Graduate Applied Probability II, Spring 2021, 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 10 pages (including this cover page) and 5 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 120 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.

Problem	Points	Score
1	10	
2	10	
3	15	
4	15	
5	20	
Total:	70	

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

[a]May 8, 2021, © 2021 Steven Heilman, All Rights Reserved.

1. (10 points) Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion. Let $\{Y(t)\}_{t\geq 0}$, and let $\{Z(t)\}_{t\geq 0}$ be two (coupled) Itô processes. So, $\sigma_1, \mu_1, \sigma_2, \mu_2 \colon \mathbf{R} \to \mathbf{R}$ are continuous functions such that

$$dY(t) = \sigma_1(t)dB(t) + \mu_1(t)dt, \qquad \forall\, t \geq 0.$$

$$dZ(t) = \sigma_2(t)dB(t) + \mu_2(t)dt, \qquad \forall\, t \geq 0.$$

Show that the following "product rule" holds for all $t > 0$:

$$d(YZ)(t) = Y(t)dZ(t) + Z(t)dY(t) + \sigma_1(t)\sigma_2(t)dt.$$

(Hint: apply Itô's formula to Y^2, Z^2 and $(Y+Z)^2$ each separately and then use the equality $YZ = (1/2)((Y+Z)^2 - Y^2 - Z^2)$.)

2. (10 points) Let $\{Y(t)\}_{t\geq 0}$, and let $\{Z(t)\}_{t\geq 0}$ be two Itô processes. Then, we define the **covariation** process $\{[Y,Z]_t\}_{t\geq 0}$ so that, for any $b>0$, $[Y,Z]_b$ is the limit in probability as $n\to\infty$ of

$$\sum_{i=0}^{n-1}(Y(b(i+1)/n)-Y(bi/n))(Z(b(i+1)/n)-Z(bi/n)),$$

if this limit exists. (In particular, $[Y,Y]_t=[Y]_t$ for all $t\geq 0$, where $[Y]_t$ denotes the quadratic variation of Y at time t.)

- Show that, for any $t>0$, with probability one we have

$$[Y+Z]_t=[Y]_t+[Z]_t+2[Y,Z]_t.$$

- Show that, for any $t>0$, with probability one we have

$$\big|[Y,Z]_t\big|\leq\sqrt{[Y]_t[Z]_t}.$$

  (This is a special case of the so-called Kunita-Watanabe inequality.)

- Using the arithmetic mean-geometric mean inequality, conclude that, for any $t>0$, with probability one we have

$$[Y+Z]_t\leq 2\big([Y]_t+[Z]_t\big),\qquad\forall\, t>0.$$

3. (15 points) Let $\{B(t)\}_{t\geq 0}$ be a Brownian motion in $\mathbf{R}^n$ (so that $B(t) = (B_1(t), \ldots, B_n(t))$ where $\{B(t)\}_{t\geq 0}, \ldots, \{B_n(t)\}_{t\geq 0}$ are n independent one-dimensional Brownian motions). For any $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, denote $||x|| := (x_1^2 + \cdots + x_n^2)^{1/2}$. Denote the open unit ball in $\mathbf{R}^n$ as

$$D := \{x = (x_1, \ldots, x_n) \in \mathbf{R}^n \colon \ ||x|| < 1\}.$$

For any $x \in D$, we use the notation $\mathbf{E}_x$ to denote that the Brownian motion is started at x (so that $B(0) = x$). Define a stopping time

$$T := \inf\{t > 0 \colon B(t) \in \partial D\}.$$

- This problem describes the distribution of the exit time of Brownian motion from the unit ball D. Let $f \colon \mathbf{R}^n \to \mathbf{R}$ be an infinitely differentiable function. Let μ denote the probability measure that is uniformly distributed in ∂D. Show that,

$$\mathbf{E}_x f(B(T)) = \int_{\partial D} \frac{1 - ||x||^2}{||x - y||^n} f(y) d\mu(y), \qquad \forall\, x \in D.$$

(Hint: let $k_y(x) := \frac{1-||x||^2}{||x-y||^n}$ for any $x, y \in D$ with $x \neq y$. You can freely use the fact that $\Delta k_y(x) = 0$, where Δ is the Laplacian on $\mathbf{R}^n$. Define

$$v(x) := \begin{cases} \int_{\partial D} \frac{1-||x||^2}{||x-y||^n} f(y) d\mu(y) & \text{, if } x \in D \\ f(x) & \text{, if } x \in \partial D. \end{cases}$$

Is it true that $\Delta v(x) = 0$? You do not need to justify moving the Laplacian inside the integral. Moreover, you may freely use that a solution of the Dirichlet problem is unique.)

