## Introduction to Econometrics Professor. Sabuj Kumar Mandal Department of Humanities and Social Sciences Indian Institute of Technology, Madras Qualitative Response Models- Probit and Tobit Models Part - 1

(Refer Slide Time: 00:15)



So, let us discuss something about goodness of fit measure or R square in qualitative response model. Now, you have to keep in mind that the standard interpretation of R square which we derived earlier in case of linear regression model is no longer applicable in qualitative response model. So, standard R square makes little sense in the context of qualitative response model.

Why this is so? Because, in the case of linear regression model, what is the R square measure? Let us say this is your xi, this is your yi and let us say you have this type of scatter plot and the regression line alpha hat plus beta hat xi is basically a representation of your data. That means, this is some kind of average line which represents these data points.

And what is your R square? Your R squared is for any given value of xi, what is your predicted value? What is your observed value? Let us say this is your observed and let us say this is my yi bar. So, this is my total variation TSS, and out of which, I have explained this much, this is y I had, this is ESS and this is RSS and your R square you have defined as ESS by TSS.

So, out of your total variation in y which is measured by yi summation yi by y whole square total variation and how much you are explained, summation yi hat minus yi bar and then your R

squared is ESS by TSS. So, when you have a scatter plot like, so that means given xi your yi can take any value and you have this proper scatter plot and you have a regression line, which basically represents the scatter plot. And what is goodness of fit? How good my regression line is in representing this scatter plot, which is nothing but raw data.

But suppose we are in a context where this scatter plot itself does not exist. Then what you will do? That means this regression line itself makes little sense. we cannot get this type of regression line which we call average line or average representation of the scatter plot, then how will you take this RSS, ESS by TSS, it does not make any sense. And that is what is happening in the context of qualitative response model. How? Let us see here.

In this axis, let us say I am measuring xi and this is yi and this is 0, this is 1. So, that means for any value of x, you are either here or here that means your probability is here or here, either the event will happen or not happen. So, either the event will happen or not happen, depending on your x value, let us say that is income. So, that means there is no scatter plot and if that is the case, if the scatter plot does not exist in the context of qualitative response model, standard R square interpretation also is no longer valid.

So, what I am saying, since there is no proper scatter plot in qualitative response model, standard interpretation of R square is no longer valid because I cannot get a line like this. And unless I get that type of regression line, I cannot apply the ESS by TSS formula to get the R square. So, what we need to know, we need to derive alternative methods of R square and let us see what type of alternative model we can derive.

(Refer Slide Time: 06:40)

 $\frac{\left|\int_{R}^{R} \frac{d^{2}}{dt} \left( \frac{d^{2}}{dt} \right) \left($ 

So, alternative measures of R square in the context of qualitative response model. So, how will you modify this? Let us say that this is R square equals to ESS by TSS and that we can write as 1 minus RSS by TSS. Now, in the context of QRM, we need to get some kind of equivalent measure of RSS and TSS. And econometrician say that it is qualitative response model, then you have incorporated certain explanatory variable and your model is explaining some part of the total variation in y and the remaining portion is RSS.

In qualitative response model when you are estimating your coefficient using log likelihood method they go for several rounds of iteration to maximize different log likelihood functions. So, the first step in that iteration is when you have not incorporated any explanatory variable and if you do not incorporate any explanatory variable, then you see if you go to Stata's output, look at how Stata is estimating the model.

## (Refer Slide Time: 08:47)



If you put the Logit command, let us say this is my model. Now, see, this is iteration 0, 1, 2, 3, 4 and this is the final. That means iteration 0 means, in your log likelihood function, you have no explanatory variable only intercept and that is your log likelihood value. And what is iteration 4, you have included all your explanatory variables in the model. Similarly, iteration 1 means 1 explanatory variable, 2 means 2 explanatory variable, 3 explanatory variable.

So, this is your initial log likelihood value, this is your final log likelihood value reported here, -417.2214. So, out of these 4, 5 levels of iteration and 5 log likelihood value something represent

your RSS, something represent your TSS and we need to identify which one should be my TSS and which is my RSS. Obviously, when you have no explanatory variable in the model there is no question of explaining anything, there is no question of remaining anything that is the reason the log likelihood corresponding to iteration zero in Stata is reported as L0 and the final one is called L1.

(Refer Slide Time: 10:37)



So, this is defined as L0 and L1, when you have no explanatory variable that is L0, that means, that time it is L0 and this is L1. So, that means, this is your RSS; this is TSS. L0 means zero-ith level of iteration in terms of Stata's outcome and L0 is -514.87. minus 514 equals to minus 514.87 or something divided by, I made a mistake I think, this is, what is L1? L1 is this. Since both are minus, minus, minus will get cancelled. 417.22 and then this is 514.87. So, this would become your R square.

So, if you calculate this, then you will get your R square. And that R square is basically the R square reported here, look at this, this is called pseudo R square which is 0.1897. So, since a standard measure of R square is not applicable in the context of qualitative response model, econometricians have derived an alternative measure of R square, which is this ESS by 1 minus 11 divided by L0.

L1 is the value of log likelihood when all explanatory variables are included and L0 is the value of log likelihood when no explanatory variables are included in the model. This is one thing you have to keep in mind.

Then how do you check overall significance of the model? So, in standard econometric model we used to do by F statistic. But here, we get a different measure. If you estimate the model and then as I said you will get L1 and L0 and L0 is the initial value of log likelihood and L1 is the final value of log likelihood when you include all the variables in your model.

And if you calculate this (2 \*( L1-L0)), then you will get an equivalent measure of F statistic in the context of qualitative response model that is called log likelihood ratio, LR chi square or likelihood ratio statistic. And that actually follows as chi square distribution where degrees of freedom equals to number of explanatory variables in the model. So, this is also known as LR chi square. So, log likelihood ratio statistic.

