Ch 6 : Independence

- Definition: $\{E_i\}_{i=1}^n \in \mathcal{F}$ are independent if for any $1 \le i_1 < i_2 \cdots < i_k \le n$

$$P(E_{i_1} \cap \cdots \cap E_{i_k}) = \prod_{j=1}^{k} P(E_{i_j}) \quad ☆1$$

Ex: Throwing two dice.

Ex: Throwing n dice.

Important calc: A and B are indep iff $\{A^c, B\}$, $\{A, B^c\}$, $\{A^c, B^c\}$ are indep.

$\Leftarrow$

$$P(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap B^c) = \mathbb{P}(A)\left(1 - \mathbb{P}(B^c)\right)$$
$$= \mathbb{P}(A)\,\mathbb{P}(B)$$

$\Rightarrow$ Similar calculations.

Conclude $\{A_i\}_{i=1}^n$ are indep iff $\{\sigma(A_i) = \{\phi, A_i, A_i^c, \Omega\}\}_{i=1}^n$

are independent.

σ-algebras are indep if every finite collection of events satisfies ☆1.

countable / uncountable / whatevs.

- The σ-algebras $\{\mathcal{F}_\alpha\}_{\alpha \in I}$ are indep if $\{E_{\alpha(i)}\}_{i=1}^n$ are independent for any $E_{\alpha(i)} \in \mathcal{F}_{\alpha(i)}$.

- $\{X_\alpha\}_{\alpha \in I}$ RVs are indep if $\{\sigma(X_\alpha)\}_{\alpha \in I}$ are independent.

  If $\{\mathcal{F}_i\}_{i=1}^n$ are σ-algebras, we sometimes write

$$\sigma\left(\bigcup_{i=1}^n \mathcal{F}_i\right) = \mathcal{F}_1 \vee \cdots \vee \mathcal{F}_n \quad \text{or} \quad \bigvee_{i=1}^n \mathcal{F}_i$$

☆ Ex: Show that $\{1_{E_i}\}_{i=1}^n$ being independent is equivalent to the above.

- An arbitrary collection $\{X_\alpha\}_{\alpha \in I}$ is independent if

Finitary condition.

$$\{X_{\alpha(i)}\}_{i=1}^n \text{ are independent } \quad \forall\ n \text{ and } \{\alpha(i)\}_{i=1}^n \in I$$

Since independence is essentially a finitary condition, we'll focus on equivalences for finite collections first.

Other conditions that imply independence

Lemma: $\{X_i\}_{i=1}^n$ are indep iff the law of $\mathbb{P}_X$ is given by

$$\mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n} \qquad \text{and}$$

$\forall \{\phi_i\}_{i=1}^n$ non-negative functions

$$\mathbb{E}_X\left[\prod_{i=1}^n \phi_i(X_i)\right] = \prod_{i=1}^n \mathbb{E}_{X_i}[\phi_i(X_i)]$$

integral over the individual laws.

Cor: $\{X_i\}_{i=1}^n$ are indep iff $\forall \{h_i\}_{i=1}^n$ that are Borel $h_i : \mathbb{R} \to \mathbb{R}$ $\{h_i(X_i)\}_{i=1}^n$ are independent.

The $\mathbb{P}_{X_i}$ are measures on $\mathbb{R}$ and are called marginals of X.

By definition $\quad \mathbb{P}_{X_i}(A) = \mathbb{P}_X(\pi_i(X) \in A) = \mathbb{P}_X(X \in \pi_i^{-1}(A))$

where π_i is the projection on the i^{th} coordinate.
Sometimes we will just write $\mathbb{P}(X_i \in A)$.

In general marginals $\mathbb{P}_{X_i}$ do not determine $\mathbb{P}_X$.

Ex: Let $X = Y \sim \text{Bernoulli}(p)$ and $X \perp Y \sim \text{Bernoulli}(p)$
work out example.

Pf: Elementary. Take any meas. rectangle

$$\mathbb{P}(X \in F_1 \times \cdots F_n) = \mathbb{P}(X_1 \in F_1 \cap X_2 \in F_2 \cdots)$$

$$= \prod_{i=1}^{n} \mathbb{P}(X_i \in F_i)$$

$$= \prod_{i=1}^{n} \mathbb{P}_{X_i}(X_i \in F_i)$$

This implies $\mathbb{P}$ and $\bigotimes_{i=1}^{n} \mathbb{P}_{X_i}$ agree on measurable rectangles.

The nonnegative meas. stuff is from indicator functions, simple approximation and MON convergence.

Note that the MC lemma is unnecessary here.

Lemma: Suppose $(X_1, \dots, X_d) = X$ is a random vector in $\mathbb{R}^d$ st $\mathbb{P}_X << \text{Leb}$. Then $\exists f : \mathbb{R}^d \to \mathbb{R}$ st

$$P_X(A) = \int_A f \, dx_1 \dots dx_d$$

Show $\mathbb{P}_{X_1}(B) = \int_B f(x_2, \dots, x_d) \, dx_1 \dots dx_d$

where $B \in \mathcal{B}(\mathbb{R})$.

Pf: Let $\pi_1 : \mathbb{R}^d \to \mathbb{R}$ be the canonical projection onto the first coordinate. Then $\forall g : \mathbb{R} \to \mathbb{R}$ bounded

$$\mathbb{E}[g(X_1)] = \mathbb{E}[g(\pi_1(X))] = \int g \circ \pi(\vec{x}) f(x_1 \cdots x_d) \, dx_1 \cdots dx_d$$

$$= \int g(x_1) \left(\int f(x_1 \cdots x_d) \, dx_2 \cdots dx_d \right) dx_1$$

by Fubini.

Ex: $(X_i)_{i=1}^n$ are independent AC random variables if

$$f(x_1 \dots x_n) = \prod_{i=1}^n f_i(x_i) \quad \text{where } f_i \text{ are the marginal densities}$$

and conversely, if $f(x_1, \dots x_n) = \prod_{i=1}^n q_i(x_i)$ then $X_1 \dots X_n$ are independent with densities

$$\frac{q_i(x)}{\int q_i(x) dx}$$

Ex: Suppose $U \sim \text{Unif}[0,1]$, $V \sim \text{Exp}(1)$

and $U \perp V$ (independent)

Find $\quad X = \sqrt{V}\cos(2\pi U) \quad Y = \sqrt{V}\sin(2\pi U)$

the joint density of X, Y.

INDEPENDENT and IDENTICALLY DISTRIBUTED (iid).

$$E[\varphi(X,Y)] = \int_0^1\int_0^\infty \varphi(\sqrt{v}\cos(2\pi u), \sqrt{v}\sin(2\pi u))\, e^{-v}\, dv\, du$$

$$x^2 + y^2 = v \qquad J = \det\begin{pmatrix} \frac{\partial x}{\partial v} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial v} & \frac{\partial y}{\partial u} \end{pmatrix} = \det\begin{pmatrix} \frac{\cos(2\pi u)}{2\sqrt{v}} & -2\pi\sqrt{v}\sin(2\pi u) \\ \frac{\sin(2\pi u)}{2\sqrt{v}} & 2\pi\cos(2\pi u)\sqrt{v} \end{pmatrix}$$

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

$$= \int_{-\infty}^{\infty}\int_{\infty}^{\infty} \varphi(x,y)\, \frac{1}{\pi}\, e^{-(x^2+y^2)}\, dx\, dy$$

$\Rightarrow X \sim N(0,2) \quad Y \sim N(0,2)$

Prop: $(X_i)_{i=1}^{n}$ are independent iff $\Phi_X(t) = E[e^{it \cdot X}]$ the cf of the vector $X = (X_1 \dots X_d)$ satisfies

$$\Phi_X(t_1, \dots, t_n) = \prod_{i=1}^{n} \Phi_{X_i}(t_i)$$

Only need to show $\Leftarrow$

$$\Phi_X(t_1, \dots, t_n) = \mathbb{E}[e^{it \cdot X}]$$

$$= \prod_{i=1}^{n} \Phi_{X_i}(t) = \prod_{i=1}^{n} \int e^{it_i x_i} \mathbb{P}_{X_i}(dt_i)$$

$$= \int \prod_{i=1}^{n} e^{it_i x_i} \mathbb{P}_{X_i}(dt_i)$$

$$= \int e^{i \sum t_i x_i} \mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n}(dt_i) \qquad \text{(Fubini)}$$

This shows the cf of $\mathbb{P}_X$ on $\mathbb{R}^d$ is the same as the cf of $\mathbb{P}_{X_1} \otimes \cdots \mathbb{P}_{X_d}$ on $\mathbb{R}^d$

<u>Prop</u> Let $\{\mathcal{F}_i\}_{i=1}^n$ be σ-algebras, and $\mathcal{C}_i \subset \mathcal{F}_i$ be π-systems (closed under finite intersections) that generate $\mathcal{F}_i$. It is sufficient to check indep on $\mathcal{C}_i$ to prove indep of $\{\mathcal{F}_i\}_{i=1}^n$.