- Let $0 < a < b < \infty$. Let U denote the annulus

$$U := \{x \in \mathbf{R}^n \colon a < ||x|| < b\}.$$

Define $u \colon \mathbf{R}^n \setminus \{0\} \to \mathbf{R}$ by

$$u(x) := \begin{cases} ||x|| & \text{, if } n = 1 \\ \log ||x|| & \text{, if } n = 2 \\ ||x||^{2-n} & \text{, if } n \geq 3. \end{cases}$$

Verify that $\Delta u(x) = 0$ for all $x \neq 0$. Define

$$T_a := \inf\{t > 0 \colon \ ||B(t)|| = a\}, \qquad T_b := \inf\{t > 0 \colon \ ||B(t)|| = b\}.$$

Show that, for any $x \in U$,

$$\mathbf{P}_x(T_a < T_b) = \frac{u(b, 0, \ldots, 0) - u(x)}{u(b, 0, \ldots, 0) - u(a, 0, \ldots, 0)}.$$

(Hint: Since u is harmonic, how is $u(x)$ related to $\mathbf{E}_x u(B(T_a \wedge T_b))$?)

- Let $x \in \mathbf{R}^n$ with $||x|| > a > 0$. Conclude that

$$\mathbf{P}_x(T_a < \infty) = \begin{cases} 1 & \text{, if } n \leq 2 \\ (a/\,||x||)^{n-2} & \text{, if } n \geq 3. \end{cases}$$

  (This is analogous to our recurrience/transience results for the simple random walk on $\mathbf{Z}^n$.)

4. (15 points) Suppose we have n finite, irreducible transition matrices $P_1, \ldots, P_n$ with n corresponding state spaces $\Omega_1, \ldots, \Omega_n$ and stationary distributions $\pi_1, \ldots, \pi_n$. Define $\Omega := \Omega_1 \times \cdots \times \Omega_n$. Let $w = (w_1, \ldots, w_n)$ be a probability distribution on $\{1, \ldots, n\}$. Consider the (discrete-time) Markov chain on Ω that, at each step, selects coordinate $1 \leq j \leq n$ with probability w_j, and then changes the state of the chain only in coordinate j according to P_j. The transition matrix P for this chain is then

$$P(x, y) := \sum_{j=1}^{n} w_j P_j(x_j, y_j) \prod_{\substack{k \in \{1, \ldots, n\}: \\ k \neq j}} 1_{\{x_k = y_k\}}, \quad \forall\, x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \Omega.$$

(You do not have to prove this.) For each $1 \leq j \leq n$, let $f_j \colon \Omega_j \to \mathbf{R}$, and define the tensor product function $f_1 \otimes f_2 \otimes \cdots \otimes f_n \colon \Omega \to \mathbf{R}$ by

$$(f_1 \otimes f_2 \otimes \cdots \otimes f_n)(x_1, \ldots, x_n) := \prod_{j=1}^{n} f_j(x_j), \qquad \forall\, x = (x_1, \ldots, x_n) \in \Omega.$$

Let $\pi := \pi_1 \otimes \cdots \otimes \pi_n$ (so that π is a probability distribution on Ω, where we consider each π_j to be a function on Ω_j in order to use the tensor product definition). Then π is stationary for P (you do not have to prove this). Assume that, for any $1 \leq j \leq n$, the transition matrix P_j has an eigenfunction $f_j \in \mathbf{R}^{\Omega_j}$ with eigenvalue λ_j.

- Show that the function $f := f_1 \otimes \cdots \otimes f_n$ is an eigenfunction of P with eigenvalue $\sum_{j=1}^{n} w_j \lambda_j$.
- Assume that, for each $1 \leq j \leq n$, $\mathcal{B}_j$ is an orthonormal basis for $\mathbf{R}^{\Omega_j}$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\pi_j}$. Show that

  $$\mathcal{B} := \{f_1 \otimes f_2 \otimes \cdots \otimes f_n \colon f_j \in \mathcal{B}_j,\ \forall\, 1 \leq j \leq n\}$$

  is an orthonormal basis with respect to the inner product $\langle \cdot, \cdot \rangle_{\pi}$.
- If P_j has spectral gap γ_j for all $1 \leq j \leq n$, show that P has spectral gap

  $$\gamma := \min_{1 \leq j \leq n} w_j \gamma_j.$$

- For any $1 \leq j \leq n$, let $\Omega_j := \{-1, 1\}$ and let $P_j(a, b) = 1$ for all $a, b \in \{-1, 1\}$ with $a \neq b$ and $P_j(a, b) = 0$ otherwise. Let $w := (1/n, \ldots, 1/n)$. Then P on Ω corresponds to the simple random walk on the discrete hypercube $\{-1, 1\}^n$. (You do not have to prove this.) Conclude that the spectral gap for the Markov chain P on Ω is

  $$\gamma = 2/n.$$

- Using the result from Exam 2, then give a bound on the mixing time of the (continuous-time) Markov chain P on $\Omega = \{-1, 1\}^n$.

