Final Exam Announcements

Final Exam: Closed-book, with three double-sided 8.5 x 11 formula sheets permitted. Please arrive early to find your seat before the prompt start at 9:00AM. Calculators are not allowed.

> *Date:* Wednesday, December 15
> *Time:* 9:00am - 12 noon
> *Location:* Walker Memorial Gymnasium (Building 50, third floor)
> *Content:* The Final will cover all the material from the current term, up to and including the material covered in the Wednesday (Dec 2nd) lecture, i.e. up to Section 9.1, with the exception of the material on confidence intervals based on the t-distribution (middle of p. 471 to the end of p. 473). However the emphasis will be on the material not covered in the first two quizzes (Chapters 5-9).

Practice Quizzes: Two past finals with full solutions are available on the OCW website (Spring05 & Spring06). An additional two finals have been posted on the course website (Spring 09 & Fall 09), which will be reviewed at the TA final review session. Please note that the material covered in the final or the course, and the course emphasis change each term. Hence past finals are not necessarily indicative of this term's final. Material presented in lecture, recitation, tutorial, and problem set exercises should be your primary source of preparation.
`http://ocw.mit.edu/OcwWeb/web/home/home/index.htm`
`http://stellar/S/course/6/fa10/6.041/materials.html`

Office Hours: Please check the course website as the final date approaches to find posted office hours pertaining to finals week.

Optional 6.041/6.431 Final Review Session: There will be a two-hour 6.041/6.431 final review session administered by two TAs. The session will consist of two parts. In the first hour, a concise overview of the theory will be presented. In the second hour, selected problems from past finals will be solved. Though completely optional, the final review is a great opportunity to reinforce your understanding of the material and perhaps gain new insight. Details for the quiz review:

> *Date:* Thursday, December 9
> *Time:* 7:30-9:30pm
> *Location:* 34-101

Problems for the final review will be selected from the 6.041 Spring 2009 and Fall 2009 Final exam (each available on the course website under `Final Exam`). We will review as many problems as time permits. Full solutions will be posted on-line following the review. We strongly recommend working through the problems before coming to the final review.

10/11
Reading

Quiz I Review
Probabilistic Systems Analysis
6.041/6.431

Massachusetts Institute of Technology

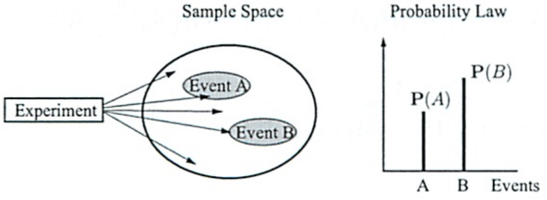
October 7, 2010

Quiz Information

- Closed-book with one double-sided 8.5 x 11 formula sheet allowed
- Date: Tuesday, October 12, 2010
- Time: 7:30 - 9:00 PM
- Location: 54-100
- Content: Chapters 1-2, Lecture 1-7, Recitations 1-7, Psets 1-4, Tutorials 1-3
- Show your reasoning when possible!

A Probabilistic Model:

- **Sample Space**: The set of all possible outcomes of an experiment.
- **Probability Law**: An assignment of a nonnegative number **P(E)** to each event E.



Probability Axioms

Given a sample space Ω:

1. **Nonnegativity:** $\mathbf{P}(A) \geq 0$ for each event A
2. **Additivity:** If A and B are disjoint events, then
$$\mathbf{P}(A \cup B) = P(A) + P(B)$$
If $A_1, A_2, \ldots,$ is a sequence of disjoint events,
$$\mathbf{P}(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$
3. **Normalization** $\mathbf{P}(\Omega) = 1$

$\cup$ = union or

$\cap$ = and

Properties of Probability Laws

part of

Given events A, B and C:

1. If $A \subset B$, then $\mathbf{P}(A) \leq \mathbf{P}(B)$
2. $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$
3. $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$
4. $\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(A^c \cap B) + \mathbf{P}(A^c \cap B^c \cap C)$

↑ draw out graphically if don't understand

Discrete Models

- **Discrete Probability Law:** If Ω is finite, then each event $A \subseteq \Omega$ can be expressed as

$$A = \{s_1, s_2, \ldots, s_n\} \qquad s_i \in \Omega$$

Therefore the probability of the event A is given as

$$\mathbf{P}(A) = \mathbf{P}(s_1) + \mathbf{P}(s_2) + \cdots + \mathbf{P}(s_n)$$

- **Discrete Uniform Probability Law:** If all outcomes are equally likely,

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|}$$

Conditional Probability

- Given an event B with $\mathbf{P}(B) > 0$, the conditional probability of an event $A \subseteq \Omega$ is given as

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

- $\mathbf{P}(A|B)$ is a valid probability law on Ω, satisfying
  1. $\mathbf{P}(A|B) \geq 0$
  2. $\mathbf{P}(\Omega|B) = 1$
  3. $\mathbf{P}(A_1 \cup A_2 \cup \cdots |B) = \mathbf{P}(A_1|B) + \mathbf{P}(A_2|B) + \cdots$, where $\{A_i\}_i$ is a set of disjoint events
- $\mathbf{P}(A|B)$ can also be viewed as a probability law on the restricted universe B.

I finally know what this symbol means

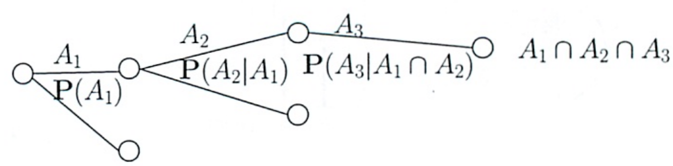
Multiplication Rule

- Let $A_1, \ldots, A_n$ be a set of events such that

$$\mathbf{P}\left(\bigcap_{i=1}^{n-1} A_i\right) > 0.$$

Then the joint probability of all events is

$$\mathbf{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1 \cap A_2) \cdots \mathbf{P}(A_n|\bigcap_{i=1}^{n-1} A_i)$$



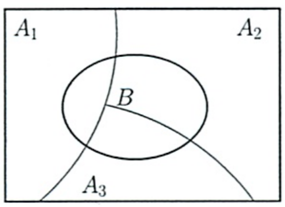
Oh - is this just multiply down the tree?

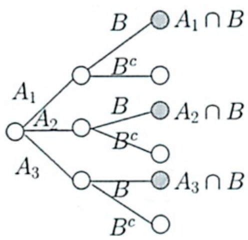
- makes much more sense now!

Total Probability Theorem

Let $A_1, \ldots, A_n$ be disjoint events that partition Ω. If $\mathbf{P}(A_i) > 0$ for each i, then for any event B,

$$\mathbf{P}(B) = \sum_{i=1}^{n} \mathbf{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbf{P}(B|A_i)\mathbf{P}(A_i)$$



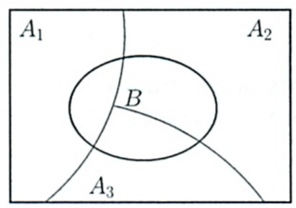


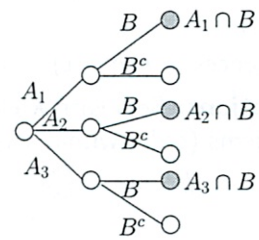
- we want P(B)
- so add up each section

Bayes Rule

Given a finite partition $A_1, \ldots, A_n$ of Ω with $\mathbf{P}(A_i) > 0$, then for each event B with $\mathbf{P}(B) > 0$

$$\mathbf{P}(A_i|B) = \frac{\mathbf{P}(B|A_i)\mathbf{P}(A_i)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B|A_i)\mathbf{P}(A_i)}{\sum_{i=1}^{n} \mathbf{P}(B|A_i)\mathbf{P}(A_i)}$$





Relationship b/w 2 conditional prob that are reverse of each other

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A|B)\,P(B) = P(A \cap B) = P(B|A)\,P(A)$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Independence of Events

- Events A and B are **independent** if and only if

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$$

  or

$$\mathbf{P}(A|B) = \mathbf{P}(A) \quad \text{if } \mathbf{P}(B) > 0$$

- Events A and B are **conditionally independent** given an event C if and only if

$$\mathbf{P}(A \cap B|C) = \mathbf{P}(A|C)\mathbf{P}(B|C)$$

  or

$$\mathbf{P}(A|B \cap C) = \mathbf{P}(A|C) \quad \text{if } \mathbf{P}(B \cap C) > 0$$

- Independence $\not\Leftrightarrow$ Conditional Independence.

Independence of a Set of Events

- The events $A_1, \ldots, A_n$ are **pairwise independent** if for each $i \neq j$

$$\mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j)$$

- The events $A_1, \ldots, A_n$ are **independent** if

$$\mathbf{P}\left(\overbrace{\bigcap_{i \in S}}^{\text{and}} A_i\right) = \overbrace{\prod_{i \in S}}^{\text{multiply}} \mathbf{P}(A_i) \quad \forall\, S \subseteq \{1, 2, \ldots, n\}$$

- Pairwise independence $\not\Rightarrow$ independence, but independence $\Rightarrow$ pairwise independence.

Counting Techniques

- **Basic Counting Principle:** For an m-stage process with n_i choices at stage i,
$$\# \text{ Choices} = n_1 n_2 \cdots n_m$$
- **Permutations:** k-length sequences drawn from n distinct items without replacement (order is important):
$$\# \text{ Sequences} = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$
- **Combinations:** Sets with k elements drawn from n distinct items (order within sets is not important):
$$\# \text{ Sets} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Counting Techniques-contd

- **Partitions:** The number of ways to partition an n-element set into r disjoint subsets, with n_k elements in the k^{th} subset:

$$\binom{n}{n_1, n_2, \ldots, n_r} = \binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\cdots-n_r-1}{n_r} = \frac{n!}{n_1! n_2!, \cdots, n_r!}$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$\sum_{i=1}^{r} n_i = n$$

Chap 2

Discrete Random Variables

- A **random variable** is a real-valued function defined on the sample space:
$$X : \Omega \to R$$
just some # outcome
- The notation $\{X = x\}$ denotes an event:
set of events
$$\{X = x\} = \{\omega \in \Omega | X(\omega) = x\} \subseteq \Omega$$
actual event
- The **probability mass function (PMF)** for the random variable X assigns a probability to each event $\{X = x\}$:
$$p_X(x) = \mathbf{P}(\{X = x\}) = \mathbf{P}(\{\omega \in \Omega | X(\omega) = x\})$$

never saw this - just more useless stuff I think

PMF Properties

- Let X be a random variable and S a countable subset of the real line
- The axioms of probability hold:
  1. $p_X(x) \geq 0$
  2. $\mathbf{P}(X \in S) = \sum_{x \in S} p_X(x)$
  3. $\sum_x p_X(x) = 1$
- If g is a real-valued function, then $Y = g(X)$ is a random variable:
$$\omega \xrightarrow{X} X(\omega) \xrightarrow{g} g(X(\omega)) = Y(\omega)$$
with PMF
$$p_Y(y) = \sum_{x | g(x) = y} P_X(x)$$

Expectation

Given a random variable X with PMF $p_X(x)$:

- $\mathbf{E}[X] = \sum_x x p_X(x)$
- Given a derived random variable $Y = g(X)$:

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x) = \sum_y y p_Y(y) = E[Y]$$

$$\mathbf{E}[X^n] = \sum_x x^n p_X(x)$$

- **Linearity** of Expectation: $\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$.

Variance

The expected value of a derived random variable $g(X)$ is

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x)$$

The variance of X is calculated as

- $var(X) = \mathbf{E}[(X - \mathbf{E}[X])^2] = \sum_x (x - \mathbf{E}[X])^2 p_X(x)$
- $var(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2$
- $var(aX + b) = a^2 var(X)$

Note that $var(x) \geq 0$

Multiple Random Variables

Let X and Y denote random variables defined on a sample space Ω.

- The **joint PMF** of X and Y is denoted by

$$p_{X,Y}(x, y) = \mathbf{P}(\{X = x\} \cap \{Y = y\})$$

- The **marginal PMFs** of X and Y are given respectively as

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$

Functions of Multiple Random Variables

Let $Z = g(X, Y)$ be a function of two random variables

- **PMF:**

$$p_Z(z) = \sum_{(x,y) | g(x,y) = z} p_{X,Y}(x, y)$$

what is this again?

- **Expectation:**

$$\mathbf{E}[Z] = \sum_{x,y} g(x, y) p_{X,Y}(x, y)$$

- **Linearity:** Suppose $g(X, Y) = aX + bY + c$.

$$\mathbf{E}[g(X, Y)] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c$$

Conditioned Random Variables

- Conditioning X on an event A with $\mathbf{P}(A) > 0$ results in the PMF:

$$p_{X|A}(x) = \mathbf{P}(\{X = x\}|A) = \frac{\mathbf{P}(\{X = x\} \cap A)}{\mathbf{P}(A)}$$

- Conditioning X on the event $Y = y$ with $\mathbf{P}_Y(y) > 0$ results in the PMF:

$$p_{X|Y}(x|y) = \frac{\mathbf{P}(\{X = x\} \cap \{Y = y\})}{\mathbf{P}(\{Y = y\})} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

Same as always
- and I do it graphically
- select only that row

Conditioned RV - contd

- Multiplication Rule: go along the branch

$$p_{X,Y}(x, y) = p_{X|Y}(x|y)p_Y(y)$$

- Total Probability Theorem:

$$p_X(x) = \sum_{i=1}^{n} \mathbf{P}(A_i)p_{X|A_i}(x)$$

$$p_X(x) = \sum_{y} p_{X|Y}(x|y)p_Y(y)$$

add up each part

Conditional Expectation

Let X and Y be random variables on a sample space Ω.

- Given an event A with $\mathbf{P}(A) > 0$

$$\mathbf{E}[X|A] = \sum_{x} x p_{X|A}(x)$$

- If $P_Y(y) > 0$, then

$$\mathbf{E}[X|\{Y = y\}] = \sum_{x} x p_{X|Y}(x|y)$$

- **Total Expectation Theorem:** Let $A_1, \ldots, A_n$ be a partition of Ω. If $\mathbf{P}(A_i) > 0\ \forall i$, then

$$\mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{P}(A_i)\mathbf{E}[X|A_i]$$

what is this again?
- just new universe
- and expectation

Independence

Let X and Y be random variables defined on Ω and let A be an event with $\mathbf{P}(A) > 0$.

- X is independent of A if either of the following hold:

$$p_{X|A}(x) = p_X(x)\ \forall x$$
$$p_{X,A}(x) = p_X(x)\mathbf{P}(A)\ \forall x$$

- X and Y are independent if either of the following hold:

$$p_{X|Y}(x|y) = p_X(x)\ \forall x \forall y$$
$$p_{X,Y}(x, y) = p_X(x)p_Y(y)\ \forall x \forall y$$

what is upside down $\forall$?
- universal quantification
- given any
- for all

Independence

If X and Y are independent, then the following hold:

- If g and h are real-valued functions, then $g(X)$ and $h(Y)$ are independent.
- $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$ (inverse is not true)
- $var(X + Y) = var(X) + var(Y)$

Given independent random variables $X_1, \ldots, X_n$,

$$var(X_1+X_2+\cdots+X_n) = var(X_1)+var(X_2)+\cdots+var(X_n)$$

Some Discrete Distributions

	X	$p_X(k)$	$\mathbf{E}[X]$	$var(X)$
Bernoulli	$\begin{cases} 1 & \text{success} \\ 0 & \text{failure} \end{cases}$	$\begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$	p	$p(1-p)$
Binomial	Number of successes in n Bernoulli trials	$\binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \ldots, n$	np	np(1-p)
Geometric	Number of trials until first success	$(1-p)^{k-1} p$, $k = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Uniform	An integer in the interval [a,b]	$\begin{cases} \frac{1}{b-a+1} & k = a, \ldots, b \\ 0 & \text{otherwise} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)(b-a+2)}{12}$

thankfully we don't need to do proofs
and I understood most of the HW
will hopefully do better

6.041/6.431 Fall 2010 Quiz 1
Tuesday, October 12, 7:30 - 9:00 PM.

DO NOT TURN THIS PAGE OVER UNTIL YOU ARE TOLD TO DO SO

Question	Score	Out of	
1.1	10	10	
1.2	3	10	−7
1.3	6 pH	10	−4
1.4	10 pH	10	
1.5	5	5	
1.6	4	10	−1
1.7	10	10	
1.8	10	10	
2.1	1	10	−9
2.2	0	10	−10
2.3	7	10	−3
Your Grade	66	105	

Name: Michael Plasmeier

Recitation Instructor: Dimitri

TA: Aliaa 3PM

- This quiz has 2 problems, worth a total of 105 points.
- You may tear apart pages 3, 4 and 5, as per your convenience, **but you must turn them in together with the rest of the booklet.**
- Write your solutions in this quiz booklet, only solutions in this quiz booklet will be graded. Be neat! You will not get credit if we can't read it.
- You are allowed one two-sided, handwritten, 8.5 by 11 formula sheet. Calculators are not allowed.
- **You may give an answer in the form of an arithmetic expression (sums, products, ratios, factorials) of numbers that could be evaluated using a calculator.** Expressions like $\binom{8}{3}$ or $\sum_{k=0}^{5}(1/2)^k$ are also fine.
- You have 90 minutes to complete the quiz.
- Graded quizzes will be returned in recitation on Thursday 10/14.

Problem 0: (0 points) Write your name, your assigned recitation instructor's name, and assigned TA's name on the cover of the quiz booklet. The Instructor/TA pairing is listed below.

