EECS 126 Probability and Random Processes — UC Berkeley, Spring 2023
Thomas Courtade — May 11, 2023, 3:10-6:00pm

Final

Last Name	First Name	SID

Left Neighbor Full Name	**Right Neighbor** Full Name	**Room Number**

- **Write your SID on every page to receive 1 point.**
- **All work you want graded should be on the fronts of the sheets in the space provided. Back sides may be used for scratch work, but will not be scanned/graded.**
- **Write all answers clearly. Answers that are not legible will not receive credit.**
- Unless otherwise stated, all your answers need to be justified and your work must be shown.
- You have 170 minutes to complete the exam. (DSP students with $X\,\%$ time accommodation should spend $170 \cdot X\,\%$ time on the exam).
- You are allowed three double-sided sheets of notes. No calculators/phones.
- Remember the Berkeley Honor Code: "*As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others.*" Violations may result in sanctions.

Problem	points earned	out of
SID		1
Problem 1		7
Problem 2		8
Problem 3		10
Problem 4		16
Problem 5		19
Problem 6		16
Problem 7		10
Problem 8		10
Problem 9		12
Problem 10		16
Problem 11		15
Total		140

1 In Summer [3+4]

a) What are you looking forward to over the summer?

b) How has your perspective on probability changed after taking this course?

2 Convergence Implies Convergence? [3+3+2]

Let $(X_n)_{n\geq 1}$ be a sequence of random variables defined on a common probability space, satisfying $\lim_{n\to\infty} \mathbb{E}[|X_n|] = 0$. Say whether each of the following statements is true or false. If true, prove it. If false, give a counterexample.

a) $X_n \to 0$ in probability.

b) $X_n \to 0$ in mean-square (i.e., $\lim_{n\to\infty} \mathbb{E}[|X_n|^2] = 0$).

c) $X_n \to 0$ in distribution.

3 Nuts and Bolts (of Probability) [5+5]

a) Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. zero-mean, unit variance random variables. In this context, precisely state the central limit theorem *and* the definition of its mode of convergence.

b) Suppose a thousand bolts are manufactured in an independent and identical fashion, and you empirically measure the diameters to have mean 1.00cm and standard deviation 0.01cm. If the manufacturing run is extended to 1 million bolts, approximately how many would you expect to find with diameters exceeding 1.05cm? You may leave your answer in terms of the standard normal cdf Φ, but explain your reasoning.

4 Counting Triangles [3+5+8]

A *triangle* in a graph $G = (V, E)$ is a collection of three vertices $u, v, w \in V$ which are joined by three edges. Let N_Δ denote the number of triangles in $G \sim \mathcal{G}(n, p)$.

a) Compute $\mathbb{E}[N_\Delta]$.

b) Explain where each of the four terms in the following expression comes from:

$$\mathbb{E}[N_\Delta^2] = \binom{n}{3}\binom{n-3}{3}p^6 + 3\binom{n}{3}\binom{n-3}{2}p^6 + 3(n-3)\binom{n}{3}p^5 + \binom{n}{3}p^3.$$

c) Find the sharp threshold $t(n)$ (of the form $t(n) = n^{-\alpha}$ for some $\alpha > 0$) such that $p(n) \gg t(n)$ ensures that G contains a triangle with high probability, and $p(n) \ll t(n)$ ensures that G is triangle-free with high probability. Show all steps in your derivation.

(The inequality $P(X = 0) \leq \frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2}$ might be helpful for part of this.)

(Extra page for Problem 4)

5 Sisyphean Chain [3+8+8]

Throughout this problem, $p_0, p_1, \ldots$ is a given sequence of numbers in the interval $[0, 1]$.

a) Let $(F_n)_{n\geq 0}$ be independent Bernoulli trials, with $F_n \sim \text{Bernoulli}(1 - p_n)$ for each $n \geq 0$. Let $N = \min\{n \geq 0 : F_n = 1\}$ be the time of the first success. Compute $\mathbb{E}[N]$ in terms of the given sequence $p_0, p_1, \ldots$.

b) Let $(X_n)_{n\geq 0}$ be a Markov chain on state space $\{0, 1, 2, \ldots\}$ with nonzero transition probabilities given by