Pf: MC Lemma and bootstrapping the MC lemma.

Do it for $n=2$

$$R = \{A_1 \cap A_2 \ , A_i \in \mathcal{C}_i \ \ i=1,2\}$$

Fix A_2.

Let $M_{A_2} = \{B \cap A_2 : B \in \mathcal{F}_1 \ , \ \mathbb{P}(B \cap A_2) = \mathbb{P}(B)\mathbb{P}(A_2)\}$

$M_{A_2} \supseteq R$, $E \subset F$, $\mathbb{P}(F \setminus E \cap A_2) = \mathbb{P}(F \cap A_2 \setminus E \cap A_2)$

$$= \mathbb{P}(F \cap A_2) - \mathbb{P}(E \cap A_2)$$

$$= (\mathbb{P}(F) - \mathbb{P}(E))\mathbb{P}(A_2) = \mathbb{P}(F \setminus E)\mathbb{P}(A_2)$$

Same for sequences. Etc.

This lemma is important. It has 2 consequences:

1) The joint cdf factors into a product iff $X_1 \cdots X_n$ are independent

2) It can be used for infinite generated σ-algebras and hence Kolmogorov 0-1 law.

Khoshnevisan's version

Lemma: let A, B be two topological spaces with Borel σ-algebras $\mathcal{B}(A)$ and $\mathcal{B}(B)$.

1) $\sigma(X) = \{ X^{-1}(A) : A \in \mathcal{B}(A) \}$

2) If $\{X_i\}_{i=1}^{\infty}$ are rvs in A. Then $Y : \Omega \to B$ is indep of $\{X_i\}_{i=1}^{\infty}$ iff it is indep of $\{X_i\}_{i=1}^{n}$ $\forall n$.

3) If $\mathcal{H}$ and $\mathcal{G}$ both generate $\mathcal{B}(A^{\infty})$ and $\mathcal{B}(B)$ and $Y^{-1}(F)$ is indep of $(X_1, \ldots)^{-1}(E)$ $\forall$ $F \in \mathcal{G}$, $E \in \mathcal{H}$ then Y is indep of $\{X_i\}_{i=1}^{\infty}$

This last bit is necessary for Kol 0-1 law.

Pf: 1) $\bigcup_{i=1}^{\infty} X^{-1}(A_i) = X^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right)$ $\quad X^{-1}(A^c) = X^{-1}(A)^c$

2) If part is obvious. Only if: $\sigma(Y)$ indep of $\sigma(\{X_i\}_{i=1}^n)$

$\Rightarrow P(A \cap B_n) = P(A)P(B_n) \quad \forall A \in \sigma(Y), B_n \in \sigma(\{X_i\}_{i=1}^n)$

Let $G_n = \sigma(\{X_i\}_{i=1}^n)$

Claim: $\sigma(\bigcup G_n) \supseteq \sigma(\{X_i\}_{i=1}^{\infty})$

By definition, $\sigma(\{X_i\}_{i=1}^{\infty})$

$= \bigcap \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra}, \mathcal{F} \supset \bigcup_{i=1}^{\infty} \bigcup_{E \in \mathcal{B}(A)} X_i^{-1}(E)\}$

It is the intersection of all σ-algebras that contain inverse images of all the Borel sets

Let $\mathcal{F}$ be a σ-algebra containing $\bigcup_n G_n$. Then it must certainly contain $X_i^{-1}(E)$ for any i and $E \in \mathcal{B}(A)$.

$\Rightarrow \mathcal{F} \supset \sigma(\{X_i\}_{i=1}^{\infty})$

Thus $\sigma(\bigcup_n G_n) \supseteq \sigma(\{X_i\}_{i=1}^{\infty})$

Now, consider $\Sigma = \{B \in \sigma(\bigcup_n G_n) : P(Y^{-1}(E) \cap B) = P(Y^{-1}(E))P(B)\}$

Σ is a λ-system.

1) $\Omega \in \Sigma$ $\quad$ 2) $A \subset B \in \Sigma$, then

$P(Y^{-1}(E) \cap B) = P(Y^{-1}(E) \cap B \setminus A) + P(Y^{-1}(E) \cap A)$ $\quad$ Law of total prob.

$\Rightarrow P(Y^{-1}(E) \cap B \setminus A) = P(Y^{-1}(E))P(B) - P(Y^{-1}(E))P(A)$

$= P(Y^{-1}(E))P(B \setminus A)$

3) Follows easily from MCT

Thus Σ is a λ-system containing $\bigcup_n G_n$ (since Y is indep of each G_n)

$\Rightarrow \quad \Sigma \supset \sigma(\bigcup_n G_n) \supset \sigma(\{X\}_{i=1}^{\infty})$.

3. Enough to check on subalgebras or π-systems. In particular, suppose $\mathcal{G}$ consists of unions of meas. rectangles. Then

$\Sigma_1 = \{ F \in \mathcal{B}(B) : P(Y^{-1}(F) \cap (X_1, \dots)^{-1}(E)) = P(Y^{-1}(F)) P((X_1, \dots)^{-1}(E)) \}$

Σ_1 is a monotone class / λ-system

$\Sigma_2 = \{ E \in \mathcal{B}(A^{\infty}) : P(Y^{-1}(F) \cap (X_1 \dots)^{-1}(E)) = P(Y^{-1}(F)) P((X_1 \dots)^{-1}(E)) \}$

is also a

★ HW

The proposition that says enough to check independence on π-systems is useful.

Lemma: If $\{\mathcal{F}_i\}_{i=1}^n$ be independ σ-fields

Then $\mathcal{F}_1 \vee \cdots \vee \mathcal{F}_k$ is independent of $\mathcal{F}_{k+1} \vee \cdots \mathcal{F}_n$

Pf: $\mathcal{F}_1 \vee \cdots \vee \mathcal{F}_k$ contains the π-system $R = \{ A_1 \cap \cdots \cap A_k, A_i \in \mathcal{F}_i \}$

(Note: This is different from the product algebra setting which was generated by measurable rectangles. This is generated by intersections:

$$A_1 \cap \cdots \cap A_k \bigcap B_1 \cap \cdots \cap B_k = \bigcap_{i=1}^{k} A_i \cap B_i \quad)$$

Then certainly, $\sigma(R) \supset \sigma(\mathcal{F}_1 \vee \cdots \vee \mathcal{F}_k)$, since R also contains $\mathcal{F}_i$.

By the previous proposition all we have to do is to check on

$$\mathcal{C}_1 = \{ A_1 \cap \cdots \cap A_k : A_i \in \mathcal{F}_i) \text{ and } \mathcal{C}_2 = \{ A_{k+1} \cap \cdots \cap A_n : A_i \in \mathcal{F}_i \}$$

But this is true by definition.

