Ordinal Regression

by James Lani

http://www.statisticssolutions.com/regression-analysis-ordinal-regression/

Click here for to get help with your Thesis or Dissertation.

Click here for FREE Thesis and Dissertation resources (templates, samples, calculators).

**Ordinal regression** is a statistical technique that is used to predict behavior of ordinal level dependent variables with a set of independent variables. The dependent variable is the order response category variable and the independent variable may be categorical or continuous. In SPSS, this test is available on the regression option analysis menu.

**Questions answered:**

Do gender and race influence happiness as categorized by the XYZ survey?

Does age relate to the level of shopping likelihood (not at all likely, somewhat likely, moderately likely, extremely likely)?

**Assumptions:**

1. One dependent variable, we cannot use multiple dependent variables.
2. Parallel lines assumption: There is one regression equation for each category except the last category. The last category probability can be predicted as 1-second last category probability.
3. Adequate cell count: As per the rule of thumb, 80% of cells must have more than 5 counts. No cell should have Zero count. The greater the cell with less count, the less reliable the chi-square test will be.

**Key terms and concepts:**

**Dependent variable:** The dependent variable is ordinal. The first category is usually considered as the lowest category and the last category is considered as the highest category; they are usually numerically coded from 0 on up). Usually in SPSS, logit function is used to predict the dependent variable category. Probit function is also used to predict the dependent variable category when the dependent variable has relatively equal categories. There is a K-1 predication where K is the number of a category in a dependent variable.

**Factor:** Factor is a categorically independent variable that must be coded as numeric in SPSS (e.g., gender coded as 0 = male and 1 = female).
**Covariate:** Covariates are continuous independent variables which are used to predict the dependent variable category (e.g., IQ score).

**Link function:** The link function is a transformation of the cumulative probabilities of the dependent ordered variable that allows for estimation of the model. However, in SPSS, five link functions are available, these link functions are as follows:

1. Logit function: Logit function is the default function in SPSS for ordinal regression. This function is usually used when the dependent ordinal variable has equal category. Mathematically, logit function equals to \( f(x) = \log(x / (1 – x)) \).
2. Probit model: This is the inverse standard normal cumulative distribution function. This function is more suitable when a dependent variable is normally distributed.
3. Negative log-log \( f(x) = -\log (-\log(x)) \): This link function is recommended when the probability of the lower category is high.
4. Complementary log-log \( f(x) = \log (-\log (1 – x)) \): This function is inverse of the negative log-log function, it is recommended when the probability of higher category is high.
5. Cauchit. \( f(x) = \tan (p(x – 0.5)) \): This link function is used when the extreme values are present in the data.

**Statistics and saved variables:** The output button in SPSS gives the flexibility to save the output. We can save predicted category, or predicted category probability by selecting this option from the output button.

**Parameter estimates, standard errors, significance levels, and confidence intervals:** In the output table of SPSS, a table called ‘parameter estimates’ appears. There is a variable named threshold, which is used for the Intercept term, and the location variable gives the coefficient for the independent variable for the specified link function. The first threshold will be used to predict the probability of the first order. Wald statistics is used to test the significance of the independent variable with degrees of freedom and standard error.

**Goodness of fit information:** Pearson chi-square test gives the information about how many predicted cell frequencies differ from observed frequencies.

**R-square estimate:** As in simple linear regression, we cannot use simple r-square in ordinal regression. R-square gives the information about how much variance is explained by the independent variable. However, variance is split into categories. Hence Cox and Snell’s, Nagelkerke’s, and McFadden’s pseudo-R2 statistics will be used in ordinal regression to estimate the variance explained by the independent variable.

*For assistance with conducting an ordinal regression or other quantitative analysis click here.*

**Resources**


**Related Pages:**

[Conduct and Interpret an Ordinal Regression](http://www.statisticssolutions.com/ordinal-regression)

[Regression](http://www.statisticssolutions.com/regression)