PROBLEMS FOR PROBABILITY: LIMIT THEOREMS I

PROF. YURI BAKHTIN

Instructions. You are allowed to work on solutions in groups, but you are required to write up solutions on your own. Please give complete solutions, all claims need to be justified. Late homework will not be accepted.

Please let me know if you find any misprints or mistakes.

1. Due by September 24, 3:30pm (Initially a different date was posted here by mistake)

1. A family $\mathcal{F}$ of subsets of a set Ω is called a σ-algebra if
   (a) $\emptyset \in \mathcal{F}$,
   (b) $A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$,
   (c) $A_n \in \mathcal{F}$ for all $n \in \mathbb{N} \Longrightarrow \bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}$.
   Prove that if one replaces (c) with
   (c')$A_n \in \mathcal{F}$ for all $n \in \mathbb{N} \Longrightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$,
   one will obtain an equivalent definition.
2. Prove that for any two sets $A, B \in \mathcal{F}$,
$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B).$$
   More generally, find and prove a similar expression (inclusion-exclusion formula) for $\mathsf{P}(\cup_{i=1}^n A_i)$.
3. Let I be a family of σ-algebras on a set Ω. Prove that $\bigcap_{\mathcal{G}\in I} \mathcal{G}$ is also a σ-algebra. Show that $\bigcup_{\mathcal{G}\in I} \mathcal{G}$ is not necessarily a σ-algebra.
4. Let P be a finitely additive function on an algebra $\mathcal{A}$ of subsets of Ω, with values in $[0, +\infty)$. Prove that the following statements are equivalent:
   (a) If $A_1, A_2, \ldots \in \mathcal{A}$ are disjoint and $\bigcup A_n \in \mathcal{A}$, then
$$\mathsf{P}(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mathsf{P}(A_n).$$
   (b) If $A_1, A_2, \ldots \in \mathcal{A}$, $A_n \subset A_{n+1}$ for all n, and $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$, then
$$\mathsf{P}(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mathsf{P}(A_n).$$
   (c) If $A_1, A_2, \ldots \in \mathcal{A}$, $A_n \supset A_{n+1}$ for all n, and $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$, then
$$\mathsf{P}(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mathsf{P}(A_n).$$
   (d) If $A_1, A_2, \ldots \in \mathcal{A}$, $A_n \supset A_{n+1}$ for all n, and $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$, then
$$\lim_{n\to\infty} \mathsf{P}(A_n) = 0.$$

5. Let $(X_n)_{n\in\mathbb{N}}$ be a bounded sequence of r.v.'s (i.e., there is a constant $C > 0$ such that $|X_n(\omega)| \le C$ for all n and ω). Prove that $\liminf_{n\to\infty} X_n$ is a r.v.
6. Give an example of a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, a sequence of r.v.'s $(X_n)_{n\in\mathbb{N}}$ and a r.v. X, such that $\{X_n \not\to X\} \ne \emptyset$ but $X_n \to X$ a.s.
7. Give an example of a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, a sequence of r.v.'s $(X_n)_{n\in\mathbb{N}}$ such that $X_n \stackrel{\mathsf{P}}{\to} 0$ but $\mathsf{P}\{\omega : X_n(\omega) \not\to 0\} = 1$.
8. Prove that if $X_n \stackrel{\mathsf{P}}{\to} X$, then there is a deterministic sequence $(n_k)_{k\in\mathbb{N}}$, $\lim_{k\to\infty} n_k = \infty$ such that $X_{n_k} \stackrel{a.s.}{\to} X$.
9. Prove the first part of the Borel–Cantelli Lemma: Denoting
$$\{A_n \text{ i.o.}\} = \limsup A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{k\ge n} A_k,$$
prove that $\sum_{n\in\mathbb{N}} \mathsf{P}(A_n) < \infty$ implies $\mathsf{P}\{A_n \text{ i.o.}\} = 0$. ("i.o." stands for "infinitely often").
10. Let $(x_n)_{n\in\mathbb{N}}$ be a number sequence. Let us define $A_n = (-\infty, x_n)$. Prove that and $A = \limsup A_n$ (the latter is defined in the previous problem) equals either $(-\infty, x)$ or $(-\infty, x]$, where $x = \limsup x_n$.
11. Find an example of $(\Omega, \mathcal{F}, \mathsf{P})$ and r.v.'s X_n such that $X_n \to 0$ a.s., but $\mathsf{E}X_n \not\to 0$.
12. The function $f(x) = \mathbb{1}_{\mathbb{Q}}(x)$ is Lebesgue-integrable and not Riemann-integrable on $[0, 1]$. Why?
13. Let X be a nonnegative r.v. Prove that $\mathsf{E}X < \infty$ iff (if and only if) $\sum_{n\in\mathbb{N}} \mathsf{P}\{X \ge n\} < \infty$. Hint: estimate both quantities in question by sums of terms of the form $\mathsf{P}\{X \in [k, k+1)\}$.

