Module 1 Discussion Reply

1. Interpreting the expectation and conditional expectation of the simple linear regression model for when x is either fixed or random.
    a. I had a similar approach as what you did. I also first checked the math to see what the result would be. That is, I plugged in the formula for y into the expectation to show mathematically what the result would be. I also made sure to note that the expectation of the error term is also 0. You say that the $E(y)$ is the expected result of our simple linear regression model. I think that this is an interesting question. On the surface, we seem to just be taking the expectation of some function. However, to me it's sort of like asking about the average value for y along the vertical axis. This is kind of a strange concept, given that the context of our problem is some linear format where y is typically increasing along the values of x.

    I do think it's correct to say that it's the expected value of the linear regression model, but I am a bit unsure about some more contextual meaning. I think an interesting idea is to use a simpler model such as $y = \beta x$, where the expectation of it is βx. It seems more obvious then that the expected value of y is just the same formula itself. I suppose then that we must be more aware then of factors such as the error term and the role that they play in linear regression.

    Regarding the conditional expectation, I also agree with your statement. It is the expected value of the model, conditioned on a single point x. I think that there's an interesting graphical representation of this, however. I mention in my discussion post how in the textbook (and also in the lectures) there is a figure showing the probability density of y for a given point x which lies "sideways" on the true regression line. I think seeing this is an interesting way to understand how linear regression combines elements of statistics into its formulation.

    I do agree with your interpretation that the difference between the two $E(y)$ and $E(y|x)$ is that the former is the expectation that considers all x's while the latter is conditioned on a single x. I think this distinction is well put and better than how I tried to describe it.

    b. I had a similar reply to the first part of problem 1b also. I similarly plugged in the formula into the expectation to get the resulting $\beta_0 + \beta_1 E(x)$. I am somewhat unsure though, since I feel like this representation looks off for some reason, but this could simply be my lack of understanding regarding linear regression. My interpretation however is similar in that I also think that it's a linear function of the β terms along with $E(x)$.

    I noticed that the question didn't give some actual point x_0 as a realization of the random variable x, so I just added it myself since that is how I think conditional expectations are usually understood. That is, it's the expectation of one random variable, given that another is known to be a specific value. Based off this, I just continued with the formula to show some conditional expectation w.r.t. x_0. I am not sure though that this is the correct methodology for answering this problem.

In the end of chapter 2 on p. 54, it shows that when x and y are bivariate normal random variables, the conditional expectation is the same as in part a. However, this is not explicitly stated in the question, so I don't think we can make that same assumption. However, my interpretation is similar

I also agree with what you think is the difference between the two. I am not sure though if there is some other more elaborate explanation in the case when x is also a random variable. I can't think of one directly at the moment, nor did I notice one from the textbook.

2. In my response to this question, I also had a similar reply. I mentioned how Δ is some change in the value of x, and that applying it would lead to some $\beta_1 * \Delta$ change in the overall formula. I used a bit of simple math to show that this is the case, when we change x to $(x + \Delta)$, the β_1 term distributes and so we are left with this additional $\beta_1\Delta$ term in the resulting equation. I also made sure to note that the change would depend on whether the slope is positive or negative. Another thing is that I sort of assumed that Δ is some positive value also. Something that I just thought of also is if β_1 is nonzero. If it is zero, then it would result in no additional change.
3. In this question I agree with your response. I mentioned also that this is not the case, unless the data truly is a straight-line relationship, which I think is atypical and wouldn't be seen too often in the real-world. I like how you specify also that it is a model of the data, and so we are basically approximating the linear function, which itself is not perfect and is not designed to be. This reminds me of the saying in statistics that all models are false, but some are useful.