EECS 126 Probability and Random Processes
Shyam Parekh

UC Berkeley, Fall 2020
December 16, 2020

Final

Last Name	First Name	SID

Rules.

- **Unless otherwise stated, all your answers need to be justified and your work must be shown. Answers without sufficient justification will get no credit.**
- You have 160 minutes to complete the exam and 10 minutes exclusively for submitting your exam to Gradescope. (DSP students with $X\%$ time accomodation should spend $160 \cdot X\%$ time on the exam and 10 minutes to submit).
- Collaboration with others is strictly prohibited.
- You may reference your notes, the textbook, and any material that can be found through the course website. You may use Google to search up general knowledge. However, searching up a question is not allowed.
- You may not use online solvers or graphing tools (ex. WolframAlpha, Desmos, Python). Simple functions (ex. combinations, multiplication) are fine.
- For any clarifications you have, please create a private Piazza post. We will have a Google Doc that shows our official clarifications.

Problem	points earned	out of
Problem 1		16
Problem 2		7
Problem 3		7
Problem 4		7
Problem 5		8
Problem 6		10
Problem 7		10
Problem 8		13
Problem 9		7
Problem 10		8
Problem 11		7
Total		100

1 True or False (4 + 4 + 4 + 4 points)

For each of the following, say whether the assertion is true or false. If it is true, provide justification, and if it is false, give a counterexample.

(a) For any finite sample space Ω and any event $A \subseteq \Omega$, $\Pr(A) = \frac{|A|}{|\Omega|}$.

> False, any sample space with non-uniform outcomes (and corresponding event) is a valid example.

(b) If $Y = X + Z$ is Gaussian, then X and Z are both marginally Gaussian.

> False. Consider $U = \mathcal{N}(0,1)$ and let $X = U1\{U \geq 0\}$ and $Z = U1\{U < 0\}$. Then, Y is Gaussian but neither X nor Z are marginally Gaussian.

(c) Michael and Kevin are playing a game where Michael scores $X_i \sim \text{Geometric}(1/m)$ points and Kevin scores $Y_i \sim \text{Geometric}(1/k)$ points at round i, independently of other rounds. Then, as the number of rounds goes to infinity, the average total number of points they score per round converges almost surely to $m + k$.

> True. This is SLLN on the iid variables $(X_i + Y_i)$.

(d) Suppose a random variable X is bounded in $[0, 1]$, and furthermore suppose that $\mathbb{E}[X] \geq \epsilon$. Then $\Pr(X \geq \epsilon/2) \geq \epsilon/2$.

> Suppose towards a contradiction this is not the case. Then we have
>
> $$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X|X \geq \epsilon/2] \Pr(X \geq \epsilon/2) + \mathbb{E}[X|X < \epsilon/2] \Pr(X < \epsilon/2) \\ &< 1 \cdot \frac{\epsilon}{2} + \frac{\epsilon}{2} \cdot 1 = \epsilon \end{aligned}$$
>
> which is a contradiction to our assumption that $\mathbb{E}[X] \geq \epsilon$.

2 Waiting Game (7 points)

The amount of time you have to wait for UC Berkeley to announce the PNP policy is exponentially distributed with parameter λ. Solve for the optimal Chernoff upper bound (by applying Markov's inequality to e^{sX}) for the probability that you have to wait for longer than some constant a.

> By memoryless property of exponentials, if X = time until the announcement, then X is exponentially distributed with parameter λ.
>
> $\Pr(X > a) = \Pr(e^{sX} > e^{sa}) \leq \frac{\mathbb{E}[e^{sX}]}{e^{sa}}$
> Using MGF of an exponential distribution, $\Pr(X > a) \leq \min_s \frac{\lambda}{(\lambda - s)e^{sa}}$
> This is equivalent to maximizing the denominator. Taking the derivative of the denominator

and setting to zero, we have

$$e^{sa}(\lambda a - 1 - as) = 0 \implies s = \frac{\lambda a - 1}{a}.$$

Hence, $\Pr(X > a) \le \frac{\lambda}{(\lambda - \frac{\lambda a - 1}{a})e^{a\lambda - 1}} = \frac{\lambda a}{e^{\lambda a - 1}}$.