2. Due by October 8, 3:30pm

1. Prove that if r.v.'s X and Y are independent and $\mathsf{E}|X| < \infty, \mathsf{E}|Y| < \infty$, then $\mathsf{E}XY = \mathsf{E}X\mathsf{E}Y$, i.e., $\mathsf{cov}(X,Y) = 0$.
2. Give an example of r.v.'s X and Y such that $\mathsf{cov}(X,Y) = 0$, but X and Y are not independent.
3. Let $0 < \alpha < \beta$. Then for any r.v. X, $\mathsf{E}|X|^\beta < \infty$ implies $\mathsf{E}|X|^\alpha < \infty$.
4. Let $0 < \alpha < \beta$. Give an example of a r.v. X such that $\mathsf{E}|X|^\beta = \infty$ and $\mathsf{E}|X|^\alpha < \infty$.
5. Prove that if a sequence $(Y_n)_{n\in\mathbb{N}}$ of r.v.'s converges in probability iff it is *Cauchy in probability*, i.e.,
$$\lim_{\substack{n\to\infty \\ m\to\infty}} \mathsf{P}\{|Y_n - Y_m| > \varepsilon\} = 0.$$
6. Suppose a family of r.v.'s $(Y_n)_{n\in\mathbb{N}}$ satisfies
$$\lim_{\substack{n\to\infty \\ m\to\infty}} \mathsf{P}\left\{\max_{m\le k\le n} |Y_k - Y_m| > \varepsilon\right\} = 0.$$
Prove that Y_n converges a.s. as $n \to \infty$.
7. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of r.v.'s. Let $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n \in \mathbb{N}$. Let $\mathcal{F}_\infty = \sigma(X_1, X_2, \ldots)$. Prove that for every set $A \in \mathcal{F}_\infty$ and every $\varepsilon > 0$ there is $n \in \mathbb{N}$ and a set $B \in \mathcal{F}_n$ such that $\mathsf{P}(A \triangle B) < \varepsilon$.
8. Prove that if X_1, X_2, X_3, X_4 are mutually independent r.v.'s and
$$Y = f(X_1, X_2), \quad Z = g(X_3, X_4),$$
for some measurable functions f, g, then Y and Z are independent.
9. Prove the second Borel–Cantelli lemma: If events $(A_n)_{n\in\mathbb{N}}$ are independent, then $\sum_{n\in\mathbb{N}} \mathsf{P}(A_n) = \infty$ implies $\mathsf{P}\{A_n \text{ i.o.}\} = 1$.
10. Let $(X_n)_{n\ge 2}$ be a sequence of independent r.v.'s with the following properties:
$$\mathsf{P}\{X_n = 0\} = 1 - \frac{1}{n\ln n}, \quad \mathsf{P}\{X_n = \pm n\} = \frac{1}{2n\ln n}.$$
Prove that $(X_2 + \ldots + X_n)/n$ converges in probability and does not converge a.s. (as $n \to \infty$) Hint: for the latter you may use the second Borel–Cantelli lemma, to prove that $\mathsf{P}\{|X_n| \ge n \text{ } i.o.\} = 1$.
11. Let $\Omega = [0,1], \mathcal{F} = \mathcal{B}([0,1]), \mathsf{P} = \mathsf{Leb}$. For each $\omega \in \Omega$ define $a_k(\omega) \in \{0,1\}$ via
$$\omega = \sum_{k=1}^{\infty} a_k(\omega) 2^{-k},$$
where $a_k = 0$ for large k if $\omega = j/2^n$ for some j, n. Prove that (a_k) is an i.i.d. sequence with $\mathsf{P}\{a_k = 1\} = \mathsf{P}\{a_k = 0\} = 1/2$.
12. Prove that if (a_k) is an i.i.d. sequence with $\mathsf{P}\{a_k = 1\} = \mathsf{P}\{a_k = 0\} = 1/2$, then $U = \sum_{k=1}^{\infty} a_k 2^{-k}$ is uniformly distributed on $[0,1]$.

