ARE 210: Sep. 28, 2022 Section Notes

Joel Ferguson

1 Conditional Expectation Calculation Example

Setup: X and Y are independent random variables that are both distributed uniform[0,1]($X,\ Y \sim U[0,1],\ X \perp Y$). Find $\mathbb{E}[X|X > Y]$.

Solution:
Strategy: To take a conditional expectation we need the conditional PDF, which can (generally) be found by deriving the conditional CDF. So the solution proceeds in three steps. First calculate the conditional CDF, then derive the conditional PDF, and finally integrate the function of interest (here the identity function) with respect to the conditional PDF to get the expectation.

Step 1: Find the Conditional CDF
We want to find $\mathbb{P}(X \leq x|X > Y)$, which in this case (as in most cases) can be found via Bayes' rule since $\mathbb{P}(X > Y)$ is greater than 0.

$$\mathbb{P}(X \leq x|X > Y) = \frac{\mathbb{P}(X \leq x, X > Y)}{\mathbb{P}(X > Y)} = \frac{\mathbb{P}(Y < X \leq x)}{\mathbb{P}(X > Y)}$$

We start with the denominator. Since X and Y are both $U[0,1]$ and independent, we know the joint PDF is $f_{X,Y}(x,y) = dxdy$. For a given value of X, say x, the probability that $Y < X$ is

$$\mathbb{P}(Y < x, X = x) = \int_0^x dxdy$$

and so the probability that Y is less than X is that probability, averaged over the possible values of x. So we have

$$\mathbb{P}(Y < X) = \int_0^1 \mathbb{P}(Y < x, X = x) = \int_0^1 \int_0^x dydx = \int_0^1 xdx = \frac{x^2}{2}\Big|_0^1 = 1/2$$

Calculating the numerator is similar, except instead of integrating over the entire support of X, we integrate up to the value at which we want to evaluate the conditional CDF, which will now be referred to as z to avoid confusion with the limit of integration for Y

$$\mathbb{P}(Y < X \leq z) = \int_0^z \int_0^x dydx = \int_0^z xdx = \frac{x^2}{2}\Big|_0^z = z^2/2$$

Combining these two results, we get

$$F_{X|X<Y}(z) = \mathbb{P}(X \leq z | X > Y) = \frac{\mathbb{P}(Y < X \leq z)}{\mathbb{P}(X > Y)} = \frac{z^2/2}{1/2} = z^2$$

Step 2: Derive the Conditional PDF

To derive the conditonal PDF we simply take the derivative of the conditional CDF

$$f_{X|X<Y}(x) = \frac{\partial}{\partial x} F_{X|X<Y}(x) = \frac{\partial}{\partial x} x^2 = 2x\ dx$$

Step 3: Take the Conditional Expectation

Now that we have the conditonal PDF, finding the conditonal expectation is simply a matter of integrating the function of interest, here the identity function $x = x$ against the conditional PDF.

$$\mathbb{E}[X|X > Y] = \int_0^1 x f_{X|X<Y}(x) = \int_0^1 x * 2x dx = \int_0^1 2x^2\ dx = \frac{2}{3} x^3 \Big|_0^1 = 2/3$$

2 Enumerating Condional Expectations as Random Variables in Small Event Spaces

Setup: We have three coins, one is fair (it gives heads and tails with equal probability), two have heads on both sides (and thus give heads with probability 1). We draw 2 coins randomly and toss them. The random variable X is defined as the number of heads after the coin toss and the random variable Y is the number of fair coins that were tossed. Find $\mathbb{E}[X|Y]$

Solution:

Strategy: The first thing to note with this problem is that we are conditioning on Y, not $Y = y$, so the answer is a random variable rather than a number (more on this in the next question). Because there are so few possible events (we can get both unfair coins and two heads, one fair coin and two heads, or one fair coin and one head), we can easily enumerate them and associate them with probabilities given what we know about the data generating process.

Step 1: Possible values of Y

First we need to enumerate the possible values that our conditioning variable can take. In this case, there are only two possibilities: we can get both unfair coins, in which case $Y = 0$, or we can get one fair coin and one unfair coin, in which case $Y = 1$.