Recitation Instructor	TA	Recitation Time
Vivek Goyal	Uzoma Orji	10 & 11 AM
Peter Hagelstein	Ahmad Zamanian	12 & 1 PM
Ali Shoeb	Shashank Dwivedi	2 PM
Dimitri Bertsekas (6.431)	Aliaa Atwi	2 & 3 PM

Summary of Results for Special Random Variables

Discrete Uniform over $[a, b]$:

$$p_X(k) = \begin{cases} \dfrac{1}{b-a+1}, & \text{if } k = a, a+1, \ldots, b, \\ 0, & \text{otherwise,} \end{cases}$$

$$\mathbf{E}[X] = \frac{a+b}{2}, \qquad \text{var}(X) = \frac{(b-a)(b-a+2)}{12}.$$

Bernoulli with Parameter p: (Describes the success or failure in a single trial.)

$$p_X(k) = \begin{cases} p, & \text{if } k = 1, \\ 1-p, & \text{if } k = 0, \end{cases}$$

$$\mathbf{E}[X] = p, \qquad \text{var}(X) = p(1-p).$$

Binomial with Parameters p and n: (Describes the number of successes in n independent Bernoulli trials.)

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n,$$

$$\mathbf{E}[X] = np, \qquad \text{var}(X) = np(1-p).$$

Geometric with Parameter p: (Describes the number of trials until the first success, in a sequence of independent Bernoulli trials.)

$$p_X(k) = (1-p)^{k-1} p, \qquad k = 1, 2, \ldots,$$

$$\mathbf{E}[X] = \frac{1}{p}, \qquad \text{var}(X) = \frac{1-p}{p^2}.$$

(did not use ?
- bad ? ? ?

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2010)

Problem 1: (75 points)

Note: All parts can be done independently, with the exception of the last part. Just in case you made a mistake in the previous part, you can use a symbol for the expression you found there, and use that symbol in the formulas for the last part.

Note: **Algebraic or numerical expressions do not need to be simplified in your answers.**

Jon and Stephen cannot help but think about their commutes using probabilistic modeling. Both of the them start promptly at 8am.

Stephen drives and thus is at the mercy of traffic lights. When all traffic lights on his route are green, the entire trip takes 18 minutes. Stephen's route includes 5 traffic lights, each of which is red with probability 1/3, independent of every other light. Each red traffic light that he encounters adds 1 minute to his commute (for slowing, stopping, and returning to speed).

1. **(10 points)** Find the PMF, expectation, and variance of the length (in minutes) of Stephen's commute.
2. **(10 points)** Given that Stephen's commute took him at most 19 minutes, what is the expected number of red lights that he encountered?
3. **(10 points)** Given that the last red light encountered by Stephen was the fourth light, what is the conditional variance of the total number of red lights he encountered?
4. **(10 points)** Given that Stephen encountered a total of three red lights, what is the probability that exactly two out of the first three lights were red?

Jon's commuting behavior is rather simple to model. Jon walks a total of 20 minutes from his home to a station and from a station to his office. He also waits for X minutes for a subway train, where X has the discrete uniform distribution on $\{0, 1, 2, 3\}$. (All four values are equally likely, and independent of the traffic lights encountered by Stephen.)

5. **(5 points)** What is the PMF of the length of Jon's commute in minutes?
6. **(10 points)** Given that there was exactly one person arriving at **exactly** 8:20am, what is the probability that this person was Jon?
7. **(10 points)** What is the probability that Stephen's commute takes at most as long as Jon's commute?
8. **(10 points)** Given that Stephen's commute took at most as long as Jon's, what is the conditional probability that Jon waited 3 minutes for his train?

Problem 2. (30 points) For each one of the statements below, give either a proof or a counterexample showing that the statement is not always true.

1. **(10 points)** If events A and B are independent, then the events A and B^c are also independent.
2. **(10 points)** Let A, B, and C be events associated with a common probabilistic model, and assume that $0 < \mathbf{P}(C) < 1$. Suppose that A and B are conditionally independent given C. Then, A and B are conditionally independent given C^c.

3. **(10 points)** Let X and Y be independent random variables. Then, $\mathrm{var}(X+Y) \geq \mathrm{var}(X)$.

Each question is repeated in the following pages. Please write your answer on the appropriate page.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

Problem 1: (75 points)

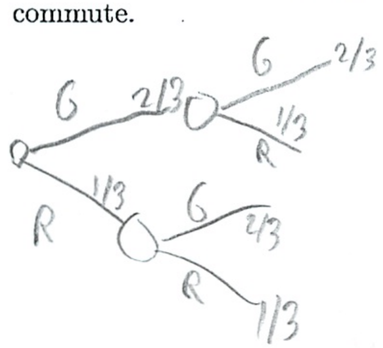
Note: All parts can be done independently, with the exception of the last part. Just in case you made a mistake in the previous part, you can use a symbol for the expression you found there, and use that symbol in the formulas for the last part.

Note: **Algebraic or numerical expressions do not need to be simplified in your answers.**

Jon and Stephen cannot help but think about their commutes using probabilistic modeling. Both of the them start promptly at 8am.

Stephen drives and thus is at the mercy of traffic lights. When all traffic lights on his route are green, the entire trip takes 18 minutes. Stephen's route includes 5 traffic lights, each of which is red with probability 1/3, independent of every other light. Each red traffic light that he encounters adds 1 minute to his commute (for slowing, stopping, and returning to speed).

1. **(10 points)** Find the PMF, expectation, and variance of the length (in minutes) of Stephen's commute.



$G_1 \cap G_2 \cap G_3 \cap G_4 \cap G_5 = 18$ min

$R_1 \cap R_2 \cap R_3 \cap R_4 \cap R_5 = 18 + 5$ min

Order does not matter

Subtract out 18 min

0 min → all green

each red 1 min

X = # Red lights

Y = minutes of commute

$P(X=0) = \left(\frac{2}{3}\right)^5$ | Y: 18

$P(X=1) = \left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^4\binom{5}{1} = 5\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^4$ | 19

$P(X=2) = \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3$ | 20

$P(X=3) = \binom{5}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2$ | 21

+5

$P(X=4) = \binom{5}{4}\left(\frac{1}{3}\right)^4\frac{2}{3}$ | 22

$P(X=5) = \binom{5}{5}\left(\frac{1}{3}\right)^5 = \left(\frac{1}{3}\right)^5$ | 23

→ over

+2

$$E[Y] = 18\cdot\left(\frac{2}{3}\right)^5 + 19\cdot 5\cdot\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right) + 20\cdot\binom{5}{2}\cdot\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2 +$$
$$21\cdot\binom{5}{3}\cdot\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^3 + 22\cdot\binom{5}{4}\cdot\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^4 + 23\cdot\left(\frac{1}{3}\right)^5$$

$$E[Y^2] = 18^2\left(\frac{2}{3}\right)^5 + 19^2\cdot 5\cdot\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right) + 20^2\cdot\binom{5}{2}\cdot\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2 +$$
$$21^2\cdot\binom{5}{3}\cdot\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^3 + 22^2\cdot\binom{5}{4}\cdot\frac{2}{3}\left(\frac{1}{3}\right)^4 + 23^2\cdot\left(\frac{1}{3}\right)^5$$

$$\mathrm{Var}(Y) = E[Y^2] - E[Y]^2 \quad +3$$

and no I am not rewriting that

See solutions. Much easier way

2. **(10 points)** Given that Stephen's commute took him at most 19 minutes, what is the expected number of red lights that he encountered?

at most 19 minutes is either 0 or 1 red light
18   19 min

$E[X \mid \text{commute at most 19 min}] = \cancel{0 \cdot \left(\frac{2}{3}\right)^5 + 1 \cdot 5 \cdot \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)}$

So all we care about is that 1 light (normalize)

$0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3}$

$\frac{1}{3}$ lights

where do there come from? 3/10

yes you do - its conditional $P(B|A) = \frac{P(A \cap B)}{P(A)}$

We don't care about prob that trip was ≤ 19 min

correct

3. **(10 points)** Given that the last red light encountered by Stephen was the fourth light, what is the conditional variance of the total number of red lights he encountered?

$E[X \mid \text{last red was 4th light}]$

- don't know how many min

know 4th = red
5th = green

don't know 1st 3 lights

$0 + 1 \cdot \frac{1}{3} \cdot \left(\frac{2}{3}\right)^4 + 2 \cdot \frac{1}{3}^2 \cdot \left(\frac{2}{3}\right)^3 \cdot \binom{3}{1} + 3 \cdot \binom{3}{2}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2$

impossible anyway and worth 0 pts

all other lights green

1 of first 3 lights green

incorrect PMF approach for E ok
no variance

$+ 4 \binom{3}{3}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 + 5 \cdot 0$

all 3 other lights red

impossible - 5th light green

$\frac{4}{10}$

$= \mathrm{var}(R_1 + R_2 + R_3)$
$= 3\,\mathrm{var}(R_i)$
$= \frac{6}{9}$

not doing formula based stuff problem

divide something out? what's total commute?

? divide anything out?

- ? is E[ ]
- total prob of all lights is 1
- just going to leave it

4. **(10 points)** Given that Stephen encountered a total of three red lights, what is the probability that exactly two out of the first three lights were red?

$P[\underbrace{2 \text{ out of 1st 3 lights red}}_{A} \mid \underbrace{X=3}_{B}]$ $\quad P(A|B) = \frac{P(A\cap B)}{P(B)}$

R _ R R — n 3 red lights can be anywhere

prob they are here

- not independent.

$P(A) = \binom{3}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)$

$P(B) = \binom{5}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2$

$P(A\cap B) = \binom{3}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)\cdot\underbrace{\binom{2}{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)}_{\text{one more in last 2 pos}}$

$$= \frac{\binom{3}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)\cdot\binom{2}{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)}{\binom{5}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2}$$

$$= \frac{\binom{3}{2}\binom{2}{1}}{\binom{5}{3}} \quad \frac{10}{10} = \frac{\frac{3!}{2!1!}\cdot\frac{2!}{1!1!}}{5!/3!2!}$$

should have reviewed that p-set a[c] more!

Jon's commuting behavior is rather simple to model. Jon walks a total of 20 minutes from his home to a station and from a station to his office. He also waits for X minutes for a subway train, where X has the discrete uniform distribution on $\{0, 1, 2, 3\}$. (All four values are equally likely, and independent of the traffic lights encountered by Stephen.)

No time on train

5. **(5 points)** What is the PMF of the length of Jon's commute in minutes?

X = time waiting for train
Y = commute time

$P(X=0) = \frac{1}{4}$ $\quad Y = 20+0 = 20$
$P(X=1) = \frac{1}{4}$ $\quad Y = 20+1 = 21$
$P(X=2) = \frac{1}{4}$ $\quad Y = 20+2 = 22$
$P(X=3) = \frac{1}{4}$ $\quad Y = 20+3 = 23$

notice how they wrote

$$P_{T_j}(l) = \begin{cases} \frac{1}{4} & \text{if } l = \{20, 21, 22, 23\} \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{3!}{\frac{5!}{3!\,2!}} = \frac{3!}{1} \cdot \frac{3!\,2!}{5!}$$

$$\frac{3! \cdot 3! \cdot 2!}{5!}$$

6. **(10 points)** Given that there was exactly one person arriving at **exactly** 8:20am, what is the probability that this person was Jon?

either Jon or Steve arrives at 8:20

P(John arrives at 8:20 | ~~Steve does not arrive 8:20~~ one person exactly at 8:20)

← can't jump to assume

$B^c = P(\text{Steve arrives } 8{:}20) = \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3$ ✓

$B = P(\text{Steve does } \underline{\text{not}} \text{ arrive } 8{:}20) = 1 - \left(\binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3\right)$

– use the beauty of $P(A|B) = \frac{P(A\cap B)}{P(B)}$

$A = P(\text{John arrives } 8{:}20) = \frac{1}{4}$

$P(A|B) = \frac{P(A\cap B)}{P(B)}$ ✓

A + B independent here

$P(A\cap B) = P(A)\,P(B)$

but this is not what you're asked to calculate

$$= \frac{\left(\frac{1}{4}\right)\left(1 - \cancel{\left(\binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3\right)}\right)}{\cancel{1 - \left(\binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3\right)}}$$ ✓

→ NO!

$= \left(\frac{1}{4}\right)$ ← prob john arrives then

right if A + B independent

$P(A|B) = P(A)$

4

~~Steve in 20 min~~ =
need a different denominator

$P\{T_J = 20\} \cap P\{T_S \neq 20\} + P\{T_J \neq 20\} \cap P\{T_S = 20\}$

Steve = drives
John = subway

7. **(10 points)** What is the probability that Stephen's commute takes at most as long as Jon's commute?

add up all of the possibilities Steve shorter or same as John

~~joint PMF ??~~ independent

total prob. Theorm?

P(Steve's commute shorter than 20 min) P(John's commute is 20)

P(John's $Y \le 20$)

P(John's $Y = 18, 19, 20$

+ P(Steve's commute is 18, 19, 20, 21) P(John's commute = 21) +

P(Steve's commute = 18, 19, 20, 21, 22) P(John's commute = 22) +

P(Steve's ≤ 23) P(John's = 23)

$=1$

$$= \left[\left(\tfrac{2}{3}\right)^5 + 5\left(\tfrac{1}{3}\right)\left(\tfrac{2}{3}\right)^4 + \binom{5}{2}\left(\tfrac{1}{3}\right)^2\left(\tfrac{2}{3}\right)^3\right] \cdot \frac{1}{4} +$$

$$\left[\;''\quad ''\quad '' \quad + \binom{5}{3}\left(\tfrac{1}{3}\right)^3\left(\tfrac{2}{3}\right)^2\right] \cdot \frac{1}{4} +$$

$$\left[\;''\quad ''\quad ''\quad '' \quad + \binom{5}{2}\left(\tfrac{1}{3}\right)^4\left(\tfrac{2}{3}\right)\right] \cdot \frac{1}{4} +$$

$$\left[\;1\;\right] \cdot \frac{1}{4}$$

John = Jon

8. **(10 points)** Given that Stephen's commute took at most as long as Jon's, what is the conditional probability that Jon waited 3 minutes for his train?

$P(A|B) = \frac{P(A \cap B)}{P(B)}$

B = Steven's commute to at most as long as Jon's - answer to previous problem

A = John waiting 3 min for train

(A, B not nessarlly independent, right?)

$P(A) = P(J\text{'s } x = 3) = \frac{1}{4}$ or $Y = 23$

~~[illegible]~~ do Bayes's Rule

$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$

$P(B|A) = 1$ since if John took 23 minutes, than Steve's commute can be anything (Ω)

$P(A|B) = \frac{1 \cdot \frac{1}{4}}{\text{that horrible thing from 1 pg back}}$

P(B) = answer to previous problem

10

Problem 2. (30 points) For each one of the statements below, give either a proof or a counterexample showing that the statement is not always true.

Show that its true or false

1. **(10 points)** If events A and B are independent, then the events A and B^c are also independent.

$B^c = 1 - B < \Omega - B$

So B and B^c have the same potential for Over lap - overlapping a little but not too much

1

$B \cap B^c = \Omega$

bad, I know

If A and B would be disjoint (non independent) than $A \cap B^c$ would completly overlap (also non independent)

independent | knowing any information about B would not help you w/ A
B^c ''

→ Since you could convert $B_c \to B$ and still would not have info

2. **(10 points)** Let A, B, and C be events associated with a common probabilistic model, and assume that $0 < \mathbf{P}(C) < 1$. Suppose that A and B are conditionally independent given C. Then, A and B are conditionally independent given C^c.

$P(A|C) \cdot P(B|C) = P(A \cap B|C)$

means

$P(A \cap B|C^c) = P(A|C^c)P(B|C^c)$

0

Same as above where

If A + B are independent than A and B^c are also independent

The conditional independence brings you to a new universe - inside general probability laws apply

Or did not really do proofs

did not study for this -proofs

? as always ? what to write

(Additional space for Problem 2.2)

3. **(10 points)** Let X and Y be independent random variables. Then, $\mathrm{var}(X+Y) \geq \mathrm{var}(X)$.

~~Var (x+y) = Var(x) + Var(y)~~

~~Y ≥ 0 so var(Y) ≥ 0 ← as a random variable~~

~~If you subtract var (X) from both sides~~

~~Var(x) + var(y) ≥ var(x)~~

~~var(y) ≥ 0~~

Var(x) = E[(X − E[x])²]
= E[x²] − E[x]²

var(x+y) ≥ var(x)

Var(x) + var(y) ≥ var(x)

Var(y) ≥ 0

Var(x+y) = var(x) + var(y)

Why is this true

Subtract var(x) from both sides ?

Var is magnitude, always ≥ 0

So theirfor it is some significant quantity ≥ 0 ?

