Probability Theory Oral Exam Notes 2008

Chaitanya Ekanadham

1 Distributions, characteristic functions, weak convegence

1.1 Definitions

probability measure, distribution function, Lebesgue-Stieltjes integral, random variable, Fourier transform/inversion, characteristic function, positive definite function, weak convergence, convergence in probability/a.s., induced distribution

1.2 Useful Theorems

Theorem 1. (1-1 correspondence between prob. measures and dbn fns)
$\{P : (\mathbb{R}, \mathbb{B}, P)$ is a probability space$\} \cong \{F : \mathbb{R} \to [0,1]$ nondec. and r-cnts s.t. $F(-\infty) = 0, F(\infty) = 1\}$.

Depends on: countable additivity, FIP

Proof idea: The basic relation is: $\forall x \in \mathbb{R}$, $P(-\infty, x] = F(x)$.
$\Rightarrow$: Clearly F is nondec, $F(-\infty) = 0$ and $F(\infty) = 1$. Right-continuity follows from countable additivity of P.
$\Leftarrow$: 3 steps.

1. STS P c.a. on field $\mathcal{F}$ generated by $\{(-\infty, x]\}$ and then apply Caratheodory Extension.
2. Take $A_j \in \mathcal{F} \downarrow \phi$ and suppose $P(A_j) \geq \delta \, \forall j$. Approx A_j by B_j obtained by shrinking slightly each interval in the finite union making up A_j to an open interval s.t. $sum_j P(A_j \cap B_j^c)$ is small.
3. Take the sequence $D_n = \bigcap_1^n \overline{B_j}$. Then $D_n \downarrow \phi$. But since the approximation is tight, and $\bigcap_1^n B_j \subseteq D_n$ have measure $> \delta/2$, this violates FIP for (D_n)

□

Theorem 2. (Alternative definitions of weak covergence)
Let $\{P_n\}$ and P be prob. measures with distribution functions $\{F_n\}$ and F, respectively. Then $\lim(P_n([a,b]) = P([a,b])\forall a, b$ non-atomic wrt $P \Leftrightarrow F_n(x) \to F(x)$ $\forall$ continuity points x of F.

Depends on:

Proof idea: follows easily from definitions. □

Theorem 3. (basic properties of char fns)
Let $X : \Omega \to \mathbb{R}$ be a r.v. with char fn $\phi(t)$ and inducing the prob. measure P and dbn fn F on $(\mathbb{R}, \mathbb{B})$. Then the following properties hold:

1. $\phi(0) = 1$, $|\phi(t)| \leq 1$ $\forall t \in \mathbb{R}$, $\overline{\phi(t)} = \phi(-t)$, and $\phi(t)$ is uniformly cnts on $\mathbb{R}$.

2. (Inversion thm): $\phi(t)$ uniquely specifies P (or equiv F) via
$$F(b) - F(a) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \phi(t) \frac{e^{-itb} - e^{-ita}}{-it} dt.$$

3. $\phi(t) \in \mathbb{R}$ $\forall t \in \mathbb{R} \Leftrightarrow X$ induces a symmetric distribution (i.e., $\forall a < b$, $F(b) - F(a) = F(-a) - F(-b)$, i.e. $-X$ also induces F).

4. $\mathbb{E}[|X|^n] < \infty \Rightarrow \phi^{(n)}(0) = i^n \mathbb{E}[X^n]$

5. If n even, $\phi^{(n)}(0)$ exists $\Rightarrow \mathbb{E}[X^n] = i^{-n}\phi^{(n)}(0)$

6. $\int |\phi(t)|dt < \infty \Rightarrow X$ has a bnded cnts density

Depends on:

Proof idea(s): Proof here

1. Follows from definition $\phi(t) = \mathbb{E}[e^{itX}]$. For unif. continuity use BCT.

2. Guess the "density" of F by the inverse Fourier transform: $F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\phi(t)dt$ and apply an FTOC argument (this is not a proof but gives the basic idea - the actual proof involves crazy integrals).

3. Use inversion formula(?)

4. Express $\phi'(0)$ as a limit, split $e^{ihx} = cos(hx) + isin(hx)$ and apply DCT with dominating function being $C_1|x| + C_2$.

5. Similar to (4) but use Fatou's lemma instead of DCT (???).

6. (?)

□

Theorem 4. Levy continuity theorem (weak convergence and char fns)
Let α_n, α be prob. measures with dbn fns F_n, F, and char fns $\phi_n(t)$, $\phi(t)$, respectively. Then:

1. $\alpha_n \to \alpha$ weakly. $\Leftrightarrow$

2. $\forall f : \mathbb{R} \to \mathbb{R}$ bnded and cnts, $lim_{n\to\infty} \displaystyle\int f(x) d\alpha_n = \int f(x) d\alpha \Leftrightarrow$

3. $\forall t \in \mathbb{R}$, $\lim\limits_{n\to\infty} \phi_n(t) = \phi(t)$.

Depends on: Riemann approximations to integrals, subdistribution functions, diagonalization, Fubini's theorem

Proof idea:

1. (1) $\Rightarrow$ (2):
   (a) Take f bnded cnts and restrict the integral to $[a, b]$ outside of which integral is negligible (can do this since f bounded).
   (b) Divide $[a, b]$ into bins whose endpoints are continuity points of F. Construct a Riemann-like sum to approximate $\int f(x) dF_n$ (need f cnts). Do the same for $\int f(x) dF$.
   (c) Use weak convergence to show that the Rieman approximations are close. Apply triangle inequality.

2. (2) $\Rightarrow$ (3): Choose $f(x) = e^{itx}$.

