MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.041 PROBABILISTIC SYSTEMS ANALYSIS
6.431 APPLIED PROBABILITY
Fall 2010

http://stellar.mit.edu/S/course/6/fa10/6.041/

Included in this opening day handout:

1. General Information (please digest it before next lecture.)

2. Syllabus with lecture subjects and quiz dates.

3. Statement on Collaboration, Honesty, etc.

4. Recitation and tutorial schedule form. Please complete it now. We know that you may have to make changes later. **We need forms returned now**, in order to have initial recitation assignments available by the morning of Thursday, September 9th.

GENERAL INFORMATION

WELCOME to 6.041/6.431! This fundamental subject is concerned with the nature, formulation, and analysis of probabilistic situations. No previous experience with probability is assumed. This course is fun, but also demanding.

6.041 and 6.431: Students intending to take the undergraduate version of the course need to sign up for 6.041, while those intending to take the graduate version should sign up for 6.431, which includes full participation in 6.041, together with some additional homework problems, additional topics, and possibly different quiz and exam questions.

6.041/6.431 has three types of class sessions: lectures, recitations, and tutorials. The lectures and recitations each meet twice a week. In addition, there will be a tutorial once a week, which is not mandatory, but is highly recommended.

LECTURES serve to introduce new concepts. They have an overview character, but also include some derivations and motivating applications. You are expected to attend. Lectures are at 12-1PM on Mondays and Wednesdays in Room 34-101. The first lecture is on Wednesday 9/8.

RECITATIONS meet on Tuesdays and Thursdays, and attendance is mandatory. In recitation, your instructor elaborates on the theory, works through new examples with your participation, and answers your questions about them. The recitation assignments will be made based on the recitation and tutorial schedule forms you complete and turn in immediately at the end of the first lecture. The recitation assignments will be available by 8AM. Thursday 9/9. The recitation assignments will be posted on the course web page (`http://stellar.mit.edu/S/course/6/fa10/6.041/`) for the entire semester.

TUTORIALS for 6.041 and 6.431 meet on Thursday afternoons and Fridays, and will be assigned in response to the recitation and tutorial schedule form, within a few days. In tutorial, you discuss and solve new examples with a little help from your classmates and your instructor. Tutorials are active sessions to help you develop confidence in thinking about probabilistic situations in real time. Tutorials are not mandatory, but are highly recommended. Past students have found them to be very helpful. **The TA who leads the tutorial you are assigned to, will be your first point of contact for questions on the problem sets.** Tutorial assignments will be posted on the course web page with the recitation assignments.

ADVANCED SECTIONS. There may be a possibility for 6.041 students to be assigned to 6.431 recitation sections. If you are interested in slightly faster paced or more advanced recitations and tutorials (while remaining responsible only for 6.041 assignments), please indicate so on the signup sheet.

RECITATION AND TUTORIAL REASSIGNMENT. We try to give everyone their first or second choice in all assignments. Unfortunately with such a large class this is not always possible. If you have a class conflict with your recitation or tutorial assignment you must submit your full class schedule so we can find an assignment which fits your schedule. Recitation and Tutorial assignments are paired, thus a reassignment in one will often require a reassignment in the other. Please submit any reassignment requests by email to the Head TA, Shashank Shekhar Dwivedi (head.ta@mit.edu). To avoid bouncing emails all day, please include your full and complete schedule with any reassignment request.

RECITATION ASSIGNMENTS PROVIDED BY THE REGISTRAR will not be followed. Please disregard them.

FIRST WEEK. There will be **no tutorials** during the first week of classes, but recitations will be held on Thursday September 9. Your updated recitation assignment will be posted by Thursday 8AM at `http://stellar.mit.edu/S/course/6/fa10/6.041/`.

INDIVIDUAL MEETINGS WITH YOUR RECITATION INSTRUCTOR AND TA are encouraged. We want to help! They will both give you their office hours at the first recitation or tutorial meeting. If you have already made a reasonable effort, your instructor or TA will be glad to help you with homework problems, before or after they are due. However, do not expect either of them to work with you if you have not yet carefully read the relevant material in both the lecture handouts and the text.

ADDITIONAL HELP FROM STAFF MEMBERS. Your tutorial TA and your recitation instructor will both have office hours every week. Optional quiz reviews are presented uniformly for the entire class, not for individual sections. Similarly, any supplementary handouts will be identical for all sections.

SPECIAL PERSONAL SITUATIONS. Unforeseen events happen to many of us during the semester. If any are likely to affect your performance, please keep your TA, recitation instructor and/or the Head TA and the lecturers aware of your situation.

ADMINISTRATIVE MATTERS. Recitation and tutorial assignments will be handled by the Head TA, Shashank Shekhar Dwivedi (head.ta@mit.edu). Copies of all material distributed can be found on the course's web site and outside the TA office in 24-312. Graded problem sets will be returned to you in your assigned tutorial. All unclaimed problem sets will be placed outside 24-312 on the metal shelves.

PREREQUISITES. The prerequisite for 6.041 and 6.431 is 18.02, or a year of college level calculus for those with undergraduate degrees from other universities. Students who have not completed the prerequisite with a grade of A, B, C or P may not enroll.

TEXT. The text for this course is *Introduction to Probability* (second edition) by Bertsekas and Tsitsiklis. It is available at the MIT Coop. Solutions to end-of-chapter problems are available at `http://athenasc.com/prob-solved_2ndedition.pdf`. We recommend that you print out these solutions. A few of these problems will be covered in recitation and tutorial. The remaining ones can be used for self-study (for best results, always try to solve a problem on your own, before reading the solution).

Additionally, the following books may be useful as references. They cover many of the topics in this course, although in a different style. You may wish to consult them to get a different perspective on particular topics:
1. A.Drake, *Fundamentals of Applied Probability Theory*
2. S. Ross, *A First Course in Probability*

PROBLEM SET questions are posted on the course website according to the schedule in the course syllabus. PSets are due at the **beginning of lecture** on their respective due date, typically Wednesday. Baskets will be placed outside 34-101, 10 minutes before lecture begins (approximately 11:55 AM), and will remain available until 12:15PM. Be sure to arrive on time the day PSets are due! **Place your solutions in the basket corresponding to your tutorial TA**. Solutions will be available on the course website, shortly after lecture. There will be 11 problem sets handed out this term, with the final PSet not collected. Your worst Pset (out of the 10 collected) will not be taken into account, which essentially allows you to miss on Pset without penalty.

Since we post PSet solutions immediately after the PSets are due, **we do NOT accept any late PSets.** Students who submit a note from Student Support Services will be excused from the appropriate PSet. Please see the head TA if you have further questions regarding this policy.

We grade homework, but often only a small, randomly chosen subset of the problems. We do post detailed solutions on the course website. Your TA is available to discuss your work with you, both before and after it is due. You may encounter difficulty figuring out where your own solution of a homework problem went astray. There are *many* ways to approach most probability problems. Just agreeing with our problem solutions may not explain why your approach didn't work. Please let your instructor or TA help you whenever such issues occur. If the intent of a question on a problem set is unclear, please email your assigned tutorial TA for clarification.

QUIZZES AND EXAMS. There will be two quizzes and a final exam this term. Quiz 1, on Tuesday October 12th, will be given in the evening from 7:30-9:00PM (venue: 54-100). Quiz 2, on Tuesday November 2nd, will be given in the evening from 7:30-9:30PM (venue: 54-100). A comprehensive final exam will be given during finals week, time and place to be determined.

CONFLICT EXAMS: Conflicts with quiz times must be submitted to the Head TA (Shashank Shekhar Dwivedi, head.ta@mit.edu) two weeks prior to the scheduled quiz date. Conflicts for the final are resolved by the Scheduling office.

THE COURSE WEB SITE at `http://stellar.mit.edu/S/course/6/fa10/6.041/` contains a wealth of information – course introduction, announcements, homework assignments and solutions, recitation and tutorial handouts, lecture slides, etc.

STUDY HABITS. In order to get the most out of the course, it is important to not fall behind. It is also important to read the text carefully before attempting to solve the Problem Sets. A very good practice is to review the transparencies handed out at lecture before attending the next lecture or recitation; this way, recitations and tutorials will be much more informative and meaningful.

Make it a point to go to staff office hours if you have any questions or just want to chat about the course; we count on seeing you during the term! Also, it is a good idea to retain a copy of your homework solutions before you turn them in. This lets you compare them with our solutions right away, rather then waiting a week until the graded solutions come back to you.

GRADES will be determined by your work in all aspects of this subject. Final grades are assigned in a meeting by the entire staff. *Your TA is not allowed to discuss likely final grades with you.*

The "formula" that will be used to determine your grade is:
First Quiz: 20%
Second Quiz: 28%
Final: 37%
Homework: 10% (Based on your best 9 out of 10 problem sets)
Attendance & Participation: 5% (Your recitation instructor's and tutorial TA's combined assessment, based primarily on their personal contact with you during recitations and tutorials.)

6.041-6.431 Statement on Collaboration, Honesty, etc.

We encourage working together whenever possible – working out problems in tutorials, discussing and interpreting reading assignments and homework. Talking about the course material is a great way to learn.

Regarding homework, the following is a fruitful (and acceptable) form of collaboration: discuss with your classmates possible approaches to solving the problems, and then have each one fill in the details and write her/his solution **independently**. An unacceptable form of dealing with homework is to copy a solution that someone else has written.

We discourage, but do not forbid, use of materials from prior terms that students may have access to. Furthermore, at the time that you are actually writing up your solutions, these materials must have been set aside; copy-editing from a bible is not acceptable.

At the top of each homework you turn in, we expect you to briefly list all sources of information you used, other than the text, books on reserve for this course, or discussions with 6.041/6.431 staff. A brief note such as "Did homework with John Thompson and Jane Appleby in study group" or "Looked at old bible for Problem 4" would be sufficient. With such a disclosure, there is no penalty or other downside to the use of sources or collaboration. On the other hand, using such sources without reference is plagiarism and is not acceptable.

After a quiz has been returned, we give students a limited amount of time to resubmit their quizzes for regrades if they feel that there is a problem with the grading on their exam. Your new grade can turn out to be higher, lower, or the same as before. (We reserve the right to regrade the entire exam.) If you submit an exam to be regraded, **do not write anything at all on the exam booklet**. Please write a note on a separate sheet of paper. We will reconsider the grade based on the explanation in your note, but TAs are not allowed to discuss the grading with you personally. Any attempt to modify an exam booklet is considered a serious breach of academic honesty. We photocopy a substantial fraction of the exams before they are returned and the probability of catching a change is high.

In general, we expect students to adhere to basic, common sense concepts of academic honesty. Presenting another's work as if it were your own, or cheating in exams will not be tolerated. The appropriate authorities at MIT will be notified in cases of academic dishonesty.

SYLLABUS

(numbers in parentheses indicate textbook sections)

	Date	Topic	due	out
W	9/8	L1: Probability models and axioms (1.1-1.2)		1
M	9/13	L2: Conditioning and Bayes' rule (1.3-1.4)		
W	9/15	L3: Independence (1.5)	1	2
M	9/20	L4: Counting (1.6)		
W	9/22	L5: Discrete rand. variables (r.v's); probability mass functions; expectations (2.1-2.4)	2	3
M	9/27	L6: Discrete r.v. examples; joint PMFs (2.4-2.5)		
W	9/29	L7: Multiple discrete r.v.'s: expectations, conditioning, independence (2.6-2.7)	3	4
M	10/4	L8: Continuous random variables (3.1-3.3)		
W	10/6	L9: Multiple continuous random variables (3.4-3.5)	4	5
M	10/11	*Columbus day – no class*		
T	10/12	**Quiz 1**, 7:30-9:00 pm; covers L1-L7; location TBA		
W	10/13	L10: Continuous Bayes rule; derived distributions (3.6; 4.1)		
M	10/18	L11: Derived distributions; convolution; covariance and correlation (4.1-4.2)	5	6
W	10/20	L12: Iterated expectations; sum of a random number of random variables (4.3; 4.5)		
M	10/25	L13: Bernoulli process (6.1)		
W	10/27	L14: Poisson process – I (6.2)	6	7
T	11/2	**Quiz 2**, 7:30-9:30pm, location TBA, covers up to L12; no class on M 11/1		
W	11/3	L15: Poisson process – II (6.2)		
M	11/8	L16: Markov chains – I (7.1-7.2)	7	8
W	11/10	L17: Markov chains – II (7.3)		
R	11/11	*Veteran's day – no recitation*		
M	11/15	L18: Markov chains – III (7.3)	8	9
W	11/17	L19: Weak law of large numbers (5.1-5.3)		
M	11/22	L20: Central limit theorem (5.4)	9	10
W	11/24	L21: Bayesian statistical inference – I (8.1-8.2)		
M	11/29	L22: Bayesian statistical inference – II (8.3-8.4)		
W	12/1	L23: Classical statistical inference – I (9.1)	10	11†
M	12/6	L24: Classical inference – II (9.1-9.4)		
W	12/8	L25: Classical inference; course overview – III (9.1-9.4)		

Final exam, during finals week

† not to be handed in

Page intentionally left blank.

6.041 First Day

9/8

~~M~~ John Tsitsiklis jnt@mit.edu

starts on time

head TA: Shashank Dwivedi head.ta@mit.edu

lecture slides but need to read book too

register for section from scratch

Quiz 1 20% 10/12

Quiz 2 28% 11/2

Final 38%

HW (9 best of 10) 9%

attendance 5%

(*) model the random aspects of the world to deal w/ it

methology + math same

can be applied in a lot of fields

Analytical framework

Uncertanity

Lecture 1

9/8

6.041 Probabilistic Systems Analysis
6.431 Applied Probability

- Staff
  - Lecturer: John Tsitsiklis, jnt@mit.edu
  - Recitation instructors: Dimitri Bertsekas (6.431), Peter Hagelstein, Ali Shoeb, Vivek Goyal
  - Head TA: Shashank Dwivedi, head.ta@mit.edu
  - Other TAs: Alia Atwi, Uzoma Orji, Sam Zamanian
- Pick up **and read** course information handout
- **Turn in recitation and tutorial scheduling form** (last sheet of course information handout)
- Pick up copy of slides
- `http://stellar.mit.edu/S/course/6/sp10/6.041/`

Coursework

– Quiz 1 (October 12, 7:30-9:00pm)	20%
– Quiz 2 (November 2, 7:30-9:30pm)	28%
– Final exam (scheduled by registrar)	38%
– Weekly homework (best 9 of 10)	9%
– Attendance/participation/enthusiasm in recitations/tutorials	5%

- Pset #1, available on Stellar, due September 15
- **Collaboration policy** described in course info handout
- Text: *Introduction to Probability*, 2nd Edition, D. P. Bertsekas and J. N. Tsitsiklis, Athena Scientific, 2008
  Read the text!

LECTURE 1

not fully certain what will happen

- **Readings:** Sections 1.1, 1.2

Lecture outline

- Probability as a mathematical framework for reasoning about uncertainty
- Probabilistic models
  - sample space
  - probability law

) not completely arbitrary

- Axioms of probability
- Simple examples

denotes

all Sample space Ω

- "List" (set) of possible outcomes
- List must be:
  - Mutually exclusive
  - Collectively exhaustive
- Art: to be at the "right" granularity

you do something
- you don't know what will happen

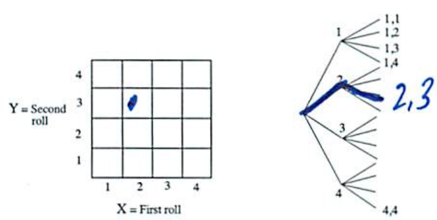
Flip a coin → H or T

When you model: choose which details to explain

Sample space: Discrete example

- Two rolls of a tetrahedral die 4 sided
  - Sample space vs. sequential description

read the bottom of the die



root → leaf
Sequence

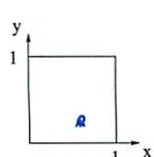
we care about the order
1st + 2nd roll

$(2,3) \neq (3,2)$

← easier

Sample space: Continuous example

$\Omega = \{(x, y) \mid 0 \leq x, y \leq 1\}$



$(\frac{1}{2}, \frac{1}{4})$

throw dart at grid
- infinite precision
- always inside square

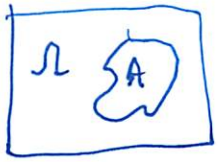
can be any real (not discrete) # inside square

probability of that exact pt: 0
every pt has 0 probability
so specify prob. of a set

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

Probability axioms

- **Event:** a subset of the sample space (half of square)
- Probability is assigned to events



subset = event

Axioms:

1. **Nonnegativity:** $P(A) \geq 0$
2. **Normalization:** $P(\Omega) = 1$ 1 = certainty
3. **Additivity:** If $A \cap B = \varnothing$, then $P(A \cup B) = P(A) + P(B)$

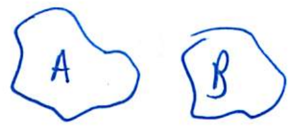
real #

terms

$A \cap B$ = "A and B"

$A \cup B$ = "A or B"

not XOR



$A \cap B = \varnothing$ "disjoined"

$P(A \cup B \cup C) =$

$P((A \cup B) \cup C)$

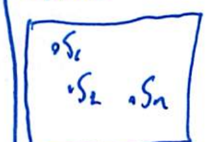
$= P(A \cup B) + P(C)$

$= P(A) + P(B) + P(C)$

∴ induction

for 4, 5, etc sets

Finite



$= P(\{s_1\}) + \dots$

$P(\{s_n\})$

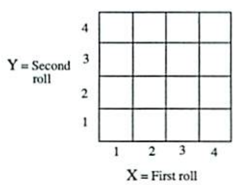
- $P(\{s_1, s_2, \ldots, s_k\}) = P(\{s_1\}) + \cdots + P(\{s_k\})$
  $= P(s_1) + \cdots + P(s_k)$ ← simplified notation
- Axiom 3 needs strengthening
- Do weird sets have probabilities?

$1 = P(\Omega) = P(A \cup A')$

A "complement"

$P(A) = 1 - P(A') \leq 1$

Probability law: Example with finite sample space



- Let every possible outcome have probability 1/16 ← no compelling reason to assume this model will be given in this class
  - $P((X, Y)$ is (1,1) or (1,2)$) =$ 1/8 (2/16)
  - $P(\{X = 1\}) =$ 1/4 (4/16)
  - $P(X + Y$ is odd$) =$ look at pic + try to visualize
  - $P(\min(X, Y) = 2) =$

Stats: is this model correct?

5/16

8/16 = 1/2

Discrete uniform law

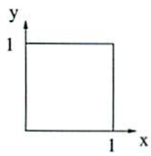
- Let all N outcomes be equally likely → $\frac{1}{N}$ sometimes applies, sometimes not
- Then,

$$P(A) = \frac{\text{number of elements of } A}{\text{total number of sample points}}$$

- Computing probabilities ≡ counting if you can count, you can solve
- Defines fair coins, fair dice, well-shuffled decks

Continuous uniform law

- Two "random" numbers in $[0, 1]$.



- **Uniform** law: Probability = Area
  - $P(X + Y \leq 1/2) = ?$ 1/8
  - $P((X, Y) = (0.5, 0.3)) = 0$

no idea of what will happen
just calculate area of set

point has 0 area

Probability law: Ex. w/countably infinite sample space

- Sample space: $\{1, 2, \ldots\}$
  - We are given $P(n) = 2^{-n}$, $n = 1, 2, \ldots$
  - Find P(outcome is even)

$$P(\{2, 4, 6, \ldots\}) = P(2) + P(4) + \cdots = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots = \frac{1}{3}$$

- Countable additivity axiom (needed for this calculation): high school series
  If $A_1, A_2, \ldots$ are disjoint events, then:

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

sum of individual pieces

Remember!

- Turn in recitation/tutorial scheduling form **now**
- Check Stellar site very late tonight or early tomorrow for recitation assignments and **attend recitation tomorrow**
- Tutorials start next week

Recitation 1
September 9, 2010

1. Give a mathematical derivation of the formula

$$\mathbf{P}((A \cap B^c) \cup (A^c \cap B)) = \mathbf{P}(A) + \mathbf{P}(B) - 2\mathbf{P}(A \cap B).$$

Your derivation should be a sequence of steps, with each step justified by appealing to one of the probability axioms.

2. Problem 1.5, page 54 in the text.

Out of the students in a class, 60% are geniuses, 70% love chocolate, and 40% fall into both categories. Determine the probability that a randomly selected student is neither a genius nor a chocolate lover.

3. A six-sided die is loaded in a way that each even face is twice as likely as each odd face. Construct a probabilistic model for a single roll of this die, and find the probability that a 1, 2, or 3 will come up.

4. Example 1.5, page 13 in the text.

Romeo and Juliet have a date at a given time, and each will arrive at the meeting place with a delay between 0 and 1 hour, with all pairs of delays being equally likely. The first to arrive will wait for 15 minutes and will leave if the other has not yet arrived. What is the probability that they will meet?

G1[†]. Problem 1.13, page 56 in the text. **Continuity property of probabilities.**

(a) Let $A_1, A_2, \ldots$ be an infinite sequence of events that is "monotonically increasing," meaning that $A_n \subset A_{n+1}$ for every n. Let $A = \cup_{n=1}^{\infty} A_n$. Show that $\mathbf{P}(A) = \lim_{n\to\infty} \mathbf{P}(A_n)$. *Hint:* Express the event A as a union of countably many disjoint sets.

(b) Suppose now that the events are "monotonically decreasing," i.e., $A_{n+1} \subset A_n$ for every n. Let $A = \cap_{n=1}^{\infty} A_n$. Show that $\mathbf{P}(A) = \lim_{n\to\infty} \mathbf{P}(A_n)$. *Hint:* Apply the result of the previous part to the complements of the events.

(c) Consider a probabilistic model whose sample space is the real line. Show that

$$\mathbf{P}([0,\infty)) = \lim_{n\to\infty} \mathbf{P}([0,n]) \qquad \text{and} \qquad \lim_{n\to\infty} \mathbf{P}([n,\infty)) = 0.$$

6.041 Recitation 1

9/9

Prof. Dimitri Bertsekas
- textbook author

TA Aliaa Atwi

grad section
since only time

~15 min lecture review
- demo concepts through problems

but having "graduated level problems"
- they have to do extra hw problem
- same quiz + final
- grading different

lots of info on web
since 1965

---

1. We build probablistic models
   - have something that is uncertain
   - dice, coins, applicants, aircraft there, measurement noise
   - using intuition + some restrictions that follows axoms
   - Could have multiple answers
   - Calculate #s associated w/ prob. of certain events
     - precise, only 1 answers
   - math is easy, understanding problems hard

2

Sample Space Ω

- outcomes
- subset of outcomes = events
  - can be entire set Ω
  - or none Ω'

Probabilities ~~are~~ for each event specified

$P(A)$

Must satisfy axioms

- not negitive $P(A) \geq 0$
- additivity If A_1, A_2, etc are disjoined events, then probability of their union $= \sum_{k=1}^{\infty} P(A_k)$
- normalization $P(\Omega) = 1$

$$1 = P(\underbrace{A \cup A'}_{\Omega}) = P(A) + P(A')$$

$$A = \Omega \quad A' = \emptyset$$

$$P(\emptyset) = 1 - \underbrace{P(\Omega)}_{1} = 0$$

- discrete probability law
  - finite # of outcomes $= \{s_1, \ldots, s_n\}$

$P(s_1), \ldots, P(s_n)$

③

For event $A = \{S_1 \ldots\ldots S_{ik}\}$

$$\underline{P}(A) = \underline{P}(S_1) + \ldots + \underline{P}(S_{ik})$$

Not true if probablity is contineous

- because each little point there is 0

Restrictions

$$\sum_{i=1}^{n} \underline{P}(S_i) = 1$$

$$\underline{P}(S_i) \geq 0$$

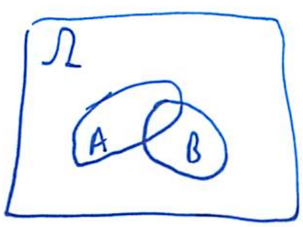
Outcome: 1st dice] label because $(1,2) \neq (2,1)$
2nd dice

$\uparrow$ 1/36

$(2,2) \neq (2,2)$ ← bad notation

↑ same 1/36

Probability for union of 2 events



$P(A \cup B)$

- trick: break it up into disjoined sets

↙ 3 sets

1. $A \cap B'$
2. $A \cap B$
3. $A' \cap B$

] sum

(4)

$$P(A\cup B) = \underbrace{P(A\cap B')}_{P(A) - \cancel{P(A\cap B)}} + \cancel{P(A\cap B)} + \underbrace{P(A'\cap B)}_{P(B) - P(A\cap B)}$$

$$P(A) + P(B) - P(A\cup B)$$

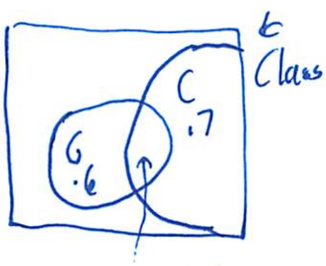
#1 Handout.

$$P(A\cup B) - P(A\cap B).$$

$$P(A) + P(B) - 2P(A\cap B)$$

tip: Think through it w/ a diagram

#2



$$P(G\cap C) = .4$$

$$P(G^c\cap C^c) = ?$$

easy: calculate the complement

$$1 - P(G\cup C)$$

$$1 - (P(G) + P(C) - P(G\cap C)$$

$$1 - .6 + .7 - .4 = .1 = 10\%$$

⑤

Uniform Probability Law

Contineous

- all equally likly

Prob (event)

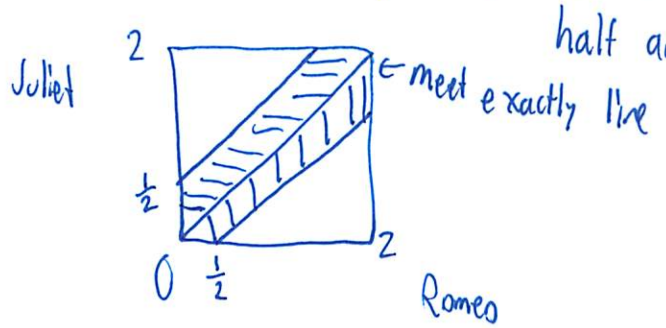
- look at area
- don't add sum of individual points
  - because that is 0
- normalized "volume" of the event
- $\dfrac{\text{Volume of } A}{\text{Volume of } \Omega}$

#34 (modified)

picture/visualize
try to find other ways of thinking

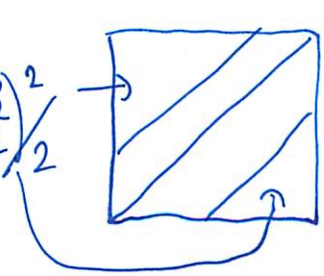
random delay b/w 0 - 2 hrs

half an' hour → meet



now need area since $\frac{A}{4}$ 2 $= 1 - \frac{\left(\frac{3}{2}\right)^2}{4} = 1 - \frac{9}{16} = \frac{7}{16}$

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



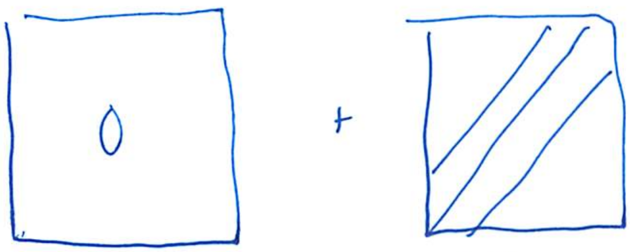
$A = 4 - 2\left(\frac{3}{2}\right)^2/2$

6

now harder

Romeo will be sick on $\frac{1}{2}$ of days + stay home

2 squares



sick

well
same as before

Same as before but $\cdot \frac{1}{2}$

$$\frac{7}{16} \cdot \frac{1}{2} = \frac{7}{32}$$

Grad level problem

- w/ simple proofs

<u>Continuity property of probability</u>

- nested sequence of events $\{A_n\}$
- increasing $A_n \subset A_{n+1}$ for all n

$\bigcup_{n=1} A_n$ <u>limit</u>

2

Show that

$$\lim_{n \to \infty} P(A_n) = P(A)$$

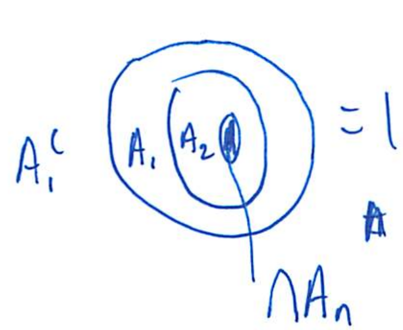
$$\bigcup_{n=1}^{\infty} A_n$$

$$\boxed{B_n = A_n \cap A_{n-1}^c}$$

$$\bigcup_{n=1}^{\infty} B_n = A$$

$$P(A) = \sum_{k=1}^{\infty} P(B_k) = \lim_{n \to \infty} \sum_{k=1}^{n} P(B_k) = \lim_{n \to \infty} P(A_n)$$

$$\underbrace{P\left(\bigcup_{k=1}^{n} B_k\right)}_{A_n}$$



$$\lim P(A_n) = P(\cap A_n)$$

$$A_n \supset A_{n+1}$$

(I think - his handwriting is messy)

Probability Reading Chap 1

9/9

Uncertain situations

Freq of occurance

or subjective belief what will happen

Set = collection of elements

S = set

x = element $\quad x \in S$

$\in$ is a member of

$\emptyset$ = empty set

$S = \{x_1, x_2, \dots x_n\}$ finite # of elements

dice $\{1,2,3,4,5,6\}$

coin $\{H, T\}$

$S = \{x_1, x_2, \dots\}$ infinite # elements

"countable infinite"

$S = \{x \mid x \text{ satisfies } P\}$

$\mid$ such that

P some property

$S = \{k \mid k/2 \text{ is integer}\}$ ← would be even #

$S \subset T$ or $T \supset S$ ← S is a subset of T

- every element of S is an element of T

If $S \subset T$ and $T \subset S$ then they are equal $S = T$

②

Ω = Universal set

all object that could be of interest in a context

we only consider sets $\boxed{S}$ that are subsets of Ω

S^c = Complement

$\{ x \in \Omega \mid x \neq S)$

$\Omega^c = \emptyset$

$\cup$ = Union

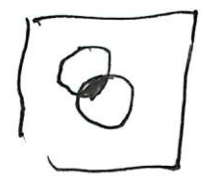
-all elements that belong to one or both

-comp sci ~~AND~~ OR (not XOR)

$\cap$ = intersection

-all elements in both

-comp sci AND



disjoined

-intersection empty

-no common element

partition

-disjoint

-union is S

(x, y) ordered pair

$\mathbb{R}$ = Real #

$\mathbb{R}^2$ = 2D

$\mathbb{R}^3$ = 3D

③

Common Algebra

$S \cup T = T \cup S$ $\qquad$ $S \cup (T \cup U) = (S \cup T) \cup U$

$S \cap (T \cup U) = (S \cap T) \cup (S \cap U)$ $\qquad$ $S \cup (T \cap U) = (S \cup T) \cap (S \cup U)$

$(S^c)^c = S$ $\qquad$ $S \cap S^c = \emptyset$

$S \cup \Omega = \Omega$ $\qquad$ $S \cap \Omega = S$

De Morgan's Law

$$\left(\bigcup_n S_n\right)^c = \bigcap_n S_n^c \qquad \left(\bigcap_n S_n\right)^c = \bigcup_n S_n^c$$

$x \in (\cup_n S_n)^c$

$x \notin \cup_n S_n \Rightarrow$ for every n we have $x \notin S_n$

x belongs to the complement of every S_n

$x \in \cap_n S_n^c$

$(\cup_n S_n)^c \subset \cap_n S_n^c$

Models

- sample space Ω
- probability law

$P(A), \ldots$

sample space = all possible outcomes

↑ subset = event

(4) also it depends on what you define as an experiment

- 3 tosses could = 1 experiment
- could be finite or infinite # of outcomes

sample space must be ~~pr~~ made of mutually exclusive events
and must be collectively exhausitive

- outcome must always be in sample space
- and avoid extra details

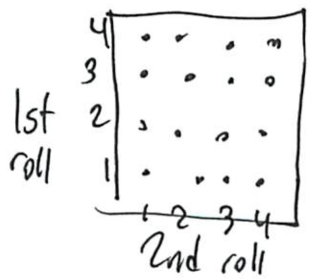
ie if the order of tosses matters (what came before)
then must build sample space

each independent

dependent

Sequential models

Or



Axioms (in lecture)

1. Not negitive
2. Additive
3. Normalization

P(A) like total mass assigned to A when mass spread out over space

⑤ $1 = P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset) = 1 + \underset{0}{P(\emptyset)}$

$$P(A_1 \cup A_2 \cup A_3) = P(A_1 \cup (A_2 \cup A_3))$$
$$= P(A_1) + P(A_2 \cup A_3)$$
$$= P(A_1) + P(A_2) + P(A_3)$$

Discrete Models

- Single coin toss

  $\Omega = \{H, T\}$

  events $\{H, T\}, \{H\}, \{T\}, \emptyset$

  if coin is fair $P(\{H\}) = P(\{T\}) = .5$

  $P(\{H, T\}) = P(\{H\}) + P(\{T\}) = 1$

  $P(\emptyset) = 0$

- 3 coin tosses

  $\Omega = \{HHH, \underline{HHT}, \underline{HTH}, HTT, \underline{THH}, THT, TTH, TTT\}$

  now probability 2 heads exactly $\frac{3}{8}$
  (tip - write set then count how many fit set)

* $P(\{s_1, s_2, \ldots, s_n\}) = P(s_1) + P(s_2) + \ldots + P(s_n)$

$$P(A) = \frac{\#\text{ of elements in } A}{n}$$

⑥ $P\{\text{sum even}\} = \frac{8}{16} = \frac{1}{2}$

$P\{\text{" odd}\} =$ " "

$P\{\text{1st roll} = \text{2nd}\} = \frac{4}{16} = \frac{1}{4}$

$P\{\text{1st roll larger than second}\} = \frac{6}{16} = \frac{3}{8}$

⋮ etc

(tip, draw, then analyze)

Contineous Models

assign probability $b-a$ to subinterval $[a,b]$ of $[0,1]$

- look at its lenght /area

(R+J example in recitation)

Property of Probability Laws

a) $A \subset B$ then $P(A) \le P(B)$

b) $P(A \cup B) = P(A) + P(B) - \underbrace{P(A \cap B)}_{\text{adjust by removing overlap}}$

c) $P(A \cup B) \le P(A) + P(B)$ ← remove overlap (if any)

d) $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) \le \sum_{i=1}^{n} P(A_i)$$

7

Models + Reality

2 stages

1. Construct Probabilistic Model
   - w/ probability law
   - w/ defined sample space
   - Some discresion in model choice
2. Find probability of certain event

(just like I have said under tip)

Full of paradoxes

- Bertrand's Paradox
- due to poor models

1.3 Conditional Probability

- When we have partial info
- ie the sum of 2 dice rolls = 9
  - how likely is that 1st roll = 9
- outcome within some given event B
  - we want likelyhood that belongs to other set A

$$\boxed{P(A \mid B)}$$

ie $P(\underset{\substack{\uparrow \\ \text{probability}}}{\text{outcome } 6} \mid \underset{\substack{\uparrow \\ \text{know}}}{\text{outcome even}}) = \frac{1}{3}$ of die

8

$$P(A \mid B) = \frac{\#\text{ elements of } A \cap B}{\#\text{ elements of } B} = \frac{P(A \cap B)}{P(B)}$$

assume $P(B) > 0$

Some laws

$$P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

$$P(A_1 \cup A_2 \mid B) = \frac{P((A_1 \cup A_2) \cap B)}{P(B)}$$

$$= \frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)}$$

$$= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)}$$

$$= P(A_1 \mid B) + P(A_2 \mid B)$$

example

....

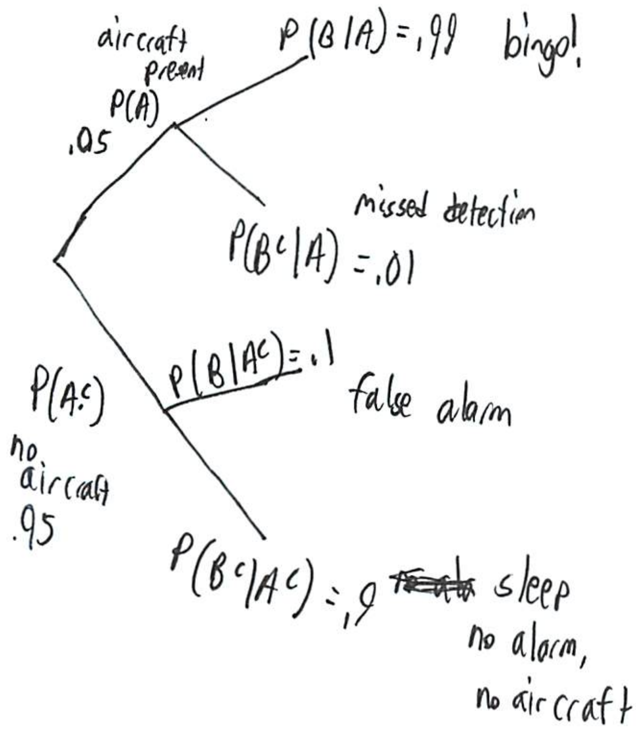
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{3/8}{4/8} = \frac{3}{4}$$

(seems straightforward)

(9)

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

Radar detection



$.05 \cdot .99 = .0495$

$.05 \cdot .01 = .0005$

$.095$

$.855$

We are dealing w/ event that occurs only if several events have occured $A = A_1 \cap A_2 \cap \dots A_n$

So have n branches

Multiplication Rule

$$P(\cap_{i=1}^{n} A_i) = P(A_1)P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \dots P(A_n \mid \cap_{i=1}^{n-1} A_i)$$

(this actually kinda makes sense)

Example: Cards

52 card deck
3 cards drawn
not replaced

* so every triplet =ly likely

find prob all 3 cards not a heart

A_i = the ith card is not a heart

$i = 1, 2, 3$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2)$$

take it step by step

$P(A_1) = \frac{39}{52}$ that are not hearts

if 1st card not a heart

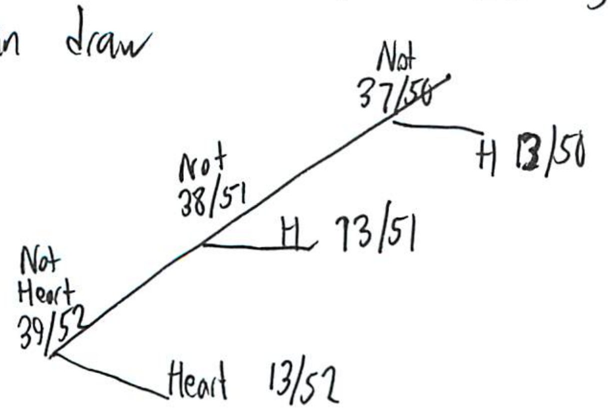
$$P(A_2 \mid A_1) = \frac{38}{51}$$

$$P(A_3 \mid A_1 \cap A_2) = \frac{37}{50}$$

$$P(A_1 \cap A_2 \cap A_3) = \frac{39}{52} \cdot \frac{38}{51} \cdot \frac{37}{50} = \frac{20387}{66300}$$

, so roughly $\frac{1}{3}$

Can draw



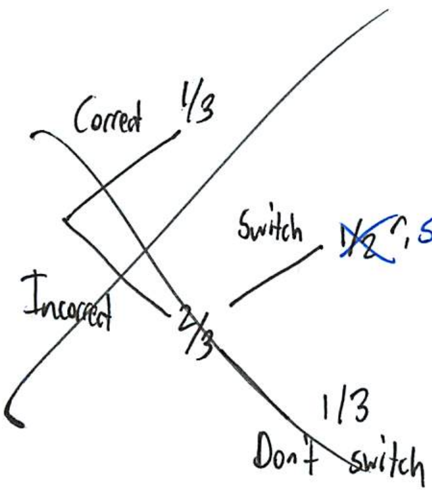
Pratice

~~Example~~: Monty Hall

Prize behind ~~#~~ ~~[illegible]~~ 1/3 doors

Can switch?

- is it worth it?



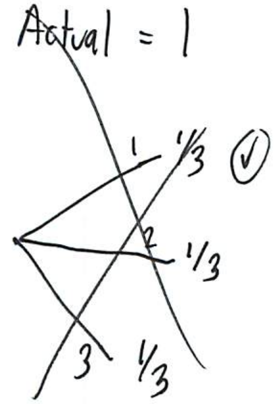
~~1/2~~?, Say $\frac{2}{3}$ - why? is beter

Why could I not figure this out?

- I think wrong tree scnario

Read problem wrong
Friend opens door w/o it

Actual = 1



Actual = 1

- Pick correct $\frac{1}{3}$
  - switch $\frac{1}{2}$ $= \frac{1}{6}$
  - not switch $\frac{1}{2}$ $= \frac{1}{6}$ ✓
- incorrect $\frac{2}{3}$
  - switch $\frac{1}{2}$ $\frac{2}{6}$ ✓
  - not switch $\frac{1}{2}$ $\frac{2}{6}$

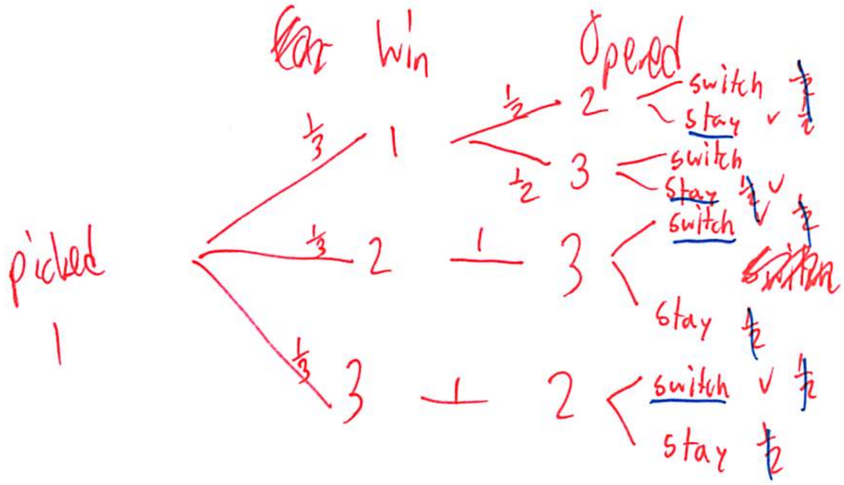
} equal

wrong too

⑫

Online Help

	1	2	3	switch	stay
picked 1	✓				✓
		✓		✓	
			✓	✓	
				↑ $\frac{2}{3}$	↑ $\frac{1}{3}$



$\frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{12}$

$\frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{12}$

$\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$

$\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$

$\frac{3}{6}$ wrong

↑

? This is not probability?

- is an action ???

- no add switching win % ??

$\frac{1}{3} + \frac{1}{3} = \frac{2}{3}$

- stay win %

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

Know what to chart

P Set 1

Thoughts

9/9

I like doing HW on large blank paper

1 to a page

plenty of room to spread out + draw

9/13

LECTURE 2

- **Readings:** Sections 1.3-1.4

Lecture outline

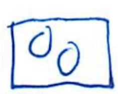
- Review
- Conditional probability
- Three **important** tools:
  - Multiplication rule
  - Total probability theorem
  - Bayes' rule

Review of probability models

- Sample space Ω
  - Mutually exclusive set 1 and only 1 outcome
    Collectively exhaustive
  - Right granularity
- Event: Subset of the sample space

set up probability law that describes situation

- Allocation of probabilities to events
1. $\mathrm{P}(A) \geq 0$
2. $\mathrm{P}(\Omega) = 1$
3. If $A \cap B = \varnothing$, disjoined
   then $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$

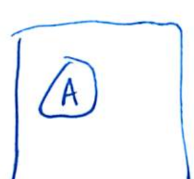


3'. If $A_1, A_2, \ldots$ are disjoint events, then:
   $\mathrm{P}(A_1 \cup A_2 \cup \cdots) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots$

"points" that covers some area

- Problem solving:

steps
  - Specify sample space
  - Define probability law
  - Identify event of interest
  - Calculate...



unified probability model

$P(\square) = 1$ whole square

$P(A) = \text{area}(A)$

$P(\{x,y\}) = 0$ (no area)

↑ single point

$\sum P(\{x,y\}) = 0$

but we just proved 1=0

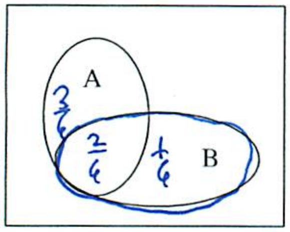
so put in a sequence

– "countable" integrate

*consider discrete + continuous separtly

Conditional probability

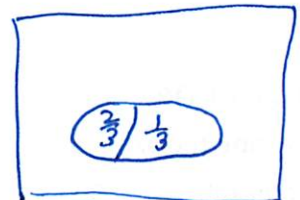
Fix B



| = "given that"

$P(B|B) = 1$

Then prob becomes



maintain proportions

- $\mathrm{P}(A \mid B) =$ probability of A, given that B occurred ← has already occured
  - B is our new universe

entire prob. centered on b

- **Definition:** Assuming $\mathrm{P}(B) \neq 0$,

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

$\frac{2/6}{3/6} = \frac{2}{3}$ same ✓

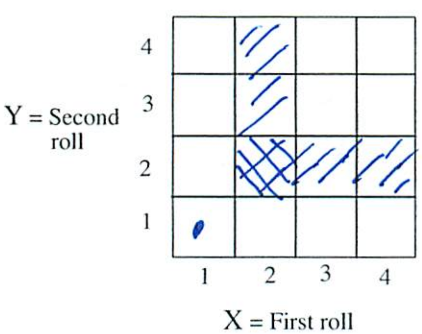
$\mathrm{P}(A \mid B)$ undefined if $\mathrm{P}(B) = 0$

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

$$P(A \cup C) = P(A) + P(C)$$
$$P(A \cup C|B) = P(A|B) + P(C|B)$$

just add the condition to almost any probability

Die roll example



$\frac{1}{16}$ each

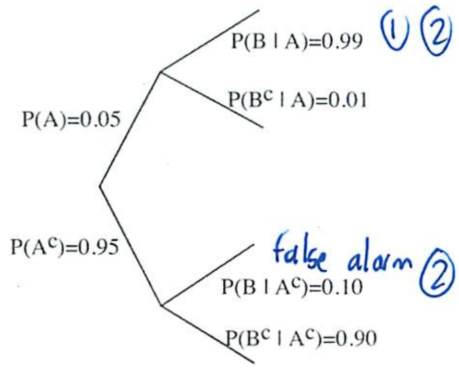
- Let B be the event: $\min(X, Y) = 2$
- Let $M = \max(X, Y)$
- $\mathrm{P}(M = 1 \mid B) =$ 0
- $\mathrm{P}(M = 2 \mid B) =$ $\frac{1/16}{5/16} = \frac{1}{5}$

Models based on conditional probabilities

in textbook?