The nice thing is that this applies to $\mathcal{F}_1 \vee \mathcal{F}_2 \cdots \vee \mathcal{F}_k$ and $\underbrace{\mathcal{F}_{k+1} \vee \cdots}_{\mathcal{F}_{k+1}^{\infty}}$

$\mathcal{F}_{k+1}^{\infty}$ contains ∞ly many σ-algebras, but the π-system

$$\mathcal{C} = \{ A_{i_1} \cap \cdots \cap A_{i_j} : \forall j, \ i_s \in \mathcal{F}_{i_s} \ \ i_s \geq k+1 \} \text{ generates } \mathcal{F}_{k+1}^{\infty}$$

(in fact this is the DEFINITION of the product ∞ algebra)

<u>Independence of infinite collectn of rvs</u>

Important for Kolmogorov 0-1 law.

Generated σ-algebras: Let $X_\alpha : \Omega \to \mathbb{R}$ be measurable $\forall\ \alpha \in I$.
$\sigma(\{X_\alpha\}_{\alpha \in I})$ is the smallest σ-algebra such that every X_α is measurable.

Independence of a collection of rvs from a σ-algebra
We say $\{X_\alpha\}_{\alpha \in I}$ is independent of a σ-algebra $\mathcal{G}$ if

$\sigma(\{X_\alpha\}_{\alpha \in I})$ is independent of $\mathcal{G}$.

Independence of 2 collections of rvs:
$\{X_\alpha\}_{\alpha \in I}$ and $\{Y_\beta\}_{\beta \in J}$ are indep if $\sigma(\{X_\alpha\}_{\alpha \in I})$

is independent of $\sigma(\{Y_\beta\}_{\beta \in I})$.

<u>Lemma</u>: Suppose E is indep of $\mathcal{F}_1 \vee \mathcal{F}_2 \cdots \vee \mathcal{F}_n \; \forall n$

then E is indep of $\bigvee_{i=1}^{\infty} \mathcal{F}_i$

Enough to show E is indep of $\mathcal{C}$, the π-system

$\mathcal{C} = \{ A_{i_1} \cap \cdots \cap A_{i_j} : \forall j, \; i_s \in \mathcal{F}_{i_s} \;\; i_s \geq 1 \}$ generates $\mathcal{F}_1^\infty$

(in fact this is the DEFINITION of the product σ algebra)

But this is true since $P(E \cap A_{i_1} \cap \cdots A_{i_j}) = \mathbb{P}(E)\mathbb{P}(A_{i_1} \cap \cdots A_{i_j})$

since $A_{i_1} \cap \cdots A_{i_j} \in \mathcal{F}_{i_1} \vee \cdots \vee \mathcal{F}_{i_j}$

Kolmo's 0-1 law

Recall that the σ-alg generated by $\{X_i\}_{i\in I}$ is the smallest σ-alg s.t. X_i is m'ble $\forall i \in I$.

<u>Defn:</u> The tail σ-alg of the RVs $\{X_i\}_{i\in\mathbb{N}}$ is the σ-alg

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(\{X_i\}_{i=n}^{\infty})$$

Ex: Let $\{X_i\}$ be indep, then $\overline{\lim}\, X_n$, $\overline{\lim}\, \frac{\sum_{i=1}^{n} X_i}{n}$ are tail measurable random variables.

$$\overline{\lim}\, X_n = \inf_{m} \sup_{n\geq m} X_n = \inf_{m\geq m_0} \sup_{n \geq m} X_n \quad \forall m_0$$

$$\Rightarrow \overline{\lim}\, X_n \in \bigvee_{k=m_0}^{\infty} \mathcal{F}_k \quad \forall \; m_0 .$$

Thm: (Kolmogorov's 0-1 law)

If $\{X_i\}_{i=1}^{\infty}$ are indep RVs, then their tail σ-alg $\mathcal{T}$ is trivial, i.e. $\forall E \in \mathcal{T}$, $P(E) = 0$ or 1. $\Rightarrow \forall$ $\mathcal{T}$-m' RV is const a.s.

Let $\mathcal{F}_i = \sigma(X_i)$

Pf: If $E \in \mathcal{T}$, then $E \in \mathcal{F}_n \vee \mathcal{F}_{n+1} \cdots$ $\forall n$

But $\mathcal{F}_1 \vee \cdots \mathcal{F}_{n-1}$ is indep of $\mathcal{F}_n \vee \mathcal{F}_{n+1} \vee \cdots$

So E is indep of $\mathcal{F}_1 \vee \cdots \vee \mathcal{F}_{n-1}$. and by previous

E is indep of $\bigvee_{i=1}^{\infty} \mathcal{F}_i$.

But E is $\mathcal{F}_1 \vee \mathcal{F}_2 \vee \cdots$ measurable, and so

$$P(E \cap E) = P(E)\,P(E) = P(E)$$

$\Rightarrow P(E) \in \{0,1\}$.

Cor: Let $X_1, \dots, X_n, \dots$ be indep. Let $A_n := \frac{X_1 + \dots + X_n}{n}$.

Then $\limsup_{n\to\infty} A_n$, $\liminf_{n\to\infty} A_n$ are both a.s. const.
Prob($\lim_{n\to\infty} A_n$ exists) is 0 or 1. If 1, $\lim_{n\to\infty} A_n$ is const a.s.

Pf: The last statement is $\mathbb{P}(\overline{\lim} A_n - \underline{\lim} A_n = 0)$

which is a difference of two tail meas. rvs equalling 0.

Dawar has a nice construction of a "Cantor set". This proves the existence of a distribution that is neither discrete nor continuous.

☆ Maybe worth constructing a Cantor set.

Weak laws of large numbers

Will prove a few results along the lines: under some independence conditions & constraints on tails or variance, averages cv. to the mean in probability.
Will start with weaker versions & slowly improve

Key tool: Under independence/uncorrelation variance is additive

Defn: A collection $\{X_i\}_{i\in I}$ of RVs w/ $E|X_i|^2<\infty$ is uncorrelated if $EX_iEX_j = EX_iX_j$ $\forall i,j\in I, i\neq j$.

Thm: If $X_1, X_2, \dots, X_n$ are uncorrelated, then

$$\mathrm{Var}(X_1+\dots+X_n) = \mathrm{Var}(X_1)+\dots+\mathrm{Var}(X_n).$$

Pf: Recall $\mathrm{Cov}(X,Y) = E((X-E(X))(Y-E(Y)))$.

$$= EXY - EXEY.$$

So uncorrelated $\iff$ $\mathrm{Cov}(X_i,X_j)=0$ $\forall i\neq j$.

Check $\mathrm{Cov}(aX+bY, Z) = a\,\mathrm{Cov}(X,Z) + b\,\mathrm{Cov}(Y,Z)$

Since $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$, get Cov is bilinear.

Also get $\mathrm{Var}(X)=\mathrm{Cov}(X,X)$ so e.g. $\mathrm{Var}(aX)=a^2\mathrm{Var}(X)$

$$\mathrm{Var}(X_1+\dots+X_n) = \mathrm{Cov}\Big(\sum_i X_i, \sum_j X_j\Big) = \sum_{i,j=1}^n \mathrm{Cov}(X_i,X_j) = \sum_{i=j=1}^n \mathrm{Cov}(X_i,X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$$

$\triangle$

Thm: (L^2 weak law)

Suppose $X_1, X_2, \dots$ are uncorrelated w/ $EX_i=\mu$ & $\mathrm{Var}(X_i)\le C<\infty$.
Let $S_n = X_1+\dots+X_n$. Then

$$\frac{S_n}{n} \xrightarrow{L^2} \mu \quad \& \quad \frac{S_n}{n} \xrightarrow{P} \mu.$$

Pf: $$E\left|\frac{S_n}{n}-\mu\right|^2 = \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) \le \frac{nC}{n^2} \xrightarrow[n\to\infty]{} 0$$ $\triangle$

Key: Under the assumptions variance of the sum grows linearly & not quadratically, so when averaging, the variance decays at rate $\frac{1}{n}$.

Thm: (Khintchine's Weak LLN)
Let $X_1, X_2, \ldots$ be iid w/ $E|X_i| < \infty$. $S_n = X_1 + \cdots + X_n$.
$\mu = EX_1$. Then $\frac{S_n}{n} \to \mu$ in L_1 & in p.

