

So $\tau \in \mathcal{F}_\infty$

(V) $\{\tau=\infty\} = \left(\bigcup_{n=0}^{\infty}\{\tau=n\}\right)^c$

Each $\{\tau=n\} \in \mathcal{F}_\infty$

hence $\bigcup\{\tau=n\} \in \mathcal{F}_\infty$.

<u>Thrm:</u> (Relations between S.T.)

Let τ_1, τ_2 be stopping times

(i) $\tau_1+\tau_2$, $\min(\tau_1,\tau_2)$, $\max(\tau_1,\tau_2)$

are stopping times

(ii) $\tau_m = \min\{\tau, m\}$ is a bounded

S.T.

(iii) $\tau_1-\tau_2$ is not necessarily a S.T.

(iv) If $\tau_n \nearrow \tau$ $\tau_n \searrow \tau$ then

τ is a stopping time

(V) If $A \in \mathcal{F}_{\tau_1}$ then

(v) If $A \in \mathcal{F}_{\tau_1}$ then

$$A \cap \{\tau_1 \leq \tau_2\} \in \mathcal{F}_{\tau_2}$$

(vi) If $\tau_1 \leq \tau_2$ then

$$\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$$

Pf: HW 9

Natural Questions

1. Martingales have constant. expectation. What about

$$\mathbb{E}(X_\tau) = ?$$

2. If $\tau_1 \leq \tau_2$ and $\{X_n\}$ a martingale

$$\mathbb{E}[X_{\tau_2} / \mathcal{F}_{\tau_1}] = X_{\tau_1} ?$$

↳ Doobs optimal sampling.

Ex: Let $\{Y_n : n \geq 1\}$ be an iid seq. of symmetric Bernoulli's.

$$X_n = \sum_{k=1}^{n} Y_k \leftarrow \text{martingale.}$$

Then

$$\mathbb{E}(X_n) = \mathbb{E}(X_m) \quad \forall n,m$$

Let $\tau = \min\{n \geq 1 : X_n = 1\}$

$$\mathbb{E}(X_\tau) = 1 \quad \text{a.s.}$$

But $\mathbb{E}(X_n) = 0$.

So X_τ cannot be a member of the martingale.

Why?

$\mathbb{E}(\tau) = \infty$ $\quad \tau$ too large

Thrm: (Doobs Optimal Sampling)

Let $Z \in L^1(\Omega)$ and let

$\{X_n : n \geq 0\}$ be a $\{\mathcal{F}_n\}$

mart. defined as

$$X_n = \mathbb{E}(Z / \mathcal{F}_n)$$

Let τ be a stopping

time then

$\{(X_\tau, \mathcal{F}_\tau), (Z, \mathcal{F}_\infty)\}$ is

a martingale and

$\mathbb{E}(X_\tau) = \mathbb{E}(Z)$

Pf: Integrable by $Z \in L^1$

and X_n is a martingale

Need to show

$$\int_A X_\tau \, d\mathbb{P} = \int_A Z \, d\mathbb{P} \quad \forall A \in \mathcal{F}_\tau$$

So

$$\int_{\Omega} X_\tau 1_A \, d\mathbb{P} = \mathbb{E}(X_\tau 1_A)$$

$$= \sum_{n\geq 0} \mathbb{E}\left(X_\tau 1_{\{A\cap\{\tau=n\}\}}\right)$$

partion from $\mathcal{F}_\tau$

$$= \sum_{n\geq 0} \mathbb{E}\left(X_n 1_{\{A\cap\{\tau=n\}\}}\right)$$

$$= \sum_{n\geq 0} \mathbb{E}\left\{\mathbb{E}(Z/\mathcal{F}_n)\, 1_{\underbrace{\{A\cap\underbrace{\{\tau=n\}}_{\in\mathcal{F}_n}\}}_{\in\mathcal{F}_n}}\right\}$$

$$= \sum_{n\geq 0} \mathbb{E}\left\{\mathbb{E}\left(Z 1_{\{A\cap\{\tau=n\}\}} / \mathcal{F}_n\right)\right\}$$

$$= \sum_{n\geq 0} \mathbb{E}\left(Z 1_{\{A\cap(\tau=n\}\}}\right)$$

$$= \mathbb{E}(Z 1_A) = \int_A Z \, d\mathbb{P}$$

For integrability

$$\mathbb{E}(Z) = \mathbb{E}(\mathbb{E}(Z / \mathcal{F}_\tau))$$

$$= \mathbb{E}(X_\tau) < \infty$$

by assumption of

$Z \in L^1(\Omega)$.

