Logistic regression

Assignment #1

Introduction

Socio-economic variables are very often categorical, rather than interval scale

logistic regression may be thought of as an approach that is similar to that of multiple linear regression, but takes into account the fact that the dependent variable is categorical.

Definition of Logistic regression

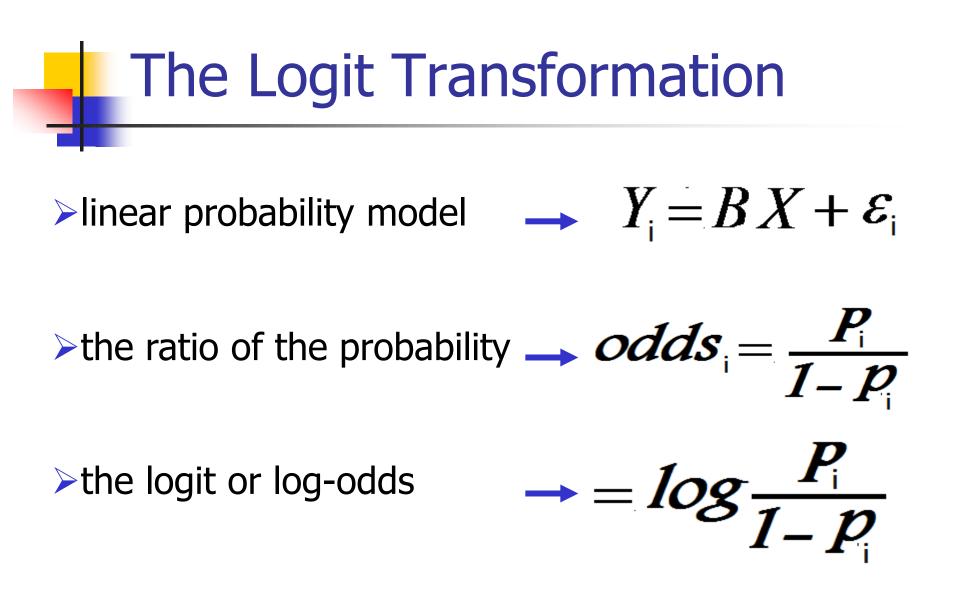
Determines the impact of multiple independent variables presented together to predict membership of one or other of the two dependent variable categories

Assumptions of logistic regression

Logistic regression does not assume a linear relationship

The dependent variable must be a dichotomy (2 categories)

➤The independent variables need not be interval, nor normally distributed, nor linearly related, nor of equal variance within each group



The logistic regression equation

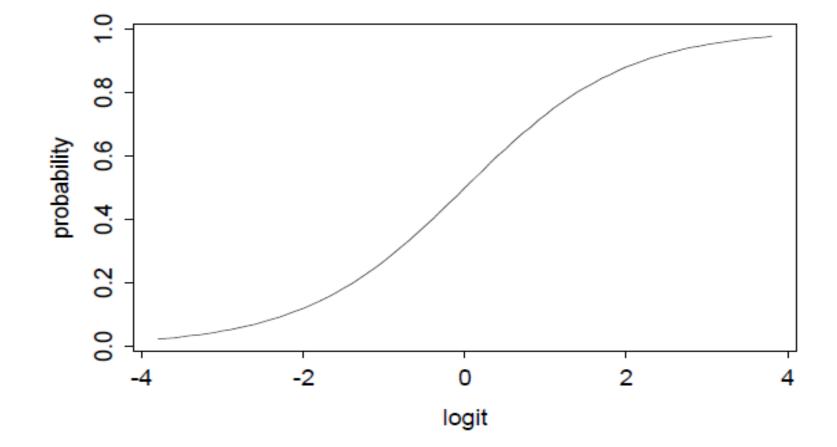
$$p = \frac{exp^{(a+b_1x_1+b_2x_2+b_3x_3...)}}{1+exp^{(a+b_1x_1+b_2x_2+b_3x_3...)}}$$

Where:

p = the probability that a case is in a particular category

- exp = the base of natural logarithms (approx 2.72),
- a = the constant of the equation and,
- b = the coefficient of the predictor variables