- Event A: Airplane is flying above
  Event B: Something registers on radar screen

prior probability beliefs



① $P(A \cap B) =$ captured plane $P(A) \cdot P(B|A) = .05 \cdot .99 =$

② $P(B) =$ $.05 \cdot .99 + .95 \cdot .10 =$

③ $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} = \ldots = .34$

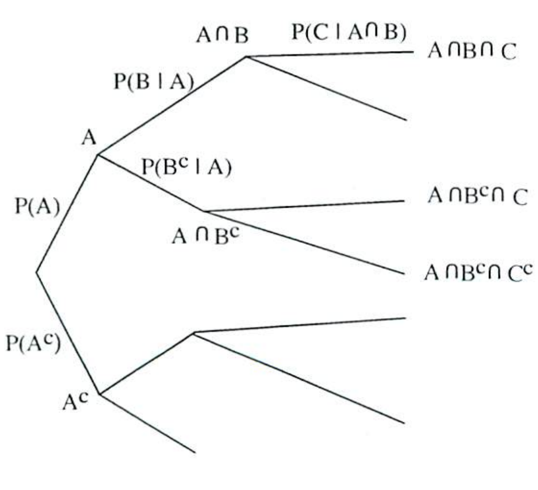
oh cool!

that A happened or not
prob when see something on radar it is a real plane

Only 34% of the time the radar alarms, it is a real plane

Multiplication rule

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

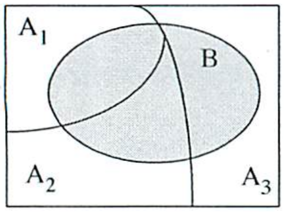


when take intersections, order does not matter

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

Total probability theorem General picture

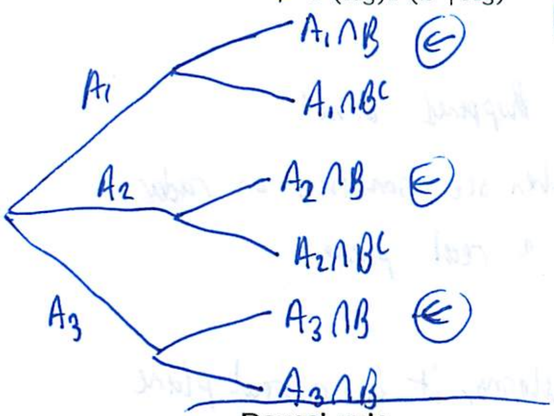
- Divide and conquer
- Partition of sample space into A_1, A_2, A_3 - all 3 compromise complete Ω
- Have $\mathrm{P}(A_i \mid B)$, for every i



- One way of computing $\mathrm{P}(B)$:

$$\begin{aligned}\mathrm{P}(B) = \quad & \mathrm{P}(A_1)\mathrm{P}(B \mid A_1) \\ + \; & \mathrm{P}(A_2)\mathrm{P}(B \mid A_2) \\ + \; & \mathrm{P}(A_3)\mathrm{P}(B \mid A_3)\end{aligned} = \begin{aligned} & P(A_1 \cap B) + \\ & P(A_2 \cap B) + \\ & P(A_3 \cap B)\end{aligned}$$

forming weighted average

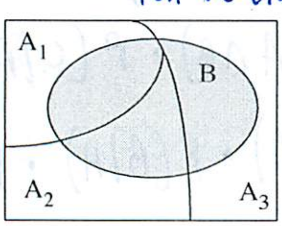


Bayes' rule

- "Prior" probabilities $\mathrm{P}(A_i)$
  - initial "beliefs"
- We know $\mathrm{P}(B \mid A_i)$ for each i
- Wish to compute $\mathrm{P}(A_i \mid B)$
  - revise "beliefs", given that B occurred

like airplane knew $P(B|A)$ wanted to infrene $P(A|B)$

run the other way



general version

$$\begin{aligned}\mathrm{P}(A_i \mid B) &= \frac{\mathrm{P}(A_i \cap B)}{\mathrm{P}(B)} \\ &= \frac{\mathrm{P}(A_i)\mathrm{P}(B \mid A_i)}{\mathrm{P}(B)} \\ &= \frac{\mathrm{P}(A_i)\mathrm{P}(B \mid A_i)}{\sum_j \mathrm{P}(A_j)\mathrm{P}(B \mid A_j)}\end{aligned}$$

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

Recitation 2
September 14, 2010

1. Problem 1.15, page 56-57 in the text.

   A coin is tossed twice. Alice claims that the event of two heads is at least as likely if we know that the first toss is a head than if we know that at least one of the tosses is a head. Is she right? Does it make a difference if the coin is fair or unfair? How can we generalize Alice's reasoning?

2. Problem 1.14, page 56 in the text.

   We roll two fair 6-sided dice. Each one of the 36 possible outcomes is assumed to be equally likely.

   (a) Find the probability that doubles are rolled.

   (b) Given that the roll results in a sum of 4 or less, find the conditional probability that doubles are rolled.

   (c) Find the probability that at least one die roll is a 6.

   (d) Given that the two dice land on different numbers, find the conditional probability that at least one die roll is a 6.

3. Example 1.13, page 29, and Example 1.17, page 33, in the text.

   You enter a chess tournament where your probability of winning a game is 0.3 against half of the players (call them type 1), 0.4 against a quarter of the players (call them type 2), and 0.5 against the remaining quarter of the players (call them type 3). You play a game against a randomly chosen opponent.

   (a) What is the probability of winning?

   (b) Suppose that you win. What is the probability that you had an opponent of type 1?

4. Example 1.12, page 27 in the text.

   The Monty Hall Problem. This is a much discussed puzzle, based on an old American game show. You are told that a prize is equally likely to be found behind any one of three closed doors in front of you. You point to one of the doors. A friend opens for you one of the remaining two doors, after making sure that the prize is not behind it. At this point, you can stick to your initial choice, or switch to the other unopened door. You win the prize if it lies behind your final choice of a door. Consider the following strategies:

   (a) Stick to your initial choice.

   (b) Switch to the other unopened door.

   (c) You first point to door 1. If door 2 is opened, you do not switch. If door 3 is opened, you switch.

   Which is the best strategy? ? I spent 30 min on this a few nights ago

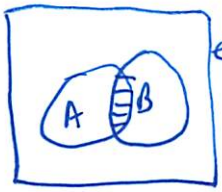
   - lets see if I can do this now

   - or he can draw it better

6.041 Recitation 2

9/14

Conditional Probability



$\leftarrow \Omega$

any set of outcome = event

Prob. Law $\rightarrow$ P(event)

rules: $P(A) \geq 0 \quad P(\Omega) = 1$

$$P\left(\bigcup A_i\right) = \sum_{n=1}^{\infty} P(A_i)$$

(countable collection of disjoined events)

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \text{ if } P(B) > 0$$

(just like we went over in lecture)

Case of Uniform Prob. Law

$\Omega = \{S_1, S_2, \ldots, S_n\}$ $\quad P(S_i) = \frac{1}{n}$ $\leftarrow$ all equally likly

$$P(B) = \frac{\#\text{ of outcomes in } B}{\text{total }\#\text{ of outcomes } (n)}$$

$$P(A \mid B) = \frac{\dfrac{\#\text{ outcomes } A \cap B}{\#\text{ outcomes } \Omega}}{\dfrac{\#\text{ outcomes in } B}{\#\text{ outcomes } \Omega}} = \frac{\#\text{ outcomes } A \cap B}{\#\text{ outcomes } B}$$

More generally

$$P(A \mid B) = \frac{\sum_{S_i \in A \cap B} P(S_i)}{\sum_{S_i \in B} P(S_i)}$$

(3)

Chess Game Example

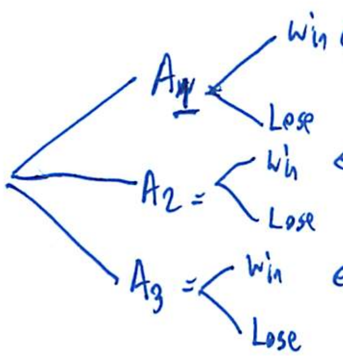
- Opponents of 3 types 1, 2, 3

$P(A_1) = .5$

$P(A_2) = .25$

$P(A_3 = .25$

↑ prob of picking oponant



$P(A_1 \cap win) = .5 \cdot .3$

$P(A_2 \cap win) = .25 \cdot .4$

$P(A_3 \cap win) = .25 \cdot .5$

total chance of winning

↓ add

$= .375 = P(B)$

P of winning for each

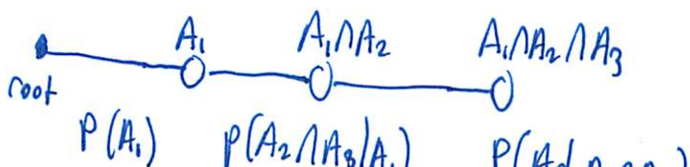
$P(B|A_1) = .3$

$P(B|A_2) = .4$

$P(B|A_3) = .5$

$\boxed{\text{* called multiplication rule}}$

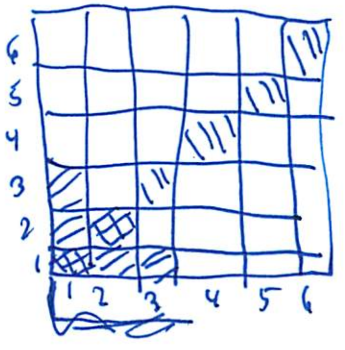
$P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2 \cap A_3 | A_1) \cdot P(A_3 | A_1 \cap A_2) \cdot$



- eaiser to do through examples

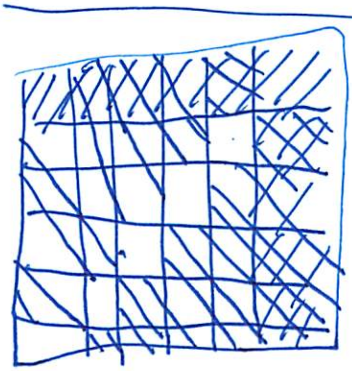
②

Examples Roll 2 Dice



A: Doubles $P(A) = \frac{6}{36} = \frac{1}{6}$

B: Sum ≤ 4 $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{6} = \frac{1}{3}$



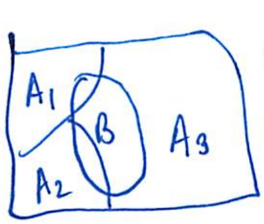
A: at least one roll is 6

B: no doubles

$$P(A \mid B) = \frac{10}{30} = \frac{1}{3}$$

$P(A \cap B \mid A) = ?$ ← will be larger, smaller denominator

$P(A \cap B \mid A \cup B) = ?$



~~$P(B) = P(B \mid A_1) + P(B \mid A_2) + P(B \mid A_3)$~~ this does not work, see

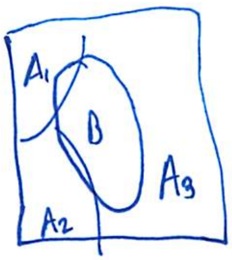
$$P(B) = P(A_1) \cdot P(B \mid A_1) + P(A_2) \cdot P(B \mid A_2) + P(A_3) \cdot P(B \mid A_3)$$

– Seems dumb, but helps you solve problem → break it down

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B)$$ ← get right notation

Many applications → chess game

Bayer's Rule



Calculate $P(A_1 | B) \dots P(A_n | B)$

$A_{1,2,3}$ = causes

B = observation

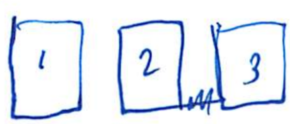
Find likly hood of each cause given observation

$$P(A_1 | B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(A_i) P(B|A_i)}{\sum_j P(A_j) P(B|A_j)}$$

Observe B — $P(A_1|B)$, $P(A_2|B)$, $P(A_3|B)$

which to pick?
- could use max likehood

Monty Hall



One door has prize

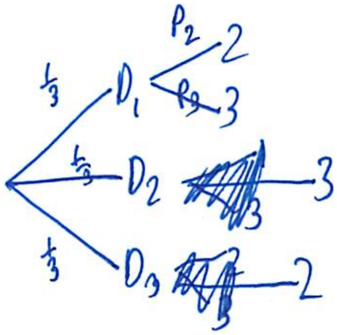
1. Pick a door (all equally likly)
2. Host opens a door w/o prize (one of the 2 you did <u>not</u> pick)
3. Host offes to switch?

$P(\text{never switch}) = \frac{1}{3}$ at start have choice b/w 3

$P(\text{always switch}) = \frac{2}{3}$ now you know prize is behind 2 doors

(5)

assuming ~~...~~ prize behind 1



D_i - prize is behind door i

O_i - Host opens door i

We don't know P_2, P_3 unless we have probabilistic model

- which does host open?

<u>Strategies</u>

(1) Never switch $= P(D_1 \cap O_2) + P(D_1 \cap O_3) = \frac{1}{3}$

$\frac{1}{3} \cdot P_2 + \frac{1}{3} \cdot P_3$ $(P_2 + P_3 = 1)$

(2) Always switch $= P(D_2) + P(D_3) = \frac{2}{3}$

$\frac{1}{3} + \frac{1}{3}$

try to solve it in the 3 step process like they do

(3) Switch if and only if door 3 open $= P(D_1 \cap O_2) + P(D_3) = \frac{1}{2}$

$\frac{1}{3} \cdot P_2$ $\quad \frac{1}{3}$

assume ~~guess~~ $\frac{1}{2}$

if no assume $\frac{1}{3}(P_2 + 1)$

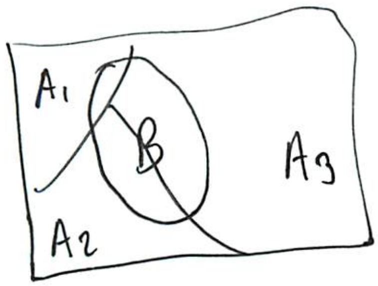
$\in \left[\frac{1}{3}, \frac{2}{3}\right]$ within

Could also think abo if host is not random/even/neutral

Prob. Reading

9/14

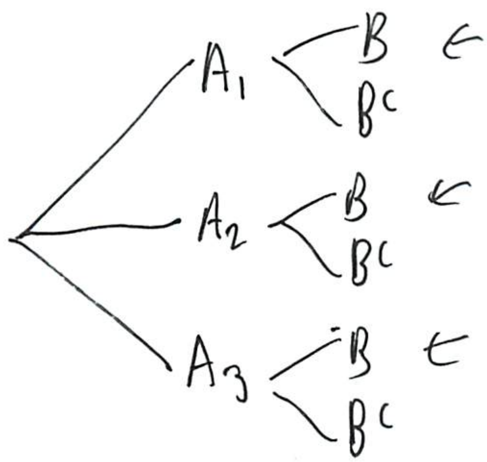
1.4 Total Prob. Theorm ~~+ Bayes' Rule~~ (did in class today)



$$P(B) = P(A_1 \cap B) + \dots + P(A_n \cap B)$$

add each piece

$$= P(A_1)\,P(B|A_1) + \dots + P(A_n)\,P(B|A_n)$$



(learn how to write/ express things)

Inference + Bayes's Rule

A_1, A_2, A_3 = disjoint events, partition of space

$P(A_i) > 0$ for all i (not a tiny point)

Then for any event B (where $P(B) > 0$)

$$P(A_i|B) = \frac{P(A_i)\,P(B|A_i)}{P(B)}$$

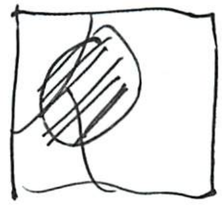
$$= \frac{P(A_i)\,P(B|A_i)}{P(A_1)\,P(B|A_1) + \dots + P(A_n)\,P(B|A_n)}$$

So what??

②

Ok: Bayes rule used to infrence <u>which</u> of the causes is real
- aka caused the effect

A_1, etc are causes
B is the effect



- (know from class, just need to know math)

~~so~~ so $P(B|A_1)$ is prob of B in A_1 - # of cancer pts w/ red hair
prior prob.

We want $P(A_1|B)$ - ~~A~~ prob ~~cancer~~ red hair causes cancer
posterior prob.

<u>Independence</u>

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

- B does not matter

"A is independent of B"

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

disjoined events are never independent

<u>Conditional Independence</u>

$\Updownarrow$ one does not imply other

- ~~items~~ events that are independent based on a condition

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

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

③

Independence of a collection of events

$$\overset{\text{↓?? symbol?}}{}$$

$P\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} P(A_i)$ for every subset of S of $\{1, 2, \dots n\}$

$P(A_1 \cap A_2) = P(A_1) P(A_2)$ ← pairwise independent

; etc each pair independent

$P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2) P(A_3)$ ← does not imply

occurance or nonoccurance does not matter

Reliability

- in complex system, eaiser to consider events unlinked
- ie network connectivity
- break system into subsystems
  - connected in series or parallel

$P(\text{series}) = p_1 p_2 \dots p_n$

$P(\text{parallel}) = 1 - (1-p_1)(1-p_2) \dots (1-p_n)$

- succeeds if any one branch fails

Independent Trials + Binomial Probabilities

- independent trials - a sequence of identical but independent steps
  - ie rolls of a die, roullette wheel

(9) Bernoulli trials - exactly 2 possible outcomes

- ie coin flip $\{H, T\}$
- can do it n times $\rightarrow A_i$ = ith toss
- any sequence has

$$p^k (1-p)^{3-k} \text{ chance}$$

$$\left.\begin{array}{l} k \text{ heads} \\ n-k \text{ tails} \end{array}\right) p^k (1-p)^{n-k}$$

p_k = P(k heads come up in n tosses)

$$\underbrace{p(k)}_{\text{binomial probabilities}} = \underbrace{\binom{n}{k}}_{} p^k (1-p)^{n-k}$$

binomial probabilities

$\hookrightarrow$ # of distinct n-toss sequences that contain k-heads

"n choose k"

bianomial coefficients

$\downarrow$

Counting argument (section 1.6)

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \quad k = 0, 1, \ldots, n$$

where for any (+) int i

$$i! = 1 \cdot 2 \cdot \ldots (i-1) \cdot i \qquad (0! = 1)$$

binomial formula (must add to 1)

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$$

5

<u>example</u> grade of service for ISP - prob that customers w/o service

$$\sum_{k=c+1}^{n} P(k)$$

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$n = 100$ customers

$p = .1$ prob that online

$c = 15$ endpoints

$.0399$ = prob one can't get endpoint

1.6 Counting

a) When Ω has finite equally likely events

$$P(A) = \frac{\# \text{ elements in } A}{\# \text{ elements in } \Omega}$$

b) When we want to calc prob of event A w/ finite outcomes each of which has known prob p

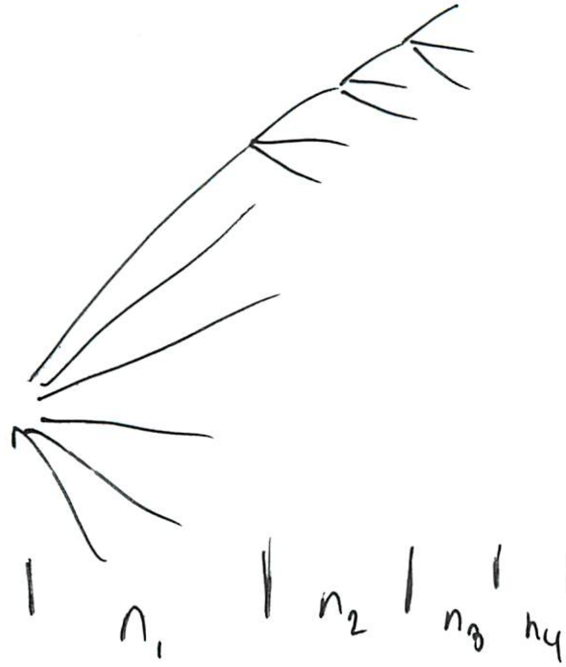
$$P(A) = p\,(\# \text{ elements of } A)$$

it is frequently challenging → combinatorics

<u>Counting Principle</u>

- divide & conquer
- break into trees + stages
  - express stages in ordered pairs

⑥



(a) There are n_1 possible results at 1st stage

(b) For every possible result at 1st stage, n_2 possible results in n_2 stage

(c) Total k stages $\rightarrow n_1 n_2 \cdots n_n$

~~How many distinct phone~~

Permutation - order of selection matters
Combination - order does not matter
partition - collection of n objects into multiple subsets

k - permutations

n - objects
k integer $k \leq n$ - want # of diff ways can pick k out of these n objects & arrange in sequence - order matters

1st: n
2nd: $n-1$ (one object gone)
3rd: $n-2$
$\vdots \quad \vdots$

(7)

last: $n-(k-1)$

\# of possible sequences

$$n(n-1)\cdots n(-k+1) = \frac{n(n-1)\cdots(n-k+1)(n-k)\cdots 2\cdot 1}{(n-k)\cdots 2\cdot 1}$$

need to get a lot better at this type of stuff

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

When $k=n$ simply "permutations"

$$n(n-1)(n-2)\cdots 2\cdot 1 = n!$$

Combinations

n - people want to form a committe of k
but here order does not matter
- remove "duplicates"

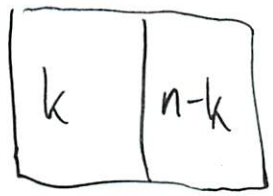
$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

leads to binomial formula (back in 1.5)

(8)

Partitions

Combos can be viewed as partition of the set of two



now can generalize more than 2 subsets

n- elements

$n = \text{sum}(n_1, n_2 \ldots n_r)$

- partition into r disjoint subsets with ith subset containing n_i elements

-?picture?

$\binom{n}{n_1}$ ways of forming 1st subset

after that $\cancel{(n-n_1}$ $\binom{n-n_1}{n_2}$

So $\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-\cdots-n_{r-1}}{n_r}$

$$\frac{n!}{n_1!\,(n-n_1)!}\cdot\frac{(n-n_1)!}{n_2!\,(n-n_1-n_2)!}\cdots\frac{(n-n_1-\cdots-n_{r-1})!}{(n-n_1-\cdots-n_{r-1}-n_r)!\,n_r!}$$

terms cancle

$$= \frac{n!}{n_1!\,n_2!\ldots n_r!}$$

→ multinomial coefficent → $\binom{n}{n_1, n_2 \ldots n_r}$

I don't get what a partiton is

9 Summarry

(lot of repeating stuff)

3 methods of solving problems

lays it out nice

a) Counting - # outcomes finite
each equally likley
Count # elements / area

b) sequential - orders matters
count branches along a tree
multiplication rule - multiply probabilities

c) divide + conquere - prob $P(B)$ are obtained from $P(B|A_i)$
use total prob theorm

$$= P(A_1 \cap B) + \dots P(A_i \cap B)$$
$$= P(A_1) P(B|A_1) + \dots P(A_i) P(B|A_i)$$

Michael Plasmeier

← Staple preffered

7/10

9/9

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Spring 2010)

Problem Set 1

Due: September 15, 2010

1. Express each of the following events in terms of the events A, B and C as well as the operations of complementation, union and intersection:

   (a) at least one of the events A, B, C occurs;

   (b) at most one of the events A, B, C occurs;

   (c) none of the events A, B, C occurs;

   (d) all three events A, B, C occur;

   (e) exactly one of the events A, B, C occurs;

   (f) events A and B occur, but not C;

   (g) either event A occurs or, if not, then B also does not occur.

   In each case draw the corresponding Venn diagrams.

2. You flip a fair coin 3 times, determine the probability of the below events. Assume all sequences are equally likely.

   (a) Three heads: HHH

   ✓ (b) The sequence head, tail, head: HTH

   (c) Any sequence with 2 heads and 1 tail

   (d) Any sequence where the number of heads is greater than or equal to the number of tails

3. Bob has a peculiar pair of four-sided dice. When he rolls the dice, the probability of any particular outcome is proportional to the sum of the results of each die. All outcomes that result in a particular sum are equally likely.

   ✓ (a) What is the probability of the sum being even?

   (b) What is the probability of Bob rolling a 2 and a 3, in any order?

4. Alice and Bob each choose at random a number in the interval $[0, 2]$. We assume a uniform probability law under which the probability of an event is proportional to its area. Consider the following events:

   A : The magnitude of the difference of the two numbers is greater than $1/3$.

   ✓ B : At least one of the numbers is greater than $1/3$.

   C : The two numbers are equal.

   D : Alice's number is greater than $1/3$.

   Find the probabilities $\mathbf{P}(B)$, $\mathbf{P}(C)$, and $\mathbf{P}(A \cap D)$.

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

---

5. Mike and John are playing a friendly game of darts where the dart board is a disk with radius of 10in.

   Whenever a dart falls within 1in of the center, 50 points are scored. If the point of impact is between 1 and 3in from the center, 30 points are scored, if it is at a distance of 3 to 5in 20 points are scored and if it is further that 5in, 10 points are scored.

   Assume that both players are skilled enough to be able to throw the dart within the boundaries of the board.

   Mike can place the dart uniformly on the board (i.e., the probability of the dart falling in a given region is proportional to its area).

   (a) What is the probability that Mike scores 50 points on one throw?

   (b) What is the probability of him scoring 30 points on one throw?

   (c) John is right handed and is twice more likely to throw in the right half of the board than in the left half. Across each half, the dart falls uniformly in that region. Answer the previous questions for John's throw.

6. Prove that for any three events A, B and C, we have

$$\mathbf{P}(A \cap B \cap C) \geq \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - 2.$$

G1[†]. Consider an experiment whose sample space is the real line.

   (a) Let $\{a_n\}$ be an increasing sequence of numbers that converges to a and $\{b_n\}$ a decreasing sequence that converges to b. Show that

$$\lim_{n \to \infty} \mathbf{P}([a_n, b_n]) = \mathbf{P}([a, b]).$$

   Here, the notation $[a, b]$ stands for the closed interval $\{x \mid a \leq x \leq b\}$. *Note:* This result seems intuitively obvious. The issue is to derive it using the axioms of probability theory.

   (b) Let $\{a_n\}$ be a decreasing sequence that converges to a and $\{b_n\}$ an increasing sequence that converges to b. Is it true that

$$\lim_{n \to \infty} \mathbf{P}([a_n, b_n]) = \mathbf{P}([a, b])?$$

Note: You may use freely the results from the problems in the text in your proofs.

Michael Plasmeier   6.041 Pset 1   9/9

1. Express + draw

office hrs: not probability of, event of

a) $\cancel{P}(A \cup B \cup C)$

b) ~~P(A ∪ B ∪ C)~~ see next sheet

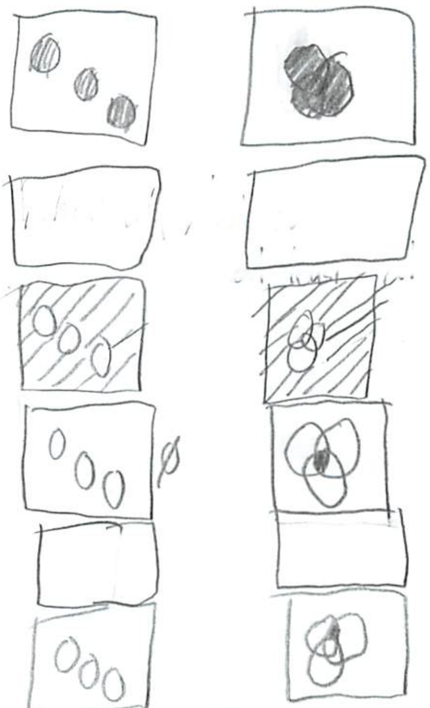
c) $1 - \cancel{P}(A \cup B \cup C)$

d) $\cancel{P}(A \cap B \cap C)$

e) see next sheet

f) $\cancel{P}(A \cap B)$: not C

g) see next sheet

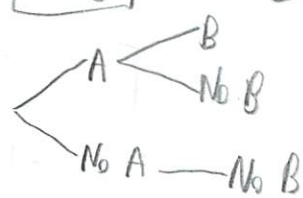


$\emptyset$

: how at most, XOR

or $P(A^c \cap B^c \cap C^c)$

: how exact 1



: how Venn

* research conditional w/ |

$\cancel{P}(A|B)$ probability of A given B

not present = A^c

1b

Office Hrs Aliaa

9/13

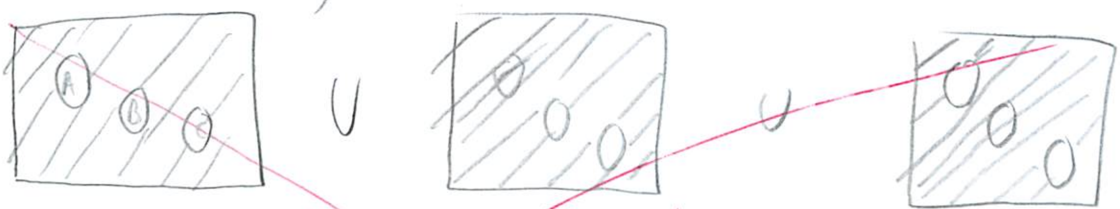
1b. or $= \cup$

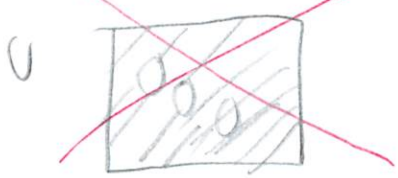
and $= \cap$

tip does not have to be most compact or 1 statement

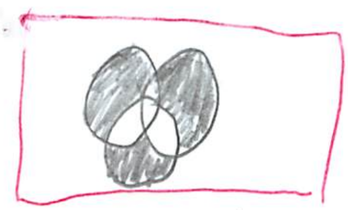
not probability of:

$\cancel{P}(A^c \cap B^c \cap C^c) \cup \cancel{P}(A \cap B^c \cap C^c) \cup \cancel{P}(A^c \cap B \cap C^c)$
$\cup \cancel{P}(A^c \cap B^c \cap C)$

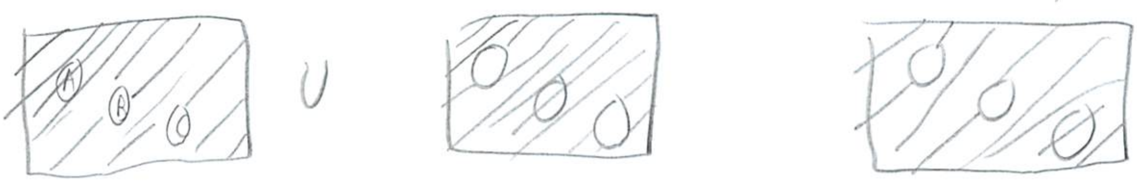




draw 1 diagram



1e. $\cancel{P}(A \cap B^c \cap C^c) \cup \cancel{P}(A^c \cap B \cap C^c) \cup \cancel{P}(A^c \cap B^c \cap C)$



↑ but thats not at most 1 even
– still confused

1g. $\cancel{P}(A \cap B) \cup \cancel{P}(A \cap B^c) \cup \cancel{P}(A^c \cap B^c)$

how can put it on graph
– multiple cases that work

(2)

2. Coin 3 times

$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

a) $\frac{1}{8}$ +0.5

b) $\frac{1}{8}$ +0.5

c) $\frac{3}{8}$ +0.5

+0.5

d) $\frac{4}{8} = \frac{1}{2}$

2/2

(3)

3. Bob → peculiar pair dice, 4 sideds

Outcome proportional to sum of results of each die

1st	4	5	8	12	16
	3	4	6	9	12
	2	3	4	6	8
	1	2	3	4	5
		1	2	3	4
			2nd		

So sum of squares = 107

a) $\dfrac{2+4+4+6+8+4+6+12+8+12+16}{107} = \dfrac{82}{107} =$ correct?

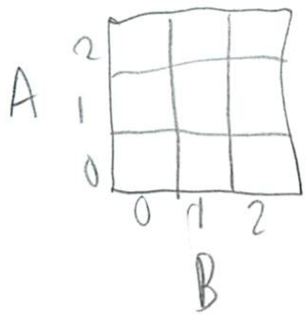
b) can be (2,3) or (3,2)

$\dfrac{6+6}{107} = \dfrac{12}{107}$

(4)

4. Alice + Bob pick # [0,2]

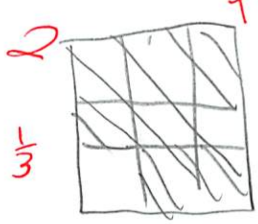
- probability proportional to area



events

+1 Right idea

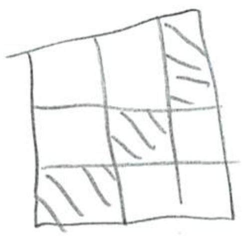
P(B)



$\cancel{\frac{8}{9}}$ (−0.5)

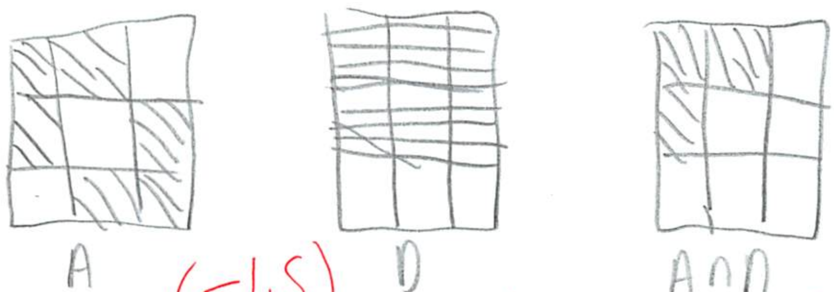
(−1) See solution

P(C)



$\frac{3}{9} = \frac{1}{3}$

P(A ∩ D)
?
AND



$\frac{3}{9} = \frac{1}{3}$

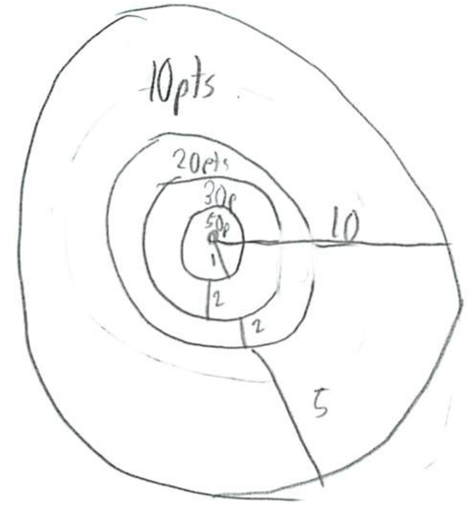
(−1.5) See solution.

We are in cont. ~~time~~. space.

1/4

(5)

5, Mike + John darts $r = 10''$



M places uniformly based on area

total $A = \pi r^2 = 100\pi$

a) 50pts = $1\pi \rightarrow \frac{1}{100}$ +1

b) 30pts = $9\pi - 1\pi = \frac{8}{100}$ +1

20pts = $25\pi - 8\pi = \frac{17}{100}$

10pts = $100\pi - 25\pi = \frac{75}{100}$

oh actually the straight forward way

c) John is 2x as likely to throw to right

? first intuition → same? since left + right same

Mark

$\frac{1}{2}$ $\frac{1}{2}$ $\times 2$ = would = 1 - can't be right

or $\frac{1}{3}$ $\frac{2}{3}$

would be twice as liky

but left and right same points

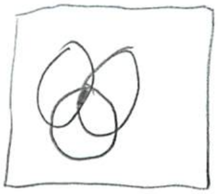
$\frac{1}{100}$, $\frac{8}{100}$ +2

4/4

(6) solved on page (6e)

6. Prove that for any 3 events A, B, C have

$$P(A \cap B \cap C) \geq P(A) + P(B) + P(C) - 2$$



must be all 3

? but how want us to prove?

thought we did this in recitation, but know

~~Well $P(A) + P(B) + P(C)$ must always be < 1~~

no can overlap

Let's say $P(A) = .7$ $\quad P(A \cap B \cap C) \geq .1$

$P(B) = .7$

$P(C) = .7$

? and makes sense w/ overlapping coverage

But how to write a proof

- I never learned

wp: Called Bonferroni einequalities

$$P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$$

? but that is union

if disjointed sets, then it is equal

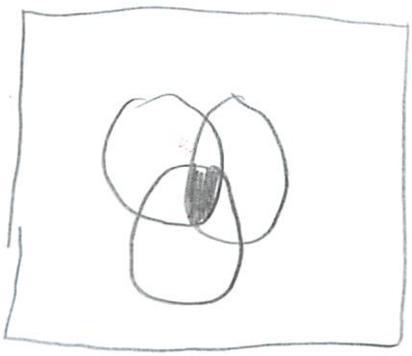
Similar to inclusion - exclusion principal

(6b)

Aliaa

6 Office Hours .

$P(A \cap B \cap C) \geq P(A) + P(B) + P(C) - 2$



$\Omega = 1 = P(A \cup A^c)$ disjoined, just add

$= P(B \cup B^c)$

$= P(C \cup C^c)$

$= P(A) + P(A^c) = P(B) + P(B^c) = P(C) + P(C^c)$

$P(A \cup B \cup C) = P(A) + P(B) + P(C)$?

– but we have double counted on right # 1,2,3

4 triple counted

So need to subtract things now

$- P(A \cap B) - P(B \cap C) - P(A \cap C)$

4 was counted 3 times

but now 0 times, add it back

$+ P(A \cap B \cap C)$

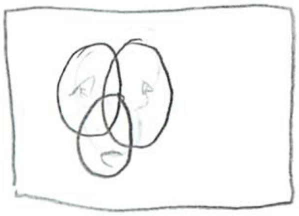
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$

Now need to switch sides

$P(A \cap B \cap C) = P(A \cup B \cup C) - P(A) - P(B) - P(C) + P(A \cap B) + P(B \cap C) + P(A \cap C)$

6C

6. Aliaa OH Restart



$$P(A\cap B\cap C) \geq P(A)+P(B)+P(C)-2$$

$$= 1 - \Big(P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + 2(ABC) + P((A\cup B\cup C)^c) \Big)$$

← student's attempt rewritten on bd

$$P(A\cap B\cap C) = P[A^c \cup B^c \cup C^c]^c$$

$$= 1 - P[A^c \cup B^c \cup C^c]$$

$$= 1 - [P(A^c) + P(B^c) + P(C^c)]$$

$$- P(A^c\cap B^c) - P(B^c \cap C^c) - P(A^c\cap C^c)$$

$$+ P(A^c \cap B^c \cap C^c)$$

6d

One Student's attempt rewritten

$$= 1 - \Big(P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C)$$
$$+ 2(A \cap B \cap C) + P[A^c \cap B^c \cap C^c] \Big)$$

but still not inequality

(6e)

Alina OH Reset 3:

want $P(A \cap B \cap C)$

so $1 - P(A \cap B \cap C) = P(A \cap B \cap C)^c$

$= P(A^c \cup B^c \cup C^c)$ Demorgan

$\boxed{\begin{array}{l} P(A \cap B)^c = P(A^c \cup B^c) \\ P(A \cup B)^c = P(A^c \cap B^c) \end{array}}$

$\leq P(A^c) + P(B^c) + P(C^c)$

can remove pieces

or just say it will be $\leq$

~~minus so flip inequality~~

$= (1 - P(A)) + (1 - P(B)) + (1 - P(C))$

moving 1 over

$P(A \cap B \cap C) \geq P(A) - 1 + P(B) - 1 + P(C) - \cancel{1} + \cancel{1}$

$\geq P(A) + P(B) + P(C) - 2$

left something out to not make a mess

- if we had all - then it would of course $<$
- but we have a plus
- but the plus we know is smaller than all of the pieces subtracting like on pg (6b)

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

Problem Set 1: Solutions
Due: September 15, 2010

1. (a) $A \cup B \cup C$

   (b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C) \cup (A^c \cap B^c \cap C^c)$

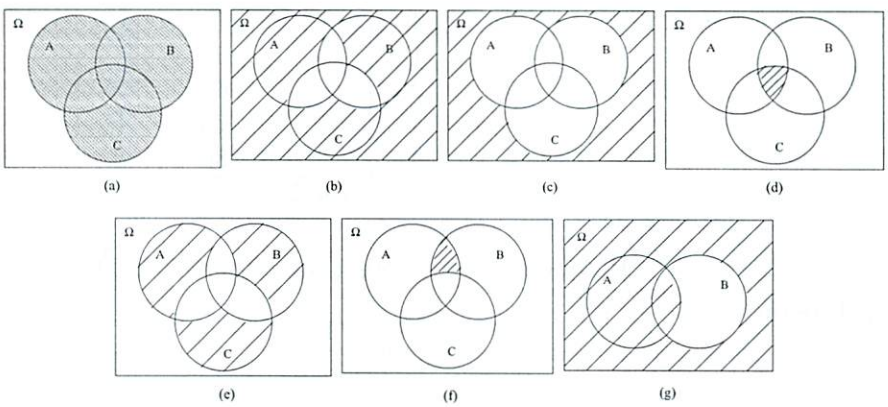
   (c) $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$

   (d) $A \cap B \cap C$

   (e) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

   (f) $A \cap B \cap C^c$

   (g) $A \cup (A^c \cap B^c)$



2. Since all outcomes are equally likely we apply the discrete uniform probability law to solve the problem. To solve for any event we simply count the number of elements in the event and divide by the total number of elements in the sample space.

   There are 2 possible outcomes for each flip, and 3 flips. Thus there are $2^3 = 8$ elements (or sequences) in the sample space.

   (a) Any sequence has probability of 1/8. Therefore $\mathbf{P}(\{H, H, H\}) = \boxed{1/8}$.

   (b) This is still a single sequence, thus $\mathbf{P}(\{H, T, H\}) = \boxed{1/8}$.

   (c) The event of interest has 3 unique sequences, thus $\mathbf{P}(\{HHT, HTH, THH\}) = \boxed{3/8}$.

   (d) The sequences where there are more heads than tails are $A : \{HHH, HHT, HTH, THH\}$. 4 unique sequences gives us $\mathbf{P}(A) = \boxed{1/2}$.

3. The easiest way to solve this problem is to make a table of some sort, similar to the one below.

Die 1	Die 2	Sum	P(Sum)
1	1	2	2p
1	2	3	3p
1	3	4	4p
1	4	5	5p
2	1	3	3p
2	2	4	4p
2	3	5	5p
2	4	6	6p
3	1	4	4p
3	2	5	5p
3	3	6	6p
3	4	7	7p
4	1	5	5p
4	2	6	6p
4	3	7	7p
4	4	8	8p
		Total	80p

$$\mathbf{P}(\text{All outcomes}) = 80p \text{ (Total from the table)}$$

and therefore

$$p = \frac{1}{80}$$

(a)

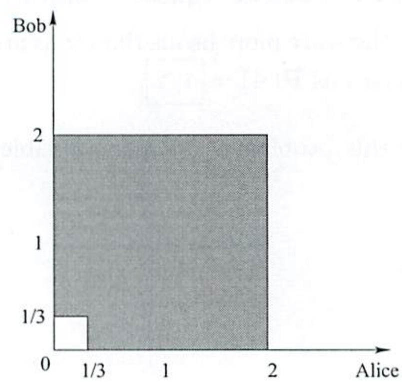
$$\mathbf{P}(\text{Even sum}) \;=\; 2p + 4p + 4p + 6p + 4p + 6p + 6p + 8p = 40p = \boxed{1/2}$$

(b)

$$\mathbf{P}(\text{Rolling a 2 and a 3}) \;=\; \mathbf{P}(2,3) + \mathbf{P}(3,2) = 5p + 5p = 10p = \boxed{1/8}$$

4. **P(B)**

The shaded area in the following figure is the union of Alice's pick being greater than 1/3 and Bob's pick being greater than 1/3.

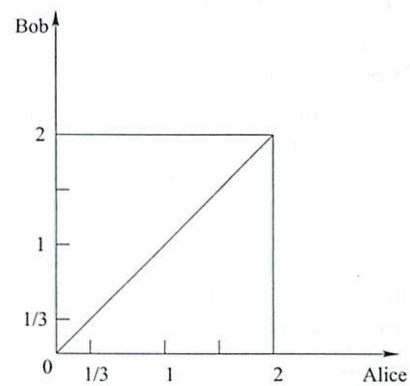


2/5

$$\begin{aligned}
\mathbf{P}(B) &= 1 - \mathbf{P}(\text{both numbers are smaller than } 1/3) \\
&= 1 - \frac{\text{area of small square}}{\text{total sample area}} \\
&= 1 - \frac{(1/3)(1/3)}{4} = 1 - \frac{1}{36} = \boxed{35/36}
\end{aligned}$$

P(C)

In the following figure, the diagonal line represents the set of points where the two selected numbers are equal.

