# Step 3 Tutorial: Obtaining equations for scoring new cases

DemoData = 'data3.sav'

In Tutorial 1, we developed a 3-class latent class model using PURPOSE, ACCURACY, UNDERSTA, and COOPERAT as indicators. In this tutorial, we will obtain regression equations for scoring new cases as a function of those same indicators. That is, the indicators will now be used as predictors in the regression equation.

Recall the output file, data3.sav, created in Tutorial 1.

➢ Open data3.sav using "File→Open"
➢ From the menu bar, click 'Model' and select 'Step3'. A dialog box will pop up
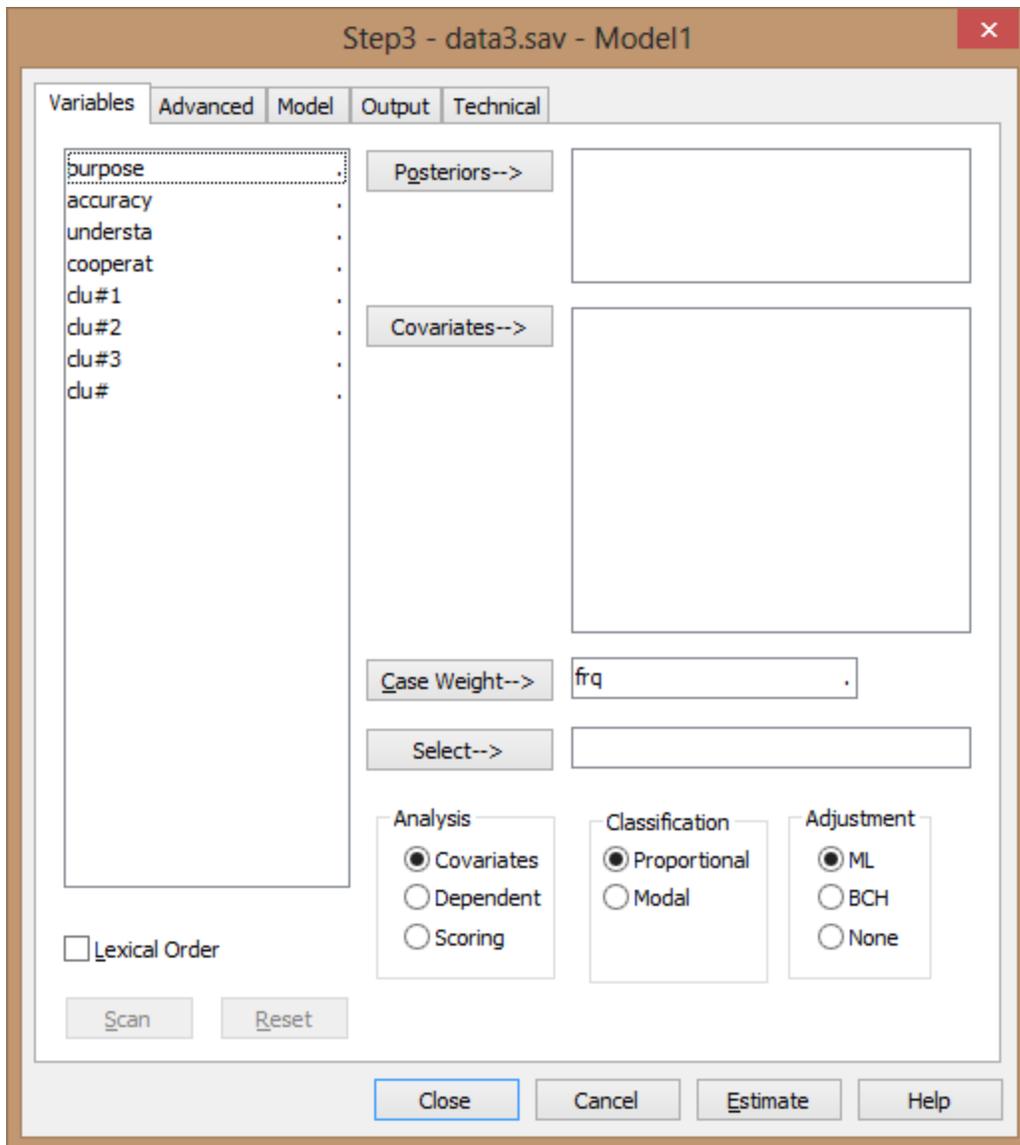
➢ For Analysis, select 'Scoring'
➢ Click on the 4 indicators and click 'Variables' to move them to the Variables box
➢ Right click on the 4 predictors and select 'Nominal'
➢ Click on clu#1, clu#2, clu#3 and click 'Posteriors' to move them to the Posteriors box