Step 2: Calculate $f_{X|Y=y}$

For $Y = 0$ this is easy, the only possible outcome is two heads, so

$$f_{X|Y=0}(x) = \begin{cases} 1 & x = 2 \\ 0 & otherwise \end{cases}$$

For $Y = 1$ we should always get at least one head from the unfair coin and one additional head half the time from the fair coin

$$f_{X|Y=1}(x) = \begin{cases} 1/2 & x = 1 \\ 1/2 & x = 2 \\ 0 & otherwise \end{cases}$$

If we want to be particularly thorough, we can actually start by finding the joint PDF and then using it to calculate the conditional PDF, but with such a small event space there's little value in doing so. Because the coin flips are independent of Y[1], it is easier to find $f_{X|Y}(x)$ and $f_Y(y)$ then calculate $f_{X,Y}(x,y)$ than it is to start from $f_{X,Y}(x,y)$. To illustrate this point, note that[2]

$$f_Y(y) = \begin{cases} 1/3 & y = 0 \\ 2/3 & y = 1 \\ 0 & otherwise \end{cases}$$

and using the conditional pdf we intuitively derived above, we get

$$f_{X,Y}(x,y) = \begin{cases} 1/3 & y = 0; x = 2 \\ 1/3 & y = 1; x = 1 \\ 1/3 & y = 1; x = 2 \\ 0 & otherwise \end{cases}$$

which we can verify is correct because it integrates to 1.

Step 3: Calulate Conditional Expectations Finally, we can use the conditional PDFs derived above to calculate the conditional expectation as a function of y. For $Y = 0$, the expectation is simply 2, since that value is realized with probability 1. For $Y = 1$, we integrate

$$\mathbb{E}[X|Y=1] = \sum_{x=1}^{2} x f_{X|Y=1}(x) = 1 * \frac{1}{2} + 2 * \frac{1}{2} = 3/2$$

So we get

$$\mathbb{E}[X|Y](y) = \begin{cases} 2 & y = 0 \\ 3/2 & y = 1 \end{cases}$$

[1] nb: X is not independent of Y

[2] If coins A and B are unfair and coin C is fair our possible pairs are AB, AC, and BC, two-thirds of which have the fair coin

3 The General Definition of Conditional Expectation Explained Through a Small Range Space

Setup: We have a probability space $(\Omega, \mathcal{F}, P)$ where $\Omega = [0,1]$,$\mathcal{F}$ is the Borel σ-algebra, and P is the Lebesgue measure. We define three random variables:

$$\begin{aligned} X(\omega) =& \mathbb{I}\{\omega \in [0, 1/2)\} \\ Y(\omega) =& \mathbb{I}\{\omega \in [0, 3/4)\} \\ Z(\omega) =& \mathbb{I}\{\omega \in [1/4, 3/4)\} \end{aligned}$$

Check that $X \perp Z$, show $\mathbb{E}[X|Y] = \frac{2}{3}\mathbb{I}\{\omega \in [0, 3/4)\}$, and calculate $\mathbb{E}[X|Y, Z]$.

Solution:
Strategy: I really like this problem because it's easy enough to solve intuitively but can really help explain some of the nuances of the general definition of conditional expectation favored in this course. In particular, it highlights that conditional expectations (in this general sense) are random variables and helps us understand what it means for e.g. $\mathbb{E}[X|Y]$ to "integrate like X over sets in $\sigma(Y)$." Here, $\sigma(Y)$ is small enough for us to easily figure out what has to be multiplied by the measure of each set to equal to integral of X over that set. It's also relatively straightforward to calculate the conditional PDF via Bayes' rule, but I find the strategy just proposed more instructive (and to some extent, they are the same).

Step 1: Check $X \perp Z$
This is relatively straightforward. We know that $X \perp Z \iff \mathbb{P}(X \wedge Z) = \mathbb{P}(X)\mathbb{P}(Z)$[3]. We can use the definitions of X and Z to find

$$X = 1 \wedge Z = 1 \Rightarrow \omega \in [0, 1/2) \wedge \omega \in [1/4, 3/4) \Rightarrow \omega \in [1/4, 1/2)$$

Thus,

$$\mathbb{P}(X = 1 \wedge Z = 1) = 1/2 - 1/4 = 1/4 = 1/2 * 1/2 = \mathbb{P}(X = 1)\mathbb{P}(Z = 1)$$

and so $X \perp Z$.

Step 2: Calculate $\mathbb{E}[X|Y]$
Our goal is to find a function $g(\omega)$ such that g is Y-measurable and $\int_A g(\omega)dP\omega = \int_A X(\omega)dP\omega$ for all $A \in \sigma(Y)$. We can explicitly accomplish this by first noting that $\sigma(Y)$ only contains four elements

[3]Here I'll only show the result for $X = 1$ and $Z = 1$. Note that this is sufficient because $\mathbb{P}(X = 0) = 1 - \mathbb{P}(X = 1)$ and $\mathbb{P}(X = 1 \wedge Z = 1) + \mathbb{P}(X = 0 \wedge Z = 1) = \mathbb{P}(Z = 1)$, so $\mathbb{P}(X = 1 \wedge Z = 1) = \mathbb{P}(X = 1)\mathbb{P}(Z = 1)$ implies $\mathbb{P}(X = 1) + \mathbb{P}(X = 0 \wedge Z = 1)/\mathbb{P}(Z = 1) = 1$, which can only be true if $\mathbb{P}(X = 0 \wedge Z = 1) = \mathbb{P}(X = 0)\mathbb{P}(Z = 1)$. Similar results hold for $Z = 0$.

