Recall we proved the following thm earlier.

Thm: (The weak law of large numbers)
Let $X_1, X_2, \ldots$ be independent, identically distributed (i.i.d.) random variables with finite expectation.
Let $S_n = X_1 + \cdots + X_n$ & $\mu = EX_1$. Then

$$\frac{S_n}{n} \xrightarrow[\text{in probability}]{P} \mu$$

i.e. $\forall \varepsilon > 0$, $P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$.

We stated a stronger version.

Thm: (Kolmogorov's strong law of large numbers)
Let $X_1, X_2, \ldots$ be iid RVs & let $S_n = X_1 + \cdots + X_n$.
If $EX_i = \mu$ is finite, then

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{a.s.} \mu \quad \text{, i.e.} \quad P\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1.,$$

Conversely, if $\limsup_{n\to\infty} \left|\frac{S_n}{n}\right| < \infty$ with positive probability,

then $EX_i = \mu$ must be finite & $\frac{S_n}{n} \xrightarrow{a.s.} \mu$.

The tool often used to "upgrade" from cv'ce in prob to almost sure cv'ce is the Borel-Cantelli lemma.

Borel-Cantelli lemmas

Let A_n be a s'ce of subsets of Ω. Given $\omega \in \Omega$, can ask how many of the A_n's occur. Let $\limsup A_n$ be the set of those ω's for which infinitely many of the A_n's occur. Formally, define

$$\limsup A_n = \lim_{m\to\infty} \bigcup_{n=m} A_n = \{\omega \text{ that are in } \infty\text{'ly many } A_n\}$$

$$\liminf A_n = \lim_{m\to\infty} \bigcap_{n=m}^{\infty} A_n = \{\omega \text{ that are in all but finitely many } A_n\}$$

We can relate these to the notions of limsup, liminf defined on functions. We have

$$\limsup_{n\to\infty} 1_{A_n} = 1_{\limsup A_n}$$

$$\liminf_{n\to\infty} 1_{A_n} = 1_{\liminf A_n}$$

We write $\omega \in \limsup A_n$ as $\omega \in A_n$ i.o.

Thm: (Borel-Cantelli lemma)

If $\sum_{n\geq 1}^{\infty} P(A_n) < \infty$ then $P(A_n \text{ i.o.}) = 0$.

Pf: 1) $P(A_n \text{ i.o.}) = \lim_{m\to\infty} \underbrace{P(\bigcup_{n=m}^{\infty} A_n)}_{\text{nested s'ce lim}} \leq \lim_{m\to\infty} \sum_{n=m}^{\infty} P(A_m) \to 0$ as the tail of a cv'nt s'ce.

Pf: 2) $N = \sum_{k=1}^{\infty} 1_{A_n} : \Omega \to [0,\infty]$.

$$EN = \sum_{k\geq 1}^{\infty} E 1_{A_n} = \sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(N=\infty) = 0$$

i.e. $P(A_n \text{ i.o.}) = 0$. Δ

Thm: $X_n \xrightarrow{P} X$ iff $\forall$ subseq $X_{n(m)}$ has a subseq $X_{n(m_k)}$ that cv. to X a.s.

Pf: 1) Suppose $X_n \xrightarrow{P} X$. Then $\forall \varepsilon > 0$ $P(|X_n - X| > \varepsilon) \to 0$

$\Rightarrow \exists$ subseq. $X_{n(m)}$ s.t.

$$\sum_{m=1}^{\infty} P(|X_{n(m)} - X| > 2^{-m}) < \infty$$

$\Rightarrow P(|X_{n(m)} - X| > 2^{-m} \text{ i.o.}) = 0$

$\forall \omega \in \{|X_{n(m)} - X| > 2^{-m} \text{ i.o.}\}^c$ we have $X_{n(m)}(\omega) \to X(\omega)$

so $X_{n(m)} \xrightarrow{a.s.} X$

2) Suppose $X_n \not\xrightarrow{P} X$.

