

Discovering Statistics Using SPSS 10th Edition

Logistic Regression

Dr. Andy Field

Slide 1 ANDY FIELD

Discovering Statistics Using SPSS 10th Edition

Aims and Objectives

- When and Why do we Use Logistic Regression?
- Theory Behind Logistic Regression
- Logistic Regression on SPSS
 - Running the analysis
 - Interpreting the Output
- Issues for Reliability of the model

Slide 2 ANDY FIELD

Discovering Statistics Using SPSS 10th Edition

When And Why

- To predict outcome variable that is a categorical dichotomy from one or more categorical or continuous predictor variables..
- Used because having a categorical dichotomy as an outcome variable violates the assumption of linearity in normal regression.

Slide 3 ANDY FIELD

Discovering Statistics Using SPSS 10th Edition

With One Predictor

$$P(Y) = \frac{1}{1 + e^{-(b_0 + b_1 X_1 + \epsilon_i)}}$$

- Outcome
 - We predict the *probability* of the outcome occurring
- b_0 and b_1
 - Can be thought of in much the same way as multiple regression
 - Note the normal regression equation forms part of the logistic regression equation

Slide 4 ANDY FIELD

Discovering Statistics Using SPSS 10th Edition

With Several Predictor

$$P(Y) = \frac{1}{1 + e^{-(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n + \epsilon_j)}}$$

- Outcome
 - We still predict the *probability* of the outcome occurring
- Differences
 - Note the multiple regression equation forms part of the logistic regression equation
 - This part of the equation expands to accommodate additional predictors

Slide 5 ANDY FIELD

Discovering Statistics Using SPSS 10th Edition

The Wald Statistic

$$Wald = \frac{b}{SE_b}$$

- Similar to *t*-statistic in Regression.
- Tests the null hypothesis that $b = 0$.
- Is biased when b is large.
- Better to look at Likelihood-ratio statistics.

Slide 6 ANDY FIELD

Exp(b)

$$\text{Exp}(b) = \frac{\text{Odds after a unit change in the predictor}}{\text{Odds before a unit change in the predictor}}$$

- Indicates the change in odds resulting from a unit change in the predictor.
 - Exp(b) > 1: Predictor ↑, Probability of outcome occurring ↑.
 - Exp(b) < 1: Predictor ↑, Probability of outcome occurring ↓.

Slide 7 ANDY FIELD

Methods of Regression

- Forced Entry:** All variables entered simultaneously.
- Hierarchical:** Variables entered in blocks.
 - Blocks should be based on past research, or theory being tested. Good Method.
- Stepwise:** Variables entered on the basis of statistical criteria (i.e. relative contribution to predicting outcome).
 - Should be used only for exploratory analysis.

Slide 8 ANDY FIELD

An Example

- Predictors of scoring a penalty in football.
- Outcome:**
 - Penalty scored (1) or missed (0)
- Predictors:**
 - Is the player a worrier? (PSWQ)
 - Ability (Previous success)
 - Anxiety (STAI)
- PSWQ and Previous are known predictors, anxiety is a new measure.

Slide 9 ANDY FIELD

Logistic Regression on SPSS

Slide 10 ANDY FIELD

Methods

Slide 11 ANDY FIELD

Residuals

Slide 12 ANDY FIELD

Options

Slide 13

ANDY FIELD

SPSS Output: Block 1

Classification Table^a

		Predicted		Percentage Correct
		Result of Penalty Kick		
Observed	Result of Penalty Kick	Missed Penalty Kick	Scored Penalty Kick	
Step 0	Result of Penalty Kick	0	35	0
	Overall Percentage	0	40	100.0
				53.3

a. Constant is included in the model.
b. The cut value is .500

Variables in the Equation

Step	Constant	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0		.134	.231	.333	1	.564	1.143

Variables not in the Equation

Step	Variables	PREVIOUS	Score	df	Sig.
0			34.109	1	.000
1		PSWQ	34.193	1	.000
	Overall Statistics		41.558	2	.000

