Notes
21-721 Probability
Name: Shashank Singh[1]

1/13/14 - 1/15/14: Branching Process Example

1/17/14: Measure Theory (Dynkin and Caratheodory)

Lemma: $\tau \subseteq \mathcal{P}(S)$ is a σ-algebra iff it is both a π-system and a d-system.

Proof: $\Rightarrow$ is trivial. Complement requires finiteness. Arbitrary union is an increasing union.

Lemma: If $\tau \subseteq \mathcal{P}(S)$ is a π-system, then $d(\tau) = \sigma(\tau)$.

Proof: Show that $\mathcal{A}_1 := \{B \in d(\tau) : B \cap C \in d(\tau), \forall C \in \tau\}$ and $\mathcal{A}_2 := \{B \in d(\tau) : B \cap C \in d(\tau), \forall C \in d(\tau)\}$ are d-systems.

Theorem: (Dynkin's Lemma) Suppose $\tau \subseteq \mathcal{P}(S)$ is a π-system. Then, for any measures $\mu_1, \mu_2 : \sigma(\tau) \to [0, \infty]$, if $\mu_1 = \mu_2$ on τ and $\mu_1(S) = \mu_2(S) < \infty$, then $\mu_1 = \mu_2$ on $\sigma(\tau)$.

Proof: Consider two extensions μ_1 and μ_2. Show $\mathcal{A} := \{F \in \Sigma : \mu_1(F) = \mu_2(F)\}$ is a d-system.

Lemma: If $\mathcal{G}_0$ is an algebra and $\lambda : \mathcal{G}_0 \to [0, \infty]$ with $\lambda(\emptyset) = 0$ then $\mathcal{L}$ is an algebra, λ is finitely additive on $\mathcal{L}$.

Proof: Tricks with clever splits to show $\mathcal{L}$ is closed under finite intersections. Obviously, $\mathcal{L}$ is closed under complement, so $\mathcal{L}$ is an algebra. More clever splits to show finite additivity.

Lemma: (Caratheodory's Lemma) Let λ be an outer measure on $(S, \mathcal{G})$. Let

$$\mathcal{L} := \{A \in \mathcal{G} : \lambda(B) = \lambda(A \cap B) + \lambda(A^c \cap B), \forall B \in \mathcal{G}\}.$$

Then, $\mathcal{L}$ is a σ-algebra, and λ is countably additive on $\mathcal{L}$.

Proof: By the previous lemma, just need to show λ is countably additive, i.e., for L_n disjoint, $L := \bigcup_n L_n \in \mathcal{L}$ and

$$\lambda\left(\bigcup_n L_n\right) = \sum_n \lambda(L_n).$$

Let $G \in \mathcal{G}$. By subadditivity, $\lambda(G) \leq \lambda(L \cap G) + \lambda(L^c \cap G)$. To show $\geq$, examine $M_n := \bigcup_{k \leq n} L_k$.

Theorem: (Caratheodory's Extension Theorem) Let Σ_0 be an algebra on S, and let $\sigma := \sigma(\Sigma_0)$. If $\mu_0 : \Sigma_0 \to [0, \infty]$ is countably additive, then $\exists$ a measure μ on (S, Σ) with $\mu = \mu_0$ on Σ_0.

Proof: Define $\lambda : \mathcal{P}(S) \to [0, \infty]$ by

$$\lambda(G) := \inf\left\{\sum_n \mu_0(F_n) : G \subseteq \bigcup_n F_n, F_n \in \Sigma_0\right\}.$$

[1] sss1@andrew.cmu.edu

There are 4 steps:

1. Show λ is an outer measure.
2. Observe that, by Caratheodory's Lemma, show $\Sigma_0 \subseteq \mathcal{L}$ and $\lambda = \mu_0$ on Σ_0.
3. Show $\Sigma_0 \subseteq \mathcal{L}$.
4. Show $\lambda = \mu_0$ on Σ_0.

1/22/14: Measurable Functions, Monotone Class Theorem

Definitions:
$m\Sigma := \{h : S \to \mathbb{R} : h^{-1} : \mathcal{B} \to \Sigma\}$ is the set of measurable functions.
$(m\Sigma)^+ := \{h \in m\Sigma : h \geq 0\}$
$b\Sigma := \{h \in m\Sigma : \exists c \in \mathbb{R}, |h| \leq b\}$
For $X \in m\Sigma$, $\sigma(X) := X^{-1}(\mathcal{B})$ (i.e., $A \in \sigma(X) \Leftrightarrow \exists B \in \mathcal{B}, A = X^{-1}(B)$).

Theorem: (Monotone Class Theorem) Suppose $\mathscr{H} \subseteq b\Sigma$ with

1. $\mathscr{H}$ is a vector space
2. $h \in (b\Sigma)^+, (h_n) \subseteq \mathscr{H} \cap (b\Sigma)^+$, and $h_n \uparrow h \Rightarrow h \in \mathscr{H}$
3. $1 \in \mathscr{H}$
4. $\exists$ a π-system $\tau \subseteq \Sigma$ with $\sigma(\tau) = \Sigma$ and, $\forall F \in \tau, 1_F \in \mathscr{H}$.

Then, $\mathscr{H} = b\Sigma$.

Proof: The main idea is to show $\mathscr{H}$ contains indicator functions (using Dynkin's Lemma), then show $(b\Sigma)^+ \subseteq \mathscr{H}$ (using (APPROX)), and then show $b\Sigma \subseteq \mathscr{H}$.

Step 1: Show, $\forall F \in \Sigma, 1_F \in \mathscr{H}$. To do this, show $\mathcal{A} := \{F \in \Sigma : 1_F \in \mathscr{H}\}$ is a d-system ($\tau \subseteq \mathcal{A}$).