Lecture 4/12

Thursday, April 12, 2018 12:30 PM

<u>Thrm</u>: Let $Z \in L^1(\Omega)$ and $\{X_n, n \geq 1\}$ be an $\mathcal{F}_n$-martingale of the form

$$X_n = \mathbb{E}(Z / \mathcal{F}_n)$$

Let τ be a S.T. Then

$\{(X_\tau, \mathcal{F}_\tau), (Z, \mathcal{F}_\infty)\}$ is a mart. and

$$\mathbb{E}(X_\tau) = \mathbb{E}(Z).$$

<u>Corollary</u>: Let $\{X_n, n \geq 1\}$ be an $\mathcal{F}_n$ mart. and τ a bounded S.T.

($\mathbb{P}(\tau \le m) = 1$)

Then

$\{(X_\tau, \mathcal{F}_\tau), (X_m, \mathcal{F}_m)\}$

is mart. and

$\mathbb{E}(X_\tau) = \mathbb{E}(X_m)$

<u>Pf:</u> Let $Z = X_m$ then

$Z \in L^1(\Omega)$ and $\forall n \le m$

$\mathbb{E}(Z \mid \mathcal{F}_n) = X_n$

As $\tau \le m$ by the

OST

$\{(X_\tau, \mathcal{F}_\tau), (Z, \mathcal{F}_m)\}$

is a martingale. ∎

<u>Corollary:</u> Let $\{X_n\}$

Corollary: Let $\{X_n\}$ be an $\mathcal{F}_n$ martingale and τ a stopping time. Then

$$\{(X_{\tau\wedge n}, \mathcal{F}_{\tau\wedge n}), (X_n, \mathcal{F}_n)\}$$

is a martingale with

$$\mathbb{E}(X_{\tau\wedge n}) = \mathbb{E}(X_n)$$

Pf: Use of the previous Corollary as

$\tau\wedge n \leq n$ $\boxtimes$

Rmk: · Martingales const. expectation

· mart. at bounded S.T. have const. expectation.

Thrm: (Characterization of mart.)

Let $\{X_n\}$ be an $\mathcal{F}_n$

of ...

Let $\{X_n\}$ be an $\mathcal{F}_n$ adapted sequence. Then $\{X_n\}$ is a martingale iff

$$\mathbb{E}(X_\tau) = \text{Constant}$$

for all bounded S.T. τ

Pf: We have already seen ($\Longrightarrow$)

($\Longleftarrow$) Assume that $\mathbb{E}(X_\tau) = \text{Const.}$ for all bounded τ.

Take $0 \le m < n$. Let $A \in \mathcal{F}_m$ and define

$$\tau(\omega) = \begin{cases} N & \omega \in A \\ m & \omega \notin A \end{cases}$$

τ is a stopping time as it is integer/positive value and the following

$$\{\tau = k\} = \begin{cases} A & k = n \\ A^c & k = m \\ \emptyset & \text{o.w.} \end{cases}$$

$(k=n)$ $\quad A \in \mathcal{F}_m \subseteq \mathcal{F}_n = \mathcal{F}_k$

$(k=m)$ $\quad A^c \in \mathcal{F}_m \Rightarrow A \in \mathcal{F}_m \subseteq \mathcal{F}_n$

$(k \neq m, k \neq n)$ $\quad \emptyset \in \mathcal{F}_n$

Also $\tau \leq n$ So τ is a bounded S.T. so

$$\mathbb{E}(X_\tau) = \text{Const.} = \mathbb{E}(X_m)$$

Bounded S.T.

