Block 16: Logit analysis

(Activity solutions can be found at the end of the document.)

Often we have a **binary response variable** which cannot be modelled using a typical linear regression model due to the categorical nature of the response. The **binary logit model** is used instead to model the probability of the outcome.

Learning Objective

- Describe the binary logit model and its advantages over discriminant and regression analysis.

Reading List

Malhotra, N.K., D. Nunan and D.F. Birks. Marketing Research: An Applied Approach. (Pearson, 2017) 5th edition [ISBN 9781292103129] Chapter 23 (from page 696).

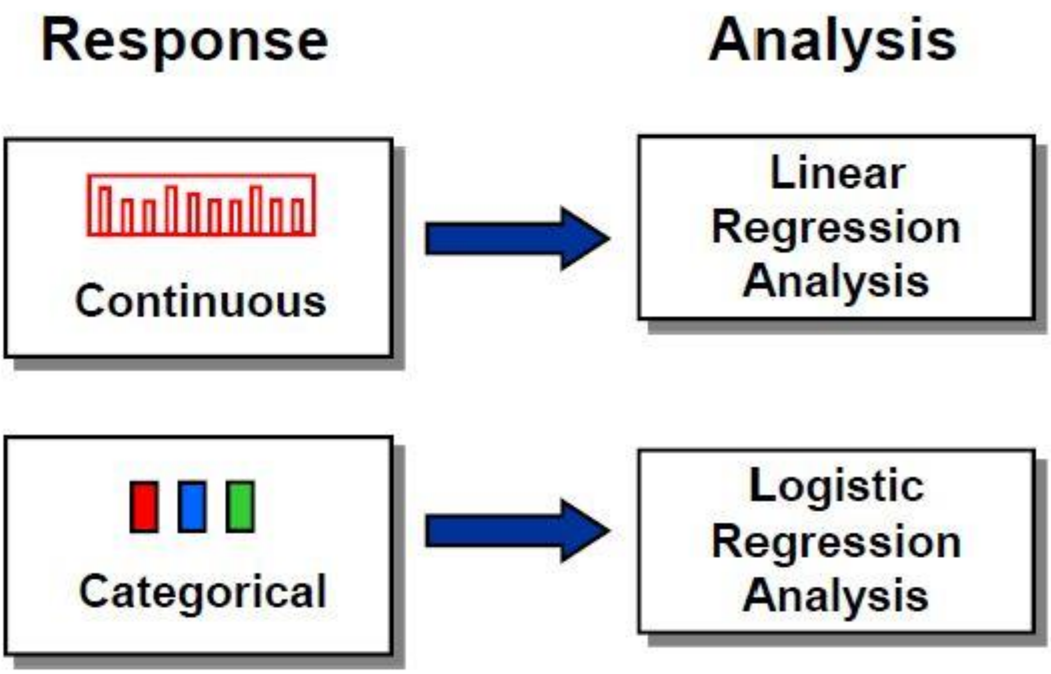
16.1 Logit analysis

For each section of *Logit analysis*, use the LSE ELearning resources to test your knowledge with the Key terms and concepts flip cards.

Overview

Regression analysis enables you to characterise the relationship between a response variable and one or more predictor variables.

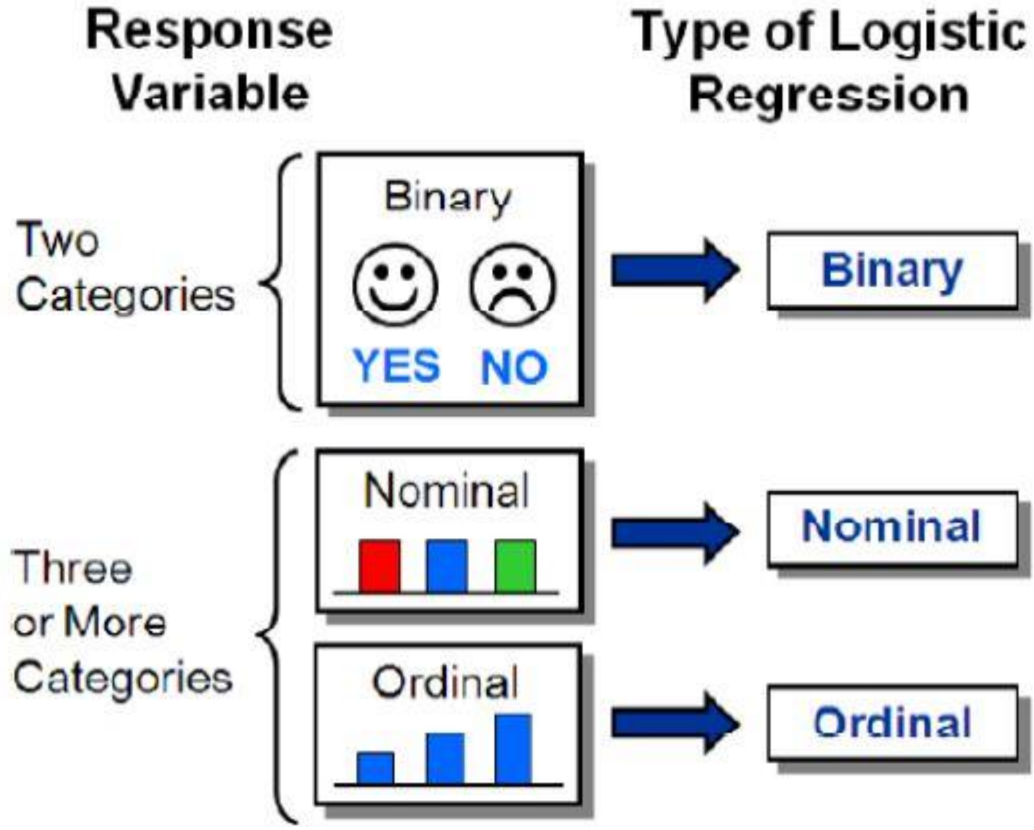
- In linear regression, the response variable is continuous.
- In logistic regression, the response variable is categorical.