Slide 14

ANDY FIELD

R^2

$$R_L^2 = \frac{\text{Model Chi-Square}}{-2\text{Log-Likelihood}}$$

$$= \frac{54.977}{103.6385}$$

$$= 0.53$$

Slide 15

ANDY FIELD

SPSS Output: Block 1

Omnibus Tests of Model Coefficients

Step	Step	Chi-square	df	Sig.
Step 1	Block	54.977	2	.000
	Model	54.977	2	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	48.662	.520	.694

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7.931	7	.339

Slide 16

ANDY FIELD

SPSS Output: Block 1

Contingency Table for Hosmer and Lemeshow Test

Step	1	Result of Penalty Kick = Missed Penalty		Result of Penalty Kick = Scored Penalty		Total
		Observed	Expected	Observed	Expected	
	1	8	7.904	0	.096	8
	2	8	7.779	0	.221	8
	3	8	6.705	0	1.295	8
	4	4	5.438	4	2.562	8
	5	2	3.945	6	4.055	8
	6	2	1.820	6	6.180	8
	7	2	1.004	6	6.996	8
	8	1	.268	7	7.732	8
	9	0	.108	11	10.892	11

Classification Table

		Predicted		Percentage Correct
		Result of Penalty Kick		
Observed	Result of Penalty Kick	Missed Penalty Kick	Scored Penalty Kick	
Step 1	Result of Penalty Kick	50	5	85.7
	Overall Percentage	7	33	84.0

a. The cut value is .500

Slide 17

ANDY FIELD

SPSS Output: The Model

Variables in the Equation

Step	PREVIOUS	B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
1	PSWQ	-.230	.080	8.309	1	.004	.784	.678	.929
	Constant	1.280	1.670	.588	1	.443	3.568		

a. Variable(s) entered on step 1: PREVIOUS, PSWQ.

Slide 18

ANDY FIELD

SPSS Output: Block 2

Omnibus Tests of Model Coefficients

Step	Step	Chi-square	df	Sig.
1	Step	1.246	1	.264
	Block	1.246	1	.264
	Model	56.223	3	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	47.416	.527	.704

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9.937	7	.192

Slide 19 ANDY FIELD

SPSS Output: Block 2

Contingency Table for Hosmer and Lemeshow Test

Step	Observed	Result of Penalty Kick = Missed Penalty		Result of Penalty Kick = Scored Penalty		Total
		Observed	Expected	Observed	Expected	
1	1	8	7.628	0	.074	8
	2	8	7.769	0	.231	8
	3	9	7.649	0	1.351	9
	4	4	5.425	4	2.575	8
	5	1	3.210	7	4.790	8
	6	4	1.684	4	6.316	8
	7	1	1.049	7	6.951	8
	8	0	.222	8	7.778	8
	9	0	.067	10	9.933	10

Classification Table^a

Observed	Result of Penalty Kick	Predicted		Percentage Correct
		Missed Penalty	Scored Penalty	
Step 1	Result of Penalty Kick	30	5	85.7
	Missed Penalty	6	34	85.0
	Scored Penalty			85.3
	Overall Percentage			

a. The cut value is .500

Slide 20 ANDY FIELD

SPSS Output: The Model

Variables in the Equation

Step	PREVIOUS	B	S.E.	Wald	df	Sig.	Exp(B)		95.0% C.I. for EXP(B)	
							Lower	Upper	Lower	Upper
1	PREVIOUS	-.203	.129	2.454	1	.117	1.225	.593	1.578	
	PSWO	-.251	.084	8.954	1	.003	.778	.660	.917	
	ANXIOUS	.276	.253	1.193	1	.275	1.318	.803	2.162	
	Constant	-11.493	11.802	.948	1	.330	.000			

a. Variable(s) entered on step 1: ANXIOUS.

Slide 21 ANDY FIELD