Moreover

$$\int_A X_\tau \, d\mathbb{P} + \int_{A^c} X_\tau \, d\mathbb{P}$$

$$=$$

$$\int_A X_m \, d\mathbb{P} + \int_{A^c} X_m \, d\mathbb{P}$$

$= S$

$$\int_A X_n \, d\mathbb{P} + \int_{A^c} X_m \, d\mathbb{P} = \int_A X_m \, d\mathbb{P} + \int_{A^c} X_m \, d\mathbb{P}$$

So for all $A \in \mathcal{F}_m$

$$\int_A X_n \, d\mathbb{P} = \int_A X_m \, d\mathbb{P}$$

and $m < n$

This is the same as saying

$$X_m = \mathbb{E}(X_n \mid \mathcal{F}_m)$$ ☒

We can extend the OST of non-decreasing R.V.

<u>Thrm:</u> Let $Z \in \mathcal{L}^1(\Omega)$ and $\{X_n, n \geq 1\}$ a $\{\mathcal{F}_n\}$ martingale of the form

$$X_n = \mathbb{E}(Z / \mathcal{F}_n) \quad n \geq 0$$

Let $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_k$ be s.t., then

$$\{(X_0, \mathcal{F}_0), (X_1, \mathcal{F}_1), \dots, (Z, \mathcal{F}_\infty)\}$$

martingale with

$$\mathbb{E}(X_0) = \dots = \mathbb{E}(Z_{\tau_k}) = \mathbb{E}(Z)$$

<u>Pf:</u> Induction Exercise. ☒

<u>Thrm:</u> $\{(X_n, \mathcal{F}_n) : n \geq 1\}$ be a martingale and let τ be a stopping time.

Then

$\{(X_{\tau\wedge n}, \mathcal{F}_n), n \geq 0\}$ is a martingale and

$$\mathbb{E}(X_{\tau\wedge n}) = \mathbb{E}(X_n)$$

<u>Pf:</u>

Measureability:

$$X_{\tau\wedge n} = X_{\tau\wedge n} 1_{\{\tau<n\}} + X_n 1_{\{\tau\geq n\}}$$

$$= \sum_{k=1}^{n-1} X_k \underbrace{1_{\{\tau=k\}}}_{\substack{\mathcal{F}_k\text{-meas} \\ \subseteq \mathcal{F}_n \text{ meas.}}} + \underbrace{X_n 1_{\{\tau\geq n\}}}_{\mathcal{F}_n\text{-meas}}$$

So $X_{\tau\wedge n} \in \mathcal{F}_n$

(Mart. Rel).

$$\mathbb{E}\left(X_{\tau \wedge (n+1)} / \mathcal{F}_n \right)$$

$$= \sum_{k=1}^{n} \mathbb{E}\left(X_k 1_{\{\tau = k\}} / \mathcal{F}_n \right)$$

$$+ \mathbb{E}\left(X_{n+1} 1_{\{\tau \geq n+1\}} \mid \mathcal{F}_n \right)$$

$$= \sum_{k=1}^{n} X_k 1_{\{\tau = k\}} + \mathbb{E}\left(X_{n+1} \underbrace{1_{\{\tau > n\}}}_{\in \mathcal{F}_n} \mid \mathcal{F}_n \right)$$

$$= \sum_{k=1}^{n} X_k 1_{\{\tau = k\}} + 1_{\{\tau > n\}} \mathbb{E}\left(X_{n+1} \mid \mathcal{F}_n \right)$$

$$= \sum_{k=1}^{n} X_k 1_{(\tau = k)} + X_n 1_{\{\tau > n\}}$$

$$= \sum_{k=1}^{n-1} X_k 1_{(\tau = k)} + X_n \left(1_{\{\tau = n\}} + 1_{\{\tau > n\}} \right)$$

$$= \sum_{k=1}^{n-1} X_k (1_{\{\tau = k\}} + 1_{\{\tau \geq n\}} X_n$$

$$= X_{\tau \wedge n}$$

∎

Martingale Inequality

Thrm: (Kolm - Doob)