Pf: Truncate: given $\alpha > 0$ let $X_i^\alpha := X_i \mathbb{1}_{X_i \leq \alpha}$

$S_n^\alpha := X_1^\alpha + \cdots + X_n^\alpha$

$$\|S_n - S_n^\alpha\|_1 \leq \sum_{i=1}^n \|X_n - X_n^\alpha\|_1 = n \|X_n \mathbb{1}_{X_n \geq \alpha}\|_1$$

$$= n E(|X_1| \mathbb{1}_{|X_1| \geq \alpha})$$

$$\left\|\frac{S_n}{n} - \mu\right\|_1 \leq \left\|\frac{S_n}{n} - \frac{S_n^\alpha}{n}\right\|_1 + \left\|\frac{S_n^\alpha}{n} - \mu_\alpha\right\|_1 + |\mu_\alpha - \mu|$$

$$\leq 2E(|X_1| \mathbb{1}_{|X_1| \geq \alpha}) + \left\|\frac{S_n^\alpha}{n} - \mu_\alpha\right\|_1$$

By the L^2 weak law $\xrightarrow[n\to\infty]{} 0$

So $\limsup_{n\to\infty} \left\|\frac{S_n}{n} - \mu\right\|_1 \leq 2E(|X_1| \mathbb{1}_{|X_1| \geq \alpha})$.

DCT applied to α gives $\xrightarrow[\alpha\to\infty]{} 0$ $\triangle$

Borel-Cantelli lemmas

Tools often used to get almost sure cv'ce from cv'ce in prob.

Let A_n be a s'ce of subsets of Ω.

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

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

Motivation for notation

$$\overline{\lim}\, A_n = \bigcap_m \bigcup_{n=m}^{\infty} A_n$$

$$\overline{\lim}\, 1_{A_n}(\omega) = 1 \iff \omega \in A_n \text{ i.o.}$$

$$= 1_{\overline{\lim} A_n}(\omega)$$

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

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

Thm: (Borel - Cantelli lemma)

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

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

Two ways of seeing the same thing.

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

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

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

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

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

not worth repeating

Thm: If f is cts, $X_n \xrightarrow{P} X$, then $f(X_n) \xrightarrow{P} f(X)$.
If f is bdd, $\mathbb{E} f(X_n) \to \mathbb{E} f(X)$.

pf: If $X_{n(m)}$ is a subseq of $X_n \Rightarrow \exists$ subseq $X_{n(m_k)} \xrightarrow{a.s.} X$
$\Rightarrow$ since f cts, $f(X_n) \xrightarrow{a.s.} f(X) \to f(X_n) \xrightarrow{P} f(X)$.
f bdd $\Rightarrow$ BCT gives $\mathbb{E} f(X_{n(m_k)}) \to \mathbb{E} f(X)$, so
$\mathbb{E} f(X_n)$ is a s'ce s.t. $\forall$ subseq has a subseq cv. to $\mathbb{E} f(x)$,
so $\mathbb{E} f(X_n) \to \mathbb{E} f(X)$.

$\triangleright$

We have seen that if $\sum P(A_n) < \infty$ then $P(A_n \text{ io}) = 0$

$\Leftrightarrow P(A_n^c \text{ ev}) = 1$ "A_n stops happening for large enough n"

Is it true that $\sum P(A_n) = +\infty \Rightarrow P(A_n \text{ io}) = 1$

Ex: $X = [0,1]$ $A_n = (0, 1/n]$ λ = Lebesgue meas.
we have $\sum_n P(A_n) = +\infty$ $\quad \mathbb{P}(A_n \text{ io}) = \mathbb{P}\left(\bigcap_n \bigcup_{m \geq n} A_m\right) = \mathbb{P}\left(\bigcap_n A_n\right)$

$$= \lim_n \mathbb{P}(A_n) = 0.$$

If A_n is decreasing certainly this doesn't have to be true.

Are there conditions under which this is true?

Thm: (2nd Borel-Cantelli lemma)
If $A_1, A_2, \dots$ 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: To use indep. have to find a product $\quad \overline{\lim} A_n = \bigcap_n \bigcup_{m=n}^{\infty} A_m$

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

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

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

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

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

Another application of BC

Let $([0,1), \mathcal{B}([0,1)), \overset{\text{Leb}}{\lambda})$ $\quad X(\omega) = \omega$. Let $\omega = 0.a_1 a_2 \cdots$ be its binary expansion. Let $b_1, b_2, \dots b_n$ (= 01101 say) be any length n pattern.

Thm: a.s. ω, ALL finite patterns appear infinitely often.

Pf: Let $X_1(\omega) = \lfloor 2\omega \rfloor - 2\lfloor \omega \rfloor$

where $\lfloor \ \rfloor$ is the floor funcn. If $\omega = \sum_{i=1}^{\infty} \frac{a_i}{2^i}$ (assuming such an expansion exists)

Then clearly $X_1(\omega) = a_1$. In general, let

$$X_n(\omega) = \lfloor 2^n \omega \rfloor - 2\lfloor 2^{n-1} \omega \rfloor \qquad \text{(Le Gall's definition)}$$

My preference is to define it as follows:

Let $X_1 = \max\{ i \in \{0,1\} : \frac{i}{2} \le \omega \}$ $\quad$ If $\frac{X_1(\omega)}{2} = \omega$ stop.

Inductively define

$$X_n = \max\left\{ i \in \{0,1\} : \sum_{i=1}^{n-1} \frac{X_i(\omega)}{2^i} + \frac{i}{2^n} \le \omega \right\}$$

If $\omega = \sum_{i=1}^{n} \frac{X_i(\omega)}{2^i}$, stop and set $X_k = 0 \quad \forall\, k \ge n+1$

Then
$$\boxed{0 \le \omega - \sum_{i=1}^{n} \frac{X_i(\omega)}{2^i} < \frac{1}{2^n}}$$

$\omega - \frac{X_1(\omega)}{2} < \frac{1}{2}$ clearly. If $\omega - \sum_{i=1}^{n-1} \frac{X_i}{2^i} \le \frac{1}{2^{n-1}}$

If $X_n(\omega)=0$ $\qquad \omega < \sum_{i=1}^{n-1} \frac{X_i}{2^i} + \frac{1}{2^n}$

$$\Rightarrow \quad \omega - \sum_{i=1}^{n} \frac{X_i}{2^i} < \frac{1}{2^n}$$

If $X_n(\omega)=1$ $\quad \Rightarrow \quad \omega - \sum_{i=1}^{n} \frac{X_i}{2^i} \leq \frac{1}{2^{n-1}} - \frac{1}{2^n} = \frac{2-1}{2^n} = \frac{1}{2^n}$

This implies $\quad \omega = \sum_{i=1}^{\infty} \frac{X_i(\omega)}{2^i}$

<u>Claim:</u> $\lambda(X_i=0) = \lambda(X_i=1) = \frac{1}{2}$

Pf: $\quad \{X_i=0\} = \bigsqcup_{\substack{a_1 \cdots a_{i-1} \\ a_j \in \{0,1\}}} \{(X_1 \cdots X_{i-1}) = (a_1 \cdots a_{i-1}),\ X_i = 0\}$

$$\{(X_1 \cdots X_{i-1}) = (a_1 \cdots a_{i-1}),\ X_i=0\} = \{\omega : \sum_{j=1}^{i-1} \frac{a_j}{2^j} \leq \omega < \sum_{j=1}^{i-1} a_j + \frac{1}{2^i}\}$$

Each of these sets has measure $\frac{1}{2^i}$ and there are 2^{i-1} of these.

$$\Rightarrow \quad \lambda(\{X_i=0\}) = \frac{1}{2} = \lambda(\{X_i=1\})$$

<u>Claim:</u> $(X_1 \cdots X_j)$ are independent.

$$\lambda(X_1=a_1 \cdots X_j=a_j) = \lambda\left(\{\omega : \sum_{i=1}^{j} \frac{a_i}{2^i} \leq \omega < \sum_{i=1}^{j} \frac{a_i}{2^i} + \frac{1}{2^j}\}\right)$$

$$= \frac{1}{2^j} = \lambda(X_1=a_1) \cdots \lambda(X_j=a_j)$$

Then for fixed $b_1 b_2 \cdots b_k$

$\underbrace{X_1 \cdots X_k}_{=}$ $\quad X_{k+1} \cdots X_{2k} \cdots$ $\qquad$ divide into blocks of k.

