Transcriber's Name: Prabhavathi
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology-Kharagpur

Lecture-45
Logistic Regression (Part-I)

Hello guys. So, in continuation of our discussion on relation analysis, today we will start a new topic that is logistic regression okay. (refer time: 00:34) So, like before going to discuss about the different concepts covered and all, I just want to bring out one example. Like, all of you I should not say all of you most of us have visited fair. Fair means mela right. We call it mela in hindi or in other languages also.

We have visit fair at some point or the other and there one very common, what is the activities there one gun is kept and some targets are kept and if we can shoot the target, then we will get some boonty right. We will get some awards, some whatever maybe it is. So, now take that example. There is some targets are there and a gun is there I have to shoot the target. So, I want to know that I want a model, which will tell me ahh like if I shoot from a particular distance, what is the probability that I will hit the target okay.

So, I am repeating again. I want a model that will tell me if I shoot the target from a particular distance maybe in feet and then that is the probably, that I will be able to hit the target. So, when I mean to answer this sort of question, first what will I have to do? First I will have to collect some observation from some data means, first say I will try to shoot the target from 1 feet, 2 feet, 3 feet, 4 feet, 5, 6, 7, 8, 9 or 4.5, 5.5 whatever it is, I will try to shoot the target from certain distance and then I will see what the results. Whether I could hit the target or I could not.

When I started when I start try it from 1 feet I could hit that, yes I could hit the target good. Then from 2 feet yes maybe again, from 3 it maybe not, from 4 feet yes. So, that way I have collected few data okay. Now I have this data. Now I want to know, given that I have started, I tried to hitting the target from particular feet, from particular distance, so, what is the probability of my hitting the target okay.

So, somehow I need a model, which will give me the answer to this. So, this this this is where the logistic regulation comes. Again though this is a simple application right, again let me take you give you an example in medical domain suppose in some medical research wants to find

out the probability, probably that a person will get cancer or whether it is highly probable or probable and not at all probable.

Means, highly some, what to say I should not say probable I will say quite a high chance of getting cancer, less chance and no chance at all maybe. These are the three categories maybe. Earlier example was these 2 categories and maybe here I am thinking 3 categories. So, suppose for that I will be taking different parameters like like maybe the age of the person, race of the person.

The genetic predisposition of the person, family history then what to say exercise history, his lifestyle history then maybe smoking habit, drinking habit many more suspects. Workplace environment at the workplace where does he stay everything put together. So, based on that whether the person is highly likely to have cancer or very less likely or not no chance at all of having a cancer, maybe 3 categories.

So, based on these different factors, so, how will I these are basically I will tell it is a predictor. These are based on these predictions, this predicted these variables how can I tell, how can I classify the person will fall in which category. So, this way again logistic regression comes. Again let us let me give you one example on say engineering applications. Engineering like better if I talk of from reliability perspective right. So, in reliability there is a term called FME a failure mode effect analysis. So, you might not be aware of this FME basically, okay you know no need to be aware of that.

What we do is that sometimes we try to find out what is the probability that the system will fail, a process will fail or a portion of the process will fail. For that what we try to do is that in what way the system can fail, we try to find out all the different ways in which the system can fail and there is a it is a very big technique okay. Let us not discuss that in that it will take 1 hour, if I start discussing FMA.

Now that there there is some some factors like, we have identified from the different ways a system or a process can fail. Now what is the severity of different we have identified different faults. Now what will what is the severity of the different faults, what is the occurrence of different faults and then what is the detection rates of different fall? All this we give some some value in a particular scale maybe from 1 to 10 scale.

This value we give it is usually subjective some experts give these values or we get this value just from the past data. So, this is where and based on this basically we try to calculate the RPN bricks probability risk probability number of the process excuse me. So, now in this case, so

where I am giving the number, the severity of the fault than the occurrence rate of default, when I am giving the number just by an expert judgment.