Let $\lambda > 0$

(i) If $\{X_n : n \geq 0\}$ is an $\mathcal{F}_n$ - Sub mart. then

$$\lambda \mathbb{P}\left(\max_{0 \leq k \leq n} X_k \geq \lambda \right) \leq \int_{\{\max_{0 \leq k \leq n} X_k > \lambda\}} X_n \, d\mathbb{P} \leq \mathbb{E}(X_n^+) \leq \mathbb{E}|X_n|$$

(ii) If $\{X_n : n \geq 0\}$ is $\{\mathcal{F}_n\}$ - mart then

$$\lambda \mathbb{P}\left(\max_{0 \leq k \leq n} |X_k| > \lambda \right) \leq \int_{\{\max_{0 \leq k \leq n} |X_k| > \lambda\}} |X_n| \, d\mathbb{P} \leq \mathbb{E}(|X_n|).$$

Pf: Let $A = \{\max_{0 \le k \le n} |X_k| > \lambda\}$

Let $\tau = \min\{k : X_k > \lambda\}$

then $A = \{X_{\tau \wedge n} > \lambda\}$

(walk through this 5)

Then

$$\lambda \mathbb{P}(A) = \int_A \lambda \, d\mathbb{P}$$

$$\le \int_A X_{\tau \wedge n} \, d\mathbb{P} \le \int_A X_n \, d\mathbb{P}$$

submart.

$$\le \int_A X_n^+ \, d\mathbb{P} \le \int_\Omega X_n^+ \, d\mathbb{P}$$

$$= \mathbb{E}(X_n^+) \le \mathbb{E}|X_n|$$

(ii) If X_n is a mart.

$\{|X_n|\}$ is a sub mart so

$\{|X_n|\}$ is a submart so

by (i)

$$\lambda\,\mathbb{P}\left(\max_{0\le k\le n}|X_n|>\lambda\right)$$

$$\le \int_{\max_{0\le k\le n}|X_n|>\lambda} |X_n|\,d\mathbb{P}$$

$$\le \mathbb{E}(|X_n|)$$

Lecture 4/17

Tuesday, April 17, 2018 12:31 PM

Martingale Inequalities

Thrm: (Burkholder Ineq.)

Let $\{(X_n, \mathcal{F}_n), n \geq 0\}$ be a martingale. with mart. diff. sequence $\{Y_n, n \geq 0\}$

Denote $X_n^* = \max\limits_{0 \leq k \leq n} |X_k|$, $n \geq 0$

Denote $S_n(x) = \sqrt{\sum\limits_{k=0}^{n} X_k^2}$

(i) Let $p > 1$ there exists Constants A_p, B_p depending only on p s.t.

$$A_p \|S_n(x)\|_p \leq \|X_n\|_p \leq B_p \|S_n\|_p$$

(ii) Let $p \geq 1$ then there exists constants

$A_p \| S_n(x) \|_p \leq \| X_n^* \|_p \leq B_p \| S_n(x) \|_p$

Convergence of Martingales

Thm: Take $\{(X_n, \mathcal{F}_n)\}$ a submartingale with

$$\sup_{n \geq 0} \mathbb{E}(X_n^+) < \infty$$

then X_n converges almost surely.

Pf: We divide the proof into cases of increasing generality.

I: Suppose $\{X_n\}$ is bounded in $\mathcal{L}^2$. Then $\{X_n^2\}$ is a submart. So

$$X_n^2 \leq \mathbb{E}(X_{n+1}^2 | \mathcal{F}_n)$$

$\mathbb{E}(X_n^2) \le \mathbb{E}(\mathbb{E}(X_{n+1}^2 \mid \mathcal{F}_n))$

So

$\mathbb{E}(X_n^2) \le \mathbb{E}(X_{n+1}^2)$

So $\{\mathbb{E}(X_n^2) : n \geq 0\}$ is

increasing and bounded so

it convergenes in $\mathbb{R}$.