3. (3) $\Rightarrow$ (1): Very hard. 4 main steps:
   (a) Construct a "skeleton" function $G(x)$ on $\mathbb{Q}$ by extracting a convergent subsequence (using compactness) $F_{n_j}(q) \to b_q$ and diagonalizing. Thus $F_{n_j}(q) \to G(q)$ on $\mathbb{Q}$.
   (b) "Fill in the holes" by defining $G(x) = \inf\limits_{q<x} b_q$ $\forall x \notin \mathbb{Q}$. Show $G(x)$ is right-cnts, nondecreasing ("subdistribution fn").
   (c) Show $G(-\infty) = 0$ and $G(\infty) = 1$. This is hard and needs Fubini's thm and continuity of $\phi(t)$ at 0.
   (d) Since ($F_{n_j} \to G$ weakly) $\Rightarrow$ ($\phi_{n_j}(t) \to \phi(t)$) it follows that $G = F$. If we take any subsequence, we can reapply this argument, so $\nexists$ a subsequence F_{n_k} not cvging weakly to G. So $F_n \to F$ weakly.

□

Theorem 5. Bochner's theorem (sufficiency for char fns)
If $\phi(t)$is cnts at $t = 0$ with $\phi(0) = 1$ and $\forall \xi_1,...\xi_n$ in $\mathbb{C}$ and $\forall t_1,...,t_n \in \mathbb{R}$, $\sum_{i,j=1}^{n} \phi(t_i - t_j)\xi_i\overline{\xi_j} \geq 0$, then $\phi(t)$ is the char fn of some probability distribution.

Depends on:

Proof idea: From the proof of previous theorem, STS that there is a sequence of char fns $\phi_n(t) \to \phi(t)$. □

1.3 Important examples

1. $\mathbb{E}[|X|]$ need not exists even if $[\phi(t)$ can be differentiable at 0: consider the density function $g_b(x) = C(x^2(ln|x|)^b)^{-1}1_{|x|>2}$. Then one can show that for $b \in (0, 1]$, the char fn corresponding to $g_b(x)$ is diff. at 0 but $\mathbb{E}[|X|] = \infty$.

2. If $\alpha_n \to \alpha$ weakly then if f is not bounded it is not necessarily true that $lim_{n\to\infty} \int f(x)d\alpha_n = \int f(x)d\alpha$. Let $C_n = \sum_{j=1}^{n} \frac{1}{j^2}$ and consider the measure $\alpha = C_\infty \sum_{j=1}^{\infty} \frac{1}{j^2}\delta_j$ and the sequence of measures $\alpha_n = (1 - \frac{C_n}{C_\infty})\delta_{-M_n} + \frac{1}{C_\infty}\sum_{j=1}^{n} \frac{1}{j^2}\delta_j$ with M_n chosen s.t. the mean of the dbn's corresp. to α_n is always 1. However, the mean of α is ∞.

   Another example is $\alpha_n = (1 - \frac{1}{n})\delta_1 + \frac{1}{n}\delta_{n^2}$ and $\alpha = \delta 1$. Then the mean of α_n is finite but $\to \infty$ as $n \to \infty$. However, the mean of α is 1.

2 Sums of random variables and limit theorems

2.1 Definitions

independent (event and r.v.), sample mean, (infinite) product measure, cylinder sets, convolution, consistent family of dbns

2.2 Useful Theorems

Theorem 1. (Generalized) Chebyshev's inequality
Let X be a random variable on $(\Omega, \mathcal{F}, P)$. Then $\forall p \geq 1$,
$P(\{\omega \in \Omega : |X(\omega) - \mathbb{E}[X(\omega)]| \geq \delta\}) \leq \frac{1}{\delta^p}\mathbb{E}[|X - \mathbb{E}[X]|^p]$

Depends on:

Proof idea: (easy - just express LHS as an integral and estimate) □

Remark: $p = 2$ gives $\frac{Var[X]}{\delta^2}$ as the upper bound. This directly implies that the sample mean of i.i.d random variables with finite mean and variance converges in probability to the true mean.

Theorem 2. Weak law of large numbers
Let (X_i) be i.i.d with finite mean m. Then the sample mean $\frac{S_n}{n} \to m$ in probability. ***Depends on:*** dependencies...

Proof idea: Write $\phi_{S_n/n}(t) = [\phi_{X_1}(\frac{t}{n})]^n$. Taylor-expand $\phi_{X_1}(\frac{t}{n})$ about 0 to the 2nd order and use the fact that $na_n \to z \Rightarrow (1 + a_n)^n \to e^z$ to show that $\phi_{S_n/n}(t)$ converges the char fn of the point mass at m. Weak convergence to a constant implies convergence in probability □

Theorem 3. Kolmogorov consistency thm: (existence of infinite product measures)
Given a consistent family of measures $\{P_n\}$ on $\mathbb{R}^n$ $\forall n \geq 1$, $\exists$ a unique measure P on $\mathbb{R}^\infty$ that agrees with P_n on n-dim cylinder sets $A_1 \times A_2 \times ... \times A_n \times \mathbb{R}^\infty$.

Depends on: Caratheodory extension theorem, approximation of Borel sets by compact subsets (in general, need a complete separable metric space with Borel sigma field) - uses the Monotone class theorem

Proof idea:

1. Define $P(A) = P_n(A)$ for any n-dimensional cylinder set A.
2. Show that P is c.a. on the field $\mathcal{F}$ generated by cylinder sets - take a $B_n \downarrow \phi$ in $\mathcal{F}$:
   (a) Show that each B_n, being the finite union of cylinder sets, is itself a cylinder set with measurable base $A_n \in \mathbb{R}^n$.
   (b) Approximate each A_n by a compact subset K_n (in general, need a separable metric space to do this) - one can show this by showing that the class of sets with this property is a monotone class containing the field of measurable rectangles and applying the Monotone class theorem. Define cylinder sets C_n with bases K_n.