Then $\exists \varepsilon > 0, \delta > 0$ & s'ce $X_{n(m)}$ s.t. $P(|X_{n(m)} - X| > \varepsilon) > \delta$ $\forall m$.

$\Rightarrow \forall$ subseq $X_{n(m_k)}$, $P(|X_{n(m_k)} - X| > \varepsilon) > \delta$ $\forall k$.

$\Rightarrow X_{n(m_k)} \not\xrightarrow{a.s.} X$ $\triangle$

Thm: (2nd Borel-Cantelli lemma)

If $A_1, A_2, \ldots$ are indep & $\sum P(A_n) = \infty$, then $P(A_n \text{ i.o.}) = 1$.

Example: $A_1 = A_2 = \cdots$, $P(A_i) \in (0,1)$ shows not true w/o indep'ce.

Pf: $P(\limsup A_n) = \lim_{m\to\infty} P(\bigcup_{n\geq m}^{\infty} A_n) = 1 - \lim_{m\to\infty} P((\bigcup_{n\geq m}^{\infty} A_n)^c)$

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

$$\prod_{n=m}^{M} P(A_n^c) = \prod_{n\geq m}^{M} (1 - P(A_n)) \leq \prod_{n\geq m}^{M} e^{-P(A_n)} = e^{-\sum_{n\geq m}^{M} P(A_n)}$$

$$\sum_{n\geq 1}^{\infty} P(A_n) = \infty \Rightarrow \sum_{n\geq m}^{M} P(A_n) \xrightarrow[M\to\infty]{} \infty \quad \forall m \text{, i.e.}$$

$$\Rightarrow \lim_{M\to\infty} \prod_{n=m}^{M} P(A_n^c) = 0 \quad \text{so} \quad P(A_n \text{ i.o.}) = 1 \ \triangleright$$

Thm: If $X_1, X_2, \dots$ are iid w/ $E|X_i| = \infty$, then $P(|X_n| > n \text{ i.o.}) = 1$

If $S_n = X_1 + \dots + X_n$, then

$$P\left(\lim_{n\to\infty} \frac{S_n}{n} \text{ exists in } (-\infty,\infty)\right) = 0.$$

Hence the strong LLN fails if $E|X_i| = \infty$

Pf: 1) By 2nd B-C lemma ETS $\sum_{n\geq 1}^{\infty} P(|X_n| \geq n) = \infty$

$$E|X_1| = \int_0^\infty P(|X_1| \geq x)\,dx = \sum_{n\geq 0}^{\infty} \int_n^{n+1} P(|X_1| \geq x)\,dx \geq \sum_{n\geq 0}^{\infty} P(|X_1| \geq n) = \sum_{n\geq 0}^{\infty} P(|X_n| \geq n).$$

2) Let $C = \{\omega : \frac{S_n(\omega)}{n} \text{ cv in } (-\infty,\infty)\}$.

$\omega \in C \Rightarrow \frac{S_n(\omega)}{n}$ cv $\Rightarrow \frac{S_n}{n} - \frac{S_{n+1}}{n+1} \to 0.$

but $\left|\frac{S_n}{n} - \frac{S_{n+1}}{n+1}\right| = \left|\frac{S_n}{n(n+1)} - \frac{X_{n+1}}{n+1}\right| \to 0$, $\quad \frac{S_n}{n(n+1)} \to 0$

$\Rightarrow \frac{X_{n+1}(\omega)}{n+1} \to 0 \Rightarrow P(C) \leq P\left(\frac{X_{n+1}}{n+1} \to 0\right)$

But $P(|X_n| > n \text{ i.o.}) = 1 \Rightarrow P\left(\frac{X_n}{n} \to 0\right) = 0$ so $P(C) = 0.$ $\triangleright$

Thm (Kolmogorov's strong law of large numbers)

Let $X_1, X_2, \dots$ be iid RVs & let $S_n = X_1 + \dots + X_n$.