Then due to this convergence

$\mathbb{E}[(X_n - X_m)^2]$

$= \mathbb{E}(X_n^2) - \mathbb{E}(X_m^2) \xrightarrow[m,n \to \infty]{} 0$

So $\{X_n : n \geq 1\}$ is Cauchy in

L^2. Therefore there exists X

s.t. $X_n \xrightarrow{L^2} X \Rightarrow X_n \xrightarrow{P} X$

$\Rightarrow X_{n_k} \xrightarrow{a.s.} X$ For

some subsequence $\{n_k\}$.

We need to show

$X_n \xrightarrow{a.s.} X$ Fix $k \geq 0$ and

let $i \geq n_k$. Then

$X_i - X_{n_k} \in \mathcal{F}_i$ So for

all $j \geq i$

$$\mathbb{E}(X_j - X_{n_k} | \mathcal{F}_i) = \mathbb{E}(X_j / \mathcal{F}_i) - X_{n_k}$$

$$= X_i - X_{n_k}$$

So $\{ X_i - X_{n_k} : i \geq n_k \}$

is a martingale and

by the Kol - Doob.

$$\mathbb{P}\left(\max_{n_k \leq i \leq n_{k+1}} |X_i - X_{n_k}| > \varepsilon\right)$$

$$\leq \frac{1}{\varepsilon^2} \mathbb{E}\left[(X_i - X_{n_k})^2\right]$$

So summing across all k

$$\sum_{k\geq 0} \mathbb{P}\left(\max_{n_k \leq i \leq n_{k+1}} |X_i - X_{n_k}| > \varepsilon \right)$$

$$\leq \frac{1}{\varepsilon^2} \sum_{k\geq 0} \mathbb{E}\left[(X_{n_{k+1}} - X_{n_k})^2 \right]$$

We have

$$\sum_{k=0}^{p} \mathbb{E}\left((X_{n_{k+1}} - X_{n_k})^2 \right)$$

$$= \sum_{k=0}^{p} \left(\mathbb{E}(X_{n_{k+1}}^2) - \mathbb{E}(X_{n_k}^2) \right)$$

$$= \mathbb{E}(X_{n_{p+1}}^2) - \mathbb{E}(X_{n_0}^2)$$

$$\leq \mathbb{E}(X_{n_{p+1}}^2) < C < \infty$$

Letting $p \to \infty$ we see

$$\sum_{k=0}^{\infty} \mathbb{E}\left[(X_{n_{k+1}} - X_{n_k})^2 \right] < \infty$$

$k=0$

and

$$\sum_{k=0}^{\infty} \mathbb{P}\left(\max_{n_k \le i \le n_{k+1}} |X_i - X_{n_k}| > \varepsilon\right) < \infty$$

So by Borel Cantelli

$$\mathbb{P}\left(\max_{n_k \le i \le n_{k+1}} |X_i - X_{n_k}| > \varepsilon, \text{ i.o.}\right) = 0$$

Thus $X_n \xrightarrow{a.s.} X$.

<u>Case II</u>: $\{X_n\}$ is a non-neg. L^2 bounded sub.mart.

By the Doob Decomp.

$$X_n = M_n + A_n \quad n \geq 0$$

We have that

$$\mathbb{E}(A_n) = \mathbb{E}(X_n) - \mathbb{E}(M_n)$$

$$= \mathbb{E}(X_n) - \mathbb{E}(M_0)$$

$$= \mathbb{E}(X_n) - \mathbb{E}(X_0)$$

$$= \mathbb{E}(X_n) - \mathbb{E}(X_0)$$

$$\leq \mathbb{E}(X_n)$$

$$\leq \sqrt{\mathbb{E}(X_n^2)} < c < \infty$$

So A_n bounded in L^1 and increasing which implies convergence in $L^1 \Rightarrow$ convergence in P. So monotone an convergence in P gives A_n conv. a.s.