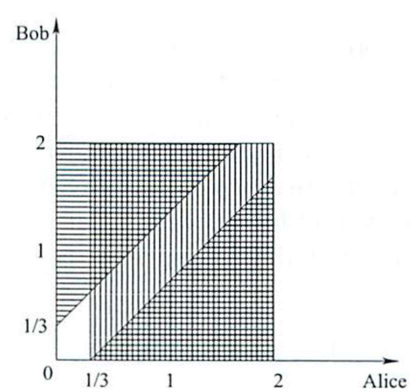


The line has an area of 0. Thus,

$$\mathbf{P}(C) = \frac{\text{area of line}}{\text{total sample area}} = \frac{0}{4} = \boxed{0}$$

P($A \cap D$)

Overlapping the diagrams we would get for $\mathbf{P}(A)$ and $\mathbf{P}(D)$,



$$\begin{aligned}\mathbf{P}(A \cap D) &= \frac{\text{double shaded area}}{\text{total sample area}} \\ &= \frac{(5/3)(5/3)(1/2) + (4/3)(4/3)(1/2)}{4} = \frac{25/18 + 16/18}{4} = \boxed{41/72}\end{aligned}$$

5. (a) The probability of Mike scoring 50 points is proportional to the area of the inner disk. Hence, it is equal to $\alpha\pi R^2 = \alpha\pi$, where α is a constant to be determined.

   Since the probability of landing the dart on the board is equal to one, $\alpha\pi 10^2 = 1$, which implies that $\alpha = 1/(100\pi)$.

   Therefore, the probability that Mike scores 50 points is equal to $\pi/(100\pi) = \boxed{0.01}$

   (b) In order to score exactly 30 points, Mike needs to place the dart between 1 and 3 inches from the origin. An easy way to compute this probability is to look first at that of scoring *more* than 30 points, which is equal to $\alpha\pi 3^2 = 0.09$.

   Next, since the 30 points ring is disjoint from the 50 points disc, probability of scoring more than 30 points is equal to the probability of scoring 50 points plus that of scoring exactly 30 points. Hence, the probability of Mike scoring exactly 30 points is equal to $0.09 - 0.01 = \boxed{0.08}$

   (c) For the part (a) question. The probability of John scoring 50 points is equal to the probability of throwing in the right half of the board and scoring 50 points plus that of throwing in the left half and scoring 50 points.

   The first term in the sum is proportional to the area of the right half of the inner disk and is equal to $\alpha\pi R^2/2 = \alpha\pi/2$, where α is a constant to be determined.

   Similarly, the probability of him throwing in the left half of the board and scoring 50 points is equal to $\beta\pi/2$, where β is a constant (not necessarily equal to α).

   In order to determine α and β, let us compute the probability of throwing the dart in the right half of the board. This probability is equal to

$$\alpha\pi R^2/2 = \alpha\pi 10^2/2 = \alpha 50\pi.$$

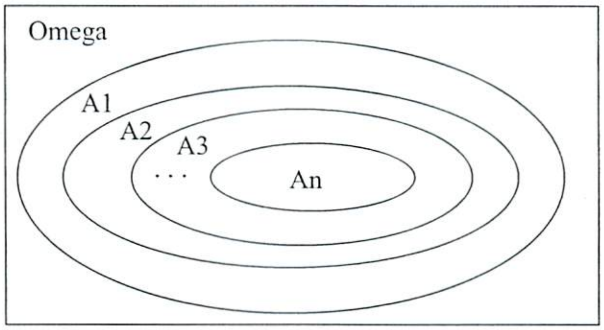
   Since that probability is equal to $2/3$, $\alpha = 1/(75\pi)$. In a similar fashion, β can be determined to be $1/(150\pi)$. Consequently, the total probability is equal to $1/150 + 1/300 = \boxed{0.01}$

   For the part (b), The probability of scoring exactly 30 points is equal to that of scoring more than 30 points minus that of scoring exactly 50. By applying the same type of analysis as in (b) above, the probability is found to be equal to $\boxed{0.08}$

   These numbers suggest that John and Mike have similar skills, and are equally likely to win the game. The fact that Mike's better control (or worst, depending on how you look at it) of the direction of his throw does not increase his chances of winning can be explained by the observation that both players' control over the distance from the origin is identical.

6. See the textbook, Problem 1.11 page 55, which proves the general version of Bonferroni's inequality.

G1†. (a) If we define $A_n = [a_n, b_n]$ for all n, it is easy to see that the sequence $A_1, A_2, \ldots$ is "monotonically decreasing," i.e., $A_{n+1} \subset A_n$ for all n:



Furthermore, $\cap_n^\infty A_n = [a, b]$.

By the continuity property of probabilities (see Problem 1.13, page 56 of the text),

$$\lim_{n\to\infty} \mathbf{P}([a_n, b_n]) = \mathbf{P}([a, b]).$$

(b) No. Consider the following example. Let $a_n = a + \frac{1}{n}$, $b_n = b - \frac{1}{n}$ for all n. Then $\{a_n\}$ is a decreasing sequence that converges to a, and $\{b_n\}$ is an increasing sequence that converges to b. If we define a probability law that places non-zero probability only on points a and b, then $\lim_{n\to\infty} \mathbf{P}([a_n, b_n]) = 0$, but $\mathbf{P}([a, b]) = 1$.

This example is closely related to the continuity property of probabilities. In this case, if we define $A_n = [a_n, b_n]$, then $A_1, A_2, \ldots$ is "monotonically increasing," i.e., $A_n \subset A_{n+1}$, but $A = (\cup_n^\infty A_n) = (a, b)$, which is an open interval whose probability is 0 under our probability law.

9/15

LECTURE 3

- **Readings:** Section 1.5
- Review
- $\boxed{\text{Independence of two events}}$
- $\boxed{\text{Independence of a collection of events}}$

Review

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},$$ assuming $P(B) > 0$

have new info -> revise beliefs

- Multiplication rule:

  $P(A \cap B) = P(B) \cdot P(A \mid B) = P(A) \cdot P(B \mid A)$ Prob of 2 things happened

  just like normal probabilities - same axiom + formulas

- Total probability theorem:

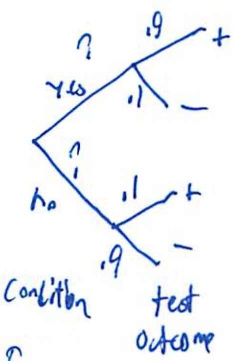
  $P(B) = P(A)P(B \mid A) + P(A^c)P(B \mid A^c)$

  add up all the diff ways B can occur
  - divide + conquer

- Bayes rule:

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)}$$

conditional prob in other direction
inference about underlying state of world
what caused an outcome?



P that person has condition

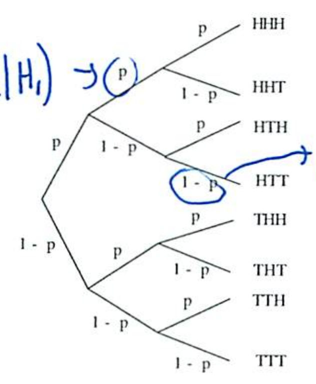
prior probability

must take into account

Models based on conditional probabilities

- 3 tosses of a biased coin:
  $P(H) = p$, $P(T) = 1 - p$

$P(H_2 \mid H_1) \rightarrow (p)$



$\rightarrow P(tH_3 \mid H_1, H_2)$

$$P(H_2) = P(H_1)\,P(H_2 \mid H_1) + P(T_1)\,P(H_2 \mid T_1)$$
$$= p \cdot p + (1-p) \cdot p = p$$
$$= P(H_2 \mid H_1) = P(H_2 \mid T_1)$$

independent
↓
prob remains same - no matter what happens

$P(THT) =$ $\cancel{p}\,(1-p) \cdot p(1-p) = p(1-p)^2$ multiply along branches

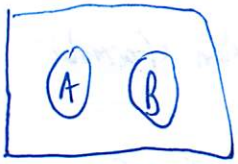
$P(\text{1 head}) =$ $(1-p) \cdot p(1-p) + (1-p) \cdot p\,(1-p) + \_\_\_ = 3p(1-p^2)$ add up all the branches that work

$P(\text{first toss is H} \mid \text{1 head}) =$ $\frac{P(H_1, \text{1 head})}{P(\text{1 head})}$ ratio of 2 quantities $= \frac{1}{3}$

Independence of two events

- **"Defn:"** $\mathrm{P}(B \mid A) = \mathrm{P}(B)$
  - "occurrence of A provides no information about B's occurrence"
- Recall that $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B \mid A)$
- **Defn:** $\boxed{\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)}$ consequence of

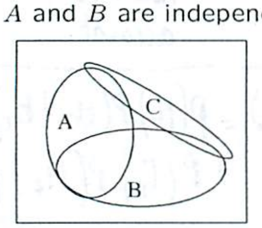
  more general ?!
- Symmetric with respect to A and B
  - applies even if $\mathrm{P}(A) = 0$
  - implies $\mathrm{P}(A \mid B) = \mathrm{P}(A)$



$P(A), P(B) > 0$ are events independent? **No** disjoint $\neq$ independent

Conditioning may affect independence

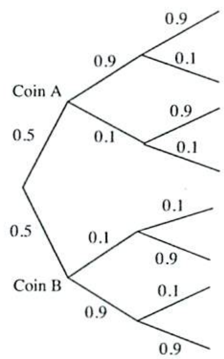
- Conditional independence, given C, is defined as independence under probability law $\mathrm{P}(\cdot \mid C)$ + may not know what will happen
- Assume A and B are independent



- If we are told that C occurred, are A and B independent? actually dependent

Conditioning may affect independence

- Two unfair coins, A and B:
  $\mathrm{P}(H \mid \text{coin } A) = 0.9$, $\mathrm{P}(H \mid \text{coin } B) = 0.1$
  choose either coin with equal probability



- Once we know it is coin A, are tosses independent? yes
- If we do not know which coin it is, are tosses independent?
  - Compare:
    $\mathrm{P}(\text{toss } 11 = H)$ symmetry $\frac{1}{2}$
    $\mathrm{P}(\text{toss } 11 = H \mid \text{first 10 tosses are heads})$ → what does this tell you
    - likely coin A
    So $P(\text{Heads}) \approx .9$
    ↑ almost

$P(H_{11}) = P(A)P(H_{11}|A) + P(B)P(H_{11}|B) = \frac{1}{2} \cdot .9 + \frac{1}{2} \cdot .1 = \frac{1}{2}$

Independence of a collection of events

- Intuitive definition:
  Information on some of the events tells us nothing about probabilities related to the remaining events
  - E.g.:

$\mathrm{P}(A_1 \cap (A_2^c \cup A_3) \mid A_5 \cap A_6^c) = \mathrm{P}(A_1 \cap (A_2^c \cup A_3))$

- Mathematical definition:
  Events $A_1, A_2, \ldots, A_n$
  are called **independent** if:

  $\boxed{\mathrm{P}(A_i \cap A_j \cap \cdots \cap A_q) = \mathrm{P}(A_i)\mathrm{P}(A_j)\cdots\mathrm{P}(A_q)}$

  for any distinct indices $i, j, \ldots, q$, (chosen from $\{1, \ldots, n\}$)

Events A_1, A_2, A_3

$$\left.\begin{aligned} P(A_1 \cap A_2) &= P(A_1)P(A_2) \\ P(A_2 \cap A_3) &= P(A_2)P(A_3) \\ P(A_3 \cap A_1) &= P(A_3)(P_1) \end{aligned}\right\} \text{pairwise indep.}$$

$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ } not necessarily true even w/ pairwise indep

⟶ see

Independence vs. pairwise independence

- Two independent fair coin tosses
  - A: First toss is H
  - B: Second toss is H $\quad P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
  - $\mathrm{P}(A) = \mathrm{P}(B) = 1/2$

HH C	HT	A
TH	TT C	
B		

  - C: ~~First and second~~ toss give same result $\quad \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$
  - $\mathrm{P}(C) =$
  - $\mathrm{P}(C \cap A) =$ $\frac{1}{4} = P(C)\,P(A)$ ← that events are independent
  - $\mathrm{P}(A \cap B \cap C) =$ $\frac{1}{4} \neq P(A)\,P(B)\,P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$
  - $\mathrm{P}(C \mid A \cap B) =$ $1 \neq P(C) = \frac{1}{2}$
- Pairwise independence **does not** imply independence

Conditional are not same as ~~independent~~ unconditional

Change of Page

The king's sibling

- The king comes from a family of two children. What is the probability that his sibling is female?

children are $\{B, G\}$ → $P(\frac{1}{2})$ for each

first assumption $P(\text{girl}) = \frac{1}{2}$

BB 1/4	BG 1/4
GB 1/4	GG 1/4

→ we know at least 1 boy

BB 1/3	BG 1/3
GB 1/3	0

$P(\text{girl}) = \frac{2}{3}$

Under what circumstances is this the right model?

Can think of alt. situations

- reproductive practices policies: we keep having children till we have a boy → so $p(\text{girl}) = 1$

or we stop at 2 boys $\quad p(\text{girl}) = 0$

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

Recitation 3: September 16, 2010

1. Example 1.20, page 37 in the text.

   Consider two independent fair coin tosses, in which all four possible outcomes are equally likely. Let

$$\begin{aligned} H_1 &= \{\text{1st toss is a head}\}, \\ H_2 &= \{\text{2nd toss is a head}\}, \\ D &= \{\text{the two tosses produced different results}\}. \end{aligned}$$

   (a) Are the events H_1 and H_2 (unconditionally) independent?

   (b) Given event D has occurred, are the events H_1 and H_2 (conditionally) independent?

2. Imagine a drunk tightrope walker, in the middle of a really long tightrope, who manages to keep his balance, but takes a step forward with probability p and takes a step back with probability $(1-p)$.

   (a) What is the probability that after two steps the tightrope walker will be at the same place on the rope?

   (b) What is the probability that after three steps, the tightrope walker will be one step forward from where he began?

   (c) Given that after three steps he has managed to move ahead one step, what is the probability that the first step he took was a step forward?

3. Problem 1.31, page 60 in the text.

   Communication through a noisy channel. A binary (0 or 1) message transmitted through a noisy communication channel is received incorrectly with probability ϵ_0 and ϵ_1, respectively (see the figure). Errors in different symbol transmissions are independent. The channel source transmits a 0 with probability p and transmits a 1 with probability $1-p$.

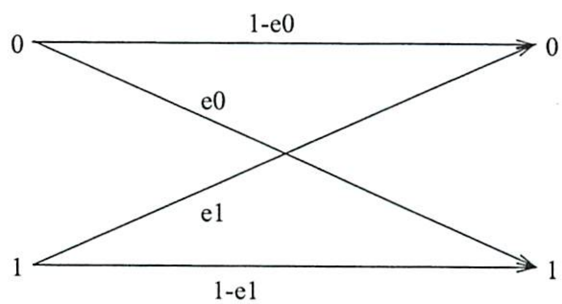


Figure 1: Error probabilities in a binary communication channel.

   (a) What is the probability that a randomly chosen symbol is received correctly?

   (b) Suppose that the string of symbols 1011 is transmitted. What is the probability that all the symbols in the string are received correctly?

(c) In an effort to improve reliability, each symbol is transmitted three times and the received symbol is decoded by majority rule. In other words, a 0 (or 1) is transmitted as 000 (or 111, respectively), and it is decoded at the receiver as a 0 (or 1) if and only if the received three-symbol string contains at least two 0s (or 1s, respectively). What is the probability that a transmitted 0 is correctly decoded?

(d) Suppose that the scheme of part (c) is used. What is the probability that a 0 was transmitted given that the received string is 101?

4. (a) Can an event A be independent of itself?

   (b) Problem 1.43(a) on page 63 in text.
   Let A and B be independent events. Use the definition of independence to prove that the events A and B^c are independent.

   (c) Problem 1.44 on page 64 in text.
   Let A, B, and C be independent events, with $\mathbf{P}(C) > 0$. Prove that A and B are conditionally independent of C.

Recitation 3 Independence

9/17

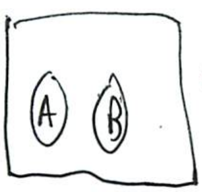
Seen a lot in real world

<u>Def</u> A + B are independent if $P(A \cap B) = P(A)P(B)$

If $P(B) > 0$ then A & B are independent if and only if

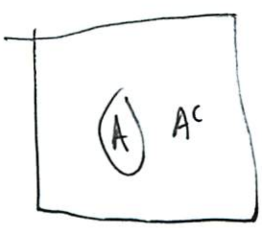
$$\underbrace{P(A \mid B)}_{\frac{P(A \cap B)}{P(B)}} = P(A)$$

- have to think about it carefully → not always intuitive

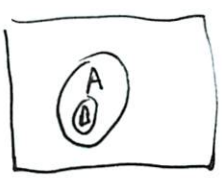


← A + B are most defentally <u>not</u> independent

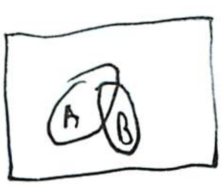
by + I mean &



$A + A^c$ <u>not</u>



$A \subset A$ <u>not</u>



A & B <u>independent</u> A+B overlap - but "not too much"

②

<u>Qu:</u> A + B are independent
is it true A + B^c are independent?
<u>Yes</u> since know true for A → must be true A^c

<u>Proof:</u> $P(A \cap B^c) = P(A) - \underbrace{P(A \cap B)}_{P(A) \cdot P(B)}$

$= P(A)(1 - P(B))$ factor

$= P(A) P^c(B)$

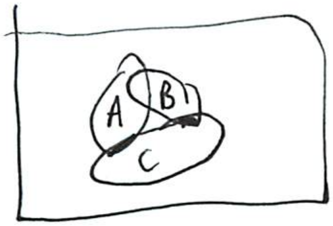
<u>Qu:</u> A + B are independent
A^c, B^c are independent?
↓ Get A + B^c are indep
A^c + B are indep } property
So A^c + B^c are inp.

<u>Def</u> Let C be such that $P(C) > 0$. We say A + B are conditionally independent (w/ respect to C) if

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

(3)

Suppose A + B are independent
- are they conditionally independent?



<u>No</u> - overlap somewhat
- so we know it can be disjoined

If A + B are conditional independent?
- are they indep.? <u>no</u>

2 coins
- blue + red → unfair coins   blue almost always heads $P=.99$
- pick 1 at random   red " " tails $P=.99$
- ~~record~~ flip twice
- record

Same as lecture

$H_1 = \{$1st toss is head$\}$

$H_2 = \{$2nd " " "$\}$

$B = \{$event blue is selected$\}$

$H_1 + H_2$ are ~~indep.~~ cond. indep relative to B.

" + " " <u>not</u> indep knowing H_1 has occured
- almost implies B was chosen + H_2 will occur

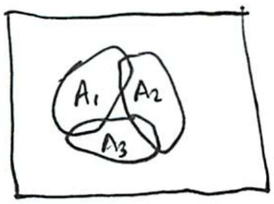
$P(H_2 | H_1) \neq P(H_2)$

justify w/ algebra - prob. tree

4

Def $A_1 \ldots A_n$ are independent if

$$P\Big(\bigcap_{i \in S} A_i\Big) = \prod_{i \in S} P(A_i) \quad \forall S \in \{1, \ldots, n\}$$



$$P(\underbrace{A_1 \cap A_2 \cap A_3}_{=\emptyset}) = 0 \neq P(A_1)P(A_2)P(A_3)$$

Pairwise independence $\neq$ Independence (of 3 or more A_s)

Conditional Independence of > 2 sets

Problem 1.44 Let A, B, C are independent ~~ev~~ events

$P(C) > 0$

$A + B$ are conditionally independent (related to C)

Proof: $$P(A \cap B \mid C) = \frac{P(A \cap B \cap C)}{P(C)} = \frac{P(A)P(B)\cancel{P(C)}}{\cancel{P(C)}}$$

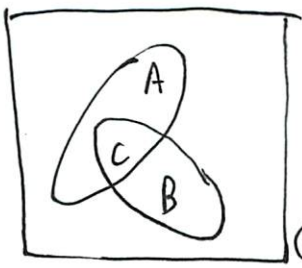
$$= P(A)P(B)$$

Unconditional P = conditional P

$$= P(A \mid C)\, P(B \mid C)$$

5

Q: On the ~~uniform~~ sample space $\Omega = \{1, \dots n\}$ w/ uniform law
What are the indep events?



m_a = # elements of A
m_b = # elements of B
c = # of common elements

$1 \le c < m_a < n$
$1 \le c < m_b < n$
n entire sample space

Independence $P(A \cap B) = P(A) \cdot P(B)$

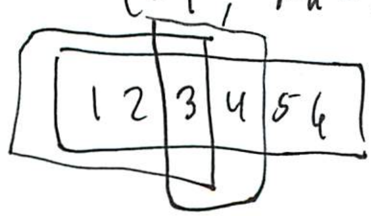
$$\frac{c}{n} = \frac{m_a}{n} \cdot \frac{m_b}{n}$$

$$n \cdot c = m_a \cdot m_b$$

7 sided If n is prime there are no indep. events that are non empty

If $n = 6$ $\quad 6 \cdot c = m_a \cdot m_b$

$c = 1, \; m_a = 2 \quad m_b = 3$



$c = 2 \quad \underset{3}{m_a} \cdot \underset{4}{m_b} = 12$

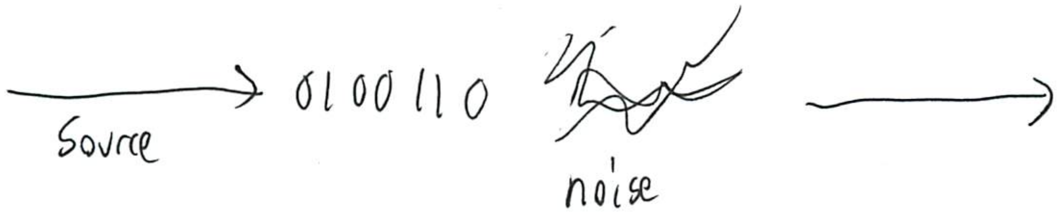
2 el, 3el → 1 overlap
3, 4, 2 overlaps

$c = 3 \quad m_a \cdot m_b = 18$ ⊗

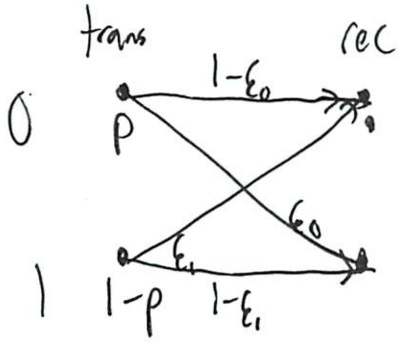
more curiosity than useful

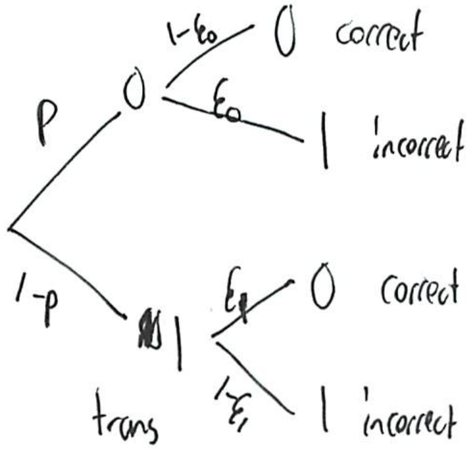
6

P. 1.31 communication theory



Some probability of error





$$P(\text{correct}, 1 \text{ symbol}) = p(1-\epsilon_0) + (1-p)(1-\epsilon_1)$$

$$P(\text{correct}, 1011) = p(\text{correct of } 1)^3 \cdot p(\text{correct of } 0)$$

$$(1-\epsilon_1)^3 \cdot (1-\epsilon_0)$$

⑦ How do we send it so to reduce errors? (coding)

- Send multiple times
  - 50 #0
  - majority rule
  - 3 times $0 \to 000$

    $1 \to 111$

So what is $P(000)$ correct?

- can be $\boxed{\begin{matrix}000\\001\\010\\100\end{matrix}}$ sum of the 4 probabilities $\begin{cases}(1-\varepsilon_0)^3\\(1-\varepsilon_0)^2\cdot\varepsilon_0\\\vdots\end{cases}$

Prob Correct transmition = total

if $\varepsilon_0 > \frac{1}{2}$ than worse reliability

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

Tutorial 1
September 16/17, 2010

1. Let A and B be events such that $A \subset B$. Can A and B be independent?

2. An electrical system consists of identical components that are operational with probability p independently of other components. The components are connected in three subsystems, as shown in the figure. The system is operational if there is a path that starts at point A, ends at point B, and consists of operational components. This is the same as requiring that all three subsystems are operational. What are the probabilities that the three subsystems, as well as the entire system, are operational?

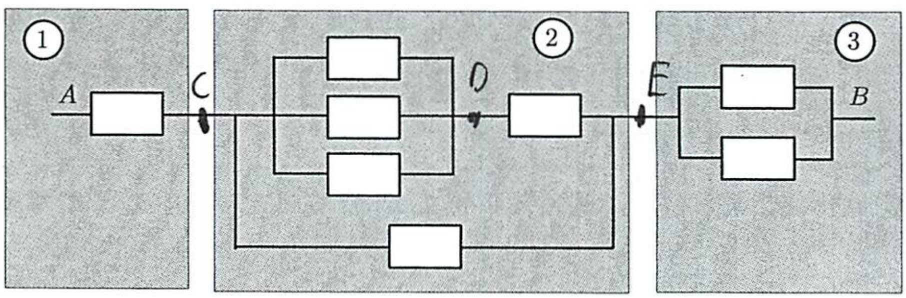


Figure 1: A system of identical components that consists of the three subsystems 1, 2, and 3. The system is operational if there is a path that starts at point A, ends at point B, and consists of operational components.

3. **The Chess Problem.** This year's Belmont chess champion is to be selected by the following procedure. Bo and Ci, the leading challengers, first play a two-game match. If one of them wins both games, he gets to play a two-game *second round* with Al, the current champion. Al retains his championship unless a second round is required and the challenger beats Al in both games. If Al wins the initial game of the second round, no more games are played.

   Furthermore, we know the following:

   - The probability that Bo will beat Ci in any particular game is 0.6.
   - The probability that Al will beat Bo in any particular game is 0.5.
   - The probability that Al will beat Ci in any particular game is 0.7.

   Assume no tie games are possible and all games are independent.

   (a) Determine the apriori probabilities that
      i. the second round will be required.
      ii. Bo will win the first round.
      iii. Al will retain his championship this year.

   (b) Given that the second round is required, determine the conditional probabilities that

i. Bo is the surviving challenger.

ii. Al retains his championship.

(c) Given that the second round was required and that it comprised only one game, what is the conditional probability that it was Bo who won the first round?

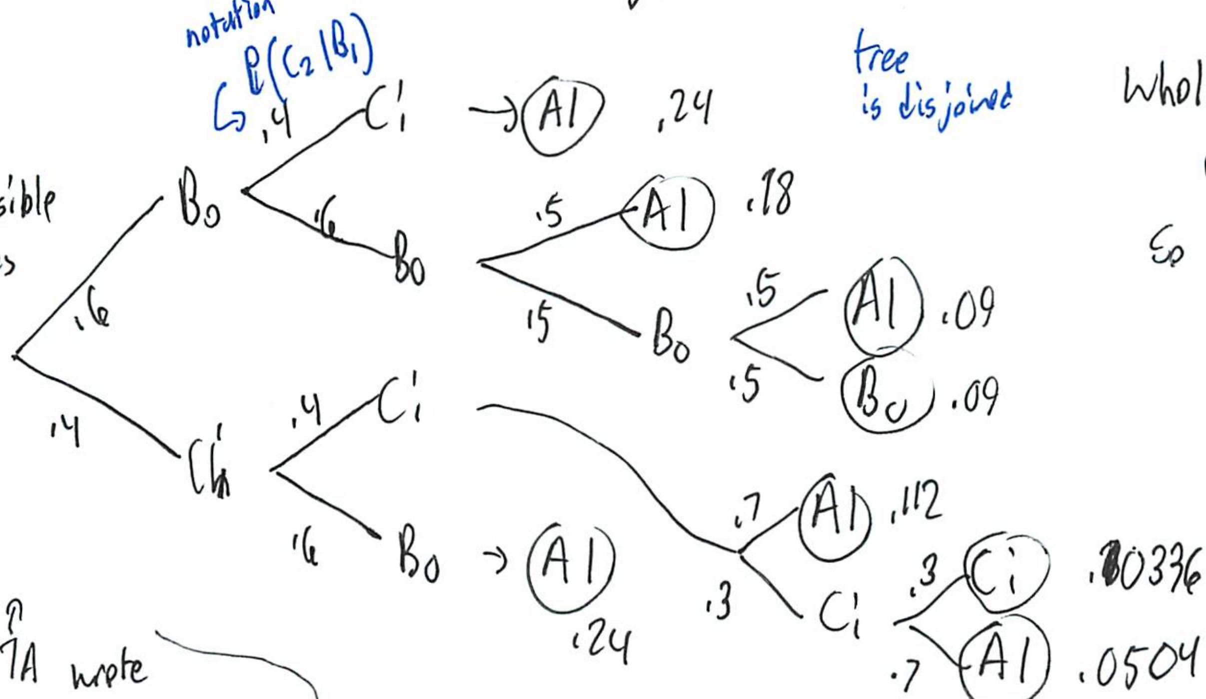
Tutorial 1

9/17

3. Round 1 Bo + Ch 2 game match

Round 2 winner R1 vs AI
1 or 2 games

Ω = all possible outcomes



tree is disjoined

Whole tornx is experiment
So sequence matters

TA wrote
$C_1 \cap C_2 \cap C_3 \cap A_4$
- ds1 more info

each endpoint is a possible outcome

adds to ~1
→ Sum = $\sum$

So then multiply a branch and add all the al wins
- multiplication rule

Now answer problems

a i) $P((B_1 \cap B_2) \cup (C_1 \cap C_2))$ ← which path does this happen

$= P(B_1 \cap B_2) + P(C_1 \cap C_2)$

$= .6^2 + .4^2 = .52$

or add my branches

(2)

b) $P[\text{B wins 1st round}] = P(P_1 \cap B_2) = .36$

$.6 \cdot .6 = .36$

c) A1 wins

- can also do (1 − Bo wins − Ci wins)

$1 - P(B) - P(C)$ ← less math to do and don't really need to draw tree

$\boxed{\text{add branches} = .9124}$

$1 - (.6^2 \cdot .5^2) - (.4^2 \cdot .3^2)$

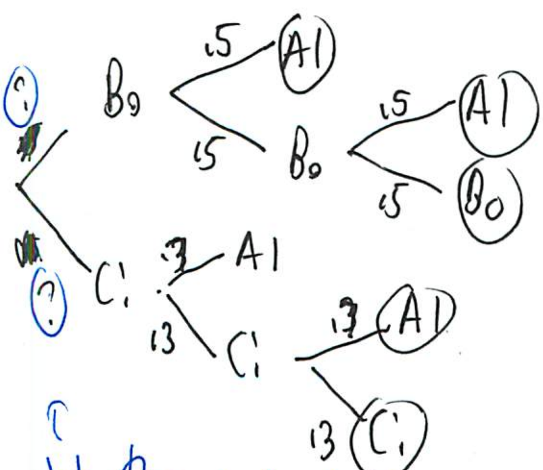
$= .8956$

---

b) Given that 2nd round is required

intuitive: cross out branches that can't happen
re distribute probabilities

or start tree from there



?
but then what here?

) I think some can work depending on situation
- need to think more about it
- be able to do it w/ just math as well

(3) Ta's way

b i) $\frac{P[B_0 \text{ survived} \cap R_2]}{P(R^2)} = \frac{P(B_1 \cap B_2)}{.52} = \frac{.36}{.52} = .69$

ii) $P[A | R_2] = \frac{P[A \cap R_2]}{P(R_2)} = \frac{\cancel{\ }}{\cancel{.52}}$

$$\frac{.6^2 \cdot .5 + .6^2 + .5^2 + .4^2 \cdot .3 + .4^2 \cdot .7 \cdot .3}{.52} = .7992$$

c) $P[B_0 \text{ survived} | R_2 \cap R_2 \text{ consisted of } 1 \text{ game}] =$

- TA way: make it an intersection

$$= \frac{P(B_0 \text{ survived} \cap R_2 \cap 1 \text{ game in } R_2)}{P(R_2 \cap 1 \text{ game in } R_2)}$$

$\boxed{\text{My way: look at } \cancel{\text{the}} \text{ tree (since I like graphical)}}$

$$= \frac{P(B_1 \cap B_2 \cap A_3)}{P(B_1 \cap B_2 \cap A_3) + P(C_1 \cap C_2 \cap A_3)}$$

$$= \frac{.6^2 \cdot .5}{.6^2 \cdot .5 + .4^2 \cdot .7} = .6164$$

(4)

If we did not do tree could do Baye's law etc

---

2. Electical System

P(whole system) depends on block being operational (opp)

Any path that is working works

$P[A\ C\ \text{opp.}] = p$

$P(E\ B\ \text{opp}) = 1 - (1-p)^2 = 1 - P\left(\underset{\text{op}}{\text{upper}}{}^c \cap \text{lower op}^c\right)$ $\boxed{\text{parallel}}$

↳ the blocks are independent

$P(CD\ \text{opp}) = 1 - (1-p)^3$

$P(CDE\ \text{opp}) = P(CD\ \text{opp} \cap DE\ \text{opp}) = p[1-(1-p)^3]$ $\boxed{\text{Series}}$

↳ independent, so multiply

$P(CE\ \text{opp}) = 1 - P(CDE\ \text{opp}^c) \times P\left(\underset{\text{op}}{\text{lower}}{}^c\right)$

$= 1 - [1 - p(1-(1-p)^3)][1-p]$

P(whole system) → multiply the 3 parts together

Building blocks

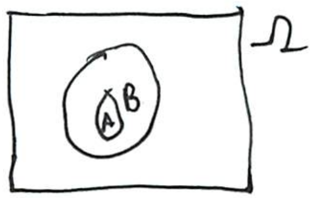
$P(\text{opp}) = P(\#1\ \text{opp} \cap \#2\ \text{opp})$

$= P(\#1\ \text{opp}) \cdot P(\#2\ \text{opp}) = p^2$

$P(\text{opp}) = P(\#1\ \text{opp} \cup \#2\ \text{opp})$

$= 1 - P(\#1\ \text{opp}^c \cap \#2\ \text{opp}^c) = 1 - (1-p)^2$

#1.



Is there any situation where $\overbrace{\text{A is a subset of B}}^{A \subset B}$ but A + B are independent?

independence → $P(A \cap B) = P(A) \cdot P(B)$

$A \subset B \Rightarrow P(A \cap B) = P(A)$

implies

Question: is their any case $P(A) \stackrel{?}{=} P(A) \cdot P(B)$

Case 1: $A = \emptyset$ so $P(A) = 0$

$0 = 0 \cdot P(B)$ → independent

Case 2: $B = \Omega$ so $P(B) = 1$

$P(A) = P(A) \cdot 1$ → independent

Can't come up w/ more

So $A = \emptyset$ or $B = 1$

9/20

LECTURE 4

- **Readings:** Section 1.6

Today: on a tangent: Counting

Lecture outline

- Principles of counting
- Many examples
  - permutations
  - k-permutations
  - combinations
  - partitions
- Binomial probabilities

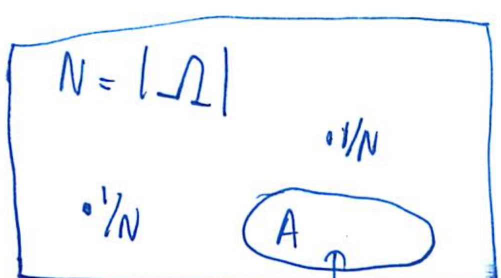
Why do we care about counting?

Discrete uniform law

- Let all sample points be equally likely
- Then,

$$\mathrm{P}(A) = \frac{\text{number of elements of } A}{\text{total number of sample points}} = \frac{|A|}{|\Omega|} = \frac{n}{N}$$

- Just count...



subset/event

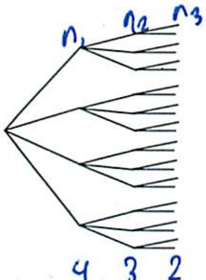
$|A| = n$

$P(A) = \frac{n}{N}$

gets complicated!

Basic counting principle

- r stages
- n_i choices at stage i



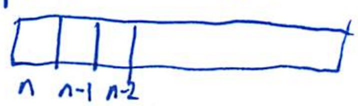
How many leaves on the tree?

$4 \cdot 3 \cdot 2 = 24$ choices

- Number of choices is: $n_1 n_2 \cdots n_r$

- Number of license plates with 3 letters and 4 digits $=$
  $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10$
- ... if repetition is prohibited $=$
  $26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7$
- **Permutations:** Number of ways of ordering n elements is:
  $n(n-1)(n-2) \cdots (1) = n!$

order matters   permutations = different ways of ordering a set



- Number of subsets of $\{1, \ldots, n\}$ $=$
  look at each item
  Choose → put it in subset or not
  $2 \cdot 2 \cdot 2 \cdots \cdot 2 = 2^n$

Example

- Probability that six rolls of a six-sided die all give different numbers? } A

cardinality of ↓

- $|A|$ – Number of outcomes that make the event happen: $6!$
  \# of permutations
- $|\Omega|$ – Number of elements in the sample space: 6^6
- – Answer: $\frac{|A|}{|\Omega|} = \frac{6!}{6^6}$

Combinations

k slots

$\binom{n}{k}$ → n n-1 n-2 ... (n-k+1)

in how many ways can we arrange them in an ordered sequence?

subset w/ k elements

Order them k!

- $\binom{n}{k}$: number of k-element subsets of a given n-element set
- Two ways of constructing an ordered sequence of k **distinct** items:
  - Choose the k items one at a time:
    $n(n-1)\cdots(n-k+1) = \boxed{\frac{n!}{(n-k)!}}$ choices
  - Choose k items, then order them ($k!$ possible orders)
- Hence: either branch - same ans

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

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

$$\sum_{k=0}^{n} \binom{n}{k} = \text{the \# of k element subsets, } 0, 1, \ldots, n = 2^n$$

?? mishead?

$$\binom{n}{n} = 1 = \frac{n!}{n!\,0!} \quad \boxed{0! = 1}$$

$$\binom{n}{0} = \frac{n!}{0!\,n!} = 1$$

Binomial probabilities

- n independent coin tosses
  - $\mathbf{P}(H) = p$

all tosses independent

- $\mathbf{P}(HTTHHH) = p(1-p)(1-p)ppp = P(TTHHHH)$

same # heads/tails So same probability

- $\mathbf{P}(\text{sequence}) = p^{\#\text{ heads}}(1-p)^{\#\text{ tails}}$

to the power

$$\mathbf{P}(k \text{ heads}) = \sum_{k-\text{head seq.}} \mathbf{P}(\text{seq.})$$ any sequence w/ exactly k heads

$$= (\#\text{ of } k\text{-head seqs.}) \cdot p^k(1-p)^{n-k}$$

$$= \boxed{\binom{n}{k} p^k (1-p)^{n-k}}$$

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$$

no matter what happens - get k heads for n

P(0 heads) + P(1 h) + P(... heads) + P(n)

disjoint, partitioned

$$\Omega = \bigcup_{k=0} \{\text{exactly } k \text{ heads}\}$$

In how many ways can we choose n # of k-head sequences

T T H H T H T T T   k=3

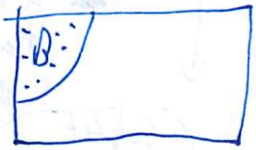
n

= # k ~~element~~ head subsets out of n element subsets

larger universe → not uniform
inside B → uniform probability

Coin tossing problem

B consists of all 3 head sequences



how many are there? $\binom{10}{3}$

- event B: 3 out of 10 tosses were "heads".
  - Given that B occurred, what is the (conditional) probability that the first 2 tosses were heads?

$P(A|B) = \frac{|A|}{|B|} = \frac{8}{\binom{10}{3}}$

- All outcomes in set B are equally likely: probability $p^3(1-p)^7$

  order of heads does not matter
  - Conditional probability law is uniform

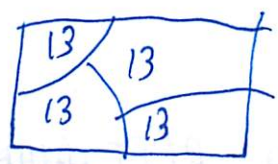
  all seq. =ly likely

  -dealing w/ counting problem
- Number of outcomes in B:
- Out of the outcomes in B, how many start with HH?

Partitions

- 52-card deck, dealt to 4 players
- Find $\mathbf{P}$(each gets an ace) A
- Outcome: a partition of the 52 cards

need $\frac{|A|}{|\Omega|}$



well shuffled → each partition is equally likely

  - number of outcomes:

$$|\Omega| = \frac{52!}{13!\,13!\,13!\,13!}$$

$|\Omega|$ = # possible partitions
-within → uniform probability law → counting problem

first player $\binom{52}{13}$ 2nd $\binom{39}{13}$ 3rd $\binom{26}{13}$ 4th $\binom{13}{13}$

$$\frac{52!}{13!\,\cancel{39!}} \cdot \frac{\cancel{39!}}{13!\,\cancel{26!}} \cdot \frac{\cancel{26!}}{13!\,13!} \cdot 1 = |\Omega|$$

$$= \frac{n!}{n_1! \dots n_4!}$$ multinomial coefficient

- Count number of ways of distributing the four aces: $4 \cdot 3 \cdot 2 \cdot 1$

  as distribute - who give it to
- Count number of ways of dealing the remaining 48 cards

$$\frac{48!}{12!\,12!\,12!\,12!}$$ rest of deck

- Answer:

$$\frac{4 \cdot 3 \cdot 2 \dfrac{48!}{12!\,12!\,12!\,12!}}{\dfrac{52!}{13!\,13!\,13!\,13!}}$$

← lots of simplifications = .105

alternative way in end of chap problem

(did not understand a lot of this)

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

Recitation 4
September 21, 2010 Counting

1. Problem 1.50, page 67 in the text.

   The birthday problem. Consider n people who are attending a party. We assume that every person has an equal probability of being born on any day during the year, independently of everyone else, and ignore the additional complication presented by leap years (i.e., nobody is born on February 29). What is the probability that each person has a distinct birthday?

2. Imagine that 8 rooks are randomly placed on a chessboard. Find the probability that all the rooks will be safe from one another, i.e. that there is no row or column with more than one rook.

3. Problem 1.61, page 69 in the text.

   Hypergeometric probabilities. An urn contains n balls, out of which exactly m are red. We select k of the balls at random, without replacement (i.e., selected balls are not put back into the urn before the next selection). What is the probability that i of the selected balls are red?

4. **Multinomial coefficient.** Derive the multinomial coefficient (the number of partitions of n distinct items into groups of $n_1, \ldots, n_r$) using a different argument than the one in class. Consider n items which can be placed into n slots and divide the group of n slots into segments of length $n_1, \ldots, n_r$ slots. Derive the multinomial coefficient by showing how many different ways can the n items be arranged into the r segments.

5. **Multinomial probabilities.** At each draw, there is a probability p_i $(i = 1, \ldots, r)$ of getting a ball of color i. Draw n objects. What is the probability of obtaining exactly n_i of each color i?

Massachusetts Institute of Technology

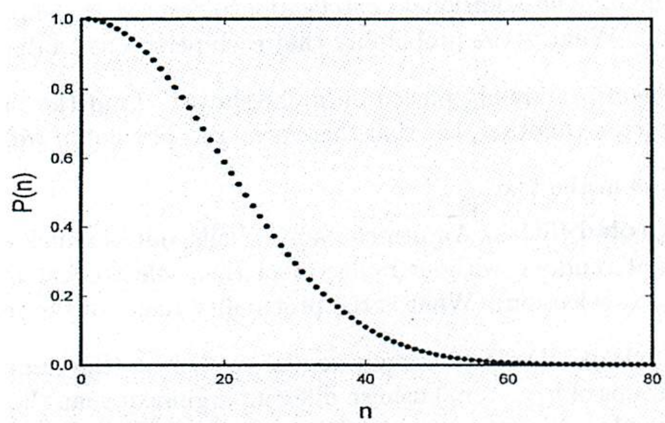
Department of Electrical Engineering & Computer Science

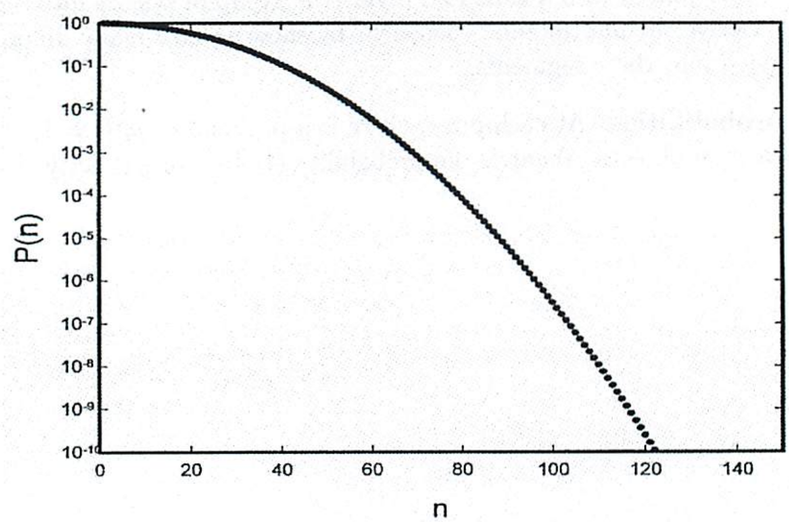
6.041/6.431: Probabilistic Systems Analysis

(Fall 2010)

Recitation 4: Extra Handout
September 21, 2010

1. As part of the solution to problem 1, plotted below are the probabilities of each person having a distinct birthday versus n the number of people present.





Recitation 4 Counting

9/21

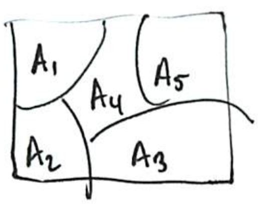
- useful → helps us calculate probability

Method to Calculate Prob.