#2

1. Always True

Show that $P(A \cap B^c) = P(A)P(B^c)$

Start w/

$P(A) = P(A \cap B) + P(A \cap B^c)$ ← how do you know to start here?

$$
\begin{aligned}
P(A \cap B^c) &= P(A) - P(A \cap B) \\
&= P(A) - P(A)P(B) \\
&= P(A)(1 - P(B)) \\
&= P(A)\,P(B^c)
\end{aligned}
$$

I am so bad at ~~posts~~ proofs

- so bad at breaking things down in small steps

2. Not always true

$C = A \cap B$

Let $P(A) > P(C)$ and $P(B) > P(C)$

So $P(A \cap B | C) = 1$

$P(A|C) = 1 \quad P(B|C) = 1$

← I will never get this

$P(A \cap B | C) = P(A|C)\, P(B|C)$

A + B are conditionally ind. given C

↓ Might get this

Given C^c A + B are disjoint so not independent

Say have 3 coins $p = 1/5$
$1/3$
$2/3$

Prob that select coin w/ p $1/5$ = C

So C^c is that you choose one of the other 2 coins

A = event 1st coin toss = heads

B = event 2nd " " "

For given coin tosses are ind such that

$P(B | A \cap C) = P(B|C)$

(2)

I like this idea of making a real world example

But given C^c A, B are <u>not</u> ind since ~~it~~ can have either coin $p = \frac{1}{3}$ or $p = 2/3$

knowing A changes belief on coin toss

$$P(B \mid A \cap C^c) = \frac{B \cap A \cap C^c}{A \cap C^c}$$

$$= \frac{\frac{1}{3}\left(\left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2\right)}{1/3\left(\frac{1}{3} + \frac{2}{3}\right)}$$

$$= \frac{5}{9}$$

<u>But</u> $P(B \mid C^c) = \frac{P(B \cap C^c)}{P(C^c)}$

$$= \frac{\frac{1}{3}\left(\frac{1}{3} + \frac{2}{3}\right)}{\frac{2}{3}}$$

$$= \frac{1}{2}$$

3. Always true

$var(x+y) = var(x) + var(y)$

var is always non neg, so

$var(x) + var(y) \geq var(x)$

Could at least have gotten that one

Think about on exam

Don't freeze up

-Well its only easy after seeing the answers

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

Quiz 1 Solutions:
October 12, 2010

Problem 1.

1. **(10 points)** Let R_i be the amount of time Stephen spends at the ith red light. R_i is a Bernoulli random variable with $p = 1/3$. The PMF for R_i is:

$$\mathbf{P}_{R_i}(r) = \begin{cases} 2/3, & \text{if } r = 0, \\ 1/3, & \text{if } r = 1, \\ 0, & \text{otherwise.} \end{cases}$$

The expectation and variance for R_i are:

$$\begin{aligned} \mathbf{E}[R_i] &= p = \frac{1}{3}, \\ \text{var}(R_i) &= p(1-p) = \frac{1}{3}\frac{2}{3} = \frac{2}{9}. \end{aligned}$$

Let T_S be the total length of time of Stephen's commute in minutes. Then,

$$T_S = 18 + \sum_{i=1}^{5} R_i.$$

T_S is a shifted binomial with $n = 5$ trials and $p = 1/3$. The PMF for T_S is then:

$$\mathbf{P}_{T_S}(k) = \begin{cases} \binom{5}{k-18}\left(\frac{1}{3}\right)^{k-18}\left(\frac{2}{3}\right)^{23-k}, & \text{if } k \in \{18, 19, 20, 21, 22, 23\}, \\ 0, & \text{otherwise.} \end{cases}$$

The expectation and variance for T_S are:

$$\begin{aligned} \mathbf{E}[T_S] &= \mathbf{E}\left[18 + \sum_{i=1}^{5} R_i\right] \\ &= \frac{59}{3}. \\ \text{var}(T_S) &= \text{var}\left(18 + \sum_{i=1}^{5} R_i\right) \\ &= \frac{10}{9}. \end{aligned}$$

2. **(10 points)** Let N be the number of red lights Stephen encountered on his commute. Given that $T_S \leq 19$, then $N = 0$ or $N = 1$. The unconditional probability of $N = 0$ is $\mathbf{P}(N = 0) = (\frac{2}{3})^5$. The unconditional probability of $N = 1$ is $\mathbf{P}(N = 1) = \binom{5}{1}(\frac{2}{3})^4(\frac{1}{3})^1$.

To find the conditional expectation, the following conditional PDF is calculated:

$$\mathbf{P}_{N|T_S\le 19}(n \mid T_S \le 19) = \begin{cases} \dfrac{(\frac{2}{3})^5}{(\frac{2}{3})^5 + \binom{5}{1}(\frac{2}{3})^4(\frac{1}{3})^1}, & \text{if } n=0, \\ \dfrac{\binom{5}{1}(\frac{2}{3})^4(\frac{1}{3})^1}{(\frac{2}{3})^5 + \binom{5}{1}(\frac{2}{3})^4(\frac{1}{3})^1}, & \text{if } n=1, \\ 0, & \text{otherwise,} \end{cases} = \begin{cases} 2/7, & \text{if } n=0, \\ 5/7, & \text{if } n=1, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$\mathbf{E}[N \mid T_S \le 19] = \frac{5}{7}.$$

3. **(10 points)** Given that the last red light encountered by Stephen was the fourth light, $R_4 = 1$ and $R_5 = 0$.

   We are asked to compute $\text{var}(N \mid \{R_4 = 1\} \cap \{R_5 = 0\})$. Therefore,

$$\begin{aligned} \text{var}(N \mid \{R_4=1\} \cap \{R_5=0\}) &= \text{var}(R_1+R_2+R_3+R_4+R_5 \mid \{R_4=1\} \cap \{R_5=0\}) \\ &= \text{var}(R_1+R_2+R_3+1+0 \mid \{R_4=1\} \cap \{R_5=0\}) \\ &= \text{var}(R_1+R_2+R_3+1) \\ &= \text{var}(R_1+R_2+R_3) \\ &= 3\text{var}(R_1) \\ &= \frac{6}{9}. \end{aligned}$$

4. **(10 points)** Under the given condition, the discrete uniform law can be used to compute the probability of interest. There are $\binom{5}{3}$ ways that Stephen can encounter a total of three red lights. There are $\binom{3}{2}$ ways that two out of the first three lights were red. This leaves one additional red light out of the last two lights and there are $\binom{2}{1}$ possible ways that this event can occur. Putting it all together,

$$\mathbf{P}(\text{two of first three lights were red} \mid \text{total of three red lights}) = \frac{\binom{3}{2}\binom{2}{1}}{\binom{5}{3}} = \frac{3}{5}.$$

5. **(5 points)** Let T_J be the total length of time of Jon's commute in minutes. The PMF of Jon's commute is:

$$\mathbf{P}_{T_J}(\ell) = \begin{cases} \frac{1}{4}, & \text{if } \ell \in \{20, 21, 22, 23\}, \\ 0, & \text{otherwise.} \end{cases}$$

6. **(10 points)** Let A be the event that Jon arrives at work in 20 minutes and let B be the event that exactly one person arrives in 20 minutes.

$$\begin{aligned} \mathbf{P}(A \mid B) &= \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \\ &= \frac{\mathbf{P}(\{T_J = 20\} \cap \{T_S \ne 20\})}{\mathbf{P}(\{T_J = 20\} \cap \{T_S \ne 20\}) + \mathbf{P}(\{T_J \ne 20\} \cap \{T_S = 20\})} \\ &= \frac{\mathbf{P}(T_J = 20)\mathbf{P}(T_S \ne 20)}{\mathbf{P}(T_J = 20)\mathbf{P}(T_S \ne 20) + \mathbf{P}(T_J \ne 20)\mathbf{P}(T_S = 20)}. \end{aligned}$$

Jon arrives at work in 20 minutes (or $T_J = 20$) if he does not have to wait for the train at the station (or $X = 0$). The probability of this event occurring is:

$$\mathbf{P}(T_J = 20) = \mathbf{P}(X = 0) = \frac{1}{4}.$$

Stephen arrives at work in 20 minutes if he encounters 2 red lights. The probability of this event is a binomial probability:

$$\mathbf{P}(T_S = 20) = \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3.$$

Thus,

$$\mathbf{P}(A \mid B) = \frac{\frac{1}{4}\left(1 - \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3\right)}{\frac{1}{4}\left(1 - \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3\right) + \frac{3}{4}\left(\binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3\right)}.$$

7. **(10 points)** The probability of interest is $\mathbf{P}(T_S \leq T_J)$. This can be calculated using the total probability theorem by conditioning on the length of Jon's commute or Jon's wait at the station. If Jon's commute is 20 minutes (or $X = 0$), then Stephen can encounter up to 2 red lights to satisfy $T_S \leq T_J$. Similarly if Jon's commute is 21 minutes (or $X = 1$), Stephen can encounter up to 3 red lights and so on.

$$\begin{aligned}
\mathbf{P}(T_S \leq T_J) &= \sum_{x=0}^{3} \mathbf{P}(T_S \leq T_J \mid X = x)\mathbf{P}(X = x) \\
&= \frac{1}{4}\sum_{x=0}^{3}\sum_{k=0}^{2+x}\binom{5}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{5-k} \\
&= 0.9352.
\end{aligned}$$

An alternative approach follows. We first compute the joint PMF of the commute times of Stephen and Jon $\mathbf{P}_{T_S,T_J}(k,\ell)$. Because of independence, $\mathbf{P}_{T_S,T_J}(k,\ell) = \mathbf{P}_{T_S}(k)\mathbf{P}_{T_J}(\ell)$.

Therefore,

$$\begin{aligned}
\mathbf{P}(T_S \leq T_J) &= \mathbf{P}(T_S = 18) + \mathbf{P}(T_S = 19) + \mathbf{P}(T_S = 20) + \mathbf{P}(\{T_S = 21\} \cap \{T_J \geq 21\}) \\
&\quad + \mathbf{P}(\{T_S = 22\} \cap \{T_J \geq 22\}) + \mathbf{P}(\{T_S = 23\} \cap \{T_J = 23\}) \\
&= \left(\frac{2}{3}\right)^5 + \binom{5}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^4 + \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3 + \binom{5}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2 \cdot \left(\frac{3}{4}\right) \\
&\quad + \binom{5}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^1 \cdot \left(\frac{2}{4}\right) + \left(\frac{1}{3}\right)^5 \cdot \left(\frac{1}{4}\right) \\
&= 0.9352.
\end{aligned}$$

8. **(10 points)** We express the conditional probability as such:

$$\mathbf{P}(X = 3 \mid T_S \leq T_J) = \frac{\mathbf{P}(\{X = 3\} \cap \{T_S \leq T_J\})}{\mathbf{P}(T_S \leq T_J)}.$$

If Jon waited 3 minutes at the train, his commute was 23 minutes and Stephen's commute takes at most as long as Jon's commute since the longest possible commute for Stephen is 23 minutes. Therefore, the numerator in the previous expression is equal to $\mathbf{P}(X = 3) = \frac{1}{4}$. The denominator was computed in the previous part.

$$\begin{aligned}\mathbf{P}(X = 3 \mid T_S \leq T_J) &= \frac{1}{\displaystyle\sum_{x=0}^{3}\sum_{k=0}^{2+x}\binom{5}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{5-k}} \\ &= 0.2673.\end{aligned}$$

Problem 2.

1. **(10 points) Always True.** We need to show that

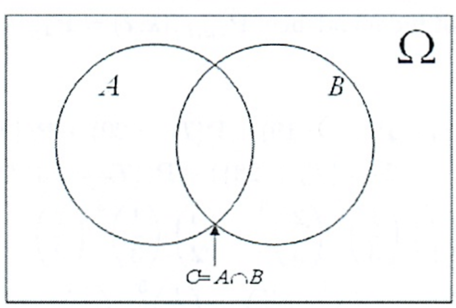
$$\mathbf{P}(A \cap B^c) = \mathbf{P}(A)\mathbf{P}(B^c).$$

We start with expressing $\mathbf{P}(A)$ as $\mathbf{P}(A \cap B) + \mathbf{P}(A \cap B^c)$. Therefore,

$$\begin{aligned}\mathbf{P}(A \cap B^c) &= \mathbf{P}(A) - \mathbf{P}(A \cap B) \\ &= \mathbf{P}(A) - \mathbf{P}(A)\mathbf{P}(B) \\ &= \mathbf{P}(A)(1 - \mathbf{P}(B)) \\ &= \mathbf{P}(A)\mathbf{P}(B^c),\end{aligned}$$

which shows that A and B^c are independent.

2. **(10 points) Not Always True.** Using the diagram below, let $C = A \cap B$ and let $\mathbf{P}(A) > \mathbf{P}(C)$ and let $\mathbf{P}(B) > \mathbf{P}(C)$. The conditional probability $\mathbf{P}(A \cap B \mid C) = 1$. Furthermore, $\mathbf{P}(A \mid C) = 1$ and $\mathbf{P}(B \mid C) = 1$. Since $\mathbf{P}(A \cap B \mid C) = \mathbf{P}(A \mid C)\mathbf{P}(B \mid C)$, A and B are conditionally independent given a third event C. Given C^c, A and B are disjoint which means that A and B are not independent.



The following is an alternative counterexample. Imagine having 3 coins with the following probability of heads: $p = 1/5$, $p = 1/3$ and $p = 2/3$, respectively. Each coin has equal probability of being selected. Let C be the event that you select the coin with $p = 1/5$. Let C^c be the event that you choose one of the other two coins. Let A be the event that the first coin toss results in heads. Let B be the event that the second coin toss results in heads. For a given coin, the tosses are independent such that:

$$\mathbf{P}(B \mid A \cap C) = \mathbf{P}(B \mid C).$$

Given C^c, A and B are not independent since we can have either the $p = 1/3$ coin or the $p = 2/3$ coin. Knowing A changes our beliefs of the result of the second coin toss.

$$\begin{aligned}
\mathbf{P}(B \mid A \cap C^c) &= \frac{B \cap A \cap C^c}{A \cap C^c} \\
&= \frac{\frac{1}{3}\left(\left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2\right)}{\frac{1}{3}\left(\frac{1}{3} + \frac{2}{3}\right)} \\
&= \frac{5}{9}.
\end{aligned}$$

However,

$$\begin{aligned}
\mathbf{P}(B \mid C^c) &= \frac{\mathbf{P}(B \cap C^c)}{\mathbf{P}(C^c)} \\
&= \frac{\frac{1}{3}\left(\frac{1}{3} + \frac{2}{3}\right)}{\frac{2}{3}} \\
&= \frac{1}{2}.
\end{aligned}$$

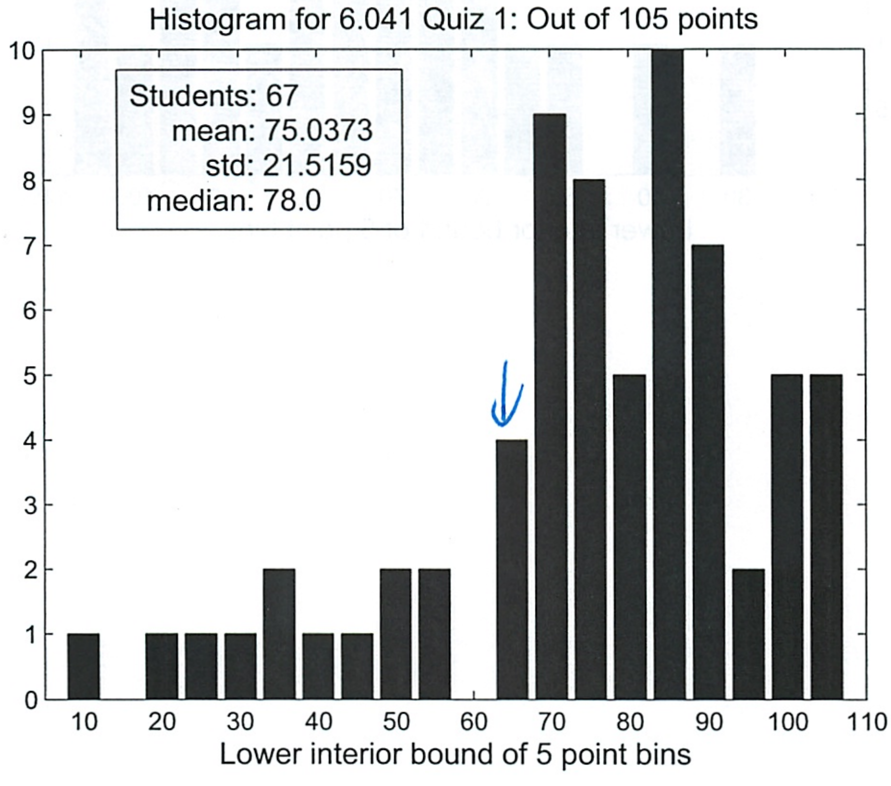
As shown, $\mathbf{P}(B \mid A \cap C^c) \neq \mathbf{P}(B \mid C^c)$.