Maybe an expert is giving or just from the past data. Somehow some of the past data I am doing some calculation. I am giving the data and so that that if I could use logistic regression here then it will be more effective. So, these are the different, so logistic regression you can see it has its application in different domain. So, that is ahh the topic that we will be discussing today. So, for ahh first I will try to bring out a difference between the regressions, because we have learned only regression analysis.

Logistic regression is also something regression right. So, irrigation what is there? So, what is the difference between this regression analysis and logistic regressions? Then we will come to the concept of logistic regression, how to go ahead, how to do this logistic equation, how to model it and then I will also mention the type of logistic regression. Here I will be just mentioning the type of logistic execution. I will not be discussing those I will be discussing in my next lecture so okay. (refer time: 06:54)

Now so we have done regression analysis. Let us consider simple linear regression okay. We have discussed different regression. Let us just consider the simple linear regression. Simple linear regression what we have tried to do, means regression analysis means, we are trying to predict the response variable from a regression variable is not it? From an independent variable we are trying to predict a dependent value of a dependent variable.

So, now for that what we have tried to what we have done. So, for basically we tried to we found out the function. The function that maps the output with the input the function that maps the output with the input, how it maps. So, this function is only y = f x. Now what is this f x? F x is a function, which has certain parameter. Now how to find out the value of the parameter? we Once we found out the value of the parameter then only a function is complete and we can say okay.

This is the function which gives the relation between the independent variable and a dependent variable. Now to find out the value of the parameter, what we have used? We have we use least square estimation method. So, basically we try to find out the ahh error between the observed variable and the predicted variable. Predicted value, observed value and the predicted value we try to find out the error, we have squared this error and then we have to we try to find see that when this error is minimized.

So, at when this error is minimized that is the point where we can consider the coefficient is that

is the best value of the coefficient, because regression there may be different function. Now function depends on the coefficient. So, this is the best value of the coefficient, where this SSE is minimum. So, now we have minimized that function and from that we found out the value of the β whatever is the parameters β 1, β 2, α, β whatever it is we found out the parameter.

So, and one more thing is that, this regression analysis it is applicable when the dependent variable it follows a normal distribution. In case of logistic regression here the dependent variable does not follow normal, here and normal distribution here, the dependent variable is just a category value. If The first example what I have given trying to suit a target, what are the two categories?

Suiting the target, success success or failure. You could successfully suit the target you could not so the success or failure there are two categories. In case of the cancer example there were three categories right. So, similarly for there may be 2 categories, 3 characters, 4 categories are more than that. So, whatever it is. So, here the dependent variable it is a it is a basically categorical variable, which may be 2, 3 or few or more outcomes.

Now one more difference is there, actually both are doing prediction. Regression analysis also predicting predicting the independent variable, logistic regression is also predicting the independent. But regression analysis is predicting the independent variable here, the independent variable is in the same same scale, what we want to predict. But in case of logistic regression here we are predicting the categories with which categories belongs.

Probability that it belongs to which category, we have basically predicting the probability of belonging to different categories, if there are 2 categories, probability of belonging to the success category, probability belonging to the failure category. (refer time: 10:15) So, now here one more thing like in regression analysis we can do regression analysis if we take a small sample of data also, but in case of logistic regression we need more bigger sample.

If the number of categories are more than, even more bigger sample we need, so as to be statistically significant. So, And one more difference is that and what to say in the regression analysis see we have used least square methods. Least square method to estimate the value of the parameter, estimating the value of the parameter that is the coding right and that once we do that we are done.

So, we use here here we use list LSE method that is least square estimation matter to estimate the parameter. But whereas in for logistic regression what we use is that we use a concept called maximum likelihood estimation MLE matter. So, what is basically MLE? In very simple

language, if I tell what is MLE? Like given a set of observed values, like that for the example, that the shooting the target, I have taken and observed values like shooting from different place and whether it is success or failure.