- sequential

~~ha~~ multiply probabilities along route

- divide + conquer
  - total probability theorm



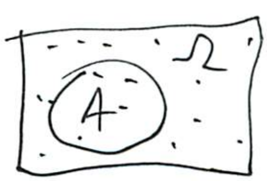
$$P(A) = P(A_i)\, P(A|A_i)$$

- Counting method
  - discrete uniform law

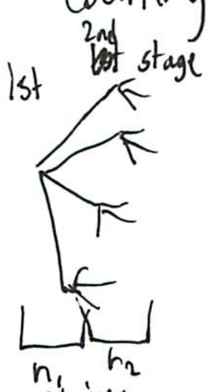
$$\Omega = \{1, \ldots, n\} \qquad P(A) = \left(\begin{matrix}\#\text{ outcomes in} \\ A\end{matrix}\right) \cdot (P)$$

prob of single outcome event

$$\underline{\text{or}} \qquad P(A) = \frac{\#\text{ outcomes in } A}{\text{Total }\#\text{ outcomes}}$$



- Counting principle



then # of leaves $= n_1 \cdot n_2 \cdots n_r$

^ # of choices

if each leaf has the same amt

(2)

Sub tricks for Counting

of k-permutations - order matters

# of k-combinations - order does not matter

# of partitions

If you are given a problem it may not fall neatly into one.

↳ may need to use a combo

Example: Birthday problem

k people attending party

P(2 have not same birthday)

365 b-days

ex: $\{251, 3, 4, 354, 71\}$

so $P(\text{no pairs in seq}) = \dfrac{\#\text{ sequences w/ distinct elements}}{\#\text{ of all possible sequences}}$

$$= \frac{365 \cdot 364 \cdot 363 \cdot \ldots 365-k+1}{365^k}$$ ← counting principle

both counts involve selection

(3)

k-permutation: selection of k objects out of n distict objects arranged in a seq where order matters

$= n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$

$= \dfrac{n!}{(n-k)!}$

Example 8x8 chess board

put 8 rooks randomly

P(rooks are safe) ie not in same row + column

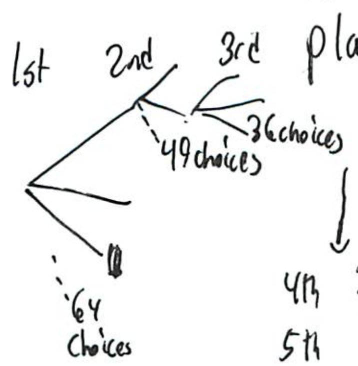
$P\left(\begin{matrix}\text{Placing safely}\\ \text{8 rooks on board}\end{matrix}\right) = \dfrac{\text{outcomes in event}}{\text{total outcomes}} = \dfrac{\#\text{ safe placements}}{\#\text{ total placements}}$

$\#$ total placements $= 64!/56! = \dfrac{n!}{(n-k)!}$ (a k-permutation)

$\#$ safe placements = break it up to 1 rook at a time

place 1 → 7x7 posibilities left → 49

2 → 6x6 → 36



4th 25 choices
5th 16
6th 4

$= 64 \cdot 49 \cdot 36 \cdot 25 \cdot 16 \cdot 4$

So answer

$$\frac{64 \cdot 49 \cdot 36 \cdot 25 \cdot 16 \cdot 4}{64!/56!}$$

---

k-combinations → order matters now

Want to find the # of sections of k out of n

4 people

Alice
Bob
Caroll
Debbie

2 of them will play - what are the possible matches?

AB BC CD
AC BD
AD

- does not matter what side of board they are on

# of choices $\frac{12}{2} = 6$

Steps

1st form k permutation $\left(\frac{n!}{(n-k)!}\right)$

Form groups around identical membership but diff. order

Reduction factor $= k!$

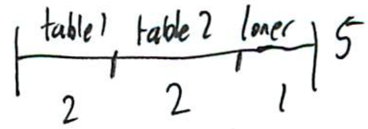
(5)

\# k combos $= \dfrac{n!}{k!(n-k)!} = \binom{n}{k}$

Suppose Emma joined → still 2 boards
- how many pairs?

$$\frac{5!}{2!\cdot 3!} = 10$$

Now they found another chess board
- 5 people
- 2 pairs of 2 + 1 longer



- One possibility for 1st stage
  - take 2 + assign them to table 1
  - Same as before

    \# of choices $= \dfrac{5!}{2!\cdot 3!} = 10 = \dfrac{n!}{n_1!(n-n_1)!}$
  - Then take remaining 3 for table 2

    \# of choices $\dfrac{(n-n_1)!}{(n-n_1-n_2)!\cdot n_2!}$

- if wanted prob of ~~loner~~ selecting loner (one anyway)

$$\frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!}$$

6

\# of partitions of n-objects into r groups of $n_1, n_2, n_3 \ldots n_r$ elements w/ no order

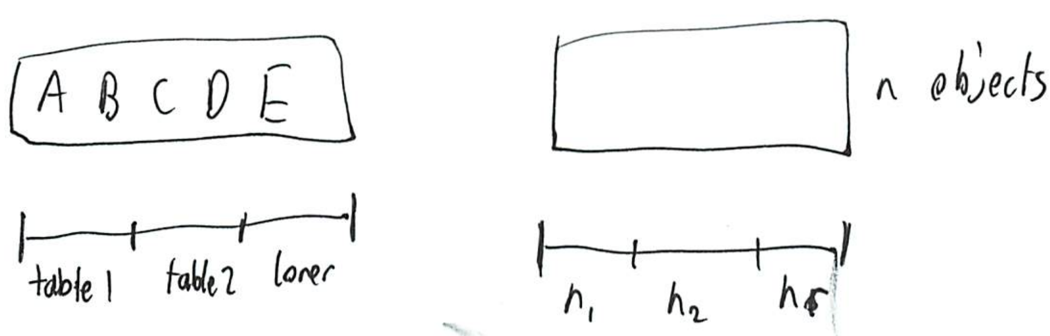
$$= \frac{n!}{n_1!(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \cdot \ldots \cdot \frac{(n-n_1 \ldots -n_{r-1})!}{n_r!(n-n_1-\ldots n_r)!}$$

a lot cancels

$$= \boxed{\frac{n!}{n_1! \cdot \ldots n_r!}}$$ multinomial formula

---

Another way to derive
- question on sheet
- shorter



AB BC
AC CA

reduction factor of 2

reduction factor of 1

reduction factor

$n_1!$ $n_2!$ $n_r!$

$$= \frac{n!}{n_1! \cdot n_2! \ldots n_r!}$$

7

$$= \frac{\text{\# of } n\text{-permutations}}{\prod \text{reduction factors}}$$ to account for lack of order

↑
the number of
multiply

Notation

9/21

Wikipedia

$\prod$ = product $\qquad \prod_{k=1}^{n} a_k$ means $a_1 a_2 \ldots a_n$

$\coprod$ = coproduct $\qquad$ Subsumes disjoined unions of sets

I guess a large $\bigcap$ $\bigcup$ just means that
- obvious

7 min late

9/22

LECTURE 5

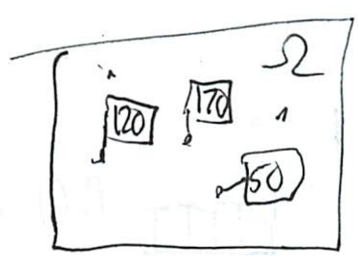
- **Readings:** Sections 2.1-2.3, start 2.4

Lecture outline

- Random variables
- Probability mass function (PMF)
- Expectation
- Variance

Random variables

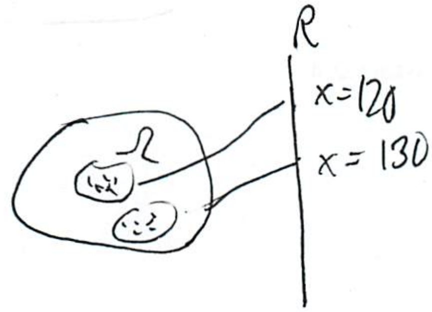
- An assignment of a value (number) to every possible outcome
- Mathematically: A function from the sample space Ω to the real numbers
  - discrete or continuous values
- Can have several random variables defined on the same sample space
- Notation:
  - random variable X
  - numerical value x



$Y(w) = g(X(w))$

$\left.\begin{matrix} X \\ \text{function } g \end{matrix}\right\} \; Y = g(X)$

Prob that a student's weight will be 120



Probability mass function (PMF)

- ("probability law", "probability distribution" of X)
- Notation:

equals what

$$p_X(x) = P(X = x)$$ ← p that takes on value

variable

$$= P(\{\omega \in \Omega \text{ s.t. } X(\omega) = x\})$$

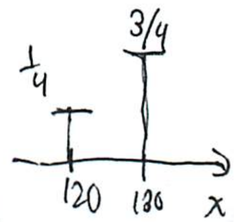
which random variable

all possible values

- $p_X(x) \geq 0 \qquad \sum_x p_X(x) = 1$

must follow rules

-Sum of all possibilities of weight =120



- **Example:** X=number of coin tosses until first head TTTTH THT HH ...

  - assume independent tosses, $P(H) = p > 0$ $X(w)=5$

$$\begin{aligned} p_X(k) &= P(X = k) \\ &= P(TT\cdots TH) \\ &= (1-p)^{k-1}p, \qquad k = 1, 2, \ldots \end{aligned}$$

  - **geometric** PMF

Since $P(T) = (1-p)$

So multiply till get to Heads (toss k)

So # of tails = k-1

* Should add up to 1 *

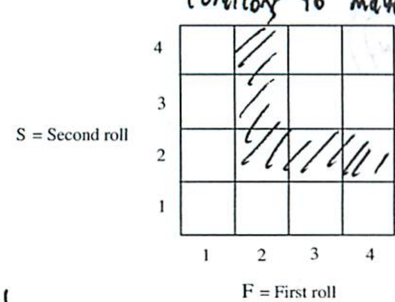
How to compute a PMF $p_X(x)$

- collect all possible outcomes for which X is equal to x
- add their probabilities
- repeat for all x

- **Example:** Two independent rolls of a fair tetrahedral die

TTTT $P_F(f)$ and $P_S(f)$

F: outcome of first throw
S: outcome of second throw
$X = \min(F, S)$ ← can also take function to make new one



$P_X(1)$

$p_X(2) =$ Collect all possible outcomes ↑ $= \frac{5}{16}$

$P_X(3)$

$P_X(4)$

Binomial PMF

- X: number of heads in $\boxed{n}$ independent coin tosses

  Constant
- $\mathrm{P}(H) = p$
- Let $n = 4$

$$
\begin{aligned}
p_X(2) &= \mathrm{P}(HHTT) + \mathrm{P}(HTHT) + \mathrm{P}(HTTH) \\
&\quad + \mathrm{P}(THHT) + \mathrm{P}(THTH) + \mathrm{P}(TTHH) \\
&= 6p^2(1-p)^2 \\
&= \binom{4}{2} p^2 (1-p)^2
\end{aligned}
$$

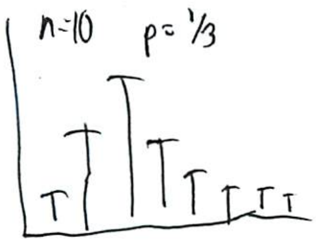
diff ways this can happen
-all have same probability

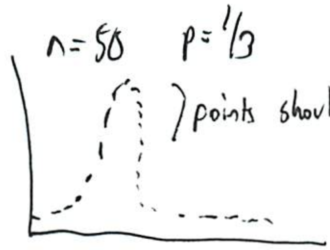
$$\binom{4}{2} = \frac{4!}{2!\,2!} = 6$$

In general:

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

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





points should be further apart

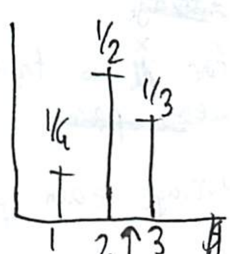
Expectation

- Definition:

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

each time you play
You get a random #

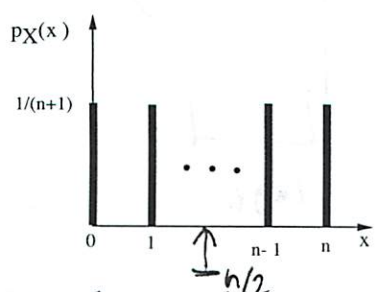
- Interpretations:
  - Center of gravity of PMF
  - Average in large number of repetitions of the experiment (to be substantiated later in this course)



what do you get on average?

$\frac{1}{6} \cdot 1 + \frac{1}{2} \cdot 2 + \frac{1}{3} \cdot 3$

- Example: Uniform on $0, 1, \ldots, n$



$\sum_x P_x(x)\, x$

$$\mathrm{E}[X] = 0 \times \frac{1}{n+1} + 1 \times \frac{1}{n+1} + \cdots + n \times \frac{1}{n+1} = \boxed{\frac{n}{2}}$$

*formula to find center of gravity
-where it would balance if on your finger

Properties of expectations

- Let X be a r.v. and let $Y = g(X)$
  - Hard: $\mathbf{E}[Y] = \sum_y y p_Y(y)$ apply def. of expectation
  - Easy: $\mathbf{E}[Y] = \sum_x g(x) p_X(x)$ for each thing that can happen - what is corresponding value of Y

  can prove formally - tomorrow
- Caution: In general, $\mathbf{E}[g(X)] \neq g(\mathbf{E}[X])$

Properties: If α, β are constants, then:

- $\mathbf{E}[\alpha] = \alpha$
- $\mathbf{E}[\alpha X] = \alpha \mathbf{E}[X]$
- $\mathbf{E}[\alpha X + \beta] = \alpha \mathbf{E}[X] + \beta$

$E[2] = 2$ always uninteresting

Variance

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

- **Second moment:** $\mathbf{E}[X^2] = \sum_x x^2 p_X(x)$
- **Variance**

$$\begin{aligned}
\text{var}(X) &= \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] \\
&= \sum_x (x - \mathbf{E}[X])^2 p_X(x) \\
&= \mathbf{E}[X^2] - (\mathbf{E}[X])^2
\end{aligned}$$

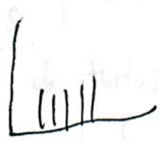
how far ~~m~~ x is from ~~average~~ average ~~each person~~

average each person's difference from avg

recitation

Properties:

- $\text{var}(X) \geq 0$
- $\text{var}(\alpha X + \beta) = \alpha^2 \text{var}(X)$


small


large

9/23

Skipped class

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

Recitation 5
September 23, 2010

1. (a) Derive the expected value rule for functions of random variables $\mathbf{E}[g(X)] = \sum_x g(x)p_X(x)$.
   (b) Derive the property for the mean and variance of a linear function of a random variable $Y = aX + b$.
   $$\mathbf{E}[Y] = a\mathbf{E}[X] + b, \qquad \text{var}(Y) = a^2\text{var}(X).$$
   (c) Derive $\text{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2$

2. A marksman takes 10 shots at a target and has probability 0.2 of hitting the target with each shot, independently of all other shots. Let X be the number of hits.
   (a) Calculate and sketch the PMF of X.
   (b) What is the probability of scoring no hits?
   (c) What is the probability of scoring more hits than misses?
   (d) Find the expectation and the variance of X.
   (e) Suppose the marksman has to pay \$3 to enter the shooting range and he gets \$2 dollars for each hit. Let Y be his profit. Find the expectation and the variance of Y.
   (f) Now let's assume that the marksman enters the shooting range for free and gets the number of dollars that is equal to the square of the number of hits. Let Z be his profit. Find the expectation of Z.

3. 4 buses carrying 148 job-seeking MIT students arrive at a job convention. The buses carry 40, 33, 25, and 50 students, respectively. One of the students is randomly selected. Let X denote the number of students that were on the bus carrying this randomly selected student. One of the 4 bus drivers is also randomly selected. Let Y denote the number of students on his bus.
   (a) Which of $E[X]$ or $E[Y]$ do you think is larger? Give your reasoning in words.
   (b) Compute $E[X]$ and $E[Y]$.

4. Problem 2.21, page 123 in the text.

   St. Petersburg paradox. You toss independently a fair coin and you count the number of tosses until the first tail appears. If this number is n, you receive 2^n dollars. What is the expected amount that you will receive? How much would you be willing to pay to play this game?

Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

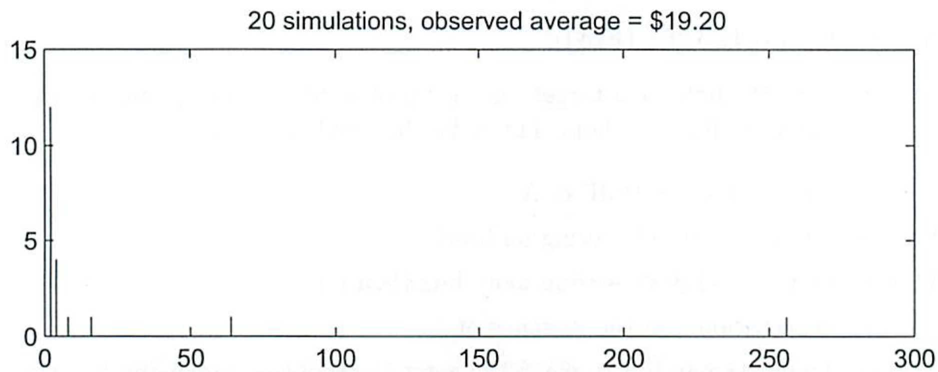
6.041/6.431: Probabilistic Systems Analysis

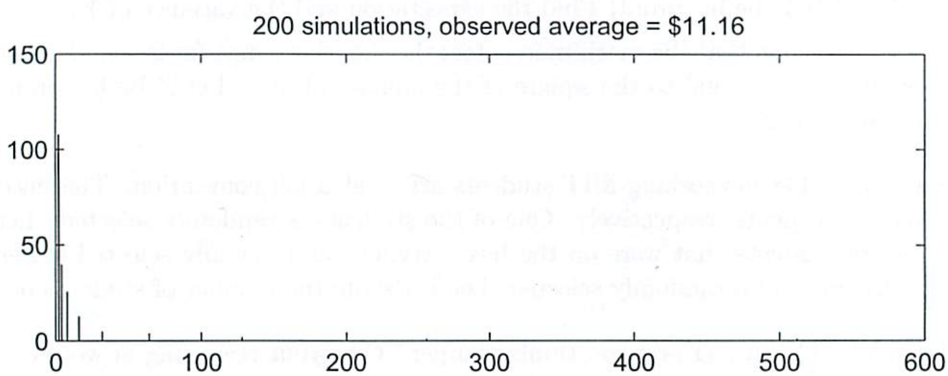
(Fall 2010)

Recitation 5: Extra Handout

September 23, 2010

1. To show some relavant computations to Problem 4, the results (plotted as histograms) of simulations of this game have been plotted below for various numbers of simulations.





Michael Plasmeier

4/10

Staple preferred.

If any pages fall apart I will only grade pages with your name.

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

Problem Set 2
Due September 22, 2010

1. Most mornings, Victor checks the weather report before deciding whether to carry an umbrella. If the forecast is "rain," the probability of actually having rain that day is 80%. On the other hand, if the forecast is "no rain," the probability of it actually raining is equal to 10%. During fall and winter the forecast is "rain" 70% of the time and during summer and spring it is 20%.

   (a) One day, Victor missed the forecast and it rained. What is the probability that the forecast was "rain" if it was during the winter? What is the probability that the forecast was "rain" if it was during the summer?

   (b) The probability of Victor missing the morning forecast is equal to 0.2 on any day in the year. If he misses the forecast, Victor will flip a fair coin to decide whether to carry an umbrella. On any day of a given season he sees the forecast, if it says "rain" he will always carry an umbrella, and if it says "no rain," he will not carry an umbrella. Are the events "Victor is carrying an umbrella," and "The forecast is no rain" independent? Does your answer depend on the season?

   (c) Victor is carrying an umbrella and it is not raining. What is the probability that he saw the forecast? Does it depend on the season?

2. You have a fair five-sided die. The sides of the die are numbered from 1 to 5. Each die roll is independent of all others, and all faces are equally likely to come out on top when the die is rolled. Suppose you roll the die twice.

   (a) Let event A to be "the total of two rolls is 10", event B be "at least one roll resulted in 5", and event C be "at least one roll resulted in 1".

      i. Is event A independent of event B?
      ii. Is event A independent of event C?

   (b) Let event D be "the total of two rolls is 7", event E be "the difference between the two roll outcomes is exactly 1", and event F be "the second roll resulted in a higher number than the first roll".

      i. Are events E and F independent?
      ii. Are events E and F independent given event D?

3. The local widget factory is having a blowout widget sale. Everything must go, old and new. The factory has 500 old widgets, and 1500 new widgets in stock. The problem is that 15% of the old widgets are defective, and 5% of the new ones are defective as well. You can assume that widgets are selected at random when an order comes in. You are the first customer since the sale was announced.

   (a) You flip a fair coin once to decide whether to buy old or new widgets. You order two widgets of the same type, chosen based on the outcome of the coin toss. What is the probability that they will both be defective?

   (b) Given that both widgets turn out to be defective, what is the probability that they were old widgets?

4. Oscar has lost his dog in either forest A (with a priori probability 0.4) or in forest B (with a priori probability 0.6).

   On any given day, if the dog is in A and Oscar spends a day searching for it in A, the conditional probability that he will find the dog that day is 0.25. Similarly, if the dog is in B and Oscar spends a day looking for it there, the conditional probability that he will find the dog that day is 0.15.

   The dog cannot go from one forest to the other. Oscar can search only in the daytime, and he can travel from one forest to the other only at night.

   (a) In which forest should Oscar look to maximize the probability he finds his dog on the first day of the search?

   (b) Given that Oscar looked in A on the first day but didn't find his dog, what is the probability that the dog is in A?

   (c) If Oscar flips a fair coin to determine where to look on the first day and finds the dog on the first day, what is the probability that he looked in A?

   (d) If the dog is alive and not found by the Nth day of the search, it will die that evening with probability $\frac{N}{N+2}$. Oscar has decided to look in A for the first two days. What is the probability that he will find a live dog for the first time on the second day?

5. In solving this problem, feel free to browse problems 43-45 in Chapter 1 of the text for ideas. If you need to, you may quote the results of these problems.

   (a) Suppose that A, B, and C are independent. Use the definition of independence to show that A and $B \cup C$ are independent.

   (b) Prove that if $A_1, \ldots, A_n$ are independent events, then

$$\mathbf{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = 1 - \prod_{i=1}^{n}(1 - \mathbf{P}(A_i)).$$

G1†. Alice, Bob, and Caroll play a chess tournament. The first game is played between Alice and Bob. The player who sits out a given game plays next the winner of that game. The tournament ends when some player wins two successive games. Let a tournament history be the list of game winners, so for example $ACBAA$ corresponds to the tournament where Alice won games 1, 4, and 5, Caroll won game 2, and Bob won game 3.

   (a) Provide a tree-based sequential description of a sample space where the outcomes are the possible tournament histories.

   (b) We are told that every possible tournament history that consists of k games has probability $1/2^k$, and that a tournament history consisting of an infinite number of games has zero probability. Demonstrate that this assignment of probabilities defines a legitimate probability law.

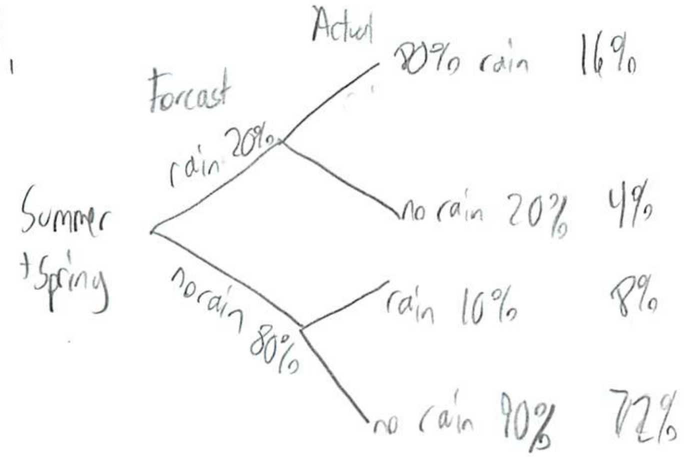
   (c) Assuming the probability law from part (b) to be correct, find the probability that the tournament lasts no more than 5 games, and the probability for each of Alice, Bob, and Caroll winning the tournament.

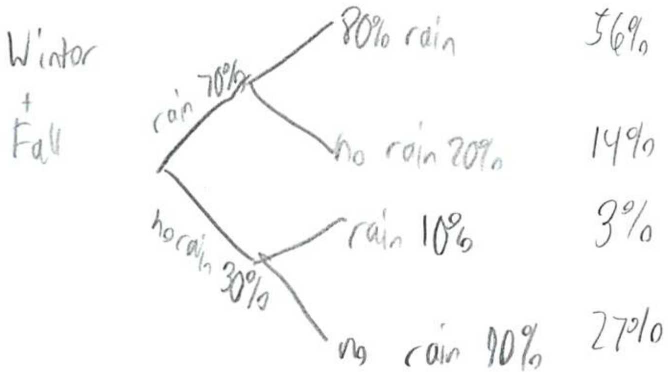
†Required for 6.431; optional for 6.041

Problem Set 2

9/20

1.





a) Missed forcast, rained

winter: $P(\text{Forcast Rain} \mid \text{Rained}) = \dfrac{P(\text{Forcast Rain} \cap \text{Rain})}{P(\text{Rained})} = \dfrac{.56}{.7} = 80\%$

summer: same $= \dfrac{.16}{.2} = 80\%$

b) $P(\text{miss forcast}) = .2 \to$ flip fair coin

No the answers are not independent

There are 2 causes to carry an umbrella

- he sees forcast (.8) and it is rain (.7 or .2 - season dependent)
- he misses " (.2) and coin is rain (.5)

((is fairly easy - but unsure of my work))

(2)

It is independent in both seasons, but the prob. he is carrying an umbrella depends on season

c) Here we go ↑ just answered that for last part

- carrying umbrella
- not raining

P(saw forcast)

Spring + Summer P(saw forcast | carring umbrella + not raining)

$$\frac{P(\text{saw forcast} \cap \text{carrying umbrella} \cap \text{not raining})}{P(\text{carring umbrella} \cap \text{not raining})} = \frac{.04}{.132} = .303$$

top: saw forcast + carrying → forcast = rain so look at tree .2 • .2

bottom: sees forcast (.8) • rain (.2) but not raining (.2)

plus misses " (.2) • coin = rain (.5)

$.8 \cdot .2 \cdot .2 + .2 \cdot .5$

$.032 + .1 = .132$

Winter

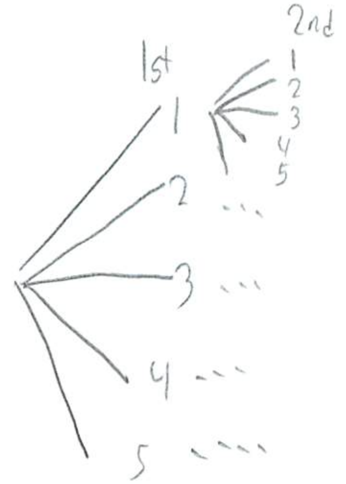
$$\frac{.7 \cdot .2}{.8 \cdot .7 \cdot .2 + .2 \cdot .5} = \frac{.14}{.128 + .1} = 61.4\%$$

(Season never up for probability)

Kinda fun if you feel ~~can~~ understand the math

(3)

2. Fair 5 sided die $\{1, 2, 3, 4, 5\}$ each roll ind. fair, roll twice



a.) A = total = 10

B = at least one roll = 5

C = at least one roll = 1

i) is A ind. of B?

does it provide any info, correct(?)

- Ind: $P(A \cap B) = P(A) \cdot P(B)$

no! if at total is 10 that can only be (5, 5) and then we know both rolls are 5, so $P(B|A) = 1$ so: $P(B \cap A) = 1$

ii) is A ind. of C?

no we know that A means (5, 5) so $P(C|A) = 0$ so? $P(C \cap A) = 0$

? can you make this assumption
- here, I think
- but not if ind.

(4)

b) D = total of 2 rolls = 7

E = difference b/w two rolls = 1

F = second roll > first roll

i) Yes → they overlap, but "not too much"

- if they were purly disjoint than that is not independence

ii) knowing sum = 7 can be

(2,5), (3,4), (4,3), (5,2)

now we know $P(E \mid D)$ is $\frac{2}{4} = .5$

$P(F \mid D)$ is $\frac{2}{4} = .5$) doesn't matter

but knowing one is greater does not tell you one is one bigger

Yes independent

(5)

3. Local Widget factory sale, everything must go

500 old widgets, 15% are defective

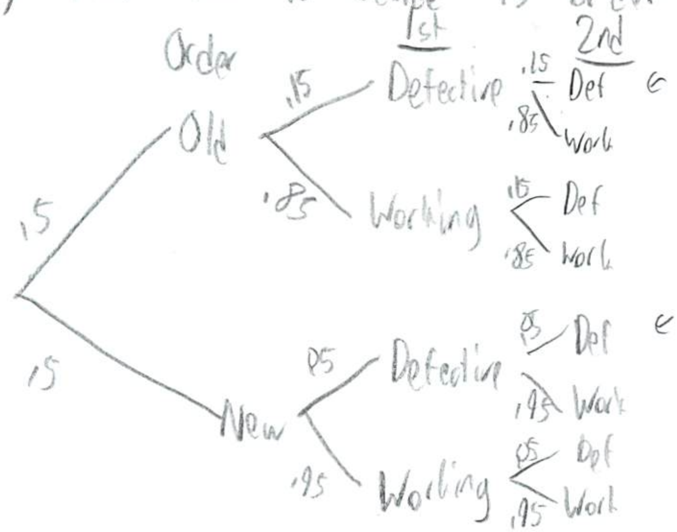
1500 new widgets, 5% " "

Wigets Selected at random ~~(old + new)~~ ~~clarified in tutorial~~ no can't be (working + broken) selected by shipping manager

You are customer # 1

YOU decide to order (old + new)

a) Fair coin to decide to order old or new



Is this the only way to display?

a) You order two of same type (based on coin toss)

P(Both Defective) → divide + conqor based on coin flip

$.5 \cdot .15 \cdot .15 + .5 \cdot .05 \cdot .05$

$.0125$

Probability is reduced.

+0.5 for being close.

(-0.5)

b) What is $P(\text{Old widgets} \mid \text{Both defective}) = \dfrac{P(\text{Old} \cap \text{Both defective})}{P(\text{Both Defective})} = \dfrac{.5 \cdot .15 \cdot .15}{.0125}$

this should be Bayes

Similar idea. +0.5

(-0.5)

$= \dfrac{.01125}{.0125} = .9$

1/2

I just do the math + seems to make sense

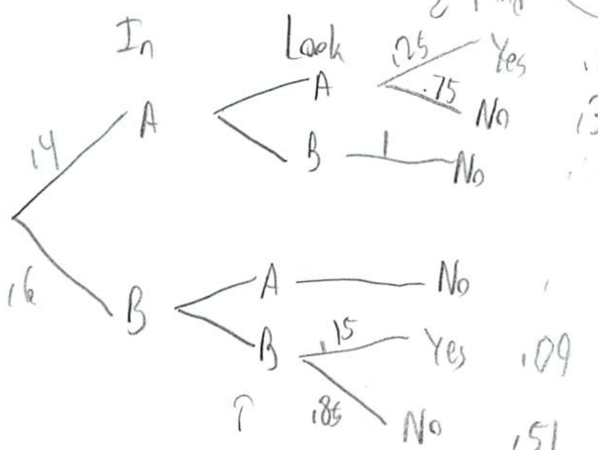
Cool - never thought how math worked - but works nice

(6)

Seems easier than last week

4. Oscar lost dog

Remember this is the conditional prob



? What is prob of this

Reminder
$P(A_i | B)$ = posterior
$P(A_i)$ = prior

? also how to represent different day dimensions

? if prob missing off tree → should I skip that one?
- seems to add up right
but P(find) depends on which he chooses to

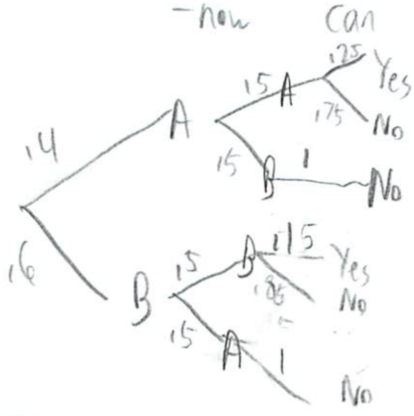
a) Function of where dog is and how easy to find
- he should look in A   .1 > .09

+1

b) $P(\text{in } A \mid \text{not found 1st day}) = \frac{P(A \cap no)}{P(no)} = \frac{.3}{.75} = .4$

.1 (-1)

c) Flips coin to decide where to look
- now can fill in



getting better at making trees

(7)

- don't know where dog is

$$P(\text{looked in } A \mid \text{found on 1st day}) = \frac{P(\text{looks in } A \cap \text{found on 1st day})}{P(\text{found 1st day})}$$

$$= \frac{.4 \cdot .5 \cdot .25}{.4 \cdot .5 \cdot .25 + .6 \cdot .5 \cdot .15} = \frac{.05}{.05 + .045} = .526$$

(−1) 9/21

d) If the dog is alive and not found by Nth day, it will die that evening w/ $P = \frac{N}{N+2}$. Oscar looks in A. P he will find live dog for 1st time on 2nd day?

$$P(\text{find dog}_{\text{Day 1}} \mid \text{looks in } A) = \frac{P(\text{find dog} \cap \text{looked in } A \cap \text{dog in } A)}{P(\text{looks in } A)}$$

- opposed of above

is this a given?

works out anyway??

$$\frac{.4 \cdot .25}{.4} = .25 \leftarrow P \text{ finds dog on 1st day}$$

$P(\cancel{\text{not}}\ \text{find dog Day 1} \mid \text{looks in } A)$

not conditional I think now

$P(\text{find day 1}) = .4 \cdot .25 = .1$

$P(\text{dog lives the night}) = \frac{1}{1+2} = .333$

$P(\text{find dog day 2}) = .4 \cdot .25 = .1$

(8)

$$\underbrace{.1}_{\text{find day } 1}$$

know add or multiply
- add diff cases
- then multiply

$$.1 + \underset{P(\text{found day } 1)^c}{.9} \cdot \underset{P(\text{dog lived})}{.333} \cdot \underset{P(\text{found day } 2)}{.1}$$

should be right, makes small effect

$$.1 + \underbrace{.02997} = .12997$$ → not very good for Oscar :'(

Oh wait they just want P he finds it on the 2nd day
So .02997

:((-1) +0

See solutions.

1/4

⑨

here is the hard qu!

5. a) A, B, C independent

Use def (ind) to show (A) and (B ∪ C) are indep

$P(A \cap B) = P(A) \cdot P(B)$ - def. independence =

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

$P(B \cup C \mid B \cap C) = P(B \cup C)$

$\hookrightarrow P(B \mid B \cap C) = P($

So I don't think it involves conditional

$P(B \cup C) = P(B) + P(C)$ No.

??? So bad at these questions

Saying events are independent means:

$$\left.\begin{array}{l} P(A \cap B) = P(A)\,P(B) \\ P(B \cap C) = P(B)\,P(C) \\ P(C \cap A) = P(A)\,P(B) \end{array}\right\} \text{pairwise ind.}$$

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

If $P(B \cap C)$ are independent than $P(B \cup C)$ must be indep.

$P(A) = P(A \cap B) + P(A \cap B^c)$

+0

See solutions.

(10)

6. Prove that if $A_1, \ldots A_n$ are independent then

$$P(A_1 \cup A_2 \cup \ldots A_n) = 1 - \prod_{i=1}^{n} (1 - P(A_i))$$

? what does that even mean product.

-definition of inp. of several events

$$P(A) = 1 - P(A^c)$$

$$1 - \prod_{i=1}^{n} P(A_i^c)$$

$$P\left(\bigcap_{i \in S} A_i\right) = P(A_1 \cap A_2 \cap \ldots \cap A_n)$$

$$P\left(\bigcup_{i \in S} A_i\right)^c = 1 - P(A_1 \cup A_2 \cup \ldots \cup A_n)$$

$$P\left(\bigcup_{i \in S} A_i\right)^c = 1 - \prod P(A_i^c)$$

? means multiply wikipedia!

Yeah makes more sense now

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\, P(A_2) \ldots P(A_n)$$

but how convert $\cup$ to $\cap$

De Morgan's Law

$$\left(\bigcup_n S_n\right)^c = \bigcap_n S_n^c$$

Proof needs to be more formal next time.

+2

$$P(\cup A_i) = \prod P(A_i^c)$$

? makes kinda more sense now

2/4

(11)

From problem #42, emailed by TA

$\Omega = \{1, \dots n\}$ all outcomes = likely

What pairs are indep

m_A = # elements in A

m_B = # " " B

m_C = # " " $A \cap B$

$1 \leq m_C < m_a < n$ $\qquad$ $1 \leq m_C < m_b < n$

which independence implies that $\frac{m_C}{n} = P(A \cap B) = P(A)P(B)$

$= \frac{m_A}{n} \frac{m_b}{n}$

or $m_C n = m_A m_B$

Conditions are nessary + sufficent for A + B to be inp

ie $\quad n = 4$

$m_C = 1 \quad m_A = 2 \quad m_B = 2$

$A = \{1, 2\} \qquad B = \{2, 3\}$

ie2 $\quad n = 6$

$C = 1 \quad m_A = 2 \quad m_b = 3$ $\qquad$ $C = 2 \quad m_A = 3 \quad m_B = 4$

$\{1, 2\} \quad \{2, 3\}$ $\qquad$ $\{1, 2, 3\} \quad \{2, 3, 4, 5\}$

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

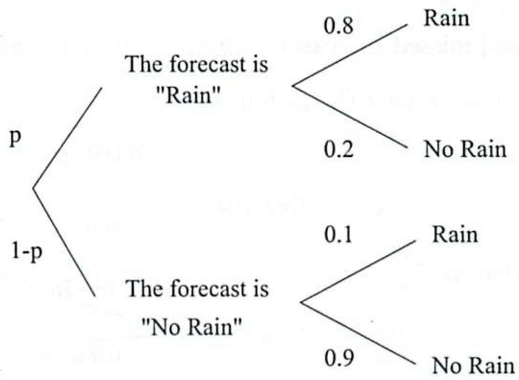
Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2010)

Problem Set 2: Solutions
Due September 22, 2010

1. (a) The tree representation during the winter can be drawn as the following:



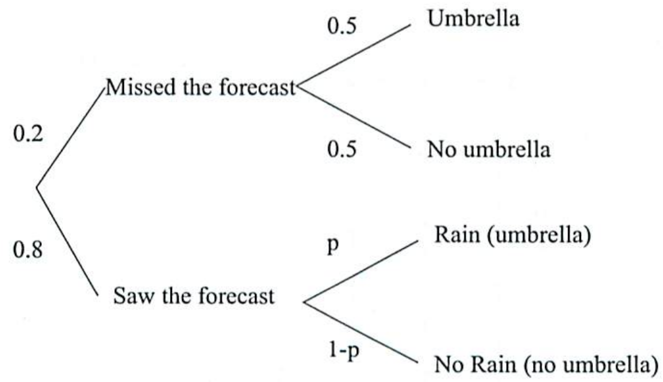
Let A be the event that the forecast was "Rain,"
let B be the event that it rained, and
let p be the probability that the forecast says "Rain." If it is in the winter, $p = 0.7$ and

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(B \mid A)\mathbf{P}(A)}{\mathbf{P}(B)} = \frac{(0.8)(0.7)}{(0.8)(0.7) + (0.1)(0.3)} = \frac{56}{59}.$$

Similarly, if it is in the summer, $p = 0.2$ and

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(B \mid A)\mathbf{P}(A)}{\mathbf{P}(B)} = \frac{(0.8)(0.2)}{(0.8)(0.2) + (0.1)(0.8)} = \frac{2}{3}.$$

(b) Let C be the event that Victor is carrying an umbrella.
Let D be the event that the forecast is no rain.
The tree diagram in this case is:



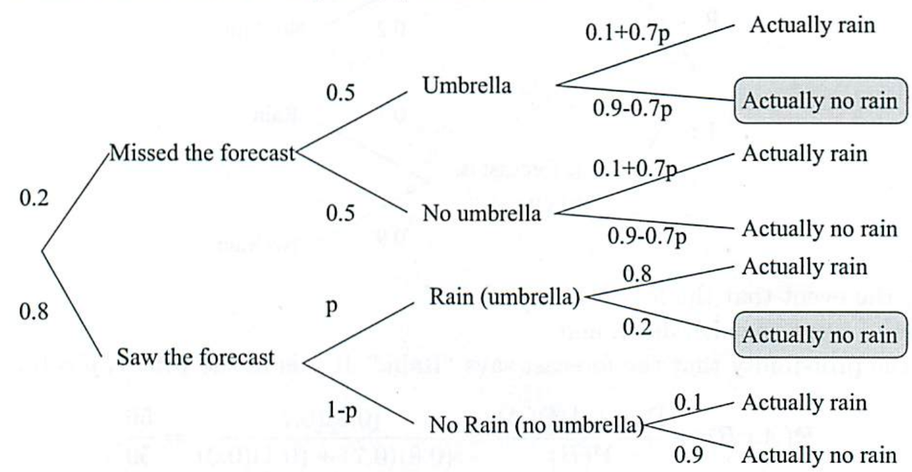
$$\begin{aligned}
\mathbf{P}(D) &= 1 - p \\
\mathbf{P}(C) &= (0.8)p + (0.2)(0.5) = 0.8p + 0.1 \\
\mathbf{P}(C \mid D) &= (0.8)(0) + (0.2)(0.5) = 0.1
\end{aligned}$$

Therefore, $\mathbf{P}(C) = \mathbf{P}(C \mid D)$ if and only if $p = 0$. However, p can only be 0.7 or 0.2, which implies the events C and D can never be independent, and this result does not depend on the season.

(c) Let us first find the probability of rain if Victor missed the forecast.

$$\mathbf{P}(\text{actually rains} \mid \text{missed forecast}) = (0.8)p + (0.1)(1-p) = 0.1 + 0.7p.$$

Then, we can extend the tree in part (b) as follows:

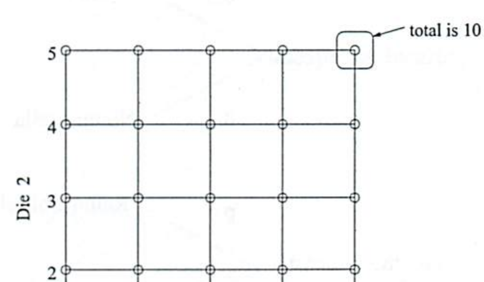


Therefore, given that Victor is carrying an umbrella and it is not raining, we are looking at the two shaded cases.

$$\mathbf{P}(\text{saw forecast} \mid \text{umbrella and not raining}) = \frac{(0.8)p(0.2)}{(0.8)p(0.2) + (0.2)(0.5)(0.9 - 0.7p)}$$

In fall and winter, $p = 0.7$, so the probability is $\frac{112}{153}$.
In summer and spring, $p = 0.2$, so the probability is $\frac{8}{27}$.

2. (a) i. $\boxed{\text{No}}$



Overall, there are 25 different outcomes in the sample space. For a total of 10, we should get a 5 on both rolls. Therefore $A \subset B$, and

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1$$

We observe that to get at least one 5 showing, we can have 5 on the first roll, 5 on the second roll, or 5 on both rolls, which corresponds to 9 distinct outcomes in the sample space. Therefore

$$\mathbf{P}(B) = \frac{9}{25} \neq \mathbf{P}(B|A)$$

ii. $\boxed{\text{No}}$ Given event A, we know that both roll outcomes must be 5. Therefore, we could not have event C occur, which would require at least one 1 showing. Formally, there are 9 outcomes in C, and

$$\mathbf{P}(C) = \frac{9}{25}$$

But

$$\mathbf{P}(C|A) = 0 \neq \mathbf{P}(C)$$

(b) i. $\boxed{\text{No}}$ Out of the total 25 outcomes, 5 outcomes correspond to equal numbers in the two rolls. In half of the remaining 20 outcomes, the second number is higher than the first one. In the other half, the first number is higher than the second. Therefore,

$$\mathbf{P}(F) = \frac{10}{25}$$

There are eight outcomes that belong to event E:

$$E = \{(1,2),(2,3),(3,4),(4,5),(2,1),(3,2),(4,3),(5,4)\}.$$

To find $\mathbf{P}(F|E)$, we need to compute the proportion of outcomes in E for which the second number is higher than the first one:

$$\mathbf{P}(F|E) = \frac{1}{2} \neq \mathbf{P}(F)$$

ii. $\boxed{\text{Yes}}$ Conditioning on event D reduces the sample space to just four outcomes

$$\{(2,5),(3,4),(4,3),(5,2)\}$$

which are all equally likely. It is easy to see that

$$\mathbf{P}(E|D) = \frac{2}{4} = \frac{1}{2}, \qquad \mathbf{P}(F|D) = \frac{2}{4} = \frac{1}{2}, \qquad \mathbf{P}(E \cap F|D) = \frac{1}{4} = \mathbf{P}(E|D)\mathbf{P}(F|D)$$

3. (a) Suppose we choose old widgets. Before we choose any widgets, there are $500 \cdot 0.15 = 75$ defective old widgets. The probability that we choose two defective widgets is

$$\begin{aligned}\mathbf{P}(\text{two defective}|\text{old}) &= \mathbf{P}(\text{first is defective}|\text{old}) \cdot \mathbf{P}(\text{second is defective}|\text{first is defective, old}) \\ &= \frac{75}{500}\frac{74}{499} = 0.02224\end{aligned}$$

Now let's consider the new widgets. Before we choose any widgets, there are $1500 \cdot 0.05 = 75$ defective old widgets. Similar to the calculations above,

$$\begin{aligned}\mathbf{P}(\text{two defective}|\text{new}) &= \mathbf{P}(\text{first is defective}|\text{new}) \cdot \mathbf{P}(\text{second is defective}|\text{first is defective, new}) \\ &= \frac{75}{1500}\frac{74}{1499} = 0.002568\end{aligned}$$

By the total probability law,

$$\begin{aligned}\mathbf{P}(\text{two defective}) &= \mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old}) \\ &\quad +\mathbf{P}(\text{new}) \cdot \mathbf{P}(\text{two defective}|\text{new}) \\ &= \frac{1}{2} \cdot 0.02224 + \frac{1}{2} \cdot 0.002568 = 0.01240.\end{aligned}$$

Note that this number is very close to what we would get if we ignored the effects of removing one defective widget before choosing the second widget:

$$\begin{aligned}\mathbf{P}(\text{two defective}) &= \mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old}) \\ &+ \mathbf{P}(\text{new}) \cdot \mathbf{P}(\text{two defective}|\text{new}) \\ &\approx \frac{1}{2} \cdot 0.15^2 + \frac{1}{2} \cdot 0.05^2 = 0.0125.\end{aligned}$$

(b) Using Bayes' rule,

$$\begin{aligned}\mathbf{P}(\text{old}|\text{two defective}) &= \frac{\mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old})}{\mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old}) + \mathbf{P}(\text{new}) \cdot \mathbf{P}(\text{two defective}|\text{new})} \\ &= \frac{\frac{1}{2} \cdot 0.02224}{\frac{1}{2} \cdot 0.02224 + \frac{1}{2} \cdot 0.002568} = 0.8965\end{aligned}$$

4. (a)

$$\mathbf{P}(\text{find in A and in A}) = \mathbf{P}(\text{in A}) \cdot \mathbf{P}(\text{find in A}|\text{in A}) = 0.4 \cdot 0.25 = 0.1$$

$$\mathbf{P}(\text{find in B and in B}) = \mathbf{P}(\text{in B}) \cdot \mathbf{P}(\text{find in B}|\text{in B}) = 0.6 \cdot 0.15 = 0.09$$

Oscar should search in Forest A first.