Step 2: $h \in (b\Sigma)^+$. For $n \in \mathbb{N}$, put

$$h_n := \sum_{k=0}^{n-1} \frac{k}{n} 1_{\frac{k}{n} < h < \frac{k+1}{n}} \in \mathscr{H},$$

by Step 1, since $\mathscr{H}$ is a vector space. Observe that $h_n \uparrow h \in \mathscr{H}$.

Step 3: If $h \in b\Sigma$, then $h = h^+ - h^- \in \mathscr{H}$. ■

Definition: Let $X \in m\mathcal{F}$ be a random variable. The *law* of X is the probability measure on $(\mathbb{R}, \mathcal{B})$ defined by $\mathcal{L}_X(B) := \mathbb{P}[X \in B]$, $\forall B \in \mathcal{B}$. The *distribution function* $F_X : \mathbb{R} \to \mathbb{R}$ is defined by $F_X(t) = \mathbb{P}[X \leq t]$, $\forall t \in \mathbb{R}$.

Lemma: $\mathcal{L}_X = \mathcal{L}_Y \Leftrightarrow F_X = F_Y$.

Proof: ($\Rightarrow$) Trivial. ($\Leftarrow$) $\mathcal{L}_X = \mathcal{L}_Y$ on a π-system. Use Dynkin's Lemma.

Lemma: (Skorohod's Construction) F is the distribution function of some random variable iff

1. $\lim_{t\to-\infty} F(t) = 0$ and $\lim_{t\to\infty} F(t) = 1$,
2. F is non-decreasing,
3. F is right-continuous ($\forall t \in \mathbb{R}, F(t) = \lim_{x\downarrow t} F(x)$).

Proof: ($\Rightarrow$) Trivial. ($\Leftarrow$) Define $X : [0,1] \to \mathbb{R}$ by

$$X(t) := \inf\{x \in \mathbb{R} : t < F(x)\}.$$

Since X is increasing, X is necessarily Borel measurable. X is a right-inverse of F except on the (at most) countable set

$$\tau_0 := \{t : X(t) - \lim_{y\downarrow t} X(y) > 0\}$$

of discontinuities of F. Hence,

$$\{[F(X(t)) = t] = 1,$$

and

$$\mathbb{P}[X(t) \leq x] = \mathbb{P}[t \leq F(x)] = F(x).$$

1/24/14: Independence

Lemma: If $\mathcal{A}$ and $\mathcal{B}$ are π-systems with $\mathcal{A} \perp \mathcal{B}$, then $\sigma(\mathcal{A}) \perp \sigma(\mathcal{B})$.

Proof: Apply Dynkin's Lemma twice.

Corollary: RV's X and Y are independent iff, $\forall x, y \in \mathbb{R}$, $\mathbb{P}[X \leq x, Y \leq y] = F_X(x)F_Y(y)$.

1/20/14: Borel-Cantelli Lemmas; Kolmogorov 0-1 Law

Lemma: (1^{st} Borel-Cantelli) $(F_n) \subseteq \mathcal{F}$, $F = \{F_n \text{ i.o.}\}$. Then,

$$\sum_n \mathbb{P}[F_n] < +\infty \Rightarrow \mathbb{P}[F] = 0.$$

Proof: $F = \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty F_k$. Then, for $G_n := \bigcup_{k=n}^\infty F_k$, $G_n \downarrow F$. Hence,

$$\mathbb{P}[F] = \lim_{n\to\infty} \mathbb{P}[G_n] \leq \lim_{n\to\infty} \sum_{k=n}^\infty \mathbb{P}[F_k] = 0. \quad \blacksquare$$

Lemma: (2^{nd} Borel-Cantelli) $(F_n) \subseteq \mathcal{F}$, $F = \{F_n \text{ i.o.}\}$, F_n independent. Then,

$$\sum_n \mathbb{P}[F_n] = +\infty \Rightarrow \mathbb{P}[F] = 1.$$

Proof: Note $\{E_n \text{ i.o.}\}^c = \{E_n^c \text{ ev}\}$, so show $\mathbb{P}\{E_n^c \text{ ev}\} = 0$, so show each $\mathbb{P}\left[\bigcap_{n=m}^\infty E_n^c\right] = 0$. Since the E_n's are independent, using the inequality $1 - x \leq e^{-x}$,

$$\mathbb{P}\left[\bigcap_{n=m}^\infty E_n^c\right] = \prod_{n=m}^\infty \mathbb{P}\left[E_n^c\right] = \prod_{n=m}^\infty \left(1 - \mathbb{P}\left[E_n\right]\right) \leq \prod_{n=m}^\infty \exp\left(-\mathbb{P}\left[E_n\right]\right) \leq \exp\left(-\sum_{n=m}^\infty \mathbb{P}\left[E_n\right]\right) = 0. \quad \blacksquare$$

Theorem: (Kolmogorov 0-1 Law) (X_n) IRV's. Define the tail σ-algebra

$$\tau := \bigcap_{m=1}^{\infty} \sigma(X_m, X_{m+1}, \dots).$$

Then

1. τ is $\mathbb{P}$-trivial (i.e., $\forall E \in \tau$, $\mathbb{P}(E) \in \{0, 1\}$),
2. every $X \in m\tau$ is $\mathbb{P}$-constant (i.e., $\exists c \in \mathbb{R}$ such that $\mathbb{P}[X = c] = 1$).

