Transcriber's Name Priya Prashanth

Applied Econometrics

Prof. Tutan Ahmed

Vinod Gupta School of Management

Indian Institute of Technology-Kharagpur

Lecture No. #09

Posterior Probability

Hello and welcome back to the lecture on Applied Econometrics. So we have been talking about probability and till now we have been talking about different concepts of prior probability like frequentist approach or classical approach. Now we will make a shift in the gear, and we are going to talk about posterior probability. And this we will see how that is significantly different from what we have learned.

And posterior probability uh by virtue of uh its advancement in the probability theory is actually something that is pretty widely used in today's world, be it in the world of physics, chemistry, or any other science. Be it in the world of machine learning, be it in the world of economics, be it in the world of trade. It has a huge, huge implication.

And that is why uh the particular example of posterior probability, which you are going to see is the Bayesian theorem. And we will see uh Bayesian theorem is perhaps one of the most most significant uh theorem that we will learn in this course or you know in in in statistics for that matter. Unfortunately, in this course though, we will talk about Bayes theorem in this lecture.

But you know the for other, you know other other topics, we will essentially, you know uh use frequentist approach. The reason being is that the Bayes theorem though, it is well known and we understand the significance of Bayes theorem vis-à-vis the other approaches of probability. This field is still in the, you know still advancing itself, it is still emerging.

So we will not be able to really use all the possible implication of Bayes theorem in all the different, you know topics of econometrics. But we we sort of give due importance to this topic in this lecture. So with this let me start, what do I understand what, what do I mean by posterior probability. Now posterior probability is something where you use, you know the evidences from the random trial.

So if suppose you are having, you know a random experiment, and you are, you know getting some probability value in one of the trial, and in the next trial, what you do is you use the result of the previous trial, you update the probability that way. So every trial you run, you take the probability of the previous trial, and you update the probability of this trial.

So as you keep on you know running multiple trials, your probability keeps on updated and it is absolutely consulting or it is absolutely taking into account the probability that you got in your previous trials, right? So that is the beauty of it. And the why is it so sort of popular or why it is so relevant or useful? The reason is that you are in one try, you are not restricted in one trial.

So in one trial, you can get a probability. But you know if you get something repeatedly, your chances of being correct is much higher, your chances increase in terms of being corrected being correct in your uh, in your sort of conclusion. So we also say it is a way to continuously update our beliefs, okay.

So we know certain thing, but the more and more experiment we run, the more and more probability values we get, what we get is actually we get a chance to update our belief about certain things. Now how it is a diff, how it is a departure or how it is similar with the, you know different concepts you have learned in prior probability. So let us say in prior probability, we have learned something called classical approach.

So we are actually going to use all the different, you know the probability rules that you have derived in in uh, you know like when you are talking about this prior probability concepts, we are also going to use the historical facts. It is not that when I say posterior probability we completely depart from you know the previous concepts of probability, it is not.

We actually are going to use all the concepts that we learned. For example, we will, we will be using the, you know the, you know the sort of priors from the historical facts, which is essentially a frequentist approach, okay. So we will see with some examples.

But before I give you example, let me also illustrate uh, let me actually tell you, you know how the Bayes theorem came and then we will be able to understand the significance of Bayes theorem and then I will give you some examples where the Bayes theorem is used. So that way we will have some understanding about the Bayes theorem before we actually jump into solving a problem. (refer time: 04:49)

So Bayes theorem, let me write down Bayes theorem and I must pay my tribute to the person Reverend Thomas Bayes, okay. So he was a minister of a charge. And you know like how this theory actually come into place. I think it is always very intriguing to understand how certain things come into place and how that thing has become so popular. How that thing has become, you know accepted by everyone, okay.

So it is story of almost, you know like many, many years back, I think back in 17, you know like, mid 18th century uh when the So reverend Thomas Bayes is a minister of a church, as I said. So there was always a debate about resurrection of Christ, right? If if you can, you know if you are going to get back, we are actually going to see, Christ to come into life, okay.