➢ Click 'Scan'
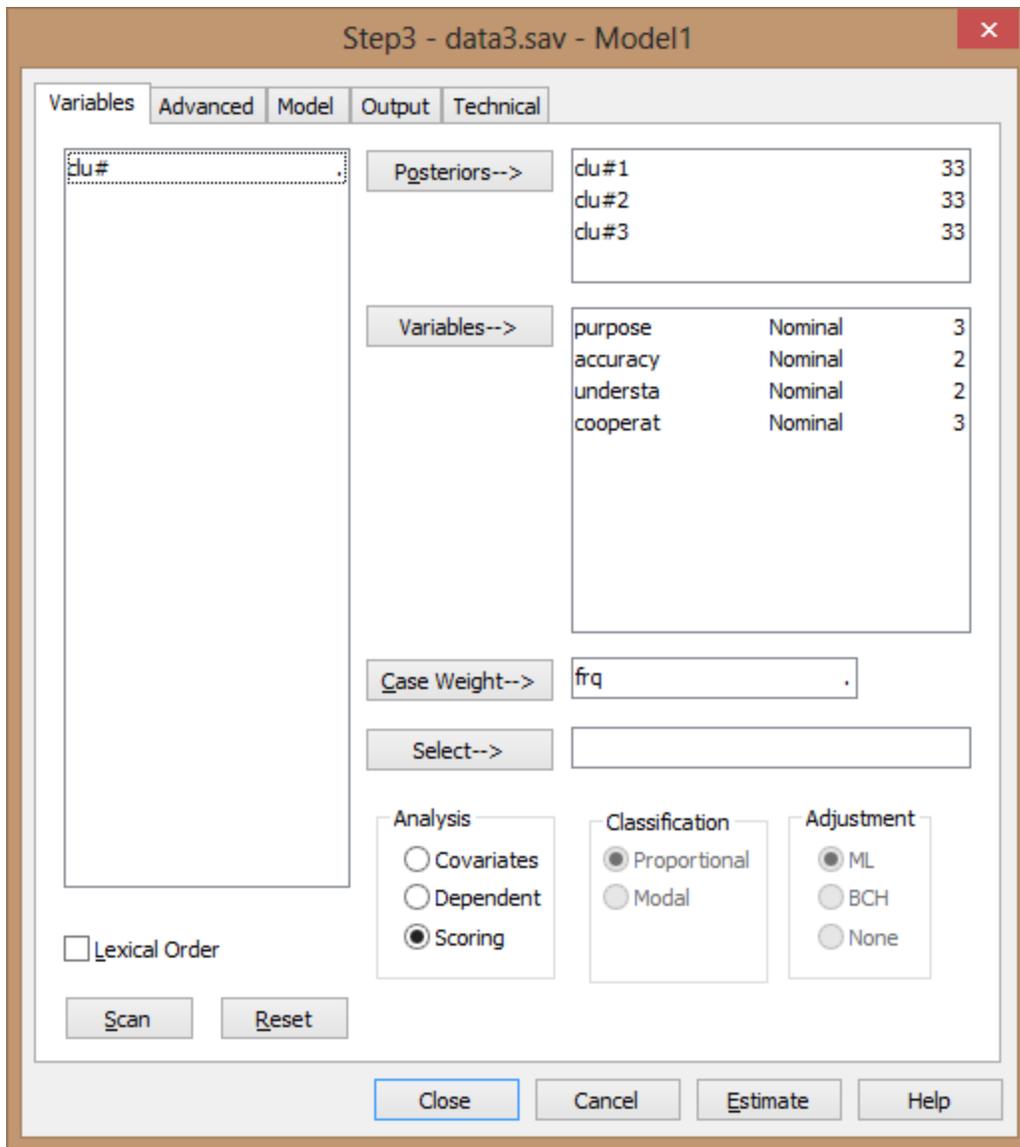
- ➢ Click on the 'Technical Tab' and set all Bayes constants to '0' (the Bayes constant option will be explained in Session 2)
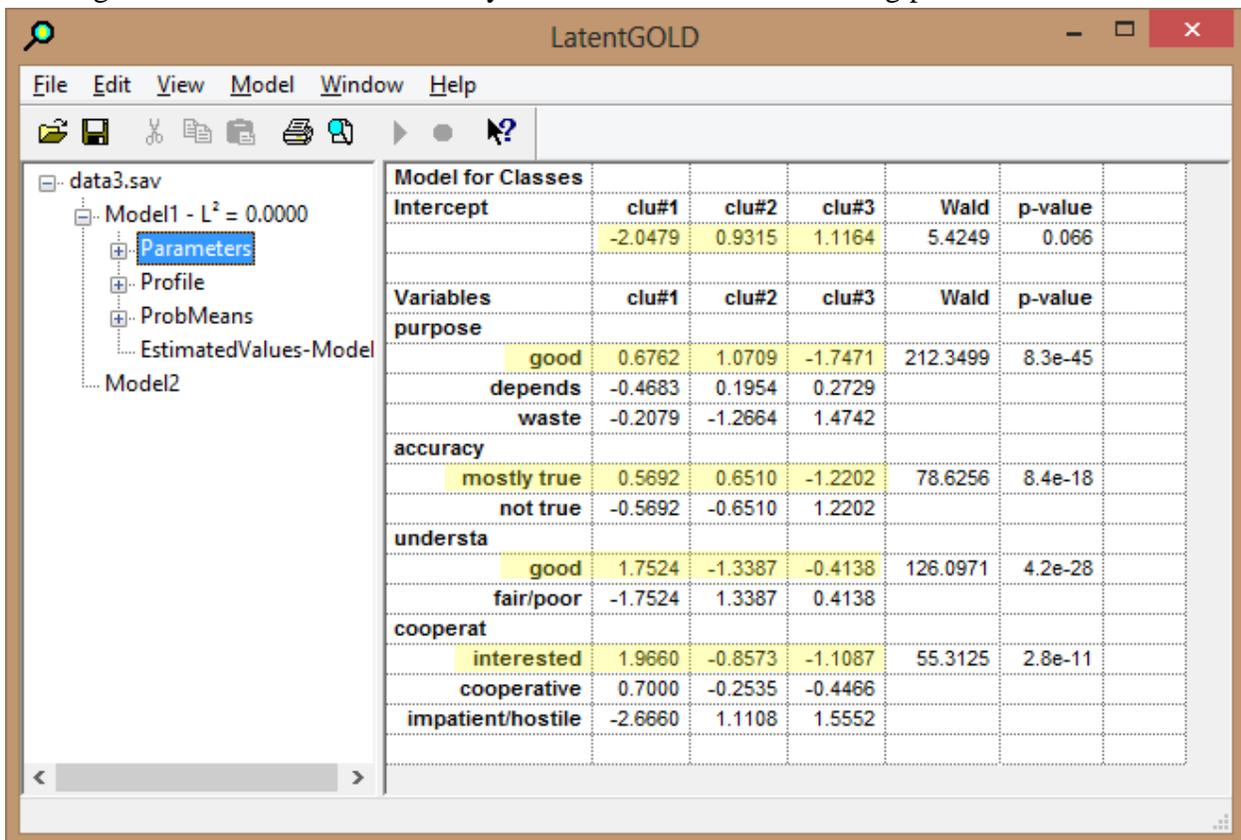- ➢ Click 'Estimate'