So, that means calculated value of chi square is nothing but 190.30. And again, you have to compare this calculated chi square with the tabulated chi square at a specific level of significance and depending on whether this calculated value is lower or higher than the tabulated value, you have to reject or do not reject your null hypothesis.

And what is your null hypothesis here? In the context of overall significance as we all know, this is basically beta 1 equals to beta 2 equals to beta 3 equals to beta 4 equals to 0. This is the null. So overall significance of the model you will get by LR chi square statistic given by Stata. Look at now, LR chi square is 195.30 with 4 degrees of freedom and Stata is also reporting the corresponding v value which is 0.000 which means LR chi square is highly significant at 1 percent level. And this shows that your model is overall significant.

And another difference between the standard model and qualitative response model is you measure the significance of individual variable instead of t here we are getting z. What is z basically? z stat is the large sample counterpart of t. As I told you earlier that in the context of qualitative response model and the Logit model, the estimation technique is MLE, maximum likelihood estimation, which requires large sample.

Because in the large sample only all these desirable properties that means, estimates are asymptotically efficient, they are consistent and efficient at the large sample that is why MLE

requires a large sample and the large sample counterpart of t is actually the z statistic. There are other differences of t and z statistic which probably you can look at.

So, what Stata is reporting is the large sample counterpart of the t to test the individual level significance of the explanatory variable. So, this is how you can estimate the Logit model and you can interpret. Same way you can estimate the model using another qualitative response model that we have discussed earlier, which is called Probit.

(Refer Slide Time: 21:35)