$$\lambda\Big(\underbrace{\{X_{mk} \cdots X_{(m+1)k} = b_1 \cdots b_k\}}^{E_m}\Big) = \frac{1}{2^k}$$

$\sum_{m=1}^{\infty} \lambda(E_m) = \infty$ but E_m are independent so

$$\lambda(E_m \text{ i.o.}) = 1$$

Another proof:

$$Z_n = \sum_{i=1}^{n} 1_{A_i}$$

Chebyshew

$$Var(Z_n) = \sum P(A_i)(1-P(A_i)) \qquad E[Z_n] = \sum_{i=1}^{n} P(A_i)$$

$$P(|Z_n - EZ_n| > c\,EZ_n) \leq \frac{1}{c^2} \frac{Var(Z_n)}{E[Z_n]^2} \leq \frac{1}{c^2} \frac{1}{E[Z_n]} \to 0$$

$$\Rightarrow P\left(Z_n \geq \frac{1}{2} E[Z_n]\right) \longrightarrow 1$$

But $P\left(Z_\infty \geq \frac{1}{2} E[Z_n]\right) \longrightarrow 1$

I like this proof more.

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

If $S_n = X_1 + \cdots + X_n$, 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 enough to show $\sum_{n=1}^{\infty} P(|X_n| \geq n) = \infty$

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

$\Rightarrow$ $X_n \geq n$ eventually.

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

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

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

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

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

Here is a cleaner proof: (due to Khoshnevisan)

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

$$\Rightarrow \quad \varlimsup \frac{|X_n|}{n} \leq 2 \varlimsup \frac{|S_n|}{n}$$

But $$E\frac{|X_1|}{\lambda} = \int_0^\infty P\left(\frac{|X_1|}{\lambda} \geq x\right) = \sum_{n=0}^{\infty} \int_n^{n+1} P\left(|X_1| \geq \lambda x\right) \geq \sum_{n=0}^{\infty} P(|X_n| \geq \lambda n)$$

$\Rightarrow$ $|X_n| > \lambda n$ eventually. This is true $\forall \lambda$.

Kolmogorov's maximal inequality

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

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

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

Pf: WLOG $EX_i = 0$.
NTS $P(\max_{1 \leq k \leq n} |S_k| \geq \lambda) \leq \frac{\mathrm{Var}(S_n)}{\lambda^2} = \frac{E[S_n^2]}{\lambda^2}$

Let A_k be the event that S_k is the 1st time $|S_i| \geq \lambda$. $i = 1, \cdots, n$
$A_1, \cdots, A_n$ are disjoint so

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

Is this true?

Certainly true $\rightarrow$ $E(S_n^2) \geq \sum_{k=1}^{n} E(S_n^2; A_k)$
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 $S_k 1_{A_k}$ $\Rightarrow$ $\cdots = 2E(S_n - S_k) E(S_k 1_{A_k}) = 0$. $\triangle$

It's quite a clever inequality and uses independence in an important

This proof is from Durrett & seems to be originally due to Etemadi.

The Etemadi and maximal inequality proofs are very similar:

1) Etemadi says let $k(n) = \alpha^n$
2) standard proof says let $k(n) = 2^n$

☆ Etemadi appears a little slicker since it does not need the maximal inequality. But the maximal inequality is used in Martingale convergence and the ergodic theorem.

<u>Thm:</u> (Kolmogorov's Strong law of large numbers)
If $X_1 \in L^1(P)$ & $X_1, X_2, \ldots$ are iid then $\frac{S_n}{n} \xrightarrow[n\to\infty]{a.s.} E X_1$ ← where $S_n = X_1 + \cdots + X_n$
Conversely, if $\limsup_{n\to\infty} \left|\frac{S_n}{n}\right| < \infty$ w/ pos. prob. $\Rightarrow$ X_j's in $L^1(P)$ (& hence ←)

(this pf from Durrett, (Khoshnevisan) Bobby's pf uses Kolmog. max ineq).

<u>Pf:</u> ($\Leftarrow$) 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) <u>Truncate</u>

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

$T_n = Y_1 + \cdots + Y_n$.

Enough to show $\frac{T_n}{n} \xrightarrow{a.s.} \mu$

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

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

$P(|S_n - T_n| \text{ is finite } \forall n) = 1$

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

2) 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

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

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

swap sums

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

Each $Var(Y_m)$ is counted as long as $k(n) \geq m$

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

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

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

$$\sum_{n \geq \lceil x \rceil} \alpha^{-2n} \leq \int_{\lceil x \rceil}^{\infty} \alpha^{-2u}\, du = \frac{\alpha^{-2\lceil x \rceil}}{\log \alpha}$$

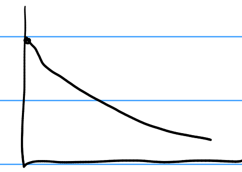
So $$\sum_{n=1}^{\infty} P(|T_{k(n)} - E\,T_{k(n)}| > \varepsilon k(n)) \le \frac{4}{(1-\alpha^{-2})\varepsilon^2} \sum_{m=1}^{\infty} \frac{E(Y_m^2)}{m^2}$$

$Y_m = X_m 1_{\{X_m \le m\}}$

$$E Y_m^2 = E\left[|X_m|^2 1_{\{|X_m| \geqslant m\}}\right] = E\left[|X_1|^2 1_{\{|X_1| \geqslant m\}}\right]$$

move sum in

So $$\sum_{m=1}^{\infty} \frac{E(Y_m^2)}{m^2} \le \sum_{m=1}^{\infty} E\left[|X_1|^2 \frac{1_{\{|X_1| \geqslant m\}}}{m^2}\right] \quad \text{— } (\star 1)$$



$m =$

Can assume $|X_1| \geqslant 2$ and deal with $|X_1| \le 2$ as a separate term.

$$\sum_{m=1}^{x} \frac{1}{m^2} \le 2\int_1^x \frac{1}{u^2}\,du = \frac{2}{x-1} \le \frac{4}{x} \qquad \frac{x}{x-1} \le 2$$

$$(\star 1) \le C + E\left[|X_1|^2 \frac{4}{|X_1|} 1_{\{|X_1| \geqslant 2\}}\right] < \infty$$

So $$\sum_{m=1}^{\infty} \frac{E(Y_m^2)}{m^2} \le C + 4\,E|X_1| < \infty,$$

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

by Borel-Cantelli.

$$\Rightarrow \quad \frac{T_{k(n)} - \mathbb{E}T_{k(n)}}{k(n)} \xrightarrow{a.s.} 0 \qquad (\text{taking } \epsilon \downarrow 0)$$

DCT gives $\mathbb{E} Y_k = \mathbb{E} X_k \mathbb{1}_{|X_k| \le k} = \mathbb{E} X_1 \mathbb{1}_{|X_1| \le k} \xrightarrow[k\to\infty]{} \mathbb{E} X_1$

so $$\frac{T_{k(n)}}{k(n)} \xrightarrow{a.s.} \mathbb{E} X_1 .$$

For $k(n) \le m < k(n+1)$

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

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

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

Proof of LLN (Khoshnevisan, using Maximal equality)

Wlog $EX_1 = 0$

<u>L^2 case</u>: By Maximal inequality

$$P\left(\max_{1\le k\le n} |S_k| > \epsilon n\right) \le \frac{Var(S_n)}{n^2 \epsilon^2} = \frac{E[X_1^2]}{n\epsilon^2}$$

$n \longrightarrow 2^n$

To get

$$\sum_{n=1}^{\infty} P\left(\max_{1\le k\le 2^n} |S_k| > \epsilon 2^n\right) < \infty$$

$$\Rightarrow \max_{1\le k\le 2^n} |S_k| \le \epsilon 2^n \qquad \text{EVENTUALLY}$$

For $2^n \le m \le 2^{n+1}$

$$|S_m| \le \epsilon 2^{n+1} \le 2\epsilon m$$

This is true for each ϵ then.

$$\overline{\lim_{m\to\infty}} \frac{|S_m|}{m} \le 2\epsilon \qquad a.s.$$

and hence $\lim \frac{|S_n|}{n} \longrightarrow 0$ a.s.

