Applied Econometrics
Prof. Tutan Ahmed
Vinod Gupta School of Management
Indian Institute of Technology-Kharagpur

Lecture - 02
Correlation and Causality

Hello, and welcome back to the lecture on Applied Econometrics. In today's class, we are actually going to talk about a very important concept, which is a fundamental concept, I would say, in all the scientific inquiry. The topic is to understand the difference between correlation and causality, alright? So we will see that all the different sciences we are trying to actually pursue a holy grail called causality.

As if you know where A leads to B, we want to understand the fact the linkage between two factors, you know with certainty. But that sometimes is really very difficult. I mean, most of the cases that is really very difficult thing to actually obtain, you know like to make sure that A actually leads to B. Now we will talk about it, and we will talk about how correlation and causality are different, and how you know they are important in their, you know understandings.

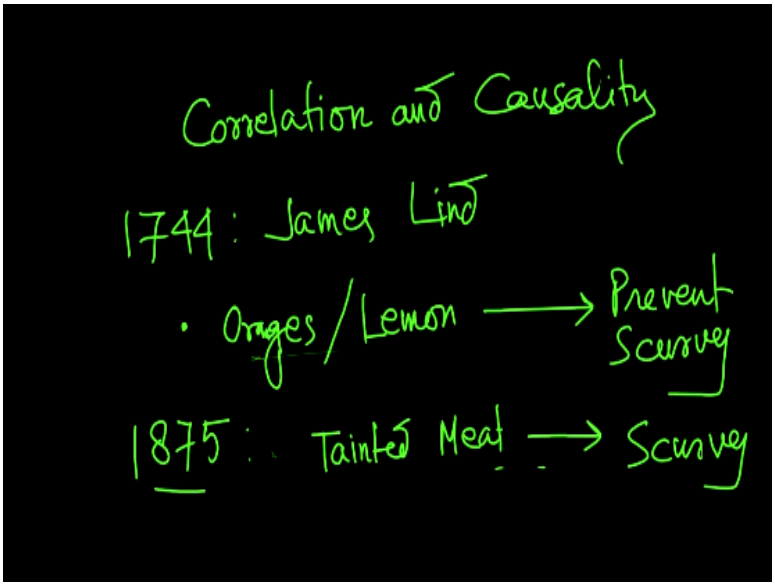
But before we get into that, I will start with a story. And this story is about a disease called scurvy. Now if you know about the, it is a very interesting story, and if you know about how, you know this scurvy disease was so prevalent in Europe during medieval time, and how it used to be a menaces for the sailors. And how, actually, you know people discovered a remedy to it, and what happened after that.

I think this is a very interesting story to begin, because that will also give us an idea about how correlation and causality are different. So let us talk about it. So scurvy is a disease that actually, you know makes your gums you know infected with fungus, and you really, it is really severe, and you can actually die. So around 2 million people died due to scurvy.

I am talking about the time when Columbus was Discovering America, and the sailors from Europe were trying to, you know go to all the corners of the world. So the sailors were really important at the time. And at the same time it was really a risky

business because of no other reason but scurvy, there could be other reasons but scurvy was one of the most, you know formidable reason why, you know the sailors were dying.

(Refer Slide Time: 02:54)



Now fast forward it is 1744. And I am talking about a person named James Lind. Now James Lind then, you know he did a sort of a controlled experiment. Now what is controlled experiment, we will talk about it later. So what happened is that somehow people got to come to this intuition that it is probably the, you know the citric fruits like the lemons and oranges, they are actually, you know helpful to curb the disease, scurvy.

So you know Captain James Lind, he is the captain, he did a controlled experiment, where he had a group of people whom he did not provide the lemons and oranges. And whereas there was another group of people whom he did provide lemons and oranges. And interestingly, it was found that the people who are given lemons and oranges they did not have scurvy.

So it means that the lemons and oranges are actually helping them to get rid of this you know this dangerous disease. So you know if I want to represent it in a diagram, so I can say that oranges, lemons actually could prevent scurvy. Now the question is, so it is alright. So you know people understood it, people actually, the sailors used to now take lots of oranges and lemons when they are traveling, you know to other parts of the world, and it was all good.

Now fast forward, we will come to around another 150 years. So it is the year 1875. Now then, there was an expedition that was supposed to happen to the Arctic regions. So you know and artic is really cold and people thought that it is perhaps a good idea if they can actually create something out of the oranges and lemons that can help them to, you know have something warm and you know like they can enjoy their drinks.

So what they did, they actually created some sort of you know lime soda and some hot drinks that they created by concentrating like they boiled oranges and lemons and you know created some juice. And they carried that juice in bottle and that is also good when you are traveling because then you the inventory, the storage that you need to keep in your ship for you know for oranges and lemons vis-à-vis this soda bottles.

The soda bottles actually take much less space, right? Now, so that is the idea. And people actually, the sailors were actually trying to, you know boil these lemons and oranges and they are sort of packing all those drinks in some bottles. Now what happened is that you know surprising for everyone, everybody found that scurvy is back, and the sailors are dying again.

Now that was something really tricky. I mean we have already discovered that oranges and lemons can actually help you to get rid of scurvy. But then, when I packed the oranges and lemons in a bottle, I see that the scurvy is back. Now what is happening really here? Now you know another so there was another argument at that time that people are also carrying meat, okay.

