CourseKata Video Transcript

Video Details

Video Title: Why Do We Need Those Greek Letters?
Video Link: https://player.vimeo.com/video/379319558

Video Transcript

Student (off screen)
Dr. Ji, what about all these Greek letters? Why do we need those?

Dr. Ji
Yeah, it's as if you weren't confused enough with all these different notations, we thought we'd add in a few more. Actually, it's because we really want you to keep a conceptual difference in mind. Right now, we've mostly been talking about the sample

[SPEAKER GESTURES TO HISTOGRAM OF THUMB LENGTHS WITH BLUE VERTICAL MEAN LINE]

up through this page in this chapter. And that's why we've really focused on understanding this data equals model, the b sub 0, plus error, right?

[WRITES "$Y_i = b_0 + e_i$" ON SCREEN]

We were really focused on that notation. But now, we really want you to remember, there is this DGP that we're trying to understand. And it's not this [POINTS TO EQUATION AND HISTOGRAM]. There's a mathematical notation for that too to help us remember. That's a different thing, and that's the really important thing we want to know about. So our notation for the DGP

[WRITES "DGP" WITH AN ARROW POINTING TO EQUATION]

is that, if we're looking for the mean of the DGP, we wouldn't want to call it b sub 0 nor do we want to call it y bar. We want to call it the mean of the DGP. How do we say that, right? We could say just thumb equals mean, in this case, of the DGP plus error, right?

[DRAWS AN ARROW FROM "DGP' POINTING TO THE RIGHT AND WRITES "Thumb = mean + error"]

But we're also gonna add in some notation to help us remember, that's what we mean, the mean of the DGP. We could call it mu [POINTS TO "mean"], a good fancy Greek letter that kind of looks like an m. But we want you to actually see that there is this model

[POINTS TO "$Y_i = b_0 + e_i$"]

that we're using to estimate this model.

[POINTS TO "Thumb = mean + error"]
So we're gonna use the general linear model notation which looks like this, y sub i equals beta sub 0, plus epsilon sub i,

[WRITES “$Y_i = \beta_0 + \varepsilon_i$” UNDERNEATH “Thumb = mean + error”]

to stand for the mean of the population [POINTS TO “β_0”] plus error. Now this beta sub 0 and this b sub 0, they won't always be the mean. Later we're gonna do different models where these things [POINTS TO “β_0” AND “b_0”] are not gonna be the mean. And so that's why we want you to keep in mind this thing

[POINTS TO “$Y_i = b_0 + e_i$”]

is about the sample, and this thing is about the DGP.

[POINTS TO “$Y_i = \beta_0 + \varepsilon_i$”]

But it won't always be the mean. And that's why we're not using y bar and mu a lot instead we're using b sub 0 and beta sub 0. That's why it's called the general linear model. It's more general than the mean. It could be other things. Now, if I had to make a guess about what this thing is,

[POINTS TO “β_0”]

I would say this is a really hard thing because this beta sub 0 is an unknown parameter. Sure, there's a mean thumb length in the world, in the population of people who have thumbs, right? But I don't know what it is. It's unknown. What would be our best guess? One way we could estimate that is by using something we do have, our data. So we know that the best-fitting parameter estimate is actually 60.1.

[WRITES “60.1” WITH AN ARROW POINTING TO “b_0”]

It's the one that balances the residuals, right? So maybe we would want to use that

[POINTS TO “b_0”]

to estimate this.

[POINTS TO “β_0”]

But is it actually 60.1 millimeters in the population? Probably not. And how do I know that? I could just imagine having taken a slightly different sample,

[A SECOND HISTOGRAM APPEARS ON SCREEN; IT SHOWS THE DISTRIBUTION OF sim.Thumb with “*$\bar{Y}$ = 62”]*

like for instance, let's say this [POINTS TO sim.Thumb HISTOGRAM] was our sample. The mean here is 62, right? If that had been our sample, maybe I would have thought my best guess is that beta sub 0 is 62. But it didn't have to be 62 either. It could have been 58.

[A THIRD HISTOGRAM APPEARS ON SCREEN; IT SHOWS THE DISTRIBUTION OF sim.Thumb with “*$\bar{Y}$ = 58”]*

In that case, I would have said, my best guess is that beta sub 0 is 58. By having a different Greek letter up here,

[POINTS TO “$Y_i = \beta_0 + \varepsilon_i$”]

we could keep in mind, even though this [POINTS TO “b_0”] is my best guess of that [POINTS TO “β_0”], I know that it's probably not that. It's probably something else. And I want to keep in mind this is an unknown thing [POINTS TO “β_0”] that I want to know about. And that's what we really need to keep these two notations separate. So whenever you see this [POINTS TO “β_0”], I want you to think, this is that unknown parameter that I want to know about. And whenever you see this [POINTS TO “b_0”], I want you to think, that's my estimate of that [POINTS TO “β_0”]. That's the best I could do with the data I have.