Proof: $\forall n \in \mathbb{N}$, define $\mathcal{X}_n := \sigma(X_1, \dots, X_n)$ and $\tau_n := \sigma(X_n, X_{n+1}, \dots,)$. Clearly, $\forall n < m$,

$$\mathcal{X}_n \perp \tau_m \quad \Rightarrow \quad \mathcal{X}_n \perp \cap_m \tau_m = \tau \quad \Rightarrow \quad \bigcup_n \mathcal{X}_n \perp \tau.$$

Thus,

$$\tau \subseteq \sigma\left(\bigcup_n \mathcal{X}_n\right) \perp \tau,$$

Hence, $\tau \perp \tau$, and so $\forall E \in \tau$, $\mathbb{P}[E] = \mathbb{P}[E \cap E] = \mathbb{P}[E]^2 \in \{0, 1\}$. ■

1/29/14: Lebesgue Integration

Define the space of simple function

$$SF^+ := \{f \in m\Sigma : f = \sum_{k=1}^{n} a_n 1_{A_n}, a_n \geq 0, A_n \in m\Sigma\},$$

and define, $\forall f \in SF^+$,

$$\mu_0(f) := \sum_{k=1}^{n} a_n \mu(A_n) \in [0, \infty] \quad \text{(where } 0 \cdot \infty := 0\text{)}.$$

We have the following properties:

1. (Well-posedness) $f, g \in SF^+, \mu\{f \neq g\} = 0 \Rightarrow \mu(f) = \mu(g)$.
2. (Positive Linearity) $f, g \in SF^+, a, b \geq 0 \Rightarrow \mu_0(af + bg) = wa\mu_0(f) + b\mu_0(g)$.
3. (Order) $f, g \in SF^+, f \geq g \Rightarrow \mu(f) \geq \mu(g)$.

1/31/14: Convergence Theorems

Monotone: $(f_n) \subseteq m\Sigma^+$ ptwise non-decreasing and $f_n \to f$ ptwise $\Rightarrow \lim_{n\to\infty} \mu(f_n) \to \mu(f)$.

Fatou: $(f_n) \subseteq m\Sigma^+ \Rightarrow \liminf_n \mu(f_n) \geq \mu(\liminf_n f_n)$ (*Proof:* by Monotone)

Reverse Fatou: $(f_n) \leq g \subseteq m\Sigma^+, \mu(g) < \infty \Rightarrow \liminf_n \mu(f_n) \geq \mu(\liminf_n f_n)$ (*Proof:* $g - f_n$)

Dominated: $(f_n) \subseteq m\Sigma, f_n \to f$ ptwise, $|f_n| \leq g \in \mathcal{L}_1 \Rightarrow \mu(|f_n - f|) \to 0$, and $\mu(f_n) \to \mu(f)$. (*Proof:* by Reverse Fatou, $\limsup_n \mu(f_n - f)| \leq \mu(0) = 0$ and

$$|\mu(f_n) - \mu(f)| = |\mu(f_n - f)| \leq \mu(|f_n - f|) \to 0.)$$

Scheffe: $f_n, f \in \mathcal{L}_1, f_n \to f$. Then, $\mu(|f_n - f|) \to 0 \Leftrightarrow \mu(|f_n|) \to \mu(|f|)$. (*Proof:* ($\Rightarrow$) trivial. ($\Leftarrow$) If all $f_n \geq 0$, then $|f_n - f| = f_n + f - 2f_n \wedge f$, and so, by (DOM), $\mu(|f_n - f|) \to 0$. In general, by Fatou, $\limsup \mu(f_n^+) \geq \mu(f^+)$ and $\limsup(\mu(f_n^-) \geq \mu(f^-)$, and, since $|f| = f^+ + f^-$, we have equality. Now use step 1.)

Bounded:

2/3/14: Inequalities

Markov's Inequality: $f \in (m\mathcal{B})^+$ non-decreasing, $X \in (m\mathcal{F})^+$. Then, $\forall c > 0$,

$$\mathbb{E}[f(X)] \geq \mathbb{E}[f(X)1_{\{X \geq c\}}] \geq f(c)\mathbb{P}[X \geq c].$$

Chebyshev's Inequality: $X \in \mathcal{L}_2$. Then, $\forall \varepsilon > 0$, by Markov's Inequality,

$$\mathbb{P}[|X - \mathbb{E}[X]| > \varepsilon] \leq \frac{1}{\varepsilon^2}\mathbb{E}[(X - \mathbb{E}[X])^2] \leq \frac{\mathbb{V}[X]}{\varepsilon^2}.$$

Jensen's Inequality: $X \in \mathcal{L}_1$, $\varphi : \mathbb{R} \to \mathbb{R}$ convex with $\varphi(X) \in \mathcal{L}_1$. Then, $\mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}[X])$.

Proof: Since φ is convex, using a tangent (subgradient) approximation, $\varphi(X) \geq \varphi(\mathbb{E}[X]) + M\mathbb{E}[X](X - \mathbb{E}[X])$, and so $\mathbb{E}[\varphi(X) \geq \varphi(\mathbb{E}[X]) + M\mathbb{E}[X] \cdot 0 = \varphi(\mathbb{E}[X])$. ∎

Hölder's Inequality (Special Case): $0 < p \leq r$. Then $\|X\|_p \leq \|X\|_r$.

Proof: Since $p \leq r$, the function $x \to x^{r/p}$ is convex. Hence,

$$\|X\|_p = \mathbb{E}[|X|^p]^{1/p} = \left(\mathbb{E}[|X|^p]^{r/p}\right)^{1/r} \leq \mathbb{E}\left[(|X|^p)^{r/p}\right]^{1/r} = \mathbb{E}[|X|^r]^{1/r} = \|X\|_r. \quad \blacksquare$$

Hölder's Inequality: $\frac{1}{p} + \frac{1}{q} = 1$. Then, $\|XY\|_1 \leq \|X\|_p\|Y\|_q$.

Proof: TODO

Minkowski's Inequality: $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$.