L^1 case:

Let $Y_i = X_i \mathbb{1}_{\{|X_i| \leq i\}}$ (Truncate)

$$S_n' = \sum_{i=1}^{n} Y_i$$

identical distribution

$$\sum_{i=1}^{\infty} P(|X_i| \geq i) = \sum_{i=1}^{\infty} P(|X_1| \geq i) \leq E[|X_1|] < \infty$$

$$\Rightarrow \quad |S_n' - S_n| = \left| \sum_{i=1}^{n} |X_i| \mathbb{1}_{\{|X_i| > i\}} \right| = o(n)$$

($Y_i = X_i$ eventually)

Since X_i have 0 mean,

$$E[S_n'] = \sum_{i=1}^{n} E[X_i] - E[X_i \mathbb{1}_{\{|X_i| > i\}}] = -\sum_{i=1}^{n} E[X_i \mathbb{1}_{\{|X_i| > i\}}]$$

$$\Rightarrow \quad |E[S_n']| \leq \sum_{i=1}^{n} E[|X_i| \mathbb{1}_{\{|X_i| > i\}}]$$

$$\overset{d}{=} E\left[|X_1| \sum_{i=1}^{n} \mathbb{1}_{\{|X_1| > i\}}\right] = E\left[\sum_{i=1}^{n} |X_1| \, i \, \mathbb{1}_{\{i < |X_1| \leq i+1\}}\right]$$

$$\leq E\left[|X_1| \min(|X_1|, n+1)\right]$$

$$\frac{|E[S_n']|}{n} \leq E\left[|X_1| \underbrace{\min\left(\frac{|X_1|}{n}, 1 + \frac{1}{n}\right)}_{\to\, 0 \text{ and bounded above}}\right] \longrightarrow 0 \quad \text{(DCT)}$$

$$\Rightarrow \quad E[S_n'] = o(n)$$

Now need to control

$$E(n) = \left\{ \max_{1 \le k \le 2^n} |S'_k - E'_k| \geq 2^n \epsilon \right\}$$

$$P(E(n)) \leq \frac{Var(S'_{2^n})}{2^{2n} \epsilon^2}$$

$$\sum_{n=1}^{\infty} P(E(n)) \leq \sum_{n=1}^{\infty} \sum_{j=1}^{2^n} \frac{E[Y_j^2]}{2^{2n} \epsilon^2}$$

bounds $Var(Y_i)$

Exchange sums.

Each $E[Y_j^2]$ will appear as long as $2^n \geq j \iff n \geq \log_2 j$

$$= \frac{1}{\epsilon^2} \sum_{j=1}^{\infty} E[Y_j^2] \underbrace{\sum_{n \geq \log_2 j} \frac{1}{4^n}}$$

Geometric series.

$$\leq \frac{C}{\epsilon^2} \sum_{j=1}^{\infty} E[Y_j^2] \frac{1}{4^{\log_2 j}}$$

$$= \frac{C}{\epsilon^2} \sum_{j=1}^{\infty} E[|X_j|^2 \mathbb{1}_{\{|X_j| \leq j\}}] \frac{1}{j^2}$$

$$= \frac{C}{\epsilon^2} \sum_{j=1}^{\infty} E[|X_1|^2 \mathbb{1}_{\{|X_1| \leq j\}}] \frac{1}{j^2}$$

$$\leq \frac{C}{\epsilon^2} E\left[|X_1|^2 \sum_{j=1}^{\infty} \mathbb{1}_{\{|X_1| \leq j\}} \frac{1}{j^2} \right]$$

$$(\star 1) \leq C + E\left[|X_1|^2 \frac{2}{|X_1|} \mathbb{1}_{\{|X_1| \geq 2\}} \right] < \infty$$

Using $\displaystyle\sum_{m=x}^{\infty} \frac{1}{m^2} \leq 2\int_x^{\infty} \frac{1}{u^2} du = \frac{2}{x}$

This shows $\max\limits_{1 \leq k \leq 2^n} |S_k' - ES_k'| < 2^n \epsilon$ eventually

Fix $2^n \leq m \leq 2^{n+1}$

$\Rightarrow$ $|S_m' - ES_m'| \leq 2^{n+1} \epsilon \leq 2 \epsilon m$ $\forall m$ large enough

$\Rightarrow$ $\overline{\lim\limits_{m \to \infty}} \left| \frac{S_m' - ES_m'}{m} \right| \leq 2\epsilon$ a.s.

If $E|X_1| = +\infty$, we had previously shown $\overline{\lim} \frac{|S_n|}{n} = +\infty$. The following is a slight upgrade based on the SLLN.

Thm: (Strong LLN w/ ∞ expectation)
Let $X_1, X_2, \ldots$ be iid w/ $EX_i^+ = \infty$, $EX_i^- < \infty$.
Let $S_n = X_1 + \cdots + X_n$. Then $\frac{S_n}{n} \xrightarrow{a.s.} \infty$

(So Strong LLN holds as long as EX_1 exists in $\mathbb{R} \cup \{\infty\}$)

Pf: Let $M > 0$, $X_i^M = \min(X_i, M)$.
X_i^M iid, $E|X_i^M| < \infty \Rightarrow \frac{S_n^M}{n} := \frac{X_1^M + \cdots + X_n^M}{n} \xrightarrow[n\to\infty]{a.s.} EX_1^M$

$X_i \geq X_i^M \Rightarrow \liminf_{n\to\infty} \frac{S_n}{n} \geq \lim_{n\to\infty} \frac{S_n^M}{n} = EX_i^M \quad \forall M.$

MCT implies $E(X_i^M)^+ \xrightarrow{M\to\infty} EX_i^+ = \infty$, so $EX_i^M = E(X_i^M)^+ - E(X_i^M)^- \to \infty$
& $\liminf_{n\to\infty} \frac{S_n}{n} = \infty$ $\triangle$

Applications (of not just the LLN)

1) Weierstrass approximation thm (constructive)

<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 $p\in[0,1]$ & $X_1, X_2, \dots$ iid Bernoulli(p)

$$S_n = X_1 + \dots + X_n$$

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

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

polynomial

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

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

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

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

2) The Asymptotic equipartition property

Let $A = \{\sigma_1, \dots, \sigma_m\}$ be an alphabet & consider words of length n in this alphabet. $\exists\, m^n$ words.

Given a word $w = (w_1, \dots, w_n)$ define the relative frequency of the letter σ_k in w to be

$$f_n(\sigma_k, w) = \# \text{ occurrences of } \sigma_k \text{ in } w.$$

Let $W = (W_1, \dots, W_n)$ be a unif. random word out of the m^n.

Eq: $W_1, \dots, W_m$ – unif. random iid letters: $P(W_i = \sigma_j) = \frac{1}{m}\ \forall i, j$.

Weak LLN

$$f_n(\sigma_k, w) = \frac{1}{n} \sum_{j=1}^{n} 1_{\sigma_k}(w_j) \xrightarrow{\text{in probability}} P(W_1 = \sigma_k) = \frac{1}{m}.$$

What can we say about the letter frequencies in long words.

Let $p_1, \dots, p_m > 0$, $p_1 + \dots + p_m = 1$.

<u>Defn</u>: If $\varepsilon > 0$, an n-letter word is ε-typical for $P = (p_1, \dots, p_m)$ if

$$|f_n(\sigma_k, w) - p_k| < \varepsilon \quad \forall k = 1, \dots, m.$$

"Frequencies are close to their correct values"

<u>Thm</u>: (Shannon)

$\forall n \geq 1$, $\varepsilon > 0$, P, let $T_n(\varepsilon)$ be the # length n ε-typical words for P.