If the response variable is *dichotomous* (i.e. binary, that is with two categories), the appropriate logistic regression model is the **binary logit model**, also known as binary logistic regression. If you have more than two categories (levels) within the response variable, then there are two possible logistic regression models as follows.

- If the response variable is nominal, you fit a *nominal logistic regression model*.

- If the response variable is ordinal, you fit an *ordinal logistic regression model*.



Key definition:

The binary logit model commonly deals with the issue of how likely an observation is to belong to each group. It estimates the probability of an observation belonging to a particular group.

Problems with linear regression models

Why not conduct a standard regression analysis?

One might be tempted to analyse a regression model with a binary response variable using ordinary least squares regression. However, there are problems with that! Besides the arbitrary nature of the coding, there is the problem that the predicted values will take on values which have no intrinsic meaning with regards to your response variable. There is also the mathematical inconvenience of not being able to assume normality and a constant variance for the error term when the response variable has only two values.

Recall the simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

If the response variable is categorical, then how do you code the response numerically? If the response is coded as 1 = ‘yes’ and 0 = ‘no’ and your regression equation predicts 0.5 or 1.1 or −0.4, what does that mean practically? If there are only two (or a few) possible response levels, is it reasonable to assume normality and a constant variance for the error term?

What about a linear probability model? Consider a linear probability model:

$$\pi_i = \beta_0 + \beta_1 X_i$$

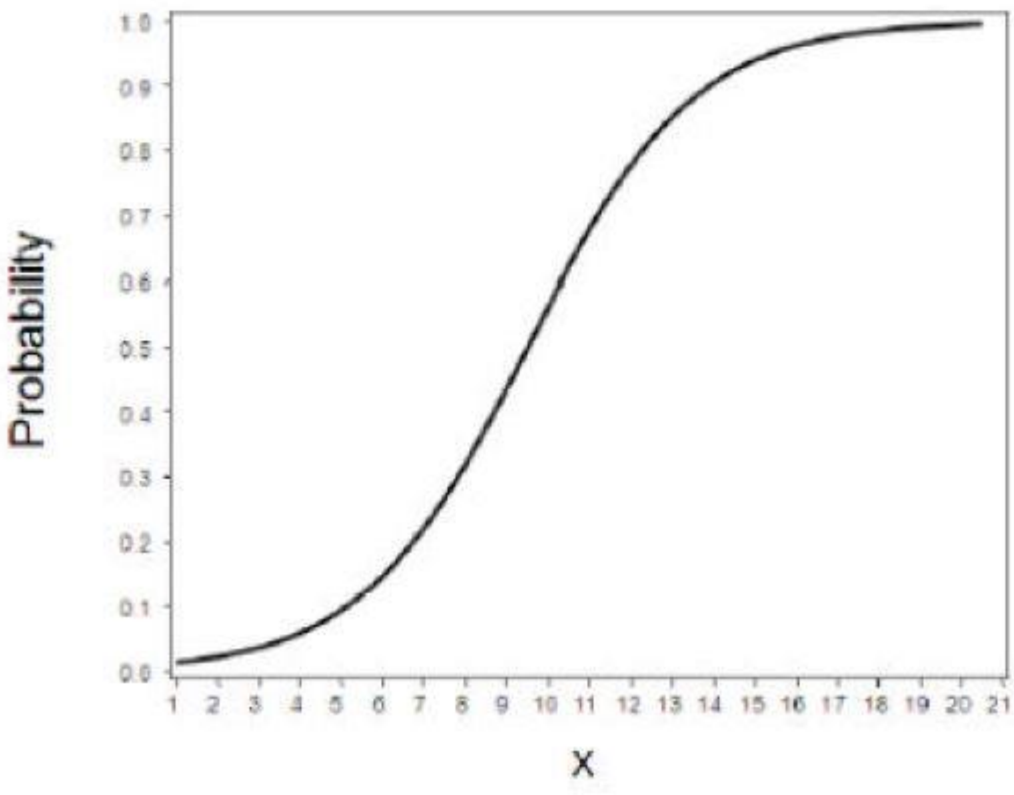
Probabilities are bounded, but linear functions can take on any value. (Once again, how do you interpret a predicted value of −0.4 or 1.1?) Given the bounded nature of probabilities, can you assume a *linear* relationship between X and π throughout the possible range of X? Can you assume a random error with normality and a constant variance? What is the observed probability for an observation?