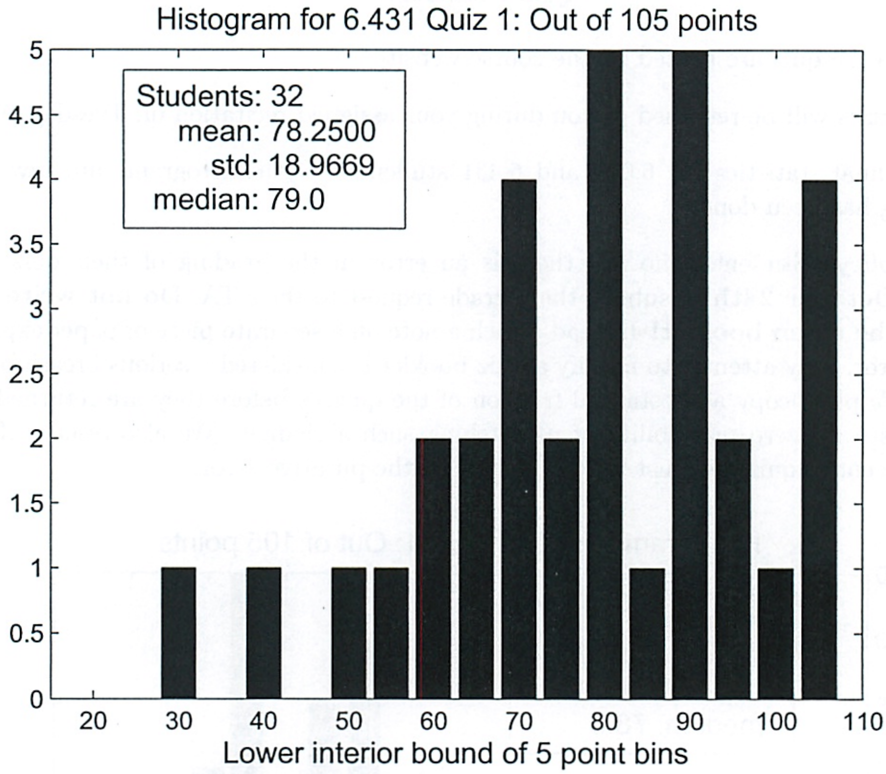
3. **(10 points) Always True.** Using independence of X and Y, $\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$. Since variance is always non-negative, $\text{var}(X) + \text{var}(Y) \geq \text{var}(X)$.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

QUIZ 1 RESULTS

- Solutions to the quiz are posted on the course website.
- Graded quizzes will be returned to you during your assigned recitation on Tuesday 10/18.
- Below are final statistics for 6.041 and 6.431 students. Both histograms are raw scores, no normalizing has been done.
- *Regrade Policy*: Students who feel there is an error in the grading of their quiz have until **Monday October 24th** to submit the regrade request to their TA. **Do not write anything at all on the exam booklet!** Instead attach a note on a separate piece of paper explaining the putative error. Any attempt to modify a quiz booklet is considered a serious breach of academic honesty. We photocopy a substantial fraction of the quizzes before they are returned, implying there exists a nonzero probability of us catching such a change. We also reserve the right to regrade the entire quiz, not just the problem with the putative error.




Histogram for 6.431 Quiz 1: Out of 105 points
Students: 32
mean: 78.2500
std: 18.9669
median: 79.0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
20
30
40
50
60
70
80
90
100
110
Lower interior bound of 5 point bins

6.041/6.431 Probabilistic Systems Analysis

Quiz II Review
Fall 2010

1 Probability Density Functions (PDF)

For a continuous RV X with PDF $f_X(x)$,

$$P(a \le X \le b) = \int_a^b f_X(x)dx$$

$$P(X \in A) = \int_A f_X(x)dx$$

Properties:

- Nonnegativity:

$$f_X(x) \ge 0 \ \forall x$$

- Normalization:

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

2 PDF Interpretation

Caution: $f_X(x) \ne P(X = x)$

- if X is continuous, $P(X = x) = 0 \ \ \forall x$!!
- $f_X(x)$ can be ≥ 1

Interpretation: "probability per unit length" for "small" lengths around x

$$P(x \le X \le x + \delta) \approx f_X(x)\delta$$

3 Mean and variance of a continuous RV

$$\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} x f_X(x)dx \\
\mathrm{Var}(X) &= E\left[(X - E[X])^2\right] \\
&= \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x)dx \\
&= E[X^2] - (E[X])^2 \ (\ge 0) \\
E[g(X)] &= \int_{-\infty}^{\infty} g(x) f_X(x)dx \\
E[aX + b] &= aE[X] + b \\
\mathrm{Var}(aX + b) &= a^2 \mathrm{Var}(X)
\end{aligned}$$

10/27

4 Cumulative Distribution Functions

Definition:

$$F_X(x) = P(X \leq x)$$

monotonically increasing from 0 (at $-\infty$) to 1 (at $+\infty$).

- Continuous RV (CDF is continuous in x):

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt$$

$$f_X(x) = \frac{dF_X}{dx}(x)$$

- Discrete RV (CDF is piecewise constant):

$$F_X(x) = P(X \leq x) = \sum_{k \leq x} p_X(k)$$

$$p_X(k) = F_X(k) - F_X(k-1)$$

5 Uniform Random Variable

If X is a uniform random variable over the interval [a,b]:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{otherwise } (x > b) \end{cases}$$

$$E[X] = \frac{b-a}{2}$$

$$\text{var}(X) = \frac{(b-a)^2}{12}$$

6 Exponential Random Variable

X is an exponential random variable with parameter λ:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \frac{1}{\lambda} \quad \text{var}(X) = \frac{1}{\lambda^2}$$

Memoryless Property: Given that $X > t$, $X - t$ is an exponential RV with parameter λ

7 Normal/Gaussian Random Variables

General normal RV: $N(\mu, \sigma^2)$:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

$$E[X] = \mu, \quad \text{Var}(X) = \sigma^2$$

Property: If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$

then $Y \sim N(a\mu + b, a^2\sigma^2)$

8 Normal CDF

Standard Normal RV: $N(0,1)$
CDF of standard normal RV Y at y: $\Phi(y)$
- given in tables for $y \geq 0$
- for $y < 0$, use the result: $\Phi(y) = 1 - \Phi(-y)$
To evaluate CDF of a general standard normal, express it as a function of a standard normal:

$$X \sim N(\mu, \sigma^2) \Leftrightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$P(X \leq x) = P\Big(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\Big) = \Phi\Big(\frac{x-\mu}{\sigma}\Big)$$

9 Joint PDF

Joint PDF of two continuous RV X and Y: $f_{X,Y}(x,y)$

$$P(A) = \int\int_A f_{X,Y}(x,y)dxdy$$

Marginal pdf: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)dxdy$
Joint CDF: $F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$

10 Independence

By definition,

$$X,\ Y \text{ independent} \Leftrightarrow f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall(x,y)$$

If X and Y are independent:

- $E[XY]$=$E[X]E[Y]$
- $g(X)$ and $h(Y)$ are independent
- $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$

11 Conditioning on an event

Let X be a continuous RV and A be an event with $P(A) > 0$,

$$\begin{aligned}
f_{X|A}(x) &= \begin{cases} \frac{f_X(x)}{P(X \in A)} & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases} \\
P(X \in B|X \in A) &= \int_B f_{X|A}(x)dx \\
E[X|A] &= \int_{-\infty}^{\infty} x f_{X|A}(x)dx \\
E[g(X)|A] &= \int_{-\infty}^{\infty} g(x) f_{X|A}(x)dx
\end{aligned}$$

If $A_1, \ldots, A_n$ are disjoint events that form a partition of the sample space,

$$\begin{aligned} f_X(x) &= \sum_{i=1}^{n} P(A_i) f_{X|A_i}(x) \text{ (}\approx\text{ total probability theorem)} \\ E[X] &= \sum_{i=1}^{n} P(A_i) E[X|A_i] \text{ (total expectation theorem)} \\ E[g(X)] &= \sum_{i=1}^{n} P(A_i) E[g(X)|A_i] \end{aligned}$$

12 Conditioning on a RV

$X,\ Y$ continuous RV

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ f_X(x) &= \int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x|y)dy \quad (\approx totalprobthm) \end{aligned}$$

Conditional Expectation:

$$\begin{aligned} E[X|Y=y] &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx \\ E[g(X)|Y=y] &= \int_{-\infty}^{\infty} g(X) f_{X|Y}(x|y)dx \\ E[g(X,Y)|Y=y] &= \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)dx \end{aligned}$$

Total Expectation Theorem:

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} E[X|Y=y] f_Y(y)dy \\ E[g(X)] &= \int_{-\infty}^{\infty} E[g(X)|Y=y] f_Y(y)dy \\ E[g(X,Y)] &= \int_{-\infty}^{\infty} E[g(X,Y)|Y=y] f_Y(y)dy \end{aligned}$$

13 Continuous Bayes' Rule

X, Y continuous RV, N discrete RV, A an event.

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x) f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y|t) f_X(t)dt} \\ P(A|Y=y) &= \frac{P(A) f_{Y|A}(y)}{f_Y(y)} = \frac{P(A) f_{Y|A}(y)}{f_{Y|A}(y)P(A) + f_{Y|A^c}(y)P(A^c)} \\ P(N=n|Y=y) &= \frac{p_N(n) f_{Y|N}(y|n)}{f_Y(y)} = \frac{p_N(n) f_{Y|N}(y|n)}{\sum_i p_N(i) f_{Y|N}(y|i)} \end{aligned}$$

14 Derived distributions

Def: PDF of a *function* of a RV X with known PDF: $Y = g(X)$.
Method:

- Get the CDF:

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{x|g(x)\leq y} f_X(x)dx$$

- Differentiate: $f_Y(y) = \frac{dF_Y}{dy}(y)$

Special case: if $Y = g(X) = aX + b$, $f_Y(y) = \frac{1}{|a|} f_X(\frac{x-b}{a})$

17

15 Convolution

$W = X + Y$, with X, Y independent.

- Discrete case:

$$p_W(w) = \sum_x p_X(x)p_Y(w - x)$$

- Continuous case:

$$f_W(w) = \int_{-\infty}^{\infty} f_X(x)f_Y(w - x)\, dx$$

18

Graphical Method:

- put the PMFs (or PDFs) on top of each other
- flip the PMF (or PDF) of Y
- shift the flipped PMF (or PDF) of Y by w
- cross-multiply and add (or evaluate the integral)

In particular, if X, Y are independent and normal, then $W = X + Y$ is normal.

19

16 Law of iterated expectations

$E[X|Y = y] = f(y)$ is a number.
$E[X|Y] = f(Y)$ is a random variable
(the expectation is taken with respect to X).
To compute $E[X|Y]$, first express $E[X|Y = y]$ as a function of y.

Law of iterated expectations:

$$E[X] = E[E[X|Y]]$$

(equality between two real numbers)

20

17 Law of Total Variance

$\text{Var}(X|Y)$ is a random variable that is a function of Y (the variance is taken with respect to X).
To compute $\text{Var}(X|Y)$, first express

$$\text{Var}(X|Y=y) = E[(X - E[X|Y=y])^2|Y=y]$$

as a function of y.

Law of conditional variances:

$$\text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E[X|Y])$$

(equality between two real numbers)

18 Sum of a random number of iid RVs

N discrete RV, X_i i.i.d and independent of N.
$Y = X_1 + \ldots + X_N$. Then:

$$\begin{aligned} E[Y] &= E[X]E[N] \\ \text{Var}(Y) &= E[N]\text{Var}(X) + (E[X])^2\text{Var}(N) \end{aligned}$$

19 Covariance and Correlation

$$\begin{aligned} \text{Cov}(X,Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

- By definition, X, Y are uncorrelated $\Leftrightarrow \text{Cov}(X,Y) = 0$.
- If X, Y independent $\Rightarrow$ X and Y are uncorrelated. (the converse is not true)
- In general, Var(X+Y)= Var(X)+ Var(Y)+ 2 Cov(X,Y)
- If X and Y are uncorrelated, Cov(X,Y)=0 and Var(X+Y)= Var(X)+Var(Y)

Correlation Coefficient: (dimensionless)

$$\rho = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} \quad \in [-1, 1]$$

$\rho = 0 \Leftrightarrow$ X and Y are uncorrelated.
$|\rho| = 1 \Leftrightarrow \quad X - E[X] = c[Y - E[Y]]$ (linearly related)

6.041 Review Session

10/27

Quiz 2

Will cover contineous
- but of course include concepts from quiz 1

Lectures 1-12
Recitations 1-13
Tutorials 1-6
P-Set 1-6

Some discrete + contineous are building blocks

Discrete

- binomial
- geometric
- Bernoulli - ~~0 or 1 sucess or failure~~ ~~$P_K(k) = \begin{cases} p & \text{if } k=1 \\ 1-p & k=0 \end{cases}$~~
- Uniform

see next pg

Contineous

- Uniform
- exponential
- normal
- poisson - not on quiz 2

(2)

Discrete

Bernoulli (p)

$$P_X(k) = \begin{cases} p & k=1 \\ 1-p & k=0 \end{cases} \qquad E[X]=p \qquad var(X)=p(1-p)$$

Binomial (n, p)

$$X = X_1 + X_2 + X_3 + \dots + X_n$$

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k=0,1,2\dots \qquad E[X]=np$$

$$var(X) = np(1-p)$$

just add up their variances

N = # of inp. Bernoulli trials to 1st sucess

Geometric (p)

$$P_X(k) = (1-p)^{k-1} p \qquad E[X] = \frac{1}{p} \qquad var(X) = \frac{1-p}{p^2}$$

$k = 1, 2, 3 \dots$

Uniform [a, b]

$$P_X(k) = \frac{1}{b-a+1} \qquad k = a, a+1, \dots b \qquad E[X] = \frac{b+a}{2} \qquad var(X) = \frac{(b-a)(b-a+2)}{12}$$

? now I know what that means!

(3)

Contineous

Uniform (a, b)

$f_x(x) = \frac{1}{b-a} \quad x \in [a,b] \quad E[x] = \frac{a+b}{2} \quad var(x) = \frac{(b-a)^2}{12}$

exponential (λ)

$f_x(x) = \lambda e^{-\lambda x} \quad x \geq 0 \quad E[x] = 1/\lambda \quad var(x) = 1/\lambda^2$

normal (μ, σ^2)

$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad E[x] = \mu \quad var(x) = \sigma^2$

(4)

$$E[X] = \int_X x f_X(x)\, dx$$

$$Var(X) = E[X - E[X]] = \int [X - E[X]]^2 f_X(x)\, dx$$

CDFs

$$P(X \le x) = \int_{-\infty}^{x} f_X(x)\, dx$$

$$= \int_{-\infty}^{x} f_X(t)\, dt$$ ← dummy variable

← need a lot of work on these

$$P(a \le X \le b) = F_X(b) - F_X(a)$$

$\lim_{x \to -\infty} F_X(x) = 0$ left ← $\qquad F_X(x) \ge 0$

$\lim_{x \to \infty} F_X(x) = 1$ right → $\qquad$ Monotonically non decreasing

exponential CDF

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

normal CDF

Use sheet w/ Φ which is for Normal(0, 1)

$Y = \dfrac{X - \mu}{\sigma}$ So now $Y = N(0, 1)$

5

joint PDF

$f_{X,Y}(x,y)$

$P(x,y \in B) = \iint_B f_{X,Y}(x,y)\,dx\,dy$

$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x)$

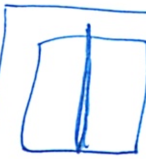
marginal pdf

integrate over all the possible values

$f_X(x) = \int_y f_{X,Y}(x,y)\,dy$

renormalize?

- no - already normalized
- all the prob adds up to 1
- squishing it to that axis



conditional
not marginal

Conditioning

- can condition on a RV
- $f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)}$ ← now must renormalize! denominator does that
- condition on an event

$f_{X|B}(x) = \dfrac{f_X(x)}{P(B)}$ ← normalize w/ denom

$X \in B$ only for!

e

Memoryless property

- exponential

$X = \exp(\lambda)$

$f_X(x) = \lambda e^{-\lambda x} \qquad x \geq 0$

$\underline{X \geq t}$

So can use conditioning to prove

$$f_{X|X\geq t}(x) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda t}}$$

← 1-CDF of exponential evaluated at t

$$= \lambda e^{-\lambda(x-t)}$$

still get exponential, but shifted

notice the exponential math

Bayes Rule

$X \rightarrow \boxed{\text{measurement}} \rightarrow Y \rightarrow \boxed{} \rightarrow X$

don't know what was sent

observe Y

know conditional pdf $f_{Y|X}(y|x)$

inference what x sent $f_{X|Y}(x|y)$

$$f_{X|Y}(x|y) = \frac{f_X(x)\, f_{Y|X}(y|x)}{\int_x f_X(x)\, f_{Y|X}(x)\, dx}$$

joint

need practice on actually <u>doing</u> these

if x is discrete

$$p_{X|Y}(x|y) = \frac{p_X(x)\, f_{Y|X}(x)}{\sum_x p_X(x)\, f_{Y|X}(y|x)}$$

⑦

Derived distributions

$Z = g(X)$

$f_Z(z) = ?$

First find CDF

$F_Z(z) = P(Z \leq z)$

$= P(g(X) \leq z)$

use fundemental theorm of calculus
differenitate

$$f_Z(z) = \frac{\partial F_Z(z)}{\partial z}$$

also if $Z = f(X, Y)$

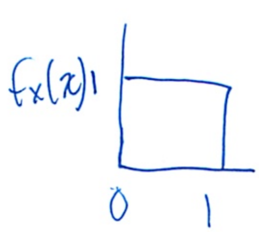
Convolution

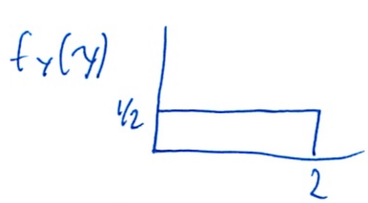
X, Y independant

$Z = X + Y$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\, f_Y(z-t)\, dt$$

flip + slide



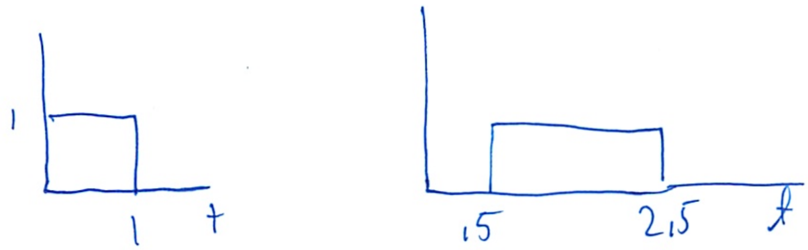


- how much sliding is there
- or was it just that one problem

$f_Z(2.5) = ?$

(8)

Flip + slide!