(c) Let $D_n = \bigcap_{j=1}^{n} C_n$ and show that $P(D_n) \geq \delta/2$ if we assume $P(B_n) \geq \delta$. But $D_n \downarrow \phi$ if $B_n \downarrow \phi$. But by compactness of the K_n's we can construct a sequence via diagonalization that is in $\bigcap_n D_n \Rightarrow$ contradiction.

3. Apply Caratheodory extension to extend P to $\sigma(\mathcal{F})$.

□

Theorem 4. Borel Cantelli lemma
Suppose (A_n) are measurable sets in $(\Omega, \mathcal{F}, P)$. Then:

1. $\sum_n P(A_n) < \infty \Rightarrow P(\{\omega \in \Omega : \omega \notin A_n \text{ eventually}\} = 1$.

2. A_n are independent and $\sum_n P(A_n) = \infty \Rightarrow P(\{\omega \in \Omega : \omega \in A_n$ for infinitely many $n\} = 1$.

Depends on: countable addivity applied to lim sup of sets, Komogorov consistency (implicitly), $(1-x) \leq e^{-x} \forall x \in [0,1]$.

Proof idea:

1. Notice that $\{\omega \in \Omega : \omega \in A_n$ for infinitely many $n\} = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n = \limsup_{n\to\infty} A_n$. Since this is an intersection of decreasing sequence of sets, use countable additivity and the fact that $\sum_k^{\infty} \to 0$ as $k \to \infty$ to show that $P(\limsup A_n) = 0$.

2. Analogously, look use countable additivity of P on $\{\omega \in \Omega : \omega \in A_n$ for finitely many $n\} = \liminf A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$ which is $\lim_{n\to\infty} \prod_{m=n} (1 - P(A_m))$ by independence. Use the fact that $1 - x \leq e^{-x}$ to show that the limit is 0.

□

Theorem 5. Kolmogorov's inequality (bound on prob. that random walk exceeds a cutoff) Suppose (X_n) are indep. with mean 0 finite variances σ_n^2. Then $P(\{T_n \equiv sup_{1\leq k \leq n}\{S_k\} \geq l\} \leq \frac{1}{l^2} \sum_{k=1}^{n} \sigma_k^2$

Depends on: chebyshev's inequality, partitioning events

Proof idea: Partition the event $\{T_n \geq l\}$ into events E_k, $k = 1, 2, ...n$ where E_k is the event that the random walk hits l for the first time at time k.
Bound $P(E_k)$ using chebyshev's inequality and use the fact that S_k is independent to the mean-0 quanitity $(S_n - S_k)$ to complete the square in the integrand: $S_k^2 \to S_k^2 + 2S_k(S_n - S_k) + S_n^2$. Conclude that $P(E_k) \leq \frac{1}{l^2} \int_{E_k} S_n^2 dP$ and sum the estimates over k to get the result. □

Theorem 6. Levy's inequality (bound on the same quantity given info about the partial tails of the sum)
If $P\{|X_i + ... + X_n| \geq l/2\} \leq \delta \forall i \in [1, n]$ then $P\{T_n \geq l\} \leq \frac{\delta}{1-\delta}$

Depends on:

Proof idea:

1. Split $\{T_n \geq l\}$ into $A = \{T_n \geq l\} \cap \{|S_n| \leq l/2\}$ and $B = \{T_n \geq l\} \cap \{|S_n| > l/2\}$.
2. Bound $P(A)$ by partitioning into E_k's as before and using the fact that in event E_k, $|S_n| \leq l/2$ is implied by $|S_n - S_k \geq l/2$ since $|S_k| \geq l$. Details:
$P(A) = \sum P(E_k \cap \{|S_n| \leq l/2\}) \leq \sum P(E_k \cap \{|S_n - S_k| \geq l/2\}) = \sum P(E_k)P(\{|S_n - S_k| \geq l/2\}) \leq \delta P(\{T_n \geq l\})$
3. $P(B) \leq P(\{|S_n| \geq l/2\}) \leq \delta$
4. Add $P(A) + P(B)$

□

Theorem 7. (cauchy-cvgnce in prob. with an extra sup condition gives a.s. cvgnce)
Let X_n be random variables on $(\Omega, \mathcal{F}, P)$. Then $\lim_{n,m\to\infty} P(\{\sup_{m \leq k \leq n} |S_k - S_m| \geq \delta\} = 0 \ \forall \delta \Rightarrow \exists S$ s.t. $S_n \to S$ a.s.

Depends on:

Proof idea: Clearly $S_n \to S$ are cauchy prob. $\Rightarrow \exists S$ s.t. $S_n \to S$ in prob and $\exists S_{n_j} \to S$ a.s.
Let $A_j = \{\omega \in \Omega : |S_n(\omega) - S_{n_j}(\omega)| \geq 2^{-j}$ for some $n \geq n_j\}$. Then it follows that $\omega \notin \limsup A_j \Rightarrow S_n(\omega) \to S(\omega)$.But one can show using the sup condition that for j suff. large, $P(A_j) < 2^{-k} \Rightarrow P(\limsup A_j) = 0$. □

Theorem 8. Levy's theorem (equiv. of modes of convergence for sums of rvs)
Suppose X_n indep. on $(\Omega, \mathcal{F}, P)$. Then:

1. the dbn of S_n cvg weakly to some dbn $\Leftrightarrow$
2. $S_n \to$ some S in prob. $\Leftrightarrow$
3. $S_n \to$ some$S$ a.s.