Instead of modelling the 0s and 1s directly, another way of thinking about modelling a binary variable is to *model the probability* of either 0 or 1. If you can model the probability of the 1 (call that $\pi\pi$), then you have also modelled the probability of the 0, which would be 1−π. Probabilities are truly continuous and so this line of thinking might sound compelling at first. One problem is that the predicted values from a linear model can assume, theoretically, *any* value. However, probabilities are by definition bounded between 0 and 1.

Another problem is that the relationship between the probability of the outcome and a predictor variable is usually *non-linear* rather than *linear*. In fact, the relationship often resembles an S-shaped curve (a 'sigmoidal' relationship). Probabilities do not have a random normal error associated with them, but rather a *binomial error* of π(1−π). Such an error is greatest at probabilities close to 0.5 and lowest near 0 and 1. They do not have a constant error associated with them.

Logistic regression curve and logit transformation

The figure below shows a model of the relationship between a continuous predictor and the probability of an event or outcome.



A linear model clearly will not fit if this is the true relationship between X and the probability. In order to model this relationship directly, you must use a *non-linear*function. One such function, the *logistic regression curve*, is displayed.

The parameter estimate of this curve determines the rate of increase or decrease of the estimated curve. When the parameter estimate is *greater* than 0, the probability of the outcome *increases* as the predictor variable values increase. When the parameter estimate is *less* than 0, the probability *decreases* as the predictor variable values increase. As the absolute value of the parameter estimate increases, the curve has a steeper rate of change. When the parameter estimate is equal to 0, the curve can be represented by a straight, horizontal line which shows an equal probability of the event for everyone.

The βiβi values cannot be computed using standard linear regression. This is not a linear model!

A logistic regression model applies a **logit transformation** to the probabilities. Two of the problems we saw with modelling the probability directly were that probabilities were bounded between 0 and 1, and that there was not likely a linear relationship between predictors and probabilities. First, deal with the problem of the restricted range of the probability. What about the range of a logit?

Logistic regression models transform probabilities called *logits*, defined as:

$$logit(\pi_i) = \ln \frac{\pi_i}{1 - \pi_i}$$

where:

- i is the index of a case (observation), for i = 1 ,…,n
- π_i is the probability the event (a sale, for example) occurs in the *i*th case
- ln is the natural log (to the base e).

As ππ approaches its maximum value of 1, the value π/(1−π) approaches infinity. As π approaches its minimum value of 0, π/(1−π) approaches 0. The natural log of something approaching infinity is something approaching infinity, and the natural log of something approaching 0 is something approaching negative infinity. So, the logit has no upper or lower bounds. If you can model the logit, then simple algebra will allow you to model the *odds* of the event, or even the *probability* of the event.

The logit transformation ensures that the model generates estimated probabilities which lie between 0 and 1. The logit is the natural log of the odds.

Conducting binary logit analysis

The process to conduct binary logit analysis is as follows:

Formulate the binary logit problem

⇓⇓

Estimate the binary logit model

⇓⇓

Determine the model fit

⇓⇓

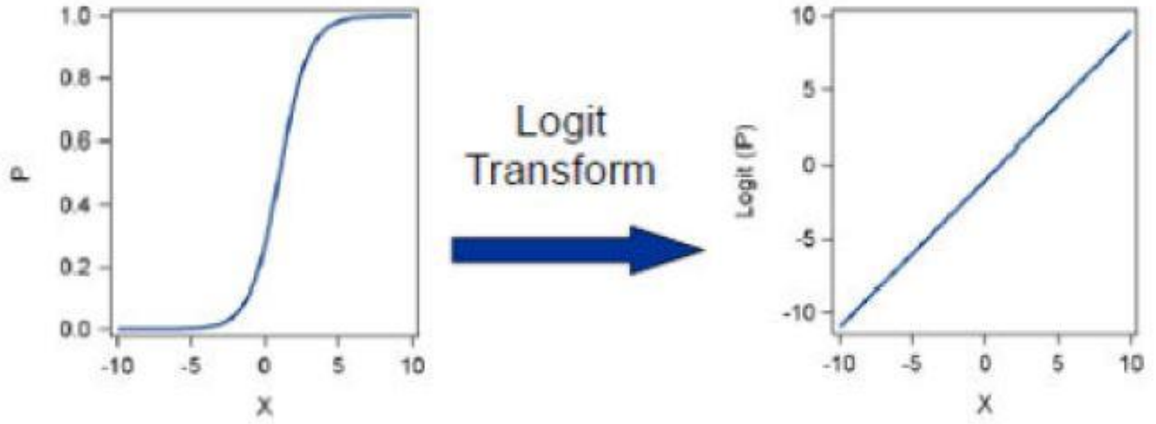
Test the significance of individual parameters

⇓⇓

Interpret the coefficients and validate

Conducting binary logit analysis

The assumption in logistic regression is that the logit transformation of the probabilities results in a *linear* relationship with the predictor variables. If the thoughts about the nature of the direct relationship between X and π are correct, then the logit will have a linear relationship with X. In other words, a linear function of X can be used to model the logit. In that way, you can indirectly model the probability. To verify this assumption, it would be useful to plot the logits by the predictor variable. An example of the logit transformation is given below.



Binary logit model

For a *binary* outcome variable, the logistic regression model with more than one predictor variable has the form:

$$logit(\pi_i) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k$$

where:

logit (πi) = the logit of the probability of the event

β0 = the intercept of the regression equation

βk = the parameter estimate of the *k*th predictor variable, X_k.

Unlike linear regression, the logit is *not* normally distributed and the variance is *not* constant. Also, logistic regression usually requires a more complex estimation method called *maximum likelihood* to estimate the parameters than linear regression (which uses the least squares procedure). Maximum likelihood finds the parameter estimates which are most likely to occur, given the data. This is accomplished by maximising the ‘likelihood function’ which expresses the probability of the observed data as a function of the unknown parameters.

An illustrative application of logistic regression using data in [Table 23.6 of the textbook] includes assessing the model fit and significance testing.

Probability and odds

An **odds ratio** indicates how much more likely, with respect to odds, a particular event occurs in one group relative to its occurrence in another group. We define odds as:

$$Odds = \frac{\pi event}{1 - \pi event}$$

The odds ratio can be used to measure the *strength of association* for 2×2 tables. Do not mistake odds for probability; odds are calculated from probabilities as shown next.

Probability versus odds

Consider the table below.

	Outcome		
	No	Yes	Total
Group A	20	60	80
Group B	10	90	100
Total	30	150	180

Probability of a **Yes outcome** in Group B = 90/100 **(0.90)**

Probability of a **No outcome** in Group B = 10/100 **(0.10)**

There is a 90% probability of having the outcome in Group B, and a 10% probability of not having the outcome. What is the probability of having the outcome in Group A? This is 75% (with a 25% probability of not having the outcome).

Odds

Consider the situation below.

Odds of Outcome in Group B

Probability of a Yes outcome in Group B ÷ Probability of a No outcome in Group B

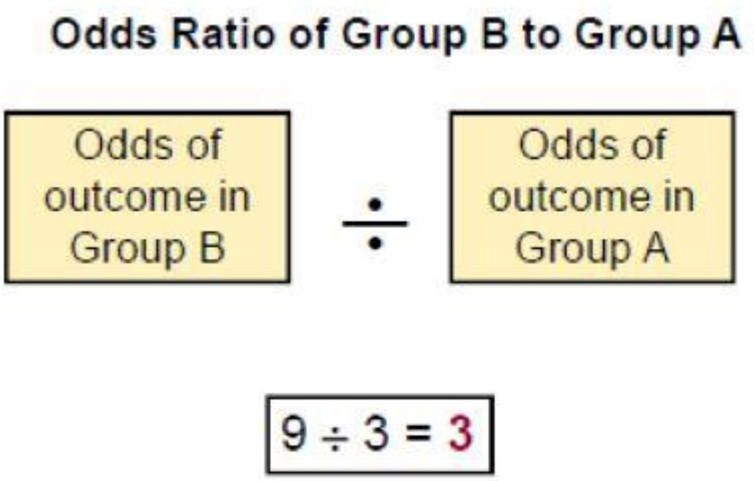
$\boxed{0.90 \div 0.10 = 9}$

The *odds* of an outcome is the ratio of the expected probability that the outcome will occur to the expected probability that the outcome will not occur.

