Ordinal logistic regression

Contents

1. Introduction

In a previous example we analyzed data with a nominal outcome variable. In some instances, the categories of a dependent variable may be ordered in a natural way. In such a case, one should try to fit a model that takes this characteristic into account un the analysis.

The file **Wellness.lsf** contains data on the mental health impairment of 40 patients given by Agresti (1990, Table 9.8). Data and syntax files can be found in the **MVABOOK\Chapter3** folder.

wellness.LSF

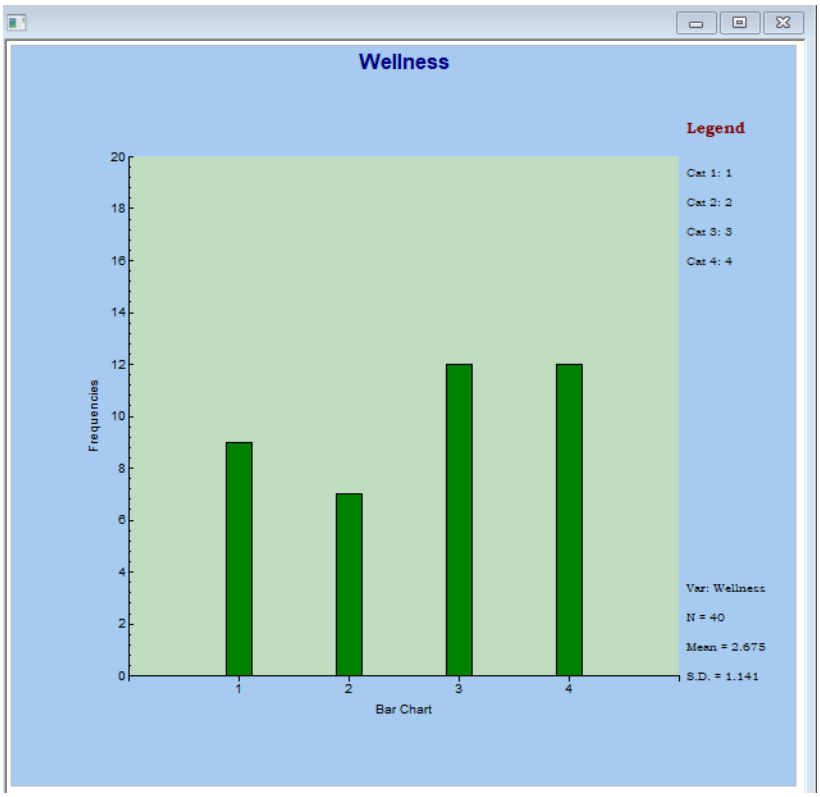
	Wellness	Ses	Events	Interact
1	4.00	1.00	1.00	1.00
2	4.00	1.00	9.00	9.00
3	4.00	1.00	4.00	4.00
4	4.00	1.00	3.00	3.00
5	4.00	0.00	2.00	0.00
6	4.00	1.00	0.00	0.00
7	4.00	0.00	1.00	0.00
8	4.00	1.00	3.00	3.00
9	4.00	1.00	3.00	3.00
10	4.00	1.00	7.00	7.00
11	4.00	0.00	1.00	0.00
12	4.00	0.00	2.00	0.00
13	3.00	1.00	5.00	5.00
14	3.00	0.00	6.00	0.00
15	3.00	1.00	3.00	3.00
16	3.00	0.00	1.00	0.00
17	3.00	1.00	8.00	8.00
18	3.00	1.00	2.00	2.00
19	3.00	0.00	5.00	0.00
20	3.00	1.00	5.00	5.00