Proof: If $X, Y \geq 0$, then by Hölder's Inequality,

$$\|X + Y\|_p^p = \mathbb{E}[|X||(X + Y)|^{p-1}] + \mathbb{E}[|X||(X + Y)|^{p-1}] \leq (\|X\|_p + \|Y\|_p)\left(\frac{\|X + Y\|_p^p}{\|X + Y\|_p}\right). \quad \blacksquare$$

2/5/14: $\mathcal{L}^p$ Spaces; Projections in $\mathcal{L}^p$

Theorem: $(\mathcal{L}_p, \|\cdot\|_p)$ is a Banach space.

Proof: Prior considerations imply that $\mathcal{L}_p$ is a linear space and $\|\cdot\|_p$ is a norm on $\mathcal{L}_p$. Hence, it suffices to show that $\mathcal{L}_p$ is complete under $\|\cdot\|_p$.

Step 1: ((X_n) converges)

Step 2: ($\lim_{n\to\infty} X_n \in \mathcal{L}_p$)

Theorem: ($\mathcal{L}^2$ projections) If $\mathcal{K}$ is a closed linear manifold in $\mathcal{L}_2$ and $X \in \mathcal{L}_2$, then $\exists! Y \in \mathcal{K}$ with (a) $\|X - Y\| = \inf\{\|X - Z\| : Z \in \mathcal{K}\}$ and (b) $X - Y \perp \mathcal{K}$.

Proof: Use the Parallelogram Law

$$\|X\|^2 + \|Y\|^2 = \frac{1}{2}\left(\|X - Y\|^2 + \|X + Y\|^2\right).$$

To show (b), use a 0 derivative argument.

2/7/14: Independence and $\mathcal{L}^p$

2/10/14: Product σ-algebra/measure; Fubini's Theorem

2/12/14: Canonical Model for Sequence of IRVs

2/14/14: Conditional Expectation

Definition: $\mathcal{G}$ sub-σ-algebra of $\mathcal{F}$, $X \in \mathcal{L}_1(\mathcal{F})$

$$Y = \mathbb{E}[X|\mathcal{G}] \Leftrightarrow Y \in \mathcal{L}_1(\mathcal{G}) \quad \text{and} \quad \mathbb{E}[X1_F] = \mathbb{E}[Y1_F], \quad \forall F \in \mathcal{G}.$$

Theorem: $X \in \mathcal{L}_1(\mathcal{F}), \mathcal{G} \subseteq \mathcal{F}$. Then, $\mathbb{E}[X|\mathcal{G}]$ exists and is unique.

2/17/14: Conditional Probability

2/19/14: Martingales; Discrete Stochastic Integral and Stopping Times

---

Definition: If (X_n) is adapted and (H_n) is predictable, then the stochastic integral of (H_n) with respect to (X_n) is $(H \cdot X)_n := \sum_{k=1}^n H_k(X_k - X_{k-1})$.

Lemma: If (X_n) is a martingale, (H_n) is predictable, and each $X_n H_n \in \mathcal{L}_1$, then $((H \cdot X)_n)$ is a martingale.

Proof: Note that $\mathbb{E}_n[H_{n+1}X_{n+1}] = H_{n+1}\mathbb{E}[X_{n+1}] = H_{n+1}X_n \in \mathcal{L}_1$. Hence,

$$\mathbb{E}_n[(H \cdot X)_{n+1}] = (H \cdot X)_n + \mathbb{E}_n\left[H_n(X_{n+1} - X_n)\right] = (H \cdot X)_n$$

Lemma: If (X_n) is a supermartingale and τ is a stopping time, then $(X_{n\wedge\tau})$ is a supermartingale.

Proof: Note that $X_{(n+1)\wedge\tau} = \sum_{k=1}^n X_k 1_{k=\tau} + X_{n+1}1_{\tau>n}$. Hence, casing on τ,

$$\mathbb{E}_n[X_{(n+1)\wedge\tau}] = \sum_{k=1}^n \mathbb{E}_n\left[X_k 1_{k=\tau}\right] + \mathbb{E}_n\left[X_{n+1}1_{\tau>n}\right] = X_{n\wedge\tau}$$

2/21/14: Martingale Convergence Theorem; Doob's Upcrossing Lemma

---

Lemma: (Doob's Upcrossing Lemma) (X_n) supermartingale. Then, $\forall a < b \in \mathbb{R}$, defining $U_n[a,b] := \#\{$up crossings of [a,b] in [0,n]$\}$

$$(b-a)\mathbb{E}[U_N[a,b]] \leq \mathbb{E}_n[X_n - a].$$

Proof: Define $H_n := 1_{\{\text{we are upcrossing at } t = N\}}$. By the supermartingale property, $\mathbb{E}[H \cdot X]_n \leq 0$. Then,

$$(b-a)\mathbb{E}[U_n[a,b]] \leq \mathbb{E}[H \cdot X]_n + (X_n - a)^- \leq \mathbb{E}(X_n - a)^-. \quad \blacksquare$$

Theorem: (Supermartingale Convergence Theorem) (X_n) supermartingale bounded in $\mathcal{L}_1$ (i.e., $\sup_n \mathbb{E}[X_n] < +\infty$). Then, $\exists X_\infty \in \mathcal{L}_1$ with $X_n \to X_\infty$ a.s.

