MATH 2B RECITATION 2/1/12

ALDEN WALKER

The midterm is due Monday at noon. It has a 4 hour time limit. The front of the exam reads:

Please use a standard bluebook and put your name and section on the outside. You may use the textbook (Pitman), handouts, solution sets, your notes and homework, and TA notes (Someone else's notes handcopied by you are OK.) Calculators and computers are allowed, but only to do elementary numerical calculations (like a *finite* sum) and to use elementary built-in functions like logs and normal distribution functions (as an alternative to using tables)-not to do programming, simulations, numerical integration, calculus, symbolic computations, or functions like ones that calculate binomial or Poisson probabilities (built in discrete probability density functions). Please indicate clearly any work done in overtime. Points will be recorded separately and considered informally in course grades.

1. Basic Probability Rules

- If A and B are independent, $P(AB) = P(A)P(B)$
- If A and B are disjoint, $P(A \cup B) = P(A) + P(B)$
- No matter what, $P(A \cup B) = P(A) + P(B) - P(AB)$, and inclusion-exclusion extends this in general

2. Conditional Probability

- $P(A|B) = \frac{P(AB)}{P(B)}$
- Bayes' rule (if B_i partition the event space):

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_i P(A|B_i)P(B_i)}$$

- Know how to make tree diagrams and how to multiply to get probabilities of interest.

2.1. **Example.** The weather report says there's a 20% chance of rain today. I carry my umbrella independently from the weather report with probability 0.1. What is the chance of me getting wet? It's just $P(AB)$, where A is the event that it rains, and B is the event that I forget my umbrella. Since they are independent, the probability is $0.2 \times 0.9 = 0.18$.

Ok now suppose that I have a bad knee and that if it will rain today, then my knee has a 0.6 probability of hurting. If it won't rain, suppose it has a 0.1 probability of hurting. Suppose there is an a priori probability of raining of 0.1. Given that my knee is hurting today, what is the probability that it will rain.

We just use Bayes' rule:

$$P(\text{rain}|\text{hurt}) = \frac{P(\text{hurt}|\text{rain})P(\text{rain})}{P(\text{hurt}|\text{rain})P(\text{rain}) + P(\text{hurt}|\text{no rain})P(\text{no rain})} = \frac{0.6 \times 0.1}{0.6 \times 0.1 + 0.1 \times 0.9} = 0.4$$

This actually happens, by the way. I wasn't sure whether knees really hurt because of rain (low pressure really I guess), but my friend injured his knee and it does in fact happen.

2.2. **Example.** Sometimes, a question might ask about elementary probability, but you might have to think carefully about how to figure out the answer.

I have two urns A and B. A has 2 red and 2 black balls, and B has 1 red and 3 black balls. I pick two balls at random from A and two from B, and I swap which urn they are in (the pair from A goes to B and the pair from B goes to A). What's the probability that A or B now contains 3 red balls?

The right way to do this is to case out how it could come to contain 3 red balls. It must be the case that we pick the 2 black balls from urn A and a red ball from B, *or* we pick the two red balls from A and two black balls from B. Thus,

$$
\begin{aligned}
P(3\text{ red}) &= P(\text{pick 2 black from A, a red from B}) + P(\text{pick 2 red from A, two black from B}) \\
&= P(\text{pick 2 black from A})P(\text{red from B}) + P(\text{pick 2 red from A})P(\text{2 black from B}) \\
&= \left[2\frac{1}{4}\frac{1}{3}\right]\left[2\cdot 3\frac{1}{4}\frac{1}{3}\right] + \left[2\frac{1}{4}\frac{1}{3}\right]\left[1 - 2\cdot 3\frac{1}{4}\frac{1}{3}\right] \\
&= \frac{1}{6}
\end{aligned}
$$

Note here $2(1/4)(1/3)$ is the probability of picking a given pair of balls out of an urn, since it's the probability of picking the first, times the probability of picking the second, times the 2 ways they could be ordered.

Anyway the point is just to be careful with these.