So that was the topic and there was an essay by one, you know that times renowned scholar Hume. So he kind of said that, you know once a man is dead he is dead, so there is no chance of coming back. And so then, you know there is a question that, is there a chance that, you know I mean, how do I say, how do I say with certainty that a man once a man is dead, he is not going to come back, right?

So for Hume, it was more like a evidence, he saw so many evidence that nobody came back. So then he he concluded that there is no chance that Chris Christ is going to come back. Now for Bayes he kind of took a mathematical angle, and he tried to understand alright, so how much evidence do I need to claim to say that well, there is a chance that Chris Christ will come back, right?

So it is basically, not on a very serious note. He kind of derived this formulae okay, to kind of calculate that, okay? How much evidence do I need to be certain about something, okay? And that is the crux of the whole thing that how given certain amount of evidence, what can you say about it, right? And that is, that is, that is all about it. And we will see the examples of Bayes theorem, and we will understand why it is so important.

So let me give you one example, one practical example. And that is actually a very recent example. And that is a, that is actually a, you know happened in the one of the flight crash in Malaysian Airlines, that flight crash in 2014. Airline 17, I think is the name. So when that flight crashed somewhere in Ukraine, so what happened is that all the people on board died, okay?

And there is nothing remained almost after the crash. And it was very difficult to identify who is who, and you know like, so what happened is like, but you know you always want to get your you know relatives and give them a proper sort of funeral, right? So then all the family members of different, you know victims, so they kind of tried to identify them. So then there was a machine called Bonaparte.

And this example, I am taking from the book, uh the Book of Why by Judea Pearl. So what happened is that the the machine Bonaparte, it kind of, you know like none of the victims had their DNA actually preserved right, because they are not criminals. So so what happened is that the DNAs of the family members near and dear ones, where the chances of DNA matching is much higher, so their DNA was taken.

And this using this Bonaparte machine, so what it did was that it continually sort of matched the DNA of the family members with the DNA of those victims, okay. And based on this repeated trials, based on repeatedly confirming if someone you know is matching the DNA of certain family member, so they are able to confirm the who exactly is that victim, which family does he belong?

And that way, I think out of 298 it detected I think around 295 and 296 victims. So that was a very, very powerful result that we obtained using Bayes theorem. There are many other examples where Bayes theorem is actually really really very helpful. (refer time: 09:09)

So for example, today we always talk about this AI ML, right. And what happens in this AI and ML algorithm is that we always try to the AI and ML algorithm tries to learn from data right. So when it tries to learn from data, it needs to confirm, it needs to confirm with certainity, you know with repeated trials, how much it is learned. And that is, you know a new, you know sort of area within this ML literature is that Bayesian network.

So that in Bayesian network, it kind of repeatedly confirms the with certainty about its learning. So that has a huge you know implication. There are many other implications of Bayes theorem. I mean, I can keep telling, you know keep on telling. But some other example, for example, in healthcare and we are going to see some examples of how Bayes theorem is used in healthcare.

Say for example, we have like, you know like, you have, you have you are detected with some disease, right? Now when you are detected with disease, how much certain you are that you actually had the disease, right. And if you do not know that, and if you if you are not sure about that, you can actually run the test again, and then you will see the drastic change in probability of whether you actually had disease or not.

And we will actually do an example to do that. For example, if you if you are tested with COVID or you are tested with maybe cancer or anything very bad. So you want to ensure that whether you had it. So you you know if you if you see that such such kind of result, you might freak out.

But then if you run second test, you know like, depending on whether you come out, you know come out as positive or negative, your probability value will drastically change and we will see that the you know in your example, I mean, there are there are n number of examples in different fields and this is how it is used basically. You understood the principle.

We will do another example, where we are going to I will see that say you know say detection of diamond or detection of you know any any kind of detection exercise, so where you kind of see in the first trial, you kind of identify something as a diamond and in the second trial if it is confirmed as diamond again, so then it is a diamond, right? But if it is not then then then the probability value will drastically decrease.