13. Let F be a distribution function. Define $F^{-1}(y) = \inf\{x : F(x) > y\}$, $y \in [0,1]$. Let $X = F^{-1}(U)$ where U is uniformly distributed on $[0,1]$. Prove that $\mathsf{P}\{X \le x\} = F(x)$ for all x. Remark: F^{-1} is often called the quantile transform since it maps *quantiles* of the two distributions involved onto each other.
14. For any sequence of distribution functions $(F_n)_{n\in\mathbb{N}}$, use the last 3 problems to construct a family of independent r.v.'s $(X_n)_{n\in\mathbb{N}}$ on $([0,1], \mathcal{B}, \mathsf{Leb})$ such that X_n has distribution function F_n.

3. Due by October 22, 3:30pm

All measures are assumed to be Borel.

1. Prove that $x_n \to x$ iff $\delta_{x_n} \Rightarrow \delta_x$, where δ_x denotes the Dirac probability measure concentrated at x, and $\Rightarrow$ denotes weak convergence.
2. Prove that if $\int_{\mathbb{R}} f\,d\mu = \int_{\mathbb{R}} f\,d\nu$ for all $f \in C_b$, then measures μ and ν coincide.
3. Explain why convergence in distribution does not, in general, imply convergence in probability.
4. Prove that if c is a nonrandom number and $X_n \xrightarrow{d} c$, then $X_n \xrightarrow{\mathsf{P}} c$.
5. Prove that $\mu_n \Rightarrow \mu$ iff $\lim_{n\to\infty} F_{\mu_n}(x) = F_\mu(x)$ for all points x where F_μ is continuous. Here we use the notation $F_\mu(x) = \mu(-\infty, x]$.
6. Prove that $\mu_n \Rightarrow \mu$ iff $\limsup_{n\to\infty} \mu_n(A) \le \mu(A)$ for all closed sets $A \subset \mathbb{R}$.
7. Prove that $\mu_n \Rightarrow \mu$ iff $\liminf_{n\to\infty} \mu_n(A) \ge \mu(A)$ for all open sets $A \subset \mathbb{R}$.
8. The following inequality was explained in class: there is $K > 0$ such that for any probability measure μ on $\mathbb{R}$ and any $a > 0$,
$$\mu((-a^{-1}, a^{-1})^c) \le \frac{K}{a}\int_0^a (1 - \operatorname{Re}\phi(t))\,dt,$$
where $\phi = \phi_\mu$ is the characteristic function of μ. Use this inequality to prove that if $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of probability measures such that their characteristic functions ϕ_{μ_n} converge to some function ϕ pointwise, and ϕ is continuous at 0, then (μ_n) is a tight family.
9. Suppose that a sequence of probability measures μ_n and a probability measure ν satisfy the following condition: for every sequence $n' \to \infty$, there is a subsequence (n'') of (n') such that $\mu_{n''} \Rightarrow \nu$. Prove that $\mu_n \Rightarrow \nu$.
10. Prove that if $a \in \mathbb{R}$ and $\sigma > 0$, then
$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$
is a probability density, i.e., $\int_{\mathbb{R}} p(x)dx = 1$. (This is the density of the Gaussian distribution $\mathcal{N}(a, \sigma^2)$.)
11. Use integration by parts to prove that if X has standard Gaussian distribution (i.e., $\mathcal{N}(0,1)$), then
$$\mathsf{E}X^n = \begin{cases} 0, & n = 2k-1,\ k \in \mathbb{N}, \\ (2k-1)!! = (2k-1)\cdot(2k-3)\cdot\ldots\cdot 3\cdot 1, & n = 2k,\ k \in \mathbb{N}. \end{cases}$$
12. Use the previous problem to compute the characteristic function of $\mathcal{N}(a, \sigma^2)$ (start with $\mathcal{N}(0,1)$).
13. Prove that if X and Y are independent r.v.'s and their distributions have densities p_X and, respectively, p_Y with respect to Lebesgue measure, then the distribution of $X+Y$ also has a density p_{X+Y} given by the *convolution* formula
$$p_{X+Y}(z) = \int_{\mathbb{R}} p_X(x)p_Y(z-x)dx, \quad z \in \mathbb{R}.$$