If $EX_i = \mu$ is finite, then

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{a.s.} \mu \text{, i.e.} \quad P\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1.$$

Conversely, if $\limsup_{n\to\infty} \left|\frac{S_n}{n}\right| < \infty$ with positive probability,

then $EX_i = \mu$ must be finite & $\frac{S_n}{n} \xrightarrow{a.s.} \mu$.

Pf: $(\Leftarrow)$ we proved that if $E|X_1| = \infty$, then $\limsup_{n\to\infty} \frac{|S_n|}{n} = \infty$ b/c $X_n \geq kn$ i.o. w/ prob 1 $\forall k$.

$(\Rightarrow)$ 1) Truncate

let $Y_k = X_k \mathbb{1}_{|X_k| \leq k}$.

$T_n = Y_1 + \dots + Y_n$.

ETS $\frac{T_n}{n} \xrightarrow{a.s.} \mu$

Pf: $$\sum_{k=1}^{\infty} P(|X_k| > k) \leq \int_0^{\infty} P(|X_1| > t)\,dt = E|X_1| < \infty$$

so $P(X_k \neq Y_k \text{ i.o.}) = 0$

$P(|S_n - T_n| \text{ bdd indep of } n) = 1$

$\Rightarrow \frac{S_n}{n} - \frac{T_n}{n} \xrightarrow{a.s.} 0$ $\triangle$

2) let $X^+ = X\mathbb{1}_{X \geq 0}$, $X^- = X\mathbb{1}_{X<0}$. We have $X = X^+ + X^-$.

Note that X_n^+ & X_n^- satisfy the assumptions, & proving the result for $X_n^{\pm}$ implies it for X_n, so WLOG assume $X_n \geq 0$

This makes S_n increasing, so can use the trick of proving cv'ce along a subseq'ce

Idea: WTS $\frac{T_n}{n} \xrightarrow{a.s.} \mu$. ETS $\forall \varepsilon > 0$ $P(|\frac{T_n}{n} - \mu| > \varepsilon \text{ i.o.}) = 0$.

Show this by showing $\sum_{n\geq 1}^{\infty} P(|\frac{T_n}{n} - \mu| > \varepsilon) < \infty$.

Unfortunately, doesn't have to be the case.

Show this along a subsequence $k(n)$:

$$\sum_{n\geq 1}^{\infty} P(|\tfrac{T_{k(n)}}{k(n)} - \mu| > \varepsilon) < \infty$$

to get $\frac{T_{k(n)}}{k(n)} \xrightarrow{as} \mu$ & use positivity of X_n's to squeeze $\frac{T_n}{n}$ between neighboring $\frac{T_{k(n)}}{k(n)}$'s.

Let $\alpha > 1$ & $k(n) = \lfloor \alpha^n \rfloor$

We have $\displaystyle\sum_{n=1}^{\infty} P(|T_{k(n)} - ET_{k(n)}| > \varepsilon k(n))$

$$\leq \varepsilon^{-2} \sum_{n=1}^{\infty} \frac{Var(T^2_{k(n)})}{k(n)^2} = \varepsilon^{-2} \sum_{n=1}^{\infty} k(n)^{-2} \sum_{m=1}^{k(n)} Var(Y_m)$$

$$= \varepsilon^{-2} \sum_{m=1}^{\infty} Var(Y_m) \sum_{n: k(n) \geq m} k(n)^{-2}$$

$$\sum_{n: k(n) \geq m} k(n)^{-2} = \sum_{n: \alpha^n \geq m} k(n)^{-2} \leq \sum_{n: \alpha^n \geq m} \left(\tfrac{1}{2}\alpha^n\right)^{-2} = 4 \sum_{n \geq \frac{\log m}{\log \alpha}} \alpha^{-2n}$$

$$= 4 \frac{\text{1st term}}{1 - \alpha^{-2}} \leq 4 \frac{m^{-2}}{1 - \alpha^{-2}}$$

