MATH 2B RECITATION 1/26/12

ALDEN WALKER

1. Random Thoughts

Without thinking too hard, which would you prefer?

(1) I give you \$500.
(2) I flip a biased coin and give you \$560 with probability 0.9.

How about this:

(1) I beat you up and take \$500 from you.
(2) I flip a biased coin and leave you alone with probability 0.1 (and beat you up and take \$570 from you otherwise).

The book "Thinking, fast and slow" by Daniel Kahneman describes how most people will prefer option (1) in the first experiment and (2) in the second experiment. Using your newfound probability knowledge, you can compute that both of these choices give a smaller expected value for your money and health. In other words, they are both the wrong choice.

2. Probability Densities (HW 1,2)

So remember the normal distribution? That was an example of a continuous distribution. While there are some things about continuous distributions that may be counter-intuitive (such that the fact that the probability of getting exactly any given particular value is 0) (for most continuous distributions), they are very nice. We extend the definition of mean to probability distributions with density function f as: $E(X) = \int_{-\infty}^{\infty} xf(x)dx$. Then we calculate $Var(X) = E(X^2) - (E(X))^2$ and $SD(X) = \sqrt{Var(X)}$.

Awesome facts:

- Note that all those rules about expectation and stuff still apply (Like $E(XY) = E(X)E(Y)$ for independent random variables X and Y, etc)
- Note that page 248 has a super cool summary of lots of facts (for instance, what is the standard deviation of a sum of identically distributed random variables? It's there!)
- If you're interested in getting the expected value of a function of a random variable, it might be better to use the tricks from page 248, rather than trying to integrate. Sometimes this works.

2.1. The Central Limit Theorem.

Says that if you have n identically distributed random variables (with finite variance), then as $n \to \infty$, the sum (or average) of the n variables converges (in distribution–that is the cumulative distribution function converges pointwise) to the normal distribution.

2.1.1. *Example.*

I have this weird random number generator on my computer, and I have no idea what distribution it's making (pseudo)random numbers from. After some experimentation (or just suppose I know) that the mean of the numbers I'm getting is 10 and the standard deviation is 5. What's the probability that after getting 100 numbers out of my generator, their sum is over 1100?

We can use the central limit theorem—even though we don't know the distribution, we know that the sum of the random numbers will approach a normal distribution. The right thing to do is the scale the problem by the standard deviation and ask about the standard normal distribution. In other words, we want to know $P(X_1 + \cdots + X_{100} \geq 1100)$. The central limit theorem says that that is the same as $P(Y \geq 1100)$ where Y is a random variable from the normal distribution with mean $10 \times 100 = 1000$ and standard deviation $5 \times \sqrt{100} = 50$. This in turn is $1 - \Phi\left(\frac{1100-1000}{50}\right) = 1 - \Phi(2) = 0.02275$. In other words, that's the probability that I'm interested in.

3. Changes of Variable! (HW 3)

Sometimes, you want to know the distribution of $g(X)$ for some random variable X (for which you know the distribution). Well, sometimes you can figure this out!

Theorem 3.1. *Let X be a random variable with density $f_X(x)$ on the range (a, b). Let $Y = g(X)$, where g is strictly increasing or decreasing on (a, b). The range of Y is then an interval $(g(a), g(b))$, and the density of Y on this interval is given by*

$$f_Y(y) = f_X(x)\frac{1}{\left|\frac{dy}{dx}\right|} \text{ where } y = g(x)$$

Of course, that's not much good unless you can get a formula for x in terms of y (ie invert g), since that's the only way you get a nice formula for $f_Y(y)$ without any x's floating around.
Way that might be more useful on the homework: If you look at how that theorem is derived, you will see that actually, you get $f_Y(y) = f_X(x(y))\left|\frac{dx}{dy}\right|$. This version will probably be more useful to you, since note that if you use the theorem version, you figure out the derivative of g, and then substitute in and simplify. It might be easier to invert first and differentiate. Just a thought.