3. Binomial Distribution

- If X_i is a random variable which is 1 with probability p and 0 otherwise and the X_i are independent, then $Y = \sum_{i=1}^{n} X_i$ is a random variable, and it has the binomial distribution, where $P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
- If p is close to 0.5 and n is fairly large, then you can approximate the binomial distribution with the normal distribution (see the central limit theorem later), where the mean is $\mu = np$ and the standard deviation is $\sqrt{np(1-p)}$, so if X is (n, p) Binomial, then $P(a \leq X \leq b) \approx \Phi(\frac{b+1/2-np}{\sqrt{np(1-p)}}) - \Phi(\frac{a-1/2-np}{\sqrt{np(1-p)}})$. Remember that little 1/2 you stick in there. It's not technically needed, since the central limit theorem gives you the convergence you want, but it makes the approximation better.
- If p is small, then you can approximate the binomial with the Poisson distribution, in which $P(Y = k) = e^{-\mu}\frac{\mu^k}{k!}$, so if X is (n, p) binomial, then $P(X = k) \approx e^{-np}\frac{(np)^k}{k!}$

4. Note

Always keep in mind (and be specific about) what your variables are. If the variable is: "the outcome of a coin flip," you cannot add it! If the variable is "the number of heads in a single coin flip," then you can certainly go ahead and add it to things. It is easy to get confused or to not say what things are exactly because you know what they are in your head, so you get the right answer without writing down something that's exactly right. Please write down what you mean! That way, everybody is happy.

4.1. **Example.** Let X be the outcome of a spin of a roulette wheel (0–37). What is the expected number of 4's in 10 rolls? This is not $10 * E(X)$! Define the indicator function Y to be 1 if X is 4 and 0 otherwise. Then $P(Y = 1) = 1/38$, so the expected number of 4's in 10 rolls is $10 * E(Y) = 10/38$.

5. Indicator Functions

If you are asked for the expected number of something and not necessarily the probability, it is a good bet that using indicator functions will be handy. The idea is that no matter how complicated the situation is, you make an indicator function for each thing you want to count and figure out the probability of it being counted. This is usually easier than figuring out the probability of counting exactly k things for all k. Let's do an example.

5.1. **Example.** Suppose 20 people each roll a die in turn; they put as many balls as the die shows into a box. What is the expected number of balls in the box after this? It might be tempting to make an indicator function for the balls, but that isn't the way to do this one. Let B_i be the number of balls placed in the box by person i. Then clearly we want to know $E(\sum B_i)$. This is $\sum E(B_i)$ by the formula, and $E(B_i)$ is just expect roll of a die, i.e. 3.5, so the expected number of balls is $20 \times 3.5 = 70$.

5.2. **Example 2.** Now 20 people each pick one of 10 boxes which each contain 6 red balls. Upon picking a box, they roll a die and color that many balls black at random. What is the expected number of red balls left after this?

Now let B_i be the indicator function for ball i being red. We want to figure out $E(B_i)$ after the procedure. For each person, there is a $\frac{9}{10}$ chance of them not picking the box in which ball i is. If they do pick the box, the probability of remaining red is 5/6 if they roll a 1, 4/6 if they roll a 2, etc, so the probability of remaining red after one person goes is

$$P(\text{stay red through one person}) = \frac{9}{10} + \frac{1}{10}\left[\frac{1}{6}\frac{5}{6} + \frac{1}{6}\frac{4}{6} + \frac{1}{6}\frac{3}{6} + \frac{1}{6}\frac{2}{6} + \frac{1}{6}\frac{1}{6} + \frac{1}{6}\frac{0}{6}\right] = \frac{113}{120}$$

Therefore, the probability of remaining red, i.e. that $B_i = 1$, is $\left(\frac{113}{120}\right)^{20} \approx 0.3$. Thus $E(\sum B_i) = \sum E(B_i) = 60 \times 0.3 = 18$.

6. Continuous Distributions, etc

Given a nice integrable function f whose total integral is 1, you can call that a probability density function, and to get $P(a < X < b)$ for a random variable from that distribution, you calculate $P(a < X < b) = \int_a^b f(x)dx$. The function $F(t) = P(X < t) = \int_{-\infty}^t f(x)dx$ is the cumulative distribution function.