(b) Using Bayes' Rule,

$$\begin{aligned}\mathbf{P}(\text{in A}|\text{not find in A}) &= \frac{\mathbf{P}(\text{not find in A}|\text{in A}) \cdot \mathbf{P}(\text{in A})}{\mathbf{P}(\text{not find in A}|\text{in A}) \cdot \mathbf{P}(\text{in A}) + \mathbf{P}(\text{not find in A}|\text{in B}) \cdot \mathbf{P}(\text{in B})} \\ &= \frac{(0.75) \cdot (0.4)}{(0.4) \cdot (0.75) + (1) \cdot (0.6)} = \frac{1}{3}\end{aligned}$$

(c) Again, using Bayes' Rule,

$$\begin{aligned}\mathbf{P}(\text{looked in A}|\text{find dog}) &= \frac{\mathbf{P}(\text{find dog}|\text{looked in A}) \cdot \mathbf{P}(\text{looked in A})}{\mathbf{P}(\text{find dog})} \\ &= \frac{(0.25) \cdot (0.4) \cdot (0.5)}{(0.25) \cdot (0.4) \cdot (0.5) + (0.15) \cdot (0.6) \cdot (0.5)} = \frac{10}{19}\end{aligned}$$

(d) In order for Oscar to find the dog, it must be in Forest A, not found on the first day, alive, and found on the second day. Note that this calculation requires conditional independence of not finding the dog on different days and the dog staying alive.

$$\begin{aligned}\mathbf{P}(\text{find live dog in A day 2}) &= \mathbf{P}(\text{in A}) \cdot \mathbf{P}(\text{not find in A day 1}|\text{in A}) \\ &\quad \cdot \mathbf{P}(\text{alive day 2}) \cdot \mathbf{P}(\text{find day 2}|\text{in A}) \\ &= 0.4 \cdot 0.75 \cdot (1 - \frac{1}{3}) \cdot 0.25 = 0.05\end{aligned}$$

5. (a) We proceed as follows:

$$\begin{aligned}
\mathbf{P}(A\cap(B\cup C)) &= \mathbf{P}((A\cap B)\cup(A\cap C)) \\
&= \mathbf{P}(A\cap B)+\mathbf{P}(A\cap C)-\mathbf{P}(A\cap B\cap C) \\
&\overset{*}{=} \mathbf{P}(A)\mathbf{P}(B)+\mathbf{P}(A)\mathbf{P}(C)-\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) \\
&= \mathbf{P}(A)\left[\mathbf{P}(B)+\mathbf{P}(C)-\mathbf{P}(B)\mathbf{P}(C)\right] \\
&= \mathbf{P}(A)\mathbf{P}(B\cup C),
\end{aligned}$$

where the equality marked with $*$ follows from the independence of A, B, and C.

(b) Proof 1: If A and B are independent, then A^c and B^c are also independent (see Problem 1.43, page 63 for the proof).

For any two independent events U and V, DeMorgan's Law implies

$$\begin{aligned}
\mathbf{P}(U\cup V) &= \mathbf{P}((U^c\cap V^c)^c)=1-\mathbf{P}(U^c\cap V^c)=1-\mathbf{P}(U^c)\cdot\mathbf{P}(V^c) \\
&= 1-(1-\mathbf{P}(U))(1-\mathbf{P}(V)).
\end{aligned}$$

We proceed to prove the statement by induction. Letting $U=A_1$ and $V=A_2$, the base case is proven above. Now we assume that the result holds for any n and show that it holds for $n+1$. For independent $\{A_1,\ldots,A_n,A_{n+1}\}$, let $B=\cup_{i=1}^n A_i$. It is easy to show that B and A_{n+1} are independent. Therefore,

$$\begin{aligned}
\mathbf{P}(A_1\cup A_2\cup\ldots\cup A_{n+1}) &= 1-(1-\mathbf{P}(B))\cdot(1-\mathbf{P}(A_{n+1})) \\
&= 1-\prod_{i=1}^{n+1}(1-\mathbf{P}(A_i)),
\end{aligned}$$

which completes the proof.

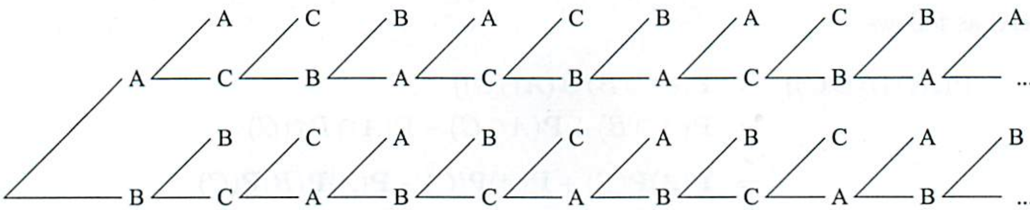
Proof 2: Alternatively, we can use the version of the DeMorgan's Law for n events:

$$\begin{aligned}
\mathbf{P}(A_1\cup A_2\cup\ldots\cup A_n) &= \mathbf{P}((A_1^c\cap A_2^c\cap\ldots\cap A_n^c)^c) \\
&= 1-\mathbf{P}(A_1^c\cap A_2^c\cap\ldots\cap A_n^c).
\end{aligned}$$

But we know that $A_1^c, A_2^c,\ldots,A_n^c$ are independent. Therefore

$$\begin{aligned}
\mathbf{P}(A_1\cup A_2\cup\ldots\cup A_n) &= 1-\mathbf{P}(A_1^c)\mathbf{P}(A_2^c)\ldots\mathbf{P}(A_n^c) \\
&= 1-\prod_{i=1}^{n}(1-\mathbf{P}(A_i)).
\end{aligned}$$

G1†. (a) The figure below describes the sample space via an infinite tree. The leaves of this tree are exactly all *finite* tournament histories; in addition, the two infinite paths represent the two *infinite* tournament histories that are possible. Note that the winner of the first game is either Alice or Bob; from then on, the winner of a game is either the winner of the previous game (in which case we have reached a leaf and the tournament has ended) or the player that sat out the previous game.The outcomes of the sample space correspond to the finite histories (which are identified with the leafs of the tree) and the two infinite histories: ACBACB... and BCABCA...



(b) The probability of an event is $1/2^k$ times the number of finite histories contained in the event. The probability of the event consisting of one or both infinite histories is 0. We have to show that this probability law satisfies the three probability axioms. It clearly satisfies nonnegativity and additivity. To check *normalization*, we have to verify that the probabilities of all tournament histories sum up to 1.
Start by noticing that two of the histories are infinite and have probability 0. Each one of the remaining histories has some finite length $k \geq 2$ (and hence is represented by one of the two leaves of the tree of the figure above at depth k) and probability $1/2^k$. Hence, summing all probabilities we get

$$2 \cdot 0 + \sum_{k=2}^{\infty} 2 \cdot \frac{1}{2^k} = \sum_{k=2}^{\infty} \frac{1}{2^{k-1}} = \sum_{k=0}^{\infty} \frac{1}{2^{k+1}} = \frac{1}{2} \sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{2} \frac{1}{1 - 1/2} = 1.$$

(c) The probability that exactly 2 games will be played is the sum of the probabilities of the two leaves at depth 2; that is,

$$P(\text{exactly 2 games}) = \frac{1}{2^2} + \frac{1}{2^2} = \frac{1}{2}.$$

Similarly, the probability that exactly i games will be played, for $i = 3, 4, 5$, is

$$\begin{aligned} P(\text{exactly 3 games}) &= \tfrac{1}{2^3} + \tfrac{1}{2^3} = \tfrac{1}{4}, \\ P(\text{exactly 4 games}) &= \tfrac{1}{2^4} + \tfrac{1}{2^4} = \tfrac{1}{8}, \\ P(\text{exactly 5 games}) &= \tfrac{1}{2^5} + \tfrac{1}{2^5} = \tfrac{1}{16}. \end{aligned}$$

Hence, the probability that the tournament lasts no more than 5 games is

$$P(\text{at most 5 games}) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}.$$

Hence, it's pretty probable that the tournament will last at most that much.
The probability that Alice wins the tournament is the sum of the probabilities of the leaves of the tree that are labeled "A"; that is,

$$\left(\frac{1}{2^2} + \frac{1}{2^5} + \frac{1}{2^8} + \cdots\right) + \left(\frac{1}{2^4} + \frac{1}{2^7} + \frac{1}{2^{10}} + \cdots\right),$$

where the first summation includes all leaves from the upper part of the tree, while the second one takes care of the leaves on the lower part. Calculating, we have

$$\frac{1}{4}\left(1 + \frac{1}{2^3} + \frac{1}{2^6} + \cdots\right) + \frac{1}{16}\left(1 + \frac{1}{2^3} + \frac{1}{2^6} + \cdots\right) = \frac{5}{16} \sum_{j=0}^{\infty} \frac{1}{8^j} = \frac{5}{16} \frac{1}{1 - 1/8} = \frac{5}{14}.$$

By symmetry (note the correspondence between the histories where Alice wins and the histories where Bob does), Bob's probability of winning is $\frac{5}{14}$, as well. Then, since the outcomes where nobody wins (these are the two infinite tournament histories) have total probability 0, Carol wins with probability $1 - \frac{5}{14} - \frac{5}{14} = \frac{4}{14}$. Hence, by not participating in the first game, Carol enters the tournament with a disadvantage.

† Required for 6.431; optional for 6.041

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

Tutorial 2
September 23/24, 2010

1. A player is randomly dealt 13 cards from a standard 52-card deck.
   (a) What is the probability the 13th card dealt is a king?
   (b) What is the probability the 13th card dealt is the first king dealt?

2. Consider a random variable X such that

$$p_X(x) = \frac{x^2}{a} \text{ for } x \in \{-3,-2,-1,1,2,3\}, \quad \mathbf{P}(X=x) = 0 \text{ for } x \notin \{-3,-2,-1,1,2,3\},$$

   where $a > 0$ is a real parameter.
   (a) Find a.
   (b) What is the PMF of the random variable $Z = X^2$?

3. 90 students, including Joe and Jane, are to be split into three classes of equal size, and this is to be done at random. What is the probability that Joe and Jane end up in the same class?

4. Draw the top 7 cards from a well-shuffled standard 52-card deck. Find the probability that the 7 cards include exactly 3 aces.

Tutorial 2

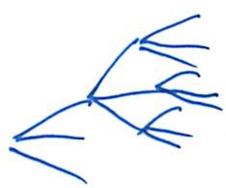
9/24

- Read the book

---

* How to describe randomness *

- Sequential method : seq of steps
  - ↳ multiplication rule
- divide + conquer
  - ↳ total prob. theorm
- Counting method
  - ↳ Counting principle



  - ↳ experiment in stages
  - ↳ each state has same substages
    - $n_1 * n_2 * n_3$
  - discrete uniform law

Ω if all pt =ly likely

$$P(A) = \frac{\#\text{ of elements in } A}{\#\text{ elements in } \Omega}$$

- or at least all pts inside are =ly likely

$$P(A) = p \cdot (\#\text{ elements in } A)$$

Shortcuts

- \# - permutations
  - order is important

fill in

$$\boxed{n}P\boxed{r} = \frac{n!}{(n-r)!}$$

permutation

(2)

Combination

↳ order does not matter

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

Partitions

- Combo of above
- Want 3 handfuls, remove all objects
  - order inside each hand does not matter
  - which hand it is in matters

$$\binom{n}{n_1\, n_2\, n_3} = \frac{n!}{n_1!\, n_2!\, n_3!}$$

#1 Player randomly delt 13 cards

a) P(13th card is king)

b) P(13th card is 1st king)

Ω = when you deal 13 cards, what are all possible outcomes?

$$\{ \underbrace{4k2A43\ldots}_{13},\ \underbrace{234k51\ldots}_{13} \}$$

Can't do actually

but think what this is

③

A = 13th card is 1st king

$P(A) = ?$

- Think using counting method

$\dfrac{\text{outcomes that satisfy}}{\text{total outcomes}}$) Since all outcomes =ly likley (discrete uniform law)

$= \dfrac{|A|}{|\Omega|} = \dfrac{\#\text{ elements in } A}{\#\ \text{"}\ \ \text{"}\ \Omega}$

$|\Omega|$ = permutation
- order matters
- it needs to be 1st king

$= \cancel{52\,P\,13}\ \ {}_{\boxed{52}}P_{\boxed{13}} = \dfrac{52!}{(52-13)!}$

filled in

divide experiment into 2 stages

no king $\underbrace{\_\ \_\ \_\ \_\ \_\ \_\ \_}_{12}$ king $\underbrace{\_}_{13\text{th}}$

$|A|$ = # of ways to deal 12 king free cards $\boxed{\times}$ # of ways to deal a king

$= {}_{48}P_{12} \cdot {}_{4}P_{1}$

$= {}_{48}P_{12} \cdot 4$

(4)

$$P(A) = \frac{{}_{48}P_{12} \cdot 4}{{}_{52}P_{13}}$$ ← can leave like this

a) B = $\underbrace{\text{any cards}}_{12 \text{ cards}}$ $\underbrace{\text{king}}_{13\text{th}}$

$|\Omega|$ = same

$|A|$ = # ways to deal 12 cards · # ways to deal a king

$= {}_{52}P_{12} \cdot {}_4P_1$

- actually not accurate, because don't know how many kings are here
- so when calculating 1st part, make allowance for 1 less king
- ~~or~~ ie pick 13th card 1st

$$= \frac{4 \cdot {}_{51}P_{12}}{{}_{52}P_{13}}$$ ← is accurate can describe it differently

$= \frac{4}{52}$ ← could just do this simple king could be anywhere
- average out all scenarios

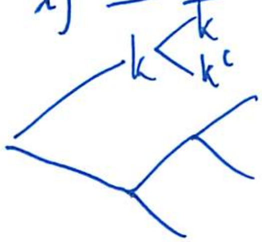
On an exam just do the sure way

Could also do sequential method
- all problems can be solved each way, may be a lot harder

(5)

b seq method) - describe outcome at each stage

i) 1st 2nd - like a tree



k_i = event that card i is a king

$A = K_1^c \cap k_2^c \cap \ldots\ k_{12}^c \cap k_{13}$

$P(A) = P(k_1^c) \cdot P(k_2^c | k_1^c) \cdot P(k_3^c | k_1^c \cap k_2^c) \cdot \ldots \cdot P(k_{13} | k_1^c \cap k_2^c \cap \ldots k_{12}^c)$

multiplication law

$$= \underbrace{\frac{48}{52} \cdot \frac{47}{51} \cdot \frac{46}{50} \cdot \ldots \cdot \frac{37}{41}}_{12 \text{ times}} \cdot \frac{4}{40}$$

both ans work on a quiz

#4 is similar
- do on own

52 cards
deal 7

A = exactly 3 aces

$|\Omega| = {}_{52}P_7$

$|A|$ = 4 ace free cards · 3 aces
do 1st

(6)

$${}_{4}P_{3} \cdot {}_{49}P_{4}$$

$$= \frac{{}_{4}P_{3} \cdot {}_{49}P_{4}}{{}_{52}P_{7}}$$

Seem really unsure this method
other method likley easier

but does order matter?
- no we don't care where aces are
- I think what I wrote assumed first 3 cards were aces

so try w/ combos

$\binom{52}{3}$ or is it $\binom{4}{3}$?

Ω = all possible 7 cards delt

* Order is not important
- would write sets, not sequences

$|\Omega| = \binom{52}{7}$

$|A| = \binom{4}{3} \boxed{\cdot} \binom{52-4}{4}$ counting principal

$P(A) = \frac{|A|}{|\Omega|}$ discrete prob law

- if you did force order it would cancle out $\frac{|A| \cdot \cancel{7!}}{|\Omega| \cdot \cancel{7!}}$

actually longer
but still cancles out

(7)

Oh can do all As up front
- like I did
- Oh but you can do it as a starting place but then ~~can~~ must shuffle $\binom{7}{3}$

---

3. 90 students

3 classes of 30

30	30	30
#1	#2	#3

A = Jane + Joe are in same class

Ω = all possible handfuls, order inside does not matter $\rightarrow$ partitions

$$|\Omega| = \binom{90}{30,30,30} = \frac{90!}{30!\,30!\,30!}$$

could also do

$$= \binom{90}{30}\cdot\binom{60}{30}\cdot\underbrace{\binom{30}{30}}_{1}$$

$|A|$ $\rightarrow$ pick a simpler problem, solve, then upgrade

Jane + Joe in | |
--- | --- | ---
1 | 2 | 3

$\leftarrow$ Jane + Joe in class 1

$$\#\text{ of ways } \cancel{\ldots} = \underbrace{\binom{2}{2}\binom{88}{28}}_{\text{Class 1}}\cdot\binom{60}{30}\cdot\binom{30}{30} = \binom{88}{28,30,30}$$

because how we defined sample space — where Alex is matters

③

$|A| = P$ above $\cdot 3$

-order does not matter of the bins

then just $\frac{|A|}{|\Omega|}$

#2

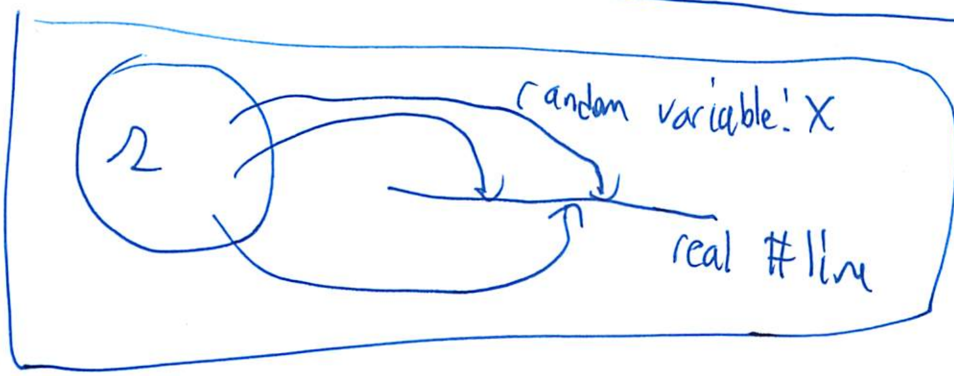
- give a random value
- P law is constant but unknow
- tip: say $P \sum = 1$
  - so do equation

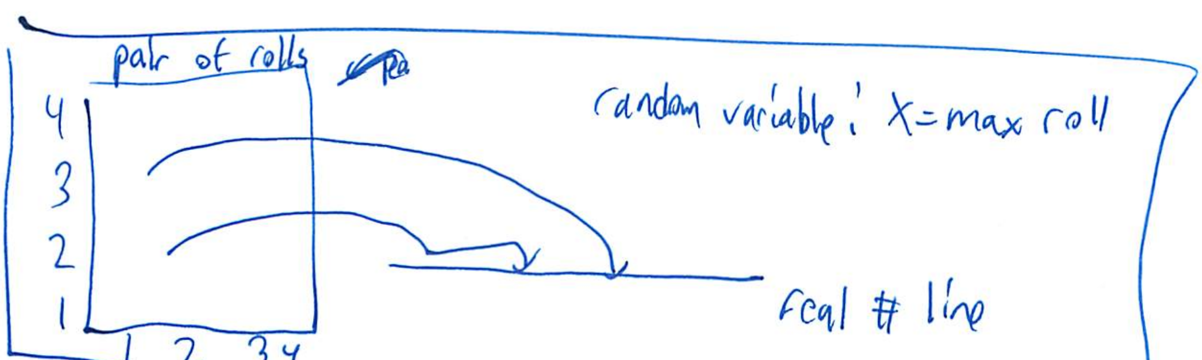
Chap 2 Discrete Random Variables

Basic Concepts

- many models → outcome numerical
- other outcomes → not numeric
  - but have some numeric attribute
  - that is determined proabatistically
  - random variables
  - - associates a particular # w/ a outcome

    ↑ numeric value/ value
- a random variable is a real-valued function of expermental outcome





visualization of random variable

(seems to be variation on same theme - but I guess whole book is like that)

②

examples

a) In 5 tosses of a coin → # heads = random variable

Seqence = **not**, no explicit # vale

b) 2 rolls of dice. Are random variables:

- Sum of the 2 rolls
- The # of 6
- $(2nd\ roll)^5$

c) Transmitting message. Are random variables

- time transmitted
- # of symbols in error
- delay

Main concepts

- random variable - real valued outcome of experiment
- function (random variable) = another random variable
- We can calculate mean / variance
- random variable can be conditioned on an event
- independence from a event or other random variable

(3)

<u>discrete</u> = ~~the~~ range (set of ~~p~~inputs) is finite or countably infinite

<u>non discrete</u> – uncountably infinite

$a \to a^2$
not discrete

This chap → all discrete

discrete values have a <u>probability mass function (PMF)</u>

- gives probability of each numerical value the random variable can take

– a function of a discrete random variable → defines another random variable
PMF can be obtained

<u>PMF Probability Mass Functions</u>

~~PMF~~ PMF of $X = p_X$

x ~~X~~ = possible value of X $= p_X(x) = P\{X = x\}$

- all outcomes where $X = x$

<u>example</u>

2 tosses of fair coin

$X = \#$ heads

$$p_X(x) = \begin{cases} 1/4 & \text{if } x=0 \text{ or } x=2 \\ 1/2 & \text{if } x=1 \\ 0 & \text{otherwise} \end{cases}$$

- Omit braces when can be no ambugity

  So $P(X=x)$ insted of $P(\{X=x\})$

- $P(X \in S)$ prob that X is in the set S

* Convention

UPPER CASE → random variables

lower case → real #, such as results of random variable,

$$\sum_{x} P_X(x) = 1$$

* sum of all possible x (values of X)
  - disjoint / partition

$$P(X \in S) = \sum_{x \in S} P_X(x)$$

So in example above

$$P(X > 0) = \sum_{x=1}^{2} P_X(x) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

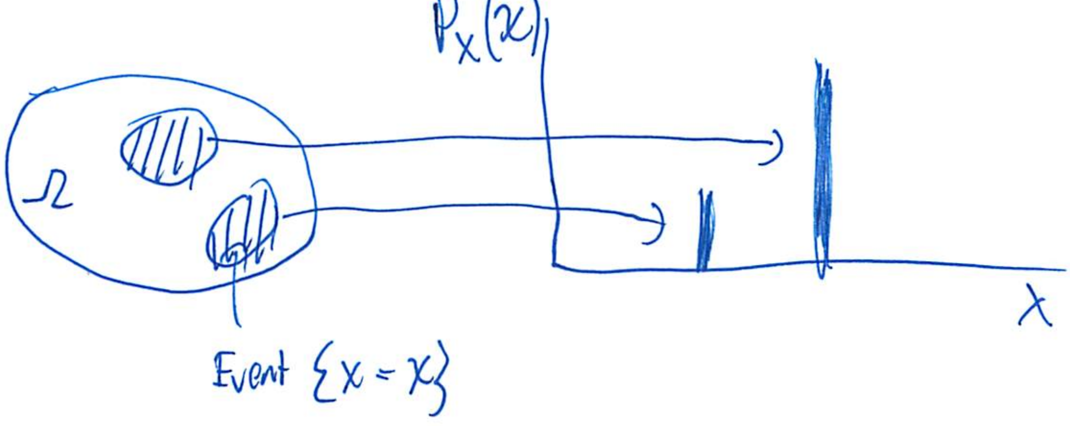
Calculation Steps

For ~~all~~ each possible ~~values~~ value x of X:

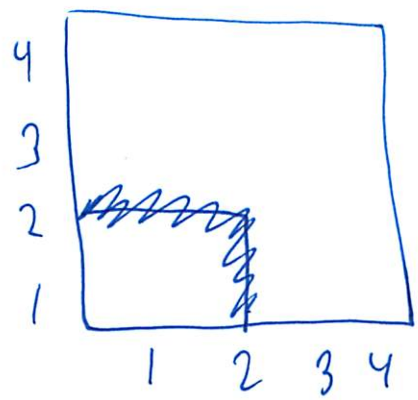
1. Collect all possible outcomes that give rise to event $\{X = x\}$
2. Add their prob. to obtain $P_X(x)$

4.b

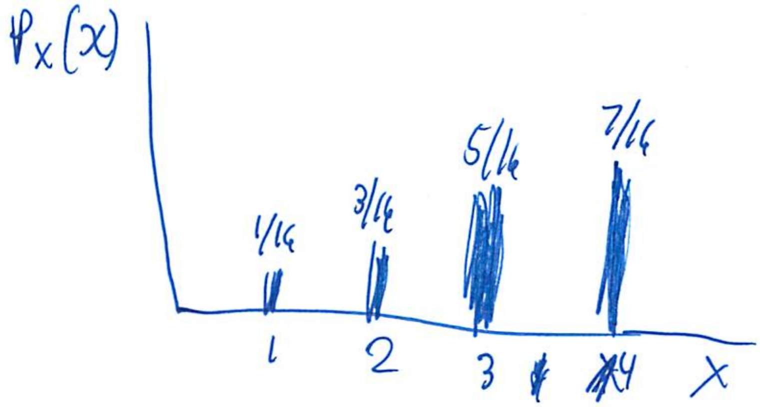
Nice Pictures



Pair of Rolls



random variable: X = max roll



(5)

So confusing both cap X and lower x
- hard for me to distinguish in handwriting

Bernoulli Random Variable

ie coin toss

Head : p

Tail : 1-p

Bernalli takes the 2 values 1, 0

$$X = \begin{cases} 1 & \text{if Head} \\ 0 & \text{if Tail} \end{cases}$$

So PMF

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

- (this seems stupid) ← me
- actually very important! ← book
- Used to model generic probablistic situations w/ 2 outcomes. ie:
  - State of telephone · free, busy
  - person and disease healthy, sick
  - person + canidate voted, againts

Can Combine Bernalli random variables to construct more Complex random variables

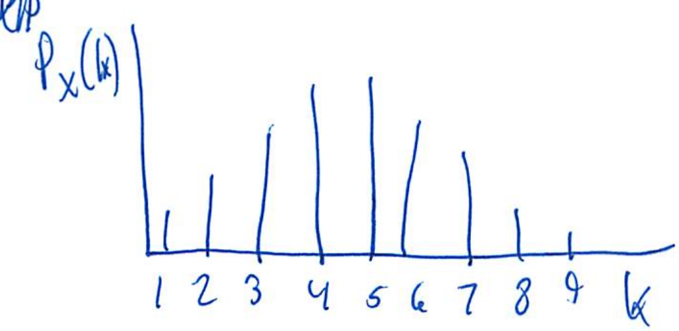
- like binomial random variable

Binomial Ranon Variable

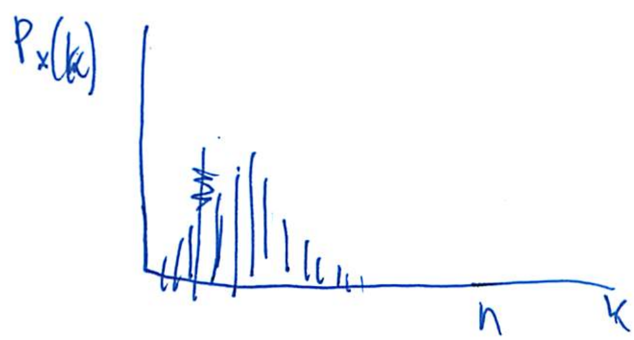
- coin tossed n times
  - Heads p
  - Tails 1-p
- X = # of heads in n tosses
  - ↑ binomial random variable w/ parametes n and p
  - PMF:

$$p_X(k) = P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k=0,1,2,\ldots,n$$

- oh used k instead of little x
  - thank goodness!



n = 9
p = 1/2



n = large
p = small

⑦ Geometric Random Variable

Suppose we can repeditly + independently toss a coin

Head → p $\quad 0 < p < 1$

geometric → number X of tosses need for a head to come up for the 1st time

PMF

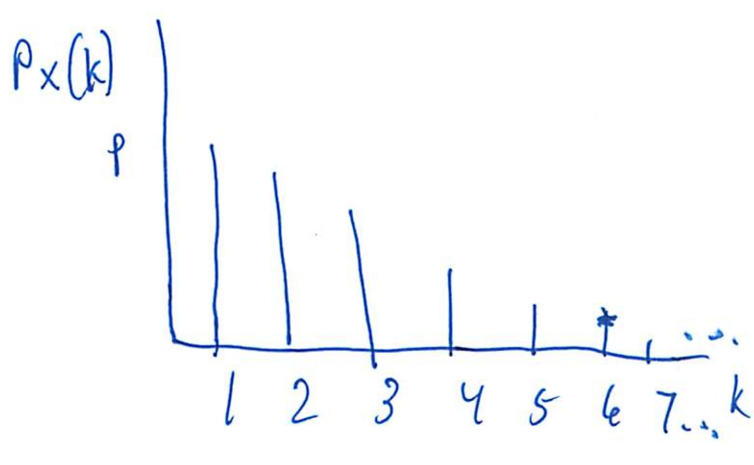
one 1

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

- probability of the sequence ~~Consistic by~~ consisting of $k-1$ sucessive fails followed by a head
  - (sounds like Tutorial 2 #1)

$$\sum_{k=1}^{\infty} p_X(k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = \sum_{k=0}^{\infty} (1-p)^k = p \cdot \frac{1}{1-(1-p)} = 1$$

- generally useful → # of trials till 1st sucess
  - Sucess is context dependant



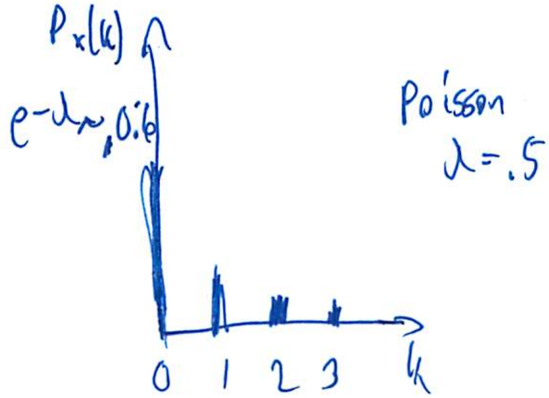
(8) Poisson Random Variable

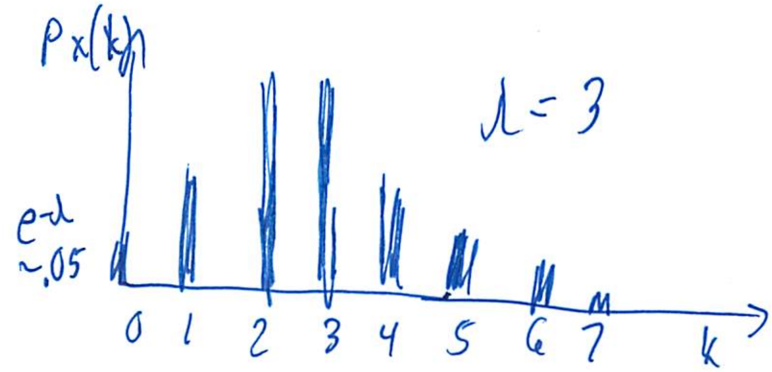
PMF

$$p_x(k) = e^{-\lambda}\frac{\lambda^k}{k!} \qquad k = 0, 1, 2 \ldots$$

λ = positive parameter characterizing the PMF

$$\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\right) = e^{-\lambda}e^{\lambda} = 1$$





- think of a binomial random variable w/ very small p + very large n

example

X = # typos

n = # of words

- X is binomial, but since P(any particular word misspelled)
  X can be modeled w/ Poisson PMF

- or # of accidents as part of # of cars in a city

(9)

- w/ λ → good approx for Binomial PMF w/ n, p

$$e^{-\lambda} \frac{\lambda^k}{k!} \approx \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \qquad k = 0, 1, \ldots, n$$

provided $\lambda = np$ ← very small
↑ very large

- ~~\$~~ by using Poisson you may get simpler models + calculations

example

$n = 100$ $\quad$ $P(k \text{ sucesses})$ in n trials $\quad$ $k = 5$
$p = .01$ $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ $n = 100$

Binomial

$$\frac{100!}{95!\,5!} \cdot .01^5 (1 - .01)^{95} = .00290$$

Poisson

$$\lambda = np = 100 \cdot .01 = 1$$

$$e^{-1} \frac{1}{5!} = .00306$$

formal justification later

↓

(10) 2.3 Functions of Random Variables

- given random variable X → can get other random variables
  - by applying transformations on X

<u>example</u>

X = today's temp in C

$Y = 1.8X + 32$ = todays temp in F

↑ linear function $Y = g(x) = ax + b$

- if X is random variable, Y will be too
  - PMF p_Y can be calculated
    - ↳ add the probabilities of all values of x where $g(x) = y$
      - basically all the inputs that give that output

$$p_Y(y) = \sum_{\{x \mid g(x) = y\}} p_X(x)$$

(skipping example - fairly obvious)

(11) ## 2.4 Expectation, Mean, and Variance

PMF of random variable X provides us w/ several #

↳ the probabilities of all the possible values of X

- but want to summarize to 1 # ⇒ the expectation of X
  - weighted (in proportion to probabilites) avg of possible values of X

example

$ wheel, each place diff amt of $
what is expected $ / spin?

$$M = \frac{m_1 k_1 + m_2 k_2 + \dots + \overset{\text{money}}{m_n} \overset{\text{area}}{k_n}}{k}$$

- well they did k as prob that the outcome will be that - a little more accurate than area or
  ⌒ amt of curve if wheel does not spin fairly

if k is large, and prob ≈ relative freq, it is reasonable m_i comes up a fraction of time that is ≈ to p_i

$$\frac{k_i}{k} \approx p_i \qquad i = 1, \dots, n$$

than

$$M \approx m_1 p_1 + m_2 p_2 + \dots + m_n p_n$$

(12)

$$\boxed{E[X] = \sum_{x} x \, p_X(x)}$$

definition of
- expected value
- expectation
- the mean

<u>example 2</u>

- 2 independent coin tosses

$P(\text{heads}) = \frac{3}{4}$

X = # of heads

so binomial random variable $n = 2$ $p = 3/4$

PMF

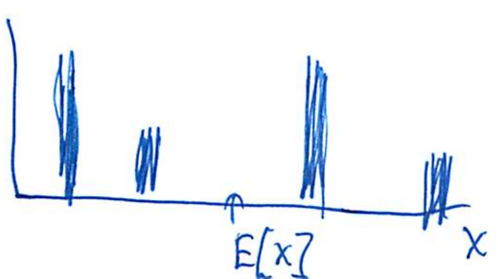
$$p_X(k) = \begin{cases} (1/4)^2 & \text{if } k = 0 \\ 2 \cdot (1/4) \cdot (3/4) & \text{if } k = 1 \\ (3/4)^2 & \text{if } k = 2 \end{cases}$$

so mean is

$$E[X] = \underbrace{0}_{\text{value}} \cdot \underbrace{\left(\tfrac{1}{4}\right)^2}_{\text{prob of it happening}} + 1 \cdot \left(2 \cdot \tfrac{1}{4} \cdot \tfrac{3}{4}\right) + 2\left(\tfrac{3}{4}\right)^2 = \frac{24}{16} = \frac{3}{2}$$

$E[X]$ is like the "representive" value of X
- near the middle of the range
- <u>Center of gravity</u>

<u>Picture</u>



13 Variance, Moments, and the Expected Value Rule

- several other quantites we can associate w/ random variable + PMF

2nd Moment of random variable

↳ expected value of X^2

nth Moment

of $E[X^n] \to X^n$

1st Moment = mean $E[X]$

Variance

$$var(x) = E[(x - E[x])^2]$$

always non negitive (+)

measure of dispersion of x around its mean

units ^{2}

Standard Deviation

$$\sigma_x = \sqrt{var(x)}$$

- eaiser to interpert b/c same units as X

Example

$$p_x(x) = \begin{cases} 1/9 & \text{if Int is in range } [-4,4] \\ 0 & \text{otherwise} \end{cases}$$

So $E[x] = 0 = \sum x \, p_x(x) = \frac{1}{9} \sum_{x=-4}^{4} x = 0$

$$Z = (X - E[x])^2 = X^2$$

(14)

$$p_Z(z) = \begin{cases} 2/9 & \text{if } z = 1, 4, 9, 16 \\ 1/9 & \text{if } z = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$Var(X) = E[Z] = \sum_z z\, p_Z(z) = 0 \cdot \frac{1}{9} + 1 \cdot \frac{1}{9} + 4 \cdot \frac{4}{9} + 9\frac{2}{9} + 16 \cdot \frac{2}{9} = \frac{60}{9}$$

~~A~~ Alt method

- does not require PMF of $(X - E[X])^2$
- X = random variable w/ PMF p_X
- $g(X)$ = function of X
- $\boxed{E[g(X)] = \sum_x g(x)\, p_X(x)}$

  - skipping verify + example

$$\boxed{Var(X) = \sum_x (x - E[X])^2 p_X(x)}$$

Properties of Mean ~~and Variance~~

$Y = aX + b$

$\boxed{E[Y]} = \sum_x (ax + b) p_X(x) = a \sum_x x\, p_X(x) + b \sum_x p_X(x) = \boxed{a\, E[X] + b}$

(so much eaiser!)

$\boxed{Var(Y)} = \sum_x (ax + b - E[aX + b])^2 p_X(x)$

$= \sum_x (ax + b - a\,E[X] - b)^2 p_X(x)$

$= a^2 \sum_x (x - E[X])^2 p_X(x) = \boxed{a^2\, Var(X)}$

(15)

Variance in Terms of Moments Expression

$$Var(x) = E[X^2] - (E[x])^2$$

- skipping verification + example

Mean + Variance of some Common Random Variables

Bernoulli

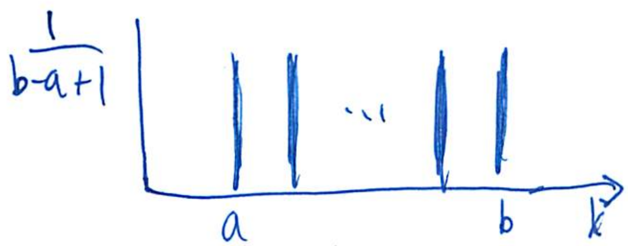
skipping some of the work

$$E[x] = p$$
$$E[x^2] = p$$
$$var(x) = p(1-p)$$

Discrete Uniform Random Variable

$$\cancel{Var(x) = E[x^2] - (E[x])^2}$$

$$E[x] = \frac{a+b}{2} \qquad var(x) = \frac{(b-a)(b-a+2)}{12}$$



- special case of discrete uniformly distributed random variable
  - aka discrete uniform
  - in the form

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

a, b are integers $a < b$

So where $a=1$ and $b=n$

$$E[x^2] = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{1}{6}(n+1)(2n+1)$$

$$Var(x) = E[x^2] - (E[x])^2$$
$$= \frac{1}{6}(n+1)(2n+1) - \frac{1}{4}(n+1)^2$$
$$= \frac{1}{12}(n+1)(4n+2-3n-3)$$
$$= \frac{n^2-1}{12}$$

- uniformly distributed over $[a,b]$ has same variance as $[1, b-a+1]$

So $Var(x) = \dfrac{(b-a+1)^2-1}{12} = \dfrac{(b-a)(b-a+2)}{12}$

Mean of Poisson

$E[x] = \lambda$ — skipping calculation

↳ note that $\sum_{m=0}^{\infty} e^{-\lambda}\frac{\lambda^m}{m!} = \sum_{m=0}^{\infty} p_x(m) = 1$

is normilization property for Poisson PMF

Decision Making Using Expected Values

- can calculated $E[x]$ of each choice and its subchoice to know what is best
- game show example

(17) 2.5 Joint PMFs of Multiple Random Variables

Probablistic models often involve several random variables

- in medical ; several tests play a role

all are part of same experiment, sample space, and prob. law

joint PMF

$p_{X,Y}$

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

$$\text{or } P(\{X=x\} \cap \{Y=y\})$$

^ union of disjointed events

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

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

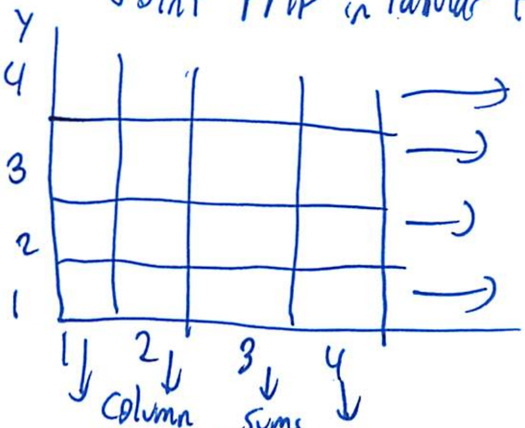
p_X and p_Y are the marginal PMFs

calculate joint PMF from marginal PMFs w/ tabular method

- 2 D table
- add row + column
- $p_Z(z) = \sum_{\{(x,y) \mid g(x,y) = z\}} p_{X,Y}(x,y)$

18

Joint PMF in tabular form



rows sum

Marginal PMF $p_Y(y)$

Column sums

Marginal PMF

$p_X(x)$

Expected/mean value extended

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

↳ like $E[aX+bY+c] = aE[X]+bE[Y]+c$

More than 2 Random Variables

$$p_{X,Y,Z}(x,y,z) = P(X=x,\ Y=y,\ Z=z)$$

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

$$E[aX+bY+cZ+d] = aE[X]+bE[Y]+cE[Z]+d$$

Mean of the Binomial

$$X_i = \begin{cases} 1 & \text{if ith student gets an A} \\ 0 & \text{otherwise} \end{cases}$$

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

↑ common mean $p = 1/3$

Since X is # of sucesses in n indep. trials it is a binomial random variable w/ parameters n + p

$$E[X] = \sum_{i=1}^{300} E[X_i] = \sum_{i=1}^{300} \frac{1}{3} = 300 \cdot \frac{1}{3} = 100$$

for n students and p of A =

$$E[x] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

2.6 Conditioning

- Conditional probabilites can capture info conveyed by varias events abot the different possible valves of a random variable
- ie given a certain events
- Same concepts as chap 1, but new notation

Conditional PMF

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

- note that events $\{X = x\} \cap A$ are disjoint for diff values of x
  Union is A

$$P(A) = \sum_x P(\{X = x\} \cap A)$$

So combo ~~[illegible]~~ $\sum_x p_{X|A}(x) = 1$

(20)

- Conditional calculated like unconditional
  - add probabilites that outcomes $X = x$ <u>and</u> belong to A
  - Then normalize by $/P(A)$

<u>Example</u>

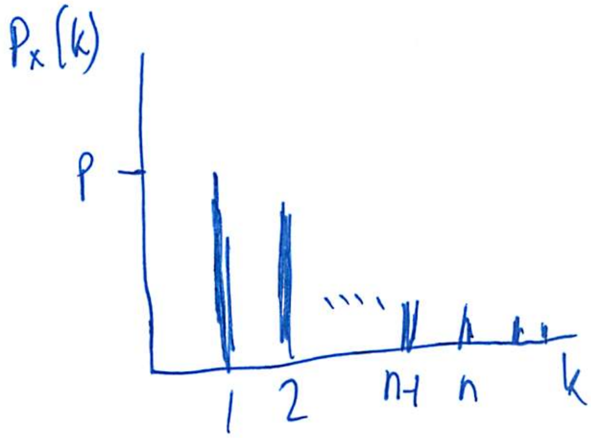
X = roll of fair sided die

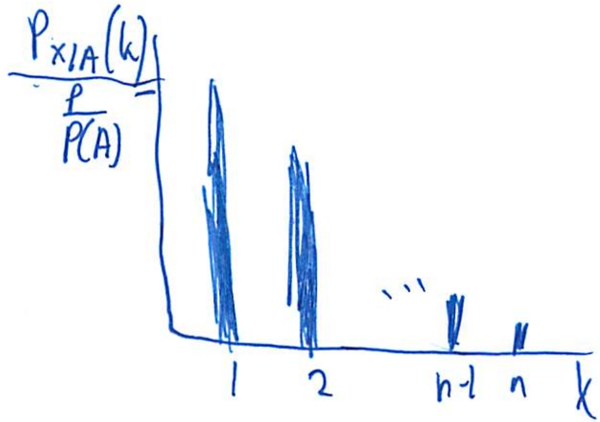
A = event that roll is even

$$P_{X|A}(k) = P(X = k \mid \text{roll even})$$

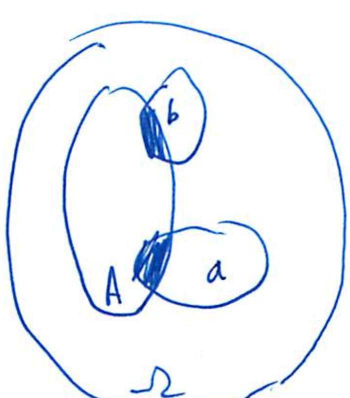
$$= \frac{P(X = k \text{ and roll is even})}{P(\text{roll is even})}$$