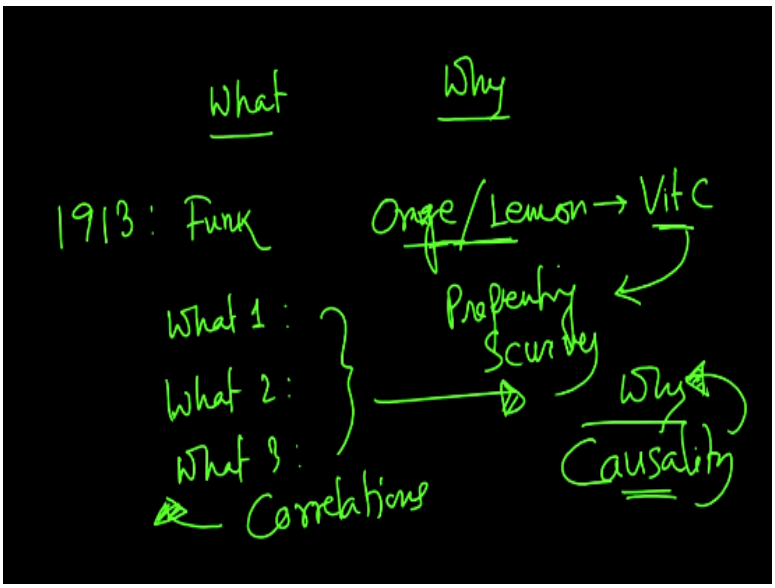
So of course when you are you know traveling in sea and you are on water on basically inside the ocean for months, so your meat is no longer fresh meat, right? It gets somewhat tainted. Now people ended up thinking and that was their worse experiment. So you know some sort of argument, there was very strong argument made, that it is not, the scurvy has nothing to do with this oranges and lemons anymore.

But it is the tainted meat, the tainted meat, which is causing scurvy. Now that is rather interesting. So then, you know so then people thought that, you know perhaps if they you know remove the diet of tainted meat from their meal, perhaps they will be you know free from scurvy. So you know sailors tried that. And it was unfortunately it was not it did not work. So scurvy was not going.

Now then that is, you know that became a very critical question then what is exactly happening here and that is where we will understand the implication of what it means by correlation and what it means by causality, okay. So when you say that, you know when Captain Lind said that oranges and lemons prevent scurvy. So it was like, you know answering what is causing what is, you know what is actually preventing scurvy.

So it is the answer is oranges and lemons, but nobody has a clue why oranges and lemons are actually preventing scurvy, alright? Or here also people are talking about it is tainted meat, which is causing scurvy. But people had no clue why tainted meat is causing scurvy. So there is a difference when I say correlation causality is the one thing we have to understand. Is that a question?

(Refer Slide Time: 07:55)



There is something called a what question and a why question, okay. So when we say correlation, we are basically speaking about what, okay. So it is the oranges and lemons that is, you know preventing scurvy. But I do not know why. Or it is perhaps,

you know some people are arguing there is a plausible explanation that it is tainted meat, which is causing scurvy, but I do not know why, okay.

So that is the problem. So up until I think it is 1913 when Casimir Funk actually discovered something called vitamin, and he is a Polish scientist. So when Casimir Funk actually discovered vitamins, so then people got to know that it is, you know some vitamin called vitamin C, which is actually preventing scurvy. So in our previous, you know diagram when we said that oranges and lemons they actually have something called vitamin C.

And that vitamin C is actually preventing scurvy, okay, preventing scurvy. So now we can explain that, you know so it is not, you know not just an answer of what is preventing scurvy? Like I can say, you know oranges and lemons are preventing scurvy, but I can also say why oranges and lemons are preventing scurvy. And the answer is vitamin C. Oranges and lemons have vitamin C.

So you see there is you know there is a what question here? What 1. So what is preventing scurvy? This is orange and lemons. That is another what question, what in oranges and lemons are preventing scurvy? So it is the vitamin C, right? Now we can further ask why vitamin C is preventing scurvy, right? That could be another what. Yeah, why vitamin C is preventing scurvy.

So you have to go to that molecular structure of the vitamin or the micronutrients, whatever there in the vitamin and you have to see how it is interacting with the human cell, right? So there could be a series of whats, alright? The series of what questions to reach to an answer why, okay. So this what questions, a series or what questions is giving you an answer to a why.

So why is much deeper than what as you can see. So you have to ask many words to go to, you know the why part, right. So that is where the causality comes. And causality actually answers or let me okay, I can write it here. Causality actually answers the why part, alright? And correlation actually helps us to discover the what part, okay.

So you can simply say that, okay, you know so and so we can see, you know when we said that oranges and lemons are, you know preventing scurvy, we can simply say that, you know this is basically a what question. So you can simply see, you know if people are having a orange and lemon diet, so they can they are actually saved from scurvy. But why?

So then we have to go much deeper, we have to ask many questions. And then that is the, you know that is something that actually, we can call causality. So that is the introduction of, you know how correlation and causality are different. And in the next lecture, I am going to further detail about the differences in the correlation and causality. Thank you.