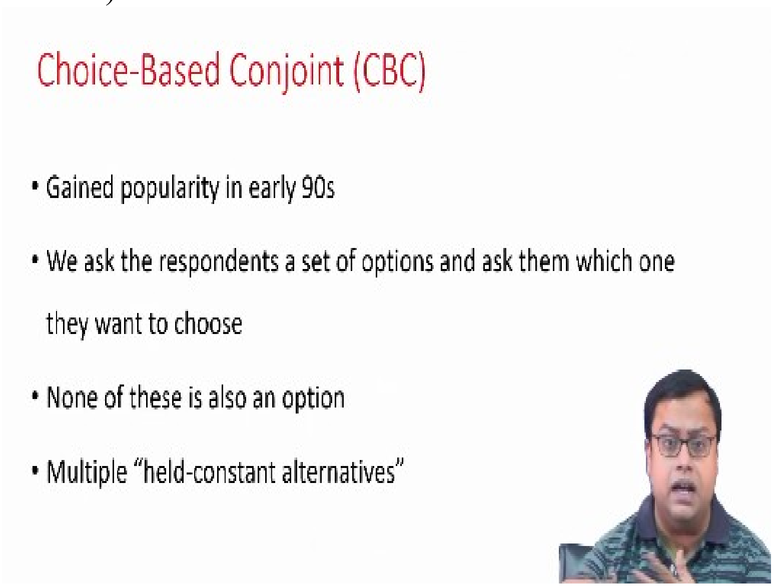
Marketing Analytics
Professor Swagato Chatterjee
Vinod Gupta School of Management
Indian Institute of Technology – Kharagpur
Lecture – 12 - What Consumers Want (Contd.)

Hello everybody! Welcome to Marketing Analytics course. This is Professor Swagato Chatterjee from VGSOM, IIT, Kharagpur who is taking this course for you. Today's topic is What Consumers Want. We are in week 2 and session 6 and today we will talk about choice based conjoint.

(Refer Slide Time: 00:39)



So till now I have talked about various other types of conjoint analysis like ranking based and rating based, but what happens is when you give lots of options to a consumer or in this case of sample, it becomes difficult for a consumer or a sample that how we can rate all of this stuff. So if he has around let us say 10 options or probably more, 16, 24 options in his hand, it is very difficult to do a one-to-one analysis and find out that whether my option x is actually higher than option y or when the options are very close to each other how I can rank them and etc that becomes difficult.

So when that is the problem, when that is an issue then what we have to do is we have to reduce the number of options which we give. So in choice based conjoint which was gained popularity in early 90s, we ask the respondents a set of options and ask them which one they want.

So let us say I give you 4 options and then I ask you to choose rather than rank. Ranking is difficult. If I ask you to rank between let us say A, B, C, D you have to do, 2 by 2 analysis for

A, you have to compare A to B, A to C, A to D then B to C, C to D and then B to D and so on. So all of these things you have to do which becomes difficult sometimes. Rather than that if you just have to find out A, B, C, D, if I know that okay, I prefer A more than B and if I know I prefer B more than C then it is automatically A is preferred more than C, so one comparison comes down.

Similarly, if you have 4 options and some more number of comparisons are coming down. So that is something will reduce the cognitive load on the mind of the consumers, that means that the results that you will be getting, the outcome that you will be getting is more close to the truth.

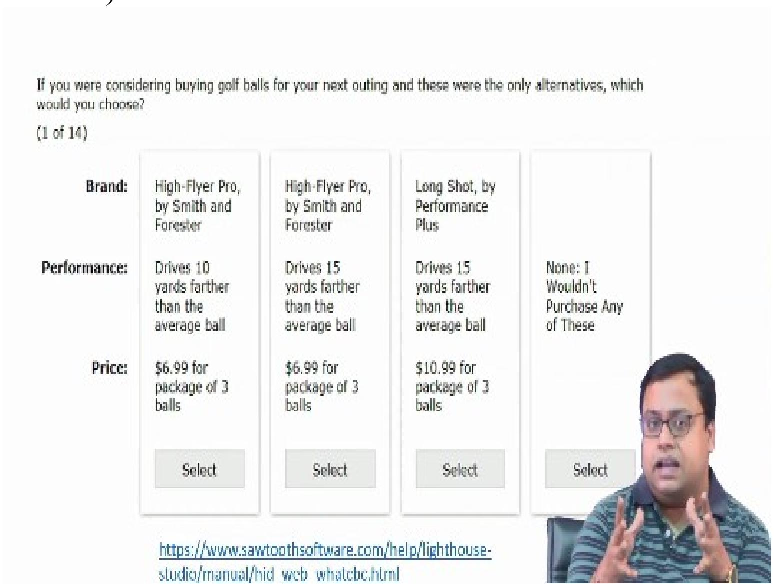
Now after that you have to do the analysis which is a little bit difficult than the normal ones which is like ranking based system or probably the rating based system. So we ask respondents a set of options and ask them which one they want to choose and so because you are giving set of options, there can be no choice options as well , I do not like any of this. That means that the cut-off utility after which I will actually buy below which I will not buy, if my all the 4 options are below that particular utility that I have in my mind, I can have the option that I will not choose also.

So that is something that is very important which you do not give in other cases. You give 16 options or 20 options, you do not know after what option people will actually think about purchasing. There can be around probably 12 to 15 options out of those 20 options which consumers do not even think about purchasing.

And if they do not even think about purchasing then the information that they are providing about those aspects might not be right because they have not experienced this, that is probably below their utility level, they do not think a lot about those kind of products. So given that has a background, we also give in this case none of these as a choice and we can also have multiple held-constant alternatives.

So you can have certain alternatives, which is held constant and you can have multiple such things when you are giving the options. So this is how it looks like.

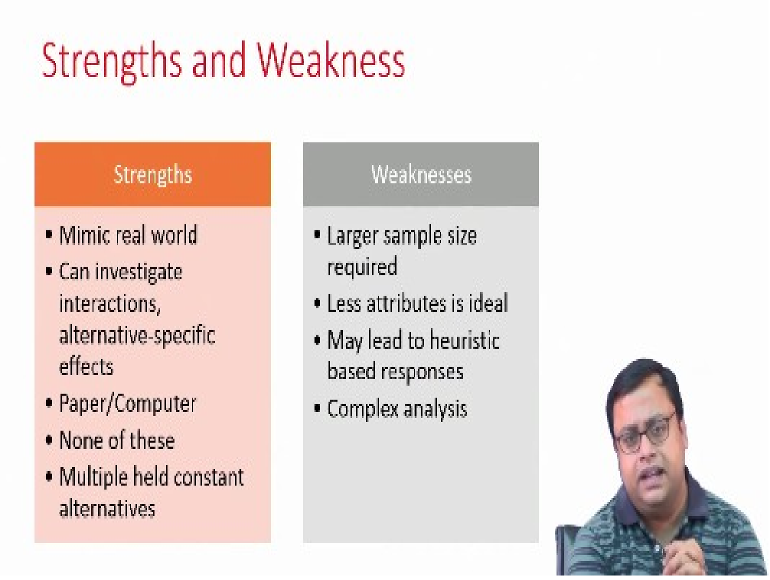
(Refer Slide Time: 04:00)



So let us say it is between that if you are considering buying golf balls for your next outing and these were the only alternatives available there, there are no other alternatives available, which one would you choose? And in this particular case, I have not given, sorry I have also given. So you will check at the fourth one is, none, I would not purchase any of this, that means so no choice option is also there.