7. Expectation

- The formula is $E(X) = \int_{-\infty}^{\infty} xf(x)dx$ where f is the density function
- $E(\sum_i X_i) = \sum_i E(X_i)$, even if the X_i are not independent.
- $E(XY) = E(X)E(Y)$, if they are independent.
- $Var(X) = E(X^2) - (E(X))^2$.

8. Central Limit Theorem

If you have a bunch of independent random variables selected from the same distribution, then no matter what the distribution is, their sum/average will be approximately normal (converge to normal as the number of variables goes up). You just need to know the mean and variance of the distribution. Formally, if $\{X_i\}_{i=1}^n$ are all picked from a distribution with mean μ and variance σ^2, then $P(\sum_i X_i < t) \approx \Phi(\frac{t-n\mu}{\sigma\sqrt{n}})$.

8.1. **Example.** I roll a single die—what is the approximate probability that I get fewer than 950 sixes in 6000 rolls? Here we are interested in indicator functions X_i for each roll, which are 1 with probability 1/6 and 0 otherwise. The distribution of each indicator is certainly not normal, but the sum of 6000 of them will be approximately normal. The mean, i.e. expected value of X_i is 1/6 and $E(X_i^2) = 1/6$, so the variance is $E(X_i^2) - (E(X_i))^2 = 5/36$ Therefore the mean of $Y = \sum_i X_i$ is 1000, and the variance is 2500/3. By the normal approximation, the probability that I roll fewer than 500 sixes is $\Phi(\frac{950-1000}{\sqrt{2500/3}}) = 0.041$. Incidentally, the probability of getting only half as many as you expect, i.e. 500, is 1.6×10^{-67}.

9. Change of Variables

If $Y = g(X)$, and you know the distribution of X, then you probably want the distribution of Y. There is a formula for it in your book, but here is a good way to do it that usually is as easy as the formula:

$$\begin{aligned} f_Y(t) &= \frac{\mathrm{d}F_Y(t)}{\mathrm{d}t} \\ &= \frac{\mathrm{d}P(g(X) < t)}{\mathrm{d}t} \\ &= \frac{\mathrm{d}P(X < g^{-1}(t))}{\mathrm{d}t} \\ &= \frac{\mathrm{d}F_X(g^{-1}(t))}{\mathrm{d}t} \\ &= f_X(g^{-1}(t))\frac{\mathrm{d}g^{-1}(t)}{\mathrm{d}t} \end{aligned}$$

The only thing you have to memorize is the original setup and the strategy that you solve for X inside the probability and then the derivative of the probability (the cumulative distribution) is the density.

10. Multiple Variables / Joint Densities

10.1. **With Independent Normal Variables.** Don't integrate! Any linear combination of normal variables is normal—look at the formula on page 363!

10.2. **With Arbitrary Distributions.** If you're interested in the density of $X+Y$, where X and Y are independent, then you can use the convolution formula. Now if you're interested in the density of something more complicated, then it is a good idea to look at the probability again:

10.3. **Example.** Suppose X and Y are uniform on $[0,1]$—what is the density function for X/Y? By definition, $f_{X/Y}(t) = \frac{d}{dt}P(X/Y < t)$, so let's try to get at that probability. It is $P(X/Y < t) = P(X < Yt) = \iint_{(x,y):x/y<t} dxdy$. There are two cases here. If $t > 1$, then it's

$$\int_0^1 \int_{x/t}^1 dydx = \int_0^1 1 - \frac{x}{t} dx = 1 - \frac{1}{2t}$$

And if $t < 1$, then we only do x from 0 until t:

$$\int_0^t \int_{x/t}^1 dydx = \int_0^t 1 - \frac{x}{t} dx = t - \frac{t}{2} = t/2$$

Therefore, differentiating, we see $f_{X/Y}(t) = \begin{cases} \frac{1}{2t^2} & t > 1 \\ 1/2 & t < 1 \end{cases}$.