14. Use the formula from the previous problem to prove that the sum of two independent r.v.'s with distributions $\mathcal{N}(a_1, \sigma_1^2)$ and $\mathcal{N}(a_2, \sigma_2^2)$ has distribution $\mathcal{N}(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$.
15. Using a discrete analogue of the above convolution formula, prove that the sum of two independent r.v.'s with Poisson distribution with parameters λ_1 and λ_2 respectively, is also a Poisson r.v., with parameter $\lambda_1 + \lambda_2$.
16. Prove the results from problems 14 and 15 using characteristic functions.
17. Prove the (weak) LLN for i.i.d. r.v.'s with finite first moment using characteristic functions.
18. Prove the following Poisson limit theorem using characteristic functions: Let $\lambda \in \mathbb{R}$ and $(X_{nk}, n \in \mathbb{N},\ 1 \le k \le n)$ be a ("triangular") array of r.v.'s such that for each n, $X_{n1}, \ldots, X_{nn}$ are i.i.d. Bernoulli with parameter $p_n \in (0,1)$ (i.e., X_{n1} takes values 1 and 0 with probabilities p_n and $1 - p_n$) such that $\lim_{n\to\infty} np_n = \lambda$. Then $X_{n1} + \ldots + X_{nn}$ converges in distribution to a Poisson r.v. with parameter λ.
19. [This problem is not for grading. This is more of a self-improvement mini-project]. Let $\mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be the space of all Borel probability distributions on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. For any $\mu, \nu \in \mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we define $\mathcal{P}(\mu, \nu)$ as the set of probability measures on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ with marginals μ and ν. $\mathcal{P}(\mu, \nu) \neq \emptyset$ since $\mu \times \nu \in \mathcal{P}(\mu, \nu)$. We define

$$d(\mu, \nu) = \inf_{P \in \mathcal{P}(\mu,\nu)} \int_{\mathbb{R}^2} (|x - y| \wedge 1) P(dx, dy).$$

Prove that d is a metric on $\mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Prove that $\lim_{n\to\infty} d(\mu_n, \nu) = 0$ iff $\mu_n \Rightarrow \nu$.

4. Due by November 5, 3:30PM

1. Prove that the Lindeberg condition holds for any i.i.d. sequence of r.v.'s with finite second moment.
2. A sequence $(X_n)_{n\in\mathbb{N}}$ of independent r.v.'s satisfies the Lyapunov condition if for some $\delta > 0$

$$\lim_{n\to\infty} \frac{1}{D_n^{2+\delta}} \sum_{j=1}^{n} \mathsf{E}|X_j - m_j|^{2+\delta} = 0,$$

where $m_j = \mathsf{E}X_j$ and $D_n^2 = \sum_{j=1}^n \mathsf{E}(X_j - m_j)^2$. Prove that the Lindeberg condition follows from the Lyapunov condition.
3. Give an example of a sequence of centered independent r.v.'s with finite second moment such that both the Lindeberg condition and CLT fail.
4. Prove that the Poisson distribution with parameter $\lambda > 0$ is infinitely divisible.
5. Suppose X has Poisson distribution with parameter $\lambda > 0$, and $(Y_n)_{n\in\mathbb{N}}$ is an i.i.d. family of r.v.'s independent of X. Prove that