Confirm that the $L^2 = 0$.

Classification Statistics

Classification errors	0.1314
Reduction of errors (Lambda)	0.6570
Entropy R-squared	0.5794
Standard R-squared	0.6116
Classification log-likelihood	-470.2322
AWE	1181.0332
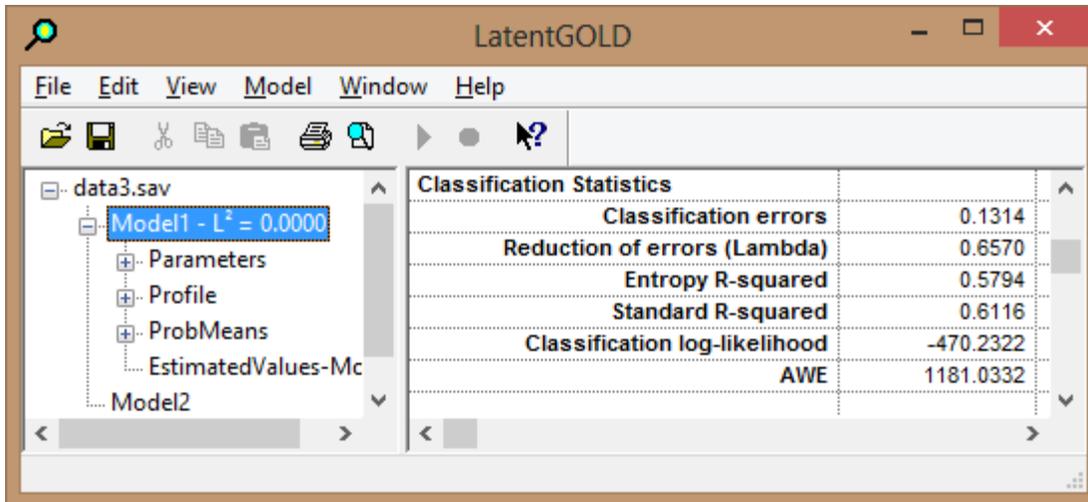
The equations are provided in the Parameters output.

➢ Click Parameters to view the coefficients in the Parameters output

Note: Ignore the other statistics as they are not relevant for the scoring procedure

Model for Classes

Intercept	clu#1	clu#2	clu#3	Wald	p-value
	-2.0479	0.9315	1.1164	5.4249	0.066

Variables	clu#1	clu#2	clu#3	Wald	p-value
purpose					
good	0.6762	1.0709	-1.7471	212.3499	8.3e-45
depends	-0.4683	0.1954	0.2729		
waste	-0.2079	-1.2664	1.4742		
accuracy					
mostly true	0.5692	0.6510	-1.2202	78.6256	8.4e-18
not true	-0.5692	-0.6510	1.2202		
understa					
good	1.7524	-1.3387	-0.4138	126.0971	4.2e-28
fair/poor	-1.7524	1.3387	0.4138		
cooperat					
interested	1.9660	-0.8573	-1.1087	55.3125	2.8e-11
cooperative	0.7000	-0.2535	-0.4466		
impatient/hostile	-2.6660	1.1108	1.5552		

Each response pattern is scored on each cluster, and is assigned to the cluster with the highest score. For example, the response pattern:

Purpose = Good, Accuracy = Mostly True, Understanding = good, Cooperation = Interested

is scored based on the coefficients highlighted above in yellow. This results in the following scores:

Cluster 1 score = 2.916, Cluster 2 score = 0.457, Cluster 3 score = -3.373.

Thus, this response pattern is assigned to Cluster 1. To obtain more meaningful scores, we can generate the posterior membership probabilities using the formula provided below. This yields the following posterior membership probabilities associated with this response pattern:

Posterior 1 = 0.9196, Posterior 2 = 0.0787, Posterior 3 = 0.0017
To generate the corresponding posteriors, use the following equation:

$Posterior(k) = \exp[score(k)]/ [\ \exp(score1) + \exp(score2) + \exp(score3)]$ k=1,2,3.