5. (20 points) This problem is a continuation of the mixing time problem from Exam 2. Suppose we have a finite, irreducible Markov chain with state space Ω, transition matrix P, and stationary distribution π. For any $0 < p < \infty$, and for any $f \in \mathbf{R}^\Omega$, define

$$||f||_{p,\pi} := \Big(\sum_{x\in\Omega} |f(x)|^p \,\pi(x)\Big)^{1/p}, \qquad \mathbf{E}_\pi f := \sum_{x\in\Omega} f(x)\pi(x).$$

Another way to bound the mixing time is to give a bound on the logarithmic-Sobolev constant of the chain (or log-Sobolev constant). Define $\alpha \geq 0$ so that $1/\alpha$ is the smallest constant $c > 0$ such that, for all $f \in \mathbf{R}^\Omega$ with $||f||_{2,\pi} \neq 0$,

$$\sum_{x\in\Omega} |f(x)|^2 \Big[\log\Big(\frac{|f(x)|^2}{||f||_{2,\pi}^2}\Big)\Big]\pi(x) \leq c \cdot \frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 P(x,y)\pi(x). \qquad (*)$$

(If no such c exists, $\alpha := 0$.) (Note that $(*)$ is dilation invariant, i.e. if $f \in \mathbf{R}^\Omega$ satisfies $(*)$, then tf also satisfies $(*)$ for all $t > 0$.) You can freely use the following bound:

$$\max_{x\in\Omega} ||H_t(x,\cdot) - \pi(\cdot)||_{\mathrm{TV}} \leq \sqrt{\frac{1}{2}\log\Big(\frac{1}{\min_{y\in\Omega}\pi(y)}\Big)}\, e^{-\alpha t}, \qquad \forall\, t > 0.$$

- Give an upper bound on the mixing time of a (continuous-time) Markov chain in terms of the log-Sobolev constant α.
- Additionally assume the Markov chain is reversible. Show that the spectral gap γ can be equivalently defined so that $1/\gamma$ is the smallest constant $b > 0$ such that, for all $f \in \mathbf{R}^\Omega$,

  $$\frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 \pi(x)\pi(y) \leq b \cdot \frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 P(x,y)\pi(x).$$

  (If no such b exists, $\gamma := 0$.) (Hint: the left side is the variance of f.)
- Additionally assume the Markov chain is reversible. Show that $2\alpha \leq \gamma$. (Hint: consider $f = 1 + \varepsilon g$ in the definition of the log-Sobolev constant, and let $\varepsilon \to 0^+$.)
- Let us re-use the notation of the previous problem to define $P_1, \ldots, P_n$, $\Omega_1, \ldots, \Omega_n$, $\pi_1, \ldots, \pi_n$, Ω, w and P. Show: if P_j has logarithmic Sobolev constant α_j for all $1 \leq j \leq n$, then P has logarithmic Sobolev constant

  $$\alpha := \min_{1\leq j\leq n} w_j \alpha_j.$$

  (Hint: just consider the case $n = 2$. Let $f\colon \Omega_1 \times \Omega_2 \to \mathbf{R}$. Define $F\colon \Omega_2 \to \mathbf{R}$ by $F(x_2) := \sqrt{\sum_{x_1\in\Omega_1} |f(x_1,x_2)|^2 \pi_1(x_1)}$, $\forall\, x_2 \in \Omega_2$. Observe that

  $$\sum_{x\in\Omega} |f(x)|^2 \Big[\log\Big(\frac{|f(x)|^2}{||f||_{2,\pi}^2}\Big)\Big]\pi(x) = \sum_{x_2\in\Omega_2} |F(x_2)|^2 \log\Big(\frac{|F(x_2)|^2}{||F||_{2,\pi_2}^2}\Big)\pi_2(x_2)$$
  $$+ \sum_{(x_1,x_2)\in\Omega_1\times\Omega_2} |f(x_1,x_2)|^2 \Big[\log\Big(\frac{|f(x_1,x_2)|^2}{|F(x_2)|^2}\Big)\Big]\pi(x_1,x_2).$$

Apply the definition of α_2 to the first term, and apply the definition of α_1 to the second term. Then, observe that $|F(a) - F(b)| \leq ||f(\cdot, a) - f(\cdot, b)||_{2,\pi_1}$ for all $a, b \in \Omega_2$ to get a bound on the first term in terms of $f(x_1, \cdot)$.)

- For any $1 \leq j \leq n$, let $\Omega_j := \{-1, 1\}$ and let $P_j(a, b) = 1$ for all $a, b \in \{-1, 1\}$ with $a \neq b$ and $P_j(a, b) = 0$ otherwise. Let $w := (1/n, \ldots, 1/n)$. Then P on Ω corresponds to the simple random walk on the discrete hypercube $\{-1, 1\}^n$. (You do not have to prove this.) You can freely use that $\alpha_j = 1$ for all $1 \leq j \leq n$. Conclude that the log-Sobolev constant for this Markov chain P on Ω is

$$\alpha = 1/n.$$

How does the resulting bound on the mixing time of the corresponding continuous-time Markov chain compare with the bound from the previous problem?

(Scratch paper)

(Extra Scratch paper)