Graduate Probability, Fall 2020, USC Instructor: Steven Heilman

Name: ______________________ USC ID: ______________ Date: ___________

Signature: ______________________.
(By signing here, I certify that I have taken this test while refraining from cheating.)

Final Exam

This exam contains 9 pages (including this cover page) and 6 problems. Check to see if any pages are missing. Enter all requested information on the top of this page.

You may use your books and notes on this exam. You cannot use a calculator or any other electronic device (or internet-enabled device) on this exam. You are required to show your work on each problem on the exam. The following rules apply:

- You have 24 hours to complete the exam.
- **If you use a theorem or proposition from class or the notes or the book you must indicate this** and explain why the theorem may be applied. It is okay to just say, "by some theorem/proposition from class."
- **Organize your work**, in a reasonably neat and coherent way, in the space provided. Work scattered all over the page without a clear ordering will receive very little credit.
- **Mysterious or unsupported answers will not receive full credit**. A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit.
- If you need more space, use the back of the pages; clearly indicate when you have done this. Scratch paper is at the end of the exam.

Do not write in the table to the right. Good luck![a]

Problem	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
Total:	60	

[a]November 23, 2020,

1. (10 points) This problem proves a dominated convergence theorem for conditional expectation. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. Let $X, Y, X_1, X_2, X_3, \ldots$ be $\mathcal{F}$-measurable random variables on $(\Omega, \mathcal{F}, \mathbf{P})$. Assume that for all $n \geq 1$, $|X_n| \leq Y$ almost surely, and $\mathbf{E}\,|Y| < \infty$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. Assume that $X_1, X_2, \ldots$ converges almost surely to X. Conclude that

$$\mathbf{E}(X_1|\mathcal{G}), \mathbf{E}(X_2|\mathcal{G}), \ldots$$

converges almost surely to $\mathbf{E}(X|\mathcal{G})$.

You can freely use the conditional monotone convergence theorem: if $0 \leq X_1 \leq X_2 \leq \cdots$ are $\mathcal{F}$-measurable random variables that converge almost surely to X, then

$$\lim_{n\to\infty} \mathbf{E}(X_n|\mathcal{G}) = \mathbf{E}(X|\mathcal{G}).$$

(Hint: formulate and prove a conditional version of Fatou's Lemma, i.e. under some assumptions, show

$$\mathbf{E}(\liminf_{n\to\infty} X_n|\mathcal{G}) \leq \liminf_{n\to\infty} \mathbf{E}(X_n|\mathcal{G}).\ )$$

2. (10 points) Let H be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$. As usual we denote $||h|| := \langle h, h\rangle^{1/2}$ for all $h \in H$. We say a Hilbert space H is *separable* if there exists a countable set $h_1, h_2, \ldots \in H$ such that $||h_i|| = 1$ for all $i \geq 1$, $\langle h_i, h_j\rangle = 0$ for all $i, j \geq 1$ with $i \neq j$, and such that, $\forall\ h \in H$,

$$\lim_{n\to\infty} \left|\left| h - \sum_{i=1}^{n} \langle h, h_i\rangle h_i \right|\right| = 0.$$

Let K be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle'$. Assume that H and K are each separable. Show that H and K are linearly isometric. That is, $\exists$ a linear function $T\colon H \to K$ such that T is injective, T is surjective, and $\forall\ g, h \in H$, we have

$$\langle T(g), T(h)\rangle' = \langle g, h\rangle.$$

3. (10 points) For any natural number n and a parameter $0 < p < 1$, define an Erdös-Renyi graph on n vertices with parameter p to be a random graph (V, E) on a (deterministic) vertex set $V = \{1, \ldots, n\}$ of n vertices (thus (V, E) is a random variable taking values in the discrete space of all $2^{\binom{n}{2}}$ possible undirected graphs one can place on V) such that the events $\{i, j\} \in E$ for unordered pairs with $i, j \in V$ are independent and each edge occurs with probability p.

   A *stable set* in a graph is a subset S of the vertices of the graph such that no two of the vertices in S are connected by an edge.

   For any $n \geq 1$, let X_n be the cardinality of the largest cardinality stable set in an Erdös-Renyi random graph on n vertices with parameter $p = 1/2$. Show that

$$\mathbf{P}(|X_n - \mathbf{E}X_n| > t) \leq 2e^{-\frac{t^2}{2n}}, \qquad \forall\, n \geq 1, \quad \forall\, t > 0.$$

   (Hint: construct an increasing sequence of σ-algebras $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n = 2^{\Omega}$ such that $\mathcal{F}_j$ corresponds to all subsets of edges going between vertices $\{1, \ldots, j\}$, and define $Y_j := \mathbf{E}(X_n | \mathcal{F}_j)$, $\forall\ 0 \leq j \leq n$. Show that $Y_0, Y_1, \ldots, Y_n$ is a martingale.)