Multiply PDF

- only nonzero values
- so where overlap

$f_Z(2.5) = \int_{.5}^{1} 1 \cdot \frac{1}{2}\, dt$ ← only 1 region here

remember that other p-set qu

Coverance (x, y)

$\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y]$

$\mathrm{Var}(X) = \mathrm{Cov}(X, Y)$ ← not always I think

$\rho(x, y)$ ← dimensionless correlation coefficent

$= \dfrac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$

if X, Y are independent → covariance = 0

9

Law of iterated expectations

$$E[X] = E[E[X|Y]]$$

Law of total variance

$$var(X) = E[var(X|Y)] + var(E[X|Y])$$

Sum of R# of RV

$$Y = X_1 + \dots + X_n$$

X_i must be inp. of each other and N

$$E[Y] = E[E[Y|N]]$$
$$= E[X_i]\,E[N]$$

$$var(Y) = \text{add variances}$$
$$= E[N\,var(X_i)] + var(N\,E[X_i])$$
$$= E[N]\,var(X_i) + (E[X_i])^2\,var(N)$$

6.041/6.431 Fall 2010 Quiz 2
Tuesday, November 2, 7:30 - 9:30 PM.

DO NOT TURN THIS PAGE OVER UNTIL YOU ARE TOLD TO DO SO

Question	Score	Out of
1.1	6	10
1.2	0	10
1.3	6 PH	10
1.4	5	10
1.5	10	10
1.6	10	10
1.7	10	10
1.8	3	10
2.1	4	10
2.2	10	10
2.3	2	5
2.4	2	5
Your Grade	68	110

Name: Michael Plasmeier

Recitation Instructor: Dimitri

TA: Aliaa

- **For full credit, answers should be algebraic expressions (no integrals), in simplified form. These expressions may involve constants such as π or e, and need not be evaluated numerically.** is so different than 1st
- This quiz has 2 problems, worth a total of 110 points.
- You may tear apart page 3, as per your convenience, **but you must turn them in together with the rest of the booklet.**
- Write your solutions in this quiz booklet, only solutions in this quiz booklet will be graded. Be neat! You will not get credit if we can't read it.
- You are allowed two two-sided, handwritten, 8.5 by 11 formula sheets. Calculators are not allowed.
- You have 120 minutes to complete the quiz.
- Graded quizzes will be returned in recitation on Thursday 11/4.

Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2010)

Problem 0: (0 points) Write your name, your assigned recitation instructor's name, and assigned TA's name on the cover of the quiz booklet. The Instructor/TA pairing is listed below.

Recitation Instructor	TA	Recitation Time
Vivek Goyal	Uzoma Orji	10 & 11 AM
Peter Hagelstein	Ahmad Zamanian	12 & 1 PM
Ali Shoeb	Shashank Dwivedi	2 PM
Dimitri Bertsekas (6.431)	Aliaa Atwi	2 & 3 PM

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

Problem 1. (80 points) In this problem:
(i) X is a (continuous) uniform random variable on $[0, 4]$.
(ii) Y is an exponential random variable, independent from X, with parameter $\lambda = 2$.

1. **(10 points)** Find the mean and variance of $X - 3Y$.
2. **(10 points)** Find the probability that $Y \geq X$.
   (Let c be the answer to this question.)
3. **(10 points)** Find the conditional joint PDF of X and Y, given that the event $Y \geq X$ has occurred.
   (You may express your answer in terms of the constant c from the previous part.)
4. **(10 points)** Find the PDF of $Z = X + Y$.
5. **(10 points)** Provide a fully labeled sketch of the conditional PDF of Z given that $Y = 3$.
6. **(10 points)** Find $\mathbf{E}[Z \mid Y = y]$ and $\mathbf{E}[Z \mid Y]$.
7. **(10 points)** Find the joint PDF $f_{Z,Y}$ of Z and Y.
8. **(10 points)** A random variable W is defined as follows. We toss a fair coin (independent of Y). If the result is "heads", we let $W = Y$; if it is tails, we let $W = 2 + Y$. Find the probability of "heads" given that $W = 3$.

Problem 2. (30 points) Let $X, X_1, X_2, \ldots$ be independent normal random variables with mean 0 and variance 9. Let N be a positive integer random variable with $\mathbf{E}[N] = 2$ and $\mathbf{E}[N^2] = 5$. We assume that the random variables $N, X, X_1, X_2, \ldots$ are independent. Let $S = \sum_{i=1}^{N} X_i$.

1. **(10 points)** If δ is a small positive number, we have $\mathbf{P}(1 \leq |X| \leq 1 + \delta) \approx \alpha\delta$, for some constant α. Find the value of α.
2. **(10 points)** Find the variance of S.
3. **(5 points)** Are N and S uncorrelated? Justify your answer.
4. **(5 points)** Are N and S independent? Justify your answer.

Each question is repeated in the following pages. Please write your answer on the appropriate page.

Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

Problem 1. (80 points) In this problem:
(i) X is a (continuous) uniform random variable on $[0, 4]$.
(ii) Y is an exponential random variable, independent from X, with parameter $\lambda = 2$.

1. **(10 points)** Find the mean and variance of $X - 3Y$. $= ?$

$X \sim \text{Uniform } [0, 4]$  $Y \sim \exp(2)$  X, Y ind

$f_X(x) = \frac{1}{b-a} = \frac{1}{4} \quad 0 \le x \le 4$ $\qquad f_Y(y) = 2e^{-2y} \quad y \ge 0$

derived distribution?

$F_X(x) = \frac{x-a}{b-a} = \frac{x}{4} \quad 0 \le x \le 4$ $\qquad F_Y(y) = 1 - e^{-2y} \quad y > 0$

$E[X] = \frac{b-a}{2} = \frac{4}{2} = 2$ $\qquad E[Y] = \frac{1}{\lambda} = \frac{1}{2}$

$\text{Var}(X) = \frac{(b-a)^2}{12} = \frac{4^2}{12} = \frac{16}{12} = \frac{4}{3}$ $\qquad \text{Var}(Y) = \frac{1}{\lambda^2} = \frac{1}{4}$

$\text{Var}(X+Y) = \text{var}(X) + \text{var}(Y) + 2\,\text{cov}(X, Y)$

$\text{var}(X - 3Y) = \text{var}(X) - 3\,\text{var}(Y)$ ind so 0

var(x)+9 Var(Y

$= \frac{4}{3} - 3\frac{1}{4}$

$= \frac{4}{3} - \frac{3}{4}$

$= \frac{16}{12} - \frac{9}{12}$

$= \frac{7}{12}$

+1

Or Linearity of Expectation

$E[X+Y] = E[X] + E[Y]$

$E[X-3Y] = E[X] - 3E[Y]$

$= \frac{b-a}{2} - 3\frac{1}{2}$

$= 2 - \frac{3}{2}$

$= \frac{1}{2}$ +5

 First qu should be ...

Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

2. **(10 points)** Find the probability that $Y \geq X$.
(Let c be the answer to this question.)

$P(Y \geq X)$

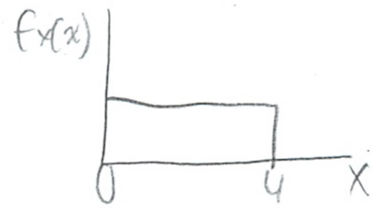
$1 - P(Y \leq X)$

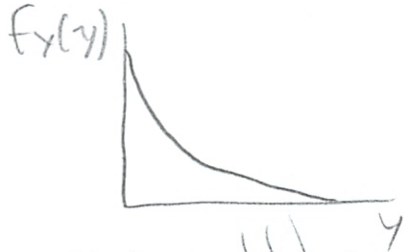
$1 - F_Y(y)$

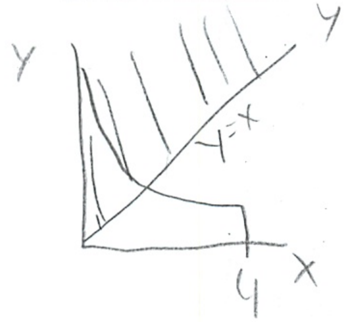
$1 - 1 - e^{-2y}$

$2 - e^{-2y}$

??
how to do?
CDF?







???? can't visualize

and otherwise how done algebraically?

if $y > 4$

$c = P(Y \geq X) = \begin{cases} 1 & y \geq 4 \\ 2 - e^{-2y} & y < 4 \end{cases}$

+0

Should be small since $E[Y] = \frac{1}{2}$ and $E[X] = 2$

3. **(10 points)** Find the conditional joint PDF of X and Y, given that the event $Y \geq X$ has occurred. $\uparrow A$
(You may express your answer in terms of the constant c from the previous part.)

X and Y independent

$$f_{X,Y|A} = \frac{f_{X,Y}(x,y)}{P(A)} = \frac{f_X(x)\, f_Y(y)}{c}$$

$x^2/2$

$$= \frac{\frac{1}{b-a} \cdot \lambda e^{-\lambda y}}{c}$$

$$= \frac{\frac{1}{4} \cdot 2e^{-2y}}{c}$$

$$= \frac{e^{-2y}}{2c} \quad 4/4$$

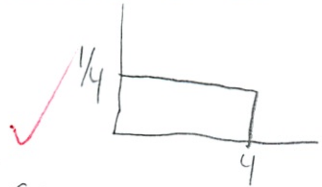
$\frac{6}{10}$

$\frac{0}{4}$ limits

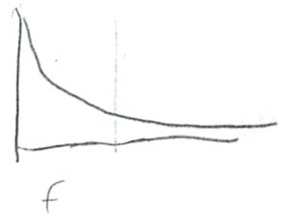
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

4. **(10 points)** Find the PDF of $Z = X + Y$.

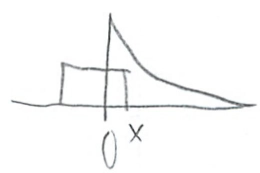
Convolution



flip + slide

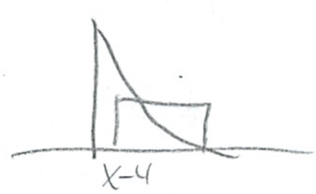


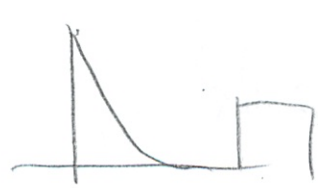
$$f_z(z) = \begin{cases} 0 & z<0 \\ \int_0^x \frac{1}{4}\cdot 2e^{-2(z-x)}\,dx & z>0,\ z<e^2 \\ \int_0^x \frac{1}{4}\cdot 2e^{-2(z-x)} & z>e^2,\ z-4<0,\ z<4 \\ \int_{x-4}^x \frac{1}{4}\cdot 2e^{-2(z-x)} & z>e^2,\ z-4>0,\ z>4 \\ 0 & \text{Never happen} \end{cases}$$



↓ same

$y = \exp(2) = e^2$





→ over

$$\begin{cases} \int_0^x \frac{e^{-2(z-x)}}{2}\,dx & 0<z<4 \\ \int_{x-4}^{x} \frac{e^{-2(z-x)}}{2}\,dx & z>4 \end{cases}$$

; but how suppose to take $\int$ of e^x
Oh its a common one, right? $\int e^x = e^x$

wrong limits & calculations

$$\int_0^x e^{-2z+2x}\,dx = \int_0^x e^{-2z} + \int_0^x e^{2x}\,dx = e^{-2z}\Big|_0^x + e^{+2x}\Big|_0^x$$

$= e^{-2z+2x}$

$= e^{-2(z-x)}$

$z>4$

$$\int_{x-4}^{x} e^{-2z+2x}\,dx = \int_0^x e^{-2z} + \int_0^x e^{2x}$$

$e^{-2z}\Big|_{x-4}^{x} + e^{2x}\Big|_{x-4}^{x}$

$e^{-2z} + e^{2x} - e^{2(x-4)}$

$e^{-2z} + e^{2x} - e^{2x-8}$

$e^{-2z+2x-2x+8}$

e^{-2z+8}

$0<z<4$

$$f_z(z) = \begin{cases} e^{-2z+8} & 0<z<4 \\ e^{-2(z-x)} & z>4 \end{cases}$$

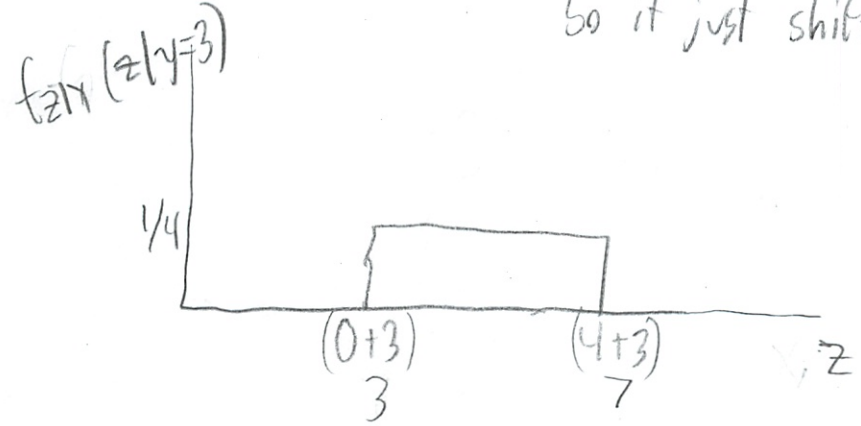
So unsure

5. **(10 points)** Provide a fully labeled sketch of the conditional PDF of Z given that $Y = 3$.

When $Y = y = 3$

Z is $X + 3$

So it just shifts X to the right



by graphing

$$f_{Z|Y}(z \mid y=3) = \begin{cases} 1/4 & 3 < z < 7 \\ 0 & \text{else} \end{cases}$$

$$= \frac{f_{Z,Y}(z,y)}{f_Y(y)}$$

← should build, this is later however

6. **(10 points)** Find $\mathbf{E}[Z \mid Y = y]$ and $\mathbf{E}[Z \mid Y]$. function of Y

$E[Z|Y=y]$ = a #

$= \int_{-\infty}^{\infty} z\, f_{Z|Y}(z|Y=y)\, dz$

intuitivly from last problem

$= \int_{0+y}^{4+y} z \, {}^1\!/_4 \; dz \quad 0+y < z < 4+y$

$= \left.\frac{z^2}{8}\right|_{0+y}^{4+y}$

$= \frac{(4+y)^2}{8} - \frac{(y)^2}{8} = 2+y$

$E[Z|Y]$ = a RV, function of Y

$= \int_{0+Y}^{4+Y} {}^1\!/_4 z \, dz \quad 0+Y < Z < 4+Y$

$\left.\frac{Z^2}{4}\right|_{0+Y}^{4+Y}$

$= \frac{(4+Y)^2}{8} - \frac{(Y)^2}{8} = 2+Y$

function of Y

Urg on all

test using $y=3$, should $=5$ try $y=0$, should $=2$

$$\frac{(4+3)^2}{8} - \frac{4^2}{8}$$

$$\frac{49}{8} - 2$$

$$6\tfrac{1}{8} - 2$$

$$4\tfrac{1}{8}$$ ⊗

close
- why the 1/8
practice test
like this

$$\frac{(4+0)^2}{8} - 0$$

$$\frac{16}{8} = 2$$ ✓

7. **(10 points)** Find the joint PDF $f_{Z,Y}$ of Z and Y.

Z and Y are not independent so can't $f_Z(z) f_Y(y)$
so how do otherwise

$f_Y(y) \, f_{Z|Y}(z|y)$

$2e^{-2y} \cdot \frac{1}{4}$ $\quad 0+y < z < 4+y$

$y \geq 0$

$f_{Z,Y}(z,y) = \frac{e^{-2y}}{2}$ $\quad y \geq 0$

$0+y \leq z < 4+y$

$y < z < 4+y$

8. **(10 points)** A random variable W is defined as follows. We toss a fair coin (independent of Y). If the result is "heads", we let $W = Y$; if it is tails, we let $W = 2 + Y$. Find the probability of "heads" given that $W = 3$.

Bayes' Rule

$P(\text{heads}) = q$

$P(\text{tails}) = 1 - q$

know $P_{W|} = \begin{cases} Y & \text{if heads}, q \\ Y+2 & \text{if tails}, 1-q \end{cases}$

want $P_{Heads|W}(\text{heads} \mid W=3) =$

$P_{W|heads} = Y$

$P_{Heads|W} = \dfrac{P(\text{heads})\, f_{W|Heads}(w)}{f_W(w)}$

derived dist?

$\dfrac{q \cdot Y}{2e^{-2Y}q}$

3

$P_Y(y) = 2e^{-2y}$

$P_W(w) = 2e^{-2y}$ if q / heads

$2+2e^{-2y}$ if $1-q$ / tails

$= \begin{cases} 2e^{-2x} \cdot q \\ (2+2e^{-2y})(1-q) \end{cases}$

I like when we get the #s

Problem 2. (30 points) Let $X, X_1, X_2, \ldots$ be independent normal random variables with mean 0 and variance 9. Let N be a positive integer random variable with $\mathbf{E}[N] = 2$ and $\mathbf{E}[N^2] = 5$. We assume that the random variables $N, X, X_1, X_2, \ldots$ are independent. Let $S = \sum_{i=1}^{N} X_i$.