So, given a set of observed values, set of values that we have observed the done we have to find out the model parameters okay. The Select the model parameter, the parameter should be such that, so that the likelihood of occurrence of that value. Whatever we have observed likelihood of occurrence of that value is maximized okay. So, the parameters should be as chosen in such a way that likelihood of the occurrence of that data set what we have observed, likelihood of the occurrence or the observed value is maximized okay. We should select the parameter in such a way.

So, that is the concept of MLE basically. So, we will see MLE in details here. See the principle of MLE is to estimate parameters but the principle of MLE is to estimate parameters by switching by choosing parameters values that give the largest possible likelihood. We have to choose the parameter in such a way that whatever the object values we get that that likelihood of getting that observed value is maximum okay.

We have to choose the parameter in such a way that technique is called MLE maximum likelihood estimation. So, now regression analyze predicts the value of a dependent variable. Logistic regression predict the probability of a given value of a dependent variable. Regression less is predictable value, logistic regression predict the probability of a given value okay. Probability of success, probability of failure whatever is the category.

If you have 4 categories probably a first category, probability of second category, probably a third category, probably a fourth category that is what logistic requisitions gives. Both estimates the respective model parameters, in both cases we have to estimate the respective model parameter, when we respect them ahh estimate the model parameter then only our model is complete essentially. (refer time: 13:15)

So, now see the model. In case of regression analysis if we consider a simple linear regression with just one independent variable, which was one regressor then we can consider its as a simple linear regression, maybe one regress or more whatever it is. We can consider it as a it will graph a straight line, if there are more number of variable then maybe it will be in a multi-dimensional plane. It will be a straight line okay, if it is a linear regression.

So, given a value of x it will give the value of y. This model will given a value of x, this model will give the value of y, if you know the parameter. Similarly for logistic regression model the model

is a sigmoid function okay. So, model is a sigmoid function, now what is a sigmoid function? It is nothing but it is it is a function, which takes an input as a ahh a real value and then outputs a probability.

Probability probably will definitely in the range of 0 and 1. It outputs a value in the range of 0 and 1 basically. That is we can tell this is a probability okay. It takes as input a real value and it outputs a value in the range of 0 and 1. So, this is the shape of a sigma sigmoid function, here you can see. It is taken as say this this it takes an as say sigmoid function okay.

Now what is the sigmoid function? If I can In a very simple way if I can take sigmoid function is nothing but f. f x = 1 by 1 + e to the power z. In a simple form, I can tell that what is the sigmoid function is 1 by 1 + e to the power z. Now this z is that z is again an expression it in terms of some parameters, in terms of some variables. So, here in the figure you see I have used, can you see the figure?

Of course it is very written in a very small font. I do not know whether you will be able to see that, you zoom it if you cannot see. So, see here, here I have used for z I have used - c 1 into x - c 2 okay. So, this c 1 and c 2 are what, c 1 and c 2 are the different parameters, c 1 and c 2 basically this will decide the shape of the curve. Shape of the IT is S shape Definitely it is S shaped, but this value of c 1 and c 2 it will decide the shape of the curve.

So, here you see if c 1 = 2 you see this blue colored line this is the blue colored line this is the blue color line this is the shape. If c 1 = 0.5 it is black solid line, c 1 = 1 oh sorry c 1 = 1 is the black solid line, c 1 = 0.5, this is this red dotted line. So, for different values of the parameter we get different shape and whatever is the shape, whatever it it outputs a value what is what will be the value of f x? Value of x will be in the range of 0 and 1.

So, we will be using the log is logistic regression model use this sigmoid function okay. Because why we use this σ because in logistic regressions we want a probability, probability will lie in the range of 0 and 1 right. So, that is why the sigmoid function comes into use in case of logistic regression model. Now we will see how we have used. (refer time: 16:41) Now for Let us see and first see an example here.