Depends on: Kolmogorov ineq., Levy's ineq., Sup condition

Proof idea: (3) $\Rightarrow$ (2) $\Rightarrow$ (1) is trivial.

1. (1) $\Rightarrow$ (2): Using char fns, the assumption implies that the infinite product $\prod \phi_j(t)$ converges to some limit that is nonzero around 0, which means the tail of the product converges to 1 around 0. One can show that this means the modulus of the tail converges to 1 $\forall t \in \mathbb{R}$ (use $1 - cos(2t) \leq 4(1 - cos(2t)))$, which in turn implies that the random variable $S_n - S_m$ converges to the point mass at 0 (i.e. cvgs in prob. to 0). Thus S_n is cauchy in prob. and thus converges in prob.

2. (2) $\Rightarrow$ (3): STS cauchy-cvg in prob. with extra sup condition. This follows from (2) (gives cauchy-cvg in prob) and Levy's inequality (which gives the sup condition).

□

Theorem 9. Kolmogorov 1-series theorem
Suppose X_n indep. on $(\Omega, \mathcal{F}, P)$ with 0 mean and $\sum Var[X_j] < \infty$. Then $S_n = \sum X_j$ converges a.s.
Depends on: Kolmogorov and sup condition ineq. OR Chebyshev's ineq. and Levy's theorem

Proof idea: Apply chebyshev ($p = 2$) to the set $A = \{|S_n - S_m| \geq \delta\}$ and show that $\lim_{m,n\to\infty} P(A) = 0$ using the variance condition. Therefore, (S_n) is cauchy in prob., and therefore converges a.s. by Levy's theorem. □

Theorem 10. Kolmogorov 2-series theorem
Suppose X_n indep. on $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}[X_j] = a_j$ and $Var[X_j] = \sigma_j$. Suppose also that $\sum a_j$ and $\sum Var[X_j]$ converge. Then $S_n = \sum X_j$ converges a.s.
Depends on: 1-series theorem

Proof idea: easy - apply 1-series to $Y_j = X_j - a_j$. □

Theorem 11. Kolmogorov 3-series theorem
Suppose X_n indep. on $(\Omega, \mathcal{F}, P)$. Then $S_n = \sum X_j$ converges a.s. $\Leftrightarrow \exists C > 0$ s.t.:

1. $\sum P(\{|X_j| > C\}) < \infty$

2. $\sum \mathbb{E}[Y_j] < \infty$ with $Y_j = X_j 1_{|X_j| \leq C}$.

3. $\sum Var[Y_j] < \infty$

Depends on: 2-series theorem, Borel-cantelli lemma, convergence of the sum of variances of uniformly bounded RV's with mean 0

Proof idea: $\Leftarrow$: By Borel-cantelli, (1) $\Rightarrow X_j = Y_j$ eventually with prob. 1. Applying 2-series to Y_j gives a.s. convergence which carries over to a.s. convergence of sums of X_j.
$\Rightarrow$: Suppose $\sum X_j$ converges a.s. Then:

1. It must be that $|X_n| \leq C$ a.s. for sufficiently large $n \Rightarrow \sum P(|X_n| > C)$ converges (by Borel Cantelli part 2)

2. WLOG we can assume $|X_j| \leq 1$ a.s. $\forall j \geq 1$. If $\mathbb{E}[X_j] = 0 \forall j \geq 0$ then $\sum \mathbb{E}[X_j]$ cvgs trivially. Otherwise if we prove the 3rd series cvgs, then since $\sum X_j$ and $\sum X - j - \mathbb{E}[X_j]$ both converge, the 2nd series must converge.

3. STS that $\sum Var[Z_j]$ converges a.s. where $Z_j = X_j - X_j'$ and X_j' is identically distributed as X_j. Note that $\sum Z_j$ converges a.s. and $|Z_j| \leq 2$ a.s. and Z_j has mean 0. Let $F_n = \{\omega : |S_j| \leq l \forall 1 \leq j \leq n\}$. Then for l large enough, since $\sum Z_j$ converges, we have $P(F_n) \geq \delta$. We observe that:

   (a) $\int_{F_{n-1}} S_n^2 dP = \int_{F_{n-1}} (S_{n-1} + X_n)^2 dP \geq \int_{F_{n-1}} S_{n-1}^2 dP = \delta \sigma_n^2$

   (b) $\int_{F_{n-1}} S_n^2 dP = \int_{F_n} S_n^2 dP + \int_{F_{n-1} \cap F_n^c} S_n^2 \leq \int_{F_n} S_n^2 + P(F_{n-1} \cap F_n^c)(l + 2)^2$

   (c) Therefore we get an estimate for σ_n^2 and summing over n gives the upper bound $\delta^{-1}(l^2 + (l+2)^2)$

□

Theorem 12. Strong law of large numbers)
Suppose X_j are i.i.d random variables with mean 0 and finite first moment. Then $\frac{S_n}{n} \to 0$ a.s.

Depends on: Borel-Cantelli, Kolmogorov 3-series

Proof idea: 3 steps:

1. Define $Y_n = X_n 1_{|X_n| \leq n}$. Show that since $\mathbb{E}[|X_1|] < \infty$, it follows that $\sum_n P(X_n \neq Y_n) = \sum_n P(|X_n| > n) = \mathbb{E}[|X_1|] < \infty$ (show this by expressing the expectation as a sum of integrals, estimating the integrals, and rearranging the sum). By Borel Cantelli, this means that $X_n = Y_n$ eventually, a.s.