4. (10 points) Let $X_0, X_1, \ldots$ be the symmetric simple random walk on the integers $\mathbf{Z}$. For any $k \in \mathbf{Z}$, let $\mathbf{P}_k$ denote the probability law on this random walk such that $X_0 := k$. Let $T_0 := \min\{n \geq 1 \colon X_n = 0\}$ be the first positive time that the random walk takes the value 0. For any positive integers j, k, r, show:

$$\mathbf{P}_k(T_0 < r,\, X_r = j) = \mathbf{P}_k(X_r = -j).$$

(Hint: draw a picture. Show that $\mathbf{P}_k(X_{T_0+(r-s)} = j \,|\, T_0 = s,\, X_s = 0) = \mathbf{P}_0(X_{r-s} = j)$ for any $j \in \mathbf{Z}$. Then deduce $\mathbf{P}_k(T_0 = s,\, X_r = j) = \mathbf{P}_k(T_0 = s)\mathbf{P}_0(X_{r-s} = j)$. For the term on the right, can we replace $-j$ with j? Combine terms and sum over all $1 \leq s < r$. Also justify that $\mathbf{P}_k(T_0 < r,\, X_r = -j) = \mathbf{P}_k(X_r = -j)$ when $j > 0$.)

5. (10 points) Let a, b be positive integers. Suppose there are c votes cast by c people in an election. Candidate 1 gets a votes and candidate 2 gets b votes. (So $c = a + b$.) Assume $a > b$. The votes are counted one by one. The votes are counted in a uniformly random ordering, and we would like to keep a running tally of who is currently winning. Suppose the first candidate eventually wins the election. We ask: with what probability will candidate 1 always be ahead in the running tally of who is currently winning the election? You are asked to show that the answer is $\frac{a-b}{a+b}$.

   To prove this, for any positive integer k, let S_k be the number of votes for candidate 1, minus the number of votes for candidate 2, after k votes have been counted. Then, define $X_k := S_{c-k}/(c-k)$. Show that $X_0, X_1, \ldots$ is a martingale with respect to the σ-algebras generated by the (reversed) ordering of $S_c, S_{c-1}, S_{c-2}, \ldots$. Then, let T such that $T = \min\{0 \leq k \leq c \colon X_k = 0\}$, or $T = c - 1$ if no such k exists. Apply the Optional Stopping theorem to X_T.

6. (10 points) For any $a = (a_1, a_2, \ldots) \in \mathbf{R}^{\mathbf{N}}$, $b = (b_1, b_2, \ldots) \in \mathbf{R}^{\mathbf{N}}$, define $\langle a, b\rangle := \sum_{i=1}^{\infty} a_i b_i$ (if it exists). Let ℓ_2 denote the Hilbert space $\{a \in \mathbf{R}^{\mathbf{N}} \colon \langle a, a\rangle < \infty\}$ with respect to the inner product $\langle \cdot, \cdot\rangle$. (You can freely use that ℓ_2 is a Hilbert space.)

   This problem proves Grothendieck's inequality: $\exists$ a constant $k > 0$ such that, $\forall\, n \geq 1$, $\forall$ real $n \times n$ matrices $(c_{ij})_{1 \leq i,j \leq n}$,

$$\max_{\substack{x^{(1)},\ldots,x^{(n)},y^{(1)},\ldots,y^{(n)}\in\ell_2\\ ||x^{(i)}||=||y^{(i)}||=1}} \sum_{i,j=1}^{n} c_{ij}\langle x^{(i)}, y^{(j)}\rangle \leq k \cdot \max_{\varepsilon_1,\ldots,\varepsilon_n,\delta_1,\ldots,\delta_n\in\{-1,1\}} \sum_{i,j=1}^{n} c_{ij}\varepsilon_i\delta_j.$$

   To prove this inequality, do the following. Let Γ denote the left side of the inequality, and let Δ denote the right side (without the k constant). We need to show that $\Gamma = O(\Delta)$.

   - First, let $g_1, g_2, \ldots$ be a fixed sequence of i.i.d. standard Gaussians. Using a previous problem, we may replace ℓ_2 with the Hilbert space $H := \{\sum_{i=1}^{\infty} a_i g_i \colon (a_i)_{i=1}^{\infty} \in \ell_2\}$ with respect to the inner product $\langle X, Y\rangle := \mathbf{E}XY$ for all $X, Y \in H$. (You can freely use that H is itself a Hilbert space.)
   - Fix $m > 0$ and let $X \in H$. Denote $X_{\leq m} := X 1_{|X| \leq m}$. Fix $\varepsilon > 0$. Pick $X^{(1)}, \ldots, X^{(n)}$ and $Y^{(1)}, \ldots, Y^{(n)}$ that come within ε of achieving the maximum in the definition of Γ. Compare $\sum_{i,j=1}^{n} c_{ij}\mathbf{E}X^{(i)}Y^{(j)}$ to $\sum_{i,j=1}^{n} c_{ij}\mathbf{E}X_{\leq m}^{(i)}Y_{\leq m}^{(j)}$ by adding and subtracting some terms inside the sum. You should be able to show that

$$\Gamma - \varepsilon \leq m^2\Delta + 100e^{-m^2/4}\Gamma.$$

     (Hint: Try dividing and multiplying some terms by m, and try dividing and multiplying some terms by $\max_{k=1,\ldots,n} ||X^{(k)} - X_{\leq m}^{(k)}||$ or $\max_{k=1,\ldots,n} ||Y^{(k)} - Y_{\leq m}^{(k)}||$.)
   - To get the $e^{-m^2/4}$ bound, note that $X^{(k)} - X_{\leq m}^{(k)} = X^{(k)} 1_{|X^{(k)}| > m}$, so you should be able to bound its L_2 norm.

(Scratch paper)

(Extra Scratch paper)