* Announcement

① HW6 due 10/7

② Midterm logistics on Piazza

(i) OH Wed 9AM ~ 6PM, 10PM ~ 11PM

(No OH on Friday)

(ii) Exam-prep section Wed 4-6 PM

(iii) Midterm Practice problems on Piazza

STAT 88: Lecture 17

Midterm Review on Wed lecture

Option 1. students write down specific questions to go over in class

2. I pick an assortment of questions for midterm practices

Contents

Section 5.6: Expectation by Conditioning

Warm up: A lost tourist arrives at a point with 2 roads. Road A brings him back to the same point after 1 hour of walk. Road B leads to the city in 2 hours. Assuming the tourist randomly chooses a road at all times, what is the expected time until the tourist arrives to the city?

T = time to arrive in city

$E(T)$?

$$R = \begin{cases} \text{road } A & \text{w.p. } \frac{1}{2} \\ \text{road } B & \text{w.p. } \frac{1}{2} \end{cases}$$

$$E(T) = \underbrace{E(T \mid R=A)}_{= 1+E(T)} \cdot \underbrace{P(R=A)}_{= 0.5} + \underbrace{E(T \mid R=B)}_{= 2} \cdot \underbrace{P(R=B)}_{= 0.5}$$

$\rightsquigarrow E(T) = 0.5 + 0.5 \cdot E(T) + 1$

$\rightsquigarrow E(T) = 3$ hrs

Last time

Expectation by conditioning

	$M = 2$	$M = 3$
$S = 2$	0	$\frac{1}{3}$
$S = 1$	$\frac{1}{3}$	$\frac{1}{3}$

$P(S=2) = \frac{1}{3}$

$P(S=1) = \frac{2}{3}$

$P(M=2) = \frac{1}{3}$ $\quad$ $P(M=3) = \frac{2}{3}$

Conditional expectation:

$$E(S|M = m) = \sum_{\text{all } s} s \cdot P(S = s|M = m).$$

$P(S=1|M=3) = \frac{P(S=1, M=3)}{P(M=3)} = \frac{1}{2}$

$P(S=2|M=3) = \frac{1}{2}$

$\rightsquigarrow E(S|M=3) = 1\cdot\frac{1}{2} + 2\cdot\frac{1}{2} = \frac{3}{2}$

Expectation by conditioning formula:

$$E(S) = \sum_{\text{all } m} E(S|M = m)P(M = m),$$

$\rightsquigarrow E(S|M=2) = 1$

$P(S=1|M=2) = 1$

$P(S=2|M=2) = 0$

or equivalently

$$E(S) = E(E(S|M)).$$

Note that $E(S|M)$ is a random variable (a function of M). When M is realized to a value $M = m$, $E(S|M)$ takes a value $E(S|M = m)$.

$M=2 \rightsquigarrow E(S|M=2) = 1$

$M=3 \rightsquigarrow E(S|M=3) = \frac{3}{2}$

$P(X=x, Y=y) = P(X=x)P(Y=y)$ for all x,y

Note: If two random variables X and Y are independent, the conditional distribution of X given $Y = y$ is the same regardless of y:

$$P(X = x|Y = y) = P(X = x).$$

$\frac{P(X=x, Y=y)}{P(Y=y)} = P(X=x)$

$\quad\;\, \shortparallel$

$P(X=x | Y=y)$

Then

$$\begin{aligned} E(X|Y = y) &= \sum_{\text{all x}} x \cdot P(X = x|Y = y) \\ &= \sum_{\text{all x}} x \cdot P(X = x) \\ &= E(X). \end{aligned}$$

Expectation of a Geometric Waiting Time

If $X \sim \text{Geom}(p)$, by our rule for finding expectations by conditioning,

$$E(X) = \underbrace{E(X|X=1)}_{=1}\underbrace{P(X=1)}_{=p} + \underbrace{E(X|X>1)}_{=1+E(X)}\underbrace{P(X>1)}_{=1-p}.$$

$E(X|X>1)$

$= E(X|\text{ first trial is failure})$

F | _ _ _ _ S ← First success

$1 + \underbrace{E(X)}$

$\shortparallel$

$E(X|X>1)$

Solve the equation for $x = E(X)$:

$$x = 1 \cdot p + (1+x) \cdot (1-p),$$

so

$$p \cdot x = 1, \text{ so } x = \frac{1}{p}.$$

5.6. Expectation by Conditioning

More practice with conditional expectation.

Example: A fair coin is tossed 3 times. Let

- X be the number of heads in the first two tosses;
- Y be the number of heads in the last two tosses.