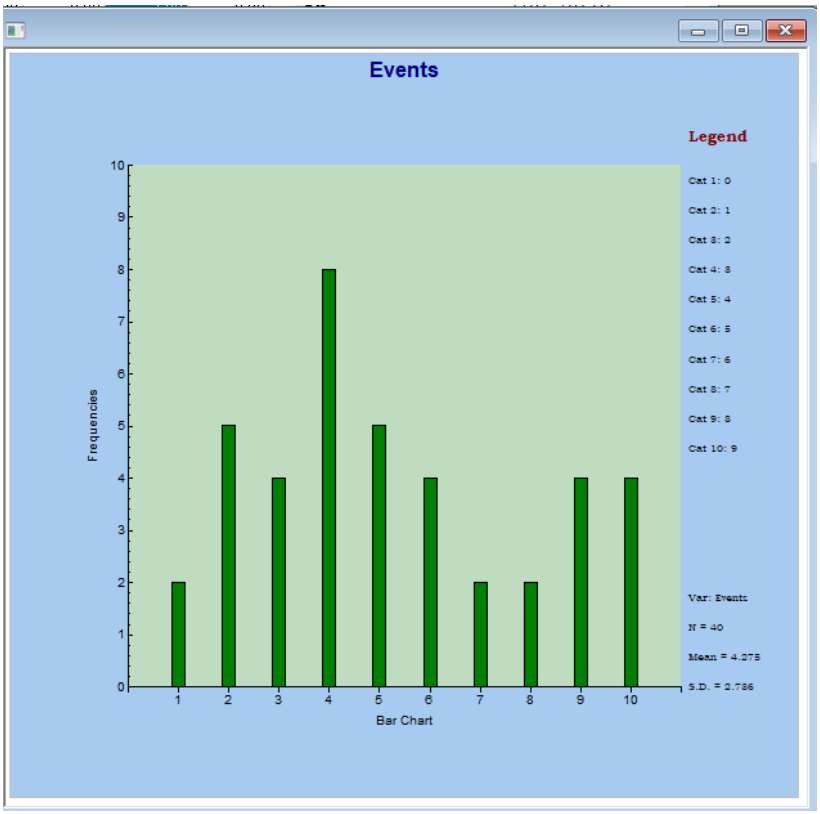
The first variable in the data set is Wellness, an ordinal variable coded as follows:

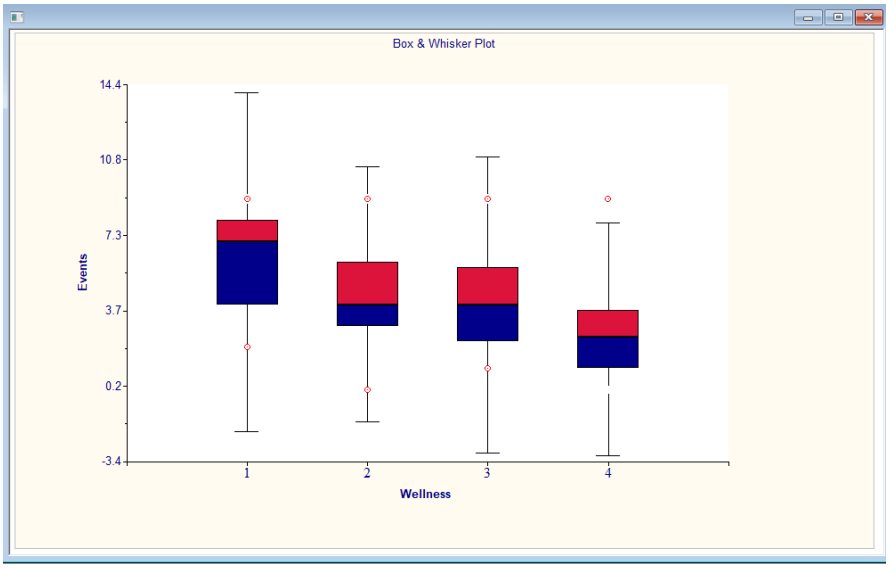
- 1: Impaired
- 2: Moderate
- 3: Mild
- 4: Well

A higher value for this variable thus indicates a higher level of mental health. Ses indicates the patient's socio-economic status and is binary in nature, assuming a value of 1 for high and 0 for low. Events is an index ranging between 1 and 9 used as measure of both number of severity of like events such as child birth, divorce, change in employment and the like. We will treat both SES and Events are continuous variables in this analysis.