Proof: Note that

$$\{X_n \text{ does not converge}\} = \bigcup_{a,b\in\mathbb{Q}} \left\{\liminf_n X_n < a < b < \limsup_n X_n\right\} = \bigcup_{a,b\in\mathbb{Q}} \{U_\infty[a,b] = \infty\}$$

It follows from Doob's Upcrossing Lemma that each

$$(b-a)\mathbb{E}[U_\infty[a,b]] \leq \sup_n \mathbb{E}[X_n] + a < +\infty$$

and hence $\mathbb{P}[U_\infty[a,b] = \infty] = 0$. Thus, $X_n \to X_\infty$ a.s. Furthermore, $X_\infty \in \mathcal{L}_1$, since, by Fatou,

$$\mathbb{E}[X_\infty] = \mathbb{E}\left[\liminf X_n\right] \leq \liminf \mathbb{E}[X_n] < +\infty. \quad \blacksquare$$

Corollary: (X_n) non-negative supermartingale. Then, $\exists X_\infty \in \mathcal{L}_1$ such that $X_n \to X_\infty$ a.s. and $\mathbb{E}[X_\infty|\mathcal{F}_n] \leq X_n$.

Proof: By the supermartingale property, $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0]$, and so $X_n \to X_\infty \in \mathcal{L}_1$ a.s. Hence, by (cFATOU),

$$X_n \geq \liminf_m \mathbb{E}_n[X_{n+m}] \geq \mathbb{E}_n\left[\liminf_m X_{n+m}\right] = \mathbb{E}_n\left[X_\infty\right]. \quad \blacksquare$$

2/24/14: Convergence of $\mathcal{L}_2$ Martingales

Theorem: If (M_n) is a martingale, then

$$\sup_n \mathbb{E}[M_n^2] < \infty \quad \Leftrightarrow \quad \exists M_\infty \in \mathcal{L}_2 \text{ such that each} M_n = \mathbb{E}[M_\infty].$$

Moreover, in this case, $M_n \to M_\infty$ in $\mathcal{L}_2$ and a.s.

Proof: ($\Leftarrow$) By the Conditional Jensen's Inequality, $M_n^2 = (\mathbb{E}_n[M_n])^2 \leq \mathbb{E}_n[M_n^2]$, and so $\mathbb{E}[M_n^2] \leq \mathbb{E}[M_\infty^2] < +\infty$.

Since expected cross terms of sum of increments are 0,

$$M_0^2 + \sum_{i=1}^n \mathbb{E}[(M_k - M_{k-1})^2] = \mathbb{E}[M_n^2] \leq \sup_k \mathbb{E}[M_k^2] < +\infty.$$

Hence, $\mathbb{E}[(M_m - M_n)^2] \to 0$ as $m > n \to \infty$. Since $\mathcal{L}_2$ is a complete metric space, $M_n \to M_\infty \in \mathcal{L}_2$ as $n \to \infty$. It remains to show $M_n = \mathbb{E}_n[M_\infty]$ a.s.

TODO

2/26/14: Convergence of Sums of IRVs in $\mathcal{L}_2$

Theorem: In $(X_n) \subseteq \mathcal{L}_2$ are IRVs with $\mu_n := \mathbb{E}[X_n]$ and $\sigma_n^2 := \mathbb{V}[X_n]$, then

(a) $\sum_n \mu_n \to$ and $\sum_n \sigma_n^2 \to$ imply $\sum_n X_n \to$ a.s.

(b) $|X_n| < K$ and $\sum_n X_n \to$ a.s. imply $\sum_n \mu_n \to$ and $\sum_n \sigma_n^2 \to$

Proof: Previously, we showed the case when each $\mu_n = 0$. Hence (a) is trivial, since

$$\sum_n X_n = \sum_n \mu_n + \sum_n (X_n - \mu_n) \to \text{ a.s.},$$

as $E[X_n - \mu_n] = 0$.

To prove (b) we use "symmetrization". TODO

Theorem: (Kolmogorov's Three Series) (X_n) IRVs. Then, $\sum_n X_n \to$ a.s. if and only if $\exists k > 0$ such that, for $X_n^k := X_n 1_{\{|X_n| \leq k\}}$, the following three series converge:

$$\sum_n \mathbb{P}[|X_n| > k], \quad \sum_n \mathbb{E}[X_n^k], \quad \text{and} \quad \sum_n \mathbb{V}[X_n^k].$$

Proof: Note that
$$\sum_n X_n = \sum_n X_n^k + \sum_n X_n 1_{\{|X_n|>k\}}.$$
By the previous result,
$$\sum_n X_n^k \to \text{ a.s.} \quad \Leftrightarrow \quad \sum_n \mathbb{E}[X_n^k] \to \quad \text{and} \quad \sum_n \mathbb{V}[X_n^k] \to .$$
By Borel-Cantelli,
$$\sum_n X_n 1_{\{|X_n|>k\}} \to \quad \Leftrightarrow \quad \sum_n \mathbb{P}[|X_n| > k]. \quad \blacksquare$$

3/5/14: SLLN for $\mathcal{L}_2$ Martingales; Lévy's Generalization of Borel-Cantelli

---

Theorem: (SLLN for $\mathcal{L}_2$ Martingales) If (X_n) is a martingale with each $X_n \in \mathcal{L}_2$, then
$$\{\langle X\rangle_\infty = +\infty\} \subseteq \left\{ \frac{X_n}{\langle X\rangle_n} \to 0 \right\}.$$
Proof: TODO

Theorem: (Lévy's Generalization of Borel-Cantelli) If $\mathbb{E}_n \in \mathcal{F}_n$,
$$Z_n := \sum_{k=1}^n 1_{E_k}, \quad \text{and} \quad Y_k := \sum_{k=1}^n \mathbb{P}_{k-1}[E_k] = \sum_{k=1}^n \mathbb{P}\left[E_k|\mathcal{F}_{k-1}\right],$$
then
$$\{Y_\infty = +\infty\} \subseteq \left\{ \frac{Z_n}{Y_n} \to 1 \right\}.$$
Proof: TODO