So this particular thing has been taken from a software called Sawtooth software who actually give this kind of a view which is very common in the market and you give brand. So brand changes like High Flyer Pro by Smith and Forester and then High Flyer Pro by Smith and Forester remains constant and Long Shot by Performance Plus is the third option and then you can change the performance also and the price also. And you want to see that what combination people are willing to choose.

(Refer Slide Time: 04:58)



There are certain strengths and weakness of this particular method. Before I jump in, so let us say I give you 8 such choices to you and if you give 8 ratings, that means I am getting, from one single person I am getting 8 results, 8 data points or probably more, we will see about that. Now when I get that, how to do the analysis? I will come to that part later, but what is the problem if I give this kind of a thing?

One thing is that the strength is that it's mimic real world. In the real world also if you remember I told it in a previous video that in marketing we have something called consumer decision making process where when consumer starts decision making, they first has to fill the need and then when they think that okay I need in this case a soft ball, in this case a golf ball so I need a golf ball.

He will look for options, what kind of golf balls are available in the market. Let us say it can be any sports equipment that you want to buy, let us say a bat if it is a cricket match and then you go on and search options that are available. Now, certain things you have in your mind, so like I will buy from x, y, z brand or a, b, c brand, some of the brand names are always top of your mind.

Some are probably not so much at the top of your mind, the moment you Google search you come to know about them. So all of these things together it creates your awareness set. Awareness set means that okay, you know about these brands. Then you collect information and then you evaluate the alternatives available and there are various ways of evaluation and you create a subset from that awareness set which is called consideration set.

Now these are the guys whom I am considering, then I will further do further probing and etcetera and from there I would create further subset of them which is choice set. So from the choice set the choice comes. So sometimes from the choice set, you decide that okay, these are the 4 brands out of which I will buy any one of them. So I will say they are sometimes much smaller.

And then you go to the market and at the last moment you see you had probably choice A, B, C, D, you would knew that okay, choice A was better than B, C, so I might go and buy A and when you went to the market, you found out that in the market you got offer of B and then you bought B. So not always whatever is the preference that you have before you purchase will actually have during the purchase.

So sometimes but whatever be the case, the choice set will define what kind of choices that you are making. So we are also giving the choice set here and then the post purchase satisfaction and etcetera thing are there in the consumer purchase decision-making process. Now here in this choice based conjoint, we are giving you those choice set, some imaginary choice sets and asking you which one will you buy to see your preference between those choice sets, choices in the choice set.

So that is something which mimics the real world consumer behavior, that is why it is more I would say more suitable for marketing research, then it can investigate interactions. Now previous models were almost linear, it can investigate interactions and alternative specific effects.

So certain alternatives have certain specific effects which you can capture. It can be run both in paper and computer which is also adaptive one cannot be done in paper, it has to be run on computer. On the other hand, the individual waiting one can be run in computer, but it is better if you run it in paper so that people can see all the options together at one time. All the options together at one time in one screen becomes very difficult which you can give in a pen and paper kind of a format when it is a ranking or rating based.

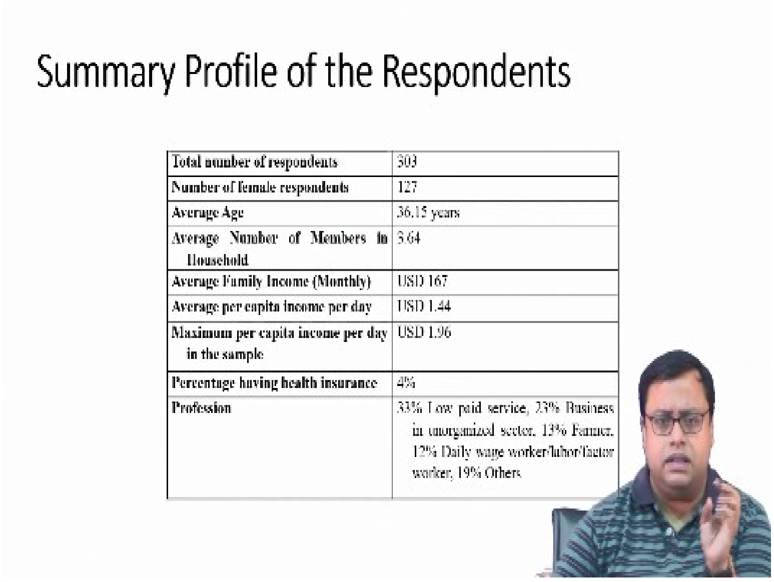
Now this one can be done in both ways, both the channel, paper channel that means paper pen and paper channel or computer based data collection both can be done and none of this is an option which was not an option in previous cases. None of this you can give as an option and multiple held-constant alternatives also. So these are some of the strains that you actually can bring in case of conjoint analysis which is choice based.

On the other hand for weakness, the first major weakness is you need is a large sample size. So remember, there one guy was creating around 16, 24 whatever was the full factorial design or half factorial design whatever that many data points. Here he is only creating the choice, one choice. So you need more data for that to create a bigger data set so that you can do analysis and the analysis method also asks for bigger data set.

Less attribute is ideal, the more attribute it becomes the more people start using heuristic methods. So less attributes is ideal, but attribute levels can be strong and then it may lead to heuristic based responses. If it is very complex task you have given then somewhat sometimes people do heuristic investments, they do not analyze it properly, they take shortcuts. So that can be one weakness and the analysis method is a little bit complex.

So when I say that, hold your thought, it will be a little bit complex than the previous analysis, it is not straight called logistic regression or linear regression, it is a little bit more than that. So hold your thought and we will see how complex it is when they are saying it is complex.

(Refer Slide Time: 10:36)



Total number of respondents	303
Number of female respondents	127
Average Age	36.15 years
Average Number of Members in Household	3.64
Average Family Income (Monthly)	USD 167
Average per capita income per day	USD 1.44
Maximum per capita income per day in the sample	USD 1.96
Percentage having health insurance	4%
Profession	33% Low paid service, 23% Business in unorganized sector, 13% Farmer, 12% Daily wage worker/labor/factor worker, 19% Others

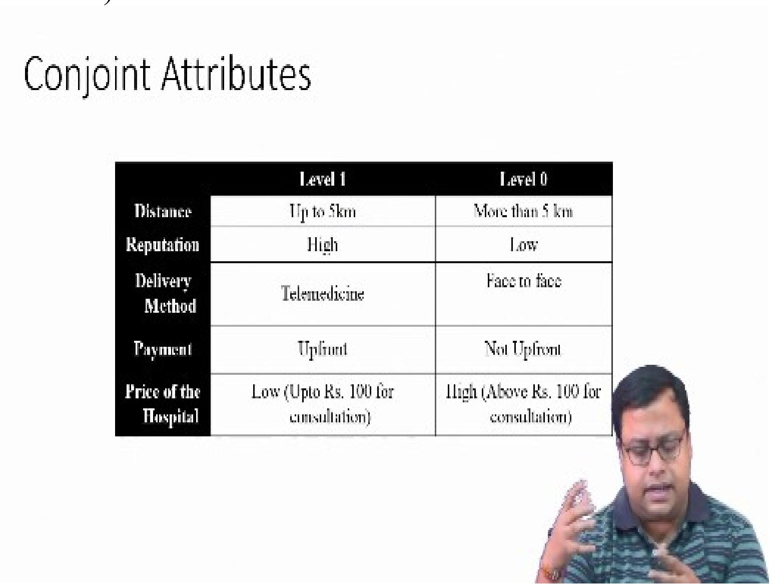
Okay so in this particular case what I have taken is, so I will talk about a research paper that we are walking on and it was on healthcare decision-making by urban slum dweller mania, urban bottom of pyramid customers. It was a special case which and the dataset is a subset of that, of the work that I am talking about. So there what we did is we actually wanted to know that we have to design a service situation for bottom of pyramid customers in a city called Bangalore mainly targeted towards the slum dwellers.