3 Terms and Conditionings (7 points)

Suppose $X \sim \min\{q, Y\}$, where $Y \sim Geom(p)$ is a geometric RV with parameter p and q is a positive integer. Calculate $\mathbb{E}[X]$ in terms of p and q.

We have that

$$\begin{aligned}
\mathbb{E}[X] &= q \Pr[Y > q] + \mathbb{E}[Y \mid Y \le q] \Pr(Y \le q) \\
&= q(1-p)^q + \mathbb{E}[Y] - \mathbb{E}[Y|Y > q] \Pr[Y > q] \\
&= q(1-p)^q + \mathbb{E}[Y] - (q + \mathbb{E}[Y])(1-p)^q \qquad \text{(Memoryless property of Geometric)} \\
&= \mathbb{E}[Y] - \mathbb{E}[Y](1-p)^q = \frac{1}{p} - \frac{(1-p)^q}{p} = \frac{1-(1-p)^q}{p}.
\end{aligned}$$

4 Homework Party (7 points)

Michael is trying to budget time for his problem set. He knows that the number of problems on the problem set is Poisson(6) distributed, and the probability that he can solve any given problem is $\frac{1}{4}$, and is independent of all other problems. If he can solve the problem, the amount of time he spends on the problem is Exponential $\left(\frac{1}{6}\right)$ distributed; otherwise, the amount of time he spends on the problem is Exponential $\left(\frac{1}{10}\right)$ distributed. What is the expected amount of time that Michael will spend on the problem set?

Let X represent the amount of time Michael will spend on the problem set, let N represent the number of problems on the problem set, and let S represent the number of problems that Michael can solve. Similarly, let the parameter of the Poisson distribution be λ, the probability of a problem being solved be p, the parameter of the first exponential be α, and the parameter of the second exponential be β.
We wish to compute $\mathbb{E}[X]$; by the law of iterated expectation, we have that

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|S]] = \mathbb{E}[\mathbb{E}[\mathbb{E}[(X|S)|N]]] = \mathbb{E}[\mathbb{E}[\mathbb{E}[X|S, N]]].$$

Given S and N, the expected amount of time that Michael will spend is $\frac{S}{\alpha} + \frac{N-S}{\beta} = S\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) + \frac{N}{\beta}$. Thus, we wish to compute

$$\mathbb{E}\left[\mathbb{E}\left[S\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) + \frac{N}{\beta}\right]\right] = \mathbb{E}\left[\mathbb{E}[S]\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) + \mathbb{E}[N]\frac{1}{\beta}\right].$$

We know that $\mathbb{E}[N]$ is λ. We compute $\mathbb{E}[S]$ again using iterated expectation:

$$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[S|N]] = \mathbb{E}[pN] = p\lambda,$$

since $S|N$ is a Binomial(N, p) distribution. Thus,

$$\mathbb{E}\left[\mathbb{E}[S]\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) + \mathbb{E}[N]\frac{1}{\beta}\right] = \left(\frac{1}{\alpha} - \frac{1}{\beta}\right)p\lambda + \frac{\lambda}{\beta} = \lambda\left(\frac{p}{\alpha} + \frac{1-p}{\beta}\right).$$

Taking $\lambda = 6, p = \frac{1}{4}, \alpha = \frac{1}{6}$, and $\beta = \frac{1}{10}$ gives us that $\mathbb{E}[X] = 54$.

Alternate Solution:

There are other ways to reason about this – the simplest is probably to note that by poisson splitting, the number of questions Michael can solve is $N_1 = \text{Poisson}(3/2)$ and the number of questions he cannot solve is $N_2 = \text{Poisson}(9/2)$. Denote X_1 as the amount of time he spends on questions he can solve, and X_2 the amount of time he spends on problems he cannot solve. Then we have that

$$\mathbb{E}[X] = \mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = \mathbb{E}[\mathbb{E}[X_1|N_1]] + \mathbb{E}[\mathbb{E}[X_2|N_2]] = 6 \cdot \frac{3}{2} + 10 \cdot \frac{9}{2} = 54.$$