Outcomes

$X=2$: HHH, HHT

$X=1$: HTH, HTT, THH, THT

$X=0$: TTH, TTT

Find $E(Y|X=2)$.

$$= \sum_{\text{all } y} y \cdot P(Y=y \mid X=2) = 2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{3}{2}$$

Find $E(Y|X=1)$.

$$= \sum_{\text{all } y} y \cdot P(Y=y \mid X=1) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{2}{4} + 2 \cdot \frac{1}{4} = 1$$

Find $E(Y|X=0)$.

$$= \sum_{\text{all } y} y \cdot P(Y=y \mid X=0) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$$

	TT	HT, TH	HH
$X=x$	0	1	2
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Find $E(Y)$.

$$= \sum_{\text{all } x} E(Y|X=x)\, P(X=x)$$

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

$$= 1$$

Example: Tamara chooses an integer N uniformly at random from 1 to 425. She then chooses an integer X uniformly at random from $1, 2, \ldots, N$. Find $E(X)$.

$$E(X) = \sum_{\text{all } n} E(X \mid N=n)\, P(N=n)$$

Fix $N=n$

$$X \mid N=n \sim \text{Unif}\{1, 2, \cdots, \underbrace{n}_{\text{little } n}\}$$

$$E(X \mid N=n) = \underbrace{\frac{n+1}{2}}_{\text{Depend on } n}$$

$$E(X) = \sum_{n=1}^{425} E(X \mid N=n) \cdot P(N=n)$$

$$= \sum_{n=1}^{425} \left(\frac{n+1}{2}\right) \cdot P(N=n)$$

$$\frac{1}{2} \sum (n+1) \cdot P(N=n)$$

$$= \frac{1}{2} \left(\sum n \cdot P(N=n) + \sum P(N=n) \right)$$

$$= \frac{1}{2} \underbrace{\sum_{n=1}^{425} n \cdot P(N=n)}_{= E(N)} + \frac{1}{2} \underbrace{\sum_{n=1}^{425} P(N=n)}_{= 1}$$

$N \sim \text{Unif}\{1, 2, \cdots, 425\}$

$$E(N) = \frac{1+425}{2}$$

$$= \frac{426}{2}$$

$$= \frac{214}{2} = 107$$

Example: Tamara chooses an integer N [redacted] at random from Pois(μ). She then picks N cards from a deck with replacement. Find the expected number of ace cards.

(1) $X = \#$ ace cards

$E(X) = \sum_n E(X|N=n)\, P(N=n)$

$X|N=n \sim \text{Binom}(n, \frac{4}{52})$

$E(X|N=n) = n \cdot \frac{4}{52}$

$E(X) = \sum_n n \cdot \frac{4}{52} \cdot P(N=n)$

$= \frac{4}{52} \underbrace{\sum_n n \cdot P(N=n)}_{= E(N) = \mu}$

$= \frac{4}{52} \cdot \mu$

(2) $X_i = \begin{cases} 1 & \text{if } i\text{th card is ace} \\ 0 & \text{else} \end{cases}$

$X = \#$ ace cards

$= X_1 + X_2 + \cdots + X_N$

$X|N=n \sim \text{Binom}(n, \frac{4}{52})$

Midterm Review

DISTRIBUTION FACTS

Name and Parameters	Values	$P(X = k)$	$E(X)$
Bernoulli (p) (also Indicator)	$0, 1$	$P(X = 1) = p$	p
Uniform on 1 through N	$1, 2, \ldots, N$	$\frac{1}{N}$	$\frac{N+1}{2}$
Binomial (n, p)	$0, 1, 2, \ldots, n$	$\binom{n}{k}p^k(1-p)^{n-k}$	np
Hypergeometric (N, G, n)	$0, 1, 2, \ldots, n$	$\frac{\binom{G}{k}\binom{N-G}{n-k}}{\binom{N}{n}}$	$n\frac{G}{N}$
Poisson (μ)	$0, 1, 2, 3, \ldots$	$e^{-\mu}\frac{\mu^k}{k!}$	μ
Geometric (p)	$1, 2, 3, \ldots$	$(1-p)^{k-1}p$	$\frac{1}{p}$

If X has the Pois(μ) distribution, Y has the Pois(λ) distributionn, and X and Y are independent, then $X + Y$ has the Pois($\mu + \lambda$).

CLASS CONVENTIONS

- Unless otherwise stated, coins are two-sided and fair, and dice are six-sided and fair.
- "At random" means uniformly at random; all outcomes equally likely.
- "The number of trials till an event occurs" means the number of trials up to and including the trial at which the event occurs for the first time.

F F S # trials until first success = 3

ACRONYMS

- i.i.d.: independent and identically distributed.
- SRS: simple random sample (drawn at random without replacement).
- CDF: The cumulative distribution function of X is the function F defined by $F(x) = P(X \leq x)$.

Preparation Study the textbook (including the Exercise at the end of each Chapter), lecture notes, homework, and quizzes, Midterm review problems. Make sure to know the conditions under which you can do the different kinds of calculations.

Ex:

- Binomial(n, p) formula is used for the number of successes out of n independent p-trials.

- Geometric(p) formula is used for the number of independent p-trials until a first success.

During the Exam

- Start with the easy questions. When you open the midterm, skim through it quickly and mark the questions that are easy to answer and do those first. Not only will this help you maximize your scoring potential but it will also bolster your confidence. Once you are done with answering the easy questions, it's time to tackle the ones you skipped. Don't over-think straightforward questions.

