ORDINAL REGRESSION

Author: Basile Marquier

Date Amended: 15/04/2019

The following resources are associated:

SPSS dataset Graduate.sav'

Ordinal regression in SPSS

Dependent (outcome) variable: ordinal

Independent (explanatory) variables: Continuous (scale) and/or Categorical

Common Applications: Regression is used to (a) *look for significant relationships* between two variables or (b) *predict* a value of one variable for a given value of the other.

Data: The data set 'Graduate.sav' looks at factors that influence the decision of whether to apply to graduate school. College juniors are asked if they are unlikely, somewhat likely, or very likely to apply to graduate school. Hence, our outcome variable has three categories.

College juniors are asked if they are unlikely, somewhat likely, or very likely to apply to graduate school. Hence, our outcome variable has three categories. Data on parental educational status, whether the undergraduate institution is public or private, and current GPA is also collected.

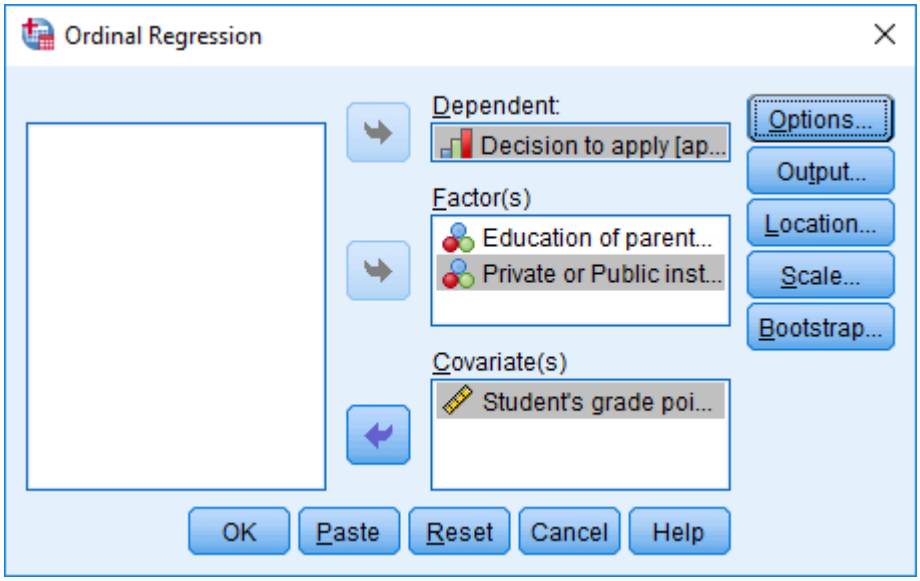
Assumptions for Ordinal regression

Assumptions		How to check
Proportional Odds		Test of parallel lines

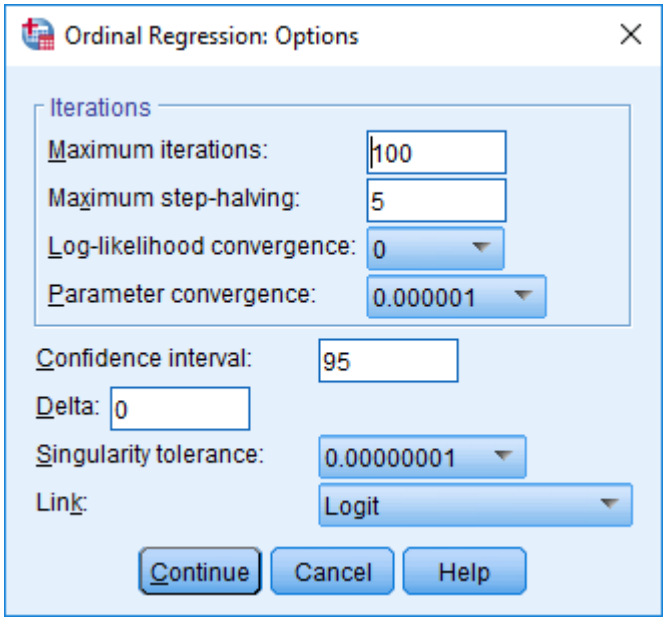
Steps in SPSS

Analyze → Regression → Ordinal

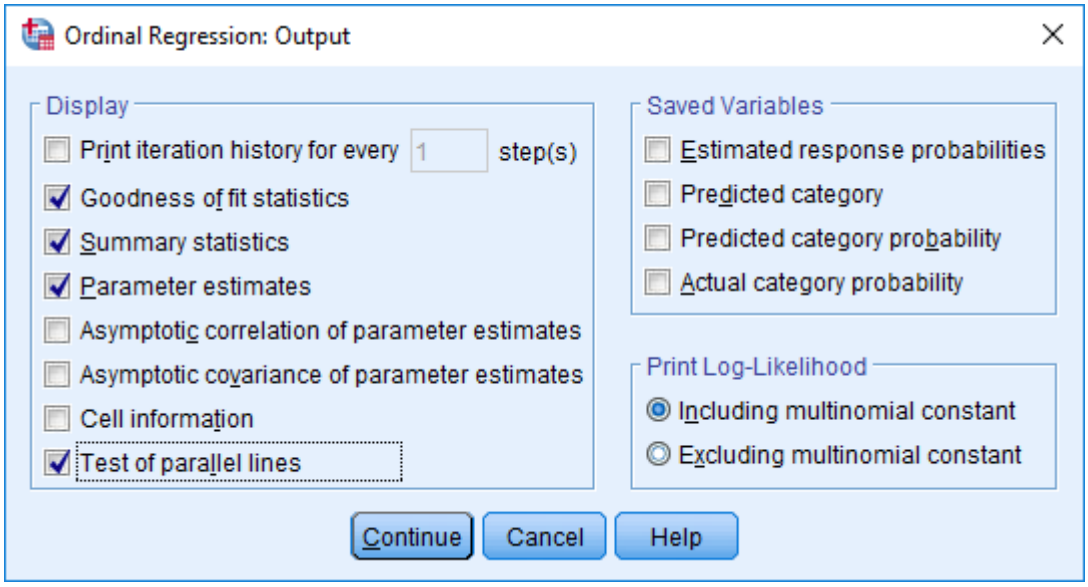
Move 'Decision to apply' to the *Dependent* box. The categorical independent variables 'Education of parents' and 'Private or Public institution' should be moved to the *Factor(s)* box. The scale independent variable GPA score should be moved to *Covariate(s)* box.



We will keep the default options:



We will tick “Test of parallel lines” as it will check the assumption of proportional odds.



Output

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	557.272			
Final	533.091	24.180	3	.000

Link function: Logit.