History	7 4 X	educ	.0410274	.0037/ 4	.08 0.000	.02384	3 .058212	12.2869	<ul> <li>Variables</li> </ul>	
K Filter commands here	0	kids1t6	3070003	.04026 -7	.63 0.000	38590	9228092	0	Filter variables	here
# Command	rc.	husage	0189197	.00274 -6	.91 0.000	01997	2901355 7 031076	45.1208	Name	Label
1 use "E\Prof. Sabui\J)	Weelridge\MB_	exper	.0139013				.,	1010300	inff	=1 if in Jab free, 19
2 ssc install margeff									hours	hours worked, 197
3 logit inlf educ kidslt6	husage	. logit inlf	educ kidslt6 h	nusage exper					kidslt6	# kids < 6 years
4 mfx		Iteration 0:	log likeliho	ood = -514.	8732				kidsge6	# kids 6-18
5 logit inlf educ kidslt6	husage	Iteration 1:	log likeliho	ood = -418.1	2083				936	woman's age in yr
6 margeff		Iteration 2:	log likeliho	ood = -417.2	2307				educ	years of schooling
7 logit inlf educ kidslt6	husage	Iteration 4:	log likeliho	od = -417.	2214				wage	est wage from ear
8 mfx compute at (kicks	slt6=0) 101								repwage	rep. wage at interv
9 mfx compute,at (kids)	sk6=0)	Logistic reg	ression		N	lumber of a	obs =	753	hushrs	hours worked by h
10 logit inlf educ kidslt6	husage exper				p	rob > chi	2 =	0.0000	husage	husband's age
11 mtx		Log likeliho	od = -417.2214	1	P	seudo R2		0.1897	4 histordian	hisband's wars of
12 mfx compute, at (kids	8115=0 100								Properties	
13 mfx compute, at (kids	k116=0)	inlf	Coef	Std. Fre.	, p		[95% Conf.	Intervall		
14 logit init educ kidsite	husage exper		coer.	Jea. 211.	. ,	/1+1	[Post cont.	Incerverj	4 Variables	
		educ	.1838658	.0394628	4.66 0	.000	.1065201	.2612115	Name	husage
		kids1t6	-1.375834	.1977922	-6.96 0	- 000	1.763499	9881681	Label	husband's age
		exper	.1142859	.012/366	-0.05 0 8.91 B	.000	.0891409	.139431	Format	5,90a
		_cons	1.032398	.7569268	1.36 0	.173 -	.4511515	2.515947	Value label	ang
									Notes	
									4 Data	dala d
100									<ul> <li>Filename</li> </ul>	MROZDIA
	-								Label	
1	1C	Command							Notes	
	2	probit inlf e	educ kidslt6 hu	sage exper					Observations	22
										36.779
2 - Sacu SE 160 - El/Ar The Edit Data Graphi	ef. SabujU Woolnidge/MRDJ icis Satujicis Liter M	ZDTA Indow Help								-
2 - Sacu SE 160 - EPHo File Edit Data Graphi P C C C C C C C C C C C C C C C C C C C	ef. SabujU Woolniege/MRD3 kas Satustas User W B • B B B • • •	ZDTA lindow Help								-
E 2-Sana/SE 163-ES/Pr Tile Edit Data Gaphi © C 2 Control Control Tile Edit Data Caphi Processor	el. Sabul U Musicipel (MIC) ics Satistics Ure W IN D D O T Z X	201A Bindow Help © Probit regres	ssion		N	lumber of a	obs =	753	* Variables	-
2 - Sana St Hol - Elym     Te Edit Das Geges     H © @ # + Li - Hotory     Her connants ter:	el. Sabul/ Woolidge/MHD) ics. Sattisics. Dier W in P. D. D. O T. 3. x	ZDTA Bindow Help © Probit regres	ssion		N L	lumber of ( R chi2(4) Prob > chi1	obs = ) = 12 =	753 193.76 0.0000	 * Variables * Fitter variables	-
2-Suna KE Mida - Elipho Tel: Eda: Data Graphi El: Bita Data	ef. Sabel/UMbolidge/MMD2 iss Satisfics liber W i • • • • • • • • • • • • • • • • • • •	201A Brobit regree Log likelihoo	ssion od = -417.99285	5	N L P P	Number of ( R chi2(4) Prob > chi Pseudo R2	obs = = :2 =	753 193,76 0.0000 0.1882	 Variables     Variables     Variables     Name	- I Label
2 - StanyOf 163 - Elyher     The East Data Graph     P     P	ef. Sabuji Wookdge/MRO3 ics Satistics Der W V 3 × V 3 × Wookidge/MR.	ZOTA Indow Hop Probit regre Log likeliho	ssion od = - <b>417.9928</b> 5	5	N L P P	lumber of ( R chi2(4) Prob > chi2 Seudo R2	obs = = 12 =	753 193.76 0.0000 0.1882	 ▲ Variables ▲ Fitter variables Name inff	here Label =1 if in lab free, 15
Constant of the second se	ef. Sahuji Wookidge/MODi es. Satinos ber W es. es. es. es. es. es. es. es. es. es. es. es. es. es. es. es.	ZDTA Indow Hop Probit regre Log likelihor inlf	ssion od = -417.99285 Coef.	s Std. Err.	N L P P Z P	iumber of ( rob > chi >seudo R2 > z	obs = = = 2 = [95% Conf.	753 193.76 0.0000 0.1882 Interval]	 Variables     Variables     Nore     Name     inf     hours	here Label = 1 if in lab free, 19 hours worked, 197
2 - Sarayi Hida - Elyer, The East Data Gapta     Hele and Annual Hida - Elyer, Hele and Annual Hida - Elyer, Hele and Annual Hida - Elyer, Stabul) 1     Sec Stabul) 1     Sec Stabul) 1     Sec Stabul) 1	ef. Sakuji Waoiridge/MRD. s Sakušis Der W S Sakušis Der W S Sakušis Der W S Sakušis Der W S Sakušis Der W Woolridge/MR. ihusage	ZDIA Bodow Hop Probit regre: Log likeliho inlf	ssion od = -417.99285 Coef.	s Std. Err.	N L P P Z P	iumber of a R chi2(4) Prob > chi Pseudo R2 Pseudo R2	obs = = = = = = [95% Conf.	753 193.76 0.0000 0.1882 Interval]	 Valables     Valables     Valables     Valables     Name     Infl     hours     Kidd85     Kidd85	ture Label =1 if in lab fore, 19 hours worked, 197 # kids < 6 years
2-Sun (H A) - Elyh      Te Est Dia Gapia      E di a Dia Gapia      Estato      Command      une "Stynot Subgl // 2 social margit      Sogia Hacksidati      Tet      Interesting a social margit	ef. Saku/U Weskdor/MED3 es. Szetiső iber W B B B B O - Y Z X O Jz Weskidge/MR. ihosage	ZDIA Bodow Hop Probit regree Log likelihor inlf educ	ssion od = -417.99285 Coef. .1993784	5 5td. Err. .0231218	I I I I I I I I I I I I I I I I I I I	iumber of a R chi2(4) Prob > chi Secudo R2 >> z    \$.000	obs = ) = 2 = [95% Conf.	753 193.76 0.0000 0.1882 Interval] .1546963 .4647643	<ul> <li>Valables</li> <li>Filter valuties</li> <li>Name</li> <li>inif</li> <li>hours</li> <li>kidatis</li> <li>kidatis</li> <li>kidatis</li> </ul>	tabel Label =1 fir lab free, 19 Not six worked, 197 k kick < c k years # kick 5-18
2-Snavit Kol-Eller      2-Snavit Kol-Eller      Tee fat: Das Graph      2-Snavit Kol-Eller      Tee command      Tee command      Command      Snavit References      Command      Command      Command      Command      Snavit References      Command	ef Salvý BholdydMCO ka Sanis Ler W I - D D O - T I X WoldgMR. Hvage	ZZTA indow Hop Probit regre: Log likeliho inlf educ kidslt6 husae	ssion od = -417.99285 Coef. .1093784 .8308753 .431716	5td. Err. .0231218 .115219 .0874651	к П П П П П П П П П П П П П П П П П П П	iumber of 6 R chi2(4) Prob > chi seudo R2 >> z    9,000 9,000	obs = = 2 = [95% Conf. .0640666 -1.0567	753 193.76 0.0000 0.1882 Interval] .1546963 6650503 805548	 ▲ Vaddes Nerrorate Infl Horrs Holge6 age edw	here [Label =1 if in Jab free, 19 hours avoided, 197 # Mode 6-10 worman's age in yrr worma's age in yrr
2 - Source SE Mide - Eighter     The East Data Design     The East Data Design     The East Data Design     The Command     The Command     The Command     Source State Provide     Source Sta	Shall () Wonkey/MICL     set Same or W	ZDTA © Probit regre: Log likelihor inlf educ kidiltő husage exper	ssion od = -417.99285 Coef. .1993764 8308753 65301716	5 5 5 6 6 7 15 7 15 7 19 7 4 7 19 7 8 7 8 7 8 7 8 7 8 7 8 7 8 7 8 7 8 7	z P 4.73 e -7.21 e -6.65 e 9.41	iumber of 4 R chi2(4) Prob > chi; seudo R2 >> z    0.000 0.000 0.000	obs = ) = 2 = = [95% Conf. .0640606 -1.0567 0651821	753 193.76 0.0000 0.1882 Interval] .1546963 .695583 .885548 .885548	 ▲ Vaskbles ▲ Farrenzable Name nil kons ködös ködös app edac app edac	tere Label =1 fil hilds free. 19 # kids < 6 yans # kids < 6 yans wernari Jage Ih yn years of clockring years of clockring