- Read each question carefully. As you know, assumptions matter. If you have misread those then your solution will be off. For example, confusing "with replacement" and "without replacement" can have a massive effect. "The fourth head is on the 20th toss" is not the same as, "There are four heads in 20 tosses". Forcing youself to read slowly, underlining key assumptions as you read, is important for doing well. If you are done with the test, check your work by reading each question afresh and solving it again instead of just reading over your answer.

- Provide reasoning or a calculation in all questions. If you did a calculation in your head, write out the calculation you did in your head.

- Don't simplify any arithmetic or algebra in your answers unless a question explicitly asks you to. The unsimplified version shows us your thought process. Simplification isn't worth the time on the test, and besides, you might mess up the simplification.

$\sum_{k=1}^{5} (B\ )(B\ )$

- If an answer is taking you numerous lines of calculation, or complicated algebra/calculus, you've probably missed something. Rethink, or move to a different problem.

- If you don't know how to do a problem, try not to leave it blank. Almost always, you will have an idea of what might be relevant. If you write that, and it is indeed important for the problem, you might get some partial credit. That said, you shouldn't expect partial credit for everything you write. We'll be looking for substantive ideas and progress towards a solution.

Problem Solving Suggestions If the experiment involves two stages, draw a tree diagram.

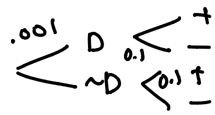
Ex D = event you have disease

+ = positive test for disease

$P(D) = 0.001$

Chance of false pos and false neg are 0.1

To find $P(D|+)$ make a tree:



If you find yourself writing something like 0.3^5, check that you have independence. You can't just assume independence, and it's wrong for situations like dealing cards etc. Either the problem has to provide an assumption of independence, or the assumption has to follow from the conditions of the experiment, e.g. sampling with replacement, or events based on separate sets of tosses, etc.

$P(A \cap B) = P(A)P(B)$ (X) - only when A and B are independent

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

To find probabilities or expectations, it sometimes helps to write down the outcome space.

Ex: A fair coin is tossed 3 times. If one of the tosses is head, what is the chance that there is at lease one tail?

HHH

HHT

HTH

HTT

THH

THT

TTH

~~TTT~~

$6/7$

When faced with a complicated probability, try and describe the outcome space in simple English.

Ex: Draw cards from a standard deck until 3 aces have appeared. Let X be the number of cards drawn. Find $P(X > x)$.

= drawing 0, 1, 2 aces in x draws.

$$P(X>x) = \frac{\binom{4}{0}\binom{48}{x}}{\binom{52}{x}} + \frac{\binom{4}{1}\binom{48}{x-1}}{\binom{52}{x}} + \frac{\binom{4}{2}\binom{48}{x-2}}{\binom{52}{x}}$$

If you are using the probability formula for one of the famous distributions, pause for a second and check the possible values. For example, this might prevent you from using the geometric (possible values $1, 2, 3, \ldots$) when you should be using the binomial (possible values 0 through n).

Ex: Let X be the number of bluberries in a bite of a blueberry muffin. What are the possible values of X? What kind of RV is X?

X in $\{0, 1, 2, \cdots\}$. $\sim$ Poisson

If a question asks for a distribution, start by listing all the possible values of the random variable. Take the time to do this carefully, identifying the minimum possible value (if there is one) and the maximum possible value (if there is one). Not only will it help focus your calculation of the probabilities, it might also get you partial credit for having understood the variable even if you didn't get the probabilities correctly.

Ex: A deck consists of 12 cards, 5 of which are blue and 7 green. I draw from the deck at random without replacement till both colors have appeared among the draws. Let D be the number of draws. Find the distribution of D.

If a question asks for an expectation, don't immediately try to find the distribution of the random variable and then apply the definition of expectation. Try properties of expectation first. The most common are additivity (see if you can write the random variable as a sum of simpler ones) and conditioning (see if you would know the expectation if you were given the result of an early stage of the experiment).

Ex: A fair die is rolled 14 times. Let X be the number of faces that appear exactly 2 times. Find $E(X)$.

The conditional distribution of Y given $X = x$ is just an ordinary distribution. You have to first recognize that "given $X = x$" means you can treat the random variable X as the constant x. You therefore have to provide the possible values of Y (under the condition that $X = x$) and the corresponding conditional probabilites given $X = x$. How you calculate those probabilities depends on the setting. Sometimes you can just see what they are because the condition $X = x$ simplifies your outcome space, and sometimes you have to use the division rule $P(Y = y|X = x) = P(X = x, Y = y)/P(X = x)$. This is not particular to finding conditional distributions. It's a feature of finding conditional probabilities in general.

Ex: Let $X_1 \sim \text{Pois}(a)$ and $X_2 \sim \text{Pois}(b)$ and they are independent. Let m be a fixed positive integer. Find the distribution of $X_1|X_1 + X_2 = m$. Recognize this as one of the famous ones and provide its parameters.

Ex: Let $X_1 \sim \text{Pois}(a)$ and $X_2 \sim \text{Pois}(b)$ and they are independent. Find $E(X_1 X_2)$.