Then

$$\left(1 - \frac{1}{n\varepsilon^2}\right) e^{n(H(P) - c\varepsilon)} \leq T_n(\varepsilon) \leq e^{n(H(P) + c\varepsilon)}$$

where $c = -\sum_{k=1}^{m} \log p_k \geq 0$ & $H(P) = \sum_{i=1}^{m} p_i \log p_i$ is the entropy of the vector $p = (p_1, \dots, p_m)$.

- There are exponentially many ε-typical words.
- They all have more or less the same probability
- $\mathbb{P}$ induces the uniform meas on ε-typical words.
- Has a vast generalization to ergodic sequences (instead of just iid sequences of letters)

pf: let $X_1, X_2, \dots$ be iid w/ $P(X_1 = \sigma_i) = p_i$.

let $W_n = (X_1, \dots, X_n)$.

Given ω – word of length n,

$$P(W_n = \omega)$$

$$= \prod P(X_1 = \omega_1) P(X_2 = \omega_2) \cdots = \prod_{k=1}^{m} p_k^{n f_n(\sigma_k, \omega)}$$

← # of times σ_k appears in ω

↑ independence ↑ σ_1 ↑ σ_1

This is how the entropy appears.

If w is ε-typical, then

$$e^{-n(H(p)+c\varepsilon)} = \prod_{k=1}^{m} p_k^{n(p_k+\varepsilon)} \le P(W_n = w) \le \prod_{k=1}^{m} p_k^{n(p_k-\varepsilon)} = e^{-n(H(p)-c\varepsilon)}$$

This is the most important. It shows ε-typical words have same prob.

so $$T_n(\varepsilon)\, e^{-n(H(p)+c\varepsilon)} \le P(W_n \text{ is } \varepsilon\text{-typical}) \le T_n(\varepsilon)\, e^{-n(H(p)-c\varepsilon)}$$

$$\nearrow \le 1 \implies T_n(\varepsilon) \le e^{n(H(p)+c\varepsilon)}.$$

For the LB, to show $P(W_n \text{ is } \varepsilon\text{-typical}) \ge 1 - \frac{1}{n\varepsilon^2}$.

$$1 - P(W_n \text{ is } \varepsilon\text{-typical}) = P(\exists k = 1, \dots, m : |f_n(\sigma_k, W_n) - p_k| > \varepsilon)$$

$$\le \sum_{k=1}^{m} P(|f_n(\sigma_k; W_n) - p_k| > \varepsilon) \quad (*)$$

$$n f_n(\sigma_k, W_n) = \sum_{j=1}^{n} \mathbb{1}_{\{X_j = \sigma_k\}} \overset{d}{=} \mathrm{Bin}(n, p_k)$$

– mean $n p_k$

var $n p_k(1-p_k) < n p_k$

so $$(*) \le \frac{n p_k}{(\varepsilon n)^2}$$

$\triangle$

Sum over p_k to get $$\sum_k \frac{p_k}{n\varepsilon^2} \le \frac{1}{n\varepsilon^2}$$

Presented this way, the SLLN appears at the end. I'm not a fan.

Cor: Note $H(P)$ maximal if P – unif $\Rightarrow$ exponentially more words w/ unif. freq than any other.

Why called AEP?

Maximize $H(p)$ over prob vectors p. Use Lagrange multipliers:

$$E(p,\lambda) = H(p) - \lambda\left(\sum_{i=1}^{n} p_i - 1\right)$$

$$\frac{\partial E}{\partial p_i} = \log p_i + 1 - \lambda = 0 \Rightarrow p_i = \text{const}. \quad H(p) = \sum \frac{1}{m}\log_e m = \log m$$

Let $\vec{q} = (q_1, \dots, q_m)$ be another prob vector. Then with $p = (1/m, \dots, 1/m)$

$$H(\lambda p + (1-\lambda)q) = -\sum \left(\lambda p_i + (1-\lambda)q_i\right)\log\left(\lambda p_i + (1-\lambda)q_i\right)$$

$$\frac{dH}{d\lambda} = -\sum \cancel{(\lambda p_i + (1-\lambda)q_i)}\frac{(p_i - q_i)}{\cancel{\lambda p_i + (1-\lambda)q_i}} - \sum (p_i - q_i)\log(\ \ )$$

★1

$$= -\sum (p_i - q_i)\left(\log(\lambda p_i + (1-\lambda)q_i)\right)\Big|_{\lambda=1} = 0$$

$$= \sum (p_i - q_i)(1 + \log m) = 0$$ which is true.

★1

$$\frac{d^2H}{d\lambda^2} = -\sum (p_i - q_i)^2 \frac{1}{(\lambda p_i + (1-\lambda)q_i)} < 0$$ as long as $\exists$ some $q_i \neq p_i$

$\Rightarrow$ H is a convex fn with a unique max at $\lambda = 1$. $\Rightarrow$ Uniform meas. unique max.

Jensen inequality proof:

$$\sum p_i \log 1/p_i \leq \log\left(\sum_{i=1}^{m} \frac{1}{p_i} p_i\right) = \log m.$$ This also gives equality iff

$\varphi(x) = cx + d$. Here $x \in \{1, \dots, m\}$ $\varphi(i) = \log 1/p_i = ci + d$

$$\Rightarrow p_i = e^{-ci+d} = Ae^{-ci} \Rightarrow \sum p_i(ci + d) = E_{p_i}[x] + d$$

Is it possible to show $c = 0$, then $d = 1/m$?

Reinterpret: Consider space of words w/ indep letters distr. according to P.
Let W be a random word.
If W is ε-typical for P, then say

$$e^{-n(H(P)+\varepsilon)} \le P(W) \le e^{-n(H(P)-\varepsilon)}$$

so

$$P\left(e^{-n(H(P)+\varepsilon)} \le P(W) \le e^{-n(H(P)-\varepsilon)}\right) \geq P(W \text{ is } \varepsilon\text{-typical}) \geq 1 - \frac{1}{n\varepsilon^2}$$

i.e.

$$P\left(\left|\frac{-\ln P(W)}{n} - H(P)\right| \le \varepsilon\right) \xrightarrow[n\to\infty]{} 1$$

so

$$\frac{-\ln(P(W))}{n} \xrightarrow[n\to\infty]{P} H(P)$$

so $\exists$ set of almost m'ze 1 s.t. on that set at the log scale all pts are "almost equally distributed".

Q: Suppose X is not discrete, then how would you generalize this theorem?

Rem: Here LLN can be applied to $\log p(X_i)$
In ergodic theory, ergodic theorem applied to $\log p(X_i)$

How do you construct dist of a r.v. if you're only allowed to sample from it?

3) Glivenko-Cantelli   Consider cdf F.

Let $\{X_i\}_{i=1}^{\infty}$ be iid. w/ cdf F.

Def: The empirical distr of $X_1, \dots, X_n$ is $F_n(x,\omega) = \frac{1}{n} \sum_{k=1}^{n} 1_{\{X_k \le x\}}(\omega)$.

i.e. F_n is distr. that gives "equal wt" to each X_i.

F_n is a Random distr fcn.

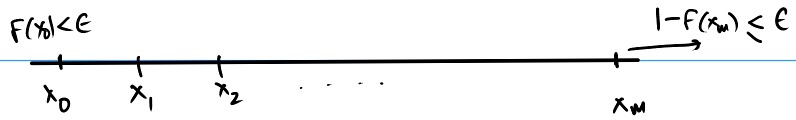
(In some sense this is a good preview for Monte Carlo integration too)

Thm: (Glivenko-Cantelli)

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = 0.$$

So the "sampled" distr cv to actual.

Pf:



For each fixed $x_1 \dots x_m$

$$\frac{1}{n} \sum_{k=1}^{n} 1_{\{X_k \le x_i\}} \to F(x_i) \qquad \text{a.s.}$$

One can pick an n large enough st

$$|F_n(x_i,\omega) - F(x_i)| < \epsilon \qquad i = 0, \dots, m$$

Only need to control intermediate x.

If $\quad x_i < x < x_{i+1}$

$$F_n(x_i) < F_n(x) < F_n(x_{i+1})$$

$$F(x_i) < F(x) < F(x_{i+1})$$

So it's clear that if we arrange for $F(x_{i+1}) - F(x_i) < \epsilon$

we must have

$$F_n(x_i) - F(x) \le F_n(x) - F(x)| \le F_n(x_{i+1}) - F(x)$$

Of course this is not possible!