So essentially, it gives an objective way for you to understand whether you can say certain things with certainty, right? So if I can claim something to be diamond, if I can claim I have COVID and you know I kind of keep on updating my belief with repeated trials until with certainly, if, if that is the case. So that is basically the theory part that you need to learn about Bayes theorem. And with this we are going to do one problem. (refer time: 11:56)

Let us say, let us say uh, you are detected with COVID. okay, with COVID. And we have couple of examples, a couple of information in our hand. As I said at the beginning that I need to have a prior and the prior means that I have information about the distribution of COVID in my population. So let us say that one person in every 1000 person has COVID. So this information I already have, okay and I call it a prior.

Let me use a different color. I already call it a prior, okay? One person in every 1000 person has COVID. So this is my prior. And I also have an information test accuracy. And this is where we see how this test, little bit of test accuracy matters so much. So let us say the test accuracy is 99%. So in 99 cases uh you actually see uh you actually get the right result.

So if I do not have COVID, so in 99 cases it will say 99% cases it will say that I do not have COVID. On the other hand, if I have COVID, in 99% cases, it will say I have COVID, okay? But in 1% cases, it will give me the contrary result, okay. So let us see how that 1% thing actually matters. So now you actually have gone to test yourself and you actually, this information you already have.

You actually have gone to test yourself and you have found yourself to be COVID positive. So you actually have COVID based on the result. Now what will happen if uh you know how certain you should be about your disease, okay? Okay. So that is that is the problem we will try to address. And let us actually see how much probability actually I have of actually getting COVID. (refer time: 14:01)

So let us say the first case, I will so there are different, you know some different graphical approach to solve this kind of problems. And I will share with you a couple of these approaches. So let us say in my population as I said, say in a population of 1000, 1 person has COVID. So let us say this is the population with COVID, okay. So let us, sorry no COVID.

This is a no COVID population and they are 999 out of 1000. So I will just you know write down the number of people considering the number of people with or without COVID considering the population of 1000. And say this is the COVID population, right? Now I have tested, I have tested myself. I can either be from COVID or I can be from no COVID, right. Now if I you now further draw this let us say the test is done.

The test is done, and I know the accuracy of test is 99%. So if the first case of no COVID if it is done accurately, if it is done accurately, so the 99% chance is what? So I will have no COVID because I am if I am in the no COVID population. So 99% cases it should tell me no COVID. Whereas, there is a 1% chance 1% chance which is the wrong thing, which is the wrong test, you know the test giving wrong result.

So that will tell me COVID. Sorry, I should use the same color. It will tell me COVID. So essentially it means that I am a no COVID person, but the moment I am tested, I am told COVID and that is because of the error in the detection, okay? Now if we if you have understood that then we should be able to actually constitute the second part of the uh graph, of the of the yeah graph.

So you I will say that you pause the video here and actually you try to constitute this. Okay. So let us actually try to constitute this. So now that I have COVID, I have COVID. So the test is 99% accurate. So what would be the outcome here? The outcome here would be COVID because now this is the, COVID is the accurate result. So I already have COVID. So in 99% cases I should get COVID, right?

Whereas for 1% cases where the test is wrong, a person who has COVID is tested wrongly would be no COVID, right? So that is now that is that is the that is the first uh level of detailing, which is done to understand how do I actually solve the problem, okay. Now I have to say, you know what is the probability that I actually have COVID? Given this prior given this information about test accuracy, I have to tell what is the probability that I actually have COVID.

Let us actually try to say, let us write it that probability probability that I have COVID, given the test is positive. So because I told you that somehow I am detected positive. I am detected positive. So now I have to tell that given that I am detected positive, what is the probability that I actually have COVID, okay? And let us actually compute that. So I am detected positive.

So that means, either from here or from here, right? So if I, let us say this is case 1, this is case 2. So if so the total number of cases will be A + B. And the true case, true case that I actually have COVID, is essentially this one essentially. I will use this color. Essentially the true case, this is the true case, right? Essentially this one, right? Because I actually have COVID and I am also tested positive.

So that has this has to happen. So if that is the case, I will write on the numerator, it will be B to simplify the whole thing. That is it. That is the simple thing about Bayesian probability. Now if I compute this, what will be the value? So B is 1 into 99 by 100. And my A is going to be 999 into 1 by 100 plus here it is a it is this this one the B right. So it is 1 into 99 by 100.