Remains to show M_n bounded in L^2 (then use part I). Recall

$$M_{n+1} - M_n = X_{n+1} - \mathbb{E}(X_{n+1} | \mathcal{F}_n)$$

So

$$\mathbb{E}(M_{n+1}^2) - \mathbb{E}(M_n^2)$$

$$= \mathbb{E}\left((M_{n+1} - M_n)^2 \right)$$

$$= \mathbb{E}\left(\left(X_{n+1} - \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) \right)^2 \right)$$

$$= \mathbb{E}(X_{n+1}^2) - \mathbb{E}\left(\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) \right)$$

$$\leq \mathbb{E}(X_{n+1}^2) - \mathbb{E}(X_n^2)$$

as $\{X_n\}$ is a submartingale.

Proceeding inductively

$M_0 = X_0$

$$\mathbb{E}(M_n^2) \leq \mathbb{E}(X_n^2) < \infty$$

So $\{M_n\}$ is L^2 bounded

and by Case I $\{M_n\}$

conv. a.s. $\Longrightarrow$ $\{X_n\}$ conv. a.s.

<u>Case 3:</u> $\{X_n\}$ nonneg.

L^1 bounded martingale

Then $\{e^{-X_n} : n \geq 0\}$ is

a sub. mart. As

$e^{-X_n} \leq 1$ $\{e^{-X_n}\}$ is L^2

bounded submart.

By case 2 $e^{-X_n} \xrightarrow{a.s} X$.

$\Rightarrow X_n \xrightarrow{a.s.} -\ln(X)$

If $X=0$ then we

would have $X_n \to \infty$

which is not possible as.

$$\mathbb{P}\left(\max_{0\leq k\leq n} |X_k| > \lambda\right) \leq \frac{\mathbb{E}|X_n|}{\lambda} < \frac{C}{\lambda}$$

So

$$\mathbb{P}\left(\sup_{n\geq 0} |X_n| > \lambda\right) \leq \frac{C}{\lambda} \xrightarrow[\lambda\to\infty]{} 0$$

and $\sup_{n\geq 0} |X_n| < \infty$ a.s.

<u>Case 4</u>: $\{X_n\}$ L^1 bounded

mart: Immediatly by

mart: Immediatly by

Krickeberg Decomp + Case 3

Case 5: $\{X_n\}$ L^1 bounded

Submat.

So $X_n = M_n + A_n$

$$\mathbb{E}(A_n) = \mathbb{E}(X_n) - \mathbb{E}(M_n)$$
$$= \mathbb{E}(X_n) - \mathbb{E}(X_0) <$$

by X_n L^1 bounded

So $\{A_n\}$ bounded in L^1

and conv. a.s.

So $M_n = X_n - A_n$ bounded in

L^1. So by case 4

$\{M_n\}$ a.s. conv. and

hence X_n conv. a.s.

Case VI: $\{X_n\}$ is a submart

Case VI: $\{X_n\}$ is a submart

with $\sup E(X_n^+) < \infty$

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

$$= 2E(X_n^+) - E(X_n)$$

$$= 2E(X_n^+) - E(M_n + A_n)$$

$$\leq 2E(X_n^+) - E(M_n)$$

$$= 2\underbrace{E(X_n^+)}_{<\infty} - \underbrace{E(X_0)}_{<\infty}$$

$$< \infty$$

So $\{X_n\}$ L^1 bounded

and by Case 5

$\{X_n\}$ conv. a.s.

Lecture 4/19

Thursday, April 19, 2018 12:33 PM

Ex: Nonconvergent Martingale

$\{X_n : n \geq 0\}$ a seq. of symmetric Bernoulli's and let

$S_n = \sum_{k=1}^{n} X_k$ then

$\{(S_n, \mathcal{F}_n) : n \geq 1\}$ is a martingale.

We know S_n is not convergent. by the law of iterated logarithm.