Bar charts of Wellness and Events are given below. We see that most of the responses on the dependent variable were either in the Mild or Well categories.







The box-and-whisker plot indicates that the distribution of Events is negatively skewed for impaired patients, and positively skewed for patients with moderate impairment (Wellness = 1, 2).

2. Ordinal logistic model

Wellness is the dependent variable of interest. Using the Impaired category as the reference category, we now use ordinal logistic regression to estimate the following:

$$\frac{\pi_1}{\pi_2+\pi_3+\pi_4} = \alpha_1 + \gamma_1 x_1 + \gamma_2 x_2$$

$$\frac{\pi_1+\pi_2}{\pi_3+\pi_4} = \alpha_2 + \gamma_1 x_1 + \gamma_2 x_2$$

$$\frac{\pi_1+\pi_2+\pi_3}{\pi_4} = \alpha_3 + \gamma_1 x_1 + \gamma_2 x_2$$

To invoke ordinal logistic regression, the link function OLOGIT should be used, as shown in the syntax file **wellness1.prl** below:

wellness1.prl

```
GlimOptions Response=Descending RefCatCode=0;
Title=Mental Impairment;
SY=wellness.LSF;
Distribution=MUL;
Link=OLOGIT;
DepVar=Wellness;
CoVars=Events Ses;
```

Results are given below.

Estimated Regression Weights

Parameter	Estimate	Standard Error	z Value	P Value
Alpha1	-0.2819	0.6231	-0.4525	0.6509
Alpha2	1.2128	0.6511	1.8626	0.0625
Alpha3	2.2094	0.7171	3.0810	0.0021
Events	-0.3189	0.1194	-2.6699	0.0076
Ses	1.1112	0.6143	1.8090	0.0704

The estimate of 1.1112 indicates that the chance of a higher level of mental wellness is associated with higher SES. The opposite trend is observed for Events: the chance of higher mental wellness increases with a higher score on the life events index.

To calculate the probability of being in the second, third or last category of wellness (remember that the first category serves as reference category here) we define

$$\lambda_j = \exp(\alpha_j + \gamma_1 x_1 + \gamma_2 x_2)$$

The probabilities for categories 2, 3 and 4 can then be expressed as

$$\pi_1 = \frac{\lambda_1}{1+\lambda_1}$$

$$\pi_2 = \frac{\lambda_2}{1+\lambda_2} - \pi_1$$

$$\pi_3 = \frac{\lambda_3}{1+\lambda_3} - \pi_2 - \pi_1$$

Using the results from the output and looking specifically at patients with an observed value of 4 on Events, we calculate the λ s and then the π s for patients with low and high SES respectively.

Event = 4, Ses = 0:

$$\hat{\lambda}_1 = \exp(-0.2819 - 0.3189(4) + 1.1112(0)) = 0.2107$$

$$\hat{\lambda}_2 = \exp(1.2128 - 0.3189(4) + 1.1112(0)) = 0.9391$$

$$\hat{\lambda}_3 = \exp(2.2094 - 0.3189(4) + 1.1112(0)) = 2.5442$$

Event = 4, Ses = 1:

$$\hat{\lambda}_1 = \exp(-0.2819 - 0.3189(4) + 1.1112(1)) = 0.6399$$

$$\hat{\lambda}_2 = \exp(1.2128 - 0.3189(4) + 1.1112(1)) = 2.8531$$

$$\hat{\lambda}_3 = \exp(2.2094 - 0.3189(4) + 1.1112(1)) = 7.7292$$

We now calculate the probabilities for each category of the outcome. Results are summarized in the table below.