Sum or random # of RVs

1. **(10 points)** If δ is a small positive number, we have $\mathbf{P}(1 \le |X| \le 1+\delta) \approx \alpha\delta$, for some constant α. Find the value of α.

interval? into pronfs?
Or is this just the constant thing? what is it asking?

aka $f_X(1)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{3\sqrt{2\pi}} e^{\frac{-(1-0)^2}{2\cdot 3^2}}$$

$$= \frac{1}{3\sqrt{2\pi}} e^{\frac{-1}{2\cdot 3^2 = 18}}$$

-ah its $f_X(x)$ ← thank you formula sheet where $x=1$

6

. But that is if the question is $P(1 \le X \le 1+\delta)$
. to get $P(1 \le |X| \le 1+\delta)$ it would be $\frac{2}{178}$
$2 \cdot P(1 \le X \le 1+\delta) \approx 2 \cdot f_X(1)$
So

2. **(10 points)** Find the variance of S.

Sum of random # of RVs

$$\mathrm{var}(S) = E[N]\,\mathrm{var}(X) + (E[X])^2\,\mathrm{var}(N)$$
$$2 \cdot 9 + 0^2 \;?$$
$$= 18$$ ✓

$E[X] = \mu = 0$
$\mathrm{var}(X) = \sigma^2 = 9$

$$E[S] = E[N]\,E[X]$$
$$= 2 \cdot 0$$
$$= 0$$

3. **(5 points)** Are N and S uncorrelated? Justify your answer.

$= E[(N-E[N])(S-E[S])]$
$= E[NS] - E[N]E[S]$
$\quad 0 \quad - \quad 2 \cdot 0$

just did today, but where ind.

$0 - 0 = 0$ = Uncorrelated

S will always center around 0, no matter how many tosses one has (N)

Not true!

+2

4. **(5 points)** Are N and S independent? Justify your answer.

No! S is $\sum_{i=1}^{N} X_i$. N is in the definition of S, thus it is ~~in~~dependent.

?

Oh wow duh

0/5

Uncorrelated does not imply independence.

Although stuff I studied I don't think helped that much
diff qu

Utter disaster
- all abstract qu
- recitation one for easier

all for partial credit
not sure on a single one
and I thought I kinda knew

except #2b
really worked it for 2 hrs
though

only 1b → no clue

6.041/6.431 Fall 2010 Quiz 2 Solutions

Problem 1. (80 points) In this problem:
(i) X is a (continuous) uniform random variable on $[0, 4]$.
(ii) Y is an exponential random variable, independent from X, with parameter $\lambda = 2$.

1. **(10 points)** Find the mean and variance of $X - 3Y$.

$$\begin{aligned}\mathbf{E}[X-3Y] &= \mathbf{E}[X] - 3\mathbf{E}[Y] \\ &= 2 - 3 \cdot \frac{1}{2} \\ &= \frac{1}{2}.\end{aligned}$$

$$\begin{aligned}\mathrm{var}(X-3Y) &= \mathrm{var}(X) + 9\mathrm{var}(Y) \\ &= \frac{(4-0)^2}{12} + 9 \cdot \frac{1}{2^2} \\ &= \frac{43}{12}.\end{aligned}$$

2. **(10 points)** Find the probability that $Y \geq X$.
(Let c be the answer to this question.)

The PDFs for X and Y are:

$$f_X(x) = \begin{cases} 1/4, & \text{if } 0 \leq x \leq 4, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_Y(y) = \begin{cases} 2e^{-2y}, & \text{if } y \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Using the total probability theorem,

$$\begin{aligned}\mathbf{P}(Y \geq X) &= \int_x f_X(x)\mathbf{P}(Y \geq X \mid X = x)\, dx \\ &= \int_0^4 \frac{1}{4}(1 - F_Y(x))\, dx \\ &= \int_0^4 \frac{1}{4} e^{-2x}\, dx \\ &= \frac{1}{8}\int_0^4 2e^{-2x}\, dx \\ &= \frac{1}{8}(1 - e^{-8}).\end{aligned}$$

3. **(10 points)** Find the conditional joint PDF of X and Y, given that the event $Y \geq X$ has occurred.
(You may express your answer in terms of the constant c from the previous part.)
Let A be the event that $Y \geq X$. Since X and Y are independent,

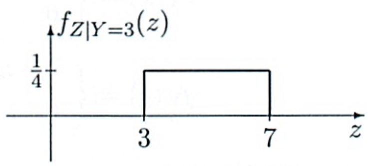
$$\begin{aligned} f_{X,Y|A}(x,y) &= \frac{f_{X,Y}(x,y)}{\mathbf{P}(A)} = \frac{f_X(x)f_Y(y)}{\mathbf{P}(A)} \text{ for } (x,y) \in A \\ &= \begin{cases} \frac{4e^{-2y}}{1-e^{-8}}, & \text{if } 0 \leq x \leq 4,\ y \geq x \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

4. **(10 points)** Find the PDF of $Z = X + Y$.
Since X and Y are independent, the convolution integral can be used to find $f_Z(z)$.

$$\begin{aligned} f_Z(z) &= \int_{\max(0,z-4)}^{z} \frac{1}{4} 2e^{-2t}\, dt \\ &= \begin{cases} 1/4 \cdot \left(1 - e^{-2z}\right), & \text{if } 0 \leq z \leq 4, \\ 1/4 \cdot \left(e^{8} - 1\right) e^{-2z}, & \text{if } z > 4, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

5. **(10 points)** Provide a fully labeled sketch of the conditional PDF of Z given that $Y = 3$.
Given that $Y = 3$, $Z = X + 3$ and the conditional PDF of Z is a shifted version of the PDF of X. The conditional PDF of Z and its sketch are:

$$f_{Z|\{Y=3\}}(z) = \begin{cases} 1/4, & \text{if } 3 \leq z \leq 7, \\ 0, & \text{otherwise.} \end{cases}$$



6. **(10 points)** Find $\mathbf{E}[Z \mid Y = y]$ and $\mathbf{E}[Z \mid Y]$.
The conditional PDF $f_{Z|Y=y}(z)$ is a uniform distribution between y and $y + 4$. Therefore,

$$\mathbf{E}[Z \mid Y = y] = y + 2.$$

The above expression holds true for all possible values of y, so

$$\mathbf{E}[Z \mid Y] = Y + 2.$$

7. **(10 points)** Find the joint PDF $f_{Z,Y}$ of Z and Y.
The joint PDF of Z and Y can be expressed as:

$$\begin{aligned} f_{Z,Y}(z,y) &= f_Y(y) f_{Z|Y}(z \mid y) \\ &= \begin{cases} 1/2 \cdot e^{-2y}, & \text{if } y \geq 0,\ y \leq z \leq y+4, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

8. **(10 points)** A random variable W is defined as follows. We toss a fair coin (independent of Y). If the result is "heads", we let $W = Y$; if it is tails, we let $W = 2 + Y$. Find the probability of "heads" given that $W = 3$.

   Let X be a Bernoulli random variable for the result of the fair coin where $X = 1$ if the coin lands "heads". Because the coin is fair, $\mathbf{P}(X = 1) = \mathbf{P}(X = 0) = 1/2$. Furthermore, the conditional PDFs of W given the value of X are:

$$\begin{aligned} f_{W|X=1}(w) &= f_Y(w) \\ f_{W|X=0}(w) &= f_Y(w-2). \end{aligned}$$

   Using the appropriate variation of Bayes' Rule:

$$\begin{aligned} \mathbf{P}(X = 1 \mid W = 3) &= \frac{\mathbf{P}(X=1) f_{W|X=1}(3)}{\mathbf{P}(X=1) f_{W|X=1}(3) + \mathbf{P}(X=0) f_{W|X=0}(3)} \\ &= \frac{\mathbf{P}(X=1) f_Y(3)}{\mathbf{P}(X=1) f_Y(3) + \mathbf{P}(X=0) f_Y(1)} \\ &= \frac{\mathbf{P}(X=1) f_Y(3)}{\mathbf{P}(X=1) f_Y(3) + \mathbf{P}(X=0) f_Y(1)} \\ &= \frac{e^{-6}}{e^{-6} + e^{-2}}. \end{aligned}$$

Problem 2. (30 points) Let $X, X_1, X_2, \ldots$ be independent normal random variables with mean 0 and variance 9. Let N be a positive integer random variable with $\mathbf{E}[N] = 2$ and $\mathbf{E}[N^2] = 5$. We assume that the random variables $N, X, X_1, X_2, \ldots$ are independent. Let $S = \sum_{i=1}^{N} X_i$.

1. **(10 points)** If δ is a small positive number, we have $\mathbf{P}(1 \le |X| \le 1 + \delta) \approx \alpha\delta$, for some constant α. Find the value of α.

$$\begin{aligned} \mathbf{P}(1 \le |X| \le 1 + \delta) &= 2\mathbf{P}(1 \le X \le 1 + \delta) \\ &\approx 2 f_X(1)\delta. \end{aligned}$$

   Therefore,

$$\begin{aligned} \alpha &= 2 f_X(1) \\ &= 2 \cdot \frac{1}{\sqrt{9 \cdot 2\pi}} e^{-\frac{1}{2} \cdot \frac{(1-0)^2}{9}} \\ &= \frac{2}{3\sqrt{2\pi}} e^{-\frac{1}{18}}. \end{aligned}$$

2. **(10 points)** Find the variance of S.

   Using the Law of Total Variance,

$$\begin{aligned} \mathrm{var}(S) &= \mathbf{E}[\mathrm{var}(S \mid N)] + \mathrm{var}(\mathbf{E}[S \mid N]) \\ &= \mathbf{E}[9 \cdot N] + \mathrm{var}(0 \cdot N) \\ &= 9\mathbf{E}[N] = 18. \end{aligned}$$

3. **(5 points)** Are N and S uncorrelated? Justify your answer.

   The covariance of S and N is

$$\begin{aligned}
\mathrm{cov}(S,N) &= \mathbf{E}[SN]-\mathbf{E}[S]\mathbf{E}[N] \\
&= \mathbf{E}[\mathbf{E}[SN \mid N]]-\mathbf{E}[\mathbf{E}[S \mid N]]\mathbf{E}[N] \\
&= \mathbf{E}[\mathbf{E}[\sum_{i=1}^{N} X_i N \mid N]]-\mathbf{E}[\mathbf{E}[\sum_{i=1}^{N} X_i \mid N]]\mathbf{E}[N] \\
&= \mathbf{E}[X_1]\mathbf{E}[N^2]-\mathbf{E}[X_1]\mathbf{E}[N] \\
&= 0
\end{aligned}$$

   since the $\mathbf{E}[X_1]$ is 0. Therefore, S and N are uncorrelated.

4. **(5 points)** Are N and S independent? Justify your answer.

   S and N are not independent.

   Proof: We have $\mathrm{var}(S \mid N) = 9N$ and $\mathrm{var}(S) = 18$, or, more generally, $f_{S|N}(s \mid n) = N(0, 9n)$ and $f_S(s) = N(0, 18)$ since a sum of an independent normal random variables is also a normal random variable. Furthermore, since $\mathbf{E}[N^2] = 5 \neq (\mathbf{E}[N])^2 = 4$, N must take more than one value and is not simply a degenerate random variable equal to the number 2. In this case, N can take at least one value (with non-zero probability) that satisfies $\mathrm{var}(S \mid N) = 9N \neq \mathrm{var}(S) = 18$ and hence $f_{S|N}(s \mid n) \neq f_S(s)$. Therefore, S and N are not independent.

6.041 Review Session

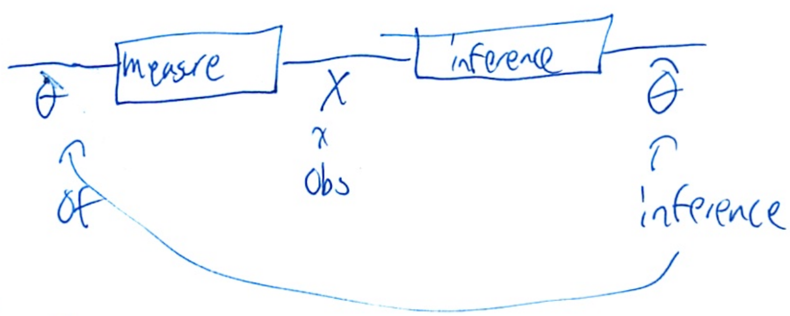
12/9

- Comprehensive but focus on "quiz 3"
- Only up to 9.1

  <u>except</u>

  - t-distribution
  - transform
  - Cont. time markov chain (almost)
  - Strong law of large #
  - convergence up to 1

Chap 8 Bayesian Inference



key factor: θ is RV

has prior dist $p_\theta(\theta)$

X has conditional dist

$p_{x|\theta}(x|\theta)$

want posterior probability

$p_{\theta|x}(\theta|x)$

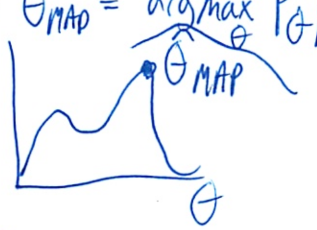
(I think on these got concepts - need practice)

② Use Bayes Rule

$$P_{\Theta|X}(\theta|x) = \frac{P_\Theta(\theta)\, P_{X|\Theta}(x|\theta)}{P_X(x)}$$

Just want 1 # (estimate)

- depends on what criteria we care about
- <u>MAP</u> - $\hat{\theta}_{MAP} = \underset{\theta}{\arg\max}\, P_{\Theta|X}(\theta|x)$



the θ that maximizes the posterior

$$\frac{d}{d\theta} P_{\Theta|X}(\theta|x) = 0$$

+ solve for θ

<u>Or</u> $= \underset{\theta}{\arg\max}\, P_\Theta(\theta)\, P_{X|\Theta}(x|\theta)$ just look at numerator

- "if solution getting too long, you are doing something wrong"

<u>Hypothesis testing</u> means minimize prob of error $(P(\hat{\theta} \neq \theta))$ esp when discrete

$$\underset{\theta}{\arg\max}\, P_{\Theta|X}(\theta|x)$$

$$\theta = \begin{cases} \theta_0 \\ \theta_1 \end{cases}$$

... or finite # of actions

$P_{\Theta|X}(\theta_0|x)$ $\quad$ $P_{\Theta|X}(\theta_1|x)$

look which is bigger

- one that is bigger is confirmed

③ Minimizing Mean-Squared Error LMS / CME

$$\hat{\theta}_{LMS} = E[\theta \mid X = x]$$

Conditional MSE

$$E[(\theta - g(x))^2 \mid X = x]$$

to minimize
pick estimate = LMS

this stuff still unclear on

resulting MSE minimum

$$\mathrm{var}(\theta \mid X = x)$$

MSE

$$E[(\theta - g(x))^2]$$

minimum $E[\mathrm{var}(\theta \mid X)]$

Linear LMS LLMS

- sometimes LMS too complex
- want computationally fast
- a little less accuracy

$$\hat{\theta} = E[\theta] + \frac{\mathrm{cov}(X, \theta)}{\mathrm{var}(X)}(X - E[X])$$

- among linear estimators, this minimizes LMS

Chap 9 Classical Estimation

- don't assume a prior dist
- θ is just an unknown, not random

$$\theta \longrightarrow \boxed{\text{Measure}} \longrightarrow X \longrightarrow \boxed{\text{Estimate}} \longrightarrow \hat{\theta}$$

(4) Regular probability but paramitrized by θ $\quad P_X(x;\theta)$

Can't use Bayes rules

Maximum Liklihood ML

$$\hat{\theta}_n = \underset{\theta}{\arg\max}\ P_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n; \theta)$$

Find θ that maximizes prob of obs Not unknown

$X_1, \ldots$ are generally given iid or at least independent

Replace joint dist w/ a product ↙ since ind

$$= \arg\max \prod_{i=1} P_{X_i}(x_i; \theta)$$

Need to take deriv and set $= 0$

Chain rule ~~...~~ would be a disaster

$$= \underset{\theta}{\arg\max}\ \log\left(\prod_{i=1} P_{X_i}(x_i; \theta)\right)$$

~~Prob~~ Take log and product becomes a sum

$$= \underset{\theta}{\arg\max} \sum_{i=1}^{n} \log\left(P_{X_i}(x_i; \theta)\right)$$

Confidence Intervals

How confident are you in your estimate?

Report an interval $[\hat{\theta}_n^-, \hat{\theta}_n^+]$ such that ~~1-α of the points are likely to fall in there~~ ← wrong wording

θ will be inside w/ prob $1-\alpha$

⑤

$$P\left[\hat{\theta}_n^- \le \theta \le \hat{\theta}_n^+\right] \ge 1-\alpha$$

$\hat{\theta}_n = \frac{\sum X_i}{n}$ true param = mean θ

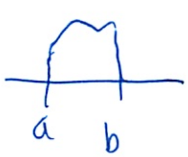
general form → $\left[\hat{\theta}_n - z\sqrt{\frac{V}{n}}\ ;\ \hat{\theta}_n + z\sqrt{\frac{V}{n}}\right]$

regular ML estimate

n = # obs
$V = var(X_1)$ ← either plug in or estimate
$z = ?$ ← use $\Phi(z) = 1 - \frac{\alpha}{2}$
go to normal table
~~work~~ backwards

Var estimation

- if X_i is bernoulli $V \le \frac{1}{4}$ } special case of
- X_i



$Var(X_i) \le \frac{(b-a)^2}{4}$

- Sample var

$$V_n = \frac{1}{n-1} \sum_i (X_i - \underbrace{M_n}_{\text{sample mean}})^2$$

- $\hat{M}_n(1-\hat{M}_n)$

} depend on problem

(6)

Poisson Process

- like a continuous time Bernoulli

$\longrightarrow \lambda$

3 properties
- time homonigenity (memoryless)
  - any intterval of same length is the same
- 2 non-overlapping intervals are ind.
- small interval probabilities
  - are almost Bernoulli

$P(0 \text{ arrivals}) \approx 1 - \lambda\tau$

$P(1 \text{ arrival}) \approx \lambda\tau$

$P(>1 \text{ arrivals}) \approx 0$

What are the RVs associated w/ ?