So by the contrapositive of yesterdays result we should be able to show it is unbounded in L^1

in L^1

$$\frac{S_n}{\sqrt{n}} \xrightarrow{D} N(0,1)$$

As the second moment

is $\mathbb{E}\left[\left(\frac{S_n}{\sqrt{n}}\right)^2\right] = 1$

$$\Rightarrow \left\{ \left|\frac{S_n}{\sqrt{n}}\right| : n \geq 0 \right\}$$

is uniformly integrable

So

$$\mathbb{E}\left(\left|\frac{S_n}{\sqrt{n}}\right|\right) \longrightarrow \mathbb{E}(|Z|) = \sqrt{\frac{2}{\pi}}$$

So

$$\mathbb{E}(|S_n|) \backsim \sqrt{\frac{2n}{\pi}} \longrightarrow +\infty$$

Hence

$\left\{ \frac{S_n}{\sqrt{n}}, n \geq 0 \right\}$ is not bounded in L^1.

Thrm: Let $Z \in L^1(\Omega)$ and let $X_n = \mathbb{E}(Z/\mathcal{F}_n)$ then $\{X_n : n \geq 0\}$ is a uniformly integrable mart.

Def: Recall $\{X_n\}$ is UI if

$$\mathbb{E}\left(|X_n| \mathbb{1}_{|X_n| > a}\right) \xrightarrow[a \to \infty]{} 0$$

uniformly in n

Prop: If $\sup_n \mathbb{E}(|X_n|^p) < \infty$ for some $p > 1$ then $\{X_n\}$ is UI.

Prop: If $\{X_n\}$ is UI then $\sup_n \mathbb{E}(|X_n|) < \infty$

Prop: Recall that if

$X_n \xrightarrow{a.s.} X$ and

$\mathbb{E}(|X_n|) < \infty \quad \forall n$

then

UI $\Longleftrightarrow$ L^1 conv.

Prop: $X \in L^1(\Omega)$ and

$\{A_n, n \geq 1\}$ seq. of events

s.t. $\mathbb{P}(A_n) \longrightarrow 0$ then

$\mathbb{E}(|X| \mathbb{1}_{A_n}) \longrightarrow 0$

Pf: By the Kolmogorov Doob

inequality

$$\mathbb{P}\left(\max_{0 \leq k \leq n} |X_k| \geq a\right) \leq \frac{1}{a} \mathbb{E}(|X_n|)$$

$$\le \frac{1}{a} \mathbb{E}(\mathbb{E}(|Z| \mid \mathcal{F}_n))$$

$$= \frac{1}{a} \mathbb{E}(|Z|)$$

Taking $n \to \infty$ then

$$\mathbb{P}(\sup_n |X_n| > a) \le \frac{1}{a} \mathbb{E}(|Z|)$$

letting $a \to \infty$ then

$$\mathbb{P}(\sup_n |X_n| > a) \le \# \to 0$$

So defining

$\{E_a : a \geq 0\}$ a seq.

of events

$$E_a = \{\sup_n |X_n| > a\}$$

then

$$\mathbb{P}(E_a) \to 0 \text{ as } a \to \infty$$

Then

$$\mathbb{E}\left[|X_n| 1_{|X_n| \geq a}\right]$$

$$\leq \mathbb{E}\left[\mathbb{E}(|Z| \mid \mathcal{F}_n) \underbrace{1_{\{|X_n| \geq a\}}}_{\in \mathcal{F}_n}\right]$$

$$= \mathbb{E}\left[\mathbb{E}(|Z| 1_{|X_n| \geq a} \mid \mathcal{F}_n)\right]$$

$$= \mathbb{E}\left(|Z| 1_{|X_n| \geq a}\right)$$

$$\leq \mathbb{E}\left(|Z| 1_{E_n}\right) \xrightarrow[a \to \infty]{} 0$$

uniformly in n.

Hence $\{X_n\}$ is U.I. $\boxed{\cdot}$

2 motivating questions:

If we have $X_n \xrightarrow{a.s.} X_\infty$

- can we look at X_∞ as the last element of the mart?

$\mathbb{E}(X_\infty / \mathcal{F}_n) = X_n$

- What about L^p convergence?

Thm: (Main Mart Thm) Taking $(X_n, \mathcal{F}_n)$ to be a mart. then the following are equivalent.