3/24/14: Doob's Maximal Inequality; Law of the Iterated Logarithm

---

Lemma: Let (X_n) be a non-negative submartingale, Then, $\forall c > 0$,
$$c\mathbb{P}[\sup_{k\le n} X_k > c] \le \mathbb{E}[X_n 1_{\{\sup_{k\le n} X_k > c\}} \le \mathbb{E}[X_n].$$
Proof: Define the stopping time $\tau := \inf\{n : X_n \ge c\}$. Then,
$$\begin{aligned} c\mathbb{P}\left[\sup_{k\le n} X_k > c\right] \le c\mathbb{P}\left[\tau \le n\right] \le \mathbb{E}[X_t 1_{\tau\le n}] &= \sum_{k=0}^n \mathbb{E}\left[X_k 1_{\tau=k}\right] \\ &= \sum_{k=0}^n \mathbb{E}\left[\mathbb{E}_k[X_n 1_{\tau=k}]\right] \\ &= \sum_{k=0}^n \mathbb{E}\left[X_n 1_{\tau=k}\right] = \mathbb{E}\left[X_n 1_{\tau\le k}\right] \le \mathbb{E}\left[X_n\right]. \end{aligned}$$

Corollary: (Doob's Maximal Inequality) Suppose (M_n) is a martingale and $\varphi : \mathbb{R} \to [0, \infty)$ is non-negative and convex such that each $\varphi(M_n) \in \mathcal{L}_1$. Then,

$$c\mathbb{P}\left[\sup_{k \leq n} \varphi(M_n) > c\right] \leq \mathbb{E}[\varphi(M_n) 1_{\{\sup_{k \leq n} \varphi(M_k) > c\}} \leq \mathbb{E}[\varphi(M_n)].$$

Theorem: (Law of the Iterated Logarithm) Suppose (X_n) are IIDRV's with each $X_n \sim \mathcal{N}(0,1)$. Define $S_n := \sum_{k=1}^n X_k$. Then, almost surely,

$$\limsup_n \frac{S_n}{h(n)} = 1, \qquad \text{where} \quad h(n) = \sqrt{2n \log\log n}.$$

Proof: Notice that the result involves an upper bound and a lower bound: $\forall \varepsilon > 0$,

$$\mathbb{P}\left[S_n > (1+\varepsilon)h(n) \text{ i.o.}\right] = 0 \qquad \text{and} \qquad \mathbb{P}\left[S_n > (1-\varepsilon)h(n) \text{ i.o.}\right] = 1.$$

Lower Bound: Let $\kappa \in (1, (1+\varepsilon)^2)$. Recall that, by Doob's Maximal Inequality, $\forall c \in \mathbb{R}, \theta = c/n$,

$$\mathbb{P}\left[\sup_{k \le n} S_k \ge c\right] = \mathbb{P}\left[\sup_{k \le n} e^{\theta S_k} \ge e^{\theta c}\right] \le e^{-\theta c}\mathbb{E}\left[e^{\theta S_n}\right] = e^{-\theta c} e^{\theta^2 n/2} = e^{-c^2/(2n)}.$$

Hence,

$$\begin{aligned} b_n = \mathbb{P}\left[S_n > (1+\varepsilon)h(n)\right] &\le \mathbb{P}\left[\sup_{k \le \kappa^n} S_k > (1+\varepsilon)h\left(\kappa^{n-1}\right)\right] \\ &\le \exp\left(-\frac{\left((1+\varepsilon)h\left(\kappa^{n-1}\right)\right)^2}{2\kappa^n}\right) \\ &= \exp\left(-\frac{(1+\varepsilon)^2(\log(n-1) + \log\log\kappa)}{\kappa}\right) \asymp n^{-(1+\varepsilon)^2/\kappa}, \end{aligned}$$

and so $\sum_{n=1}^\infty a_n < +\infty$. The upper bound follows by the 1^{st} Borel-Cantelli Lemma.

Upper Bound: Let $N \in \mathbb{N}$ be sufficiently large such that

$$\gamma_N := \frac{1 - \varepsilon + \sqrt{1/N}}{\sqrt{1 - 1/N}} \in (0,1)$$

(since $\gamma_N \to 1 - \varepsilon$ as $N \to \infty$). The Lower Bound gives $\mathbb{P}\left[S_n + 2h(n) > 0 \text{ i.o.}\right] = 0$, so we show

$$\mathbb{P}\left[S_{N^{n+1}} - S_{N^n} \ge (1-\varepsilon)h(N^{n+1}) + 2h(N^n) \text{ i.o.}\right] = 1.$$

Notice that $S_{N^{n+1}} - S_{N^n} = \sqrt{N^{n+1} - N^n}\,\xi$, where $\xi \sim \mathcal{N}(0,1)$. A standard lower bound gives,

$$\mathbb{P}[\xi > x] \ge \left(x + x^{-1}\right)^{-1} \varphi(x), \quad \forall x > 0$$

(where $\varphi(x) = (2\pi)^{-1/2}\exp\left(-x^2/2\right)$). Also notice that

$$x(n) := \frac{(1-\varepsilon)h(N^{n+1}) + 2h(N^n)}{\sqrt{N^{n+1} - N^n}} = \gamma_N \sqrt{2\log\log n},$$

and thus,

$$\mathbb{P}\left[\xi \ge x(n)\right] = \left(\gamma_N\sqrt{2\log\log n} + \left(\gamma_N\sqrt{2\log\log n}\right)^{-1}\right)^{-1} \varphi(\gamma_N\sqrt{2\log\log n}) \asymp n^{\gamma_N},$$

and so, for the obvious choice of b_n, $\sum_{n=1}^\infty b_n = +\infty$. $\blacksquare$