	$E[\,]$	$var(\,)$
- exponential (λ) - interarrival time $t \geq 0$ - time until 1st arrival $\underline{f_T(t) = \lambda e^{-\lambda t}}$ don't forget!	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
- poisson RV ($\lambda\tau$) $= f_T(t) = \ldots$	$\lambda\tau$	$\lambda\tau$
- erlang (k) $Y_k = T_1 + T_2 + \ldots + T_k$ time until kth arrival time sum of exponentical $f_T(t) = \frac{\lambda^k y^{k-1} e^{-\lambda y}}{(k-1)!}$ $y \geq 0$	$\frac{k}{\lambda}$	$\frac{k}{\lambda^2}$

(is this in right order?)

- Fresh start - what ever happend before does not matter if ~~stop~~ startin watching at random time

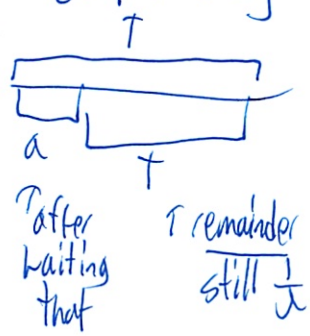
7

Memorylessness – time until 1st arrival

– time to wait 2nd has same distribution

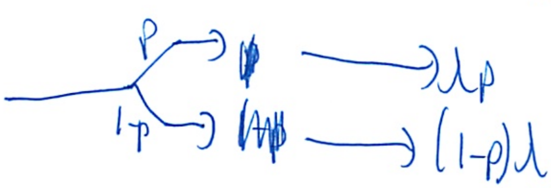
$T = \exp \lambda$ time till 1st arrival ...

$E[T \mid T > 2] = 2 + E[T]$

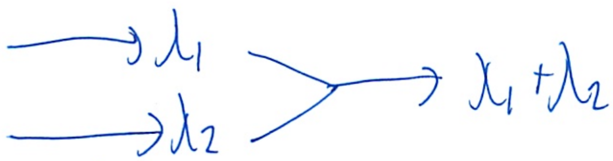


Splitting + Merging

x x → each arrival accept w/ prob p
reject w/ prob $1-p$



) ind (Bernoulli would not have been ind)



$E[\text{time until 1st arrival}] = \frac{1}{\lambda_1 + \lambda_2}$ (✓ knew)

Ok have 1st arrival, prob from $\lambda_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2}$ (✓ knew as well)

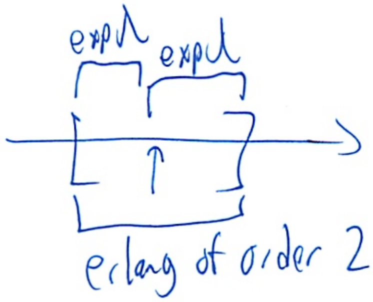
Random Incident Paradox

If you start at a point, poisson both ways you look would be twice as long as you think
Since more likely to land on large time

I am surprised how much I've learned this semester

(8)

My key: tell everything apart!



Limit Theorems

$X \geq 0$ $\quad$ X_i : is RV mean $= \mu$, var $= \sigma^2$

Markov Inequality

$$P(X \geq a) \leq \frac{E[X]}{a} \quad a > 0$$

need to review these concepts

gives a very lose band

Chebisher Inequality

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2} \quad \forall c > 0$$

Sample Mean

- estimate E[ ]
- take samples, add them, divide by n

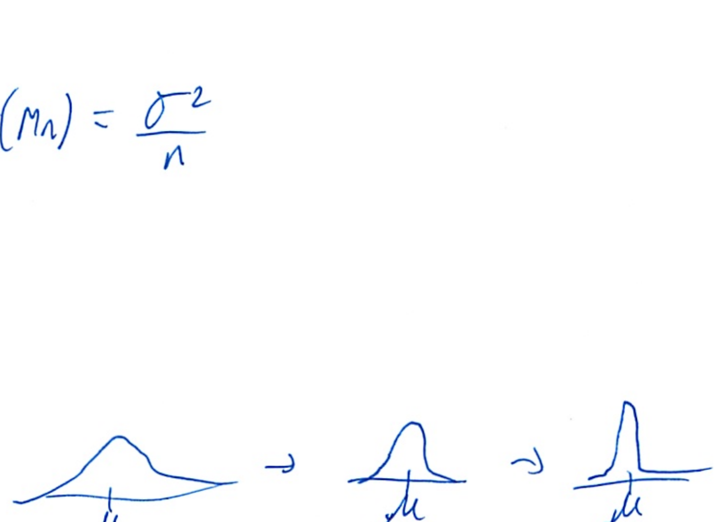
$$M_n = \frac{X_1 + \dots + X_n}{n}$$

$$E[M_n] = E[X_i] = \mu \qquad var(M_n) = \frac{\sigma^2}{n}$$

Weak Law Large #s

$$M_n \xrightarrow[\text{in prob}]{} \mu$$

Converges in Prob



$$P(|M_n - \mu| \geq \epsilon) \to 0 \text{ as } n \to \infty$$

↑ prob that RV will deviate from mean decreases as n ↑ to ∞

(9) WLL# if $n \to \infty$

But if $n \not\to \infty$ can use chebshev

Central Limit Theorm CLT

$X_1, \ldots, X_n$ X_is are iid $E[X_i] = \mu$

$Var(X_i) = \sigma^2$

$X_1 + X_2 + \ldots + X_n \approx N(n\mu, n\sigma^2)$

↑ normal

as $n \to \infty$

DeMore-LaPlace

- just put on cheat sheet + put on exam

6.041 Fall 2009 Final Exam
Tuesday, December 15, 1:30 - 4:30 PM.

DO NOT TURN THIS PAGE OVER UNTIL YOU ARE TOLD TO DO SO

Name: ______________________

Recitation Instructor: ______________________

TA: ______________________

Question	Score	Out of
1		
2 (a)		5
2 (b)		5
2 (c)		5
2 (d)		5
3 (a)		5
3 (b)		5
3 (c)		5
3 (d)		5
3 (e)		5

Question	Score	Out of
4 (a)		5
4 (b)		5
4 (c)		5
4 (d)		5
4 (e)		5
4 (f)		5
5 (a)		5
5 (b)		5
5 (c)		5
5 (d)		5
5 (e)		5
Your Grade		100

- This exam has 5 problems, worth a total of 100 points.
- **When giving a formula for a PDF, make sure to specify the range over which the formula holds.**
- Write your solutions in this quiz booklet, only solutions in this quiz booklet will be graded. Be neat! You will not get credit if we can't read it.
- You are allowed three two-sided, handwritten, formula sheets plus a calculator.
- You may give an answer in the form of an arithmetic expression (sums, products, ratios, factorials) of numbers that could be evaluated using a calculator. Expressions like $\binom{8}{3}$ or $\sum_{k=0}^{5}(1/2)^k$ are also fine.
- The last page of this final contains a standard normal table.

Problem 1: (incorrect answers: -1 point) Write your name, your assigned recitation instructor's name, and assigned TA's name on the cover of the quiz booklet. The Instructor/TA pairing is listed below.

Recitation Instructor	TA	Recitation Time
Jeffrey Shapiro	Jimmy Li	10 & 11 AM
Danielle Hinton	Uzoma Orji	1 & 2 PM
William Richoux	Ulric Ferner	2 & 3 PM
John Wyatt (6.431)	Aliaa Atwi	11 & 12 PM

Problem 2. (20 points)
A pair of jointly continuous random variables, X and Y, have a joint probability density function given by

$$f_{X,Y}(x,y) = \begin{cases} c, & \text{in the shaded region of Fig. 1} \\ 0, & \text{elsewhere.} \end{cases}$$

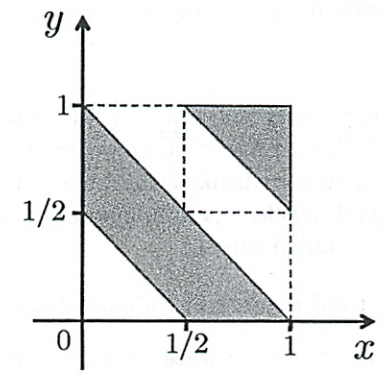


Figure 1: The shaded region is the domain in which $f_{X,Y}(x,y) = c$.

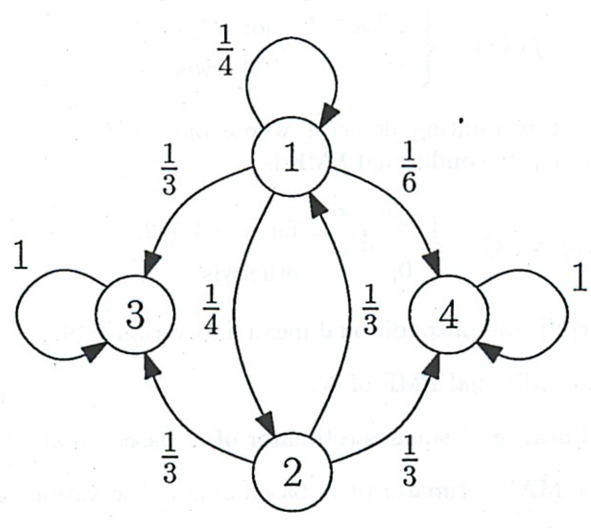
(a) (5 points) Find c.

(b) (5 points) Find the marginal PDFs of X and Y, i.e., $f_X(x)$ and $f_Y(y)$.

(c) (5 points) Find $\mathbf{E}[X \mid Y = 1/4]$ and $\text{Var}[X \mid Y = 1/4]$, that is, the conditional mean and conditional variance of X given that $Y = 1/4$.

(d) (5 points) Find the conditional PDF for X given that $Y = 3/4$, i.e., $f_{X|Y}(x \mid 3/4)$.

Problem 3. (25 points)
Consider a Markov chain X_n whose one-step transition probabilities are shown in the figure.



(a) (5 points) What are the recurrent states?

(b) (5 points) Find $\mathbf{P}(X_2 = 4 \mid X_0 = 2)$.

(c) (5 points) Suppose that you are given the values of $r_{ij}(n) = \mathbf{P}(X_n = j \mid X_0 = i)$. Give a formula for $r_{11}(n+1)$ in terms of the $r_{ij}(n)$.

(d) (5 points) Find the steady-state probabilities $\pi_j = \lim_{n\to\infty} \mathbf{P}(X_n = j \mid X_0 = i)$, or explain why they do not exist.

(e) (5 points) What is the probability of eventually visiting state 4, given that the initial state is $X_0 = 1$?

Problem 4. (30 points)
Al, Bonnie, and Clyde run laps around a track, with the duration of each lap (in hours) being exponentially distributed with parameters $\lambda_A = 21$, $\lambda_B = 23$, and $\lambda_C = 24$, respectively. Assume that all lap durations are independent. At the completion of each lap, a runner drinks either one or two cups of water, with probabilities 1/3 and 2/3, respectively, independent of everything else, including how much water was consumed after previous laps. (The time spent drinking is negligible, assumed zero.)

(a) (5 points) Write down the PMF of the total number of completed laps over the first hour.

(b) (5 points) What is the expected number of cups of water to be consumed by the three runners, in total, over the first hour.

(c) (5 points) Al has amazing endurance and completed 72 laps. Find a good approximation for the probability that he drank at least 130 cups. (You do not have to use 1/2-corrections.)

(d) (5 points) What is the probability that Al finishes his first lap before any of the others?

(e) (5 points) Suppose that the runners have been running for a very long time when you arrive at the track. What is the distribution of the duration of Al's current lap? (This includes the duration of that lap both before and after the time of your arrival.)

(f) (5 points) Suppose that the runners have been running for 1/4 hours. What is the distribution of the time Al spends on his second lap, given that he is on his second lap?

Problem 5. (25 points)
A pulse of light has energy X that is a second-order Erlang random variable with parameter λ, i.e., its PDF is

$$f_X(x) = \begin{cases} \lambda^2 x e^{-\lambda x}, & \text{for } x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

This pulse illuminates an ideal photon-counting detector whose output N is a Poisson-distributed random variable with mean x when $X = x$, i.e., its conditional PMF is

$$p_{N|X}(n \mid x) = \begin{cases} \dfrac{x^n e^{-x}}{n!}, & \text{for } n = 0, 1, 2, \ldots, \\ 0, & \text{otherwise.} \end{cases}$$

(a) (5 points) Find $\mathbf{E}[N]$ and $\text{Var}[N]$, the unconditional mean and variance of N

(b) (5 points) Find $p_N(n)$, the unconditional PMF of N.

(c) (5 points) Find $\hat{X}_{\text{lin}}(N)$, the linear least-squares estimator of X based on an observation of N.

(d) (5 points) Find $\hat{X}_{\text{MAP}}(N)$, the MAP estimator of X based on an observation of N.

(e) (5 points) Instead of the prior distribution in Eq. (1), we are now told that

$$\mathbf{P}(X=2) = 3^3/35, \qquad \mathbf{P}(X=3) = 2^3/35.$$

Given the observation $N = 3$, and in order to minimize the probability of error, which one of the two hypotheses $X = 2$ and $X = 3$ should be chosen?

Useful integral and facts:

$$\int_0^{\infty} y^k e^{-\alpha y}\, dy = \frac{k!}{\alpha^{k+1}}, \quad \text{for } \alpha > 0 \text{ and } k = 0, 1, 2, \ldots \text{ (recall that } 0! = 1)$$

The second-order Erlang random variable satisfies:

$$\mathbf{E}[X] = 2/\lambda, \qquad \mathrm{Var}(X) = 2/\lambda^2.$$

Each question is repeated in the following pages. Please write your answer on the appropriate page.

Problem 2. (20 points)
A pair of jointly continuous random variables, X and Y, have a joint probability density function given by

$$f_{X,Y}(x,y) = \begin{cases} c, & \text{in the shaded region of Fig. 1} \\ 0, & \text{elsewhere.} \end{cases}$$

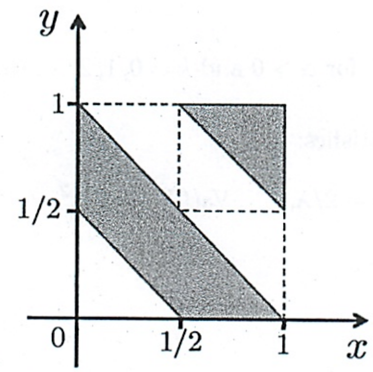


Figure 2: The shaded region is the domain in which $f_{X,Y}(x,y) = c$.

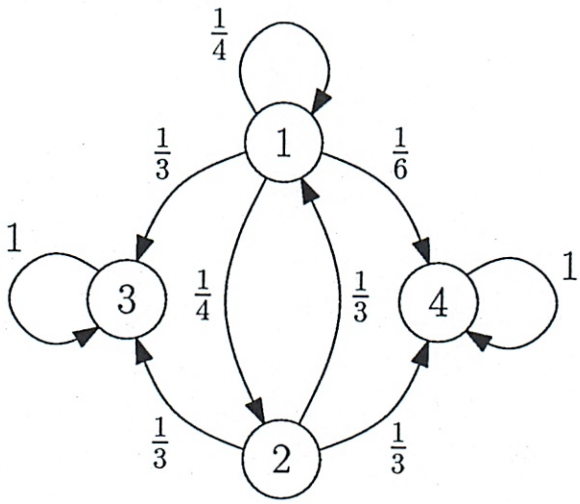
(a) (5 points) Find c.

(b) (5 points) Find the marginal PDFs of X and Y, i.e., $f_X(x)$ and $f_Y(y)$.

(c) (5 points) Find $\mathbf{E}[X \mid Y = 1/4]$ and $\text{Var}[X \mid Y = 1/4]$, that is, the conditional mean and conditional variance of X given that $Y = 1/4$.

(d) (5 points) Find the conditional PDF for X given that $Y = 3/4$, i.e., $f_{X|Y}(x \mid 3/4)$.

Problem 3. (25 points)
Consider a Markov chain X_n whose one-step transition probabilities are shown in the figure.



(a) (5 points) What are the recurrent states?

(b) (5 points) Find $\mathbf{P}(X_2 = 4 \mid X_0 = 2)$.

(c) (5 points) Suppose that you are given the values of $r_{ij}(n) = \mathbf{P}(X_n = j \mid X_0 = i)$. Give a formula for $r_{11}(n+1)$ in terms of the $r_{ij}(n)$.

(d) (5 points) Find the steady-state probabilities $\pi_j = \lim_{n\to\infty} \mathbf{P}(X_n = j \mid X_0 = i)$, or explain why they do not exist.

(e) (5 points) What is the probability of eventually visiting state 4, given that the initial state is $X_0 = 1$?