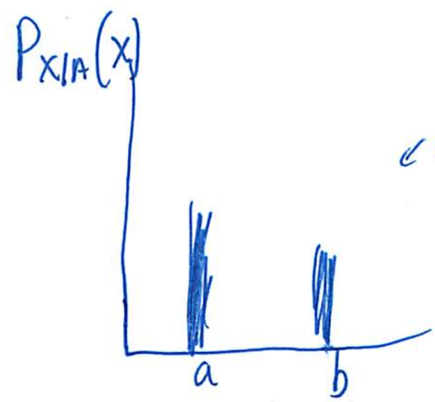
$$= \begin{cases} 1/3 & \text{if } k = 2, 4, 6 \\ 0 & \text{Otherwise} \end{cases}$$





Abstract visualization





← area that overlaps

21

Conditioning One Random Variable on Another

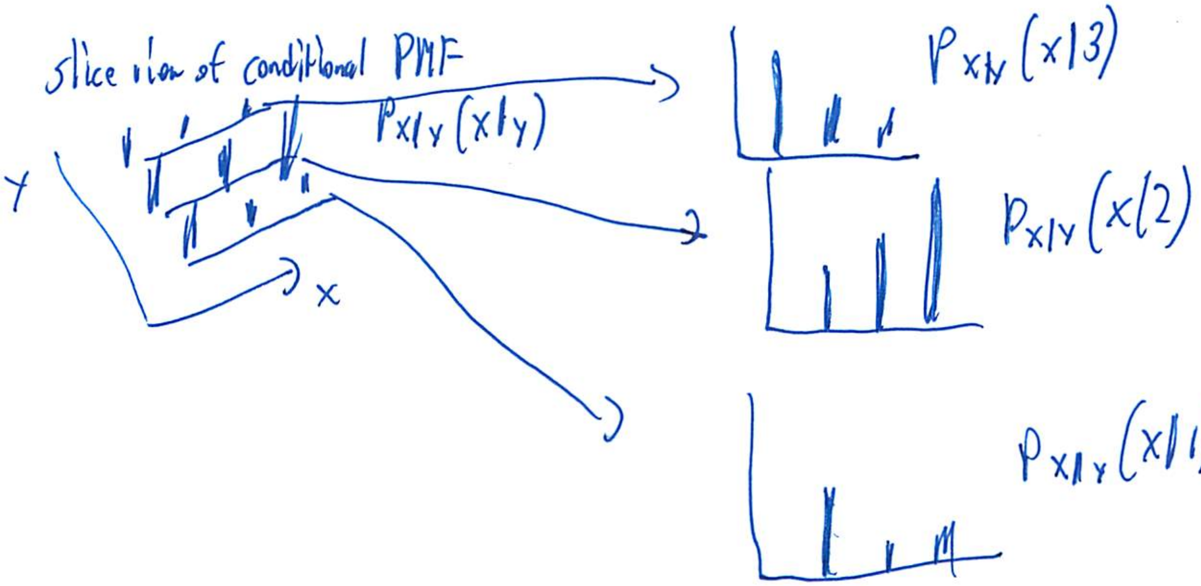
Conditional PMF

$$p_{X|Y}(x|y) = P(X=x \mid Y=y)$$

$$= \frac{P(X=x,\ Y=y)}{P(Y=y)}$$

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

$$\sum p_{X|Y}(x|y) = 1$$



conditional PMF often helpful to calc joint PMF

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

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

(22)

Example

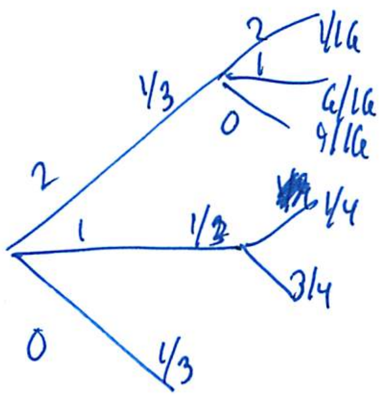
- prof answers qu correct $\frac{3}{4}$
- prof asked 0, 1, 2 qu (1/3 prob each)

X = # qu asked
Y = # q wrong

Calculate joint PMF $p_{X,Y}(x,y)$
by calculating $P(X=x, Y=y)$ for all combos x and y

ie if 1 qu asked and wrong

$$p_{X,Y}(1,1) = p_X(x) \cdot p_{Y|X}(y|x) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}$$



Y \ X	0	1	2
2	0	0	1/48
1	0	4/48	6/48
0	16/48	12/48	9/48

$\boxed{X}$ # qu asked

~~correct~~ wrong $\boxed{Y}$

can also make this 2D table
calc p of any event of interest

$$P(\text{at least 1 wrong}) = p_{X,Y}(1,1) + p_{X,Y}(2,1) + p_{X,Y}(2,2)$$
$$\frac{4}{48} + \frac{6}{48} + \frac{1}{48}$$

Conditional PMF can also be used to calculate marginal PMF

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

(23)

- provides a divide + conqur approach to calc marginal PMF

Conditional Exception

- can think of conditional PMF as ordinary PMF over a new universe determined by the conditioning event (like python w/ env)

let X + Y be random variables associated w/ same experiment

- expectation of X given A, $P(A) > 0$

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

- w/ function $g(x)$

$$E[g(x)|A] = \sum_x g(x) p_{X|A}(x)$$

Conditional expectation of X given value y of Y

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

Total Expectation theorm

If $A_1, \ldots, A_n$ are disjoint / partition $P(A_i) > 0$ for all i

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

for any event B with $P(A_i \cap B) > 0$ for all i

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

Lastly $$E[X] = \sum_y p_Y(y) E[X|Y=y]$$

- the last 3 follow from the total probability theorm
  "the unconditional average can be obtained by averaging the conditional averages"
- Skipping verification + examples

2.7 Independence

- similar to Chap 1

Independence of a Random Variable From an Event

- independence of a random variable is similar to independence of 2 events
- Conditioning event provides no additional info

X is ind. of A if

$$P(X = x \text{ and } A) = P(X = x)P(A) = p_X(x)P(A) \text{ for all } x$$

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

$$\text{if } P(A) > 0 \Rightarrow p_{X|A}(x) = p_X(x) \text{ for all } x$$

Independence of Random Variables

- 2 variables are inp. if

$$p_{X,Y}(x, y) = p_X(x)\, p_Y(y) \text{ for all } x, y$$

(25)

- same as req. $\{X = x\}$ and $\{Y = y\}$ be independent for every x and y

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

equivilant to

$$p_{X|Y}(x|y) = p_X(x) \text{ for all } y \text{ w/ } p_Y(y) > 0 \text{ and all } x$$

- aka independence of Y presents no info about X
- Similar notion of conditional independence of 2 random variables given an event A w/ $P(A) > 0$
  - conditioning event A defines a new universe
    - all prob/PMF are redefined by conditional
- X and Y are conditionally independent given A, if

$$P(X = x, Y = y | A) = P(X = x | A)\, P(Y = y | A) \text{ for all } x, y$$

$$p_{X|Y,A}(x|y) = p_{X|A}(x) \text{ for all } x + y \text{ such that } p_{Y|A}(y) > 0$$

- Conditional independence may not imply unconditional independence

$$E[XY] = E[X]\, E[Y]$$

$$E[g(X)\, h(Y)] = E[g(X)]\, E[h(Y)]$$

- (I should try prooving, to help my proof skills)

26

$$Var(X+Y) = var(\tilde{X}+\tilde{Y})$$
$$= \ldots$$
$$= Var(\tilde{X}) + var(\tilde{Y})$$
$$= var(X) + var(Y)$$

the variance of the sum of 2 indp. random variables
= sum of their variances

Independence of Several Random Variables

3 variables are inp if

$$p_{X,Y,Z}(x,y,z) = p_X(x)\,p_Y(y)\,p_Z(z) \text{ for all } x, y, z$$

- (duh)
- if 3 indp. variables, can't interperate anything
- book actually says intuitivly obvious
  - but tedious to verify

Variance of the Sum of Ind. Random Variables

- important in many contexts
  - like averaging something
- deal w/ cumulitive effects of several sources of randomness

$$\boxed{Var(X_1 + X_2 + X_3 + \ldots + X_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n)}$$

- built on formula $var(X+Y) = var(X) + var(Y)$

↓

27

2.8 Summary + Discussion

Random variables provide natural tools for dealing w/ probablistic models in which the outcome determins certain numerical values of interest

PMF, mean, + variance describe the discrete random variable

-not going to copy summary now

9/27

LECTURE 6

- **Readings:** Sections 2.4-2.6

More Discrete Random Variables

→ numerical outcomes

Lecture outline

- Review: PMF, expectation, variance
- Conditional PMF
- Geometric PMF
- Total expectation theorem
- Joint PMF of two random variables

– little concepts new this week

– mostly notation – PMF

– E[ ] and var( ) new

~~PMF~~

– easiest material

Review

← words eupimisim

- Random variable X: function from sample space to the real numbers

→ describes random outcome of experiment

- PMF (for discrete random variables): $p_X(x) = \mathbf{P}(X = x)$

X = random variable

x = the value that was picked

$\sum_x p_x(x) = 1$

- Expectation:

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

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

$\neq g(E[x])$ in general except linear functions

$$\mathbf{E}[\alpha X + \beta] = \alpha \mathbf{E}[X] + \beta$$

- $\mathbf{E}\Big[X - \mathbf{E}[X]\Big] =$

how far is it from the mean?

$E[x] - E[x] = 0$ how much from left + right – average = 0

$$\begin{aligned} \text{var}(X) &= \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] \\ &= \sum_x (x - \mathbf{E}[X])^2 p_X(x) \\ &= \mathbf{E}[X^2] - (\mathbf{E}[X])^2 \end{aligned}$$

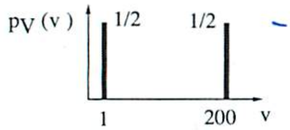
← how far from mean, abs.

Standard deviation: $\sigma_X = \sqrt{\text{var}(X)}$

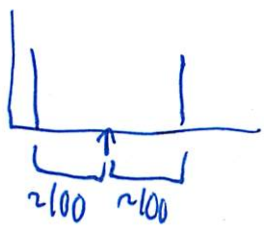
exactly how spread out distro. is

Random speed

- Traverse a 200 mile distance at constant but random speed V — either 1 or 200
  - flip a fair coin

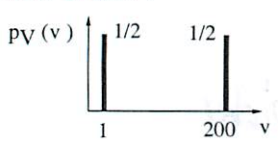


- $d = 200$, $T = t(V) = 200/V$
  ↑ function of random variable
- $\mathbf{E}[V] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 200 = 100.5$ avg speed
- $\text{var}(V) = \frac{1}{2}(1 - 100.5)^2 + \frac{1}{2}(200 - 100.5)^2 \approx 10{,}000$ hard to interperat
- $\sigma_V = \sqrt{10{,}000} = 100$ more meaningful



Average speed vs. average time

- Traverse a 200 mile distance at constant but random speed V



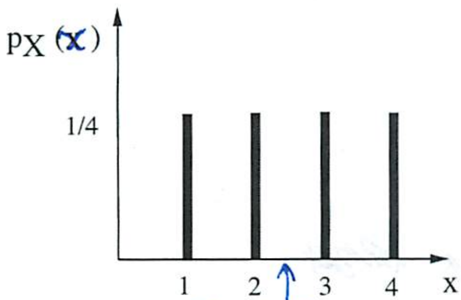
- time in hours $= T = t(V) =$ $200/V$
- time $\mathbf{E}[T] = \mathbf{E}[t(V)] = \sum_v t(v) p_V(v) =$ $\frac{1}{2} \cdot 200 + \frac{1}{2} \cdot 1 = 100.5$
- $\mathbf{E}[TV] = 200 \neq \mathbf{E}[T] \cdot \mathbf{E}[V]$ no matter what
  10,000
- $\mathbf{E}[200/V] = \mathbf{E}[T] \neq 200/\mathbf{E}[V]$.
  ~100 ~$\frac{200}{100}$

$$\mathbf{E}[g(V)] \neq g(\mathbf{E}[V])$$

always a conditional counter part

Conditional PMF and expectation

- $p_{X|A}(x) = \mathbf{P}(X = x \mid A)$
- $\mathbf{E}[X \mid A] = \sum_x x p_{X|A}(x)$

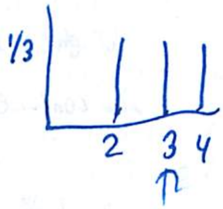


E[X]=2.5

- Let $A = \{X \geq 2\}$ ← know event has occured → now in new universe

$p_{X|A}(x) = \frac{1}{3}$ $x = 2, 3, 4$

$\mathbf{E}[X \mid A] = 3$



$$E[g(x) \mid A] = \sum_x g(x) P_{X|A}(x)$$

Geometric PMF

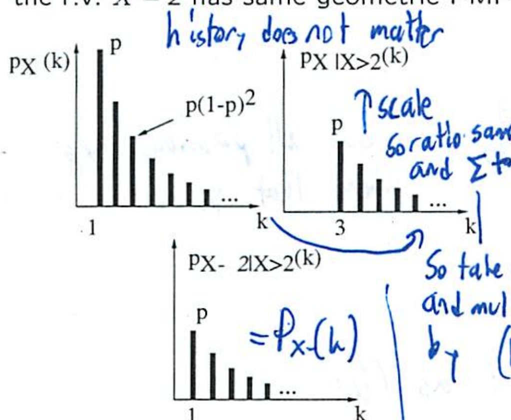
- X: number of independent coin tosses until first head for 1st time

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

$$\mathbf{E}[X] = \sum_{k=1}^{\infty} k p_X(k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p$$

- Memoryless property: Given that $X > 2$, the r.v. $X - 2$ has same geometric PMF

history does not matter



↑ scale so ratio same and Σ to 1

So take diagram and multiply (1) by $(1-p)^2$

$= P_X(k)$

# of remaining coin tosses

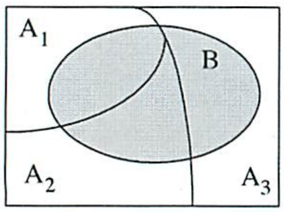
Lets say person got 2 tails already

- P(heads | 2 tails already) = p
- same! - past history does not matter "memoryless"
- $X-2$ = # of remaining coin flips ↳ given $X > 2$

$$P_X(k) = P_{X-2|X>2}(k)$$

Total Expectation theorem

- Partition of sample space
  into disjoint events $A_1, A_2, \ldots, A_n$



$B = \{X = x\}$

divide + conqur

$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B \mid A_n)$

$p_X(x) = \mathbf{P}(A_1)p_{X|A_1}(x) + \cdots + \mathbf{P}(A_n)p_{X|A_n}(x)$ } rewrite in chap 2 notation

$\mathbf{E}[X] = \mathbf{P}(A_1)\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X \mid A_n]$

weighted average — average of whole class = weighted avg of each section

~~forget~~

- Geometric example:
  $A_1: \{X = 1\}$, $A_2: \{X > 1\}$
  heads stop, tails contine

$$\mathbf{E}[X] = \mathbf{P}(X=1)\mathbf{E}[X \mid X=1] + \mathbf{P}(X>1)\mathbf{E}[X \mid X>1]$$

$p \cdot 1 + (1-p) \cdot$ ____

- Solve to get $\mathbf{E}[X] = 1/p$

$E[X] = p \cdot 1 + (1-p)(1 + E[X])$
$= 1/p$

short cut, based on divide + conqur and intuition

$E[X \mid X>1] = E[X-1 \mid X>1] + 1$
$= E[X] + 1$ proved on last sheet

recursive!

Joint PMFs

- $p_{X,Y}(x,y) = \mathbf{P}(X = x \text{ and } Y = y)$

have $P_X(x)$ and $P_Y(y)$
how do they relate to one another?
need to know joint PMFs
– in order to answer $P(X \le Y)$

some Ω
X, Y are $f(\Omega)$
– but really disappear in these problems

y \ x	1	2	3	4
4	1/20	2/20	2/20	0
3	2/20	4/20	1/20	2/20
2	0	1/20	3/20	1/20
1	0	1/20	0	0

- $\sum_x \sum_y p_{X,Y}(x,y) = 1$

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

$P_X(4) = \frac{2}{20} + \frac{1}{20} = \frac{3}{20}$ sum over all possible ys where that x

- $p_{X|Y}(x \mid y) = \mathbf{P}(X = x \mid Y = y) = \dfrac{p_{X,Y}(x,y)}{p_Y(y)}$
  conditional PMF
  – old concept
  – new notation

- $\sum_x p_{X|Y}(x \mid y) =$

– function of 2 variables
– fix a Y and think of as $f(x)$
ie $y = 2 \Rightarrow P_x(x) = \begin{cases} 1/5 & \text{if } x = 2, 4 \\ 3/5 & \text{if } x = 3 \\ 0 & \text{otherwise} \end{cases}$

must add to 1

based on that assumption, rescale

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

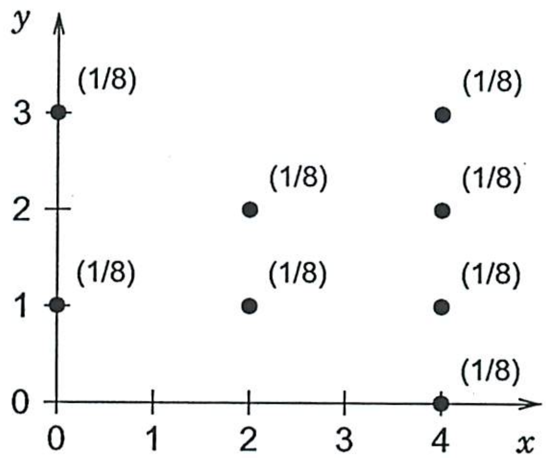
Recitation 6
September 28, 2010

1. Consider an experiment in which a fair four-sided die (with faces labeled 0, 1, 2, 3) is thrown once to determine how many times a fair coin is to be flipped. In the sample space of this experiment, random variables N and K are defined by

   - $N =$ the result of the die roll
   - $K =$ the total number of heads resulting from the coin flips

   (a) Determine and sketch $p_N(n)$
   (b) Determine and tabulate $p_{N,K}(n, k)$
   (c) Determine and sketch $p_{K|N}(k \mid 2)$
   (d) Determine and sketch $p_{N|K}(n \mid 2)$

2. Consider an outcome space comprising eight equally likely event points, as shown below:



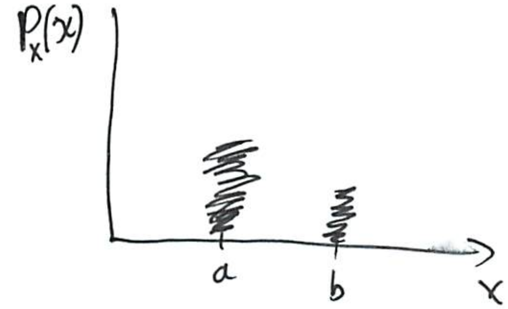
   (a) Which value(s) of x *maximize(s)* $\mathbf{E}[Y \mid X = x]$?
   (b) Which value(s) of y *maximize(s)* $\text{var}(X \mid Y = y)$?
   (c) Let $R = \min(X, Y)$. Prepare a neat, fully labeled sketch of $p_R(r)$,
   (d) Let A denote the event $X^2 \geq Y$. Determine numerical values for the quantities $\mathbf{E}[XY]$ and $\mathbf{E}[XY \mid A]$.

3. **Example 2.17. Variance of the geometric distribution.** You write a software program over and over, and each time there is probability p that it works correctly, independent of previous attempts. What is the variance of X, the number of tries until the program works correctly?

Recitation 6

9/26

Discrete Random Variables



collect, add all probabilities that event a occures
or value of a

$E[X]$

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

$$E[g(X)] = \sum_x g(x) P_X(x)$$ expected value rule

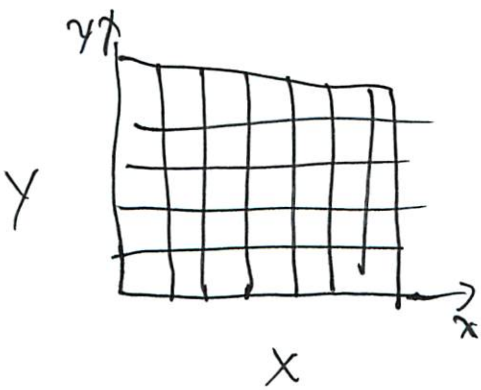
X, Y related values

$$P_{X,Y}(x,y) = P(\{X=x, Y=y\})$$

Marginal PMFs

$$P_X(x) = \sum_y P_{X,Y}(x,y),$$ column sum

$$P_Y(y) = \sum_x P_{X,Y}(x,y)$$ row sum



(2)

$$\boxed{E[g(X,Y)] = \sum_{x,y} g(x,y)\, P_{X,Y}(x,y)}$$

- talk more about later, next time

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

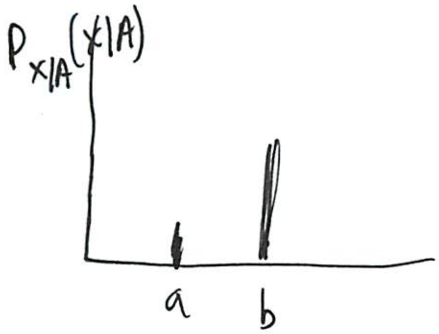
Condition on Event A

$P(A) > 0$

Define $P_{X|A}(x|A) = P(X=x|A) = \dfrac{P(\{X=x\} \cap A)}{P(A)}$

$E[X|A] = \{x\, P_{X|A}(x|A)\}$

$\mathrm{var}(X|A) = \ldots$



Only where overlap

divisor is related to A

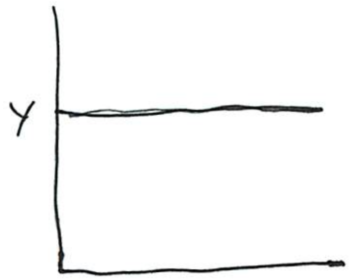
---

2 random variables ; one conditional on the other

$$\underbrace{}_{X} \qquad \underbrace{}_{Y}$$

$$P_{X|Y}(x|y) = P(\{X=x\} | \underbrace{\{Y=y\}}_{A}) = \frac{P(\{X=x, Y=y\})}{P(\{Y=y\})} = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

③ How to visualize this in terms of a table?



Must be scaled by common value

- PMF of Y

Slice at y the joint PMF

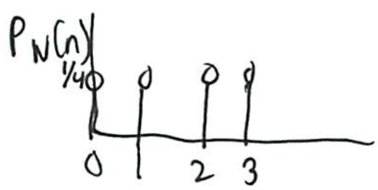
- take that row

Normalize by dividing by $P_X(y)$

- Shape remains the same
- move up and down

Example: roll a 4 sided die

$N = 0, 1, 2, 3$ ← 4 sides of die [map outcomes to #]



Now toss a fair coin N times

k = # of heads

2 random variables

- what is joint PMF

k \ n	0	1	2	3
3	0	0		
2	0	0		
1	0	1/2		
0	1	1/2		

P don't toss, so certain 1 head

#(4)

<u>But</u> all # must add up to 1

- I thought it was strange
- Must rescale

$$P_{N,K}(n,k) = \frac{1}{4} \cdot \text{Binomial}_n(k)$$
$$= P(N) \cdot P(k|N)$$

↑ by definition

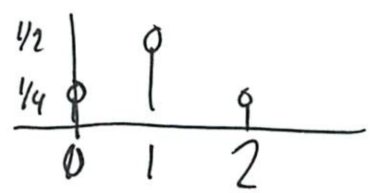
3	0	0	0	1/32
2	0	0	1/16	3/32
1	0	1/8	1/8	3/32
0	1/4	1/8	1/16	1/32
	0	1	2	3

Now want

$$P_{k|n}(k|2)$$

- slice + normalize

$$\begin{matrix} 0 \\ 1/16 \\ 1/8 \\ 1/16 \end{matrix} \cdot 4 = \begin{matrix} 0 \\ 1/4 \\ 1/2 \\ 1/4 \end{matrix}$$



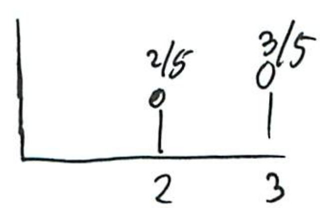
$$P_{n|k}(n|2)$$

- not binomial

$0 \quad 0 \quad 1/16 \quad 3/32$

$\cdot$

$\frac{2}{5} \quad 3/5$



(5)

Divide + Conquer

Helpful

Divide into smaller sets

$$A = A_1 \cup A_2 \cup \cdots \cup A_n \quad \text{Partition}$$

$$P_{X|A}(x|A) = P(A_1)P_{X|A_1}(x|A_1) + \cdots P(A_n)P_{X|A_n}(x|A_n)$$

$$P(\{X=x\}|A) =$$

- total prob. theorem
- for special events

TPT for 2 random variables
- marginal PMF as weighted sum of conditional PMFs

$$P_X(x) = \sum_{y} P_Y(y)\, P_{X|Y}(x|y)$$

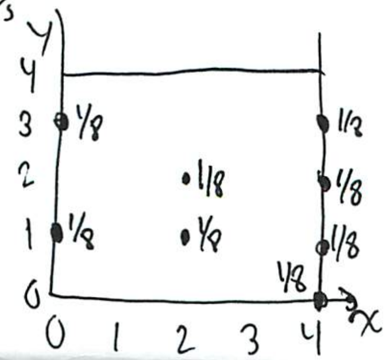
partition values into individual components

$$A_i = \{Y = y\}$$

to calc marginal PMFs

Example X, Y

joint PMFs



Calculate conditional expectation

$$E[Y|X=x] = \begin{cases} 2 & \text{if } x=0 \\ \text{und} & \text{if } x=1,3,\text{others} \\ 1.5 & \text{if } x=2 \\ 1.5 & \text{if } x=4 \end{cases}$$

Calc conditional PMY of Y given $X=0$, $X=1\ldots$

(e) $var(X \mid Y=y)$ - for given values of y, calc $var(x)$
marginal PMF

~~[illegible]~~

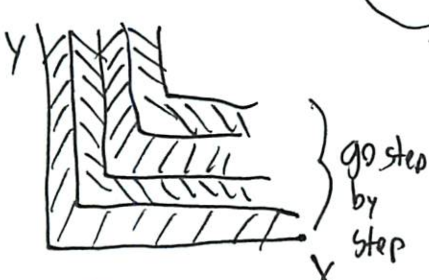
Slice + normalize

$$= \begin{cases} 0 & \text{if } y=0 \\ 8/3 & y=1 \\ 1/4 & y=2 \\ 4 & y=3 \\ \text{und} & y=4 \end{cases}$$

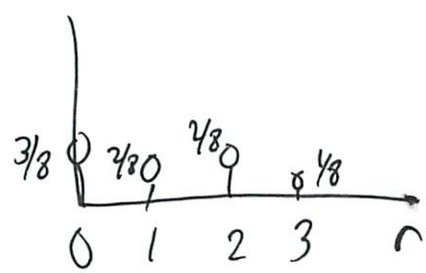
note going

Calc $P_R(r)$ $\quad R = \min(X, Y)$

$\{$ ← generate w/ collect + record



refers to existing data



$E[XY] = \ldots$ do on own

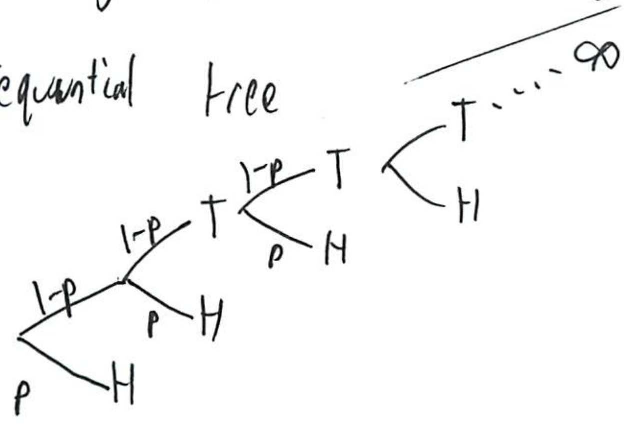
$E[XY \mid R] = \ldots$ just apply the formula

⑦

Last Problem:

Calculate geometric random variable

- Sequential tree



$P(\text{head}) = p$

$X =$ # of tosses until 1st head
↳ k is actual #

$P_X(3) = (1-p)^2 p$

$$P_X(k) = \begin{cases} (1-p)^{k-1} p & \text{if } k = 1, 2, 3 \ldots \\ 0 & \text{else} \end{cases}$$

(*) Make sure to define (*) domain

$E[X] = \frac{1}{p}$

- ~~part of the tree~~
- Can see in tree!
- Whereever you are when you look forward it is the same as if you are starting for 1st time

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

↑ actually easier to calc here

$$E[X^2] = P(X=1) \rightarrow E[X^2 \mid X=1] + P(X>1) \rightarrow E[X^2 \mid X>1]$$

8

$$= p \cdot 1 + (1-p) \cdot E[x^2 \mid x > 1]$$

looks like could do just

$E[x^2]$ —

but can't, need to shift by 1

$$= p \cdot 1 + (1-p) \cdot E[(x+1)^2]$$

$$= p \cdot 1 + (1-p) \cdot E[x^2 + 2x + 1]$$

$$= p \cdot 1 + (1-p) + 2(1-p)\frac{1}{p} + (1-p)\underbrace{E[x^2]}_{\text{only unknown left}}$$

$$E[x^2] = 2/p^2$$

- recursive
- refers to itself
- but group terms to find ans
  - duh
  - that was obvious

$$\text{var}(x) = E[x^2] - (E[x])^2$$

$$= \frac{2}{p^2} - \frac{1}{p^2}$$

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

---

know the binomial, etc

should be able to just use, don't need to reprove

Michael Plasmeier 7/10 9/25

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

Problem Set 3
Due September 29, 2010

1. The hats of n persons are thrown into a box. The persons then pick up their hats at random (i.e., so that every assignment of the hats to the persons is equally likely). What is the probability that
   (a) every person gets his or her hat back?
   (b) the first m persons who picked hats get their own hats back?
   (c) everyone among the first m persons to pick up the hats gets back a hat belonging to one of the last m persons to pick up the hats?

   Now assume, in addition, that every hat thrown into the box has probability p of getting dirty (independently of what happens to the other hats or who has dropped or picked it up). What is the probability that

   (d) the first m persons will pick up clean hats?
   (e) exactly m persons will pick up clean hats?

2. Alice plays with Bob the following game. First Alice randomly chooses 4 cards out of a 52-card deck, memorizes them, and places them back into the deck. Then Bob randomly chooses 8 cards out of the same deck. Alice wins if Bob's cards include all cards selected by her. What is the probability of this happening?

3. (a) Let X be a random variable that takes nonnegative integer values. Show that

$$\mathbf{E}[X] = \sum_{k=1}^{\infty} \mathbf{P}(X \geq k).$$

   Hint: Express the right-hand side of the above formula as a double summation then interchange the order of the summations.

   (b) Use the formula in the previous part to find the expectation of a random variable Y whose PMF is defined as follows:

$$p_Y(y) = \frac{1}{b-a+1}, \qquad y = a, a+1, \ldots, b$$

   where a and b are nonnegative integers with $b > a$. Note that for $y = a, a+1, \ldots, b$, $p_Y(y)$ does not depend explicitly on y since it is a uniform PMF.

4. Two fair three-sided dice are rolled simultaneously. Let X be the difference of the two rolls.
   (a) Calculate the PMF, the expected value, and the variance of X.
   (b) Calculate and plot the PMF of X^2.

5. Let $n \geq 2$ be an integer. Show that

$$\sum_{k=2}^{n} k(k-1)\binom{n}{k} = n(n-1)2^{n-2}.$$

   Hint: As one way of solving the problem, following from Example 1.31 in the text, think of a committee that includes a chair and a vice-chair.

G1†. A candy factory has an endless supply of red, orange, yellow, green, blue, black, white, and violet jelly beans. The factory packages the jelly beans into jars in such a way that each jar has 200 beans, equal number of red and orange beans, equal number of yellow and green beans, one more black bean than the number blue beans, and three more violet beans than the number of white beans. One possible color distribution, for example, is a jar of 50 yellow, 50 green, one black, 48 white, and 51 violet jelly beans. As a marketing gimmick, the factory guarantees that no two jars have the same color distribution. What is the maximum number of jars the factory can produce?

†Required for 6.431; optional for 6.041

Michael Plasmeier

6.041 Pset 3

IF YOU DON'T STAPLE YOUR PSET, I WILL GRADE ONLY THE FIRST PAGE.

— office hrs

9/26

1/4

1. n hats thrown into a box

each person picks a hat at random

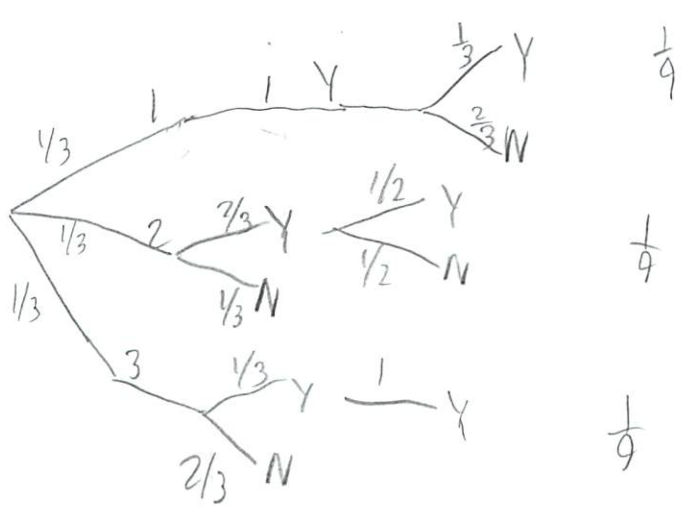
(ie the assignment of hats to each person is equal)

a) P(each person gets hat back) $\frac{1}{n}$

- assume they go 1st $\to \frac{1}{n}$
- assume they go 2nd
  - first person could have it $P = \frac{1}{n}$
  - then they have P of picking it $P = \frac{1}{n-1}$, if the person does not have it
  - So total $P = \frac{1}{n} + \frac{1}{n-1} \cdot \left(1 - \frac{1}{n}\right)$
- so just $\frac{1}{n}$?

−1

for n=3 picks still there gets it?



just draw the chart!
- so much easier to visualize
- but more writing
- need to learn formulas

$$\text{Sum} = \frac{3}{9} = \frac{1}{3} = \boxed{\frac{1}{n}}$$

−1

②

b) The first m persons who pick get hats back

- geometric?
- yeah since 1st is p
- 2nd is less than that $p \cdot p$?

p^m

and $p = \frac{1}{n}$ $\rightarrow$ $\boxed{\left(\frac{1}{n}\right)^m}$ -1

So for $n = 3$

$m = 1 \rightarrow \frac{1}{3}$
$m = 2 \rightarrow \frac{1}{9}$
$n = 3 \rightarrow \frac{1}{27}$

makes sense

c) Every one among the 1st m persons to pick up their hat gets one of the last m people

- now it can be any of the m people
- think w/ # $\rightarrow$ $n = 10$ $m = 5$
  - first 5 get the last 5
  - combinations or partitions

? want handful of m and $n-m$ objects

$\binom{n}{m \;\; n-m} = \boxed{\frac{n!}{m!\,(n-m)!}}$ ✓ confirmed in ~~the~~ OH

-0.5

? oh same thing anyway!

(3)

Now assume that every hat has $P = p$ of getting dirty (independent of anything else)

$1-p$ = clean

d) P (first m people will pick up clean hats)

- geometric
- like heads/tails

$$\prod_{i=1}^{n} P(A_i)$$

$$P_X(m) = (1-p)^m$$ +0.5

$$\prod_{i=1}^{n} (1-p)$$

related by not geometric

e) exactly m persons will get clean hats

- combination

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

✓ confirmed

−0.5

(4)

2. Alice randomly choses 4 cards from 52 deck, memorizes + replaces
Bob randomly chooses 8
P(All of 4 of Alice's cards are in Bob's hand)

- assuming Alice shuffles after replacing her cards
- like tutorial 2 #4
- Combo

- Ω - all possible 8 cards delt

$|\Omega| = \binom{52}{8}$ ← why did I write 7 ??

↑ fixed here

$|A| = \binom{4}{4} \cdot \binom{52-4}{4}$

given 4 cards choose all 4

need to pick rest of cards

$$P = \frac{\frac{4!}{4! \cdot 0!} \cdot \frac{48!}{4!\,44!}}{\frac{52!}{8!\,45!}} = \frac{\binom{48}{4}}{\binom{52}{8}} = \frac{9}{6188} = .00145$$

$= .145\%$

↑ can leave

↑ continue

- makes sense → very low

tip watch sizes and probabilities
which are you working with
don't multiply/add 2 different things

(5)

3. a) X = random value

Show $E[X] = \sum_{k=1}^{\infty} P(X \geq k)$

Hint: Express right side as double summation than interchange order of summation

$$E[X] = \sum_k k P_X(k)$$

$=$

- but how double summination?,
- there is only 1 variable

and. $\sum_{k=1}^{\infty} k = X$ so $X \geq k$

and $E[X]$ is adding all the k values

so $E[X] = \sum_k k + \sum_k P_X(k)$

(perhaps this is 2x summation

$E[X] = \sum k + P(X)$ ← where $X \geq k$
since $X = \sum k$

$E[X] = X + P(X)$

? but $E[X] \neq \sum P(X)$

? ask in OH

I think I got the hang of actual probability

- Just always bad at the symbolic/ math aspects
- w/ formulas

3. in OH: Proove $E[X] = \sum_{k=1}^{\infty} P(X \geq k)$

know $P(X \geq k) = \sum_{i=k}^{\infty} P_X(i)$

X discrete

$X > 0$

Probability of any value $= P_X(i)$ same

that function

$$\sum_{k=1}^{\infty} f(\,) = \sum_{k=1}^{\infty} \sum_{i=k}^{\infty} P_X(i)$$

split into pieces - actually more like a substitution, not splitting a sum

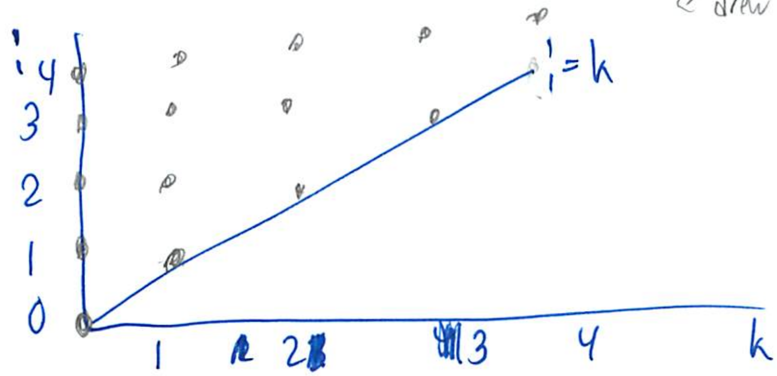
can interchange summations

$$= \sum_{i=1}^{\infty} \sum_{k=1}^{i} P_X(i)$$

evaluates to

$$= \sum i P(i) = E[X]$$

< draw poorly



Could say for every i sum over k

or for every k sum over i

How tie in w/ expectation
- just one step we did

$$\sum_{k=1}^{\infty} \sum_{i=k}^{\infty} = \sum_{i=1}^{\infty} \sum_{k=1}^{i}$$

summing vertically =
" horizontally

(Oh makes a lot more sense now!)

each sum adds PMF

next step

$$\sum_{i=1}^{\infty} P_X(i) \underbrace{\sum_{k=1}^{i} 1}_{= i}$$

+2

now
you have the expectation!

3b. Follow through same technique
- express prob. as sum

$Y \geq k$
- uniform prob
- k has certain values it takes

- line w/ uniform P

(6)

b) Use the formula in the previous part to find the expectation of a random variable Y whose PMF =

$$P_Y(y) = \frac{1}{b-a+1} \qquad Y = a, a+1, a+2, \ldots b$$

$a, b > 0$

$b > a$

- From book?:
- p89: special case of discrete uniformly distributed random variable aka discrete uniform

$$E[X] = \frac{a+b}{2}$$ ← PMF symmetric around

verify by induction

- $E[X^2] = \frac{1}{n} \sum_{k=1}^{n} k^2 = \frac{1}{6}(n+1)(2n+1)$

Ask in OH

+2

4/4

(6b)

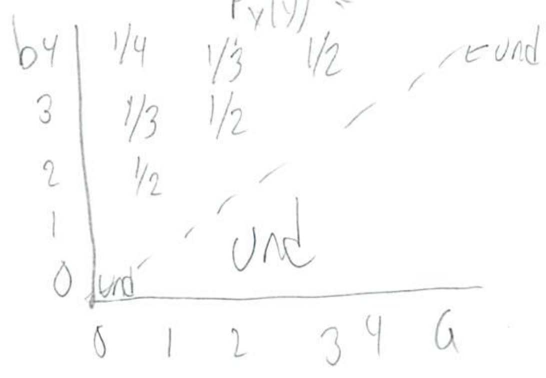
After Office Hrs On Own

$$E[X] = \sum i \, P(i)$$

$=$

? can draw

$P_Y(y) =$



? but what is the value of y?

- or should we look at these inputs?

how are a, b, y related

- or use that formula
- There is no k - do they mean y?

$$\sum_{b=1}^{\infty} \sum_{a=0}^{b-1} \frac{1}{b-a+1} \cdot y$$

✓ does not depend on y "uniform pmf"

? so how does that help?

$$\sum_{b=1}^{\infty} \sum_{a=0}^{b-1} \frac{1}{b-a+1} \cdot \sum_{i=0}^{\infty} a+i \cdot b$$

? is that right I never know when I have it

(7)

4. 2 fair 3-sided dice
X = difference of 2 cells

a) PMF

$P_X(k)$ = make chart 1st

	1	2	3
3	2	1	0
2	1	0	1
1	0	1	2

2 $\frac{2}{9}$
1 $\frac{4}{9}$
0 $\frac{3}{9} = \frac{1}{3}$

$$P_X(k) = \begin{cases} 2 & \text{if } x=3, y=1 \text{ or } x=1, y=3 \\ 1 & \text{if } \ldots \text{ (4 cases) } |x-y|=1 \\ 0 & \text{if } x=y \end{cases}$$

? circular definition

or is x the difference

$$P_X(k) = \begin{cases} \frac{2}{9} & \text{if } x=2 \\ \frac{4}{9} & \text{if } x=1 \\ 3/9 & \text{if } x=0 \end{cases}$$

← yeah I think thats it

a2) $E[X]$

$$E[X] = \sum_x x\, p_X(x)$$

mean value

$$= 2 \cdot \frac{2}{9} + 1 \cdot \frac{4}{9} + 0 \cdot \frac{3}{9}$$
$$= 8/9$$

(8)

a3) Var(x)

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

$$Z = (X - E[x])^2 = X^2$$

-Square of random variable

$$P_z(z) = \begin{cases} 2/9 & \text{if } z = 2^2 = 4 \\ 4/9 & \text{if } z = 1^2 = 1 \\ 3/9 & \text{if } z = 0^2 = 0 \end{cases}$$

↑ this stays same

→ in this case, since Z don't "overlap" when you square

→ if $X = 10 \to \frac{1}{2}$

$X = -10 \to \frac{1}{2}$

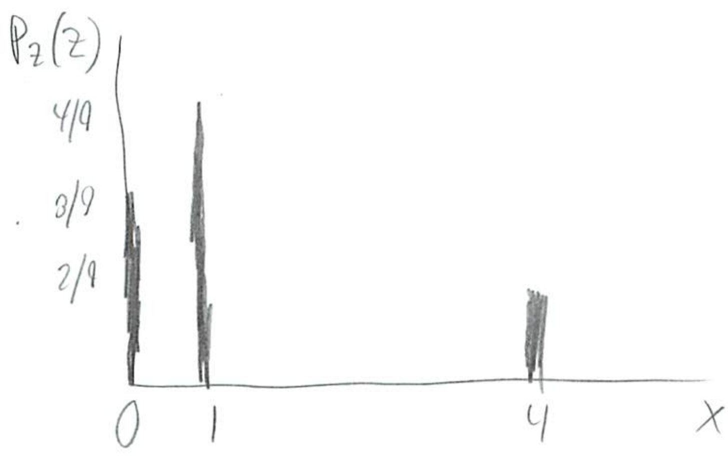
then would have $p = 1 \to x^2 = 100$

now $E[Z]$

$$4 \cdot \frac{2}{9} + 1 \cdot \frac{4}{9} + 0 \cdot \frac{3}{9} = \frac{12}{9} = Var(x)$$

b) PMF of X^2 did

b2) Plot -show



9

5. Let $n \geq 2$ be an integer. Show that

$$\sum_{k=2}^{n} k(k-1)\binom{n}{k} = n(n-1)2^{n-2}$$

- Look at example 1.31, think of committe w/ chair + vice chair

- n choices for leader
- for vice-leader → have n-1 canidates, can choose 2^{n-1} subsets

- or for fixed k we can choose a k-person club from n people $\binom{n}{k}$ possible choices
- add over all possible club sizes to find total # of clubs
- so for leader its permutation ${}_nP_2$
- but that is not problem, its proving the above

$$\sum_{k=2}^{n} k(k-1)\binom{n}{k} = n(n-1)2^{n-2}$$

Sum over club sizes from 2 → n people

pick leader out of k choices

pick vice leader w/ who is left

pick leader

vice leader

all of the rest possible sizes

additional people

1.31 :

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}$$

- pick leader
- then fill the rest
- we added permutations for all the diff. vice leaders

(10)

So what to do? Problem also talks about binomial formula

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$$

or when $p = 1/2$

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

- So that is all related, moving people from combo (order does not matter) to permutation (leader + vice matter)

$\underbrace{\text{leader} \quad \text{vice}}_{\text{permutation}}$ $\underbrace{\text{the rest}, n}_{\text{combo}}$ ← any # up to n

↳ gets right size

$_nP_2$

left counting graps in terms of size

$$\frac{n!}{(n-2)!} \cdot \frac{n!}{(n-2)!\,(n-n-2)!}$$

Binomial formula → must add to 1

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$$

- which is close to what we have
- but what do you write?