Maybe even a better way of thinking about it is that $f_Y(y) = d(F_Y(y))/dy$, where F_Y is the cumulative distribution function, and then plug in what F_Y is in terms of probability, and then solve. You get the same thing.

3.1. Example.

You give me \$1, and I pick a number randomly between 0 and 1 uniformly, and I square it. If the number is above 3/4, I give you \$8. What is the amount you expect to win every time?

Note that if you didn't square the number, your expected profit is $-3/4 + 7/4 = 1$, so the game is in your favor. However, squaring clearly lowers your chances of winning, so we don't know if you're still expected to win.

Well, let X be the number picked uniformly at random from the interval. You are interested in the expected value of your profit $Y = g(X)$, where $g(x) = \begin{cases} -1 & \text{if } x^2 < 3/4 \\ 7 & \text{otherwise} \end{cases}$ You can use the formula on page 263 to calculate $E(Y)$ as:

$$E(Y) = \int_{-\infty}^{\infty} g(x)f_X(x)dx = \int_0^{\sqrt{3/4}} (-1)1dx + \int_{\sqrt{3/4}}^1 7dx = -\sqrt{3/4} + (1 - \sqrt{3/4})7 \approx 0.072$$

So you still expect to win! Let's look at this a little bit differently. Suppose we want the density function of X^2. Let $h(x) = x^2$, and let $Y = h(X)$, i.e. $X = \sqrt{Y}$. Then by the formula above, we can calculate

$$f_Y(y) = f_X(x(y))\left|\frac{dx}{dy}\right| = f_X(x(y))\frac{1}{2\sqrt{Y}}$$

Between 0 and 1 (and $f_X \equiv 1$). Whoops it's not defined at zero, but we don't care because that's a set of measure zero! We can just go ahead and find the expected value of our profit Z:

$$E(Z) = (-1)P(Y \le 3/4) + 7P(Y > 3/4) = -\int_0^{3/4} \frac{1}{2\sqrt{y}}dy + 7\int_{3/4}^1 \frac{1}{2\sqrt{Y}}dy = -\frac{\sqrt{3}}{2} + 7\left(1 - \frac{\sqrt{3}}{2}\right) \approx 0.072$$

Wasn't that more fun?

Yet another way is to plug in what we want:

$$f_Y(y) = \frac{\mathrm{d}F_Y(y))}{\mathrm{d}y}$$

And $F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(X \le \sqrt{y}) = F_X(\sqrt{y})$, so

$$f_Y(y) = \frac{\mathrm{d}F_x(\sqrt{y})}{\mathrm{d}y} = f_X(\sqrt{y})\frac{\mathrm{d}\sqrt{y}}{\mathrm{d}y} = f_X(\sqrt{y})\frac{1}{2\sqrt{y}}$$

Notice this is the same as we got with the formula, but maybe now we understand it better.

3.2. **Convergence of Integrals.** How do you show that an improper integral diverges? A method that we will accept on the homework is if one of the one-sided integrals diverges, then the integral doesn't exist. I.e. if you want to show that $\int_{-\infty}^{\infty} f(x)dx$ diverges, you can show that $\int_0^{\infty} f(x)dx$ diverges.

4. Continuous Joint Distributions (HW 4)

These are just like regular joint distributions. That is, if f is the joint distribution of X and Y, then

$$P((X,Y) \in B) = \iint_B f(x,y)dxdy$$

Similarly, if you want the marginal density, for example:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

4.1. **Example.** X and Y are uniform random variables on $[0,1]$. What is the probability that $X > Y$?
Let $B = \{(x,y) \,|\, x > y\}$, then the probability we want is

$$\iint_B 1dxdy = \int_0^1 \int_y^1 1dxdy = \int_0^1 (1-y)dy = 1 - 1/2 = 1/2$$

Or it's just the area of the triangle above the diagonal below the line $y = 1$.

5. Independent Normal Variables (HW 5,6)

If you combine two normal variables, you get a joint density function which looks basically like a normal density function, except 3D (p. 357). This is pretty cool. We can also use the nice formulas on p.363 to combine two normal variables into one to answer questions like on your homework.