So, here is an example where the hours a student the number of hours a student is put on the studies that is given as well as whether he has passed the exam or he he could not pass the exam is given here right. So, say similar type of example we have discussed earlier where the it is in the data was number of hours that a student put on for studies and the other data is the independent variable is the stay score what you got based on the number.

We try to find out is there any relation between the number of hours studying and that the score of the student. So, we found here there is a relation, there is a linear relation and we could find a really linear relation. Now here it is the student the score is not given, but whether the student has passed or failed two categories, 0, 1, 0 is fail and 1 is pass okay. So, you see here is the data.

So, all this green color dot is the past data and category 1 that is pass and red color dot is the fail, for different values this is hours of study and this is the probability okay. This side will get probability and this is the hours of study. And here till now we did not get the probability. We Just from the data whatever this is our observed data, this data what is given this data what is given this is our observed data fine.

So, based on this observed data we will have to find out the given the given the student has put say ahh recep 4 is given which value is not given here. Let me take a value, which is not given okay 0.25 or whatever it is. Maybe 7 hours, some value, which is not given. Maybe say 1.8 hours okay or 1.9 hours. Maybe a student has put 1.9 hours what is the probability that he will pass the exam. We want to find that. So, first this is the observed value okay.

So, to predict that given the the student has put x hours whether he will what is the probability that we will pass to find that we will be using the sigmoid function. Now this is the what to say first we have plotted the data plotted. The data plotting will not will not be using the sigmoid curve there; σ will give us the output. So, data is this like for whatever hours of study just to what students who have failed we have given it with a red dot.

Where pass we have given you the green dot. Now to calculate the given x we need to calculate the probability whether the probability of passing or probably failing whatever it is. For that now we will be using the sigmoid function. Sigmoid function means f x = 1 by 1 - e to the power z. Now what is this z now? We will have to use the sigmoid function to give to get the result. Now what is this z for us? (refer time: 19:42)

So, basically in this ahh what to say, this concept was developed and popularized by primarily by Joseph Berkson, where he coined a new term logit okay. Now what is logic we will see? you This actually logistic regression has come from this logit only logit regression, logit regression or logistic regression. So, this Joseph Berkson on his coins this new term logit. But some this logit basically ahh logit will is somehow related with z we will see what it is okay.

So, logistic regression is a statistical method, it uses a logistic function to model a binary dependent variable. Now okay, it may be binary or it may be more than one. So, if suppose now we are looking for binary in this example because, it is student pass or fail, pass the exam or fail could not pass the exam. So, it is a binary, so it uses a logistic function to model a binary dependent variable or although many more complex extensions exist.

Complex extension may be more than two categories okay. So, our job here is to find logistic regression estimates the parameter of the logistic model. We found that it will be using the sigmoid graph. Now we will have to estimate, what will be the parameter means z is something, that there are some parameters to it. Now we will have to estimate the parameter. First what is the sigmoid function for us and then what are the parameters okay. (refer time: 21:05)

So, see a binary logistic model has 2 dependent variables. If we are considering binary it has this 2 dependent possible variable, with 2 possible values that is pass or fail, happy or sad whatever it is success, failure okay. So, it is represented by an indicator variable, where the two values are labeled 1 or 0. If it is binary and repeating it again and again. (refer time: 21:29) So, now the logistic function, takes the following form.

So, logistic function, so what is the law this logistic function means we want to find out probability that the student will pass the exam? So, that let me write it this way, probability that the student will pass the exam given certain values of x, x may be the hours. So, that I am writing it this way or let me call it as p x this is let me call it as p (x). Given x what is the probability that he will pass. So, this I am writing is this way given x what is the probability he will pass. So, to find this, I will be using the sigmoid function. So, what is my sigmoid function? (refer time: 22:06)

Sigmoid function is So, what is my sigmoid function, so P x = 1 by 1 + e to the power z. So, this if I multiply both sides by e to the power z but will I get e to the power z by 1 + e to the power z is not it? If I multiply both sides by both numerator and denominator by e to the power z this remains sense I will just get e to the power z by 1 + e to the power z.

