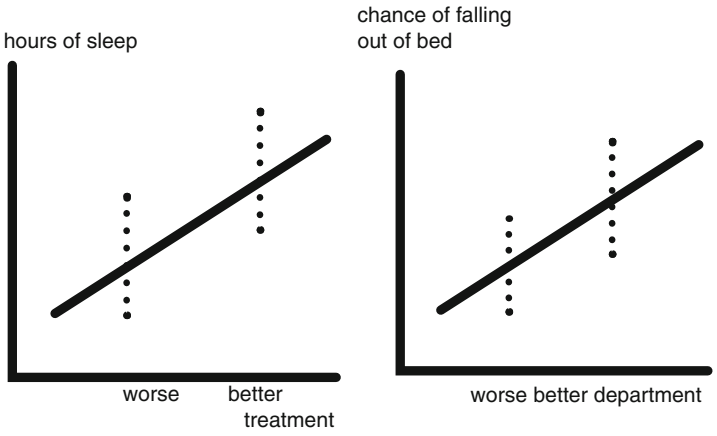
Chapter 36
Logistic Regression with a Binary Predictor (55 Patients)

1 General Purpose

Similarly to chi-square tests, logistic regression can be used to test whether there is a significant difference between two treatment modalities. To see how it works review the linear regression example from Chap. 5. The linear regression model with treatment modality as independent variable (x-variable), and hours of sleep as dependent variable (y-variable = outcome variable) showed that the treatment modality was a significant predictor of the hours of sleep, and, thus, that there was a significant difference between the two treatments. If your treatment is not a medicine but rather a type of hospital department, and your outcome is not hours of sleep, but, rather, the chance of falling out of bed, then we will have a largely similar situation.



T.J. Cleophas, A.H. Zwinderman, *SPSS for Starters and 2nd Levelers*,
DOI 10.1007/978-3-319-20600-4_36

The type of department is assumed to predict the risk of falling out of bed, and is defined as a binary x-variable. The risk of falling out of bed is the y-variable, but, unlike hours of sleep like in Chap. 6, falling out of bed is not a continuous variable, but rather a binary variable: you either fall or you don't. With binary y-variables linear regression is impossible, and logistic regression is required. Otherwise, the analysis and interpretation is pretty much similar to that of the linear regression.

2 Schematic Overview of Type of Data File

Outcome binary	predictor binary
•	•
•	•
•	•
•	•
•	•
•	•
•	•

3 Primary Scientific Question

In clinical research the predictor is often a treatment modality or the presence of a patient characteristic, the outcome is often responding yes or no. If your chance of responding is large, then your treatment is excellent. With logistic regression the chance of responding is calculated as the odds of responding (= ratio of number of responders / number of nonresponders) or rather the log odds (logarithmically transformed odds). The larger the logodds of responding, the better the treatment.

4 Data Example

The example of Chap. 35 is used once more. In 55 hospitalized patients the risk of falling out of bed was assessed. The question to be answered was: is there a significant difference between the risk of falling out of bed at the departments of surgery and internal medicine. The first 10 patients of the 55 patient file is underneath.

Fall out of bed	Department
1,00	,00
1,00	,00
1,00	,00
1,00	,00
1,00	,00
1,00	,00
1,00	,00
1,00	,00
1,00	,00
1,00	,00

fall out bed 0 = no, 1 = yes
department 0 = surgery, 1 = internal medicine

5 Crosstabs

The data file is in extras.springer.com, and is entitled "chapter35unpairedbinary". We will start by opening the data in SPSS.

Command:
Analyze....Descriptive Statistics....Crosstabs....Row(s): enter department.... Column(s): enter falloutofbed....click OK.

Department * falloutofbed crosstabulation
Count

		Falloutofbed		
		,00	1,00	Total
Department	,00	20	15	35
	1,00	5	15	20
Total		25	30	55

The output sheet shows a 2×2 contingency table. It shows that at both departments the same numbers of patients fall out of bed. However, at the department of surgery many more patients do *not* fall out of bed than at the internal department.

6 Logistic Regression

For analysis the statistical model Binary Logistic Regression in the module Regression is required.

Command:
Analyze....Regression....Binary Logistic Regression....Dependent: enter falloutofbed.... Covariates: enter department....click OK.

Variables in the equation

		B	S.E.	Wald	Df	Sig.	Exp(B)
Step 1[a]	Department	1,386	,619	5,013	1	,025	4,000
	Constant	−,288	,342	,709	1	,400	,750

[a]Variable(s) entered on step 1: department

The above results table of the logistic regression shows that B (the regression coefficient) for the variable department (which is the hospital department) is a significant predictor of the chance of falling out of bed with a p-value of 0,025. This is a p-value largely similar to that of the chi-square test from Chap. 35. The meaning of this logistic regression is also largely the same as that of the chi-square test.

A nice thing about logistic regression is that, unlike with chi-square tests, an odds ratio is given. The odds ratio, equals 4,000, which can interpreted as follows. The chance of falling out of bed is four times larger at the department of surgery than it is at the department of internal medicine. The odds ratio equals e^B, with e = Euler's number = mathematical constant = 2,783 and B = regression coefficient), and is written in the table as "exp (B)".

The significant correlation between the type of department and the risk of falling out of bed can be interpreted as a significant difference in safety at the two departments.

7 Conclusion

Similarly to chi-square tests, logistic regression can be used to test whether there is a significant difference between two treatment modalities. E.g, a better and worse treatment on a better and worse outcome can be tested. Or the effect of a better or worse hospital department on a better or worse chance of falling out of bed. A nice thing about logistic regression is that it does not only provide p-values but also, unlike chi-square testing, odds ratios, which can be interpreted as the ratio of success in the better as compared to that of the worse response group.

8 Note

More background, theoretical, and mathematical information about logistic regression is given in Statistics applied to clinical studies 5th edition, Chaps. 17 and 65, Springer Heidelberg Germany, 2012, from the same authors.