Now I have to calculate this but I think it should be something around 9 point I think it is 99 on the numerator and in the denominator 1098. Just let me calculate it here. 99 by 1098, it should be something around yeah point 9.02 or 01, something like that. 9.02%. So essentially, it means it is a very interesting result we got.

Essentially it means that though I am tested positive, though I am tested positive, given the fact that I have some number about the accuracy of the test and I have some information about the prior, I actually have a chance of only 9.02% of COVID. So is it not interesting? So even if I am tested positive, I should not be so worried because the actual probability of me having COVID is only 9% 9.02%.

And that is it. So this is how we should try to understand you know like using Bayes theorem about what is the real scenario. And this is the beauty because we can actually this is the first level of you know like, first level of trial we got and we got the probability and we see that it is actually low. And we will see in the next example, how we can continually update.

But before we get into that, let me actually, you know do the same problem with a different approach. So we will do the same graphical, we will take the same graphical approach, but in a slightly different manner, okay. So let us do that. So the same example we are going to do. So let us say uh my info my data was that I have only in 1 only 1000 person 1 has COVID right. (refer time: 20:29)

So let us say, let me actually take, so if it is 1 in 1000, I just want to sort of make a little bigger number so that uh it is easier for me to draw. So 1 in 1000 I can also approximately write, basically write 10 in 10,000 right, the same thing. Now let us have the full population here um. So when I have this this 10 in 10,000, so that means this part this part is actually people who have COVID right, who have COVID.

So essentially, that is the number 10 right? And here the rest of the people who do not have COVID, so that is 9990, I am counting out of out of 10,000. So this will do not have COVID okay? Now my test accuracy, my test accuracy is 99% okay. So if my if I you know use this test for this on this 10 people, so what I will have is that, if I multiply if I multiply this test accuracy, you know with the total number of individuals who actually have COVID I will get 9.9 right?

So given you know given the test accuracy is 99% actual number of people who have COVID let us say or who are identified as COVID is say this red color strip, which is which is 9.9. This red color is 9.9 right? In the same vein, since the test accuracy is 99%. So it means the inaccuracy test, inaccuracy is 1%.

And if my test in accuracy is 1% what I will have for this people, now I have to calculate inaccuracy on this people because this people only if the test is inaccurate, they will be diagnosed as COVID right? So that means that will mean 9990 into 1 by 100. So which would mean 99.9. So 99.9 individuals, they will have COVID wrongly. So let us say this is the 90, this is the 99.9.

This is the 99.9 individuals who are diagnosed with COVID wrongly, okay? Now in my data set, what I have now is this two number. One is 9.9 who are actually who actually have COVID and the other 99.9 were diagnosed with COVID but who do not have COVID, right? So then the probability that someone diagnosed with COVID actually has COVID is that will be reflected by this number because they are the true cases right, this is the true cases.

So the true cases the probability, let me use a different color. The probability probability of someone with COVID given the cases of you know tested positive, so tested positive, this is tested positive, this is tested positive, but this is the true cases okay. So then actual COVID is, you know the true cases are 9.9 in the numerator and in the denominator 9.9 + 99.9, which is essentially 9.9 by how much would this be, 18 109 I guess right?

Now if I do this 9.9 by 108, let me enter that. 9.9 by 108, let me see what was that 109.8. Sorry 109.8 is going to give me the same number, the same number here, the same probability which is 9.01 or the 9.02. Essentially, it is 9.02% right? So you can either do this you know branching approach and graphical approach or you can actually calculate the area under this curve.

And you know and that that way, you can actually um calculate the probability of actual, you know probability of something to happen actually in reality, and there you use the probability of other thing. So here we are using, you know probability of someone you know who does not have COVID, but diagnosed as COVID and who, who has COVID diagnosed as COVID.

So we are basically considering all the facts. So this is how we, you know these are two different approaches I have shown to calculate Bayesian probability. What we will do is we will do a couple of more, some of similar problems. And we will also uh we will see this is the first level we actually get the probability value and we will try to actually update the probability to see what happens in the next trial if I get the same result, okay? So with this, we will conclude the first lecture on Bayesian theorem and the posterior probability.