Select Reading Question Responses (9/3/2021)

> On page 120 it explains what a linear combination is. Are X and Y always random? Could they not be a set calculated value?

Linear combinations make sense in a *lot* of different situations. If you've taken (or are planning to take) linear algebra, it might be useful to point out that you can form linear combinations of any two elements of a vector space. The set of random variables is one example of a vector space, so we can form linear combinations of them. But other "non-random things" are also vector spaces (like numbers, or lists of numbers), and you can also take linear combinations of those (as long as both things that you're taking linear combinations of are of the same type).

> What does the k mean in the expected value and general variance formulas?

The k in these formulas represents the number of possible outcomes of the random variable. For example, if your random variable models a coin flip, then $k = 2$, and if your random variable models a standard dice roll, then $k = 6$.

> When learning about integrals in calculus we used a very similar method of calculating areas under a curve. I imagine that you could use concepts from calculus to help compute probabilities from continuous distributions. Is this something that is use in probability? I think it would be pretty cool to try some applications of that if so.

> The book also mentioned how the area under the curve could be used to calculate the probability and I was wondering if that was a practical application of how to use integrals from calculus.

Yes, this is definitely an application of integrals! At the higher levels, courses in probability theory usually begin with extended discussions of integrals. We won't discuss this since calculus is not a prerequisite for this course. We'll just get computers to calculate integrals for us (without using the word "integral"). Note that knowing calculus wouldn't help you avoid this use of computers: the types of integrals that underlie probability calculations for normal or chi-squared or t-distributions are are typically impossible to do by hand.

> I understand that a negative average out come in problem 3.31 means that it is expected to lose money, but what would a negative outcome mean in a scenario that does not involve money or objects?

I suppose it really depends on the situation! For example, if you're thinking about daily low temperature in Celsius as a random variable, a negative outcome (or expected value) would mean that the low temperature is below freezing. Or, if you're thinking about left-pinky-length minus right-pinky-length as a random variable, a negative outcome means that someone's right pinky is longer than their left pinky.

I think in most concrete situations, it should be pretty clear how to interpret negative numbers, even if it's hard to say anything generalizable.

> I was stumped by 3.43(b). I arrived at the mean easily enough, but calculating the standard deviation for the whole week of purchases was perplexing. First, I multiplied the standard deviation for one day (.3354) by seven (since there are seven days in a week), and upon getting \$2.35, I checked my answer in the back of the book. Unfortunately, my response was incorrect. Thus, I went back and checked my work, using the formula $\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$ since variance is just s squared. I squared the standard deviations of the muffin and the cup of coffee given in the question, and plugged them into the equation as $\mathrm{Var}(X)$ and $\mathrm{Var}(Y)$. Yet the answer I received wasn't the correct one; in fact, it was the same answer I got when I multiplied .3354 times 7. What is the correct way to calculate the standard deviation for a whole week?

Let C be the random variable representing the price of a cup of coffee, and M the random variable representing the price of a muffin. We are told that

$$\begin{aligned} E[C] &= 1.40 \quad SD[M] = 0.30 \\ E[M] &= 2.50 \quad SD[M] = 0.15 \end{aligned}$$

We are told that C and M are independent.

For part (a), let D be the random variable representing the amount that Sally spends on breakfast on one day. Since Sally buys exactly one cup of coffee and one muffin every day, we have $D = C + M$. Then applying the formula on linearity of expectated value, we have

$$E[D] = E[C + M] = E[C] + E[M] = 1.40 + 2.50 = 3.90.$$

Also, since C and M are independent, we can apply the formula for variance given on

p. 123 as follows:

$$\begin{aligned}
\mathrm{Var}[D] &= \mathrm{Var}[C + M] \\
&= \mathrm{Var}[C] + \mathrm{Var}[M] \\
&= \mathrm{SD}[C]^2 + \mathrm{SD}[M]^2 \\
&= 0.30^2 + 0.15^2 \\
&= 0.1125 \\
\mathrm{SD}[D] &= \sqrt{\mathrm{Var}[D]} \\
&= \sqrt{0.1125} \\
&\approx 0.3354
\end{aligned}$$

For part (b), let W represent the amount that Sally represents on breakfast in one week. Let $D_1, D_2, \ldots, D_7$ be the amount that Sally spends on breakfast for each day of the week (D_1 on day 1, D_2 on day 2, etc). Then $D_1, D_2, \ldots, D_7$ are all independent and

$$W = D_1 + \cdots + D_7,$$

so

$$\mathrm{Var}[W] = 1^2 \cdot \mathrm{Var}[D_1] + \cdots + 1^2 \cdot \mathrm{Var}[D_7] = 7 \cdot 0.1125 \approx 0.7875$$

so $\mathrm{SD}[W] = \sqrt{0.7875} \approx 0.89$.

> I was very confused on how to compute variance of a linear combination of variables when the coefficient is not 1? What makes the coefficient change?

> I was very confused on how to solve 3.45–3.47 conceptually? Why are you able to square the denominator when removing it from the parentheses for variance?

It's because of the formula on p. 123 for "Variability of linear combinations of random variables." Notice the a^2 and b^2 that appear on the right-hand side of this formula.