11

$$\sum_{k=a}^{n} \underbrace{k(k-1)(k-2)\cdots}_{a \text{ times}} \binom{n}{k} = \underbrace{n(n-1)(n-2)\cdots}_{a \text{ times}} 2^{n-a}$$

+2/2

5 in OH. How do you show a proof?

right → how many groups can you break people into
left → count up groups you can form

right

$\underline{n}, \underline{n-1}, \underline{\quad --- \;\; ---- \quad}$

subset size m

$m \geq n-2$

- lecture → calced total # of subsets from size n

binomial 2^n

- either in or not, so 2 choices for everything

$\underline{2}\ \underline{2}\ \underline{2}\ \underline{2}\ \underline{2}\ \underline{2}$

or think you have $n-2$ slots

can

- include them
- discard

← big point I missed

2^{n-2}

left

add up the groups

k = possible group size (total # you are including in group) you are considering now

- add up all of the k possible group sizes

TA can not demonstrate it mathmatically

- only describe in counting principles

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

Problem Set 3 Solutions
Due September 29, 2010

1. *The hats of n persons are thrown into a box. The persons then pick up their hats at random (i.e., so that every assignment of the hats to the persons is equally likely). What is the probability that*

   (a) *every person gets his or her hat back?*
   Answer: $\frac{1}{n!}$.
   Solution: consider the sample space of all possible hat assignments. It has $n!$ elements (n hat selections for the first person, after that $n-1$ for the second, etc.), with every single-element event equally likely (hence having probability $1/n!$). The question is to calculate the probability of a single-element event, so the answer is $1/n!$

   (b) *the first m persons who picked hats get their own hats back?*
   Answer: $\frac{(n-m)!}{n!}$.
   Solution: consider the same sample space and probability as in the solution of (a). The probability of an event with $(n-m)!$ elements (this is how many ways there are to disribute the remaining $n-m$ hats after the first m are assigned to their owners) is $(n-m)!/n!$

   (c) *everyone among the first m persons to pick up the hats gets back a hat belonging to one of the last m persons to pick up the hats?*
   Answer: $\frac{m!(n-m)!}{n!} = \frac{1}{\binom{n}{m}} = \frac{1}{\binom{n}{n-m}}$. .
   Solution: there are $m!$ ways to distribute m hats among the first m persons, and $(n-m)!$ ways to distribute the remaining $n-m$ hats. The probability of an event with $m!(n-m)!$ elements is $m!(n-m)!/n!$.

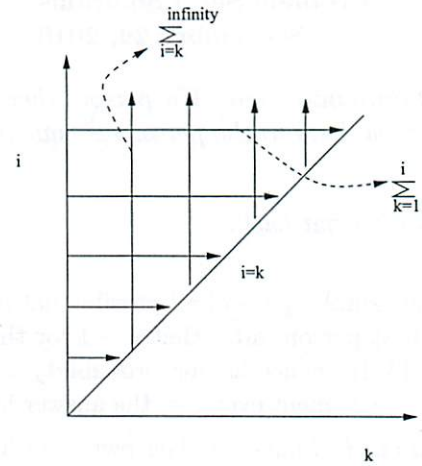
   Now assume, in addition, that every hat thrown into the box has probability p of getting dirty (independently of what happens to the other hats or who has dropped or picked it up). What is the probability that

   (d) *the first m persons will pick up clean hats?*
   Answer: $(1-p)^m$.
   Solution: the probability of a given person picking up a clean hat is $1-p$. By the independence assumption, the probability of m selected persons picking up clean hats is $(1-p)^m$.

   (e) *exactly m persons will pick up clean hats?*
   Answer: $(1-p)^m p^{n-m} \binom{n}{m}$.
   Solution: every group G of m persons defines the event "everyone from G picks up a clean hat, everyone not from G picks up a dirty hat". The events are disjoint. Each has probability $(1-p)^m p^{n-m}$. Since there are $\binom{n}{m}$ such events, the answer follows.

2. Since 4 cards are fixed, Bob can only choose 4 more cards out of 48 remaining cards, so total number of hands Bob can have such that they include Alice's cards is $\binom{4}{4}\binom{48}{4}$. The total number of ways Bob can choose any 8 cards is $\binom{52}{8}$. So the probability is $\frac{\binom{4}{4}\binom{48}{4}}{\binom{52}{8}}$

3. (a) The picture below illustrates the double sum needed to prove the statement of this problem:



We first note that

$$\mathbf{P}(X \geq k) = \sum_{i=k}^{\infty} p_X(i)$$

and proceed as follows:

$$\sum_{k=1}^{\infty} \mathbf{P}(X \geq k) = \sum_{k=1}^{\infty}\sum_{i=k}^{\infty} p_X(i) = \sum_{i=1}^{\infty}\sum_{k=1}^{i} p_X(i) = \sum_{i=1}^{\infty} i\, p_X(i) = \mathbf{E}[X].$$

(b) We first compute

$$\mathbf{P}(Y \geq k) = \begin{cases} 1 & k \leq a \\ \frac{b-k+1}{b-a+1} & a+1 \leq k \leq b \\ 0 & k \geq b+1 \end{cases}$$

So

$$\begin{aligned} \sum_{k=1}^{\infty} \mathbf{P}(Y \geq k) &= \sum_{k=1}^{a} 1 + \sum_{k=a+1}^{b} \frac{b-k+1}{b-a+1} \\ &= a + \frac{1}{b-a+1}\sum_{k=1}^{b-a} k \\ &= a + \frac{1}{b-a+1}\frac{(b-a+1)(b-a)}{2} \\ &= a + \frac{b-a}{2} \\ &= \frac{b+a}{2} \end{aligned}$$

Therefore $\mathbf{E}[Y] = \frac{b+a}{2}$.

4. (a) For each value of X, we count the number of outcomes which have a difference that equals that value:

$$p_X(x) = \begin{cases} 1/9 & x = -2, 2 \\ 2/9 & x = -1, 1 \\ 3/9 & x = 0 \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbf{E}[X] = \sum_{x=-2}^{2} x p_X(x) = -2\frac{1}{9} + -1\frac{2}{9} + 0\frac{3}{9} + 1\frac{2}{9} + 2\frac{1}{9} = \boxed{0}.$$

We can also see that $\mathbf{E}[X] = 0$ because the PMF is symmetric around 0.
To find the variance of X, we first compute

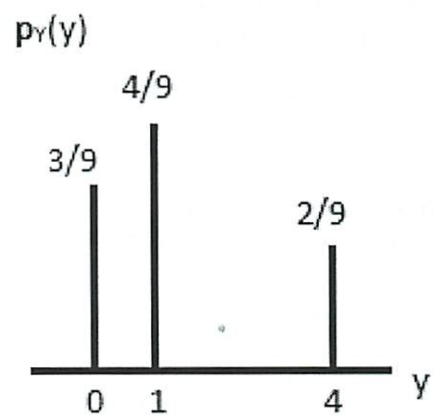
$$\mathbf{E}[X^2] = \sum_{x=-2}^{2} x^2 p_X(x) = 4\frac{1}{9} + 1\frac{2}{9} + 0\frac{3}{9} + 1\frac{2}{9} + 4\frac{1}{9} = \boxed{\tfrac{4}{3}}.$$

and

$$\operatorname{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \boxed{\tfrac{4}{3}}.$$

(b) Let $Z = X^2$. By matching the possible values of X and their probabilities to the possible values of Z, we obtain

$$p_Z(z) = \begin{cases} 2/9 & z = 4 \\ 4/9 & z = 1 \\ 3/9 & z = 0 \\ 0 & \text{otherwise.} \end{cases}$$



5. Consider k out of n persons forming a club, with one being designated as the leader and another as the treasurer. We can first choose the leader (n choices), then the treasurer ($n-1$ choices), and then a subset of the remaining $n-2$ persons. Thus, there are $n(n-1)2^{n-2}$ possible clubs.

   Alternatively, for any given k, there are $\binom{n}{k}$ choices for the members of the club. There are $k(k-1)$ choices for the leader and treasurer, so that there are $k(k-1)\binom{n}{k}$ k-member clubs. Summing over all k, we see that there is a total of $\sum_{k=2}^{n} k(k-1)\binom{n}{k}$ possible clubs.

G1†. *A candy factory has an endless supply of red, orange, yellow, green, blue, black, white, and violet jelly beans. The factory packages the jelly beans into jars in such a way that each jar has 200 beans, equal number of red and orange beans, equal number of yellow and green beans, one more black bean than the number blue beans, and three more violet beans than the number of white* *beans. One possible color distribution, for example, is a jar of 50 yellow, 50 green, one black, 48 white, and 51 violet jelly beans. As a marketing gimmick, the factory guarantees that no two jars have the same color distribution. What is the maximum number of jars the factory can produce?*

Answer: $\binom{101}{3} = 166650$.

Solution: Let $N_1, N_2, N_3, N_4, N_5, N_6, N_7, N_8$ denote, respectively, the numbers of red, orange, yellow, green, blue, black, white, and violet jelly beans in a jar. There is a one-to-one correspondence

$$x = (x_1, x_2, x_3, x_4) \mapsto N = (x_1, x_1, x_2, x_2, x_3, x_3 + 1, x_4, x_4 + 3)$$

between the non-negative integer solutions $x = (x_1, x_2, x_3, x_4)$ of the equation

$$x_1 + x_2 + x_3 + x_4 = 98,$$

and the sequences $N = (N_1, N_2, N_3, N_4, N_5, N_6, N_7, N_8)$ of non-negative integers N_i satisfying the conditions

$$N_2 = N_1,\ N_4 = N_3,\ N_6 = N_5 + 1,\ N_8 = N_7 + 3,\ \sum_{i=1}^{8} N_i = 200$$

(i.e. possible color arrangements). The number of possible solutions x is $\binom{101}{3}$ according to the solution of the more general problem given below:

Given a non-negative integer n and a positive integer k, consider the equation

$$x_1 + x_2 + \ldots + x_k = n,$$

to be solved with respect to non-negative integer variables $x_1, x_2, \ldots, x_k$. Find the total number of solutions (solutions $x_1 = 1,\ x_2 = 0$ and $x_1 = 0,\ x_2 = 1$ to the equation $x_1 + x_2 = 1$ are considered as different).

Answer: $\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$.

Solution: there is a one-to-one correspondence between non-negative integer solutions of equation $x_1 + \ldots + x_k = n$ and sequences of $n+k-1$ symbols (n "o" and $k-1$ "|"), where a solution $x = (x_1, \ldots, x_k)$ maps to the sequence in which the i-th "|" (where $i \in \{1, 2, \ldots, k-1\}$) is in the $x_1 + \ldots + x_i + i$th place: in this bijection, the numbers of "o" between the consecutive "|" correspond to the values of x_i. Hence the total number of solutions equals the number of ways of selecting $k-1$ places for the "|" symbols in a sequence of length $n + k - 1$.

†Required for 6.431; optional for 6.041

9/29

LECTURE 7

- **Readings:** Finish Chapter 2

Quiz in 10 days

Up to today's lecture

Lecture outline

- Multiple random variables
  - Joint PMF
  - Conditioning
  - Independence
- More on expectations
- Binomial distribution revisited
- A hat problem

Review

diff notation

single $p_X(x) = P(X = x)$ — prob of diff values X can take

joint $p_{X,Y}(x, y) = P(X = x, Y = y)$ - that pair

conditional $p_{X|Y}(x \mid y) = P(X = x \mid Y = y)$ -one depends on the other $\boxed{\sum_x P_{X|Y}(x/y) = 1}$

fix value of Y → slice → normalize

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

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

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

$X=x$ $Y=y$

3 random variables: similar

$$P_X'(x) = \sum_Y \sum_Z P_{X,Y,Z}(x, y, z)$$

Independent random variables

$$p_{X,Y,Z}(x,y,z) = p_X(x)p_{Y|X}(y \mid x)p_{Z|X,Y}(z \mid x,y)$$

- Random variables X, Y, Z are independent if:

  $$p_{X,Y,Z}(x,y,z) = p_X(x) \cdot p_Y(y) \cdot p_Z(z)$$

  for all x, y, z

having knowledge of x does not change beliefs about y, z

- conditionals = marginals

joint

y \ x	1	2	3	4
4	1/20	2/20	2/20	
3	2/20	4/20	1/20	2/20
2		1/20	3/20	1/20
1		1/20		

like A, B are ind. say

$P(A \cap B) = P(A)P(B)$ $\rightarrow$ $X = \begin{cases} 1 & \text{if } A \\ 0 & \text{if } A^c \end{cases}$

$Y = \begin{cases} 1 & \text{if } B \\ 0 & \text{if } B^c \end{cases}$

A, B ind $P_{X,Y}(1,1) = P_X(1)P_Y(1)$

- also need to check things factor out

$P_{X,Y}(1,0) = P_X(1)P_Y(0)$

- Independent? No — info about Y relevant to X

- What if we condition on $X \leq 2$ and $Y \geq 3$? look at new universe

	1	2	
4	1/9	2/9	= 1/3
3	2/9	4/9	= 2/3
	= 1/3	= 2/3	

is any product = to x of marginal prob

- all are Yes

no matter what y is, your beliefs ~~of~~ about x don't change

Expectations

average value you will see if you do experiment a lot

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

$$\mathrm{E}[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y)$$

- In general: $\mathrm{E}[g(X,Y)] \neq g\big(\mathrm{E}[X], \mathrm{E}[Y]\big)$ *

except

- $\mathrm{E}[\alpha X + \beta] = \alpha \mathrm{E}[X] + \beta$
- $\mathrm{E}[X + Y + Z] = \mathrm{E}[X] + \mathrm{E}[Y] + \mathrm{E}[Z]$ picking 1 student - their quiz score

- If X, Y are independent:
  - $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$ $= \sum_x \sum_y xy \, P_{X,Y}(x,y) = \sum_y \sum_x P_X(x) P_Y(y)$ $= E[X] \, E[Y]$
  - $\mathrm{E}[g(X)h(Y)] = \mathrm{E}[g(X)] \cdot \mathrm{E}[h(Y)]$

similar argument

$g(X)$, $h(Y)$ are independent

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

Variances

- $\text{Var}(aX) = a^2\text{Var}(X)$
- $\text{Var}(X + a) = \text{Var}(X)$
- Let $Z = X + Y$.
  If X, Y are independent:

  $$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

assume $E[X] = E[Y] = 0$ — not independent

$$\text{var}(X+Y) = E[(X+Y)^2] = E[X^2] + E[Y^2] + 2E[XY]$$
$$= \text{Var}(x) + \text{var}(y) + 0$$

- Examples:
  - If $X = Y$, $\text{Var}(X + Y) =$ $\text{Var}(2X) = 4\text{Var}(X)$
  - If $X = -Y$, $\text{Var}(X + Y) =$ 0

  not independent

  - If X, Y indep., and $Z = X - 3Y$,
    $\text{Var}(Z) =$ $\text{Var}(X) + \text{var}(3y) = \text{Var}(X) + 9\,\text{var}(Y)$

    $X, -3Y$ are ind.

Binomial mean and variance

- $X = \#$ of successes in n independent trials
  - probability of success p

    $$E[X] = \sum_{k=0}^{n} k \underbrace{\binom{n}{k} p^k (1-p)^{n-k}}_{P_X(k)}$$

- $X_i = \begin{cases} 1, & \text{if success in trial } i, \\ 0, & \text{otherwise} \end{cases}$

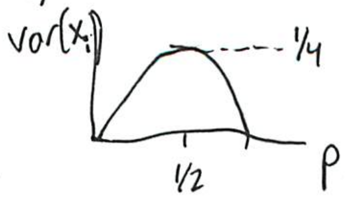
  indicator variables

- $E[X_i] =$ $p \cdot 1 + (1-p) \cdot 0 = p$

\# of successes $= \sum_{i=0}^{n} X_i$

- $E[X] =$ $E[X_1 + X_2 + \dots X_n] = np$
- $\text{Var}(X_i) =$ $p(1-p)^2 + (1-p)(0-p)^2 = p(1-p)$
- $\text{Var}(X) =$ $\text{Var}(X_1) + \dots + \text{Var}(X_n)$

  $= np(1-p)$



if $p \approx 0, 1$ then not much randomness so var is small

$p = 1/2$ = fair coin

The hat problem

- n people throw their hats in a box and then pick one at random.
  - X: number of people who get their own hat
  - Find $\mathbf{E}[X]$

$$X_i = \begin{cases} 1, & \text{if } i \text{ selects own hat} \\ 0, & \text{otherwise.} \end{cases}$$ indicator variable

- $X = X_1 + X_2 + \cdots + X_n$
- $\mathbf{P}(X_i = 1) = \frac{1}{n}$
- $\mathbf{E}[X_i] = 1 \cdot \frac{1}{n} \cdot 0 \cdot (1 - \frac{1}{n}) = \frac{1}{n}$
- Are the X_i independent? don't care that hats out → picking but if know everyone except last one got their hat back then we would know they would get their hat back NO
- $\mathbf{E}[X] =$ linearity of expectations true if inp or not

$= E[X_1] + \cdots + E[X_n]$
$= 1$

Variance in the hat problem

- $\text{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \mathbf{E}[X^2] - 1$

$$X^2 = \sum_i X_i^2 + \sum_{i,j:i \neq j} X_i X_j$$ cross terms → $n(n-1)$ cross terms

- $\mathbf{E}[X_i^2] \Rightarrow \begin{cases} 0^2 \\ 1^2 \end{cases} = \begin{cases} 0 \\ 1 \end{cases} \Rightarrow E[X_1] = \frac{1}{n}$

$\mathbf{P}(X_1 X_2 = 1) = \mathbf{P}(X_1 = 1) \cdot \mathbf{P}(X_2 = 1 \mid X_1 = 1)$
$= \underbrace{\frac{1}{n}}_{\text{1st person}} \cdot \underbrace{\frac{1}{n-1}}_{\text{2nd person}}$

- $\mathbf{E}[X^2] = n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)}$
- $\text{Var}(X) = 2 - 1 = 1$

Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2010)

Recitation 7

September 30, 2010

1. **Problem 2.35, page 130 in the text.** Verify the expected value rule

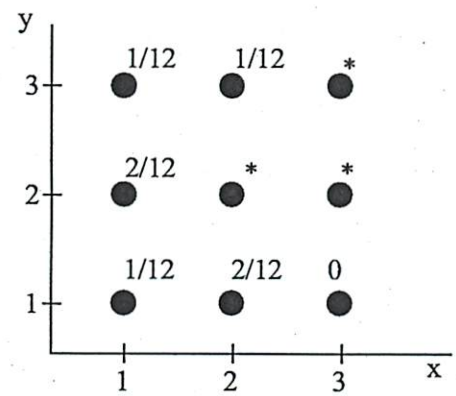
$$\mathbf{E}[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y),$$

using the expected value rule for a function of a single random variable. Then, use the rule for the special case of a linear function, to verify the formula

$$\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y],$$

where a and b are given scalars.

2. Random variables X and Y can take any value in the set $\{1, 2, 3\}$. We are given the following information about their joint PMF, where the entries indicated by a * are left unspecified:



(a) What is $p_X(1)$?

(b) Provide a clearly labeled sketch of the conditional PMF of Y given that $X = 1$.

(c) What is $\mathbf{E}[Y \mid X = 1]$?

(d) Is there a choice for the unspecified entries that would make X and Y independent?

Let B be the event that $X \le 2$ and $Y \le 2$. We are told that conditioned on B, the random variables X and Y are independent.

(e) What is $p_{X,Y}(2,2)$?
(If there is not enough information to determine the answer, say so.)

(f) What is $p_{X,Y|B}(2, 2 \mid B)$?
(If there is not enough information to determine the answer, say so.)

3. **Problem 2.33, page 128 in the text.** A coin that has probability of heads equal to p is tossed successively and independently until a head comes twice in a row or a tail comes twice in a row. Find the expected value of the number of tosses.

Recitation 7

9/30

Different instructor

1. Verify expected value rule → Prove
   - originally defined w/ PMF
     - sometimes too hard to bother w/ that
   - trial + error process
     - suppose to take time
   - assume true for functions of 1 random variable

$$E[g(X,Y)] =$$ - can do w/ conditioning on $Y = y$

$$= \sum_y p_Y(y)\, E[g(X,Y) \mid Y = y]$$

- means lock Y to y

$$= \sum_y p_Y(y)\, E[g(X,y) \mid Y = y]$$

$$= \sum_p p_Y(y) \cdot \sum_x g(x,y)\, p_{X|Y}(x,y)$$

- be careful, use right weighting

manipulate

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

$$= \sum_x \sum_y g(x,y) \underbrace{p_Y(y)\, p_{X|Y}(x|y)}_{\text{is joint PMF}}$$

← can do this here

"don't need to worry about that nonsense"

Done

I like this TA

② $$E[\underbrace{aX+bY}_{\substack{\text{is Function} \\ g(x,y)}}] = \sum_{X}\sum_{Y} (ax+by)\, P_{X,Y}(x,y)$$

$$= a\sum_{x} x \underbrace{\underbrace{\sum_{y} P_{X,Y}(x,y)}_{P_X(x)}}_{E[X]} + b \underbrace{\sum_{y} y \underbrace{\sum_{x} P_{X,Y}(x,y)}_{P_Y(y)}}_{E[Y]}$$

marginal = word of emphsis
only talking about 1 random variable
<u>not</u> like econ

2.

Y \ X	1	2	3
3	1/12	1/12	*
2	2/12	*	*
1	1/12	2/12	0

* = unspecified

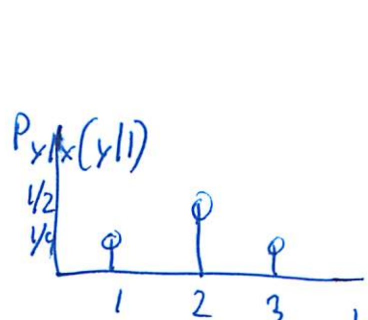
a) $P_X(1) = \frac{4}{12} = \sum_{y=1}^{3} P_{X,Y}(1,y) = \frac{1}{3} = \frac{4}{12}$

b)

Y	$X = 1$
3	1/4
2	1/2
1	1/4

$P_{Y|\{X=1\}}(y)$ or $P_{Y|X}(y|1)$
- same thing
- diff notation

$$\frac{P_{X,Y}(1,y)}{P_X(1)}$$



③

c) $E[Y \mid X=1]$

↳ mean of distribution

center of mass

$$\sum_{y=1}^{3} y P_{Y|X}(y \mid 1) = 2$$

Variables independent

- know info about one does not help you w/ other
- $P_{X,Y}(x,y) = P_X(x)\, P_Y(y)$ for all (x,y)

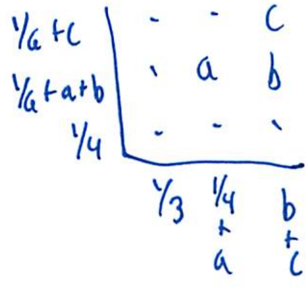
$$P(\{X=x\} \cap \{Y=y\}) = P(\{X=x\})\, P(\{Y=y\})$$

When $P(\{X=x\}) > 0$ ie $P_X(x) > 0$

$$P_{Y|X}(y \mid x) = P_Y(y)$$

d) <u>One option</u>

Name the unlabled pts



and sum up the sides

- get 4 non-linear eq.
- seems so painful

- Demitri's research
- How to check <u>if</u> a solution

(One pen already running out - ~~3~~ 4 weeks in)

(B) (4)

2nd option

know $P_{X,Y}(3,1) = 0$

and since $y=1$ is $1/4$

Know ~~$X=3 = 0$~~

$\hookrightarrow P_{X,Y}(3,y) = 0$

$X=2 \rightarrow 2/3$

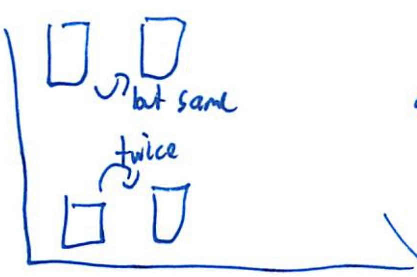
$Y=2 \rightarrow 7/12$

So a must $= 5/12$

3rd

Ok must be more insightful way of doing

Look at



so know X and Y are not ind!

Contridiction to

observed $\rightarrow \dfrac{P_{X,Y}(2,1)}{P_{X,Y}(1,1)} \neq \dfrac{P_{X,Y}(2,3)}{P_{X,Y}(1,3)}$

if X and Y were independant

would have to $= \dfrac{P_X(2)\cancel{P_Y(1)}}{P_X(1)\cancel{P_Y(1)}} \stackrel{?}{=} \dfrac{P_X(2)\cancel{P_Y(3)}}{P_X(1)\cancel{P_Y(3)}}$

but contridicted

5

Does that give us enough to assign a joint PMF to (2,2)

2/12	4/12
1/12	2/12

each column + row must be a scalar multiple of 4/12

f) $P_{X,Y|B}(2,2)$

4/9

- normalize probabilities inside the box B

---

3. Coin tossed $\begin{cases} \text{heads } p \\ \text{tail } p-1 \end{cases}$

tossed till head twice in a row or tails twice in a row

- feels their should be a geometric random variable
  - but <u>no</u>
  - prob. of sucess not same for every trial
  - but not fixed prob. for each trial p vs 1-p
  - could look at disjoint pairs

$\underbrace{HT}\,\underbrace{HT}$

but what about?

H $\underbrace{\boxed{TH}}$/T

- 2 good techniques in this class on this section
  - indicator random values
  - recursive somehow

6 If know what happens on 1st toss

~~[illegible]~~ ✓

- Total ~~Probability~~ Expectation Theorm something w/

$$E[X] = \underbrace{P(H_1)}_{p} E[X \mid H_1] + \underbrace{P(T_1)}_{1-p} E[X \mid T_1]$$

- glad did not have to find PMF of X

expand more

$$E[X \mid H_1] = \underbrace{P(H_2 \mid H_1)}_{p} \underbrace{E[X \mid H_1 \cap H_2]}_{2} + \underbrace{P(T_2 \mid H_1)}_{1-p} \underbrace{E[X \mid H_1 \cap T_2]}_{1 + E[X \mid T_1]}$$

continue as though had tail in toss 1 - (recursive)

$$E[X \mid T_1] = \underbrace{P(H_2 \mid H_1)}_{p} \underbrace{E[X \mid T_1 \cap H_2]}_{1 + E[X \mid H_1]} + \underbrace{P(T_2 \mid T_1)}_{1-p} \underbrace{E[X \mid T_1 \cap T_2]}_{2}$$

2 linear eq - 2 unknowns

Solution

$$E[X \mid H_1] = \frac{2 + (1-p)^2}{1 - p(1-p)}$$

$$E[X \mid T_1] = \frac{2 + p^2}{1 - p(1-p)}$$

$$E[X] = \frac{2 + p(1-p)}{1 - p(1-p)}$$

7

$P = 0, 1$ $\quad E[X] = 2$

Otherwise $\quad E[X] \in (2, 3)$

$P = \frac{1}{2}$ $\quad E[X] = 3$

when $p = 1/2$, After 1st toss, identical distributed Bernulli trials

p of each trial is $\frac{1}{2}$

- independent

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

Tutorial 3
September 30/October 1, 2010

1. Let X and Y be independent random variables. Random variable X has mean μ_X and variance σ_X^2, and random variable Y has mean μ_Y and variance σ_Y^2. Let $Z = 2X - 3Y$. Find the mean and variance of Z in terms of the means and variances of X and Y.

2. Problem 2.40, page 133 in the text.
A particular professor is known for his arbitrary grading policies. Each paper receives a grade from the set $\{A, A-, B+, B, B-, C+\}$, with equal probability, independently of other papers. How many papers do you expect to hand in before you receive each possible grade at least once?

3. The joint PMF of the random variables X and Y is given by the following table:

$y = 3$	c	c	$2c$
$y = 2$	$2c$	0	$4c$
$y = 1$	$3c$	c	$6c$
	$x = 1$	$x = 2$	$x = 3$

   (a) Find the value of the constant c.

   (b) Find $p_Y(2)$.

   (c) Consider the random variable $Z = YX^2$. Find $\mathbf{E}[Z \mid Y = 2]$.

   (d) Conditioned on the event that $X \neq 2$, are X and Y independent? Give a one-line justification.

   (e) Find the conditional variance of Y given that $X = 2$.

Tutorial 3

1. Find $E[X]$

-has to be linear → linearity of expectation

$$\boxed{E[aX+bY+c] = aE[X]+bE[Y]+c}$$

always holds

$$E[Z] = E[2X-3Y] = 2E[X]-3E[Y]$$
$$= 2\mu_X - 3\mu_Y$$

$$Var(Z) = Var(2X-3Y)$$

$$\boxed{Var(aX+bY+c) = a^2 var(X) + b^2 var(Y)} \rightarrow X+Y \text{ indep}$$

$$= 4Var(X) + 9var(Y)$$

because $var(-3x)$

$$= 4\sigma_X^2 + 9\sigma_Y^2$$

---

3. Picture w/ joint PMFs

Y=3	c	c	2c
Y=2	2c	0	4c
Y=1	3c	c	6c
	X=1	X=2	X=3

②

a) Normalization $\sum_x \sum_y P_{x,y}(x,y) = 1$

$c + 2c + 3c + c + c + 2c + 4c + 6c = 1$

$\boxed{c = 1/20}$

easy qu

b) $P_y(2) = P[\{y=2\}]$ for discrete random variables

$= \sum_x P_{x,y}(x,2)$

$= 2c + 4c$

$= \frac{6}{20} = \frac{3}{10}$

c) $z = yx^2$

$E[z | y=2] = E[yx^2 | y=2]$

substitute in

$= E[2x^2 | y=2]$

must keep conditional, unless is independant

$= 2\, E[x^2 | y=2]$

$= 2 \sum_x x^2 p_{x|y}(x|2)$

table is giving joint PMF, but want conditional PMF

- Scale / normalize

$p_{x|y}(x|2) = \frac{p_{x,y}(x,2)}{p_y(2)} = \frac{p_{x,y}(x,2)}{3/10}$

3

$$P_{X|Y}(x|2) = \begin{cases} 1/3 & \text{if } x=1 \\ 2/3 & \text{if } x=3 \\ 0 & \text{otherwise} \end{cases}$$

$$E[Z|Y=2] = 2\left[1^2 \cdot \frac{1}{3} + 3^2 \cdot \frac{2}{3}\right] = \frac{38}{3}$$

d) Are the lines independent?

- are scalar multiples of each other
- 2 variables are indp.
- always know ~~…~~ x=3 is 2x as likely as x=1 no matter what y is (Y=y)
- if it holds for one, it holds for the other

e) $Var(Y|X) = 2$

↳ ~~…~~

$$\sum_{y} (y - E[X|X=2])^2 \cdot P_{Y|X}(y|2)$$

or

$$E[Y^2|X=2] - E[Y|X=2]^2$$

$E[Y|X=2] = 2$ because the pmf given X=2 is symetrical around 2

$$\sum_{y} (y-2)^2 \cdot P_{Y|X}(y|2)$$

- what they are scaled do

$$P_{Y|X}(y|2) = \begin{cases} 1/2 & \text{for } y=1 \\ 1/2 & \text{for } y=3 \end{cases}$$

4

$$= (1-2)^2 \cdot \frac{1}{2} + (3-2)^2 \cdot \frac{1}{2} = 1$$

- very standard problem
- make sure comfortable doing this

---

pg 116 - memorize

~~Bon~~ - wrote on seperate sheet

---

2. Brain Teaser Style qu

Prof picks grades at random

How many papers do you need to submit till you've seen every grade for 1st time $= X$

$E[X] = ?$

- go through PMFs
  - bad idea in complex problems
- try to use linearity of expectation problem
- recursion ← least common

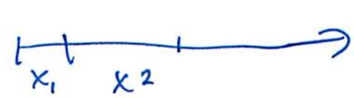
roughly $X = X_1 + X_2 + \ldots$

Don't name the grades

- do # of papers until new grade

X_i = # of papers b/w ith and $(i+1)$st sucess

Sucess = new grade



$X = 1 + X_1 + X_2 + X_3 + X_4 + X_5$

(5)

$$E[X] = 1 + E[X_1] + E[X_2] + E[X_3] + E[X_4] + E[X_5]$$

problem has been divided into sub problems

- geometric random variable

$p = \text{prob(sucess)}$

X_1 = geometric r.v. w/ $p = 5/6$

$$E[X_1] = 6/5$$

X_2 = geometric r.v. $p = \frac{4}{6}$

$$E[X_2] = 6/4$$

... ...

X_i = " " $p = \frac{6-i}{6}$

$$E[X_i] = 6/(6-i)$$

$$= 1 + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} = 14.7$$

harder to get a new grade

Key Probs

Bernoulli p:

$$P_x(k) = \begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$$

$E[X] = p$

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

p. sucess in 1 trial

Binomial p, n: $\sum_{i=1}^{n} X_i$

$$P_x(k) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$$

$E[X] = np$

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

$k = 0, 1, 2, 3 \ldots n$ ← pnt!

Sucesses can be in any order

P(sucess in n trials) k = # of sucesses $= \sum_{i=1}^{n} X_i$

Geometric p

$$p_x(k) = \underbrace{(1-p)^{k-1}}_{\text{each failure}} p \quad (k = 1, 2, 3 \ldots)$$

# of trials until 1st sucess

$E[X] = 1/p$

$Var(X) = \frac{1-p}{p^2}$

---

p 116 - make sure know devation

(think I ~~explain~~ know how to explain)

- not many dirvations on quiz
  - like p-sets

Michael Plasmeier

2/10

Do not include 50 pages of office hour notes with your pset.

Staple does not work

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

Problem Set 4
Due October 6, 2010

1. Random variables X and Y have the joint PMF

$$p_{X,Y}(x,y) = \begin{cases} c(x^2+y^2), & \text{if } x \in \{1,2,4\} \text{ and } y \in \{1,3\}, \\ 0, & \text{otherwise.} \end{cases}$$

   (a) What is the value of the constant c?
   (b) What is $\mathbf{P}(Y < X)$?
   (c) What is $\mathbf{P}(Y > X)$?
   (d) What is $\mathbf{P}(Y = X)$?
   (e) What is $\mathbf{P}(Y = 3)$?
   (f) Find the marginal PMFs $p_X(x)$ and $p_Y(y)$.
   (g) Find the expectations $\mathbf{E}[X]$, $\mathbf{E}[Y]$ and $\mathbf{E}[XY]$.
   (h) Find the variances $\text{var}(X)$, $\text{var}(Y)$ and $\text{var}(X+Y)$.
   (i) Let A denote the event $X \geq Y$. Find $\mathbf{E}[X \mid A]$ and $\text{var}(X \mid A)$.

2. The newest invention of the 6.041/6.431 staff is a three-sided die with faces numbered 1, 2, and 3. The PMF for the result of any one roll of this die is

$$p_X(x) = \begin{cases} 1/2, & \text{if } x = 1, \\ 1/4, & \text{if } x = 2, \\ 1/4, & \text{if } x = 3, \\ 0, & \text{otherwise.} \end{cases}$$

   Consider a sequence of six independent rolls of this die, and let X_i be the random variable corresponding to the ith roll.

   (a) What is the probability that exactly three of the rolls have result equal to 3?
   (b) What is the probability that the first roll is 1, given that exactly two of the six rolls have result of 1?
   (c) We are told that exactly three of the rolls resulted in 1 and exactly three resulted in 2. Given this information, what is the probability that the sequence of rolls is 121212?
   (d) Conditioned on the event that at least one roll resulted in 3, find the conditional PMF of the number of 3's.

3. Suppose that X and Y are independent, identically distributed, geometric random variables with parameter p. Show that

$$\mathbf{P}(X = i \mid X + Y = n) = \frac{1}{n-1}, \qquad \text{for } i = 1, 2, \ldots, n-1.$$

4. Consider 10 independent tosses of a biased coin with a probability of heads of p.

   (a) Let A be the event that there are 6 heads in the first 8 tosses. Let B be the event that the 9th toss results in heads. Show that events A and B are independent.

   (b) Find the probability that there are 3 heads in the first 4 tosses and 2 heads in the last 3 tosses.

   (c) Given that there were 4 heads in the first 7 tosses, find the probability that the 2nd head occurred during the 4th trial.

   (d) Find the probability that there are 5 heads in the first 8 tosses and 3 heads in the last 5 tosses.

5. Consider a sequence of independent tosses of a biased coin at times $t = 0, 1, 2, \ldots$. On each toss, the probability of a 'head' is p, and the probability of a 'tail' is $1 - p$. A reward of one unit is given each time that a 'tail' follows immediately after a 'head.' Let R be the total reward paid in times $1, 2, \ldots, n$. Find $\mathbf{E}[R]$ and $\mathrm{var}(R)$.

G1†. A simple example of a random variable is the *indicator* of an event A, which is denoted by I_A:

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Prove that two events A and B are independent if and only if the associated indicator random variables, I_A and I_B are independent.

   (b) Show that if $X = I_A$, then $\mathbf{E}[X] = \mathbf{P}(A)$.

†Required for 6.431; optional for 6.041

Michael Plasmeier   <u>6.041 P-Set 4</u>   10/3

1. $p_{X,Y}(x,y) = \begin{cases} c(x^2+y^2) & \text{if } x \in \{1,2,4\} \text{ and } y \in \{1,3\} \\ 0 & \text{otherwise} \end{cases}$

a.) What is value of c?

- must add to 1

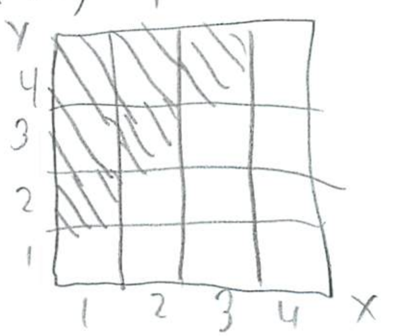
Y \ X	1	2	3	4
4	0	0	0	0
3	10	13	0	25
2	0	0	0	0
1	1	5	0	17

$\Sigma = 71$

$c = 1/71$

actually surprisingly simple

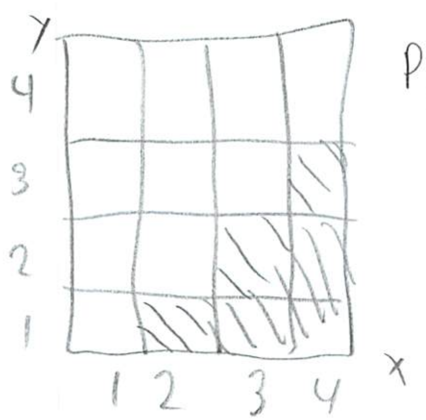
c) $P(Y > X)$?



$\dfrac{0+0+0+10+13+0}{71} = \dfrac{23}{71} = .323$

note flipped

b) $P(Y < X)$

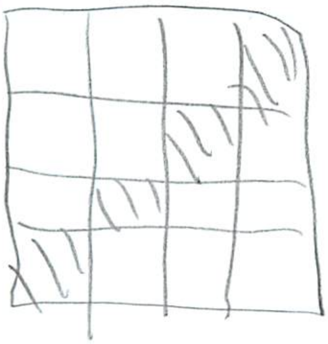


$\dfrac{25+0+0+5+0+17}{71} = \dfrac{47}{71} = .66$

(2)

d) $P(Y=X)$

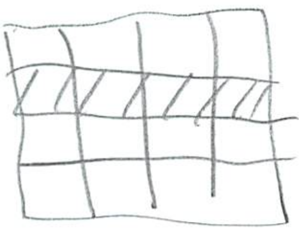
Y



X

$0 + 0 + 0 + 1 / 71 = \frac{1}{71} = .014$

e) $P(Y=3)$



$16 + 13 + 0 + 25 / 71 = \frac{48}{71} = .676$

f) Find marginal PMFs $p_X(x)$ and $p_Y(y)$

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

$$p_X(x) = \begin{cases} 11/71 & \text{if } x=1 \\ 18/71 & \text{if } x=2 \\ 0 & \text{if } x=3\text{, otherwise} \\ 42/71 & \text{if } x=4 \end{cases}$$

$$p_Y(y) = \begin{cases} 23/71 & \text{if } y=1 \\ 0 & \text{if } y=2 \\ 48/71 & \text{if } y=3 \\ 0 & \text{if } y=4 \end{cases} = \begin{cases} 23/17 & \text{if } y=1 \\ 48/71 & \text{if } y=3 \\ 0 & \text{otherwise} \end{cases}$$

(3)

this is easy + fun
Just decode notation

g) Find expectations

$$E[X] = \frac{11}{71} \cdot 1 + \frac{18}{71} \cdot 2 + 0 \cdot 3 + 42/71 \cdot 4 = \frac{215}{71} = 3.02$$

$$E[Y] = \frac{23}{71} \cdot 1 + 0 \cdot 2 + \frac{48}{71} \cdot 3 + 0 \cdot 4 = \frac{167}{71} = 2.35$$

$E[XY]$ = value of multiplying the two

$$= \frac{1}{71} \cdot 1 + \frac{5}{71} \cdot 2 + \frac{17}{71} \cdot 4 + \frac{10}{71} \cdot 3 + \frac{13}{71} \cdot 6 + \frac{25}{71} \cdot 12$$

$$= \frac{487}{71} = 6.86$$

h) Find Vars

$$\text{var}(X) = E[(X - E[X])^2]$$

$$= \sum_x (x - E[x])^2 p_X(x)$$

$$(1 - 3.02)^2 \cdot \frac{11}{71} + (2 - 3.02)^2 \cdot \frac{18}{71} + (4 - 3.02)^2 \cdot \frac{42}{71}$$

$$= 1.46$$

$$\text{var}(Y) = (1 - 2.35)^2 \cdot \frac{23}{71} + (3 - 2.35)^2 \cdot \frac{48}{71}$$

$$= .876$$

$\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$ if independent

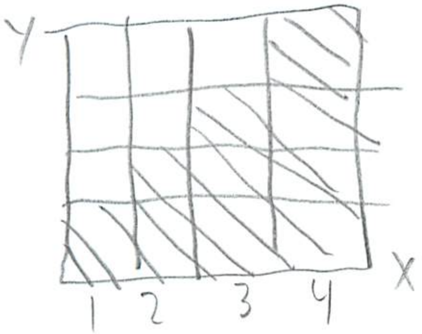
$$= 2.34$$

- could verify independance by calculating var the long way

(4)

i) Let A denote $X \geq Y$.

Fun when I actually understand it → good teaching



0
0 25
0 0 0
1 5 0 17 / 48

rescale

$$E[X|A] = \frac{1}{48}\cdot 1 + \frac{5}{48}\cdot 2 + 0\cdot 3 + \frac{42}{48}\cdot 4 = \frac{179}{48} = 3.729$$

$$\mathrm{var}(X|A) = (1-3.729)^2\cdot\frac{1}{48} + (2-3.729)^2\cdot\frac{5}{48} + (4-3.729)^2\cdot\frac{42}{48}$$

$$= .53$$

(5)

2. 3 sided die

$$P_X(x) = \begin{cases} 1/2 & \text{if } x=1 \\ 1/4 & \text{if } x=2 \\ 1/4 & \text{if } x=3 \\ 0 & \text{otherwise} \end{cases}$$

Sequence of 6 rolls of a die

X_i = Value of ith roll

a) What is prob 3 of the rolls all = 3

$\frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{64} = .015625$ (−1)

b) What is prob that 1st roll is 1 given that 2 of 6 rolls have been 1

combo - order does not matter

- oh no → joint

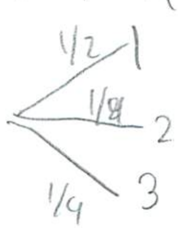
$\underbrace{\boxed{1}}_{\text{must be a 1}} \quad \underbrace{\_\ \_\ \overset{1}{\_}\ \_\ \_}_{\text{one must be a 1, any position}}$

$|A| = \frac{1}{2} \cdot \binom{5}{1}$ - but each chance not equal!

$|A| = \frac{1}{2} \cdot \frac{5!}{(4!)\,1!}$ ← so can't do that, right?

$P(\text{1st is a 1}) + P(\text{of the next 5, one roll} = 1)$

$\frac{1}{2} \cdot$



⑥

$\cancel{P(X_2 = 1 \mid X_1 = 1)} = \frac{1}{2} \cdot \frac{1}{2}$

- ? perhaps one of those 3 patterns we should learn
- Binomial
  - # of sucesses in n trials
  - $n = 5$
  - $k = 1$
  - $P_X(x) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$

$\frac{1}{2} \cdot \binom{5}{1} .5^1 (1 - .5)^{5-1}$

$\frac{1}{2} \cdot \frac{5!}{4!\,1!} \cdot .5 \cdot (.5)^4$ (-1)

c) We are told 3 rolls = 1, 3 rolls = 2

$P(\text{seq } 121212)$?

- so what combo is this ?