(i) X_n is UI

(ii) X_n conv. in L^1

(iii) $X_n \xrightarrow{a.s.} X_\infty$ $X_\infty \in L^1$

$\{(X_n, \mathcal{F}_n) : n = 0, 1, 2, \dots, \infty\}$

is a mart. and X_∞ closes the mart.

(iv) There exists $Y \in L^1(\Omega)$

$X_n = \mathbb{E}(Y / \mathcal{F}_n)$ $\forall n \geq 0$

Pf: We will prove

(i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i)

((i) $\Rightarrow$ (ii)) UI $\Rightarrow$ L^1 boundness ((i) => (ii)) UI => L^1 boundness

then by the conv. thrm.

we see we have a.s. conv.

so

a.s. con. + UI => L^1 conv.

(ii => iii) L_1 conv => L_1 bounded

$\overset{\text{conv.}}{\underset{\text{thrm.}}{\Longrightarrow}}$ $X_n \xrightarrow{a.s.} X_\infty$, $X_\infty \in L^1(\Omega)$.

us a.s. conv + L^1 conv + $\mathbb{E}(|X_n|) < \infty$

=>

$\{X_n\}$ is U.I.

Let $A \in \mathcal{F}_n$ $n > m$ then

$$\int_A X_m \, d\mathbb{P} = \int_A X_n \, d\mathbb{P}$$

So taking limits

$$\int_A X_m \, d\mathbb{P} = \lim_{n \to \infty} \int_A X_n \, d\mathbb{P}$$

then by UI

$$\int_A X_m \, d\mathbb{P} = \int_A X_\infty \, d\mathbb{P}$$

$A \in \mathcal{F}_m$ was arb. so

$$X_m = \mathbb{E}(X_\infty / \mathcal{F}_m)$$

Hence $X_\infty \in L^1(\Omega)$ and

$$\{(X_n, \mathcal{F}_n) \quad 0 \leq n \leq \infty\}$$

is a martingale.

(iii $\Rightarrow$ iv) Take $Y = X_\infty$

(iv $\Rightarrow$ i) See the thm

proved earlier today.

proved earlier today.

for $X_n = \mathbb{E}(X_\infty \mid \mathcal{F}_n)$. ∎

Wald Equations

Ex: Let $\{X_n : n \geq 0\}$ be a seq. of iid R.V.

Let N be a nonneg. integer valued R.V. indep. of $\{X_n\}$. Then if

$\mathbb{E}(N) < \infty$ $\mathbb{E}|X_1| < \infty$

and $S_n = \sum_{i=1}^{n} X_i$

We have

$$\mathbb{E}(S_N) = \sum_{n=0}^{\infty} \mathbb{E}(S_n)\, \mathbb{P}(N=n)$$

$$= \sum^{\infty} n\, \mathbb{E}(X_1)\, \mathbb{P}(N=n)$$

$n=0$

$= \mathbb{E}(X_1)\,\mathbb{E}(N)$

Thrm: (Wald Equation)

Let $\{X_n, n \geq 1\}$ be a seq. of iid R.V. with finite mean μ and let τ be a stopping time.

If $\mathbb{E}(\tau) < \infty$ then

$$\mathbb{E}(S_\tau) = \mu\,\mathbb{E}(\tau)$$

Ex: Let X_n be iid symmetric Bernoullis let $S_n = \sum_{k=1}^{n} X_k$ and let

$\tau = \min\{n \geq 1 : S_n = 1\}$

then $(S_n)_{n=1}^{\infty}$ is

a mart. $S_\tau = 1$ a.s.

$\mathbb{E}(S_\tau) = 1 \neq \mathbb{E}(S_1)$

So $\{S_1, S_\tau\}$ is not

a martingale. This is

due to the fact that

$\mathbb{E}(\tau) = +\infty$. To see

why, assume $\mathbb{E}(\tau) < \infty$

then by the W.E.

$$\mathbb{E}(S_\tau) = \mathbb{E}(X_1)\,\mathbb{E}(\tau)$$

$$\mathbb{E}(S_\tau) = 0$$

but $\mathbb{E}(S_i) = 1$ so

$\mathbb{E}(\tau) = \infty$.