And we want to know that what kind of healthcare services they have and what kind of services I mean probably they want and so this was our sample size, there were 303 respondents. Out of them 127 were female, the average age was 36.15 years, average number of members in household in the family was 3.64, average family income was $167 monthly which turns out to be around 10,000 rupees monthly probably.

And then for average per capita income was $1.44, maximum per capita income per day was $1.96. So these are all bottom of pyramid customers and percentage having healthy insurance is only 4 percent, only 4 percent people out of this guy have certain kind of health insurance which is obvious if they are bottom of pyramid customers and at time the central government policies of health insurance for poor has not been introduced when the data collection was done.

And the profession wise, there were 33 % was low paid service, so some job 23 % was business in unorganized sector, 13 % farmer, 12 % daily wage worker and 19 % others. So these are the, so the data was collected for I would say from a government hospital where lots of people who are, I would say coming from weaker economic background generally go there and they sorted it out. Some doctors helped us in the data collection part.

(Refer Slide Time: 12:51)



Now what we have given as the choice to these guys were decided based on these five levels. So distance, reputation, delivery method, payment and price of the hospital were the five aspects and distance was up to 5 kilometer or more than 5 kilometer. So we did not write high or low. Now reputation is something which was not very constant in this particular case, not very concrete in this particular case, it was abstract because I do not know how people perceive reputation.

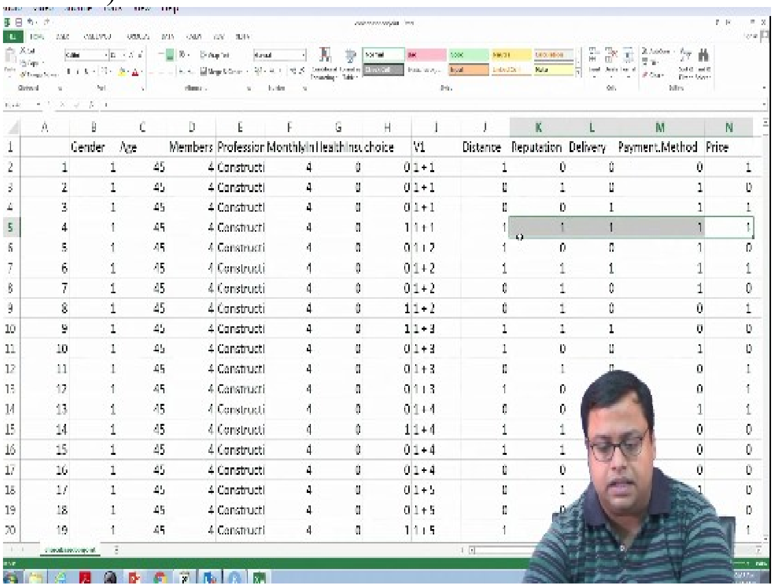
So we just kept it a little bit hazy which is high or low. In a case which is let us say for people who are not bottom of pyramid customers who are let us say middle-class or upper class customer who have access of internet, for them the reputation can be 4 star or 3 star or 2 star, this kind of rating that you get in websites called Practo let us say or some other Google reviews or something like that.

So there that can be one aspect, but here for bottom of pyramid customers we thought that they will not access of those kind of information. So for them high and low was the only thing and then delivery method is either telemedicine or so we are trying to create a new product, so we thought that telemedicine can be one option. So telemedicine or face-to-face which one they prefer.

So because we are creating new products, that is why we kept another option of pricing which is not upfront, upfront means you pay and you take the service, you go away. Non-upfront means you can pay through EMI or something else, small staggered payment services and price of the hospital was low or high. Low means up to 100 rupees per consultation and high means above.

So here this option is only for your OPD service not anything else, not for surgery and etcetera. So out of that so there are five attributes and two levels. So 2^5 that means 32, I hope I am right, 32 options that can be created. Out of that every customer were given 8 choices if I am not wrong, 8 choices, 4 attributes, that are four these things at a time, four choices at a time and they wanted to, they were asked to choose any one of them options and we were trying to see that how people are reacting to this. So I will come to this result.

(Refer Slide Time: 15:18)



The dataset look like this. If you have the files choice based conjoint dot CSB file, so the dataset looks like this. We have collected the data, then we have done the data input in such a way such that a certain library in R can handle the dataset. So the dataset will look like this, I will just put it on. So the first one is serial number, then gender, age, number of members.

And then profession of the person, monthly income, health insurance, the choice that they are making. Now this is something that I will come to that and then distance of the case reputation, delivery, payment method and price, something like this has been written here and then the level 0 and level 1 are here, so level 1 and level 0 accordingly we will understand which one is which level, fair enough.

And now you see that this is 1 + 1, 1 + 1 means first customer first choice set. Each choice set has 4 options, that is why there are 4 observations where 1 + 1 is there. Each customer has 4 options and out of those 4 options they are choosing one, they are not choosing the other 3, any one they are choosing. See, 1+1, there were 4 options, out of them they are choosing the fourth one; 1+2 that means first customer second choice set.

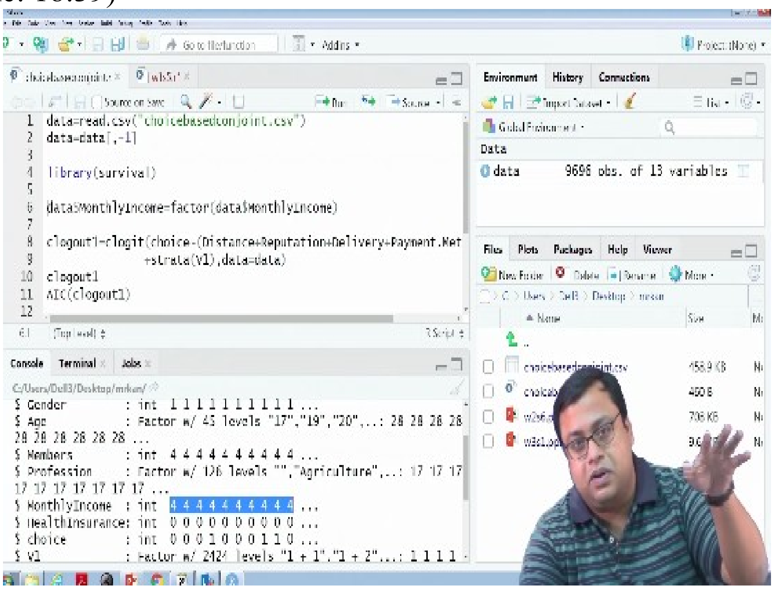
There are again four options, again he is choosing the last one. 1+3, there are four options, first customer third choice set; there are 4 options, now he is choosing the first one and how I can know which option? The option can be known by this. So this first choice set, first customer first choice set first option is nothing but this,that means distance level 1, reputation level 0, delivery level 0, payment matrix level 0 and price level 0.

And you can go here and find out what that means. So if you spend a little bit of time, you will understand what the data is. Again once more. This is 1 + 1, 1 + 2, 1+ 3 the first one means the customer number plus the choice set number. Each customer is getting eight choice sets, there are 32 options in total, eight choice set, each choice set has four options, so we are covering all 32 options.

That has been given to the customers, customers are seeing four options at a time and choosing one of them and again four options another time and choosing next one. Now what are those four options? The first option 1 + 1, this guy has chosen the fourth option because this one you see this one is 1, this one is 1 and then what did he choose? He chose actually out of these four options he chose the last one.