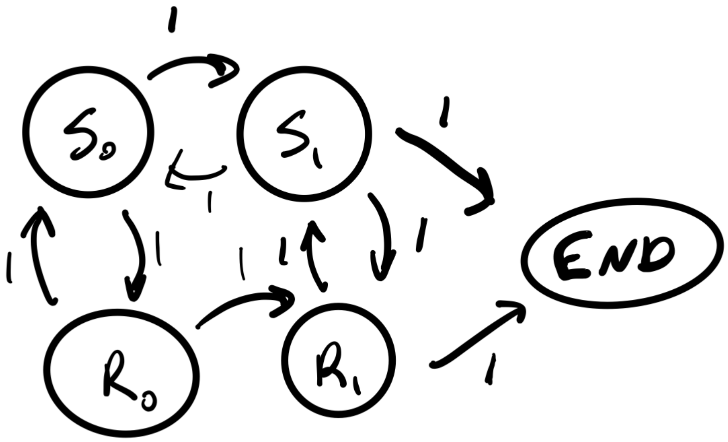
5 Seasonal Depression (8 points)

Kevin is a store owner in a part of town with strange weather. The weather alternates between two states: sunny and rainy. The length of each sunny period and rainy period is determined by an independent exponential with rate 1.

- If the weather is sunny, customers arrive at a store according to a Poisson process with rate 1 and each one leaves with rate 1.
- If the weather is rainy, customers arrive at a store according to a Poisson process with rate 1 and do not leave.

Suppose it is currently sunny and there are no customers in the store. What is the expected amount of time before there are at least 2 people in the store?

The key to this problem here is that we can set this up as a markov chain with 4 non-terminal states. There will be one state for each combination or sunny/rainy and number of customers. State S_i represents sunny with i customers, and R_i represents rainy with i customers.



We set up the hitting time equations, where x_T represents the expected time to reach END starting from state T.

$$\begin{aligned} x_{S_0} &= \frac{1}{2} + \frac{1}{2}x_{S_1} + \frac{1}{2}x_{R_0} \\ x_{S_1} &= \frac{1}{3} + \frac{1}{3}x_{R_1} + \frac{1}{3}x_{S_0} \\ x_{R_0} &= \frac{1}{2} + \frac{1}{2}x_{R_1} + \frac{1}{2}x_{S_0} \\ x_{R_1} &= \frac{1}{2} + \frac{1}{2}x_{S_1}. \end{aligned}$$

We solve and get the solution

$$x_{S_0} = \frac{5}{2}.$$

6 Geometric Perspective on Variance (3 + 5 + 2 points)

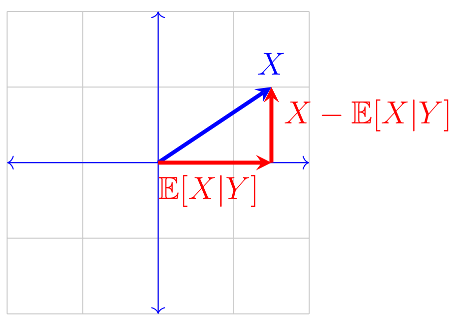
In this problem, we prove the identity $\mathrm{var}(X) = \mathbb{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbb{E}[X|Y])$. Assume X and Y are zero mean. You may use that $\mathrm{var}(X|Y) = \mathbb{E}[(X - \mathbb{E}[X|Y])^2|Y]$.

(a) First, show that $\mathbb{E}[\mathrm{var}(X|Y)] = \mathrm{var}(X - \mathbb{E}[X|Y])$.

$$\begin{aligned} \mathbb{E}[\mathrm{var}(X|Y)] &= \mathbb{E}[\mathbb{E}[(X - \mathbb{E}[[X|Y])^2|Y]] \\ &= \mathbb{E}[(X - \mathbb{E}[X|Y])^2] \\ &= \mathrm{var}(X - \mathbb{E}[X|Y]). \end{aligned}$$

The last inequality holds because $X - \mathbb{E}[X|Y]$ is zero mean.