So $\displaystyle\sum_{n=1}^{\infty} P(|T_{k(n)} - ET_{k(n)}| > \varepsilon k(n)) \leq \frac{4}{(1-\alpha^{-2})\varepsilon^2} \sum_{m=1}^{\infty} \frac{E(Y_m^2)}{m^2}$

$$EY_m^2 = \int_0^{\infty} 2y\, P(|Y_m| > y)\, dy \leq \int_0^{m} 2y\, P(|X_1| > y)\, dy$$

So $\displaystyle\sum_{m=1}^{\infty} \frac{E(Y_m^2)}{m^2} \leq \sum_{m=1}^{\infty} \frac{1}{m^2} \int_0^{\infty} \mathbb{1}_{y<m}\, 2y\, P(|X_1| > y)\, dy$

$$= \int_0^{\infty} \left(\sum_{m=1}^{\infty} \frac{1}{m^2} \mathbb{1}_{y<m} \right) 2y\, P(|X_1| > y)\, dy$$

Look at $\displaystyle 2y \sum_{m=\lceil y \rceil}^{\infty} \frac{1}{m^2} \leq 2y \int_{\lceil y \rceil - 1}^{\infty} \frac{1}{x^2}\, dx = 2y \left(-\frac{1}{x}\right)\Big|_{\lceil y \rceil - 1}^{\infty} = \frac{2y}{\lceil y \rceil - 1} \leq 4$

if $y > 1$.

$\sum_{m=1}^{\infty} \frac{1}{m^2}$ cv'nt, so $\exists C>0$ s.t. $\forall y>0$

$$2y \sum_{m=\lceil y \rceil}^{\infty} \frac{1}{m^2} < C$$

So $\displaystyle\sum_{m=1}^{\infty} \frac{E(Y_m^2)}{m^2} \leq C \int_0^{\infty} P(|X_1|>y)\,dy = C\,E|X_1| < \infty$,

So $\forall \varepsilon>0 \quad P(\,|T_{k(n)} - E T_{k(n)}| > \varepsilon k(n) \text{ i.o.}) = 0$

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

(Dominated convergence thm)

Now $\quad E Y_k = E X_k 1_{|X_k| \leq k} = E X_1 1_{|X_1| \leq k} \xrightarrow[k\to\infty]{} E X_1$

So $\quad \dfrac{T_{k(n)}}{k(n)} \xrightarrow{a.s.} E X_1$

For $\quad k(n) \leq m < k(n+1)$

$$\underbrace{\frac{T_{k(n)}}{k(n)}}_{\xrightarrow{a.s.} E X_1} \underbrace{\frac{k(n)}{k(n+1)}}_{\to \frac{1}{\alpha}} = \frac{T_{k(n)}}{k(n+1)} \leq \frac{T_m}{m} \leq \frac{T_{k(n+1)}}{k(n)} = \underbrace{\frac{T_{k(n+1)}}{k(n+1)}}_{\xrightarrow{a.s.} E X_1} \underbrace{\frac{k(n+1)}{k(n)}}_{\to \alpha}$$

So almost surely $\frac{1}{\alpha} E X_1 \leq \liminf \frac{T_m}{m} \leq \limsup \frac{T_m}{m} \leq \alpha E X_1$

$\forall \alpha > 1$. Sending $\alpha \to 1$ get $\quad \dfrac{T_m}{m} \xrightarrow{a.s.} E X_1 \qquad \triangle$

Applications of the WLLN

1) Weierstrass approximation thm

<u>Thm</u> let $f : [0,1] \to \mathbb{R}$ be cts.