So 1, 1, 1, 1, 1 is something that he has chosen. Similarly, we have collected data of if I am not wrong around 303 customers.

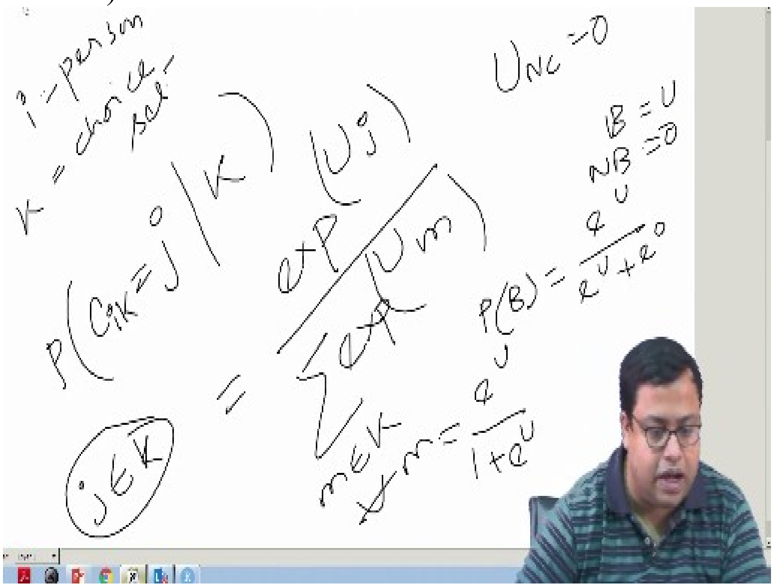
(Refer Slide Time: 18:39)



Now out of these 303 customers we are doing some analysis. So the first job is to set working directory, choose source file location and then read the data. If the data has 9,696 observations which is nothing, but 303 into 32 basically, each guy has 32 rows. Now I am removing the first one, serial number I do not need, so I am removing that. Then there is a library called survival, I am calling that library.

And if you see in my data if I just check my data, somehow the monthly income was integer and 4, so that was a categorical variable. We wanted do 10,000 to 15,000 or 15,000 to 20,000 and so on, so that has to be a factor variable, that we converted to a factor variable. Now what is my job? My job is to see that how people are reacting to this. So what I am doing is I am writing c log output. So here we are doing something called conditional logit.

(Refer Slide Time: 19:55)



What is conditional logit? Conditional logit says that the probability that somebody will choose the choice of let us say ith person from the kth choice set is j where i is the person and k is the choice set and j is the, j belongs to k, so j ∈ k actually, okay. Given the other choices or given the choice set is k given the choice set is k is basically exponential of utility from j divided by summation of exponential of utility of m where m ∈ k for all m.

i= person

k= choice set

$P(C_{ik} = j/k) = exp^{(uj)} / \Sigma\ exp^{(Um)}$

So you just multiply, so what do you do? You find out the utility of jth option and utility of all other options and the formula is like this. So exponential, so exp means exponential actually, exponential of U, or e^{Uj}. So that means that by chance if none of this is an option, for none of this which is also one choice set once, for that utility of no choice is equal to 0.

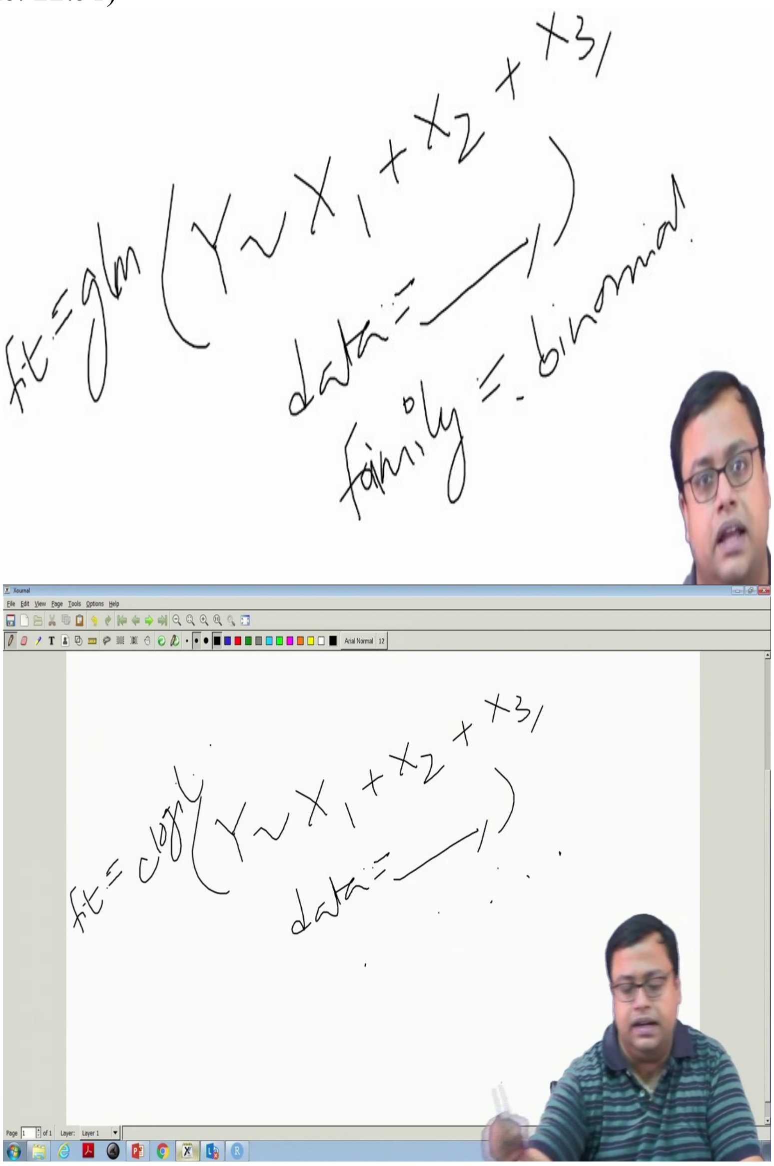
Now this actually mimics the logistic regression, normal logistic regression. In case of normal logistic regression, we have buy or not buy. So let us say buy option is U and not buy option is 0 then what will be probability of buying? Probability of buying will be $e^U /(e^U + e^0)$ comes up to be $e^U / (1+e^U)$ something like this, so that is logistic regression.

$P(B) = e^U /(e^U + e^0) = e^U / (1+e^U)$

It actually mimics logistic regression, so we have done that. This is called conditional logit, why conditional? So it is condition to the choice set. If the k changes, k changes to something else let us say $k^{/}$, the formula, the expectations and etcetera will change. So that is something that we did in this particular case. So I will run a c logit out which is equal to c logit, it's exactly same.

I told you before that most of our models in R where you are doing a predictive model will follow this function.