3/26/14: Doob's Maximal $\mathcal{L}_p$-Inequality

Theorem: If (M_n) is a martingale, $p > 1$, $q = \frac{p}{p-1}$, then, for $M_n^* := \sup_{k \le n} |X_n|$,

$$\mathbb{E}\left[\sup_{k\le n} |M_n|^p\right]^{1/p} = \|M_n^*\|_p \le q\|M_n\|_p \quad \text{and} \quad \|M_\infty^*\| \le q \sup_n \|M_n\|_p.$$

Proof: Observe that the second inequality follows from the first by Monotone Convergence. Since (M_n) is a martingale and the function $x \mapsto |x|$ is convex, Doob's Maximal Inequality gives

$$\mathbb{P}\left[M_n^* \ge c\right] \le \frac{1}{c}\mathbb{E}\left[|M_n| 1_{\{M_k^* \ge c\}}\right].$$

Hence,

$$\begin{aligned}
\|M_n^*\|_p^p = \mathbb{E}\left[(M_n^*)^p\right] = \int_0^\infty \mathbb{P}\left[(M_n^*)^p \ge t\right] dt &= \int_0^\infty \mathbb{P}\left[M_n^* \ge t^{1/p}\right] dt \\
&= p\int_0^\infty \mathbb{P}\left[M_n^* \ge x\right] x^{p-1}\, dx \\
&\le p\int_0^\infty \mathbb{E}\left[|M_n| 1_{\{M_n^* \ge x\}}\right] x^{p-2}\, dx \\
&= p\mathbb{E}\left[|M_n| \int_0^{M_n^*} x^{p-2}\, dx\right] \\
&= \frac{p}{p-1}\mathbb{E}\left[|M_n|(M_n^*)^{p-1}\right] \le q\mathbb{E}\left[|M_n|^p\right] = q\|M_n\|_p^p. \quad \blacksquare
\end{aligned}$$

3/28/14: Kukutani's Theorem

Theorem: (Kukutani's Theorem) (X_n) non-negative IRV's with each $\mathbb{E}[X_n] = 1$. Then,

$$M_n := \prod_{k=1}^n X_k \to M_\infty \in \mathcal{L}_1$$

a.s., and, furthermore, for $a_n := \mathbb{E}[\sqrt{X_n}]$,

$$\begin{aligned}
&\prod_n a_n > 0 \Leftrightarrow \mathbb{E}[M_\infty] = 1 \\
\text{and} \quad &\prod_n a_n = 0 \Leftrightarrow M_\infty = 0 \quad \text{a.s.}
\end{aligned}$$

Proof: Define

$$N_n := \prod_{k=1}^n \frac{\sqrt{X_k}}{\mathbb{E}[\sqrt{X_k}]} = \frac{\sqrt{M_n}}{\prod_{k=1}^n a_k}.$$

Note that N_n is a martingale. Since (X_n) are independent and N_n is non-negative, $\sup_n \mathbb{E}[N_n] \leq 1$, and hence $N_n \to N_\infty \in \mathcal{L}_1$ a.s. Also, since each $a_k \leq 1$,

$$\mathbb{E}[N_n^2] = \frac{\mathbb{E}[M_n]}{\prod_{k=1}^n a_k} = \prod_{k=1}^n a_k^{-2} \uparrow \prod_k a_k^{-2},$$

and hence, if $\prod_n a_n > 0$, then

$$\prod_{k=1}^n a_k^{-2} = \mathbb{E}[N_n^2] \uparrow \mathbb{E}[N_\infty^2] = \frac{\mathbb{E}[M_\infty]}{\prod_n a_n^2} \quad \Rightarrow \quad \mathbb{E}[M_\infty] = 1.$$

On the other hand, if $\prod_n a_n = 0$, then

$$0 \leftarrow N_n \prod_{k=1}^n a_k = \sqrt{M_n} \to \sqrt{M_\infty} \quad \text{a.s.}$$

3/31/14: Radon-Nikodym Theorem

4/14/14: Inversion Theorems for Characteristic Functions

Theorem: (Lévy's Inversion Theorem) Let X be a random variable with characteristic function φ. For $a < b$,

$$\lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-i\theta a} - e^{-i\theta b}}{i\theta} \varphi(\theta)\, d\theta = \frac{1}{2}\left(\mathbb{P}(\{a\}) + \mathbb{P}(\{b\})\right) + \mathbb{P}((a,b)).$$

Proof: Since,

$$\left|\frac{e^{-i\theta a} - e^{-i\theta b}}{i\theta}\right| = \left|\frac{1}{\theta}\int_{\theta a}^{\theta b} e^{-ix}\, dx\right| \leq \frac{1}{\theta}\int_{\theta a}^{\theta b} |e^{-ix}|\, dx = b - a,$$

the integrand in question is bounded and so we can use Fubini's Theorem:

$$\lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-i\theta a} - e^{-i\theta b}}{i\theta} \varphi(\theta)\, d\theta = \mathbb{E}G_T(X; a, b),$$

where, since cos is even and sin is odd,

$$\begin{aligned} G_T(x; a, b) := \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{i\theta(x-a)} - e^{i\theta(x-b)}}{i\theta}\, d\theta &= \frac{1}{\pi}\int_0^T \frac{\sin(\theta(x-a)) - \sin(\theta(x-b))}{i\theta}\, d\theta \\ &= \operatorname{sign}(x-a)S(|x-a|T) - \operatorname{sign}(x-b)S(|x-b|T), \end{aligned}$$

for

$$S(u) := \frac{1}{\pi}\int_0^u \frac{\sin(x)}{x}\, dx.$$

By a homework problem, $S(u) \to 1/2$ as $u \to \infty$, and hence

$$G_T(x; a, b) \to \frac{1}{2} \left(\operatorname{sign}(x-a) - \operatorname{sign}(x-b) \right).$$

as $T \to \infty$. By the Bounded Convergence theorem, as $T \to \infty$,

$$\mathbb{E} G_T(X; a, b) \to \mathbb{E} \frac{1}{2} \left(\operatorname{sign}(x-a) - \operatorname{sign}(x-b) \right) = \frac{1}{2} \left(\mathbb{P}(\{a\}) + \mathbb{P}(\{b\}) \right) + \mathbb{P}((a,b)). \quad \blacksquare$$

Theorem: Suppose $\varphi_X \in \mathcal{L}_1$. Then, F_X is continuously differentiable and $f_X = F_X'$ is bounded. Furthermore, f_X is the Fourier transform of φ_X.