$$Z = \sum_{k=1}^{X} Y_k$$

is a r.v. Prove that the distribution of Z is infinitely divisible. (The previous problem is a specific case where $Y_n \equiv 1$.)
6. Prove that Gaussian distributions are infinitely divisible.
7. Suppose r.v.'s $X_1, \ldots, X_n$ are independent and each one has an infinitely divisible distribution. Is the distribution of $X_1 + \ldots + X_n$ infinitely divisible?
8. Prove that the Cauchy distribution with density $p(x) = \frac{1}{\pi}\frac{1}{1+x^2}$, $x \in \mathbb{R}$ is stable.
9. Suppose r.v.'s $X_1, \ldots, X_n$ are independent and each one has a stable distribution. Is the distribution of $X_1 + \ldots + X_n$ stable?
10. Prove that if a distribution is stable it is infinitely divisible.
11. Give an example of an infinitely divisible distribution that is not stable.
12. Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d. r.v.'s with uniform distribution on $[-1, 1]$. Let $Y_n = 1/X_n$ for all n. Find $\alpha \in (0, \infty)$ such that

$$\frac{\sum_{j=1}^n Y_j}{n^\alpha}$$

converges in distribution to a nonconstant r.v.
13. For random vectors there is a theory of characteristic functions parallel to that for random variables. Let $X = (X_1, \ldots, X_d)$ be a random vector in $\mathbb{R}^d$, i.e., each of $X_1, \ldots, X_d$ is a r.v. The characteristic function of X is defined then as

$$\phi_X(t) = \phi_X(t_1, \ldots, t_d) = \mathsf{E}e^{i\langle t, X\rangle} = \mathsf{E}e^{i(t_1 X_1 + \ldots + t_d X_d)}, \quad t = (t_1, \ldots, t_d) \in \mathbb{R}^d.$$

(The angular brackets stand for the standard inner product in $\mathbb{R}^d$). Just as in the one-dimensional case, it turns out that for any $d \in \mathbb{N}$, weak convergence of distributions in $\mathbb{R}^d$ is equivalent to convergence of their characteristic functions. Use this to prove the following Cramér-Wold theorem: a sequence of random d-dimensional vectors $(X^{(n)})_{n\in\mathbb{N}}$ converges in distribution (as $n \to \infty$) to a random vector Y iff for any nonrandom vector $t \in \mathbb{R}^d$, $\langle t, X^{(n)} \rangle$ converges in distribution to $\langle t, Y \rangle$.

14. [Not for grading] Use Problem 13 to state a multi-dimensional version of CLT.