The p-value of less than 0.001 shows that the model is a very good finding on how well does the model fits the data.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	400.843	435	.878
Deviance	400.749	435	.879

Link function: Logit.

For this table, we would like to fail to reject the null hypothesis in order for our model to be a good fit. There are two tests (Pearson and Deviance). Both reject the null hypothesis.

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	533.091			
General	529.077	4.014	3	.260

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

This tests the assumption of proportional odds and we want it to be greater than 0.05. This is the case here (p-value = 0.26). The main assumption of the ordinal regression is checked.

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[apply = 1]	1.214	.939	1.671	1	.196	-.627	3.055
	[apply = 2]	3.310	.952	12.076	1	.001	1.443	5.177
Location	gpa	.616	.263	5.499	1	.019	.101	1.130
	[pared=.00]	-1.048	.268	15.231	1	.000	-1.574	-.522
	[pared=1.00]	0[a]	.	.	0	.	.	.
	[public=.00]	.059	.289	.041	1	.839	-.507	.624
	[public=1.00]	0[a]	.	.	0	.	.	.

Link function: Logit.

a. This parameter is set to zero because it is redundant.

The most interesting aspect is the locations. We have a statistically significant result for GPA score (p-value = 0.019). We also have a statistically significant result for variable “pared” (p-value < 0.001). On the other hand, the fact that the undergraduate institution is public or private does not have any statistically significant impact on the dependent variable “apply”.

The value of the GPA coefficient is positive (0.616), which suggests that as GPA increases the likelihood of applying to a graduate school will increase. The value of [pared=0] is negative (-1.048), which suggests that if you are in the category “pared=0”, i.e. no parents have a degree, you are less likely to apply to a graduate school.