(Refer Slide Time: 22:51)



So for example for lm regression, you wrote lm(Y~X1+X2+X3), data is equal to data set name and this you wrote fit is equal to this. This is what you have done in case of normal regression linear regression.

fit= lm(Y~X1+X2+X3)

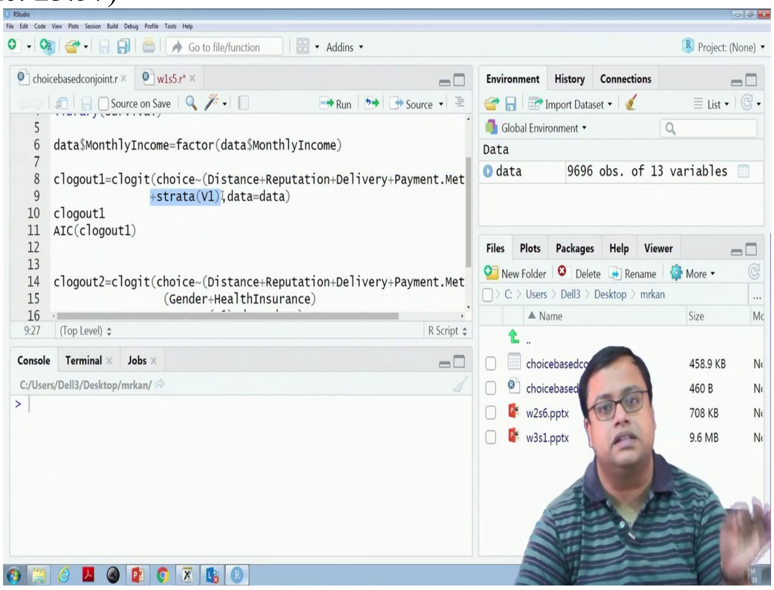
The moment I converted it to let us say logistic regression, it became glm, that is the only difference that we had and along with that I wrote family is equal to binomial or something like that.

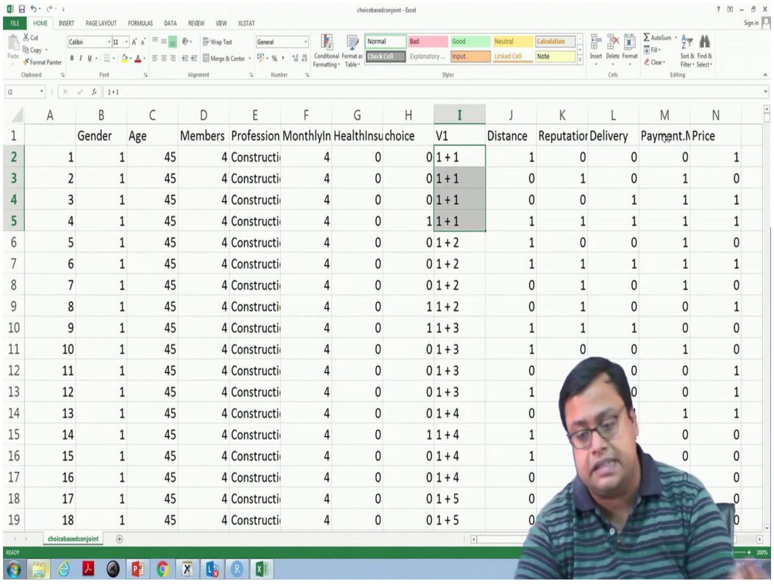
fit= glm(Y~X1+X2+X3)

So we have, I have wrote this, so that is where the extra I have written when it is logistic regression. Here instead of this and instead of this probably, this part is also not there. I am writing c logit that is the only difference that I have over this thing. So c logit and then (Y~X1+X2+X3) I have written that.

fit= clogit(Y~X1+X2+X3)

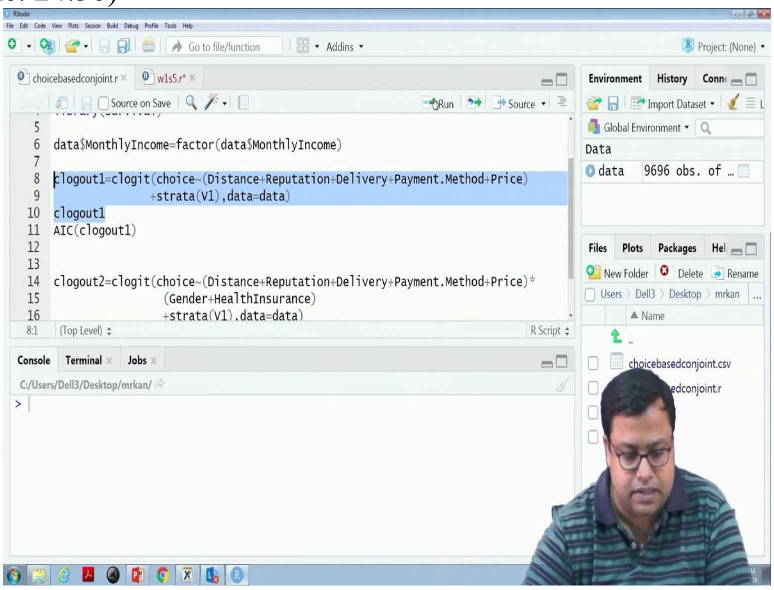
(Refer Slide Time: 23:57)

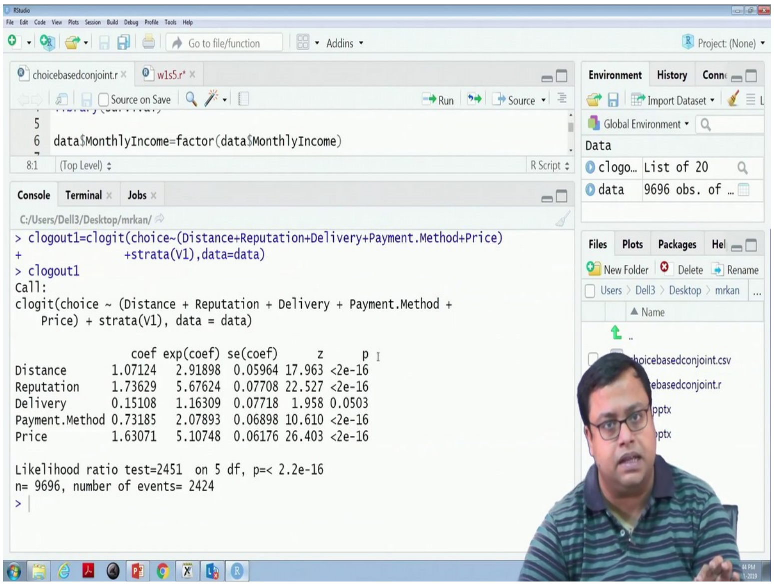




So, choice tilde distance plus reputation plus delivery plus payment method plus price. The extra thing that I am writing is strata is equal to v1, strata v1. What is strata? The strata is this particular column where I will come to know and this has to be written in this format 1 plus something, 1 plus something this v1 column has to be because, this c logit the survival library or the c logit function follows this pattern. That is why you have to write it in this way 1 + 1, 1+1, 1+1, 1 + 1 these are the 4 options, so this gives you the strata.

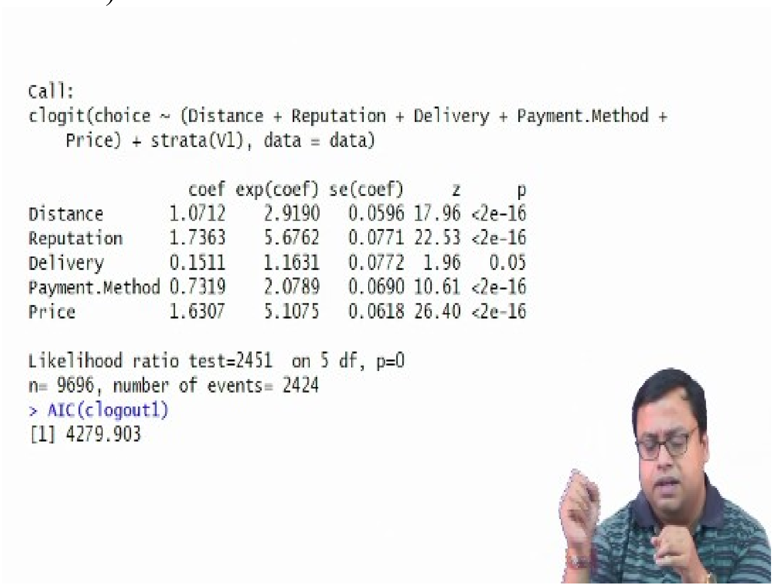
(Refer Slide Time: 24:38)





So I have done that and I will run this thing, the result comes out to be this. So let us see the result first. So this is the result that comes up and I will explain the result now properly.