$$\sigma(Y) = \{\emptyset, [0, 3/4), [3/4, 1], \Omega\}$$

So we can write our objective as finding a function that satisies

$$g \in \sigma(Y)$$
$$\int_{\emptyset} g(\omega)dP(\omega) = \int_{\emptyset} X(\omega)dP(\omega)$$
$$\int_{[0,3/4)} g(\omega)dP(\omega) = \int_{[0,3/4)} X(\omega)dP(\omega)$$
$$\int_{[3/4,1]} g(\omega)dP(\omega) = \int_{[3/4,1]} X(\omega)dP(\omega)$$
$$\int_{\Omega} g(\omega)dP(\omega) = \int_{\Omega} X(\omega)dP(\omega)$$

The first condition implies that g can only take two unique values, one on $[0, 3/4)$ and one on $[3/4, 1]$. Together with the other conditions, this gives us a strategy for solving for g, we simply need to find values of the integrals of X on those two sets and set g equal to those values divided by the measures of the respective sets[4].

$$g(\omega) = \begin{cases} \frac{\int_{[0,3/4)} X(\nu)dP(\nu)}{\int_{[0,3/4)} dP(\nu)} & \omega \in [0, 3/4) \\ \frac{\int_{[3/4,1]} X(\nu)dP(\nu)}{\int_{[3/4,1]} dP(\nu)} & \omega \in [3/4, 1] \end{cases}$$

Let's start with the numerators. Since $X(\omega) = \mathbb{I}\{\omega \in [0, 1/2)\}$ we have

$$\int_{[0,3/4)} X(\nu)dP(\nu) = 1 * P([0, 1/2)) + 0 * P([1/2, 3/4)) = 1/2$$
$$\int_{[3/4,1]} X(\nu)dP(\nu) = 0 * P([2/4, 1]) = 0$$

The denominators are simply the lengths of the intervals

$$\int_{[0,3/4)} dP(\nu) = 3/4$$
$$\int_{[3/4,1]} dP(\nu) = 1/4$$

So we get

$$g(\omega) = \begin{cases} \frac{1/2}{3/4} = 2/3 & \omega \in [0, 3/4) \\ \frac{0}{1/4} = 0 & \omega \in [3/4, 1] \end{cases}$$

[4]Note that this looks a lot like Bayes' rule. The numerators simplify to $\mathbb{P}(X \wedge Y = y)$ and the denominators to $\mathbb{P}(Y = y)$

Or more compactly, $g(\omega) = 2/3\mathbb{I}\{\omega \in [0, 3/4)\}$.

Step 3: Calculate $\mathbb{E}[X|Y, Z]$
The first step here is to calculate the values that (Y, Z) can take

$$(Y, Z) = \begin{cases} (1, 0) & \omega \in [0, 1/4) \\ (1, 1) & \omega \in [1/4, 3/4) \\ (0, 0) & \omega \in [3/4, 1] \end{cases}$$

So $g(\omega) = \mathbb{E}[X|Y, Z](\omega)$ can take on three unique values (since it has to be (Y, Z)-measureable), one on $[0, 1/4)$, a second on $[1/4, 3/4)$, and a third on $[3/4, 1]$. We can use the exact same strategy as above to find

$$g(\omega) = \begin{cases} \frac{\int_{[0,1/4)} X(\nu)dP(\nu)}{\int_{[0,1/4)} dP(\nu)} & \omega \in [0, 1/4) \\ \frac{\int_{[1/4,3/4)} X(\nu)dP(\nu)}{\int_{[1/4,3/4)} dP(\nu)} & \omega \in [1/4, 3/4) \\ \frac{\int_{[3/4,1]} X(\nu)dP(\nu)}{\int_{[3/4,1]} dP(\nu)} & \omega \in [3/4, 1] \end{cases}$$

Which we can calculate as above to find

$$g(\omega) = \begin{cases} 1 & \omega \in [0, 1/4) \\ 1/2 & \omega \in [1/4, 3/4) \\ 0 & \omega \in [3/4, 1] \end{cases}$$

This result is interesting because $X \perp Z$. Since it is not the case that $X \perp (Y, Z)$, we can learn more about the value of X by incorporating our knowledge of the value of Z than if we only used information on the value of Y.

4 Characteristic Function of Normal and Taking Expectations with it

Setup: X is a standard normal random variable ($X \sim N(0, 1)$). Find $\Psi_X(t)$ and calculate $\mathbb{E}[X^3]$.

Solution:
Strategy: Deriving the characteristic function is mostly about finding the PDF of a shifted normal variable after setting up the integration problem. Calculating the third moment is relatively easy once we have the characteristic function, it just requires some careful application of the product rule.