SES	Impaired	Moderate	Mild	Well
Low	0.3703	0.1740	0.3103	0.2335
High	0.2595	0.3902	0.3502	0.1449

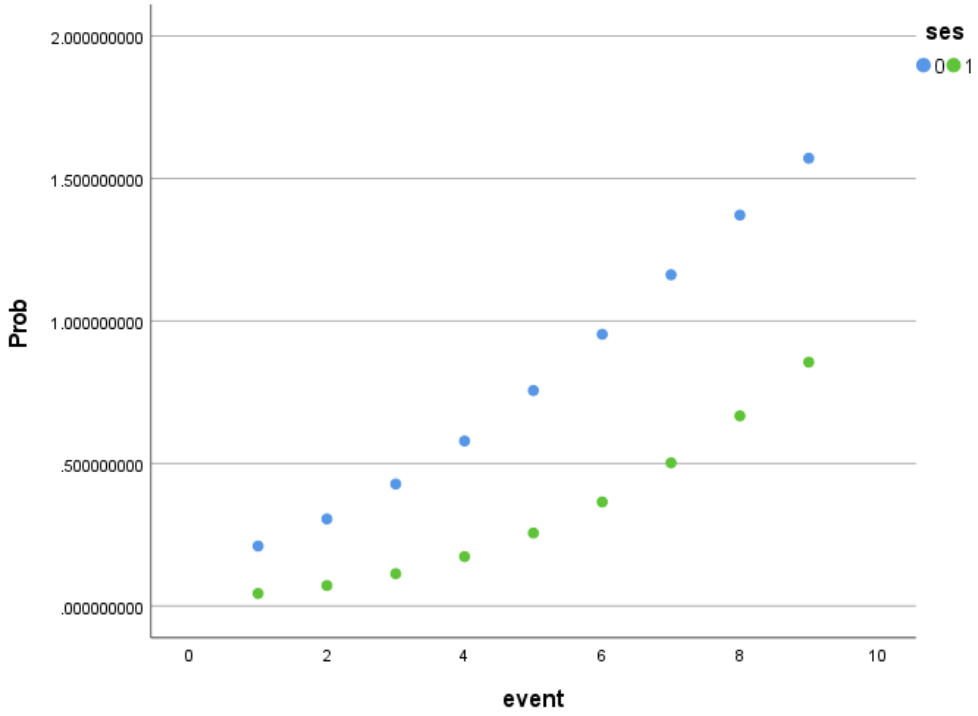
The probability of being mentally impaired (Wellness = 1) is noticeably higher for patients with low SES when compared to that of patients with high SES (0.3703 vs 0.2595). Patient with high SES (and remember, a score of 4 on the life event index) are most likely to be in the moderate or mild categories on the mental health spectrum.

The results above give us a snapshot of probabilities for the four categories of the outcome relative to patient's SES, but do not provide insight into the impact of change in the number of life events. To look at that, we need to repeat these calculations for various values of Events.

We may want to consider looking at the probability of having a mild impairment or "well" score on the outcome compared to being classified as impaired or moderate. To do so, we calculate

$$\Pr(y>2)=\exp(-\hat{\alpha}_2-\hat{\gamma}_1 x_1-\hat{\gamma}_2 x_2)\Pr(y\le 2)$$
$$=\frac{1}{\hat{\lambda}_2}(\hat{\pi}_1+\hat{\pi}_2)$$
$$=\frac{1}{\hat{\lambda}_2}\frac{\hat{\lambda}_2}{1+\hat{\lambda}_2}$$
$$=1+\exp[-(1.2128-0.31891x_1+1.1112x_2)]$$

The results of calculating this probability for the two levels of socio-economic status and various values of the life event index are given as a scatterplot below:



We note that higher levels on the event index are associated with a higher probability to be in either the “mild” or “well” category of the mental health category, regardless of the socio-economic status of the patient. However, it is also clear that patients with a high socio-economic status were consistently more likely to have a mild or well diagnosis than patients with a low socio-economic status, regardless of the score on the life event index.