The odds for Group B are 9 indicating that you expect 9 times as many occurrences as non-occurrences in Group B. What are the odds of having the outcome in Group A? These are 0.75/0.25 = 3.

Odds ratio

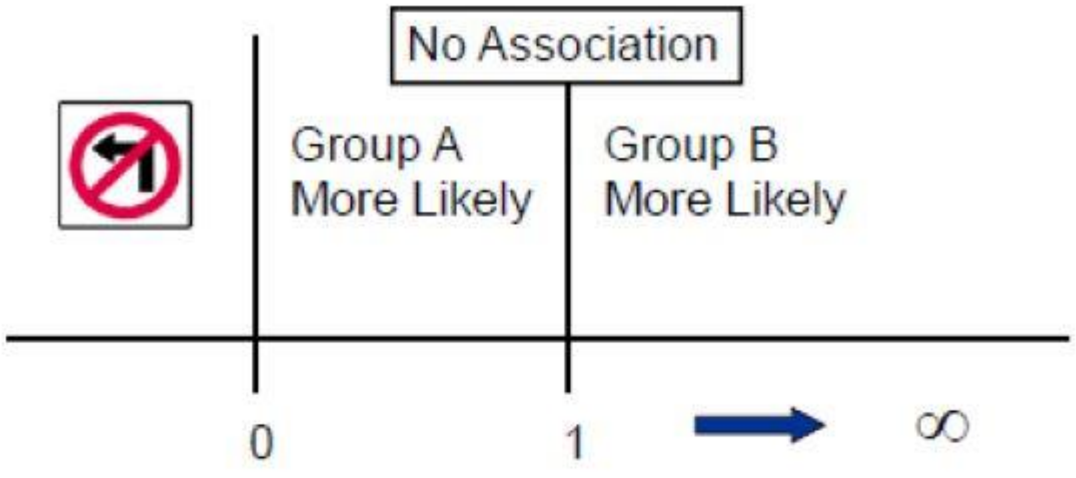
Consider the situation below.



The *odds ratio* of Group B to Group A equals 3, indicating that the odds of getting the outcome in Group B are 3 times those in Group A.

Properties of the odds ratio, B to A

The odds ratio shows the *strength of the association* between the predictor variable and the outcome variable. If the odds ratio is 1, then there is no association between the predictor variable and the outcome. If the odds ratio is greater than 1, then Group B is more likely to have the outcome. If the odds ratio is less than 1, then Group A is more likely to have the outcome. This is shown below.



Odds ratio calculation

Suppose we had the following binary logit model, where gender is the predictor variable for some binary outcome:

$$logit\ (\hat{\pi}) = \log(odds) = \beta_0 + \beta_1 \times Gender$$

Remember that in logistic regression you model the natural log of the odds and not the odds or probability directly. For interpretation, often the parameter estimates are converted into something more interpretable - an *odds ratio*.

In order to understand this, write out the linear model predicting the natural log of the odds. In order to see that in terms of odds, the natural log is 'undone' by *exponentiation*.

Exponentiation of the right-hand side of the equation must also be done to maintain equality. Therefore, you can look at the model in terms of odds and can estimate odds for females or males. The odds ratio is then the ratio of the odds of one group to the odds of another group.

The odds ratio of females to males would be:

$$odds_{females} = e^{\beta_0+\beta_1} \quad and \quad odds_{males} = e^{\beta_0}$$

hence:

$$odds\ ratio = \frac{e^{\beta_0+\beta_1}}{e^{\beta_0}} = \frac{e^{\beta_0+\beta_1}}{e^{\beta_0}} = e^{\beta_1}$$

Activity

1. Activity on the block's topics

1. The data file *Honours.sav* contains data on the following variables for a sample of 200 students. (An Excel version of the dataset is *Honours.xlsx*.) Use the LSE ELearning resources to practice the exercise or watch a video walkthrough.
    - Gender - '0 = male' and '1 = female'.
    - Reading - metric of a student's reading ability.
    - Science - metric of a student's scientific ability.
    - Honours - '0 = no honours' and '1 = honours'.

A researcher wants to use the binary logit model to build a model to help investigate how the odds that a student achieves honours is affected by certain predictors. Perform a binary logistic regression in SPSS to assist the researcher.

Use **Analyze** >> **Regression** >> **Binary Logistic**……. Make 'Honours' the (binary) dependent variable and the others as 'Covariates'. Click 'Options……' and select 'CI for exp(B)'.