Why I bought this in this form. It becomes easier for calculation. We will see that. So, here, so this is in this form. So, my P x given x my probability that the student will pass that is I can tell it y. Y = P x is equals to nothing but e to the power x by 1 + e to the power x okay. Now let us not write x it is let us write as z okay. Now what is this z? It is nothing but a if if there are just one variable.

In in our example, what is given? It is just one variable. Hours put on study, there may be more

than one variable okay. So, for each variable there will be a parameter is not it? So, z is nothing but a linear combination of all these parameters along with an intercept, it is nothing but a linear combination of all those parameters along with an intercept. This is the same thing what we got to see in the linear regression okay.

So, here z is nothing but this, suppose there are m parameters x 1 to x m. So, it will be my z will be β 0 + β 1 x 1 + β 2 x 2 + β 3 x 3 up to + β m x m, if there are more than one parameters. Now in this case there is just one parameter the example hours to study and whether he has pass and fail. So it will be β 0 + β1 x e to the power β 0 + β 1 x by 1 + e to the power β 0 + β 1 x okay. So, this is the probability that the student will pass.

Now we will introduce a new term here that is called odds. It is it is not new, many of you may be ever of this term is what is odd? Odd is the ratio probability of success and failure. Odd is basically a ratio of something happening to the something happening probability of something happening by something not happening. Usually when you try to find out probability, what is this? Probability is this ratio of something happening by means the sorry not ratio when we try to find out probably we how we write?

As the number of times something has happened divided by the total number of times that anything that it can happen. The happening plus not having put together the whole sample space that is how we find out the probability, but in case of odds that is the difference is that there in the denominator it is not the total sample space, the denominator is just the number of ways it can fail, number of ways it can pass by number of ways it can fail okay.

So, odd is the ratio of the probability of success and failure. So, that is it. So, now what is odd? Odd is in case in this example what is this? Probably that the student passed by probably that the students fail P x by okay P x by 1 - P x. This is the probability that the students fail. So, if I write P x by 1 - this is P x, P x by 1 - P x if I use this is P x and if you do simplification you will get odd. this is the value of odd.

You can do it yourself. Say very simple calculation okay. So, this is the value of odd. Now what will do is that this logit logit is basically the log of this odd. This new term is Joseph this particular person Joseph Berkson, who has coined this term logit. Logit is nothing but the log of this odds okay. So, log of odds. If we try to find out the log of odds what we will get? Log of odd is nothing but this term.

This log of odd is nothing but this is called a logit and this is the z for us okay. This is nothing but z here what we have used z, here I am using the term t z or t anything instead of z if I write here

t, it is a variable whatever name you give okay. So, this is the logit. That is t is the logit, so value of t is this okay. So, now we know P x. What is P x? P x is this. And what is this? This is nothing but ahh sorry this b 0 plus this portion not e this portion.

What is this portion? This portion is nothing but t then P x I can write e to t e to the power t by 1 + e to the power t fine, that is my P x probability that the student will pass okay. So, now this e this t has β 0, β 1 ahh terms are there. Now so if I can find out this value of t, how can how I can find out the value of t. If I know the value of all the parameters x value I will get from the data, x is the independent variable, which I want to know.

If I put this many hours of study will a student pass or will I pass? So, x value I know. The only things I do not know is the β the parameters. So, if I can find out this value of the parameters, then I will be able to find out the t. If I can find out the value of t, then I will be fine able to find out the probability is not it? This is nothing but the sigmoid function is not it. This is nothing, but the sigmoid function. Sigmoid function was Our sigmoid function how we have written 1 by 1 - e to the power of sorry 1 + e to the power 3.