$P(3\text{rolls} = 1) = \binom{6}{3} \cdot \frac{1}{2}^3 \left(1 - \frac{1}{2}\right)^{6-3}$? double counting

$P(3 \text{ rolls} = 2) = \binom{6}{3} \cdot \frac{1}{4}^3 \left(1 - \frac{1}{4}\right)^{6-3}$

$P(121212) = \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{1}{4}$

$$P(121212 \mid 3\text{rolls} = 1 \text{ and } 3 \text{ rolls} = 2) = \frac{P(121212 \cap 3\text{rolls} = 1 \cap 3\text{rolls} = 2)}{P(3\text{rolls} = 1 \cap 3 \text{ rolls} = 2)}$$

⑦

$$\frac{\binom{6}{3}\cdot\frac{1}{2}^3\left(\frac{1}{2}\right)^3\cdot\binom{6}{3}\frac{1}{4}^3\left(\frac{3}{4}\right)^3\cdot\frac{1}{2}\cdot\frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{4}}{\binom{6}{3}\cdot\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^3\cdot\binom{6}{3}\cdot\frac{1}{4}^3\left(\frac{3}{4}\right)^3}$$

$$=\frac{1}{2}\cdot\frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{4}=.00195 \quad (-1)$$

d) Find conditional PMF of $x=3$ | At least one roll $=1$

Each roll is independant so $P_X(x=3 \mid \text{at least one in 6 rolls} = 3) = 1/4$

$$P_X(x \mid \text{at least one in 6 rolls} = 3) = \begin{cases} 1/2 & \text{if } x=1 \\ 1/4 & \text{if } x=2 \\ 1/4 & \text{if } x=3 \\ 0 & \text{otherwise} \end{cases}$$

(−1)

0/4

7b

2c) So if had $1\ 2\ 1\ 2\ 1\ 1\ 2$

$$\left(\frac{1}{2}\right)^4 \left(\frac{1}{4}\right)^3$$

So $1\ 1\ 1\ 2\ 2\ 2$

$$\underbrace{\binom{6}{3}}_{\text{how many sequences}} \underbrace{\left(\frac{1}{2}\right)^3 \left(\frac{1}{4}\right)^3}_{\text{one sequence}}$$

~~more choices than 2~~ $\lt$ H, T

actually not binomial strictly

---

then not indedendant

- if B is true → A has to be true
- so just P(B)

$B = 121212$

$A = 3\text{"}1\text{"}\ 3\text{"}2\text{"}$

$$\boxed{\left(\frac{1}{2}\right)^3 \left(\frac{1}{4}\right)^3}$$

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B)}{P(A)} = \frac{\cancel{\left(\frac{1}{2}\right)^3 \left(\frac{1}{4}\right)^3}}{\binom{6}{3}\cancel{\left(\frac{1}{2}\right)^3 \left(\frac{1}{4}\right)^3}} = \frac{1}{\binom{6}{3}}$$

- pick <u>one</u> of the 3"1" 3"2" sequences

(8)

3. Suppose X and Y are independant, identically distributed, geometric random variables w/ paramter p.

Show that $P(X=i \mid X+Y=n) = \frac{1}{n-1}$ for $i = 1, 2, \ldots, n-1$

$\boxed{\text{geometric is number of tosses for head to come up 1st} \quad P_X(x) = (1-p)^{k-1} p \quad k = 1, 2, 3 \ldots}$

Show w/ numbers

$X=4$
$Y=5$ $\quad n=9$

$P(X \in (1,8)) = \frac{1}{9-1} = \frac{1}{8}$ makes sense

- but how to prove w/ letters?
- OH + study

$P(X=i \mid X=n) = \frac{1}{n}$ for $i = 1, 2, 3 \ldots, n$ ← identically distributed

$P(X=i \mid X=n) = \frac{1}{n-1}$ for $i = 1, 2, 3 \ldots, n-1$ ← one less chance

as each outcome has an equal chance of occuring

Y is also identically distributed

and they are independent, having 1 gives no info about other

So

$P(X=i \mid X+Y=n) = \frac{1}{n-1}$ for $i = 1, 2, 3 \ldots n-1$

$n-1$ because minimum n is 1 so

$n-1+1 = n$

So would <u>not</u> work $P(X=i \mid X+Y=n) = \frac{1}{n}$ for $i = 1, 2, 3 \ldots n$

See in OH

I just don't know how to know how legal ans

8a

OH: can write proof

* and only = 0 if independent

$$P(x=i \mid x+y=n) = \frac{P(x=i \cap x+y=n)}{P(x+y=n)}$$

if $x=i$, $i+y=n$, so $n-i=y$

independent so can multiply

$$\frac{P(x=i) \cdot P(y=n-i)}{P(x+y=n)}$$

A_1, A_2, A_3

$P(B) = \sum$

$P(B) = \sum_i P(A_i \cap B)$

$= \sum P(B|A_i) P(A_i)$ total prob theorm

each piece adds to whole

$P(i+n-i=n)$??

$x+y=n$

$(1, n-1)$
$(2, n-2)$ } sum of pairs
$(3, n-3)$

$\vdots$ $n-1$ patterns

$\sum_i^{n-1} (i, n-i)$

each poring has a prob $\frac{1}{n-1}$ each =ly likyy

$\sum_1^{n-1}{}_{(x,y)} (x=i, y=n-i)$

8b.

Independent

$$\sum_{i}^{n-1} P_X(X=i)\, P_Y(Y=n-1)$$

each is a geometric

-sub in geometric random variable

Geometric $(1-p)^{k-1}\,p$

prob head on $n-i$ toss

$X = (1-p)^{i-1}\,p \qquad Y = (1-p)^{n-1-1}\,p$

$$\frac{\left[(1-p)^{i-1}p\right]\left[(1-p)^{n-(i-1)}p\right]}{\sum_{i=1}^{n-1}\left[(1-p)^{i-1}p\right]\left[(1-p)^{n-(i-1)}p\right]}$$

den $\displaystyle\sum_{i=1}^{n-1}(1-p)^{n-(i-1)+i-1}\,p^2$ $\qquad \displaystyle\sum_{i=1}^{k} a^k =$

$$p^2\sum_{i=1}^{n-1}(1-p)^n$$

$$p^2(1-p)^n\sum_{i=1}^{n-1}1$$

$$p^2(1-p)^n\cdot(n-1)$$

8c)
Numerator

$$\left[(1-p)^{i-1}p\right]\left[(1-p)^{n-(i-1)}p\right]$$

$$(1-p)^{i-1+n-(i-1)}p^2$$

$$(1-p)^n p^2$$

So togetor

$$\frac{(1-p)^n p^2}{(1-p)^n p^2 \cdot (n-1)} = \frac{1}{n-1}$$

(9)

4. Consider 10 independent coin tosses w/ biased coin $p(\text{Heads}) = p$

a) A = 6 heads in 8 tosses

B = 9th toss = heads

Show A + B are independant

Each toss of a coin is independant

$$P(H) = P(H|H)$$

$$p = \frac{P(H \cap H)}{P(H)} = \frac{p \cdot p}{p} = p$$

? again whats a legal ans?

need to do w/ events A + B

A = 6 heads in 8 tosses

$$\binom{8}{6} p^6 (1-p)^2$$

B = 9th toss = heads

$$P(H_9) = p$$

$P(A \cap B) = P(A) P(B)$ — both only true independant

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

So demenstrate one

9b

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

$\binom{8}{6} p^6 (1-p)^2 \cap p = \binom{8}{6} p^6 (1-p)^2 \cdot p$

✓ asserted independence

^ could have also argued it was disjoint up here

- try other form

- or disjoint

- tosses are indp.

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

here can put $\cdot$ sign

- disjoint + independent

$P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{\cancel{\binom{8}{6} p^6 (1-p)^2} \cdot p}{\cancel{\binom{8}{6} p^6 (1-p)^2}} = p$

+1

(10)

b) $P($3 heads in 1st 4 tosses$) \cap P($2 heads in last 3 tosses$)$

binomial - k sucesses in n trials

$$\binom{4}{3} p^3 (1-p)^{4-3} \cdot \binom{3}{2} p^2 (1-p)^{3-2}$$

$$\binom{4}{3} p^3 (1-p) \cdot \binom{3}{2} p^2 (1-p)$$

+1

c) Given that there were 4 heads in 1st 7 tosses, prob that 2nd head was in 4th trials

Condition $\binom{7}{4} p^4 (1-p)^3$

- geometric # of trials to 1st sucess

$(1-p)^{k-1} p$ ← need to modify

- or

$\underbrace{\text{1 head}}_{\text{3 trials}} \cap \underbrace{H}_{} \cap \underbrace{\text{2 heads}}_{\text{3 trials}}$ ← or just this and then no need to condition

$$\binom{3}{1} p (1-p)^2 \cdot p \cdot \binom{3}{2} p^2 (1-p)$$

↑ this is both want condition

$P(E|F) = \dfrac{P(E \cap F)}{P(F)}$ ← what I did ← need $= \dfrac{\binom{3}{1} p(1-p)^2 \cdot p \binom{3}{2} p^2 (1-p)}{\binom{7}{4} p^4 (1-p)^3}$

(-1)

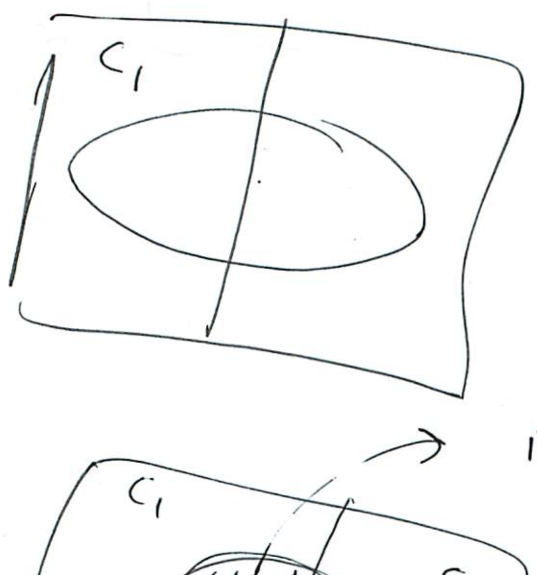
$F =$ 4H in 7

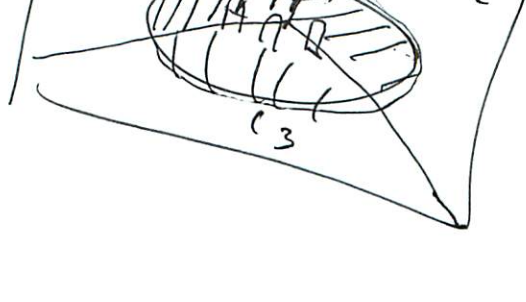
$E =$ Event 2nd on 4th

$$P(E|F) = \frac{P(E \cap F)}{P(?)}$$

Office Hrs

TA wrote

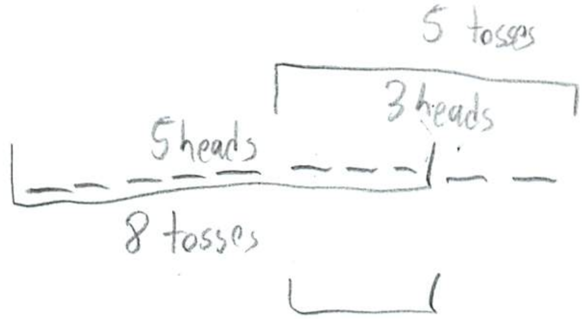




$$P(A \cap B \cap C_1) = P(A \cap B \mid C_1)\,P(C_1)$$

(11)

d) prob 5 heads in 1st 8 tosses and 3 heads in last 5 tosses



? how to deal w/ overlap ?!

- add the probabilities - but add is <u>or</u>
- ? confused - ask

first 5 bre heads of 8

A $\binom{8}{5} p^5 (1-p)^3$

$P(A \cap B)$

last 3/5 heads

B $\binom{5}{3} p^3 (1-p)^2$

Count way can overlap

$\underbrace{T\,T\,T\,H\,H}\,\underbrace{H\,H\,H,\,T\,T}$

? is this the best way

Want conditionally

$P(A \cap B \mid \text{center 3 are heads})$

$\sum_i P(A \cap B \mid C_i)\, P(C_i)$ ← same as before sum of intersections

$\boxed{\Phi}$

(11b)

3/

$P(A\cap E|C_3)$ $\binom{5}{2} p^2 (1-p)^3 \cdot 1 \cdot \cancel{\binom{2}{0}}^{1} (1-p)^2$ $\cdot$

$P(C_3)$ $\cancel{\binom{3}{3}}^{1} p^3$

2/ $\underbrace{H\ H\ T\ H\ T\ \overbrace{H\ H\ T}}, \overbrace{H\ T}$

$\binom{5}{3} p^3 (1-p)^2 \cdot 1 \cdot \binom{2}{1} p (1-p)$ $\cdot$

$P(C_2)$ $\binom{3}{2} p^2 (1-p)$

1/

$\underbrace{H\ H\ H\ H\ T\ \overbrace{H\ T\ T}}, \overbrace{H\ H}$

$\binom{5}{4} p^4 (1-p) \cdot 1 \cdot \binom{2}{2} p^2$ $\cdot$

$P(C_1)$ $\binom{3}{1} p (1-p)^2$

$\sum$ each term reduce

Write it out box it.

(-1)

2/4

(12)

5. Independent tosses of a biased coin

$t = 0, 1, 2, \ldots$ ← time

$P(\text{head}) = p$

$P(\text{tail}) = 1-p$

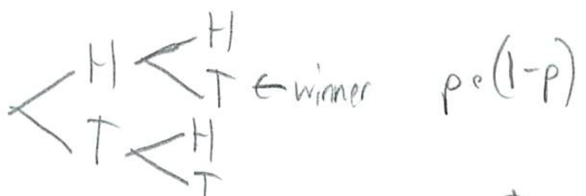
Reward when HT

R = total reward paid out at time $1, 2, 3 \ldots n$

$E[R]$ $\quad$ $\text{var}(R)$

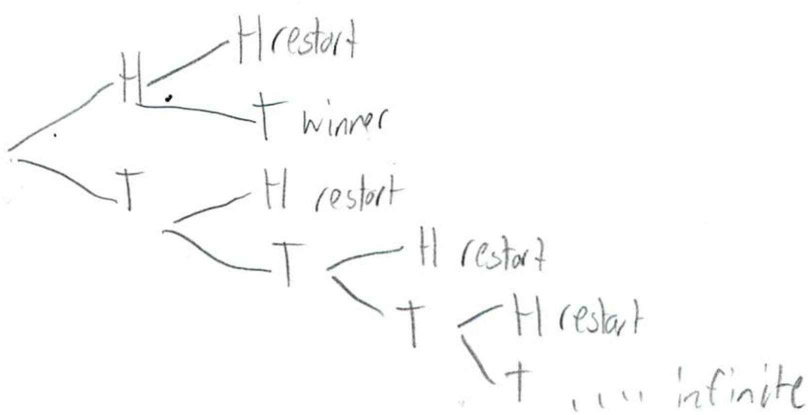
time = 1 $\quad$ $E = 0$

time = 2



did something like this in Thur recitation

- not geometric since prob is not even
- can't do HT, HT what about H,T H,T
- indicator random values
- recursive somehow



(13)

$$P(x \mid H_1) = p \cdot P(x \mid H_1) + (1-p) \cdot 1$$

restart

1 point

$$P(x \mid T_1) = p \cdot P(x \mid H_1) + (1-p) \cdot 0$$

where to go from here?

don't care

2 linear eq, 2 unknowns

handwaving in recitation

Solved in book w/o handwaving

- group terms to find ans

$$P(x) = p\big(p \cdot P(x \mid H_1) + (1-p) \cdot 1\big) + (1-p)\big(p \cdot p(x \mid H_1)\big)$$

then repeat after the 1st trial

- do I need to signify that seperatly

Use expectated values

$$E[R] = E[R \mid H_1]\, P(H_1) + E[R \mid T_2]\, P(T_1)$$

contine / Continue

$$E[R \mid H_1] = \underbrace{E[R \mid H_1 \cap H_2]}_{\text{restart}}\, p \cdot p + \underbrace{E[R \mid H_1 \cap T_2]}_{1} \cdot p(1-p)$$

$$= E[R \mid H_1]\, p^2 + p(1-p)$$

Solve for $E[R \mid H_1]$

$$x = x p^2 + p(1-p)$$

$$x - x p^2 = p(1-p)$$

$$x(1-p^2) = p(1-p)$$

need do for tails

$$x = \frac{p(1-p)}{(1-p^2)}$$

(14)

$$E[R \mid T_1] = E[R \mid T_1 \cap H_2] \cdot p(1-p) + E[R \mid T_1 \cap T_2](1-p)^2$$

$$= E[R \mid H_1] \; p(1-p) + E[R \mid T_1](1-p)^2$$

$$= \left[E[R \mid H_1] \, p^2 + p(1-p)\right] p(1-p) + E[R \mid T_1](1-p)^2$$

$$x = \left[\frac{p(1-p)}{(1-p^2)} p^2 + p(1-p)\right] p(1-p) + x(1-p)^2$$

$$x - x(1-p)^2 \quad \ldots$$

$$x(1-(1-p)^2)$$

$$E[R \mid T_1] = \frac{\left[\frac{p(1-p)}{(1-p^2)} p^2 + p(1-p)\right] p(1-p)}{(1-(1-p)^2)}$$

$$E[R] = \frac{p(1-p)}{(1-p^2)} + \frac{\left[\frac{p(1-p)}{(1-p^2)} p^2 + p(1-p)\right] p(1-p)}{(1-(1-p)^2)}$$

$$(1-p)^2 = (1-p)(1-p)$$
$$1-2p+p^2$$

(15)

$$= \frac{p-p^2}{1-p^2} + \frac{\left[\left(\frac{p-p^2}{1-p^2}\right)\cdot p^2 + p-p^2\right] p-p^2}{1-2p+p^2}$$

$$= \frac{p-p^2}{1-p^2} + \frac{\left[\left(\frac{p^3-p^4}{p^2-p^4}\right) + p-p^2\right] p-p^2}{1-2p+p^2}$$

Such a mess - not reducing!.

0/2

See solutions

Additional Help

- not for grading

$\cancel{P(A|B_1) + P(A|B_2)}$

not legal

2 different universes!

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

always

$= P(A)\, P(B)$ if independant

$E[R] = E[R|H_1]P(H_1) + E[R|T_1]P(T_1)$

$E[R|H_1] = E[R|H_1 \cap H_2] + E[R|H_1 \cap T_2] P.$

$1 + E[R|H_1$

$1 \quad E[X] = \sum_i E[X|A_i]P(A_i)$

$H_1 \quad H_2 \qquad T_1 \quad T_2$

$$E[X] = p\,E[X \mid H_1] + 1-p\,E[X \mid T_1]$$

$$E[X \mid H_1] = p\,E[X \mid H_1 \cap H_2] + (1-p)\,E[X \mid \cancel{H_1} \cap T_2]$$

$$E[X \mid H_1] = p + (1-p)(1+E[X \mid T_1])$$

$$E[X \mid T_1] = q + (1-q)(1+E[X \mid H_1])$$

$$\cancel{p + (1-q)(1+}$$

$$E[X \mid T_1] = q + (1-q)\left[1 + p + (1-p)(1+E[X \mid T_1])\right]$$

$q + p = ?$

$$E[X \mid T_1] = q + (1-q) + p(1-q) + (1-p)(1-q) + (1-p)(1-q)\,E[X \mid T_1]$$

$$E[X \mid T_1]\big(1 - (1-p)(1-q)\big)$$

solve for like a variable

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Fall 2010)

Problem Set 4: Solutions

1. (a) From the joint PMF, there are six (x, y) coordinate pairs with nonzero probabilities of occurring. These pairs are (1, 1), (1, 3), (2, 1), (2, 3), (4, 1), and (4, 3). The probability of a pair is proportional to the sum of the squares of the coordinates of the pair, $x^2 + y^2$. Because the probability of the entire sample space must equal 1, we have:

$$(1+1)c + (1+9)c + (4+1)c + (4+9)c + (16+1)c + (16+9)c = 1.$$

Solving for c, we get $c = \boxed{\frac{1}{72}}$.

(b) There are three sample points for which $y < x$:

$$\mathbf{P}(Y < X) = \mathbf{P}(\{(2,1)\}) + \mathbf{P}(\{(4,1)\}) + \mathbf{P}(\{(4,3)\}) = \frac{5}{72} + \frac{17}{72} + \frac{25}{72} = \boxed{\frac{47}{72}}.$$

(c) There are two sample points for which $y > x$:

$$\mathbf{P}(Y > X) = \mathbf{P}(\{(1,3)\}) + \mathbf{P}(\{(2,3)\}) = \frac{10}{72} + \frac{13}{72} = \boxed{\frac{23}{72}}.$$

(d) There is only one sample point for which $y = x$:

$$\mathbf{P}(Y = X) = \mathbf{P}(\{(1,1)\}) = \boxed{\frac{2}{72}}.$$

Notice that, using the above two parts,

$$\mathbf{P}(Y < X) + \mathbf{P}(Y > X) + \mathbf{P}(Y = X) = \frac{47}{72} + \frac{23}{72} + \frac{2}{72} = 1$$

as expected.

(e) There are three sample points for which $y = 3$:

$$\mathbf{P}(Y = 3) = \mathbf{P}(\{(1,3)\}) + \mathbf{P}(\{(2,3)\}) + \mathbf{P}(\{(4,3)\}) = \frac{10}{72} + \frac{13}{72} + \frac{25}{72} = \boxed{\frac{48}{72}}.$$

(f) In general, for two discrete random variable X and Y for which a joint PMF is defined, we have

$$p_X(x) = \sum_{y=-\infty}^{\infty} p_{X,Y}(x,y) \qquad \text{and} \qquad p_Y(y) = \sum_{x=-\infty}^{\infty} p_{X,Y}(x,y).$$

In this problem the ranges of X and Y are quite restricted so we can determine the marginal PMFs by enumeration. For example,

$$p_X(2) = \mathbf{P}(\{(2,1)\}) + \mathbf{P}(\{(2,3)\}) = \frac{18}{72}.$$

Overall, we get:

$$p_X(x) = \begin{cases} 12/72, & \text{if } x = 1, \\ 18/72, & \text{if } x = 2, \\ 42/72, & \text{if } x = 4, \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad p_Y(y) = \begin{cases} 24/72, & \text{if } y = 1, \\ 48/72, & \text{if } y = 3, \\ 0, & \text{otherwise.} \end{cases}$$

(g) In general, the expected value of any discrete random variable X equals

$$\mathbf{E}[X] = \sum_{x=-\infty}^{\infty} x p_X(x).$$

For this problem,

$$\mathbf{E}[X] = 1 \cdot \frac{12}{72} + 2 \cdot \frac{18}{72} + 4 \cdot \frac{42}{72} = \boxed{3}$$

and

$$\mathbf{E}[Y] = 1 \cdot \frac{24}{72} + 3 \cdot \frac{48}{72} = \boxed{\tfrac{7}{3}}.$$

To compute $\mathbf{E}[XY]$, note that $p_{X,Y}(x,y) \neq p_X(x)p_Y(y)$. Therefore, X and Y are not independent and we cannot assume $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$. Thus, we have

$$\begin{aligned}\mathbf{E}[XY] &= \sum_x \sum_y xy p_{X,Y}(x,y) \\ &= 1 \cdot \frac{2}{72} + 2 \cdot \frac{5}{72} + 4 \cdot \frac{17}{72} + 3 \cdot \frac{10}{72} + 6 \cdot \frac{13}{72} + 12 \cdot \frac{25}{72} = \boxed{\tfrac{61}{9}}.\end{aligned}$$

(h) The variance of a random variable X can be computed as $\mathbf{E}[X^2] - \mathbf{E}[X]^2$ or as $\mathbf{E}[(X - \mathbf{E}[X])^2]$. We use the second approach here because X and Y take on such limited ranges. We have

$$\text{var}(X) = (1-3)^2 \frac{12}{72} + (2-3)^2 \frac{18}{72} + (4-3)^2 \frac{42}{72} = \boxed{\tfrac{3}{2}}$$

and

$$\text{var}(Y) = (1 - \frac{7}{3})^2 \frac{24}{72} + (3 - \frac{7}{3})^2 \frac{48}{72} = \boxed{\tfrac{8}{9}}.$$

X and Y are not independent, so we cannot assume $\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$. The variance of $X+Y$ will be computed using $\text{var}(X+Y) = \mathbf{E}[(X+Y)^2] - (\mathbf{E}[X+Y])^2$. Therefore, we have

$$\mathbf{E}[(X+Y)^2] = 4 \cdot \frac{2}{72} + 9 \cdot \frac{5}{72} + 25 \cdot \frac{17}{72} + 16 \cdot \frac{10}{72} + 25 \cdot \frac{13}{72} + 49 \cdot \frac{25}{72} = \frac{547}{18}.$$

$$(\mathbf{E}[X+Y])^2 = (\mathbf{E}[X] + \mathbf{E}[Y])^2 = \left(3 + \frac{7}{3}\right)^2 = \frac{256}{9}.$$

Therefore,

$$\text{var}(X+Y) = \frac{547}{18} - \frac{256}{9} = \boxed{\tfrac{35}{18}}.$$

(i) There are four $(x,\ y)$ coordinate pairs in A : (1,1), (2,1), (4,1), and (4,3). Therefore, $\mathbf{P}(A) = \frac{1}{72}(2 + 5 + 17 + 25) = \frac{49}{72}$. To find $\mathbf{E}[X \mid A]$ and $\text{var}(X \mid A)$, $p_{X|A}(x)$ must be calculated. We have

$$p_{X|A}(x) = \begin{cases} 2/49, & \text{if } x = 1, \\ 5/49, & \text{if } x = 2, \\ 42/49, & \text{if } x = 4, \\ 0, & \text{otherwise,} \end{cases}$$

$$\begin{aligned}
\mathbf{E}[X \mid A] &= 1 \cdot \frac{2}{49} + 2 \cdot \frac{5}{49} + 4 \cdot \frac{42}{49} = \boxed{\frac{180}{49}}, \\
\mathbf{E}[X^2 \mid A] &= 1^2 \cdot \frac{2}{49} + 2^2 \cdot \frac{5}{49} + 4^2 \cdot \frac{42}{49} = \frac{694}{49}, \\
\text{var}(X \mid A) &= \mathbf{E}[X^2 \mid A] - (\mathbf{E}[X \mid A])^2 = \frac{694}{49} - \left(\frac{180}{49}\right)^2 = \boxed{\frac{1606}{2401}},
\end{aligned}$$

2. Consider a sequence of six independent rolls of this die, and let X_i be the random variable corresponding to the ith roll.

   (a) What is the probability that exactly three of the rolls have result equal to 3? Each roll X_i can either be a 3 with probability 1/4 or not a 3 with probability 3/4. There are $\binom{6}{3}$ ways of placing the 3's in the sequence of six rolls. After we require that a 3 go in each of these spots, which has probability $(1/4)^3$, our only remaining condition is that either a 1 or a 2 go in the other three spots, which has probability $(3/4)^3$. So the probability of exactly three rolls of 3 in a sequence of six independent rolls is $\boxed{\binom{6}{3}(\frac{1}{4})^3(\frac{3}{4})^3}$.

   (b) What is the probability that the first roll is 1, given that exactly two of the six rolls have result of 1? The probability of obtaining a 1 on a single roll is 1/2, and the probability of obtaining a 2 or 3 on a single roll is also 1/2. For the purposes of solving this problem we treat obtaining a 2 or 3 as an equivalent result. We know that there are $\binom{6}{2}$ ways of rolling exactly two 1's. Of these $\binom{6}{2}$ ways, exactly $\binom{5}{1} = 5$ ways result in a 1 in the first roll, since we can place the remaining 1 in any of the five remaining rolls. The rest of the rolls must be either 2 or 3. Thus, the probability that the first roll is a 1 given exactly two rolls had an outcome of 1 is $\boxed{\frac{5}{\binom{6}{2}}}$.

   (c) We are now told that exactly three of the rolls resulted in 1 and exactly three resulted in 2. What is the probability of the sequence 121212? We want to find

   $$\mathbf{P}(121212 \,|\, \text{exactly three 1's and three 2's}) = \frac{\mathbf{P}(121212)}{\mathbf{P}(\text{exactly 3 ones and 3 twos})}.$$

   Any particular sequence of three 1's and three 2's will have the same probability: $(1/2)^3(1/4)^3$. There are $\binom{6}{3}$ possible rolls with exactly three 1's and three 2's. Therefore,

   $$\mathbf{P}(121212 \,|\, \text{exactly three 1's and three 2's}) = \boxed{\frac{1}{\binom{6}{3}}}.$$

   (d) Conditioned on the event that at least one roll resulted in 3, find the conditional PMF of the number of 3's. Let A be the event that at least one roll results in a 3. Then

   $$\mathbf{P}(A) = 1 - \mathbf{P}(\text{no rolls resulted in 3}) = 1 - \left(\frac{3}{4}\right)^6.$$

   Now let K be the random variable representing the number of 3's in the 6 rolls. The (unconditional) PMF $p_K(k)$ for K is given by

   $$p_K(k) = \binom{6}{k}\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{6-k}.$$

We find the conditional PMF $p_{k|A}(k \mid A)$ using the definition of conditional probability:

$$p_{K|A}(k \mid A) = \frac{\mathbf{P}(\{K = k\} \cap A)}{\mathbf{P}(A)}.$$

Thus we obtain

$$p_{K|A}(k \mid A) = \begin{cases} \frac{1}{1-(3/4)^6}\binom{6}{k}\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{6-k} & \text{if } k = 1, 2, \ldots, 6, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $p_{K|A}(0 \mid A) = 0$ because the event $\{K = 0\}$ and the event A are mutually exclusive. Thus the probability of their intersection, which appears in the numerator in the definition of the conditional PMF, is zero.

3. By the definition of conditional probability,

$$\mathbf{P}(X = i \mid X + Y = n) = \frac{\mathbf{P}(\{X = i\} \cap \{X + Y = n\})}{\mathbf{P}(X + Y = n)}.$$

The event $\{X = i\} \cap \{X + Y = n\}$ in the numerator is equivalent to $\{X = i\} \cap \{Y = n - i\}$. Combining this with the independence of X and Y,

$$\mathbf{P}(\{X = i\} \cap \{X + Y = n\}) = \mathbf{P}(\{X = i\} \cap \{Y = n - i\}) = \mathbf{P}(X = i)\mathbf{P}(Y = n - i).$$

In the denominator, $\mathbf{P}(X + Y = n)$ can be expanded using the total probability theorem and the independence of X and Y:

$$\begin{aligned} \mathbf{P}(X + Y = n) &= \sum_{i=1}^{n-1} \mathbf{P}(X = i)\mathbf{P}(X + Y = n \mid X = i) \\ &= \sum_{i=1}^{n-1} \mathbf{P}(X = i)\mathbf{P}(i + Y = n \mid X = i) \\ &= \sum_{i=1}^{n-1} \mathbf{P}(X = i)\mathbf{P}(Y = n - i \mid X = i) \\ &= \sum_{i=1}^{n-1} \mathbf{P}(X = i)\mathbf{P}(Y = n - i) \end{aligned}$$

Note that we only get non-zero probability for $i = 1, \ldots, n - 1$ since X and Y are geometric random variables.

The desired result is obtained by combining the computations above and using the geometric

PMF explicitly:

$$\begin{aligned}
\mathbf{P}(X = i \mid X + Y = n) &= \frac{\mathbf{P}(X = i)\mathbf{P}(Y = n - i)}{\sum_{i=1}^{n-1} \mathbf{P}(X = i)\mathbf{P}(Y = n - i)} \\
&= \frac{(1-p)^{i-1}p(1-p)^{n-i-1}p}{\sum_{i=1}^{n-1}(1-p)^{i-1}p(1-p)^{n-i-1}p} \\
&= \frac{(1-p)^n}{\sum_{i=1}^{n-1}(1-p)^n} \\
&= \frac{(1-p)^n}{(1-p)^n \sum_{i=1}^{n-1} 1} \\
&= \frac{1}{n-1}, \quad i = 1, \ldots, n-1.
\end{aligned}$$

4. (a) Since $\mathbf{P}(A) > 0$, we can show independence through $\mathbf{P}(B) = \mathbf{P}(B \mid A)$:

$$\mathbf{P}(B \mid A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} = \frac{\binom{8}{6}p^6(1-p)^2 p}{\binom{8}{6}p^6(1-p)^2} = p = \mathbf{P}(B).$$

Therefore, A and B are independent.

(b) Let C be the event "3 heads in the first 4 tosses" and let D be the event "2 heads in the last 3 tosses". Since there are no overlap in tosses in C and D, they are independent:

$$\begin{aligned}
\mathbf{P}(C \cap D) &= \mathbf{P}(C)\mathbf{P}(D) \\
&= \binom{4}{3}p^3(1-p) \cdot \binom{3}{2}p^2(1-p) \\
&= 12p^5(1-p)^2.
\end{aligned}$$

(c) Let E be the event "4 heads in the first 7 tosses" and let F be the event "2nd head occurred during 4th trial". We are asked to find $\mathbf{P}(F \mid E) = \mathbf{P}(F \cap E)/\mathbf{P}(E)$. The event $F \cap E$ occurs if there is 1 head in the first 3 trials, 1 head on the 4th trial, and 2 heads in the last 3 trials. Thus, we have

$$\begin{aligned}
\mathbf{P}(F \mid E) &= \frac{\mathbf{P}(F \cap E)}{\mathbf{P}(E)} = \frac{\binom{3}{1}p(1-p)^2 \cdot p \cdot \binom{3}{2}p^2(1-p)}{\binom{7}{4}p^4(1-p)^3} \\
&= \frac{\binom{3}{1} \cdot 1 \cdot \binom{3}{2}}{\binom{7}{4}} = \frac{9}{35}.
\end{aligned}$$

Alternatively, we can solve this by counting. We are given that 4 heads occurred in the first 7 tosses. Each sequence of 7 trials with 4 heads is equally probable, the discrete uniform probability law can be used here. There are $\binom{7}{4}$ outcomes in E. For the event $E \cap F$, there are $\binom{3}{1}$ ways to arrange 1 head in the first 3 trials, 1 way to arrange the 2nd head in the 4th trial and $\binom{3}{2}$ ways to arrange 2 heads in the first 3 trials. Therefore,

$$\mathbf{P}(F \mid E) = \frac{\binom{3}{1} \cdot 1 \cdot \binom{3}{2}}{\binom{7}{4}} = \frac{9}{35}.$$

(d) Let G be the event "5 heads in the first 8 tosses" and let H be the event "3 heads in the last 5 tosses". These two events are not independent as there is some overlap in the tosses (the 6th, 7th, and 8th tosses). To compute the probability of interest, we carefully count all the disjoint, possible outcomes in the set $G \cap H$ by conditioning on the number of heads in the 6th, 7th, and the 8th tosses. We have

$$\begin{aligned}
\mathbf{P}(G \cap H) &= \mathbf{P}(G \cap H \mid 1 \text{ head in tosses 6–8})\mathbf{P}(1 \text{ head in tosses 6–8}) \\
&\quad + \mathbf{P}(G \cap H \mid 2 \text{ heads in tosses 6–8})\mathbf{P}(2 \text{ heads in tosses 6–8}) \\
&\quad + \mathbf{P}(G \cap H \mid 3 \text{ heads in tosses 6–8})\mathbf{P}(3 \text{ heads in tosses 6–8})
\end{aligned}$$

$$\begin{aligned}
&= \binom{5}{4} p^4(1-p) \cdot p^2 \cdot \binom{3}{1} p(1-p)^2 \\
&\quad + \binom{5}{3} p^3(1-p)^2 \cdot \binom{2}{1} p(1-p) \cdot \binom{3}{2} p^2(1-p) \\
&\quad + \binom{5}{2} p^2(1-p)^3 \cdot (1-p)^2 \cdot p^3.
\end{aligned}$$

$$= 15p^7(1-p)^3 + 60p^6(1-p)^4 + 10p^5(1-p)^5.$$

5. Let I_k be the reward paid at time k. We have

$$\mathbf{E}[I_k] = \mathbf{P}(I_k = 1) = \mathbf{P}(\text{T at time } k \text{ and H at time } k-1) = p(1-p).$$

Computing $\mathbf{E}[R]$ is immediate because

$$\mathbf{E}[R] = \mathbf{E}\left[\sum_{k=1}^{n} I_k\right] = \sum_{k=1}^{n} \mathbf{E}[I_k] = np(1-p).$$

The variance calculation is not as easy because the I_ks are not all independent:

$$\begin{aligned}
\mathbf{E}[I_k^2] &= p(1-p) \\
\mathbf{E}[I_k I_{k+1}] &= 0 \quad \text{because rewards at times } k \text{ and } k+1 \text{ are inconsistent} \\
\mathbf{E}[I_k I_{k+\ell}] &= \mathbf{E}[I_k]\mathbf{E}[I_{k+\ell}] = p^2(1-p)^2 \quad \text{for } \ell \geq 2
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}[R^2] &= \mathbf{E}[(\sum_{k=1}^{n} I_k)(\sum_{m=1}^{n} I_m)] = \sum_{k=1}^{n}\sum_{m=1}^{n} \mathbf{E}[I_k I_m] \\
&= \underbrace{np(1-p)}_{n \text{ terms with } k=m} + \underbrace{0}_{2(n-1) \text{ terms with } |k-m|=1} + \underbrace{(n^2-3n+2)p^2(1-p)^2}_{n^2-3n+2 \text{ terms with } |k-m|>1}
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(R) &= \mathbf{E}[R^2] - (\mathbf{E}[R])^2 \\
&= np(1-p) + (n^2-3n+2)p^2(1-p)^2 - n^2p^2(1-p)^2 \\
&= np(1-p) - (3n-2)p^2(1-p)^2.
\end{aligned}$$

G1†. (a) We know that I_A is a random variable that maps a 1 to the real number line if ω occurs within an event A and maps a 0 to the real number line if ω occurs outside of event A. A similar argument holds for event B. Thus we have,

$$I_A(\omega) = \begin{cases} 1, & \text{with probability } \mathbf{P}(A) \\ 0, & \text{with probability } 1 - \mathbf{P}(A) \end{cases}$$

$$I_B(\omega) = \begin{cases} 1, & \text{with probability } \mathbf{P}(B) \\ 0, & \text{with probability } 1 - \mathbf{P}(B) \end{cases}$$

If the random variables, A and B, are independent, we have $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$. The indicator random variables, I_A and I_B, are independent if, $\mathbf{P}_{I_A,I_B}(x,y) = \mathbf{P}_{I_A}(x)\mathbf{P}_{I_B}(y)$
We know that the intersection of A and B yields.

$$\begin{aligned}
\mathbf{P}_{I_A,I_B}(1,1) &= \mathbf{P}_{I_A}(1)\mathbf{P}_{I_B}(1) \\
&= \mathbf{P}(A)\mathbf{P}(B) \\
&= \mathbf{P}(A \cap B)
\end{aligned}$$

We also have,

$$\begin{aligned}
\mathbf{P}_{I_A,I_B}(1,1) &= \mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}_{I_A}(1)\mathbf{P}_{I_B}(1) \\
\mathbf{P}_{I_A,I_B}(0,1) &= \mathbf{P}(A^c \cap B) = \mathbf{P}(A^c)\mathbf{P}(B) = \mathbf{P}_{I_A}(0)\mathbf{P}_{I_B}(1) \\
\mathbf{P}_{I_A,I_B}(1,0) &= \mathbf{P}(A \cap B^c) = \mathbf{P}(A)\mathbf{P}(B^c) = \mathbf{P}_{I_A}(1)\mathbf{P}_{I_B}(0) \\
\mathbf{P}_{I_A,I_B}(0,0) &= \mathbf{P}(A^c \cap B^c) = \mathbf{P}(A^c)\mathbf{P}(B^c) = \mathbf{P}_{I_A}(0)\mathbf{P}_{I_B}(0)
\end{aligned}$$

(b) If $X = I_A$, we know that

$$\mathbf{E}[X] = \mathbf{E}[I_A] = 1 \cdot \mathbf{P}(A) + 0 \cdot (1 - \mathbf{P}(A)) = \mathbf{P}(A)$$

†Required for 6.431; optional for 6.041

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

Name: Michael Plasmeier

Recitation Instructor: Dimitri

TA: Aliaa 3PM

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	

- 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.

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

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 (((

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, $\text{var}(X+Y) \geq \text{var}(X)$.

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

Problem 1: (75 points)

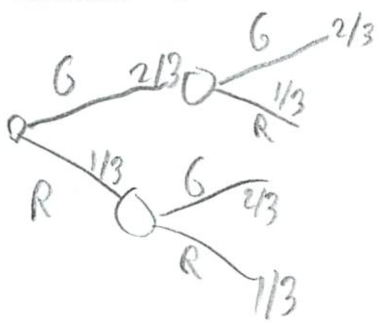
Note: All parts can be done independently, with the exception of the last part. Just in case you made a mistake in the previous part, you can use a symbol for the expression you found there, and use that symbol in the formulas for the last part.

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

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

Order does not matter

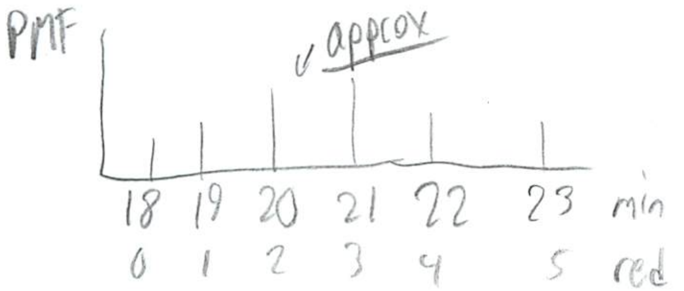
Subtract out 18 min

0 min → all green

each red 1 min

X = # Red lights

Y = minutes of commute



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

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

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

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

+5

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

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

Y
18
19
20
21
22
23

→ over

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

+2

2/2

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

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

+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 these come from? 3/10

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 don't know 1st 3 lights
5th = green

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

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

4/10

not doing formula based stuff problem

divide something out what's total commute = 1 ī divide anything out?

- ī 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?

$P[\overset{A}{\text{2 out of 1st 3 lights red}} \mid \overset{B}{X=3}]$

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

− not independent.

$\underbrace{R \;\_\; R}_{\text{prob they are here}} \; R$ — ∩ 3 red lights can be anywhere

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

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

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

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

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

should have reviewed that p-set a∪ 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

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

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

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

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)

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

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

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

$P(A|B) = \frac{P(A \cap B)}{P(B)}$ ✓

A + B independent here

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

but this is not what you're asked to calculate

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

NO!

$= \frac{1}{4}$

4

right if A + B independent

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

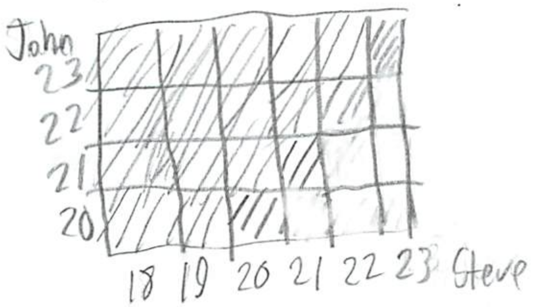
Steve = drives
John = subway

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(Steven'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(Steven's ≤ 23) P(John's = 23)

1

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

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

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

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

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?

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

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

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}$$

$P(B \mid A) = 1$ since if John took 23 minutes, than Steve's commute can be any thing (Ω)

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

$P(B)$ = 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 tre or false

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

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

1

So B and B^c have te same potential for overlap - overlapping a little but not too much

$B \cap B^c = \Omega$

bad, I know

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

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

Ur did not really do proof

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~~

$$\mathrm{Var}(x) = E[(X - E[x])^2] = E[X^2] - E[X]^2$$

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 ?

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:

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

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:

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

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:

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

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

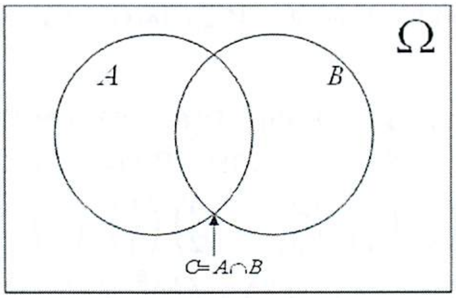
$$\mathbf{P}(A \cap B^c) = \mathbf{P}(A)\mathbf{P}(B^c).$$

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,

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

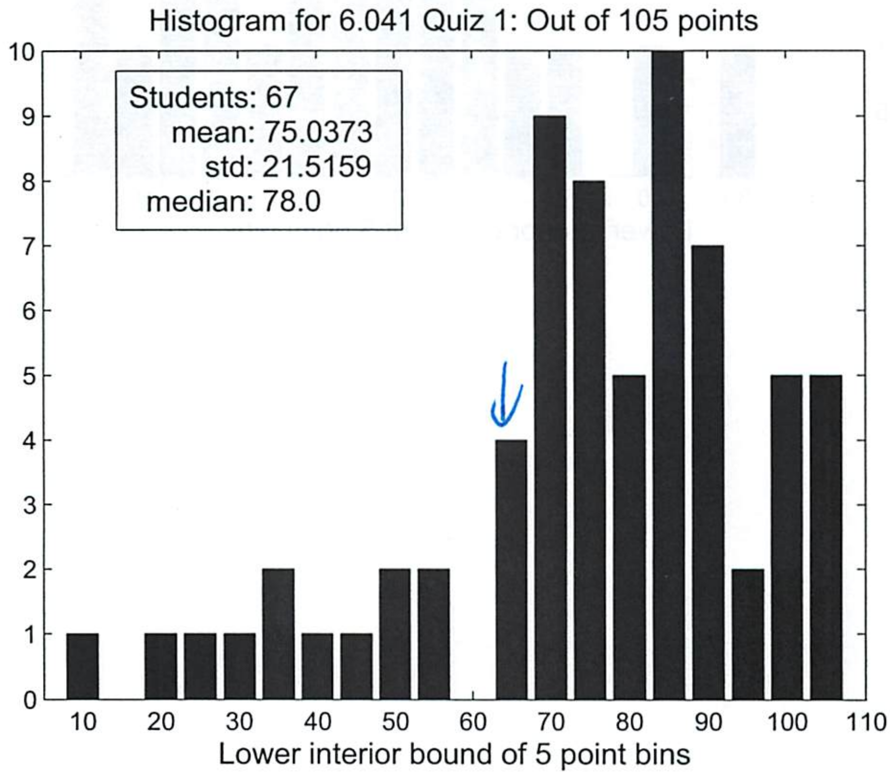
As shown, $\mathbf{P}(B \mid A \cap C^c) \neq \mathbf{P}(B \mid C^c)$.

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.



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