2. Apply the 3-series theorem to show that $\sum_n \frac{Y_n - \mathbb{E}[Y_n]}{n}$ converges a.s. (each summand is bounded by 1, has mean 0 and variance can be bounded by $C\mathbb{E}[|X|]$).

3. Conclude by Borel-Cantelli that $\sum_n \frac{X_n - \mathbb{E}[Y_n]}{n}$ converges a.s. $\Rightarrow \frac{\sum_{j=1}^n X_j}{n} - \frac{\sum_{j=1}^n \mathbb{E}[Y_j]}{n} \to 0$ a.s. (the last part uses 'Kronecker' lemma which can be proved using a simple $\epsilon/2$ argument).

□

Theorem 13. Kolmogorov 0/1 Law
Let P be any infinite product measure on the space $(\Omega, \mathcal{B})$ where Ω is the space of all infinite sequences $(x_j)_{j=1}^{\infty}$ and $\mathcal{B}$ is the infinite product sigma field. Let $\mathcal{B}_n = \sigma(X_1, ..., X_n)$ and $\mathcal{B}^n = \sigma(X_n, X_{n+1}, ...)$ and let $\mathcal{B}^\infty = \cap_{j=n}^{\infty} \mathcal{B}^j$ be the tail field. Then $\forall A \in \mathcal{B}^\infty$, $P(A) = 0$ or 1.

Depends on: Kolmogorov consistency, independence of events, Monotone class theorem

Proof idea: 3 steps:

1. Show that A is independent of the sigma-field $\mathcal{B}_n$, $\forall n \Rightarrow A$ is independent of sets in the field $\mathcal{F} = \cap \mathcal{B}_n$ (need P to be a product measure).

2. Show that the class $\mathcal{A}$ of sets independent of A is a monotone class: $P(\cup(A_j) \cap A) = P(\cup(A_j \cap A)) = \sum P(A_j \cap A) = P(A) \sum P(A_j) = P(A)P(\cup(A_j))$. Therefore, $\mathcal{A} \supseteq \sigma(\mathcal{F}) = \mathcal{B} \supseteq \mathcal{B}^\infty$.

3. Conclude that A is independent of itself

□

Theorem 14. Central limit theorem
Suppose (X_j) are i.i.d random variables with mean 0 and finite variance σ^2. Then $\frac{S_n}{\sqrt{n}} \to N(0, \sigma^2)$ in distribution (weakly).

Depends on: characteristic functions, limit property of e

Proof idea: Exercise in characteristic functions. The char fn of $\frac{S_n}{\sqrt{n}}$ is $\phi(t) = [\phi_{X_1}(t/\sqrt{n})]^n = \mathbb{E}[e^{itX_1/\sqrt{n}}]^n = [1 + it0 - \frac{t^2\sigma^2}{2n} + o(\frac{1}{n})]^n \to e^{-t^2\sigma^2/2}$. □

Theorem 15. Lindeberg condition
Suppose (X_j) are independent with distributions (α_j). Suppose that $\forall \epsilon > 0$, $\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x|\geq \epsilon s_n} x^2 d\alpha_j$. Then the CLT holds.

Depends on: char fns, triangular arrays

Proof idea: 3 steps (long - just try to know outline):

1. Reformulate the problem using triangular arrays as showing that $\prod_{j=1}^{n} phi_{n,j}(t) \to e^{-t^2/2}$ where $\phi_{n,j}(t)$ is the char fn of X_j/s_n.

2. The trick is to replace $\phi_{n,j}(t)$ by $\psi_{n,j}(t) = exp(\phi_{n,j}(t) - 1)$ and show that $|\prod_{j=1}^{n} \phi_{n,j}(t) - \prod_{j=1}^{n} \psi_{n,j}(t)|$ is small for large n and for $|t| < T$ for any finite T.

   Details: To do this, we take the absolute difference of the logs and use a taylor expansion of log. The absolute difference can be bounded by $\sup_{|t|\leq T} |\sum_{j=1}^{n} |\log \phi_{n,j}(t) - (\phi_{n,j}(t) - 1)| \leq \sup_{t\leq T} C \sum_{j=1}^{n} |\phi_{n,j}(t) - 1|^2 \leq C\{\sup_{|t|\leq T} \sup_{1\leq j\leq n} |\phi_{n,j}(t) - 1|\}\{\sup_{|t|\leq T} \sum_{j=1}^{n} |\phi_{n,j}(t) - 1|\}$. Show that:

   (a) The first factor $\to 0$ as $n \to \infty$:
   take $|\sup_{|t|\leq T} |\phi_{n,j}(t) - 1|$ and express it as an integral wrt α_j. Use taylor expansion of exponential to bound this integral by $C_T \int \frac{x^2}{s_n^2} d\alpha_j$ (need mean 0 condition) Then split the integral into sets $\{|x| \geq \epsilon s_n\}$ and its complement. Bound the integral over the complement by a constant and show that the second one goes to 0 uniformly

in j by the Lindeberg condition. Thus the entire thing goes to 0 uniformly in j and t.

(b) The second factor is uniformly bounded in n:
Use the same integral estimate $C_T \int \frac{x^2}{s_n^2} d\alpha_j$ for each summand and bring C_T/s_n^2 outside the sum.

3. Show the result for the ψ's by taking logs, i.e. that $\sup_{|t|\leq T} |[\sum_{j=1}^{n}(\phi_{n,j}(t) - 1)] + \frac{t^2}{2}| \leq C_T\epsilon$. Do this by bringing $t^2/2$ into the sum as $\sigma_j^2 t^2/(2s_n^2)$, writing each summand as an integral (using the mean 0 condition), and splitting up the integral over the same ϵ set and its complement. Taylor expand the exponential to 3rd order in the complement and 2nd order in the regular set. Use Lindeberg condition for the regular set.