2-Stautif 160-Elph     Terrard     2-Stautif 160-Elph     Terrard     Ter	C Sale J Monole (SMC)     S Stands for M     S     S Stands for M     S     S Stands for M     S	ZDIA Bactor Hop Log likeliho inif educ kidslife husage exper- _cons	ssion od = -417.99285 Coef. .1993784 8180733 65111716 .65599393	5 5 5 60231218 .115219 .0065999 .4533481	Z P 4.73 0 -7.21 0 -6.85 0 9.41 0 1.48 0	iumber of a R chi2(4) Prob > chi seudo R2 >> z    0.000 0.000 0.000 0.000 0.000	obs = = 2 = = [95% Conf. .066406806 -1.0567 06521821 .2186067	753 193.76 0.0000 0.1882 Interval] .1546963 085482 0558482 1.558485	 <ul> <li>Variables</li> <li>Variables</li></ul>	here Lubert = 1 fin Jub Inte, 15 Hours worked, 159 H Kick 6 years # Kick 6 years work 3 kick 6 years work 3 kick 6 years work 3 kick 6 years work 9
2 - Servel 163 - 109-     Tre Sate Date Output     2 - Servel 163 - 109-     Tre Sate Date Output     10 - 109-     109-	4         5.64/J Workdy/MROL           sis Stands         Der W           III - DD         O           III - DD         III           III - DD         IIII	ZDIA indow Hop Probit regree Log likelihor inif educ kidslt6 husage eaper cos	ssien od = -417.99285 Coef. .1093784 8380733 659915 .6699393	5td. Err. .0231218 .0074651 .0069998 .4533481	z P 4,73 0 -6,85 0 9.41 0 1.48 0	iumber of ( R chi2(4) rob > chi seudo R2 > z    0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000	obs = ) = 2 = = [95% Conf. .0640606 .0521821 .2186067	753 193.76 0.0000 0.1882 Interval] 0565963 0555402 0756209 1.558465	 ➤ Vaddis Varddis Vare Vare Vare Vare Vare Vare Vare Vare	Inter Label = 1 fit in kich tete, 19 4 kick - 64 yance 4 kick - 64
2 - Stand Kid- Other     in the final data of the second data     in the second data of the second data	C Sabu / Monitor / MODI     ss Standar L bor 10         T & X X         T         T & X X         T         T & X X         T         T         T	201A Docume Hop Probit regree Log likelihor inlf educ kidslt6 husage exper cons - mfx	ssion d = -417.99285 Coef. .1093784 -851073 -8511716 .6699393	5 5td. Err. .0231218 .115219 .0074651 .0069998 .4533481	I P P P 4.73 0 -7.21 0 9.41 0 9.41 0 1.48 0	iumber of ( .R chi2(4) rob > chi seudo R2 >> z    .000 .000 .000 .000 .000 .000 .000 .000 .000 .000	obs = = = = [95% Conf. .0640606 -1.0567 06588 .06521821 .2186067	753 193.76 0.0800 0.1882 Interval] .1546963 695983 085485 .0755209 1.558485	 ★ Vasibles ▼ statutes Nare statutes Ril locri Bill born Bill Bill Bill Bill Bill Bill Bill Bil	htte [Label] = 1 fil hib http: 19 # labd < 6 yas boars worked. 197 # labd < 6 yas wormark 3 gait ya di schooling est, wage fram can eys wage di schooling tray wage di schooling hours worked by h hours worked by h
2 - Secol 163 - 169-     if is to Bo O Over     if is to Bo O Over     if is to Bo O Over     if is to Bo Ove	A Child J Wooledy MRCC so Status Dar W - T T X -	ZDTA indow Hop Probit regre: Log likelihor inlf educ kidslfor husage exper- cons . mfx	ssion od = -417.99285 Coef. .1093764 -8309733 .6699333	5 5td. Err. .0231218 .115219 .0074551 .0069998 .4533481	Z P 4,73 0 -7.21 0 -6.53 9.41 0 1.48 0	iumber of i R chi2(4) seudo R2 >> z  .000 .000 .000 .000 .000 .000 .000	cbs = ) = 22 = [95% Conf. .0640606 -1.0567 06588 .0521821 .2186067	753 193.76 0.0000 0.1882 Interval] .1546963 .085540 2.085540 2.695620 1.558485	 Validade     Validade	ter Lubel to ar worked. 197 # lub < 6 years # lub < 6 years years of aboling est, wage from can rep, wage a district hours worked by National sage Mathematics and Mathematics and Mathematics Mathema
2 - Search 443 - EPA 10 - Command 10 - Com		LCEA Indow Hep Probit regre: Log likeliho inlf educ kidiltő hutage exper cos - mfx Harginal eff Probit regre:	ssion od = -417.99285 Coef. .1993784 8108753 .6599393 icts after prob	5 5td. Enr. .0231218 .115219 .0074651 .4533481	N L P P P 4.73 0 -6.55 0 9.41 0 1.48 0	iumber of ( R ch2(4) rob > ch1hi seudo R2 > [2] .000 .000 .000 .000 .000 .000 .000 .0	obs = = 2 = [95% Conf. .0640606 .06521821 .2186067	753 193.76 0.0000 0.1882 Interval] .1546963 0855482 0555482 1.558485	 ➤ Vaddes ▼ Vaddes Name Name Notes Notes Name Name Name Name Name Name Name Name	Inter Label = 1 if in lab lots, 15 # lish < 5 years = 16 shocking 10 years of shocking years of shocking to shocking year hours worked by 1 hours worked by
J-Seculit 16-100-     More table to device the secular test of the secular test o	45.64-()         WorkstynMECO           as Stands for W         0 - 4           F         0 - 7         2 ×           Weckley/MEC         -         -           *huage         -         -         -           *huage         -         -         -           *huage         -         -         -           *huage         -         -         -	201A Probit regres Log likeliho inif educ kidilő huisage exper cons - mfx Harginal effi y = Pi	ssion od = -417.99285 Coef. .1093784 8080735 .6093933 acts after prob (silf) (prediction S28239552	5 5td. Err. .0231218 .115219 .007651 .00669998 .4533481 oit tt	N L L P P P P P - 7.721 0 9.41 0 9.41 0 1.48 0	iunber of ( R ch2(4) rob ) chinh seudo R2 t>[2] 2,000 .000 .000 .000 .000 .000	obs = = = = [95% Conf. .0640606 -1.0567 .06521821 .2186067	753 193.76 0.0000 0.1882 Interval] .1546963 695638 6956209 1.558485	 Vankiles     Vankiles	Inter Label a f if is bis http: 15 hours nooled, 159 # kick = 6 years # kick = 6 years hours nooled by 1 hours nooled by 1 hour
Source 163 - 150-     The Set Data Capital     Set Data     Set Data Capital     Set Data     Set Data Capital     Set Data     Set     Set Data     Set	4.5 key JWooldey MBCO           ss Statis	ZDTA Indow Hop 0 Probit regre: Log likeliho durage educ kidsit5 hurage educ kidsit5 hurage educ kidsit5 hurage educ spar 	ssien od = -417.99285 Coef. .1993784 .8580753 .6699393 iects after prob r(ialf) (predi .5823955	; Std. Err. .15219 .007461 .0065998 .4533481 .011 tt	N L P P 7.73 0 4.73 0 -7.21 0 1.45 0 1.45 0	iunber of A ch2(4) A ch2(4) rob > ch1 seedo R2 >>[1] \$,000\$,000\$,	cbs = = = (95% Conf. -1.0567 -65382 -2.186467	753 9.0600 0.1882 154953 .6935485 .955485 1.558485	 × VacUales	Inter [Lade] =1 fit hab here, 19 hours worked, 197 # kick < 5 years # kick < 5 years wormark 3 gait price years of chooledge hours worked by h hours worked by h hours worked by h