(Refer Slide Time: 24:54)



So just running that I am getting something like this. So it is giving me this particular and the AIC is 4279. It is giving me that see all my p values other than delivery probably, I can still consider delivery to be significant, all guys have P value is smaller than 0.05. So that means each of my coefficient is significantly different from 0.

So if you do not understand what I am trying to say, in this part you have to go back and check linear regression what those P value, T value coefficient, the standard data etcetera mean. So here I have done that and coefficient is positive for all the cases. That means that distance goes up. If the distance increase, then, so not increase actually from distance level 0 to level 1 because we have, if you check the dataset you have almost coded as 0, 1; 1 means level 1, 0 means level 0.

So when I can change to level 0 to level 1, my overall utility increases. So e^{Ui}, that Ui is actually summation of beta x, that summation of beta x is something that I am talking about here.

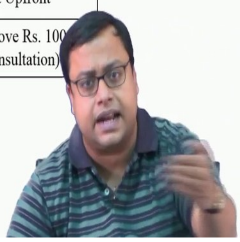
(Refer Slide Time: 26:25)

$$U_i = \beta_1 \times Dist + \beta_2 \times Pric + \beta_3 \times —$$

```
Call:
clogit(choice ~ (Distance + Reputation + Delivery + Payment.Method +
    Price) + strata(V1), data = data)

                coef  exp(coef) se(coef)     z      p
Distance       1.0712    2.9190   0.0596 17.96 <2e-16
Reputation     1.7363    5.6762   0.0771 22.53 <2e-16
Delivery       0.1511    1.1631   0.0772  1.96   0.05
Payment.Method 0.7319    2.0789   0.0690 10.61 <2e-16
Price          1.6307    5.1075   0.0618 26.40 <2e-16

Likelihood ratio test=2451  on 5 df, p=0
n= 9696, number of events= 2424
> AIC(clogout1)
[1] 4279.903
```

Conjoint Attributes

	Level 1	Level 0
Distance	Up to 5km	More than 5 km
Reputation	High	Low
Delivery Method	Telemedicine	Face to face
Payment	Upfront	Not Upfront
Price of the Hospital	Low (Upto Rs. 100 for consultation)	High (Above Rs. 100 consultation)



So what is the meaning of this thing? The meaning of this thing is that if there was a formula that I have given there then Ui is this whatever I have written here, Ui is something into distance plus something into price plus something, this something is let us say β_1, β_2 , β_3 and so on, that is what I get here.

$U_i = \beta_1$*Dist+ β_2*Repu+ β_3*...........

Now if my price level or distance level converts from level 0 to level 1, my coefficient is 1.0712 which is positive and significant, that means my utility increases by that much.

What does that mean? That means level 0, distance level 0 is you see more than 5 kilometer. Distance level 1 is up to 5 kilometer. That means as distance decreases level 0 to level 1 means distance is decreasing. As distance decreases, my preference, my utility increases. Fair enough, obvious. Can you interpret the next one reputation? Please pause it for one minute and go back thus in the previous slide and check what does the reputation say.

So if you have done that, you will understand the reputation 0 to 1, level 0 to level 1 preference increases. That means reputation low to high preference increases and that is probably the most important jump. So the Part-Worth of 0 to 1 that distance is more. So most important factor in that case is reputation. So reputation matters for everybody because it is a credence product, it is a healthcare.

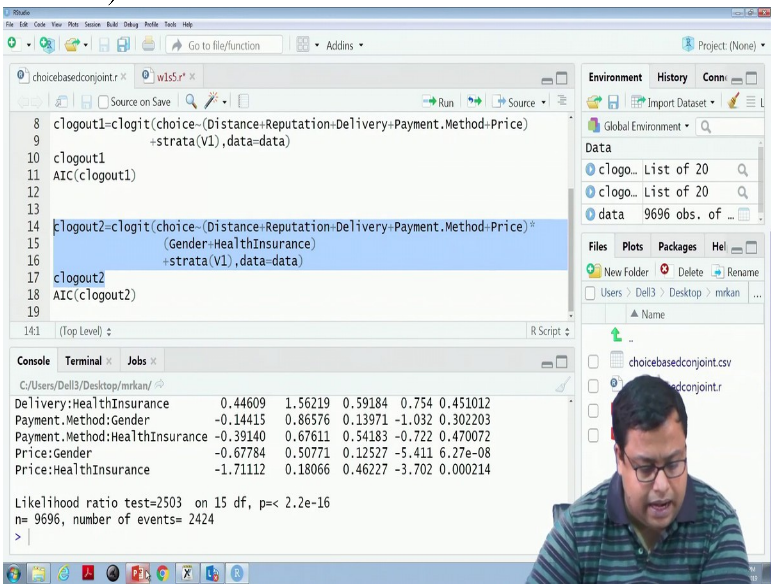
So healthcare is a credence product. If you do not understand what is credence, you can search probably, there is something called search good, there is something called experience good and there is something called credence good. Search good are those kind of products where you can search about the product before the purchase and you will know that how much is the quality, the quality perception happens before the purchase.

Experience is that quality perception happens only when you experience it, but healthcare is such a service where even after going to the doctor and getting the consultation from the doctor, you do not know whether the doctor has done the right job because you have little knowledge about that thing. Similarly is lawyer service. So this is not an experiential service, this is not a search product, you will not know about the quality of the service that you got.

Even after the service happened, this kind of services are called credence service and there external source of information, for example, reputation matters a lot. So that is why 1.73. Similarly, delivery, now delivery method I would say not upfront to face-to-face. Sorry delivery method is face-to-face to telemedicine. The preference is slight preference 0.15, that is significant or not significant, I do not know whether it is 0.05 higher than 0.05.

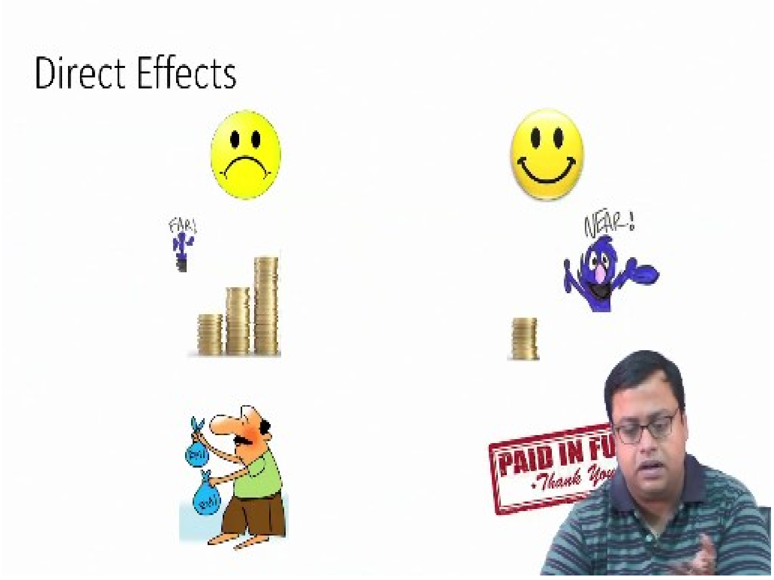
But whatever be the case the distance is not much strong and then what about the payment method? People actually is liking upfront payment. So they were preferring from level 0 to level 1, so this is something that is very important and then price obviously, they are preferring low price. From high, high is level 0, low is level 1, so they are preferring because 1.63. The most important factor is reputation and then price, so it makes sense till now.