$\square$

2.3 Important examples

1. One 'tail event' would be $S_A = \{\omega \in \Omega : \lim_{n\to\infty} S_n(\omega)/n \in A\}$ for some borel set A. The SLLN shows that the limit converges a.s. to a constant. which agress with the conclusion of the 0/1 law that $P(S_A) = 0$ or 1. Something not in the tail field would be the event $\{\omega \in \Omega : \sup |X_j(\omega)| = 1\}$.

2. Suppose Y_j are independent and $Y_j = \pm 1$ each w.p. 1/2 and let $X_j = \sigma_j Y_j$ with $\sigma_j = 1/j$. Let $S_n = \sum_{j=1}^{n} X_j$ and let $s_n^2 = \sum_{j=1}^{n} \sigma_j^2$. Then since $s_\infty^2 = \lim_{n\to\infty} s_n^2 < \infty$, by the 1-series theorem, S_n converges a.s. Therefore, $S_n/s_n \to \sum_{j=1}^{\infty} \sigma_j Y_j / s_\infty$ a.s., which is *not* normally distributed (can show this with char fns), so the CLT cannot apply. In general if the sum of the variances is finite, the CLT may not apply.

3. Even if the sum of the variances is infinite, the CLT can stil fail. Consider $X_j = \pm j$ each w.p. $p_j/2$ and 0 w.p. $1 - p_j$ with $p_j = 1/j^2$. Then since $\sum p_j < \infty$, X_j is eventually 0 all the time, i.e. $\sum X_j$ converges a.s. to a finite value. But $s_\infty^2 = \sum Var[X_j] = \infty \Rightarrow S_n/s_n \to 0$ a.s. Intuitively, even though $s_n \to \infty$, the contribution to s_n for large n comes

from too large values of X_j with small probabilities. This motivates Lindeberg's condition.

4. One triangular array example shows that the limit of the sum of n indep. Bernoulli(p_n) trials converges in dbn to Poisson(λ) whenever $\lambda = \lim np_n \in (0, \infty)$. To see this , just write $S_n = X_{1,n} + ... + X_{n,n}$ where $\{X_{j,n}\}_{j=1}^n$ are indep. Bernoulli(p_n) r.v.s's. Write out the characteristic function S_n.

3 Dependent random variables

3.1 Definitions:

conditional expectation, version, conditional probability, regular conditional expectation

3.2 Useful theorems

Theorem 1. (Existence and a.e. uniqueness of conditional expectation)
Let X be a r.v. on $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}[|X|] < \infty$ and suppose $\Sigma \subseteq \mathcal{F}$ is a σ-field. Then $\mathbb{E}[X|\Sigma]$ exists and is unique a.s.

Depends on: Radon-Nikodym theorem

Proof idea: Write $X(\omega) = X^+(\omega) - X^-(\omega)$ (difference of nonneg fns). For $* \in \{+, -\}$, define $\lambda^*(A) = \displaystyle\int_A X^*(\omega) dP$ on $\mathcal{F}$ and let $\tilde{\lambda}^*$ be the restriction to Σ. Clearly $\tilde{\lambda} << \tilde{P} = P|_\Sigma \Rightarrow \exists \frac{d\tilde{\lambda}^*}{d\tilde{P}}$. Let $Y(\omega) = \frac{d\tilde{\lambda^+}}{d\tilde{P}} - \frac{d\tilde{\lambda^-}}{d\tilde{P}}$. Verify that Y satisfies all the defining properties of conditional expectation.

□

Theorem 2. (Properties of conditional expectation)
Let X be a r.v. on $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}[|X|] < \infty$. Then the following properties hold.

1. $\mathbb{E}[X|\Sigma_2] = \mathbb{E}[\mathbb{E}[X|\Sigma_1]|\Sigma_2]$ if $\Sigma_2 \subseteq \Sigma_1$.

2. $X \geq 0$ a.s. $\Rightarrow \mathbb{E}[X|\Sigma] \geq 0$ a.s.

3. $\mathbb{E}[aX_1 + bX_2|\Sigma] = a\mathbb{E}[X_1|\Sigma] + b\mathbb{E}[X_2|\Sigma]$

4. $\mathbb{E}[XY|\Sigma] = Y\mathbb{E}[X|\Sigma]$ a.s. for any Y Σ-measurable,bounded.

5. ϕ convex on $\mathbb{R} \Rightarrow \mathbb{E}[\phi(X)|\Sigma] \geq \phi(\mathbb{E}[X|\Sigma])$ a.s.

Depends on: definitions, manipulations

Proof idea: All of these are easy except for conditional Jensen's. One nice way to see it is that every convex function $\phi(x)$ can be written $\phi(x) = sup_i\{a_i x + b_i\}$ (countable i). Then by monotonicity,taking sup, and linearity, we have $\mathbb{E}[\phi(X)|\Sigma] \geq \mathbb{E}[a_i X + b_i|\Sigma]\ \forall i \Rightarrow \mathbb{E}[\phi(X)|\Sigma] \geq \sup_i \mathbb{E}[a_i X + b_i|\Sigma] = \phi(\mathbb{E}[X|\Sigma])$ □

4 Markov chains

4.1 Definitions:

irreducible, positive recurrence, null recurrence, transcience, periodicity, invariant distribution

4.2 Useful theorems

Theorem 1. (Irreducible markov chains) Let χ be a countable state space for an irreducible markov chain $(X_j)_{j\geq 1}$. Then all states $x \in \chi$ are positive recurrent, null recurrent, or transient.