2 - Stankel Kell - Office     1 - Stankel Kell     1 - Stankell     1 - Stankell     1 - Stankell     1 - Stankell     1 - S	41 Sabaj 2 Mulokaj AMICO           11 Sabaj 2 Mulokaj AMICO           12 Sabas         10 m           11 Novel         10 m           12 Sabas         10 m           12 Mulokaj 2 MICO         10 m           12 Mulokaj 2 MICO         10 m           14 MICO         10 m           14 MICO         10 m	201A Indow Hop Probit regree Log likeliho inlf educ kidslifo huisage esper cons . afx Marginal eff, y = P veriable	ssion d = -417.99285 Coef. .1993764 .6959152 .6959393 ictis after produce r(is14) (predic .5829552 dy/dx 51	5 Std. Err. 40331218 .0005998 .0074651 .0005998 .0074651 .0005998 .0074651 .0005998 .0074651 .0005998 .00746510	N 4,73 @ -7.721 @ 9,41 @ 1.455 @ 2. P> z	iumber of i R ch2(4) seudo R2 seudo R2	obs = = 2 = = [955 Conf. 654805 	753 193.76 0.0000 0.1882 154652 0.9563482 0.9563482 0.9563485 1.558445 1.558445	 ✓         Valdes           Valdes         Vale           Name         Intervention           Intervention         Nors           Kidopé         apr           able         able           Name         Intervention           Name         Intervention           Name         Intervention           Name         Intervention           Name         Intervention           Name         Intervention	los (del = 1 fit lab lice, 19 = 1 fit lab lice, 19 = 4 lab, 6 - 9 = 9 control, 19 - 19 = 9 control, 19 - 19 = 10 - 10 - 10 =
2 - Secol 163 - 169- 16 -	4.5bcl / Worksdy-MMCD           ss Stands         Der W	ZZTA. Holow Hop Probit regre: Log likeliho: inif educ idilfo inif educ idilfo y = Pi y = Pi y = Pi educ	ssien od = -412.99245 Coef. .3093764 -8508753 .6659933 Acts after prob (s147) (predic .5429952 dy/dx 51 .0427617	5 Std. Err. 40321218 40732118 4074511 407451 407	N L P P P 4,73 0 9,41 0 1,48 0 2 P <sub>2</sub>  z  2 P <sub>2</sub>  z  0,000	iumber of 1 R ch2(4) seudo R2 seudo R2	005 = = = = [95% Conf. -1.0667 -0.65883 -2.218667 	753 159.76 9.182 Interval] 1.558953 655958 655958 1.558485 1.558485 X 1.258485	 A Vanidate	her Label = 1 fit bib too. 19 = 4 bib too. 19 = 4 bib + bib too. 19 = 4 bib + bib = 10 bib + bib + bib = 10
2 - Servel 143 - Difference 144     2 - Servel 144	An and a second	201A monw Mop Probit regree Log likeliho kinif kinif hulus huper 	ssion od = -417.99285 .093764 .093764 .8980733 .6839735 .6839735 .6659933 icts after prob r(inif) (predic .58239552 dy/dx 51 .6427917 -3243762	; Std. Err. .03212 8 .097451 .15219 .097451 .453481 .453481 .51t .t) .td. Err.	I         I           I         I           P         P           I         I           I         I           I         I           I         I           I         I           I         I           I         I	iumber of A R ch2(4) Yeb > ch1 >izi >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >izi >ieudo R2 >ieudo R2	obs         =           2         =           [935 Conf.         -	753 153.76 0.0000 0.005563 1550563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.095563 0.09556 0.09556 0.0055 0.0055 0.0050 0.0050 0.0055 0.0050 0.0050 0.0050 0.0050 0.0050 0.0055 0.0050 0.0050 0.0050 0.0055 0.0050 0.0055 0.0050 0.00555 0.00555 0.00555 0.00555 0.005	 <ul> <li>✓ Valides</li> <li>✓ Encrosolite</li> <li>✓ Encrosoli</li></ul>	The control of the second seco
J - Securit 163 - 100-     More tasks to account of the task to	4f         Shekaja (Millo)           as         Stanis           bil         0           r         0           r         0           r         0           r         0           r         0           r         0           r         0           r         0           r         1           r         1           stage         1	201A hotew hep Probit regre: Log likeliho inif educ kidsit6 husage exper oos - mfx Harginal eff; y = P y = P variable educ kidsit6 husage educ hus	ssion d = -417.99285 Coef. .899754 -8589753 .6959784 .6959785 .6959983 with after production (s17) (prediction s2239552 dy/dx 51 .0427017 .3427075 .3427075	5 Std. Err. .0059998 .0069998 .4533481 .0069998 .4533481 .0069998 .4533481 .0069998 .4533481 .0069998 .4533481 .0069998 .4533481 .0069998 .4533481 .0069998 .0069998 .0069998 .0069998 .006998 .006998 .00698 .006988 .0069888 .0069888 .0069888 .0069888 .00698888 .0069888 .0069888 .0069888 .0069888 .0069888 .00698888 .0069888888 .006988888 .00698888 .00698888 .0069888 .006	R L L P P P 4,73 6 9,41 6 9,41 6 9,41 6 1,48 0 2 2 7,73 0,000 1,48 0 0 1,48 0 0 5 9,60 0 1,50	iunber of ( R ch2(4) rob > ch1 rob > ch2 rob = ch2 rob	005 = = = [933 Conf	753 159.76 9.766 9.1882 Interval] 1.559693 	▲ Valuation     ↓     ✓ Valuation     ↓     ✓ Particulation     ↓     ✓ Particulation     ↓	here Label = 1 if it labition; 19 High - 1 is labition; 19 = 1 id labition; 10 = 1 id la
Source 161 - 160-     To See Date Output	Ar (Sted) Wester (MOC) as Status for W 	2000 top 0 Proble regre ishelloo ishelloo regre ishelloo regre ishelloo ishel	asien od = -417.99285 Coef. .1993784 8194753 .4639378 .4639378 .6639333 icts after prob r(inif) (predic .5823552 dy/dx 51 343762 3193775 .03597283	; Std. Err. 40321218 409745511 409745511 40974551 514 1 40590 - 4 40591 - 6 40591 - 6 40591 - 6 40591 - 6 40591 - 7 40591 - 7 40	E L L L L L L L L L L L L L L L L L L L	umber of A R ch2(4) reb > ch1 seudo R2 r>[2]   .000 .000 .000 .000 .000 .000 .000 .0	bbs =	753 153.76 9.0600 9.1882 155495 1.559485 1.559485 1.559485 1.559485 1.559485 1.559485 2.37716 2.37716 2.37916	A Valdels	Inter Label I abel I dick of the the to I dick of the to I dick of the to when the total dick I dick of the total I dick of tota
Second 16-16-16-     The test the owner     Second 16-16-16-     The test the owner     Second 16-16-16-16-     The test the owner     Second 16-16-16-16-16-     Second 16-16-16-16-16-     Second 16-16-16-16-16-16-     Second 16-16-16-16-16-16-     Second 16-16-16-16-16-16-16-     Second 16-16-16-16-16-16-16-16-16-16-16-16-16-1	41         Scholdy (Weinker) (MEO)           100         100	200 more top Proble regree Log likelihoo lidslife educ ons - wfx Parginal eff Parginal eff y = p - veriable keinis boxage eoper	ssion d = -417.99285 Coef. .899776 .899776 .699939 acts after prob (s17) (predic .82329552 dy/dx 51 .427077 .329775 .0257283	5 516.Err. .0051218 .0065998 .0056999 .4533481 .0069998 .4533481 .0069998 .4533481 .0069998 .4533481 .0069998 .006997 .5 .0069998 .006998 .0069998 .006998 .006998 .0069998 .00698 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .006988 .0069888 .0069888 .0069888 .006988	N         L           μ         P         P           σ         σ         σ         σ           σ         σ         σ         σ         σ           σ         σ         σ         σ         σ         σ           σ         σ         σ         σ         σ         σ         σ         σ           σ	Lumber of ( R chi2(4) >> z    >.000 .000 .000 .000 .000 .000 .000 .0	obs         =           2         =           2         =           933 Conf.	753 193,766 0,1882 11547963 1,1549563 0,955429 0,955429 0,955429 1,553485 1,553485 1,553485 1,553485	 ✓         Valutions           Valutions         Name           Name         Name           Nons         Kologofi           age         edic           Valutions         Kologofi           age         edic           Name         Name	Inter Label at file label loss, 19 Anter State label loss, 19 Al label 4, 19 Al label 4, 19 Al label 4, 19 Anter State Jack 4, 19 Anter State 4, 19 An