2. Solution to activity on the block's topics

1. SPSS initially describes a 'null model', which is a model with no predictors and just the intercept. This output is:

Block 0: Beginning Block

Classification Table[a, b]

Observed	Predicted

			Honours		Percentage Correct
			No honours	**Honours**	
Step 0	**Honours**	**No honours**	147	0	100.0
		Honours	53	0	.0
	Overall Percentage				73.5

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S. E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1.020	.160	40.540	1	.000	.361

Variables not in the Equation

			Score	df	Sig.
Step 0	**Variables**	**Gender**	3.871	1	.049
		Reading	47.906	1	.000
		Science	34.862	1	.000
	Overall statistics		58.609	3	.000

In the first box, we see that SPSS has predicted that all cases are 0 (no honours), since the majority of students in the sample were non-honours students. Remember, this model has no predictors so, unsurprisingly, if you had to predict the type of student you would look for the larger proportion group (here non-honours students) and predict this type. The overall percentage gives the percentage of cases for which the dependent variable was correctly predicted given the model. Here it is 73.5% as 147/200 = 0.735.

The 'B' is the coefficient for the constant (i.e. the intercept) in the null model, which here is −1.020−1.020. To test the null hypothesis that the actual constant is 0, we use the Wald test statistic. The test statistic value is 40.540 and the pp-value is 0.000, so the result is highly significant.

‘Exp(B)’ is the exponentiation of the B coefficient, which is an odds ratio. This value is given by default because odds ratios can be easier to interpret than the coefficient, which is in log-odds units. This is the odds: 53/147 = 0.361.

‘Score’ is a score test that is used to predict whether or not an independent variable would be significant in the model. Looking at the pp-values (located in the column labelled ‘Sig.’), we can see that each of the predictors would be statistically significant, although gender is very borderline (the pp-value is 0.049).

We now turn to the most interesting part of the output: the overall test of the model (in the ‘Omnibus Tests of Model Coefficients’ table) and the coefficients and odds ratios (in the ‘Variables in the Equation’ table). This output is:

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	**Step**	71.052	3	.000
	Block	71.052	3	.000
	Model	71.052	3	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	160.236[a]	.299	.436

c. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Classification Table[a]

Observed			Predicted		
			Honours		Percentage Correct
			No honours	Honours	
Step 1	**Honours**	**No honours**	135	12	91.8
		Honours	26	27	50.9

	Overall Percentage			81.0

d. The cut value is .500

Variables in the Equation

		B	**S. E.**	**Wald**	**df**	**Sig.**	**Exp(B)**
Step 1[a]	Gender	1.482	.447	10.980	1	.001	4.404
	Reading	.104	.026	16.147	1	.000	1.109
	Science	.095	.030	9.688	1	.002	1.099
	Constant	-12.777	1.976	41.818	1	.000	0.000

e. Variable(s) entered on step 1: Gender, Reading, Science.

Variables in the Equation

		95% C.I.for EXP(B)	
		Lower	**Upper**
Step 1[a]	**Gender**	1.832	10.584
	Reader	1.054	1.167
	Science	1.036	1.167
	Reader		

f. Variable(s) entered on step 1: Gender, Reading,Science.

‘Chi-square’ is a test statistic value and ‘Sig.’ is its pp-value. In this example, the statistics for the Step, Block and Model are the same because we have not used stepwise logistic regression or blocking. The value given in the Sig. column is the probability of obtaining the chi-square statistic or a more extreme value given that the null hypothesis is true. In other words, this is the probability of obtaining the test statistic value of 71.052 or a more extreme (higher) value if there is in fact no effect of the independent variables, taken together, on the dependent variable. In this case, the model is statistically significant because the pp-value is 0.000.

‘−2−2 Log likelihood’ is not very informative here. However, it can be used to compare nested (reduced) models. Larger (absolute) values suggest superior models.

‘Cox & Snell R Square’ and ‘Nagelkerke R Square’ are pseudo R-squares. Logistic regression does not have an exact equivalent to R2R2 found in linear regression. There are many pseudo R-square statistics (these are only two of them).

‘Observed’ indicates the numbers of 0s and 1s that are observed in the dependent variable. ‘Predicted’ gives values of the dependent variable based on the full logistic regression model. This table shows how many cases are correctly predicted (135 cases are observed to be 0 (no honours) and are correctly predicted to be 0; 27 cases are observed to be 1 (honours) and are correctly predicted to be 1), and how many cases are not correctly predicted (12 cases are observed to be 0 but are predicted to be 1; 26 cases are observed to be 1 but are predicted to be 0).

‘Overall Percentage’ is the hit ratio giving the overall percentage of cases that are correctly predicted by the model (in this case, the full model that we specified). As you can see, this percentage has increased from 73.5% for the null model to 81.0% for the full model.