Depends on: Manipulations, and intuitive probability arguments

Proof idea: Proof here □

Theorem 2. (Necessary/sufficient conditions for transience) Let $(X_j)_{j\geq 1}$ be an irreducible MC on countable state space χ. Then:

1. (X_j) is transient $\Leftrightarrow \forall x, y, \in \chi$, $G(x,y) = \displaystyle\sum_{n=0}^{\infty} \pi^{(n)}(x,y) < \infty$.
2. $\forall x, y \in \chi$, $G(x,y) = f(x,y)G(y,y)$ and $G(x,x) = \frac{1}{1-f(x,x)}$ where $f(x,y) = P_x(\tau_y < \infty)$ and τ_x is the first hitting time of state x.

Depends on: dependencies...

Proof idea: Proof here □

Theorem 3. (Recurrence and invariant distributions) Let (X_j) be a recurrenet irreducible MC on countable χ. Then:

1. (X_j) is null recurrent $Rightarrow \lim_{n\to\infty} \pi^{(n)}(x,y) = 0 \ \forall x, y$.

2. (X_j) is positive recurrent $Rightarrow \lim_{n\to\infty} \pi^{(n)}(x,y) = \frac{1}{\mathbb{E}_{P_y}[\tau_y]} = Q(y)$. Q is an invariant distribution for (X_j).

Depends on: dependencies...

Proof idea: Proof here □

4.3 Important examples:

1. Recurrence for SSRW on $\mathbb{Z}^d$:

2.

3.

5 Martingales

5.1 Definitions:

martingale, super/sub-martingale

5.2 Useful theorems

Theorem 1. Doob's inequality ('Chebychev-type inequality for the sup of martingale sequence'): Suppose $\{X_j\}_{j=1}^n$ is a martingale on $(\Omega, \mathcal{F}, P)$. Then $P(\{\omega \in \Omega : \sup_{1\le j\le n} |X_j(\omega)| \ge l\}) \le \frac{1}{l}\int_{\{\sup_{1\le j\le n}|X_j|\ge l\}} |X_n| dP \le \frac{\mathbb{E}[|X_n|]}{l}$.

Depends on: Chebyshev, a convex function of a martingale is a submartingale

Proof idea: Let $S(\omega) = \sup_{1\le j\le n} |X_j(\omega)$ and partition the event that $S \ge l$ into events $\{E_j\}_{j=1}^n$ where the martingale sequence exceeds l for the first time at time j. Apply Chebychev to bound $P(E_j)$ and then use the fact that $|X_j|$ is convex together with the martingale property. □

Theorem 2. (L_1 convergence to martingale limits)
Suppose $\{X_n\}$ is a martingale on $(\Omega, \mathcal{F}, P)$ and $\exists X \in L_p(\Omega, \mathcal{F}, P)$ s.t. $\forall n$, $X_n = \mathbb{E}[X|\mathcal{F}_n]$. Then $\|X_n - X\|_p \to 0$.

Depends on: approximation of L_p fns by bounded fns, Minkowski (triangle) inequality, completeness in L_2

Proof idea: Show it first for X bounded. In this case, $X_n(\omega)$ is uniformly bounded in n and ω. In particular, X_n is uniformly bounded in the L_2 norm. By viewing X_n as the sum of martingale differences, show that (X_n) is a L_2-cauchy sequence and so has an L_2 limit. Since L_2 convergence is equivalent to L_p convergence for all p for a uniformly bounded sequence of functions, the result holds.

For general $X \in L_p$, approximate X by a bounded fn $X' \in L_\infty$ and consider the martingale $X'_n = \mathbb{E}[X'|\mathcal{F}_n]$. It can be using cond. expectation properties that $\|X'_n - X_n\|_p$ is small if $\|X' - X\|_p$ is.

□

Theorem 3. (Sufficient conditions for an L_p martingale limit)
Suppose $\{X_n\}$ is a martingale on $(\Omega, \mathcal{F}, P)$ and for some $p > 1$, $\sup_n \|X_n\|_p < \infty$. Then $\exists X \in L_p(\Omega, \mathcal{F}, P)$ s.t. $\forall n$, $X_n = \mathbb{E}[X|\mathcal{F}_n]$. Furthermore, $\|X_n - X\|_p \to 0$

Depends on: L_p duality and weak compactness of bounded sets, previous theorem

Proof idea: Let M be the bound. Since $p > 1$ the dual is L_q and since the closed ball of radius M is weakly compact, we can extract a weakly compact subsequence X_{n_j} with weak limit $X \in L_p$. One can use this and the martingale property to show that $\forall A \in \mathcal{F}_n$, $\mathbb{E}[X 1_A] = \mathbb{E}[X_n 1_A]$, from which it follows that $X_n = \mathbb{E}[X|\mathcal{F}_n]$. Apply preceding theorem to get L_p convergence.

□

Theorem 4. (Sufficient conditions for an L_1 martingale limit)
Suppose $\{X_n\}$ is a martingale on $(\Omega, \mathcal{F}, P)$ and that X_n are uniformly integrable. Then $\exists X \in L_1(\Omega, \mathcal{F}, P)$ s.t. $\forall n$, $X_n = \mathbb{E}[X|\mathcal{F}_n]$.