Notice that the expected value of this distribution does not exist.

11. Additional Examples

11.1. **Example.** Let Z be the random variable with distribution:

$$f_Z(x) = \begin{cases} \frac{1}{2} & \text{if } x < 1 \\ \frac{3}{2x^4} & \text{if } x \geq 1 \end{cases}$$

Suppose I have 300 independent Z_i each distributed identically, and I add them up. What's the approximate probability that this sum is greater than 310?

Notice that we are asked for an approximate probability in a problem with lots of iid random variables, so we think that the central limit theorem will be useful.

First, let's find the expected value:

$$\begin{aligned} E(Z) &= \int_0^1 \frac{x}{2} dx + \int_4^\infty \frac{3}{2x^3} dx \\ &= \left(\frac{x^2}{4} \right|_0^1 + \left(\frac{-3}{4x^2} \right|_1^\infty \\ &= 1 \end{aligned}$$

Now we also need the distribution of Z^2. Doing a change of variables gives

$$\begin{aligned} f_{Z^2}(x) &= \frac{\mathrm{d}P(Z^2 < x)}{\mathrm{d}x} \\ &= \frac{\mathrm{d}P(Z < \sqrt{x})}{\mathrm{d}x} \\ &= f_Z(\sqrt{x}) \frac{1}{2\sqrt{x}} \end{aligned}$$

Thus

$$f_{Z^2}(x) = \begin{cases} \frac{1}{4\sqrt{x}} & \text{if } x < 1 \\ \frac{3}{4x^{5/2}} & \text{if } x > 1 \end{cases}$$

So

$$\begin{aligned} E(Z^2) &= \int_0^1 \frac{\sqrt{x}}{4} dx + \int_1^\infty \frac{3}{4x^{3/2}} dx \\ &= \left(\frac{2}{3} x^{3/2} \frac{1}{4}\right|_0^1 + \left(\frac{-3}{2\sqrt{x}}\right|_1^\infty \\ &= \frac{1}{6} + \frac{3}{2} \\ &= \frac{5}{3} \end{aligned}$$

Thus $Var(Z) = E(Z^2) - (E(Z))^2 = \frac{2}{3}$. Therefore, we can approximate a sum of 300 of these variables as a $(300, 200)$ normal distribution, so the approximate probability that this sum is greater than 310 is $1 - \Phi(\frac{310-300}{\sqrt{200}}) = 0.24$

What is the probability that this sum is greater than Y, where Y is a $(290, 20)$ normal variable? From this setup, it's clear what do to: use the normal approximation that we have: the difference between $\sum Z_i$ and Y is approximately equal to the difference between a normal $(300, 200)$ variable and a normal $(290, 20)$ variable. By page 363, this is a normal $(10, 220)$ variable. Therefore $P(\sum Z_i - Y > 0) \approx 1 - \Phi(\frac{0-10}{\sqrt{220}}) = 0.75$.

11.2. **Example.** The stock market has crashed three times in the last 50 years (note: I don't know if this is true, and I haven't even defined "crashed", but whatever). What it the probability that it happens tomorrow?

Well, there are $50 \cdot 365 = 18250$ days in 50 years (forget leap years), so let's assume that it happens on a given day with probability $p = 3/18250 = 1.6 \times 10^{-4}$. Then that's the probability that it happens tomorrow.

What's the probability that it crashes some time in the next 5 years? We model this as $5 \cdot 365 = 1825$ trials with $p = 1.6 \times 10^{-4}$. We can thus use the Poisson approximation with $\mu = 1825p = 0.3$. Not a huge surprise, since this is 1/10 of the original time frame. Anyway, this means that

$$P(\text{crash}) = 1 - P(\text{no crahes}) = 1 - e^{-\mu} = 0.259$$

So over a 1/4 chance of a crash.

If you become a quant, make sure to take this into account: the probability of a crash tomorrow is so small as to be negligable. Therefore, you might simply discount it in your model. However, if you use your short-term model over the long run, you will very likely be *very* wrong at some point, which could result in you going bankrupt, economic collapse, etc.