5. Due by November 26, 3:30 PM

1. Prove Jensen's inequality: if $f : \mathbb{R} \to \mathbb{R}$ is a convex function and X is a random variable such that $\mathsf{E}|X| < \infty$, then either $\mathsf{E}f(X) = \infty$ or
$$f(\mathsf{E}X) \le \mathsf{E}f(X) < \infty.$$
To do this, use the following property of convex functions: for every $x_0 \in \mathbb{R}$ there is $a(x_0)$ such that for all $x \in \mathbb{R}$, $f(x) \ge f(x_0) + a(x_0)(x - x_0)$.
2. Prove Jensen's inequality for conditional expectations: if $f : \mathbb{R} \to \mathbb{R}$ is a convex function and X is a random variable such that $\mathsf{E}|X| < \infty$ and $\mathsf{E}|f(X)| < \infty$, then for any σ-algebra $\mathcal{G} \subset \mathcal{F}$,
$$f(\mathsf{E}[X|\mathcal{G}]) \le \mathsf{E}[f(X)|\mathcal{G}].$$
3. Prove that if $\mathsf{E}X^2 < \infty$, then $\mathsf{E}[(\mathsf{E}(X|\mathcal{G}))^2] < \infty$ for any σ-algebra $\mathcal{G}$. (In class we interpreted conditional expectations as orthogonal projectors in L^2, but we did not check the statement of this problem). One way to do this is to use Jensen's inequality.
4. Let $X_n \xrightarrow{L^p} X$ for some $p \ge 1$, i.e., $\mathsf{E}|X_n - X|^p \to 0$. Show that $\mathsf{E}[X_n|\mathcal{G}] \xrightarrow{L^p} \mathsf{E}[X|\mathcal{G}]$.
5. Suppose X, Y are i.i.d. r.v.'s such that $\mathsf{E}|X| < \infty$. Prove that
$$\mathsf{E}(X|X+Y) = \mathsf{E}(Y|X+Y) = \frac{X+Y}{2}.$$
6. Prove that a r.v. X and a sigma-algebra $\mathcal{G}$ are independent (i.e., that for every $B \in \mathcal{G}$, r.v.'s X and $\mathbb{1}_B$ are independent) iff $\mathsf{E}(g(X)|\mathcal{G}) = \mathsf{E}g(X)$ for every bounded measurable function g.
7. Let $\mathcal{G}$ be a sigma-algebra, and X, Y be r.v.'s such that X is $\mathcal{G}$-measurable and Y is independent of $\mathcal{G}$. Let F be a bounded function measurable with respect to $\mathcal{B}(\mathbb{R}^2)$ and let $a(x) = \mathsf{E}F(x, Y)$. Prove that
$$\mathsf{E}(F(X,Y)|\mathcal{G}) = a(X).$$
(Hint: start with functions $F(x,y) = \mathbb{1}_A(x)\mathbb{1}_B(y)$)
8. Let the random point (X, Y) be uniformly distributed in $0 < x < 1$, $0 < y < x$. Find the distribution of Y conditioned on $X = x$, for every $x \in (0,1)$. Find the distribution of $\mathsf{E}[Y|X]$.
9. Recall that we defined a Markov process with initial distribution ρ and transition probability $P(\cdot,\cdot)$ via
$$\mathsf{P}\{X_0 \in A_0, \ldots, X_m \in A_m\}$$
$$= \int_{A_0} \rho(dx_0) \int_{A_1} P(x_0, dx_1) \ldots \int_{A_{m-1}} P(x_{m-2}, dx_{m-1}) \int_{A_m} P(x_{m-1}, dx_m).$$
Prove that this definition is equivalent to
$$\mathsf{E}f(X_0, \ldots, X_m)$$
$$= \int_{\mathbb{R}} \rho(dx_0) \int_{\mathbb{R}} P(x_0, dx_1) \ldots \int_{\mathbb{R}} P(x_{m-1}, dx_m) f(x_0, \ldots, x_m),$$

for any m and any bounded Borel function $f : \mathbb{R}^{m+1} \to \mathbb{R}$.

10. Prove that if $X = (X_n)_{n\geq 0}$ is a Markov process with transition kernel $P(\cdot,\cdot)$, then for any $m \geq 0, n \in \mathbb{N}$, and any sets $A_1, \ldots, A_n$,

$$\begin{aligned}
&\mathsf{P}(X_{m+1} \in A_1, \ldots, X_{m+n} \in A_n | X_0, \ldots, X_m) \\
\stackrel{a.s.}{=} &\mathsf{P}(X_{m+1} \in A_1, \ldots, X_{m+n} \in A_n | X_m) \\
\stackrel{a.s.}{=} &\int_{A_1} P(X_m, dx_1) \ldots \int_{A_{n-1}} P(x_{n-2}, dx_{n-1}) \int_{A_n} P(x_{n-1}, dx_n).
\end{aligned}$$

In particular,

$$\begin{aligned}
\mathsf{P}(X_{m+n} \in A | X_0, \ldots, X_m) &\stackrel{a.s.}{=} \mathsf{P}(X_{m+n} \in A | X_m) \\
\stackrel{a.s.}{=} \int_{\mathbb{R}} P(X_m, dx_1) \ldots &\int_{\mathbb{R}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, A).
\end{aligned}$$

for any Borel set A.

11. Use Problem 7 to prove the following: Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is a Borel function. Let X_0 be a r.v. with distribution ρ. Let $(W_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence of r.v.'s. Define inductively $X_n = f(X_{n-1}, W_n)$ for $n \in \mathbb{N}$. Prove that X_n is a Markov process with initial distribution ρ and transition probability $P(\cdot,\cdot)$ defined by

$$P(x, A) = \mathsf{P}\{f(x, W_1) \in A\}.$$

12. Prove that if (X_n) is an i.i.d. sequence, then

$$S_n = \begin{cases} 0, & n = 0, \\ X_1 + \ldots + X_n, & n \in \mathbb{N} \end{cases}$$

is a Markov process. Find a transition kernel for this process. (Hint: you may use Problem 11).