Support (4) - 100- Support (4) - 100- S	4. Shelly Wooksdy-MMCD           cs Stands         Der W           e         Der W           b         Der W	ZDA homov big 0 Probit regre ligitation initf reduc idition reduc idition reduc idition reduc idition reduc	ssion od = -417.99245 Coef. .1993764 -8080753 .0937764 .093776 .0937764 .0937764 .0937852 .0459933 Ss2 3952 .04720617 .04270617 .04270617 .04270617 .0437062 .0457062 .0457062 .0457062 .0457062 .0457062 .0457062 .0457062 .0	5 Std. Err. 4032128 4074551 105219 4074551 105219 105451 105555 1055555 1055555 1055555 1055555 1055555 1055555 1055555 1055555 1055555 1055555 1055555 10555555 10555555 10555555 10555555 105555555 10555555555 10555555555 105555555555	E P Z P Z P Z P Z P Z P Z P Z P Z P Z P	iunber of A R. 612(4) R. 612(4) >iseudo R2 >i[]   i.600 .600 .600 .600 .600 .600 .600 .60	cbs         =           =         =           =         =           (95% Cenf.         =	753 159.76 9.3882 2015 - 2015 9.3882 2015 - 2015 2015	Autobac     Vandaba     Vandabaa     Vandabaa     Vandabaa     Vandabaa     Vandabaa     Vandabaaa     Vandabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	her Lebel a f if i bib too. 19 a f if i bib too. 19 a f id i bib too. 19 a f id i bib too. 19 a f ido f - 10 kooran socked, 197 a f ido f - 10 kooran socked, 197 horbandri sogi kooran socked, 197 horbandri sogi horbandri sogi horbandri sogi horbandri sogi horbandri sogi horbandri socked, 197 horbandri sogi horbandri sogi horbandri sogi horbandri socked, 197 horbandri sogi horbandri sogi horbandr
2 - Servel 143 - Other 10 - Servel 143 - Other 10 - Servel 143 - Other 10	41 Saku(2) Wuoka(44802) Ins Stanta for Min <b>v v v</b> <b>v</b> <b>v</b> <b>v</b> <b>v</b> <b>v</b> <b>v</b> <b>v</b>	2000 Probli regree Log likeliho Log likel	stion d = -417.99285 Coef. .1093764 -8581735 .6599393 iets after prob (11f) (prediction .6299393 dy/dx 51 .427977 .927975 .8293752	5 Std. Err. .4331218 .4873451 .4873451 .4533481 .4533481 .46997 .46999 .4895959 .4895959 .4895959 .4895959 .4895959 .4895959 .48959559 .48959555555555	x         P           y         P	iumber of ( R chi2(4) seeds R2 b)(800	005 = 2 = 2 = (935 Corf 4.6667 65383 3 .6667 3 .6693 216667 3 .61465 216667	753 153.76 9.0880 9.1825 1.1546953 1.695269 2.695269 2.695269 2.28695 2.28695 2.28695 2.28695 2.28695 2.28695 2.28695 2.28716 5.22815 2.28716 5.22815 2.28716 2.29716	<ul> <li>✓ Valides</li> <li>✓ Enroration</li> <li>Value</li> <li>Value<!--</td--><td>Inter (date) a 1 if in lab frace, 19 a 1 if in lab frace, 19 a 1 idia 4 idia 4 idia 4 idia 4 idia 4 idia 4 idia 1 idia 4 idia 4 idia 4 idia 1 idia 4 idia 4 idia 4 idia 1 idia 4 idia 4 idia 4 idia 4 idia 1 idia 4 idi</td></li></ul>	Inter (date) a 1 if in lab frace, 19 a 1 if in lab frace, 19 a 1 idia 4 idia 4 idia 4 idia 4 idia 4 idia 4 idia 1 idia 4 idia 4 idia 4 idia 1 idia 4 idia 4 idia 4 idia 1 idia 4 idia 4 idia 4 idia 4 idia 1 idia 4 idi
2 - Secold 163 - 1040     the factor of	4f. Sheldy J Monkey (MBC)           sis Stands         Der W           • • • • • • • • • • • • • • • • • • •	2000 Proble regree Log likelikoo Log likelikoo Lidilis Hardisl eff edic Lidilis Hardisl eff edic Lidilis Hardisl eff edic Lidilis hold edic Lidilis hold edic Lidilis hold edic Lidilis hold edic Lidilis hold edic Lidilis hold edic Lidilis hold edic Lidilis hold edic Lidilis hold edic	ssien od = -417.99285 Coef. .899776 -8198753 .0659815 .0659815 .0659815 .0659815 .0659815 .0659815 .0659815 .0659815 .0659815 .0659815 .0427017 .9427017 .9427017 .9427017 .047707 .042707 .047707 .04	5 516.Err. .0232218 .026599 .0	E P Z P Z P Z P Z P Z P Z P Z P Z P Z P	iunber of R ch2(4) *seudo R2 *>[2]   .600 . .600 .	005 = 2 = 1935 Corf. 	753 153.76 9.0600 9.1882 155953 .693562 .095529 .095529 1.553485 X 12.2869 .33776 6.5.1288 39.6.388	Vankbe     Vankb	here Luded Luded T fin biols for 19 Luded L
Source 163 - 1504     Source 164 - 1504     To Source 164	A Subj Ubube (s4800) as Standa for the P D D D D D D D D D D D D D D D D D D D	2000 may any any any any any any any any any a	ssien od = -417.99285 Coef. .1993784 8189753 .4699393 ects after prob r(inlf) (predic .5823552 dy/dx 51 343742 319975 .0397783	5 5 5td. Err. 115219 80974501 .0069998 .4533481 .66999 .4533481 .0069992 .4533481 .0069992 .4533481 .006992 .0072	N         L           p         p           7         7.11           9         9.13           9         9.14           9         9.54           9.55         0.000           1.45         0.000	iumber of 4 R chi2(4) seudo R2 seudo R2	005 = 2 = [935 Cerf .1.657 - .95300 - .95300 - .95300 - .215667 - .21567	753 153.76 9.0800 9.1882 155953 .4695582 .975299 1.558485 1.558485 1.558485 1.558485 1.558485 1.558485 1.558485 1.558485 1.558485	<ul> <li>✓ Valides</li> <li>✓ Terrenden</li> <li>Name</li> <li>Name</li></ul>	Inter Inter I Inter Inter Inter I Inter Inter
Source 16-16-16-     Source 16-16-16-     Source 16-16-16-16-     Source 16-16-16-16-16-     Source 16-16-16-16-16-16-16-     Source 16-16-16-16-16-16-16-16-16-16-16-16-16-1	4f. Shidy (J) Worksty (MICO)           as Stands for W           1         1           1         1           1         1           1         1           1         1           1         1           1         1           1         1           1         1           1         1           1         1           1         1           2         1           1         1           2         1           1         1           2         1           2         1           2         1           2         1           2         1           2         1           2         1           3         1           3         1           3         1           3         1           3         1           3         1           3         1           3         1           3         1           3         1           3         1	200 top 100 to	ssien ad = -417.99285 Coef. .809734 .809736 .0093764 .809736 .009378 .009378 .009378 .019776 .019775 .019776 .0197776 .019776 .0197776 .019776 .0197776 .0197776 .01	5 Std. Err. .0231218 .135219 .0069998 .4533481 .006999 .4533481 .006999 .4533481 .006999 .015 .0	н р р р р р р р р р р 1.48 0 0.000 1.48 0 0.000 0 0.000 0 0.000 0 0.000 0 0.000 0 0.000 0 0.000 0 0 0 0 0 0 0 0 0 0 0 0	Lumber of R ch32(4) seudo R2 > 2   000 .0000 .000 .000 .000 .0	005 = [ 2 = [ [93X Corf. ] (0640665 -1.0667 -05380 -05380 -05380 -054800 -054800 -05480 -05480 -0548000 -0548000 -0548000 -0548000 -0548000 -054800000000000000000000000000000000000	753 135.76 9.0609 9.382 Interval] 1.354653 	<ul> <li>Valuations</li> </ul>	here Label  T if it lab from 19 hours worked. 197 High C + 10 hours worked. 197 ho