Step 1: Derive $\Psi_X(t)$
We can start this problem by writing down the definition of the characteristic function

$$\Psi_X(t) = \mathbb{E}[exp(itx)] = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

We can then combine the exponential terms into $e^{itx - x^2/2}$ and note that

$$itx - x^2/2 = \frac{-(x^2 - 2itx + (it)^2)}{2} - t^2/2 = \frac{-(x-it)^2}{2} - t^2/2$$

since $(it)^2 = -t^2$. Thus, we can pull $e^{-t^2/2}$ outside of the integral to get

$$\Psi_X(t) = \mathbb{E}[exp(itx)] = e^{-t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-it)^2}{2}} dx$$

Finally, we can use a change of variable to show that the integral is equal to 1. Let $y = x - it$, so $dy = dx$. We can rewrite the integral as

$$e^{-t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy = e^{-t^2/2}$$

since the integral is now of the PDF of a standard normal.

Step 2: Calculate $\mathbb{E}[X^3]$

We can use the property of the characteristic function that $\Psi^{(3)}(0) = i^3\mathbb{E}[X^3] = -i\mathbb{E}[X^3]$ to solve this problem. Really, this amounts to careful application of the chain and product rules

$$\begin{aligned}
\Psi'(t) &= -te^{-t^2/2} \\
\Psi''(t) &= -e^{-t^2/2} + t^2e^{-t^2/2} \\
\Psi^{(3)}(t) &= te^{-t^2/2} + 2te^{-t^2/2} - t^3e^{-t^2/2}
\end{aligned}$$

So $\Psi^{(3)}(0) = 0$ and thus $\mathbb{E}[X^3] = 0$.

5 Deriving the Characteristic Function of a Transformation of IID Random Variables

Setup: We have a sequence of n iid random variables $\{X_i\}_{i=1}^n$ with mean μ ($\mathbb{E}[X_i] = \mu$) and variance σ^2 ($\mathbb{V}[X_i] = \sigma^2$). We also define the sample mean random variable $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. We want to find the characteristic function of a random variable Y defined as

$$Y = \frac{\bar{X}_n - \mathbb{E}[\bar{X}_n]}{\sqrt{\mathbb{V}[\bar{X}_n]}}$$

Solution:

Strategy: The key to this problem is using a lot of features of independent random variables. In particular, we use the fact that the expectation (variance) of a sum of independent random variables is the sum of the expectations (variances) to find some of the constants in Y and then that the characteristic function of a sum of independent variables is the product of their characteristic functions to find the last piece.

Step 1: Find $\mathbb{E}[\bar{X}_n]$
Since $\{X_i\}_{i=1}^n$ are iid, we know that the expectation of their sum is the sum of their expectations, which are all μ. This combined with the linearity of the expectation operator gives us

$$\mathbb{E}[\bar{X}_n] = \mathbb{E}[\frac{1}{n}\sum_{i=1}^n X_i] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \frac{1}{n}\sum_{i=1}^n \mu = \mu$$

Step 2: Find $\sqrt{\mathbb{V}[\bar{X}_n]}$
Since the variance of a sum of independent random variables is the sum of their variances, we similarly have

$$\mathbb{V}[\bar{X}_n] = \mathbb{V}[\frac{1}{n}\sum_{i=1}^n X_i] = \frac{1}{n^2}\mathbb{V}[\sum_{i=1}^n X_i] = \frac{1}{n^2}\sum_{i=1}^n \mathbb{V}[X_i] = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}$$

So we have $\sqrt{\mathbb{V}[\bar{X}_n]} = \sigma/\sqrt{n}$.

Step 3: Find $\Psi_{\sum_{i=1}^n X_i}(t)$
Since the characteristic function of a sum of independent random variables is the product of their characteristic functions and all of the X_is have the same characteristic function we get

$$\Psi_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n \Psi_{X_i}(t) = \Psi_{X_1}^n(t)$$

Step 4: Put it all together We now use the fact that $\Psi_{aA+b}(t) = exp(itb)\Psi_A(at)$ to find Ψ_Y. In particular, $a = \sqrt{n}/\sigma$, $b = \mu\sqrt{n}/\sigma$, and $A = \sum_{i=1}^n X_i$.

$$\Psi_Y(t) = \Psi_{\frac{\sqrt{n}}{\sigma}\sum_{i=1}^n X_i - \frac{\sqrt{n}\mu}{\sigma}}(t) = e^{-it\frac{\sqrt{n}\mu}{\sigma}}\Psi_{\sum_{i=1}^n X_i}(\frac{\sqrt{n}}{\sigma}t) = e^{-it\frac{\sqrt{n}\mu}{\sigma}}\Psi_{X_1}^n(\frac{\sqrt{n}}{\sigma}t)$$