‘B’ are the values for the logistic regression coefficients for predicting the dependent variable from the independent variables. They are in log-odds units. Similar to linear regression, the prediction equation is:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

where p is the probability of being an honours student. Expressed in terms of the variables used in this example, the estimated logistic regression equation is:

$$\ln\left(\frac{p}{1-p}\right) = 12.777 + 1.482 Gender + 0.104 Reading + 0.095 Science$$

These estimates tell you about the relationship between the independent variables and the dependent variable, where the dependent variable is on the logit scale. These estimates tell you the amount of increase in the predicted log odds of being an honours student which would occur for a one-unit increase in a specific predictor, holding all other predictors constant. The columns labelled ‘Wald’ and ‘Sig.’ give the significance tests of each independent variable (here they are all highly significant). Because these coefficients are in log-odds units, they are often difficult to interpret, so they are often converted into odds ratios. You can do this by hand by exponentiating the coefficient, or by looking at the right-most column in the ‘Variables in the Equation’ table labelled ‘Exp(B)’.

This fitted model says that, holding reading and science metrics at a fixed value, the odds of getting honours for females (female = 1) over the odds of getting honours for males is exp(1.482) = 4.404. In terms of percentage change, we can say that the odds for females are 340.4% higher than the odds for males. The coefficient for reading says that, fixing gender and holding science at a fixed value, we will see a 10.9% increase in the odds of getting honours for a one-unit increase in reading score since exp(0.104) = 1.109. The coefficient for science says that, fixing gender and holding reading at a fixed value, we will see a 9.9% increase in the odds of getting honours for a one-unit increase in science score since exp(0.095) = 1.099. Note that 95% confidence intervals for the odds ratios are also reported. All exclude 1 (the lower endpoints are all greater than 1), hence there is evidence of an association.

Discussion forum and activity

To access the solutions to these questions and case study, click here to access the printable Word document or click here to go to LSE's Elearning resources.

Activity on the block's topics

1. The data file *Honours.sav* contains data on the following variables for a sample of 200 students. (An Excel version of the dataset is *Honours.xlsx*.)

    - Gender - '0 = male' and '1 = female'.
    - Reading - metric of a student's reading ability.
    - Science - metric of a student's scientific ability.
    - Honours - '0 = no honours' and '1 = honours'.

A researcher wants to use the binary logit model to build a model to help investigate how the odds that a student achieves honours is affected by certain predictors. Perform a binary logistic regression in SPSS to assist the researcher.

Use **Analyze** >> **Regression** >> **Binary Logistic**……. Make 'Honours' the (binary) dependent variable and the others as 'Covariates'. Click 'Options……' and select 'CI for exp(B)'.

Video of walkthrough of activity 1.

Learning outcomes checklist

Use this to assess your own understanding of the chapter. You can always go back and amend the checklist when it comes to revision!

- Describe the binary logit model and its advantages over discriminant and regression analysis.

Block 16: Logit analysis

Solution to exercise on the block's topics

1. SPSS initially describes a 'null model', which is a model with no predictors and just the intercept. This output is:

Block 0: Beginning Block

Classification Table[a, b]

Observed			Predicted		
			Honours		Percentage Correct
			No honours	Honours	
Step 0	Honours	No honours	147	0	100.0
		Honours	53	0	.0
	Overall Percentage				73.5

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S. E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1.020	.160	40.540	1	.000	.361

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	Gender	3.871	1	.049
		Reading	47.906	1	.000
		Science	34.862	1	.000
	Overall statistics		58.609	3	.000

In the first box, we see that SPSS has predicted that all cases are 0 (no honours), since the majority of students in the sample were non-honours students. Remember, this model has no predictors so, unsurprisingly, if you had to predict the type of student you would look for the larger proportion group (here non-honours students) and predict this type. The overall percentage gives the percentage of cases for which the dependent variable was correctly predicted given the model. Here it is 73.5% as 147/200 = 0.735.

The ‘B’ is the coefficient for the constant (i.e. the intercept) in the null model, which here is −1.020−1.020. To test the null hypothesis that the actual constant is 0, we use the Wald test statistic. The test statistic value is 40.540 and the pp-value is 0.000, so the result is highly significant.

‘Exp(B)’ is the exponentiation of the B coefficient, which is an odds ratio. This value is given by default because odds ratios can be easier to interpret than the coefficient, which is in log-odds units. This is the odds: 53/147 = 0.361.

‘Score’ is a score test that is used to predict whether or not an independent variable would be significant in the model. Looking at the pp-values (located in the column labelled ‘Sig.’), we can see that each of the predictors would be statistically significant, although gender is very borderline (the pp-value is 0.049).