This is just in writing in a different format. I just got this e to the power t by 1 + e to the power t. This is nothing but a sigmoid function. So, we use the sigmoid function to find out the probability that the student will pass given that he has put x hours in study that is e to the power t by 1 + e to the power t. Now if I am if I can calculate t, for calculating t I need to know my the value of the parameters right. (refer time: 28:29)

So, a small example is here. A sample is collected to examine the effect of toxic substance on tumor okay. A subject is examined for the toxic content in the body under the presence or absence of tumors. We have some data, first to do anything in any any sort of statistical studies we need some data first right. So, what is the some subjects if we call this some people are examine and they try to find out what is the concentration of particular toxic substance and the person has tumor or not.

So, these are the datas, which you got. So, concern for around see here. The number of subject at each concentration N and the number of number having tumors is shown in the table. So, for concentration of 0 we found total 50 subjects, total 50 subjects as concentration 0 out of that two had tumor. Similarly for concern 54 subject had concentration 2.1 and 5 had tumor.

Similarly 46 persons had concentration 5.4 out of them 5 at tumor similarly okay likewise for the outer values. Now the first thing is we will have to find out the logic that is the value of t. If we can find out the value of t then we will be able to solve the find out the probability right. So, to

find out the probability what we need? We need to find out the odds. How What is what is odds? Odds will be probability of having tumor by not having tumor okay.

So, what is this? How many person is having tumor? Two. So, it will be odd will be 2 by 48. So, here odd will be 2 by 48 for the first term this is this. Second is the odd will be 5 by 49, likewise we found the odds. Then logit is what? Log of odds and log of odds is that this value okay. So, here we got everything right. We found a logit we found the logic value. This is the logit value that is the value of t okay. (refer time: 30:40)

So, now we can find out the value of the parameters. So, what is t? T is nothing but is there is only one independent variable. So, there will be two parameters $\beta 0$ and $\beta 1$. So, $t = \beta 0 + \beta 1 x$. So, this logit is a linear function and with this data we will be able to find out the $\beta 0$ and $\beta 1$ value. Here I will have to explain something first listen to this. So, with this with this data, because just see we know t, we know x just we do not know $\beta 0$ and $\beta 1$.

Let us take two of this and we will be able to find out $\beta 0$ and $\beta 1$ okay or we may plot the data what is $\beta 0$? B 0 is the intercept, $\beta 1$ is what? B1 is the slope. We can plot the data, from there also we can find out the $\beta 0$ and $\beta 1$. Once we know the $\beta 1$, $\beta 0$ and $\beta 1$ then we know the t value. Then given any x will be able to tell whether the person has tumor or not okay.

So, now in this question see, we could easily found $\beta 0$. First of all if the points fall in a straight line just by finding out the intercept under slope slope we could find out the $\beta 0$ and $\beta 1$ or by since there is two unknown parameters, just take two values and will be able to find out the $\beta 0$ and $\beta 1$. But the question is if I take these two values I will get one set of $\beta 0$ and $\beta 1$.

If I take these two values maybe I will get some other set of $\beta 0$ and $\beta 1$ maybe if I take this value and this value maybe I get another set of $\beta 0$ and $\beta 1$. So, there may be different values of $\beta 0$ and $\beta 1$. Now, which is the best value? okay, To find out which is the best value of $\beta 0$ and the best value of $\beta 0$ and $\beta 1$ is that which will make the likelihood of getting the value what we have observed is maximum.

It will maximize the likelihood of occurrence of the observed value understood. So, for that we will have to use the MLE method. Now let us not use that. Now let us just stick to this we got the $\beta 0$, $\beta 1$ value. (refer time: 32:29) and then given that for example a subject is exposed to a concentration of 10 has an estimated probability of tumor. So, we could found its estimated probability of tumor is 0.36 okay. (refer time: 32:41)

So, now the log the log-odds that is the logarithm of odd is a linear combination of one or more independent variable that is that we call it predictors okay. There in the earlier example the concentration is just only one independent variable. Concentration is the predictor. The independent variable can each be a continuous variable it can be any real value okay. The probability of the value can vary between 0 such as certainly false and once this is certainly true, if it is just two categories. (refer time: 33:19)

Increasing one of the independent variable, if in the example what we have seen there was just one independent variable. If we increase one of the independent if there is more independent variable, multiplicatively scales the odds of the given outcome at a constant rate. If we increase one more independent variable if there is one more independent variable, which also decides the dependent variable then what happens.