(b) Draw X, $\mathbb{E}[X|Y]$, and $X - \mathbb{E}[X|Y]$ in the Hilbert space of random variables. Specify and justify the angle between $\mathbb{E}[X|Y]$ and $X - \mathbb{E}[X|Y]$.



The angle between $X - \mathbb{E}[X|Y]$ and $\mathbb{E}[X|Y]$ is $\pi/2 = 90$ degrees due to the orthogonality of the MMSE.

(c) Use the previous two parts to conclude the identity $\mathrm{var}(X) = \mathbb{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbb{E}[X|Y])$.

This is the Pythagorean theorem.

7 Coldplay (2 + 4 + 2 + 2 points)

You have a funny clock, which displays a number $x \in \{1, ..., 12\}$ at any given time, that behaves differently than a normal clock. First, it only shows the current hour. For every hour $x \in \{1, ..., 11\}$, the clock behaves normally, meaning $x \mapsto x + 1$ with probability 1. However, upon reaching 12 (AM or PM), the next hour it shows is 7 with probability $p \in (0, 1)$, or 1 otherwise (i.e. it will read a sequence (11, 12, 7, 8, 9, ...) or (11, 12, 1, 2, 3, ...), respectively).

(a) Suppose the clock currently reads 10. What is the expected time until the clock reads 3?

This can be seen to be $5 + 6 \cdot \mathbb{E}[\text{Geometric}(1 - p) - 1] = \frac{6}{1-p} - 1$.

(b) Compute the stationary distribution π. (*Hint: think about the relationship between the stationary distribution in a state and expected return time to a state*).

By symmetry, $\pi_{12} = \pi_7 = ... = \pi_{11}$ and $\pi_1 = ... = \pi_6$. We have that

$$\pi_{12} = \frac{1}{\mathbb{E}_{12}[T_{12}^{+}]} = \frac{1}{6p + 12(1-p)} = \frac{1}{12 - 6p}.$$

Then, we must that have $\pi_1 = ... = \pi_6 = (1 - 6\pi_{12})/6 = \frac{1-p}{12-6p}$.

(c) Suppose, infinitely long ago, your great$^{\infty}$-grandparents initialized the clock according to the initial distribution ψ. Can you use the stationary distribution from the previous part to say what the probability that the clock is currently in state 3 is? Justify your answer.

No. The chain has period 6, so we are not guaranteed to converge.

(d) Draw or describe a continuous time Markov chain with the same stationary distribution.

One example could be $Q(i, i+1) = 2$ for $i = 1, ..., 11$, and then $Q(12, 1) = Q(12, 6) = 1$.

8 Hidden Markovs Among Us (3 + 3 + 3 + 2 + 2 points)

Let the number of people infected by COVID on day n be denoted by X_n. Each day, X_n increases by 1 with probability $\frac{2}{3}$ or decreases by 1 with probability $\frac{1}{3}$. If $X_n = 0$, it stays the same with probability $\frac{1}{3}$ or increases with probability $\frac{2}{3}$. Then, let $Y_n \sim \text{Binomial}(X_n, \frac{3}{4})$ represent the number of people who test positive for COVID, i.e. that we report having COVID. Assume $X_0 = 1$.

(a) What is the MAP estimate of X_2 given that we observe $Y_2 = 1$?

There are three possible ways we can observe $Y_2 = 1$. Enumerating them and then evaluating using conditional probability:

(1) $X_1 = 0, X_2 = 1, Y_2 = 1$, which happens with probability $\frac{1}{3} \cdot \frac{2}{3} \cdot \frac{3}{4} = \frac{1}{6}$

(2) $X_1 = 2, X_2 = 1, Y_2 = 1$, which happens with probability $\frac{2}{3} \cdot \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{6}$

(3) $X_1 = 2, X_2 = 3, Y_2 = 1$, which happens with probability $\frac{2}{3} \cdot \frac{2}{3} \cdot 3 \cdot (1 - \frac{3}{4})^2 \cdot \frac{3}{4} = \frac{1}{16}$

Normalizing, we get $P(X_2 = 1 | Y_2 = 1) = \frac{\frac{1}{6} + \frac{1}{6}}{\frac{1}{6} + \frac{1}{6} + \frac{1}{16}} = \frac{16}{19}$. Our MAP estimate is $\hat{X}_2 = 1$.