So, that means, instead of Logit you can put Probit also and you will get this. Again mfx command if you put, you will get the estimated value. So, now, if you compare the Logit and Probit model, look at the coefficient here, education coefficient is 0.10 in the context of Probit. But that was 0.18 in the context of Logit. So, that means, the coefficients of Probit model and the coefficients of Logit model they are not directly comparable. But the question is how to choose between Logit and Probit?

(Refer Slide Time: 22:47)

I Notel - Windows Journal File Edit View Inset Actions Tools Help D D D D D D D D D D D D D D D D D D D	- 0 x
NPTEL	$\frac{\mathcal{F}(\sigma, \tau)}{\left[ \begin{array}{c} \log it \ & \operatorname{Pot} \int it \\ \operatorname{Log} it \ & \operatorname{Log} it \ & \operatorname{Pot} \int it \\ \operatorname{Log} it \ & \operatorname{Log} it \ & \operatorname{Pot} \int it \\ \operatorname{Log} it \ & \operatorname{Log} it \$
	b ar an

How to select between Logit and Probit? So, in Logit, once again, if you look at what we assume, Pi equals to 1 by 1 plus e to the power minus alpha plus beta xi. And in the context of Probit, Probit Pi equals to f(alpha plus beta xi). And this f(alpha plus beta xi) assumes the cumulative distribution function of the logistic distribution, that means fi equals to integration minus infinity to alpha plus beta xi and then, alpha plus beta xi fz.dz and z equals to 1 by root 2 pi sigma square into e to the power minus zi square by 2 and this zi equals to zi minus mu by sigma whole square, the standard normal variable.