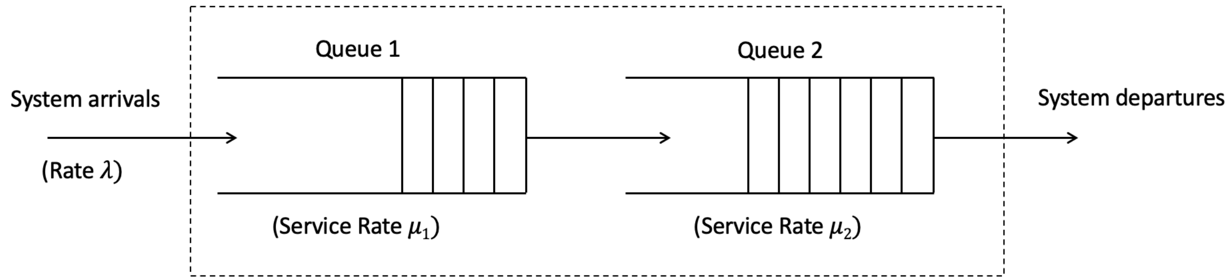
$$P_{n,0} = 1 - p_n, \quad P_{n,n+1} = p_n, \quad n \geq 0.$$

Find necessary and sufficient conditions on $p_0, p_1, \ldots$ for this chain to be irreducible and positive recurrent.

c) Assuming the conditions of part (b) hold, what is the stationary distribution for the chain $(X_n)_{n\geq 0}$ in terms of $p_0, p_1, \ldots$?

(Extra page for Problem 5)

6 Waiting in Line, Again [6+4+6]



Consider a system of two M/M/1 queues connected as illustrated above. The arrival rate to the combined system is assumed to be a Poisson process of rate λ, and the service times in the first and second queues are i.i.d. $\mathrm{Exp}(\mu_1)$ and $\mathrm{Exp}(\mu_2)$, respectively. Assume that $0 < \lambda < \min\{\mu_1, \mu_2\}$.

Let $N_{i,t}$ be the number of customers in queue $i \in \{1, 2\}$ at time $t \geq 0$. The state of the system at time $t \geq 0$ is represented by the pair $X_t := (N_{1,t}, N_{2,t})$.

a) Model the process $(X_t)_{t\geq 0}$ as a CTMC. In particular, draw a state transition diagram with clearly labeled transition rates between states.

b) Let $q_{(m,n),(m',n')}$ denote the transition rate from state (m, n) to (m', n'). Argue that if a distribution π satisfies

$$\pi_{(m,n)}q_{(m,n),(m+1,n)} = \pi_{(m+1,n)}q_{(m+1,n),(m,n+1)} = \pi_{(m,n+1)}q_{(m,n+1),(m,n)} \quad \forall m, n \geq 0,$$

then it is a stationary distribution.

(Hint: With the help of your diagram from part (a), you can answer this question without doing any math.)

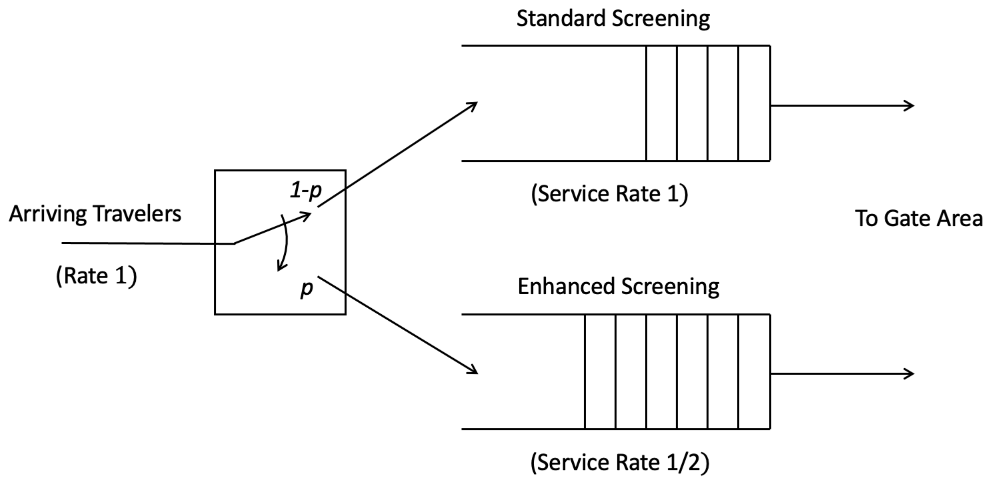
c) Compute the stationary distribution π, and formulate a simple probabilistic model for the system in steady state.

(Hint: The stationary distribution for this chain takes the form $\pi_{(m,n)} = f(m)g(n)$ for some functions f, g.)