Define the Bernstein poly $B_n f$ by

$$(B_n f)(x) := \sum_{j=0}^{n} \binom{n}{j} x^j (1-x)^{n-j} f\left(\tfrac{j}{n}\right).$$

$$\lim_{n\to\infty} B_n f = f \quad \text{unif. on } [0,1]$$

<u>Pf:</u> let $p \in [0,1]$ & $X_1, X_2, \ldots$ iid Bernoulli(p)

$$S_n = X_1 + \cdots + X_n$$

Then $B_n f(p) = E\left(f\left(\frac{S_n}{n}\right)\right)$. so NTS

$$\lim_{n\to\infty} E\left(f\left(\tfrac{S_n}{n}\right) - f(p)\right) = 0 \quad \text{unif. on } [0,1]$$

$$\sup_{0\le p\le 1} \left|(B_n f)(p) - f(p)\right| \le \sup_{0\le p\le 1} E\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right|$$

$$(*) = E\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right| = E\left(\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right| 1_{\left|\frac{S_n}{n} - p\right| < \delta}\right) + E\left(\left|f\left(\tfrac{S_n}{n}\right) - f(p)\right| 1_{\left|\frac{S_n}{n} - p\right| \ge \delta}\right)$$

f cts on cpct $[0,1] \Rightarrow$ unif. cts $\Rightarrow s(\delta) := \sup_{\substack{x,y\in[0,1] \\ |x-y|<\delta}} |f(x) - f(y)|$

is finite & $s(\delta) \xrightarrow[\delta\to 0]{} 0$.

get $(*) \le s(\delta) + 2\max_{x\in[0,1]} f(x) \cdot P\left(\left|\frac{S_n}{n} - p\right| \ge \delta\right)$

goes to 0 but need a uniform bd on p.

$$P\left(\left|\tfrac{S_n}{n} - p\right| \ge \delta\right) \le \frac{\mathrm{Var}\left(\frac{S_n}{n}\right)}{\delta^2} = \frac{n \cdot p(1-p)}{n^2\delta^2} \le \frac{1}{4n\delta^2}$$

so $(*) \le s(\delta) + \dfrac{2\max_{x\in[0,1]} f(x)}{4n\delta^2}$

So $\underbrace{\limsup_{n\to\infty} \sup_{0\le p\le 1} \left|(B_n f)(p) - f(p)\right|}_{} \le s(\delta) \quad \forall \delta > 0.$

$\delta \to 0$ get $= 0$ so $B_n f \to f$ unif on $[0,1]$ $\triangle$

Kolmogorov's maximal inequality

$S_n = X_1 + \dots + X_n$, X_j's indep & in $L^2(P)$.
Then $\forall \lambda > 0$, $n \geq 1$

$$P\left(\max_{1\le k\le n} |S_k - ES_k| \geq \lambda\right) \le \frac{\mathrm{Var}(S_n)}{\lambda^2}$$

Rmk: Chebyshev gives the weaker bound $P(|S_n - ES_n| \geq \lambda) \le \frac{\mathrm{Var}\, S_n}{\lambda^2}$.

Pf: WLOG $EX_i = 0$.

NTS $P(\max_{1\le k\le n} |S_k| \geq \lambda) \le \frac{\mathrm{Var}(S_n)}{\lambda^2}$.

Let A_k be the event that S_k is the 1st of $|S_k| \geq \lambda$.

$A_1, \dots, A_n$ are disjoint so

$$P(\max_{1\le k\le n} |S_k| \geq \lambda) = P\left(\bigcup_{k\ge1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k) \le \sum_{k=1}^{n} \frac{E(S_k^2; A_k)}{\lambda^2} \overset{?}{\le} \frac{E(S_n^2)}{\lambda^2}$$

$E(S_n^2) \geq \sum_{k\ge1}^{n} E(S_n^2; A_k)$

STS $E(S_n^2; A_k) \overset{?}{\geq} E(S_k^2; A_k)$

$(S_n - S_k)^2 \geq 0 \Rightarrow \quad S_n^2 \geq 2(S_n - S_k) S_k + S_k^2$

so $\quad E(S_n^2; A_k) \geq E(2(S_n - S_k) S_k; A_k) + E(S_k^2; A_k)$

$S_n - S_k$ indep of $S_k 1_{A_k} \Rightarrow \quad {}^{\uparrow} = 2E(S_n - S_k) E(S_k 1_{A_k}) = 0.$ $\triangle$