Now both normal and the logistic distribution have zero mean because here we are assuming standard normal variable. So, mean is 0 but their variance and standard deviation are different. So zi follows a normal distribution with 0 mean and variance 1. And here, the zi follows in the context of logistic distribution, the mean is again 0 the variance actually pi square by 3, it is not 1.

Because of the difference in their variance, here it is 1, here it is pi square by 3, where pi equals 22 by 7, so that means here it is sigma equals to pi by root 3 equals to 1.81 or something. So, if you multiply the Probit equation by 1.81, that means beta hat Probit multiplied by 1.81 equals to beta hat Logit. If you look at your model what is the Probit estimate? Probit estimate is for a particular variable 0.10 and here, it is 0.10 here it is 0.18.

So, Probit estimate is point 0.10 multiplied that by 1.81 equals to 0.18. So, Probit estimate is 0.10 and if you multiply that by 1.81 then you will get the Logit estimate and that is the idea. Now the question is how do you select between Probit and Logit model?

(Refer Slide Time: 27:59)



Compared to Probit, Logit has a flatter tail. What does it mean? That means if there are too many observations at the upper or lower tail then from the diagram itself it is very clear that Logit is a better representation of that sigmoid ace or non-linear relationship than the Probit because Logit it is a little sour. And in the middle range Logit equals to Probit, then Logit is preferred than Probit.

Many times, we use Logit and Probit exchangeabley. And we think that we may use either Logit or Probit it does not matter. But the question is what is your data set? If in your sample you see for example, in the car owner or house owner relationship, say that there are 90 percent individuals who are actually having car that means you are at the upper end or there are only 10 percent people who are having car are at the lower end, that situation is better captured by Logit model not the Probit one.

But if your data says around 50 percent, 50-50, 40-60 or 45-55 something like that, that means actually you are in the middle range, you have equal number of zeros and ones. And then, we can say that Logit is equivalent to Probit. Most of the cases we are actually in the middle range. That is why we use Logit or Probit alternatively. But actually, that is not the case.

If your sample says that you have too many observations in the tail either upper or lower, then you must use Logit model and not the Probit model. That is why we need to know this type of diagrammatic representation to know which one to select between Logit and Probit. Logit is preferred than Probit and in the middle they are equivalent.

So, with this, we are basically closing our discussion on Logit and Probit. In our next class for you we will discuss another qualitative response model. Thank you.