BUT this proof is wrong in Khoshnevisan too, which is what I originally followed.

He claims, at least in the textbook version that if

$$\sup_{x_i \le x < x_{i+1}} |F(x) - F(x_i)| < \epsilon \quad \text{then} \quad F(x_i) \le F(x_{i-1}) + \epsilon .$$

This is not true, of course. What if x_i is at a jump?

However, we can ensure that $\frac{1}{n}\sum_{j=1}^{n} 1_{\{x_i \le X_j < X_{i+1}\}} < 2\epsilon$

Again using the SLLN. This is true for each $i = 0, \dots, m-1$, and all large enough n (a.s. ω). Then

$$|F_n(x,\omega) - F_n(x_i,\omega)| = \left|\frac{1}{n}\sum_{i=1}^{n} 1_{\{x_i \le X_j < X_{i+1}\}}\right| < 2\epsilon$$

$$\forall \; x \in [x_i, x_{i+1})$$

Therefore $|F_n(x) - F(x)| \le |F_n(x_i) - F_n(x)| + |F_n(x_i) - F(x)|$

$$\le 2\epsilon + |F_n(x_i - F(x)|$$

$$\le 2\epsilon + |F_n(x_i) - F(x_i)| + |F(x_i) - F(x)| \le 4\epsilon \quad \forall\; x \in [x_i, x_{i+1})$$

For $x < x_0$ we have $\frac{1}{n}\sum_{j=1}^{n} 1_{\{X_j \le x\}} = F_n(x,\omega) \le F_n(x_0,\omega) < 2\epsilon$

by monotonicity. $\Rightarrow$ $|F_n(x,\omega) - F(x)| \le 3\epsilon$ since $F(x) \le \epsilon$.

The case of $x > x_m$ is similar.

How can one ensure that such x_i can be found:

There can only be countably many pts of discontinuity of F since it is non-decreasing.

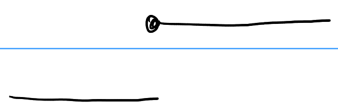
☆ Let $[0,1] = [0,\epsilon) \cup [\epsilon, 2\epsilon) \dots \cup [m\epsilon, 1]$ many disjoint pieces

where $m = \left\lfloor \frac{1}{\epsilon} \right\rfloor$. Let $x_m = \inf\{x : F(x) \geq m\epsilon\}$

Inductively define $x_i = \inf\{x < x_{i+1} : (m-i)\epsilon \leq F(x) < (m-i+1)\epsilon\}$

By right continuity these x_i exist, unless it is $-\infty$, in which case stop, and relabel them st $x_0 = x_{i+1}, x_1 = x_{i+2} \dots, x_{m'} = x_m$

Example:



$\epsilon = 1/2 \Rightarrow x_2 = 1 \quad x_1 = \inf\{x < x_2 : \frac{1}{2} \leq F(x) < 1\}$

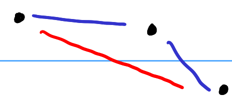
$= -\infty$.

Erdős bound on Ramsey numbers

K_m = complete graph

Fix n. $R_n = \inf\{N :$ any bichromatic coloring of K_N yields a $K_n \subseteq K_N$ st all edges have the same color$\}$

$R_2 = 3$



$R_3 = 6$

Ramsey introduced these in 1930 in consistency checking of logical formulas. (Something to do with k-sat?)

Ramsey in fact proved that $R_n < \infty$ $\forall n \geq 1$.

Thm: As $n \to \infty$ $R_n \geq (c + o(1))\, n 2^{n/2}$ $\quad \frac{1}{c} = e\sqrt{2}$

Pf: The proof is adversarial. If for some N, you find a coloring st all subsets K_n are not monochromatic, then $R_n \geq N$.

How do you find such a coloring? 1) write an algorithm?

2) COLOR the edges of K_N Randomly (N is some fixed #)

Whats the prob that a fixed subgraph of size n is edge-monochrome?

A: $\frac{1}{2^{\binom{n}{2}}} \cdot 2$ ← Blue or red

$P(\exists$ a subset of n vertices having same color$) \le \binom{N}{n} \frac{1}{2^{\binom{n}{2}}}$

$= E\left[\#\text{ of subgraphs that are monochrome}\right] < 1$

$\Rightarrow P(\text{All subgraphs are NOT monochromatic}) > 0$

$\Rightarrow \exists$ a coloring of K_N st all subgraphs are NOT monochrome.

$$\binom{N}{n} \frac{1}{2^{\binom{n}{2}}} = \frac{N!}{n!\,(N-n)!} \frac{1}{2^{\binom{n}{2}}} \le \frac{N^n}{2^{n^2/2 - n/2}} \frac{1}{n^{n+1/2} e^{-n} \sqrt{2\pi}}$$

$$= \left(\frac{\sqrt{2} N e}{2^{n/2} n} \right)^n \frac{1}{\sqrt{n}} \frac{1}{\sqrt{2\pi}}$$

Simply choose $\quad N = \left\lfloor 2^{n/2} n \frac{1}{\sqrt{2}\, e} \right\rfloor$

4) <u>Percolation</u>: $\mathbb{Z}^d$-graph $p \in (0,1)$

delete edges independently w/ prob $1-p$.

We say there is percolation if $\exists$ infinite connected subgraph.

(physical interpretation as water passing through - percolating)

If E - set of edges $X_e(p) := 1$ or 0 w/ prob $p, 1-p$ resp.

$\uparrow$ not $\quad \uparrow$ deleted

percolation - tail event wrt to $\{X_e\}_{e \in E}$ so Kolmogorov's 0-1 law says $P(\text{percolation}) = 0$ or 1.

Which is it?

<u>Thm</u>: $\forall d \geq 2$ $\exists$ critical prob. $p_c(\mathbb{Z}^d) \in [0,1]$ s.t. if $p < p_c(\mathbb{Z}^d)$, $P(\text{perco}) = 0$ & if $p > p_c(\mathbb{Z}^d)$, $P(\text{perco}) = 1$.

($d=1$ trivial)

$d=2$ $\quad p_c = \frac{1}{2}$ $\quad$ other lattices or dimensions hard.

<u>pf</u>: <u>Monotonicity argument</u>: construct all $X_e(p)$'s for diff. p's on the same sp.

let $\{U_e\}_{e \in E}$ be iid unif. on $[0,1]$.

let $X_e(p) := \mathbb{1}_{U_e \in [0,p]}$. Given p, $\{X_e(p)\}_{e \in E}$ are iid $Ber(p)$.

However $X_e(p) \leq X_e(r) \; \forall e \in E$ if $p \leq r$.

let $\Gamma(p) := \{ e \in E \,;\, X_e(p) = 1 \}$.

Then $\Gamma(p)$ is a subgraph of $\Gamma(r)$ if $p \leq r$.

$\Rightarrow$ if $\Gamma(p)$ has ∞ comp, so does $\Gamma(r)$ $\forall r \geq p$.

$\inf_p \{ \Gamma(p) \text{ has } \infty \text{ comp} \}$ is the crit. prob $p_c(\mathbb{Z}^d)$ $\square$

This sort of idea ties in well to Hausdorff measure if you gave it as an exercise.

5) Monte-Carlo integration

Given $\phi : \mathbb{R}^n \to \mathbb{R}$, NTF $I(\phi) = \int_{[0,1]^n} \phi(x)\,dx$.

Can be very hard. Approximate as follows.

$X_1, \dots, X_n$ - iid uniform from $[0,1]^n$. $\phi(X_1), \dots, \phi(X_n)$ - random, iid.

$E\,\phi(X_i) = I(\phi)$.

$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} \phi(X_i) \xrightarrow[n\to\infty]{a.s.} I(\phi)$

So if pick large # of $X_1, \dots, X_n$, $\frac{1}{n}\sum_{i=1}^{n}\phi(X_i)$ good approx. of $I(\phi)$

Best if n - large. easy to compute