6. Practice problems on Markov chains. Not for grading. Some of these problems will be given on the in-class final exam

1. Let τ be a stopping time w.r.t. a filtration $(\mathcal{F}_n)_{n\ge 0}$. Prove that
$$\mathcal{F}_\tau = \{A \in \mathcal{F} : \text{ for all } n \ge 0, A \cap \{\tau \le n\} \in \mathcal{F}_n\}$$
is a σ-algebra. Prove that τ is $\mathcal{F}_\tau$-measurable. Prove that if (X_n) is a process adapted to $(\mathcal{F}_n)$ (i.e. X_n is $\mathcal{F}_n$-measurable for all n), then X_τ is a $\mathcal{F}_\tau$-measurable r.v.
2. In our proof of the fact that all states of an irreducible (and countable state space) Markov chain have the same type, we denoted $p = \mathsf{P}_i\{\tau_j < \tau_i\}$, where $\tau_i = \min\{n \ge 1 : X_n = i\}$ stands for the hitting time for state i, and from the inequality
$$\mathsf{E}_i\tau_j \le \mathsf{E}_i\tau_i + (1-p)\mathsf{E}_i\tau_j \tag{6.1}$$
we derived that
$$\mathsf{E}_i\tau_j \le \frac{\mathsf{E}_i\tau_i}{p} < \infty. \tag{6.2}$$
But, in fact, there is no contradiction in having $\mathsf{E}_i\tau_j = \infty$ in (6.1). Fix this by considering truncated times $\tau_j^N = \tau_j \wedge N$, proving estimates on τ_j^N analogous to (6.1),(6.2), and letting $N \to \infty$.
3. In class, we proved that a state i of a Markov chain is transient iff
$$\sum_{n=1}^{\infty} p_{ii}^N < \infty.$$
Prove that if the Markov chain is irreducible and i is transient, then for any other state j
$$\sum_{n=1}^{\infty} p_{ji}^N < \infty.$$
4. Consider the following simple random walk on $\mathbb{Z}$. For a number $p \in (0,1)$, set
$$p_{i,i+1} = p, \quad p_{i,i-1} = 1-p, \quad i \in \mathbb{Z},$$
and set $p_{ij} = 0$ if $|i-j| \ne 1$. Prove that if $p = 1/2$, then this Markov chain is null-recurrent. Prove that if $p \ne 1/2$, then it is transient.
5. Consider simple random walk in $\mathbb{Z}^d$. At each step it jumps to one of the $2d$ nearest neighbors of the current state, i.e.,
$$p_{ij} = \begin{cases} \frac{1}{2d}, & |i-j| = 1, \\ 0, & \text{otherwise.} \end{cases}$$
Prove that for $d = 2$ this Markov chain is recurrent.
6. Prove that simple random walk in $\mathbb{Z}^d$ is transient for $d = 3$.

7. The simple random walk on a graph is a Markov chain on the vertices of the graph. At each step the MC chooses one of the neighbors (vertices connected to the current one by an edge) uniformly among all the neighbors, and jumps to that vertex.

   Let $\mathbb{T}$ be an infinite tree such that all vertices have degree 3 (in other words, every two vertices in this graph are connected by a path consisting of edges of the graph, there are no loops, and every vertex has exactly 3 neighbors). Prove that the simple random walk on $\mathbb{T}$ is transient.
8. Suppose X is a recurrent irreducible Markov chain. Recall that $\tau_h = \min\{n \geq 1 : X_n = h\}$. Prove that the average time spent in state i during one excursion from a state h
$$\rho_i = \mathsf{E}_h \sum_{k=1}^{\infty} \mathbb{1}_{\{X_k=i,\ k<\tau_h\}}$$
   is finite.
9. Consider the following Markov chain on $\mathbb{N} \cup \{0\}$:
$$p_{ij} = \begin{cases} 1/3, & j = i+1, \\ 2/3, & j = i-1,\ i > 0, \\ 2/3, & i, j = 0, \\ 0, & \text{otherwise.} \end{cases}$$
   Is it recurrent? transient? positive-recurrent? Find $\lim_{n\to\infty} p_{ij}^{(n)}$ for all i, j.
10. Suppose the number of states in the Markov chain is finite. Show that there are positive recurrent states then. Derive that if there is only one communication class there is a unique invariant distribution.
11. Suppose an irreducible Markov chain on $\mathbb{N}$ defined by transition probabilities p_{ij} admits a stationary distribution $\pi = (\pi_1, \pi_2, \ldots)$, i.e. $\pi_i \geq 0$ for all i, $\sum \pi_i = 1$ and
$$\pi_i = \sum \pi_j p_{ji}.$$
   Prove that the chain is positive recurrent.