(b) What is the MLE estimate of X_2 given $Y_2 = 1$? Multiple values are fine.

This is the same as question 1e from Midterm 1. The pmf of the binomial given $Y_2 = 1$ is proportional to $X_2 \cdot (\frac{1}{4})^{X_2}$. Checking the first couple values of X_2:

(1) $X_2 = 1$: .25

(2) $X_2 = 2$: .125

(3) $X_2 = 3$: .057

So our MLE estimate is also 1.

(c) What is the LLSE estimator of X_2 given $Y_2 = 1$? (You may use $\text{Var}(Y_2) \approx 1.056$)

First we solve for the covariance to apply the LLSE formula:

$$\begin{aligned}
\mathrm{Cov}(X_2, Y_2) &= \mathbb{E}[X_2 Y_2] - \mathbb{E}[X_2]\,\mathbb{E}[Y_2] \\
&= \mathbb{E}[X_2\,\mathbb{E}[Y_2|X_2]] - \frac{3}{4}\,\mathbb{E}[X_2]^2 \\
&= \mathbb{E}[X_2 \frac{3}{4} X_2] - \frac{3}{4}\,\mathbb{E}[X_2]^2 \\
&= \frac{3}{4}\mathrm{Var}(X_2)
\end{aligned}$$

We calculate $P(X_2 = 3) = \frac{4}{9}, P(X_2 = 1) = \frac{4}{9}, P(X_2 = 0) = \frac{1}{9}$. This gives us $\mathbb{E}[X_2] = \frac{16}{9}$ and $\mathrm{Var}(X_2) \approx 1.2839$.

As a result our LLSE is $L[X_2|Y_2] \approx .9117Y_1 + .5627$. For Y_2, this is equal to 1.4744.

(d) What is the MMSE estimator of X_2 given $Y_2 = 1$?

We can reuse the probabilities from part (a). We get $\hat{X}_2 = \frac{3}{19} \cdot 3 + \frac{16}{19} \cdot 1 = \frac{25}{19} \approx 1.3158$.

(e) Is the Markov chain $\{X_t\}_{t=0}^{\infty}$ positive recurrent, null recurrent, or transient?

Transient. Ignoring the self-loop, it is the birth-death chain.

9 Dungeons and Dragons (7 points)

In a game of Dungeons and Dragons, Aditya suspects Catherine is using a loaded 20-sided die. However, he doesn't want to risk falsely accusing her, so he conducts a hypothesis test to upper-bound the probability of a false alarm. Suppose if $X = 0$, the die is fair and a roll Y is distributed according to $\Pr(Y = y|X = 0) = 0.05$ for $1 \le y \le 20$. If $X = 1$, the die is loaded, and rolls have the distribution

$$\Pr(Y = y|X = 1) = \begin{cases} 0.025 & 1 \le y \le 10 \\ 0.075 & 11 \le y \le 20 \end{cases}$$

(or more simply, the probability of being greater than 10 is three times the probability of being less than or equal to 10). Construct a Neyman-Pearson decision rule to maximize the probability Aditya is correct if he accuses Catherine of cheating, while constraining the probability Aditya falsely accuses Catherine to be ≤ 0.05.

Notice that the cases 1-10 and 11-20 have equivalent probabilities in either case, so it's equivalent to consider the indicator random variable $Z = \mathbb{1}\{Y > 10\}$, i.e. $Z = 1$ if $11 \le Y \le 20$ and $Z = 0$ otherwise. The likelihood ratio in terms of Z is

$$L(z) = \frac{\Pr(Z=z|X=1)}{\Pr(Z=z|X=0)} = \begin{cases} \frac{1}{2} & z=0 \\ \frac{3}{2} & z=1 \end{cases}$$