Problem 4. (30 points)
Al, Bonnie, and Clyde run laps around a track, with the duration of each lap (in hours) being exponentially distributed with parameters $\lambda_A = 21$, $\lambda_B = 23$, and $\lambda_C = 24$, respectively. Assume that all lap durations are independent. At the completion of each lap, a runner drinks either one or two cups of water, with probabilities 1/3 and 2/3, respectively, independent of everything else, including how much water was consumed after previous laps. (The time spent drinking is negligible, assumed zero.)

(a) (5 points) Write down the PMF of the total number of completed laps over the first hour.

(b) (5 points) What is the expected number of cups of water to be consumed by the three runners, in total, over the first hour.

(c) (5 points) Al has amazing endurance and completed 72 laps. Find a good approximation for the probability that he drank at least 130 cups. (You do not have to use 1/2-corrections.)

(d) (5 points) What is the probability that Al finishes his first lap before any of the others?

(e) (5 points) Suppose that the runners have been running for a very long time when you arrive at the track. What is the distribution of the duration of Al's current lap? (This includes the duration of that lap both before and after the time of your arrival.)

(f) (5 points) Suppose that the runners have been running for $1/4$ hours. What is the distribution of the time Al spends on his second lap, given that he is on his second lap?

Problem 5. (25 points)
A pulse of light has energy X that is a second-order Erlang random variable with parameter λ, i.e., its PDF is

$$f_X(x) = \begin{cases} \lambda^2 x e^{-\lambda x}, & \text{for } x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

This pulse illuminates an ideal photon-counting detector whose output N is a Poisson-distributed random variable with mean x when $X = x$, i.e., its conditional PMF is

$$p_{N|X}(n \mid x) = \begin{cases} \frac{x^n e^{-x}}{n!}, & \text{for } n = 0, 1, 2, \ldots, \\ 0, & \text{otherwise.} \end{cases}$$

Useful integral and facts:

$$\int_0^\infty y^k e^{-\alpha y}\, dy = \frac{k!}{\alpha^{k+1}}, \quad \text{for } \alpha > 0 \text{ and } k = 0, 1, 2, \ldots \text{ (recall that 0!=1)}$$

The second-order Erlang random variable satisfies:

$$\mathbf{E}[X] = 2/\lambda, \qquad \mathrm{Var}(X) = 2/\lambda^2.$$

(a) (5 points) Find $\mathbf{E}[N]$ and $\mathrm{Var}[N]$, the unconditional mean and variance of N

(b) (5 points) Find $p_N(n)$, the unconditional PMF of N.

(c) (5 points) Find $\hat{X}_{\text{lin}}(N)$, the linear least-squares estimator of X based on an observation of N.

(d) (5 points) Find $\hat{X}_{\text{MAP}}(N)$, the MAP estimator of X based on an observation of N.

(e) (5 points) Instead of the prior distribution in Eq. (1), we are now told that

$$\mathbf{P}(X = 2) = 3^3/35, \qquad \mathbf{P}(X = 3) = 2^3/35.$$

Given the observation $N = 3$, and in order to minimize the probability of error, which one of the two hypotheses $X = 2$ and $X = 3$ should be chosen?

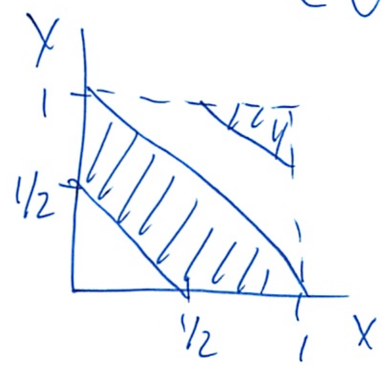
Fall 2009 Review Session

12/9

- quicker teaser than Spring 09
- same prof as us
- but he may give long + hard exams

2. Pair of jointly continuous RV

$$f_{X,Y}(x,y) = \begin{cases} c & \text{in shaded} \\ 0 & \text{otherwise} \end{cases}$$



a) Find c

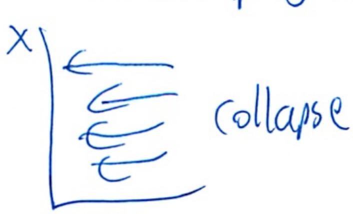
Find area
Divide up into triangles

~~area~~

area $= \frac{1}{2}$

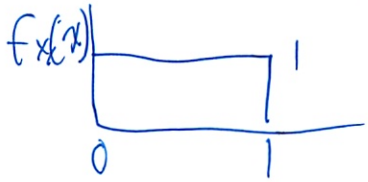
$c = \frac{1}{\text{area}} = \frac{1}{1/2} = 2$

b) Find marginal PDF
- do from graph
- same as collapsing ~~it~~ it to axis looking at

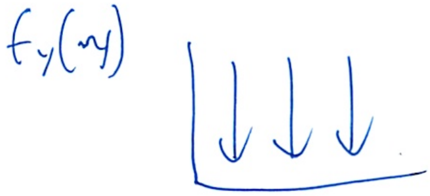


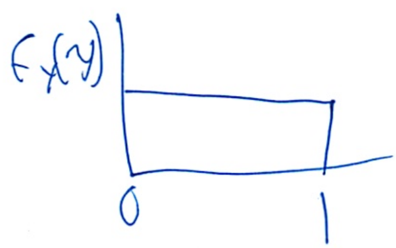
②

See has uniform dist [0,1]



height is $\frac{1}{2}$ everywhere
multiply by $c=2$

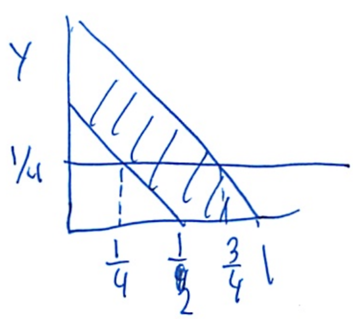


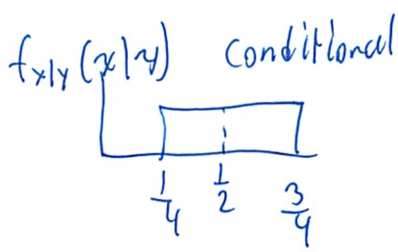


Think about what would happen if got rid of small triangle

c) Find $E[X \mid Y=\frac{1}{4}]$

Look at graph where $y = 1/4$





So $\dfrac{\frac{3}{4}-\frac{1}{4}}{2} = \frac{1}{2}$

$Var(X \mid y=\frac{1}{4}) = \dfrac{b-a}{12}$ since uniform

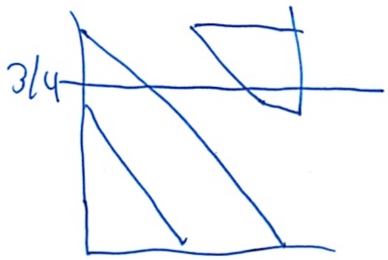
$\dfrac{\frac{3}{4}-\frac{1}{4}}{12} = \dfrac{1}{48}$

Seems easier now

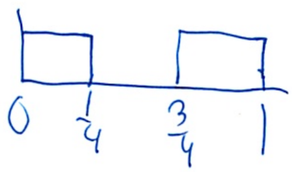
3

d) Find conditional PDF for X given $Y = 3/4$

$f_{X|Y}(x | 3/4)$



Shape will be preserved
- but normalize, so sum to 1

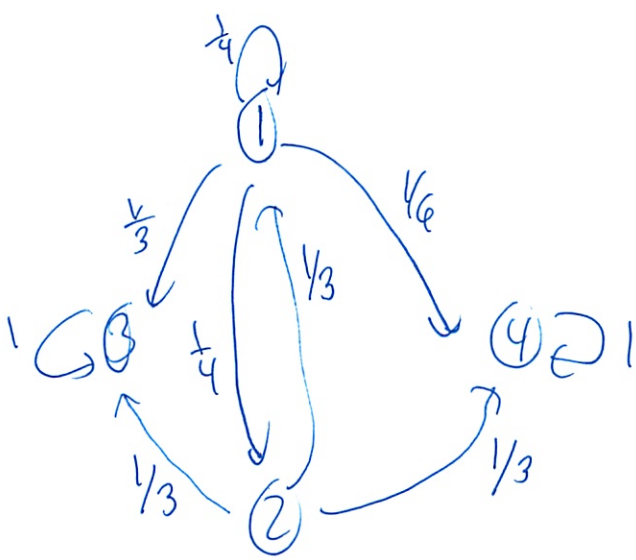


So now what is height?
So area of boxes $\Sigma = 1$

2

Chap 3 problem

3. Markov Chain



(4)

a) What are the recurrent states

3, 4 recurrent

asorbing as well ~~only~~ (just them)

each forms its own recurrenting class

~~Others~~ are transient → 1, 2

b) $P(X_2 = 4 \mid X_0 = 2)$

still need to review

Can write in terms of n-transition possibilities

$r_{24}(2)$ = just look

$= \cancel{r_{24}(1) \cdot 1} +$

$= r_{24}(1) \cdot r_{44}(1) + r_{21}(1)\, r_{14}(1) + r_{22}(1)\, r_{24}(1)$

$= \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot \frac{1}{6}$ 0

$= \frac{7}{18}$

c) Given all n step transition possibilities $r_{ij}(n)$

$r_{11}(n+1) = r_{11}(n) \cdot 1 + r_{12}(n) \cdot \frac{1}{3}$

$r_{11}(1)$? $r_{21}(1)$

says these are given, so this is ans

5

d) Find steady state prob

Must have single recurrent class that is aperiodic
 - else need init state
 ↳ ie not cycling

Whole markov chain could be a single markov chain

e) Now a starting state given

= P(asorbed into 4)

Must actually solve prob of asorbtions

$P(X_\infty = 4 \mid X_0 = 1)$

$a_1 = ?$

Write system of eq + solve

boundry conditions
$$\begin{cases} \underline{S_4} \\ a_4 = 1 \\ a_3 = 0 \end{cases}$$
where you start

$$a_1 = \frac{1}{4}a_1 + \frac{1}{4}a_2 + \frac{1}{3}a_3 + \frac{1}{6}a_4$$
$a_3 \uparrow 0$, $a_4 \uparrow 1$

$$a_2 = \frac{1}{3}a_1 + \frac{1}{3}a_4 +$$
$a_4 \uparrow 1$

Solve, plug a_2 into a_1

$$a_1 = \frac{1}{4}a_1\left(\frac{1}{3}a_1 + \frac{1}{3}\right) + \frac{1}{6}$$

6

$a_1 = \frac{1}{3}a_1 + \frac{1}{4}$

Subtract a_1 from both sides

$\frac{2}{3}a_1 = \frac{1}{4}$

$a_1 = \frac{2}{3} \cdot \frac{1}{4}$

$= \frac{3}{8}$

$a_2 = \frac{1}{3}\left(\frac{3}{8}\right) + \frac{1}{3}$

#2+3? straightforward crank the wheel problems
- the ones I like :)

---

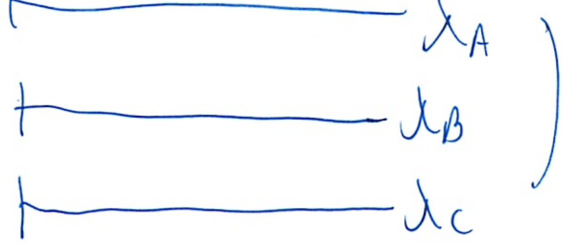
4. $\lambda_A = 21$ $\lambda_B = 23$ $\lambda_C = 24$

$P(\text{One cup}) = 1/3$

$P(\text{Two cup}) = 2/3$

What problem is it? Poisson problem

(might not have guessed the λ clued it off)



total $= \lambda = \lambda_A + \lambda_B + \lambda_C$

merged process

"arrival" = lap complete

a) Arrivals in 1 hr?

- poisson w/ param $(\lambda_A + \lambda_B + \lambda_C)$

7

$p_K(k) = p_K(k, 1)$

$= \frac{68^k}{k!} e^{-68}$ $k = 0, 1, \ldots$ ↓to ∞

↑always for Poisson

b) $E[\text{cups in } 1\text{ hr}]$

$N = \#$ cups in 1 hr $E(N|k]$

$E[N] = E[E[N|k]]$

↑law of iterated expectations

$N = X_0 + X_1 + \cdots + X_k$ = sum of random # of RVs

$X_i \doteq \begin{cases} 1 & \text{w/ prob } \frac{1}{3} \\ 2 & \text{w/ prob } \frac{2}{3} \end{cases}$

↑for ith lap completed

(these compound problems confuse me)

((review this one)

$E[N|k] = k\,E[x_i]$

$E[X_i] = 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3}$

$= \frac{5}{3}$

$E[N] = E[\frac{5}{3}k]$

$= \frac{5}{3} \cdot 68$

c) Al has great endurance. Completed 72 laps

Want P(drank at least 130 cups)

The hint clues us in that it is CLT

8

$N_{AL} = X_1 + X_2 + \dots + X_{72}$ ← iid rv

$P(N_{AL} > 130)$

Need $E[N_{AL}]$ $var(N_{AL})$

$E[N_{AL}] = 72 \cdot E[X_i]$ ← linearity of expectations
must have same expectation

$= 72 \cdot \frac{5}{3}$

$= 120$

$Var(N_{AL}) = 72 \cdot var(X_i)$ ← must be uncorrelated
have same var

$Var(X_i) = E[X_i^2] - (E[X_i])^2$

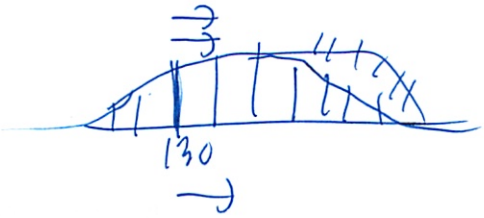
$= \frac{1}{3} + \frac{2}{3} \cdot 4 - \left(\frac{5}{3}\right)^2$

$= \frac{2}{9}$

$Var(N_{AL}) = 72 \cdot \frac{2}{9}$

$= 16$

$\sigma_{N_{AL}} = 4$



~~Normal~~ Convert to normal, use CLT

$P\left(\frac{N_{AL} - 120}{4} > \frac{130 - 120}{4}\right)$

⑨

$= P(Z_n \geq 2.5)$

$= 1 - \Phi(2.5)$

look at table

d) What is prob Al finishes 1st?

$\lambda_A + \lambda_B + \lambda_C = 68$

What is prob it came from Als process?

$$\frac{\lambda_A}{\lambda_A + \lambda_B + \lambda_C} = \frac{21}{68}$$

(seems so easy need to know how to do)

e) You arrive at track

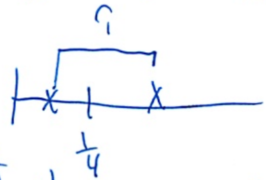
- this is paradox thing → random incidence
- erlang order 2

$R \sim \text{erlang}(2)$

$$f_R(r) = \frac{(21)^2 r}{1!} e^{-21r} \quad r \geq 0$$

⑩

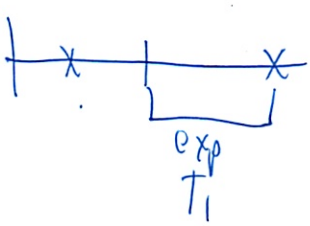
f. This is the hard qu.



Just worry about A1

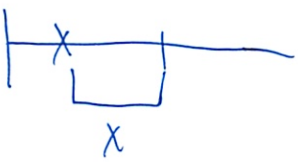
Look at process at time $\frac{1}{4}$

- kinda random incident



But looking backwards - will not be exponential

So can't be erlang of order 2



Need dist of X

- do via derived distribution
- $F_X(x) = P(X \le x)$

$F_X(x) = P(X \le x) = P(1 \text{ arrival in } [\frac{1}{4} - x, \frac{1}{4}] \cap 0 \text{ arrivals } [0, \frac{1}{4} - x])$

poisson w/ 1   poisson w/ 0

$P(\text{One arrival in } [0, 1/4])$

Because subset poisson w/ 1

Or more intuitive $= \begin{cases} 0 & x < 0 \\ \cdots & \\ 1 & x \ge \frac{1}{4} \end{cases}$

(4)

If know had an arrival

- all canidate arrivals are =ly likly
- Some ratio $\frac{x}{1/4}$ ← interval you want, ← time interval

- Continuous uniform prob law

$$= \begin{cases} 0 & x < 0 \\ \frac{x}{1/4} & 0 \le x \le 1/4 \\ 1 & x \ge 1 \end{cases}$$

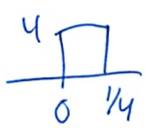
that is CDF

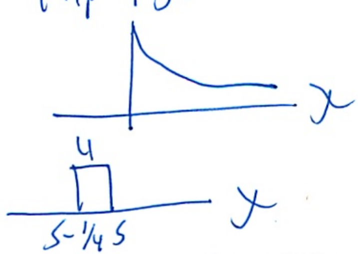
Differentiate to get PDF

$$f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} 4 & x \in [0, 1/4] \\ 0 & \text{otherwise} \end{cases}$$

$S = X + T_1$

$f_S(s) =$ can't just add dist
ind.
must convolve
flip + shift





(12)

$$f_S(s) = \begin{cases} 0 & s < 0 \\ \int_0^s 4 \cdot 21 e^{-21\tau} d\tau & s \in [0, 1/4] \\ \int_{s-1/4}^{s} 4 \cdot 21 e^{-21\tau} d\tau & s > \frac{1}{4} \end{cases}$$

whole thing overlaps now

Shift a little
look at overlap

~~NA~~