Depends on: Density of $\mathcal{F}_n$-measurable functions in L_1, Doob's inequality, Jensen's inequality

Proof idea: Proof here □

Theorem 5. (Doob decomposition theorem)
Suppose $(X_n, \mathcal{F}_n)$ is a submartingale on $(\Omega, \mathcal{F}, P)$. Then $\forall n \geq 1$, X_n can be expressed uniquely as $Y_n + A_n$ such that:

1. $(Y_n, \mathcal{F}_n)$ is a martingale
2. $A_{n+1} \geq A_n$ a.s. $\forall n \geq 1$ and $A_1 \equiv 0$.
3. $\forall n \geq 2$, A_n is $\mathcal{F}_{n-1}$-measurable.

Depends on: dependencies...

Proof idea: □

Theorem 6. (Optional stopping-time theorem)
If $(X_n, \mathcal{F}_n)$ is a martingale and $0 \leq \tau_1 \leq \tau_2 \leq C$ are bounded stopping times. Then $\mathbb{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] = X_{\tau_1}$ a.s. ***Depends on:*** Doob's decomposition theorem

Proof idea: Proof here □

Remark: A common application is when $\tau_1 \equiv 0$. Then for τ bounded we get $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ - 'Gambler's ruin'

Theorem 7. (Upcrossing inequality - controlling oscillation of a martingale sequence)
Let $(X_n, \mathcal{F}_n)$ be a martingale. Let $U_n(a, b)$ be the number of upcrossings from $a \to b$ by the sequence $X_1, ..., X_n$. Then $\forall n$, $\mathbb{E}[U_n(a, b)] \leq \frac{\mathbb{E}[a-X_n]^+}{b-a} \leq \frac{|a|+\mathbb{E}[|X_n|]}{b-a}$.

Depends on: dependencies...

Proof idea: proof □

Theorem 8. (Martingale CLT) Let $(\xi_j)_{j\geq 1}$ be a stationary ergodic sequence of L_2 martingale differences. Then $\frac{\xi_1+...+\xi_n}{\sqrt{n}} \to \mathcal{N}(0, Var[\xi_1])$ in distribution.

Depends on: dependencies...

Proof idea: Proof here □

5.3 Important examples

1.

6 Ergodic theorems

6.1 Definitions:

invariant measure, measure-preserving maps, stationary process, shift operator, ergodic measure, extremal measure

6.2 Useful theorems

Theorem 1. (Ergodic theorem - 'generalization of SLLN')
Let $(\Omega, \mathcal{F}, P)$ be a probability space where $\omega \in \Omega$ corresponds to a sequence $(x_j = \omega(j))_{j \geq 1}$.Suppose T is a measure-preserving map on Ω (1-1, invertible, and $P(T^{-1}(A)) = P(A)$, $\forall A \in \mathcal{F}$. Then $\forall f \in L_1(\Omega, \mathcal{F}, P)$, $\lim_{n\to\infty} \frac{f(\omega) + f(T\omega) + ... + f(T^{n-1}\omega)}{n} = \mathbb{E}_P[f|\mathcal{L}]$ a.s. P and also in $L_1(P)$, where $\mathcal{L} = \{A : TA = A\}$ is the *invariant* σ-field. If P is ergodic for T (i.e. $P(A) \in \{0, 1\}$ $\forall A \in \mathcal{L}$) then the limit is equivalent to $\mathbb{E}_P[f(\omega)]$ a.s. and in L_p whenever $f \in L_p$.

Depends on: dependencies...

Proof idea: Proof here □

Theorem 2. (Extremal measures are exactly the T-invariant ergodic measures)
Let $(\Omega, \mathcal{F})$ be a measurable space and let $T : \Omega \to \Omega$ be 1-1,invertible. Consider the family of T-invariant measures $\mathcal{M}$ on $(\Omega, \mathcal{F}$. Then:

1. $\mathcal{M}$ is a convex set. Thus $\exists$ a subset $\mathcal{M}_e$ of extremal measures that are not convex combinations of other measures in $\mathcal{M}$.

2. A T-invariant measure $P \in \mathcal{M}$ is in $\mathcal{M}_e \Leftrightarrow P$ is ergodic for T (i.e. $P(A) \in \{0, 1\}$ $\forall A \in \mathcal{F}$ s.t. $T(A) = A$).

3. $P_1, P_2 \in \mathcal{M}_e$ distinct $Rightarrow$ P_1, P_2 are orthogonal i.e. $existsE \in \mathcal{F}$ with $T(E) = E$ s.t. $P_1(E) = 1$ and $P_2(E) = 0$.

4. Let $\pi(x, dy)$ be a fixed transition measure on $(\chi, \mathcal{F})$. Consider a measure μ on $(\chi, \mathcal{B})$ which (along with the transition measure π) induces a measure P_μ on the infinite product space $(\chi^\infty, \mathcal{B}^\infty)$ invariant w.r.t. the shift operator T on χ^∞. Then P_μ is extremal (ergodic wrt T) $Leftrightarrow \mu$ is extremal for $\tilde{\mathcal{M}}$, the set of all stationary measures on $(\chi, \mathcal{B}$ for π.

5. If $\mu_1, \mu_2 \in \tilde{mathcalM}$ are distinct extremal measures, then they are orthogonal.

6. $\nu \in \tilde{\mathcal{M}} \Rightarrow existsA \in \mathcal{F}$ of ν-measure 1 s.t. $\forall x \in A$, the MC starting at x satisfies: $\lim_{n\to\infty} \frac{f(X_1) + ... + f(X_n)}{n} = \mathbb{E}_\nu[f(X)]$ a.s.ν.

7. If $(\Omega, \mathcal{F})$ is a complete separable metric space with $\mathcal{F}$ being the Borel field, then every $P \in \mathcal{M}$ arises from a unique convex combination of extremal (ergodic) measures.

Depends on: dependencies...

Proof idea: Proof here $\square$