Don't do double integrals for this question on the homework! Any linear combination of normal variables is normal, so if you want the difference between normal variables, use p. 363 to answer the question using a 1-dimensional normal distribution.

5.1. **Little Note.** The stuff on page 363 is (I believe) all that you need to do the homework, plus some little facts like the distribution of $-X$ takes the negative of the mean and leaves the standard deviation alone. It also tells you what happens if you multiply a $(0,1)$ normal distribution by a factor c (you get a $(0,c^2)$ distribution). But what happens if you multiply any normal distribution by a factor c? We can use the change of variables above: we are interested in the distribution of cX, where X is a (μ,σ^2) normal random variable. Here $y(x) = cx$, so $x(y) = (1/c)y$. Therefore the distribution function is

$$\begin{aligned} f_{cX}(y) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{1}{c}y-\mu)^2/\sigma^2} \frac{1}{c} \\ &= \frac{1}{\sqrt{2\pi}c\sigma} e^{-\frac{1}{2}\frac{1}{c^2}(y-c\mu)^2/\sigma^2} \\ &= \frac{1}{\sqrt{2\pi}c\sigma} e^{-\frac{1}{2}(y-c\mu)^2/(c\sigma)^2} \end{aligned}$$

Which we recognize as a $(c\mu,(c\sigma)^2)$ normal distribution.

5.2. **Another Little Note.** Note that usually when you are using a normal distribution you care about the mean and the standard deviation, but when you write it down as (μ,σ^2), you are writing down the mean and variance.

5.3. **Adding Isn't Multiplying.** Note that when you add n normal random variables, the distribution of their sum has variance multiplied by n, so the standard deviation is multiplied by $\sqrt{n}$, as we saw in the first example. However, if you multiply a single random variable by n, you get variance multiplied by n^2, as we just saw. This sort of makes sense intuitively: it is easier to get a single random variable larger—to get a sum large, you must have all the summands large, which is more difficult. In any case, it's true.

5.4. **Example.** I pick 2 numbers from a standard normal distribution and I find their difference. What is the probability that their difference is bigger than 1?

Ok, page 363 tells you that the distribution of $X+Y$ for standard normal independent X and Y is a normal $(0,2)$ distribution. However, we want $X-Y$. But the distribution of $-Y$ is the same as Y because it's symmetric. Therefore, the distribution of $X-Y$ is a normal $(0,2)$ distribution. Then to find the probability that (the absolute value of) this is bigger than one is just $\Phi(-1/\sqrt{2}) + (1-\Phi(1/\sqrt{2})) \approx 0.4795$.

6. Operations On Stuff

Sometimes, you just have to operate on stuff. For instance, what if we have two random variables X and Y (with densities f_X and f_Y), and we want the density of $X+Y$. We can convolute! I believe this shows up in signal processing and stuff, especially in relation to the Fourier transform, so you may have seen it there. Anyway the formula is what you would write down if you thought about it:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx$$

6.1. **Example.** What's the distribution of the sum of two uniform random variables X and Y on $(0,1)$? We know that $f_X(x) = f_Y(x) = \chi_{[0,1]}$. (If you haven't seen that before, the characteristic function of A, usually denoted χ_A or $\mathbf{1}_A$, is the function which is 1 on A and 0 elsewhere).

Clearly, $f_{X+Y}(z) = 0$ if $z \notin [0,2]$, so we will assume z is in that interval:

$$\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx \\
&= \int_{\max(z-1,0)}^{\min(1,z)} f_X(x) f_Y(z-x)dx \\
&= \int_{\max(z-1,0)}^{\min(1,z)} dx \\
&= \min(1,z) - \max(z-1,0)
\end{aligned}$$

The graph of that is just z from 0 to 1 and $2-z$ from 1 to 2. In this particular case, I was able to shove a bunch of different cases in to the max and min. On your homework , I believe you will need to case out z (i.e. suppose z is less than 1, then here is the answer, etc).