Computer Application

Untitled1 [DataSet0] - Sl	PSS Data Edi	itor			
<u>F</u> ile <u>E</u> dit <u>V</u>	∕iew <u>D</u> ata	<u>T</u> ransform	<u>A</u> nalyze	<u>G</u> raphs	Utilities	s Add- <u>o</u> ns <u>Wi</u> ndow <u>H</u> elp
🗁 🗏 🔔	📴 🔶 🔿	· 🔚 🖬 🛛	Repor	ts		
1:			D <u>e</u> scr	iptive Statisti	cs	•
	var	var	Ta <u>b</u> les	3		Var var var var
1			Compa	are Means		•
2			<u>G</u> ener	al Linear Mo	del	•
З	1		Gener	ali <u>z</u> ed Linear	Models	ils 🕨
4	1		Mi <u>x</u> ed	Models		•
5			<u>C</u> orrel	ate		•
6			<u>R</u> egre	ssion		▶ R Linear
7			Loglin			د <u>Curve Estimation</u>
8			Neura	l Net <u>w</u> orks		Partial Least Squares
9			Classi			Binary Logistic
10			_	Reduction		• R. Multinomial Logistic
11			Sc <u>a</u> le			B Or <u>d</u> inal
12				rametric Tes	ts	Ros Probit
13			Time S			R Nonlinear
14			<u>S</u> urviv ⊠® ⊷			B Mericht Extinction
15				g Value Ana	l¥sis	R 2 Stone Least Sauares
16	<u> </u>			e Response		
17				ex Samples		Optimal Scaling
18				/ Control		
19	<u> </u>		ROC C	Cur <u>v</u> e		
20						

Computer Application

C	Logistic Regression		×
	household income in ,0 for years old [age] for monthly mortgage paym for number of persons in h	Dependent: take solar panel offer [Block 1 of 1 Previous Covariates: Mortgage Famsize Method: Enter Selection Variable: Prule	OK <u>P</u> aste <u>R</u> eset Cancel Help
		Categorical <u>Save</u> <u>Options</u>	

Computer Application

Logistic Regression: Options Statistics and Plots ✓ Classification plots ✓ Hosmer-Lemeshow goodness-of-fit ✓ Casewise listing of residuals ✓ Outliers outside 2 std. dev. ✓ All cases	 Correlations of estimates Iteration history CI for exp(B): 95 % 	Continue Cancel Help
 Display At each step 	C At Jast step	
Probability for Stepwise E <u>n</u> try: .05 Remo <u>v</u> al: .10 ✓ Include con <u>s</u> tant in model	Classification c <u>u</u> toff: .5 <u>M</u> aximum Iterations: 20	

Table (1.1) Classification table

			Predicted				
			Take solar p				
	Observed		Decline offer	Take offer	Percentage correct		
Step 0	take solar panel	decline offer	0	14	.0		
	offer	take offer	0	16	100.0		
	Overall Percentage				53.3		

^b The cut value is .500.

Table 1.3 Variables not in the equation table

Variables	not in	the	Equation
-----------	--------	-----	----------

			Score	df	Sig.
Step 0	Variables	Famsize	14.632	1	.000
		Mortgage	6.520	1	.011
	Overall Statistics		15.085	2	.001

Table 1.5 Model Summary

Model Summary					
Step	–2 Log likelihood	Cox & Snell R square	Nagelkerke R square		
1	17.359°	.552	.737		

^a Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

Table 1.8 Classification table

Classification Table^a

			Predicted				
			Take solar panel offer				
	Observed		Decline offer	Take offer	Percentage correct		
Step 1	take solar panel	decline offer	13	1	92.9		
	offer	take offer	2	14	87.5		
	Overall Percentage				90.0		

^a The cut value is .500.

Table 1.9 Variables in the equation

Variables in the Equation								
		В	S.E.	Wald	df	Sig.	Exp(B)	
Step 1ª	Famsize	2.399	.962	6.215	1	.013	11.007	
	Mortgage	.005	.003	3.176	1	.075	1.005	
	Constant	-18.627	8.654	4.633	1	.031	.000	
Probał	oility of a ca	se =		ly size) + (.(nily size) +		~ ~ /		