(Refer Slide Time: 29:48)



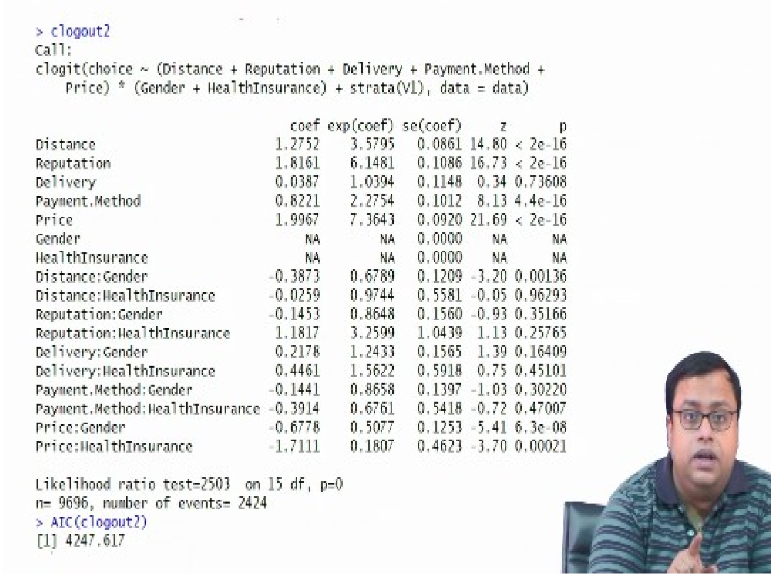
Then what we did is we have done a regression again. Now I have taken the moderating effect, so moderating effect means interaction terms and interaction terms is written by multiplication. So multiplication, gender and healthcare. So I want to see that when gender changes from 0 to 1 or healthcare insurance, availability of healthcare insurance changes whether thus preference change or not. So I run this, it will take some time because, okay so it did and I get actually….

(Refer Slide Time: 30:21)



So these are the some things that I already discussed when your direct effects is when you are happy, when you are not happy.

(Refer Slide Time: 30:27)

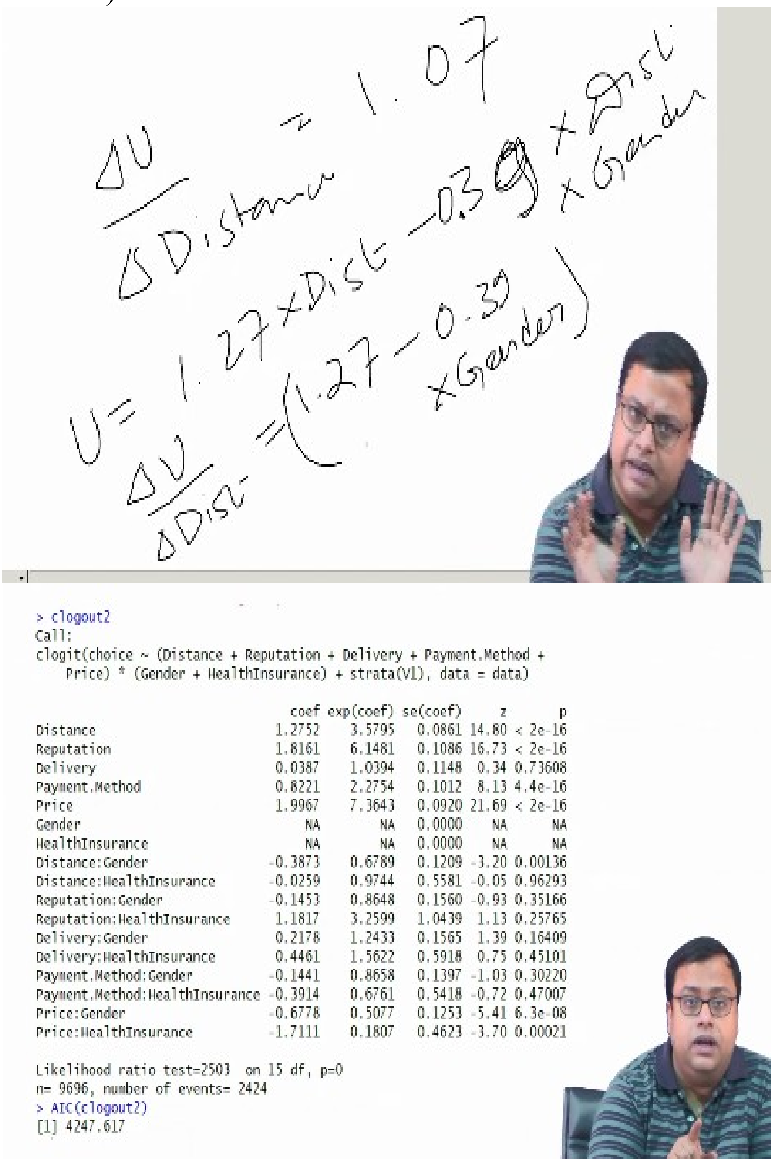


```
> clogout2
Call:
clogit(choice ~ (Distance + Reputation + Delivery + Payment.Method +
    Price) * (Gender + HealthInsurance) + strata(V1), data = data)

                                  coef exp(coef) se(coef)     z        p
Distance                        1.2752    3.5795   0.0861 14.80  < 2e-16
Reputation                      1.8161    6.1481   0.1086 16.73  < 2e-16
Delivery                        0.0387    1.0394   0.1148  0.34  0.73608
Payment.Method                  0.8221    2.2754   0.1012  8.13  4.4e-16
Price                           1.9967    7.3643   0.0920 21.69  < 2e-16
Gender                              NA        NA   0.0000    NA       NA
HealthInsurance                     NA        NA   0.0000    NA       NA
Distance:Gender                -0.3873    0.6789   0.1209 -3.20  0.00136
Distance:HealthInsurance       -0.0259    0.9744   0.5581 -0.05  0.96293
Reputation:Gender              -0.1453    0.8648   0.1560 -0.93  0.35166
Reputation:HealthInsurance      1.1817    3.2599   1.0439  1.13  0.25765
Delivery:Gender                 0.2178    1.2433   0.1565  1.39  0.16409
Delivery:HealthInsurance        0.4461    1.5622   0.5918  0.75  0.45101
Payment.Method:Gender          -0.1441    0.8658   0.1397 -1.03  0.30220
Payment.Method:HealthInsurance -0.3914    0.6761   0.5418 -0.72  0.47007
Price:Gender                   -0.6778    0.5077   0.1253 -5.41  6.3e-08
Price:HealthInsurance          -1.7111    0.1807   0.4623 -3.70  0.00021

Likelihood ratio test=2503  on 15 df, p=0
n= 9696, number of events= 2424
> AIC(clogout2)
[1] 4247.617
```

Now next one is this part. Now the distance if you see carefully, the distance again so it is still significant and then reputation is most important. Now after controlling for lot of this price has become most important. Now gender and healthcare, probably there were some correlation or something that is why they are coming in but what is more important to see here is this one. Let us say 0.00136, this one is significant.

And this is a negative, that means when this gender becomes 0 to 1, means for male gender the distance says effect comes down, minus 0.3873. Initially it was 1.2752, but it comes down when your gender is male, it comes down from 0.0, so what does this mean? This is the interaction effect.

(Refer Slide Time: 31:33)



So if I just write in the first model, in the first model if I just write delta utility / delta distance, so distance changes, how much utility changes? It was 1.07 or something.