Proof: Since F_X is monotone, it is continuous a.e. Hence, let $a < b$ such that F_X is continuous at both a and b. Then, by the previous result (Lévy Inversion),

$$F_X(b) - F_X(a) = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-i\theta a} - e^{-i\theta b}}{i\theta} \varphi_X(\theta) \, d\theta,$$

and hence

$$F(b) - F(a) \le \frac{(b-a)}{2\pi} \int_{\mathbb{R}} |\varphi_X(\theta)| \, d\theta.$$

Thus, F_X is (Lipschitz) continuous. Furthermore, since

$$\frac{e^{-i\theta a} - e^{-i\theta b}}{i\theta(b-a)} \to \begin{cases} e^{-i\theta a} & \text{if } b \downarrow a \\ e^{-i\theta b} & \text{if } a \uparrow b \end{cases},$$

we have

$$f(x) = F'(x) = \int_{\mathbb{R}} e^{-i\theta x} \varphi_X(\theta) = \mathcal{F}[\varphi]. \quad \blacksquare$$

4/21/14: Characteristic Functions

Theorem: (Lévy's Convergence Theorem) (μ_n) probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with characteristic functions (φ_n). Suppose $\varphi_n \to \varphi$ pointwise and φ is continuous at 0 (i.e., $\lim_{\theta \to 0} \varphi(\theta) = 1$). Then, $\exists$ a probability measure μ with characteristic function φ and $\mu_n \to \mu$ weakly.

Proof: **Step 1 (verify tightness):** ($\Rightarrow \exists \mu$ with Characteristic Function φ). Let $\varepsilon > 0$. Since φ is continuous at 0, $\exists \delta > 0$ with $|\theta| < \delta \Rightarrow |1 - \varphi(\theta)| < \varepsilon/4$, so that

$$|2 - \varphi(\theta) - \varphi(-\theta)| \le |1 - \varphi(\theta)| + |1 - \varphi(-\theta)| < \varepsilon/2.$$

Hence,

$$\frac{1}{\delta} \int_0^{\delta} |2 - \varphi(\theta) - \varphi(-\theta)| \, d\theta < \varepsilon/2.$$

Since $\varphi_n \to \varphi$ pointwise, by (BDD), $\exists N \in \mathbb{N}$ such that, $\forall n > N$,

$$\begin{aligned}\varepsilon > \frac{1}{\delta}\int_0^\delta |2 - \varphi_n(\theta) - \varphi_n(-\theta)|\,d\theta &= \frac{2}{\delta}\int_0^\delta |1 - \mathbb{E}[\cos(\theta X)]|\,d\theta \\ &= \int_{\mathbb{R}} d\mu_n(x)\frac{1}{\delta}\int_0^\delta (1-\cos(\theta x))\,d\theta \\ &\geq \int_{\{|dx|\geq 2\}} d\mu_n(x)\left(1 - \frac{\sin(\delta x)}{\delta x}\right) \\ &\geq \mu_n(|\delta x| \geq 2) = \mu_n(|X| \geq K),\end{aligned}$$

for $X \sim \mu$, using Fubini's Theorem, and the non-negativity of the integrand, for $K := 2/\delta$.

Step 2 (verify subsequence stuff): TODO

Theorem: (Central Limit Theorem) (X_n) IID RV's, with each $\mathbb{E}[X_n] = 1$ and $\sigma^2 := \mathbb{E}[X_n^2] < +\infty$. Then, for $G_n := \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^n X_i$,

$$\mathbb{P}[G_n \leq x] \to \varphi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-y^2/2}\,dy.$$

Proof: The characteristic function of the Standard Gaussian is

$$\varphi(\theta) = \int_{-\infty}^\infty e^{i\theta t}\frac{e^{-t^2/2}}{\sqrt{2\pi}}\,dt = e^{-\theta^2/2}\int_{-\infty}^\infty \frac{e^{-(\theta-t)^2/2}}{\sqrt{2\pi}}\,dt = e^{-\theta^2/2}$$

(since $t^2/2 - i\theta t - \theta^2/2 = (\theta - t)^2/2$). Hence, by the Method of Characteristic Functions, it suffices to show that, $\forall \theta \in \mathbb{R}$, $\varphi_{G_n}(\theta) \to \varphi(\theta)$. We first perform a second-order Taylor estimate:

$$e^{ix} - 1 = i\int_0^x e^{it}\,dt = ix + \int_0^x te^{it}\,dt = ix - x^2/2 + R_2(x),$$

where $|R_2(x)| \leq \int_0^x t^2/2\,dt = |x|^3/6$. (since

$$|e^{ix} - (1 + ix)| \leq \left|\int_0^x t\,dt\right| = x^2/2),$$

and hence

$$|e^{ix} - (1 + ix - x^2/2)|$$

Thus,

$$\varphi_{G_n}(\theta) = \left(o(1/n) + 1 - \frac{\theta^2}{2n}\right)^n \to e^{-\theta^2/2}. \quad \blacksquare$$