We will be using one more parameter and what is its multiplicity multiplicatively scales the odds of the given outcome at a constant rate, plus β 2, x 2. Earlier it was β 0 + β 1 x. Now it is β 2, x 2. Another one variable x 2 and as a constant rate that is β 2, with each independent variable with each independent variable having its own parameter fine. For a binary dependent these generalize the odds ratio.

For binary we have seen just it is just two, binary logistic the dependent variable has two level, multinomial, outputs with more than two nomial ah. more than two levels and see if it is two then we call it binary logistic regression. If it is more than two we call it multinomial logistic regression. Then we have something called ordinal logistic regression. If the multiple categories are ordered, ordinal its ranking you knows?

So, if we can order the different categories then it is called ordinal logistic regression. Here also we will be using the sigmoid function only, which is just give the probability maybe we will be using some threshold value. In case of binary maybe the threshold value is 0.5. If the probability of occurrence is ahh greater than or equal to 0.5 then maybe it is success. If the probability of occurrence is less than 0.5 then it is failure. We can consider it that way.

Similarly if the 3 levels, accordingly will set the threshold for each level okay. So, similarly we can have an ordinal logistic regression where there can be ranking among the the categories. (refer time: 35:06) Logistic regression models the probabilities for classif ah is Logistic regression a classifier now that is the question okay. So, logistic regression it models the probability for classified classification problem with possible outcomes.

What does logistic equation gives? It gives the probability of the possible outcomes. What is the

probability of being in this class, what is the probability of being in this class being in this class, this class, this class. What is the probability of being in different classes? Logistic regression gives that it is basically an extension of linear regression model for classification problem.

But it is not a classifier. Classifier, classifies into different classes, it is this gives the probability of being in a different classes okay. The logistic regression simply models probably of output in terms of input and does not perform statistical classification. It is not a classifier. It is strictly not a classifier. It just gives the probability of being in different classes. In that way we can say it is a classifier actually.

But it is actually not a classifier. It gives the probability of being in different classes. Although it can be used to make a classifier, how we can use to make a classifier. That is what that is what just now what I told by choosing a cut-off value and classifying inputs with probably greater than the cut-off as one class, below the cut-off as the other. Logistic regression, give us the probability of occurrence of a particular value, probability of passing the exam, probability of failing the exam.

Now if we choose a particular result say suppose 0.5 we have choose that result. Now if we get a probability more than 5 or more 0.5 or more than that then we can this it belongs to success class. If it is less than 0.5 it belongs to the failure class. That way we have you so means we have used this to make to act as a classifier. But actually what the logistic regression gives, it gives us the probability of belonging to different classes. It is strictly not a classifier, but we can use it to make a classifier. (refer time: 37:04)

So, there are different logistic regression techniques. So, one is binary logistic regression. Under binary logistic regression, there are definitely two categories of outputs. And there may be one explanatory variable there may be many explanatory variable. For multinomial there are many explanatory variables and many categories or maybe one explanatory variable, then many categories that is not possible right okay. So, these are the different types of logistic regression, so, which we will be discussing in the next coming few lectures. (refer time: 37:25)

So, now coming to the conclusion, so what this lecture provide an introduction to logistic regression, that comprises a background of the development of logistic regressions, compression between linear and logistic regression, how logistic regression is used as a classifier, how we can use logistic regression as a classifier by making the ah by having a threshold value right. so and different types of logistic regression that also we have just mentioned here okay. We did not discuss. In the next lecture we will learn more about logistic regression. (refer time: 37:55) With that this is the reference. And thank you guys.