$\Delta U / \Delta Distance = 1.07$

Now in the second model if you see carefully, the second model is basically utility is equal to 1.27 into distance minus 39, so 1.27 into distance minus 3.9 or 0.39 sorry, minus 0.39 into distance into gender keeping other things constant.

$U = 1.27 * Distance – 0.39 * Distance * Gender$

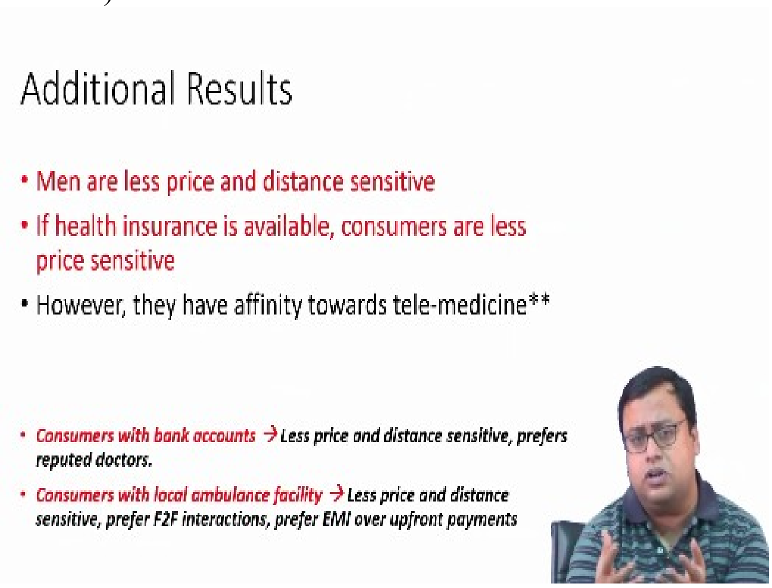
$\Delta U / \Delta Distance = 1.27 – 0.39 * Gender$

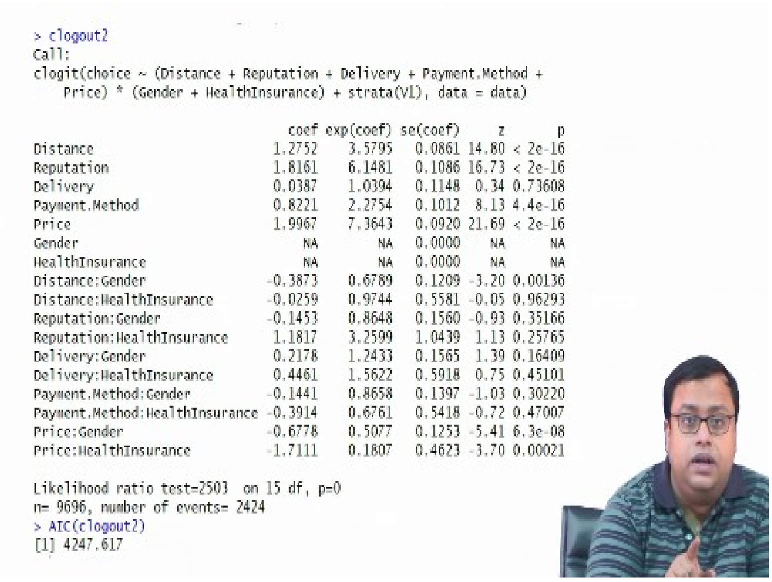
So then what is delta u delta dist? This is 1.27 minus 0.39 into gender. So when gender is 1 that means gender is male, delta u, delta dist is low. When gender is female, ΔU/ ΔDistance is high, what does that mean? That means male guys are less sensitive towards distance, female people are more sensitive towards distance, makes sense, I hope so. So which should not be the case which I do not think is something that should happen in country like India.

But this is a reality. Similarly, gender increases the price sensitivity drops; gender not increases from 0 to 1, gender changes from level 0 to level 1, that means your gender is female versus male. When you are male, your price sensitivity drops, this one is also significant and the last one is when you have health insurance then also your price sensitivity drops and it drops like hell, it drops like 1.7.

So here price is, utility is 1.99 from level 0 to level 1, but that is very different when you have health insurance or not health insurance though in our data set only 4 percent had health insurance. So this data was, this interaction effect is I do not know how much we can believe on this, but still that is something that is we should focus. 303 is 4 percent means 12 data points which is not significant enough, we should probe it in a little bit more in this area. But it gives you some hint that health insurance matters for people who are bottom of pyramid for their healthcare choices.

(Refer Slide Time: 33:54)





```
> clogout2
Call:
clogit(choice ~ (Distance + Reputation + Delivery + Payment.Method +
    Price) * (Gender + HealthInsurance) + strata(V1), data = data)

                                  coef exp(coef) se(coef)     z       p
Distance                        1.2752    3.5795   0.0861 14.80 < 2e-16
Reputation                      1.8161    6.1481   0.1086 16.73 < 2e-16
Delivery                        0.0387    1.0394   0.1148  0.34 0.73608
Payment.Method                  0.8221    2.2754   0.1012  8.13 4.4e-16
Price                           1.9967    7.3643   0.0920 21.69 < 2e-16
Gender                              NA        NA   0.0000    NA      NA
HealthInsurance                     NA        NA   0.0000    NA      NA
Distance:Gender                -0.3873    0.6789   0.1209 -3.20 0.00136
Distance:HealthInsurance       -0.0259    0.9744   0.5581 -0.05 0.96293
Reputation:Gender              -0.1453    0.8648   0.1560 -0.93 0.35166
Reputation:HealthInsurance      1.1817    3.2599   1.0439  1.13 0.25765
Delivery:Gender                 0.2178    1.2433   0.1565  1.39 0.16409
Delivery:HealthInsurance        0.4461    1.5622   0.5918  0.75 0.45101
Payment.Method:Gender          -0.1441    0.8658   0.1397 -1.03 0.30220
Payment.Method:HealthInsurance -0.3914    0.6761   0.5418 -0.72 0.47007
Price:Gender                   -0.6778    0.5077   0.1253 -5.41 6.3e-08
Price:HealthInsurance          -1.7111    0.1807   0.4623 -3.70 0.00021

Likelihood ratio test=2503  on 15 df, p=0
n= 9696, number of events= 2424
> AIC(clogout2)
[1] 4247.617
```

So this is what we found out. Men are less price sensitive if health insurance is available, consumers are less price sensitive. However, they have affinity towards telemedicine which was also seen here, no which was not seen here. So probably this was not so strong of an effect, that is why double star I have written. There were some other variables that we took and we have actually, I have not shown in this result like consumers who have bank accounts or not, we took one variable.

And we saw in the actual paper that less, there people who have bank account are less price and distance sensitive and prefers reputed doctors. Often times it happens that people who have bank accounts have some amount of financial capital. So they can actually handle all these things, a little bit of more distance they can go for doctor service or they can actually give a little bit of more price and etcetera.

And when there is local ambulance facility available, so this is a social capital that you can generate. If you can generate a little bit of social community kind of feeling where people can come together and have a local ambulance facility then also they have lesser price and distance sensitive and they prefer F2F, that is face-to-face interactions and prefer EMI over upfront payments. So here our major goal was that is it more necessary to create a new design of service in healthcare.

Or was it more necessary to see that what kind of things will make people less price and distance sensitive and they will prefer more reputed doctors and etcetera even when they are in the bottom of pyramid. So this is the case study and method was like that where conjoint analysis helped. So that is all for this particular video, we will continue with the new topic, fresh new topic called Customer Segmentation in the next week. Thank you for being with me.