We now turn to the most interesting part of the output: the overall test of the model (in the ‘Omnibus Tests of Model Coefficients’ table) and the coefficients and odds ratios (in the ‘Variables in the Equation’ table). This output is:

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	**df**	**Sig.**
Step 1	**Step**	71.052	3	.000
	Block	71.052	3	.000
	Model	71.052	3	.000

Model Summary

Step	**-2 Log likelihood**	**Cox & Snell R Square**	**Nagelkerke R Square**
1	160.236[a]	.299	.436

c. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Classification Table[a]

Observed			Predicted		
			Honours		Percentage Correct
			No honours	Honours	
Step 1	Honours	No honours	135	12	91.8
		Honours	26	27	50.9
	Overall Percentage				81.0

d. The cut value is .500

Variables in the Equation

		B	S. E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Gender	1.482	.447	10.980	1	.001	4.404
	Reading	.104	.026	16.147	1	.000	1.109
	Science	.095	.030	9.688	1	.002	1.099
	Constant	-12.777	1.976	41.818	1	.000	0.000

e. Variable(s) entered on step 1: Gender, Reading, Science.

Variables in the Equation

		95% C.I.for EXP(B)	
		Lower	Upper
Step 1[a]	**Gender**	1.832	10.584
	Reader	1.054	1.167
	Science	1.036	1.167
	Reader		

f. Variable(s) entered on step 1: Gender, Reading,Science.

‘Chi-square’ is a test statistic value and ‘Sig.’ is its pp-value. In this example, the statistics for the Step, Block and Model are the same because we have not used stepwise logistic regression or blocking. The value given in the Sig. column is the probability of obtaining the chi-square statistic or a more extreme value given that the null hypothesis is true. In other words, this is the probability of obtaining the test statistic value of 71.052 or a more extreme (higher) value if there is in fact no effect of the independent variables, taken together, on the dependent variable. In this case, the model is statistically significant because the pp-value is 0.000.

‘−2−2 Log likelihood’ is not very informative here. However, it can be used to compare nested (reduced) models. Larger (absolute) values suggest superior models.

‘Cox & Snell R Square’ and ‘Nagelkerke R Square’ are pseudo R-squares. Logistic regression does not have an exact equivalent to R2R2found in linear regression. There are many pseudo R-square statistics (these are only two of them).

‘Observed’ indicates the numbers of 0s and 1s that are observed in the dependent variable. ‘Predicted’ gives values of the dependent variable based on the full logistic regression model. This table shows how many cases are correctly predicted (135 cases are observed to be 0 (no honours) and are correctly predicted to be 0; 27 cases are observed to be 1 (honours) and are correctly predicted to be 1), and how many cases are not correctly predicted (12 cases are observed to be 0 but are predicted to be 1; 26 cases are observed to be 1 but are predicted to be 0).

‘Overall Percentage’ is the hit ratio giving the overall percentage of cases that are correctly predicted by the model (in this case, the full model that we specified). As you can see, this percentage has increased from 73.5% for the null model to 81.0% for the full model.

‘B’ are the values for the logistic regression coefficients for predicting the dependent variable from the independent variables. They are in log-odds units. Similar to linear regression, the prediction equation is:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

where *p* is the probability of being an honours student. Expressed in terms of the variables used in this example, the estimated logistic regression equation is:

$$\ln\left(\frac{p}{1-p}\right) = 12.777 + 1.482 Gender + 0.104 Reading + 0.095 Science.$$

These estimates tell you about the relationship between the independent variables and the dependent variable, where the dependent variable is on the logit scale. These estimates tell you the amount of increase in the predicted log odds of being an honours student which would occur for a one-unit increase in a specific predictor, holding all other predictors constant. The columns labelled ‘Wald’ and ‘Sig.’ give the significance tests of each independent variable (here they are all highly significant). Because these coefficients are in log-odds units, they are often difficult to interpret, so they are often converted into odds ratios. You can do this by hand by exponentiating the coefficient, or by looking at the right-most column in the ‘Variables in the Equation’ table labelled ‘Exp(B)’.

This fitted model says that, holding reading and science metrics at a fixed value, the odds of getting honours for females (female = 1) over the odds of getting honours for males is exp(1.482) = 4.404. In terms of percentage change, we can say that the odds for females are 340.4% higher than the odds for males. The coefficient for reading says that, fixing gender and holding science at a fixed value, we will see a 10.9% increase in the odds of getting honours for a one-unit increase in reading score since exp(0.104) = 1.109. The coefficient for science says that, fixing gender and holding reading at a fixed value, we will see a 9.9% increase in the odds of getting honours for a one-unit increase in science score since exp(0.095) = 1.099. Note that 95% confidence intervals for the odds ratios are also reported. All exclude 1 (the lower endpoints are all greater than 1), hence there is evidence of an association.