This is monotonically increasing (it has to be as there are only two elements), so next we check whether we need to reduce the false-positive or false-negative probability. The "naive" PFA is $\Pr(Z=1|X=0) = 0.5$, so the test needs to reduce it, i.e. the optimal decision rule has the form

$$r(z) = \begin{cases} 0 & z=0 \\ 1 \text{ w.p.} \gamma & z=1 \end{cases}$$

To set γ we look at the PFA under the decision rule:

$$\begin{aligned} \Pr(r(Z)=1|X=0) &= 0.05 \\ \Pr(r(Z)=1|Z=1)\Pr(Z=1|X=0) &= 0.05 \\ \gamma \cdot 0.5 &= 0.05. \end{aligned}$$

Therefore $\gamma = 0.1$, so the overall decision rule is

$$r(y) = \begin{cases} 0 & 1 \le y \le 10 \\ 1 \text{ w.p.} \frac{1}{10} & 11 \le y \le 20 \end{cases}$$

10 Delayed Kalman Filter (6 + 2 points)

Consider a process with the transition rule $x_{n+1} = ax_n + v_n$ where $v_n \sim \mathcal{N}(0, \sigma_v^2)$. We can only observe the process at even-numbered times, i.e. we see $y_{2n} = x_{2n} + w_{2n}$, where $w_n \sim \mathcal{N}(0, \sigma_w^2)$.

1. Find a recurrence relation for the MMSE of the even states $\hat{x}_{2n} = \mathbb{E}[x_{2n}|y_0, y_2, \ldots, y_{2n}]$ in terms of $\hat{x}_{2n-2}$.

We can describe this as a normal Kalman filtering problem with a different transition rule:

$$x_{2n} = ax_{2n-1} + v_{2n-1} = a(ax_{2n-2} + v_{2n-2}) + v_{2n-1} = a^2 x_{2n-2} + (av_{2n-2} + v_{2n-1}),$$

i.e. transitions with a^2 in place of a and with noise from a $\mathcal{N}(0, (a^2+1)\sigma_v^2)$ distribution. The recurrence for the even state MMSE is therefore

$$\hat{x}_{2n} = a^2\hat{x}_{2n-2} + k_{2n}(y_{2n} - a^2\hat{x}_{2n-2}),$$

where the Kalman gain k_{2n} is given by

$$k_{2n} = \frac{a^4\sigma^2_{2n|2n} + (a^2+1)\sigma_v^2}{a^4\sigma^2_{2n|2n} + (a^2+1)\sigma_v^2 + \sigma_w^2}$$

and the estimator variance obeys the recurrence

$$\sigma^2_{2n+2|2n+2} = (1-k_n)a^4\sigma^2_{2n|2n}.$$

2. Find a recurrence relation for the MMSE of the odd states $\hat{x}_{2n+1} = \mathbb{E}[x_{2n+1}|y_0, y_2, \ldots, y_{2n}]$ in terms of $\hat{x}_{2n}$.

$$\hat{x_{2n+1}} = a_{2n}\hat{x_{2n}}$$

11 Oh Yeahhhhh (2 + 5 points)

Suppose an infinitely large bucket is being filled with kool-aid continuously with rate 1 Liters/min. Two bartenders serve drinks according to independent Poisson Processes with rates 2 drinks/min (Bartender A) and 3 drinks/min (Bartender B). Whenever they serve a drink, they empty the shared bucket into the glass and serve that. Say this process started infinitely in the past.

1. Suppose you come in at a random time, cut to the front of the line, and take the next drink that is served by either bartender. What is the expected volume of the drink you get?

This is RIP of the merged process, so the answer is 2/5.

2. Suppose you come in at a random time, again cut to the front, but insist on taking the next drink served by Bartender A. What is the expected volume of the drink you receive?

For bartender A, who is serving drinks at rate 2 drinks/min. Then let X be the amount of kool-aid you receive. We have

$$\mathbb{E}[X|\text{you take bartender A}] = \Pr[\text{Next arrival is bartender A}]\,\mathbb{E}[\text{RIP of merged process}] +$$

$$\Pr[\text{next arrival is bartender B}]\,\mathbb{E}[\text{interval length of merged PP}]$$
$$= \frac{2}{5}\frac{2}{5} + \frac{3}{5}\frac{1}{5} = \frac{7}{25}.$$