7. Practice problems on martingales. Not for grading. Some of these problems will be given on the in-class final exam

1. Prove that if τ and σ are stopping times w.r.t. a filtration $(\mathcal{F}_n)$, then $\tau \wedge \sigma$ and $\tau \vee \sigma$.
2. Prove that if τ and σ are stopping times w.r.t. a filtration $(\mathcal{F}_n)$ and $\tau \le \sigma$, then $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$.
3. Let $(X_k)_{k\in\mathbb{N}}$ be i.i.d. $N(0,1)$ r.v.'s. Let $S_n = X_1 + \ldots + X_n$. Find $\alpha \in \mathbb{R}$ such that $Z_n = e^{S_n + \alpha n}$ is a martingale w.r.t. its natural filtration.
4. Suppose (M_n) is a martingale and τ is a stopping time w.r.t. a filtration $(\mathcal{F}_n)$. Let $M_n^\tau = M_{\tau\wedge n}$ for all n. Prove that $(M_n^\tau, \mathcal{F}_n)$ is a martingale.
5. Suppose $(M_n, \mathcal{F}_n)$ is a square-integrable martingale (i.e., $\mathsf{E}M_n^2 < \infty$ for all n). Prove that this process has orthogonal increments:
$$\mathsf{E}(M_{n_2} - M_{n_1})(M_{n_4} - M_{n_3}) = 0$$
for all $n_1, n_2, n_3, n_4 \in \mathbb{N}$ satisfying $n_1 \le n_2 \le n_3 \le n_4$.
6. Suppose $(Y_n, \mathcal{F}_n)$ is a martingale and $(V_n, \mathcal{F}_{n-1})$ is a bounded predictable sequence. We define the process $((V\cdot Y)_n, \mathcal{F}_n)$ via
$$(V\cdot Y)_n = V_0 Y_0 + \sum_{i=1}^n V_i \Delta Y_i, \quad n \ge 0,$$
where $\Delta Y_i = Y_i - Y_{i-1}$. Prove that $((V\cdot Y)_n, \mathcal{F}_n)$ is a martingale.
7. Let N_n be the size of a population of bacteria at time n. At each time each bacterium produces a number of offspring and dies. The number of offspring is independent for each bacterium and is distributed according to the Poisson law with rate parameter $\lambda = 2$. Assuming that $N_1 = a > 0$, find the probability that the population will eventually die, i.e., find
$$\mathsf{P}\{\text{there is } n \text{ such that } N_n = 0\}.$$
Hint: Express the answer in terms of a and a number $c > 0$ such that $\exp(-cN_n)$ is a martingale (prove that such a number exists).
8. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be $[0,1)$ with Borel σ-algebra and Lebesgue measure. Let $f \in L^1(\Omega, \mathcal{F}, \mathsf{P})$. Let
$$f_n(x) = 2^n \int_{(k-1)2^{-n}}^{k2^{-n}} f(y)dy, \quad \text{for } x \in [(k-1)2^{-n}, k2^{-n}), \quad k \in \mathbb{N}.$$
Prove that $f_n(x) \to f(x)$ for Lebesgue-almost all $x \in [0,1)$. Hint: Prove that $(f_n, \mathcal{F}_n)$ is a martingale where $\mathcal{F}_n$ is the σ-algebra generated by intervals $[(k-1)2^{-n}, k2^{-n})$.