 Student ID: ____________________

(Extra page for Problem 6)

7 TSA Checkpoint [10]



Airport security can be modeled as two parallel M/M/1 queues, with the first queue corresponding to standard screening, and the second queue corresponding to enhanced screening. The arrival rate of travelers to the checkpoint is 1/min, the average service time for standard screening is 1 min, and the average service time for enhanced screening is 2 min. (As always, *service time* only counts the time "in service", and not time spent waiting in the screening queue.)

Security officers direct travelers to enhanced screening with probability p, independent of all other travelers. What value of $p \in [0, 1]$ minimizes the expected time a customer spends (waiting and in service) in airport security? Assume the system is in steady state for your computation.

8 MAP Estimation [10]

Let X have density

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Conditioned on $\{X = x\}$, you observe $Y \sim \text{Geom}(x)$ (supported on $\{1, 2, \dots\}$). What is the MAP estimate of X given observation $Y = k$?

9 Rectangle or Triangle? [2+8+2]

Consider two densities f_0, f_1 as defined below:

$$f_0(y) = \begin{cases} \frac{1}{2} & -1 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}, \quad \text{and} \quad f_1(y) = \begin{cases} 1 - |y| & -1 \leq y \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

You observe Y, and would like to discriminate between the hypotheses:

$$H_0 : Y \sim f_0, \quad \text{and} \quad H_1 : Y \sim f_1.$$

a) What is likelihood ratio $L(y)$ for this hypothesis testing problem?

b) What is the Type II error rate of the optimal test, subject to Type I error rate at most α?

c) Sketch the error curve, and place an ‘X’ corresponding to the performance of the MLE test. Clearly label the x-y coordinates of your ‘X’.

 Student ID: ____________________

(Extra page for Problem 9)

10 Gaussian Parameters and Estimation [6+10]

The following subparts are independent of one another.

a) Let (X, Y) be jointly Gaussian vectors, with respective marginal means μ_X, μ_Y, marginal covariance matrices Σ_X, Σ_Y, and covariance $\Sigma_{XY} = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)^T]$. For (constant) matrices A, B of compatible dimensions and a vector ζ, determine the distribution of

$$Z = AX + BY + \zeta.$$

b) Suppose you observe i.i.d. $X_1, X_2, \ldots X_n$, which are assumed to be $N(\mu, \Sigma)$ with (μ, Σ) unknown. What is the maximum likelihood estimate of the pair (μ, Σ) given the observations $X_1, \ldots, X_n$?

(You may may assume μ, Σ are scalars for full credit. However, the multidimensional case is no more difficult if you are familiar with basic multivariable calculus. In particular, for a $d \times d$ positive definite matrix A and a vector $x \in \mathbb{R}^d$, you may freely use the following gradient expressions: $\nabla_A \log\det(A) = A^{-1}$. For $f(A, x) := x^T A x$, we have the partial derivatives $\nabla_A f(A, x) = xx^T$ and $\nabla_x f(A, x) = 2Ax$.)

 Student ID: ______________________________

(Extra page for Problem 10)

11 Hacking the Kalman Filter [5+10]

Consider the scalar state-space and observation model

$$X_n = aX_{n-1} + V_n$$
$$Y_n = X_n + W_n, \quad n \geq 1,$$

with $X_0 = 0, \mathrm{Var}(V_n) = \sigma_V^2, \mathrm{Var}(W_n) = \sigma_W^2$, and the usual assumption of uncorrelated, zero-mean noise processes. Assume $a \neq 0$.

Suppose someone has already implemented a Kalman filter for you. That is, at iteration n, you have the quantities Y_n, $(\hat{X}_{n|n}, \sigma^2_{n|n})$, and $(\hat{X}_{n-1|n-1}, \sigma^2_{n-1|n-1})$ available to you.

a) What extra updates should you add to do one-step prediction? I.e., what equations should you add to the Kalman filter to compute $\hat{X}_{n+1|n}$ on iteration n?

b) What extra updates should you add to do one-step smoothing? I.e., what equations should you add to the Kalman filter to compute $\hat{X}_{n-1|n}$ on iteration n?

Answers should be in terms of Y_n, $(\hat{X}_{n|n}, \sigma^2_{n|n})$, and $(\hat{X}_{n-1|n-1}, \sigma^2_{n-1|n-1})$, and any parameters of the state-space model. You may use $\mathrm{Var}(\tilde{Y}_n) = a^2\sigma^2_{n-1|n-1} + \sigma^2_V$, from lecture.

 Student ID: ________________________

(Extra page for Problem 11)