# Step 3 Tutorial #3: Obtaining equations for scoring new cases in an advanced example with quadratic term

**DemoData = 'diabetes.lgf', 'diabetes.dat', 'data5.dat'**

We begin by opening a saved 3-class latent class model using GLUCOSE, Insulin, and SSPG as indicators, where the variances each of these indicators was specified to be class-dependent, and a direct effect between INSULIN and GLUCOSE was included in the model ('model 5'). In this tutorial, we will show how to use the Step3 module in Latent GOLD 5.0 to obtain an algorithm (equations) and related SPSS syntax for scoring new cases based on this model. That is, in these equations the 3 indicators will be used as predictors.

Open the saved model definition

➢ Open diabetes.lgf using "File→Open"
➢ Double click on Model 5. A dialogue box will pop up.
➢ Click on the ClassPred tab.
➢ Check the 'Classification – Posterior' box to request the posterior membership probabilities to be output to a file (see Figure 1)
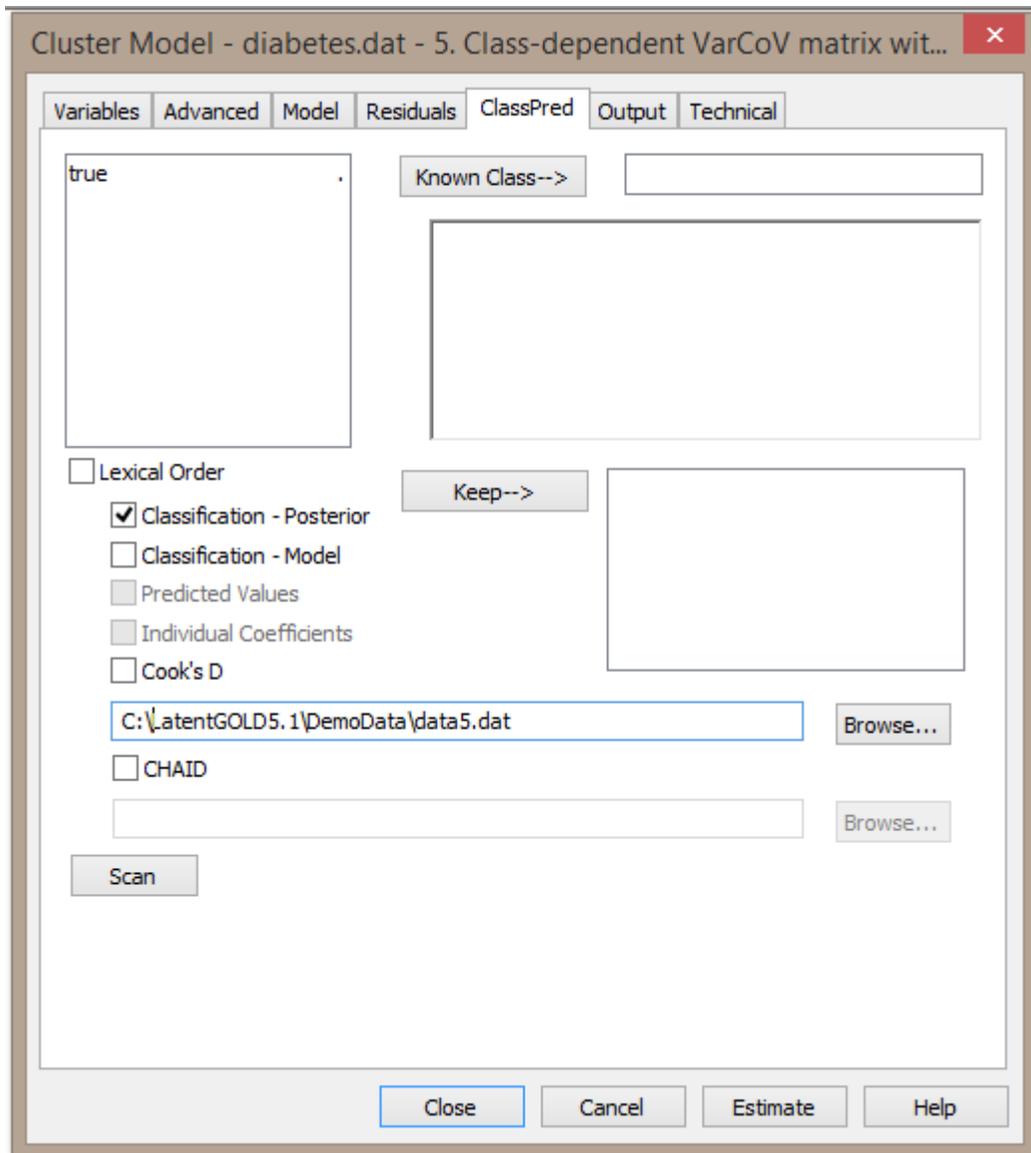
**Figure 1.** Requesting the posterior membership probabilities to be output to a file.

➤ Click Estimate

➤ Open data5.dat using "File→Open"
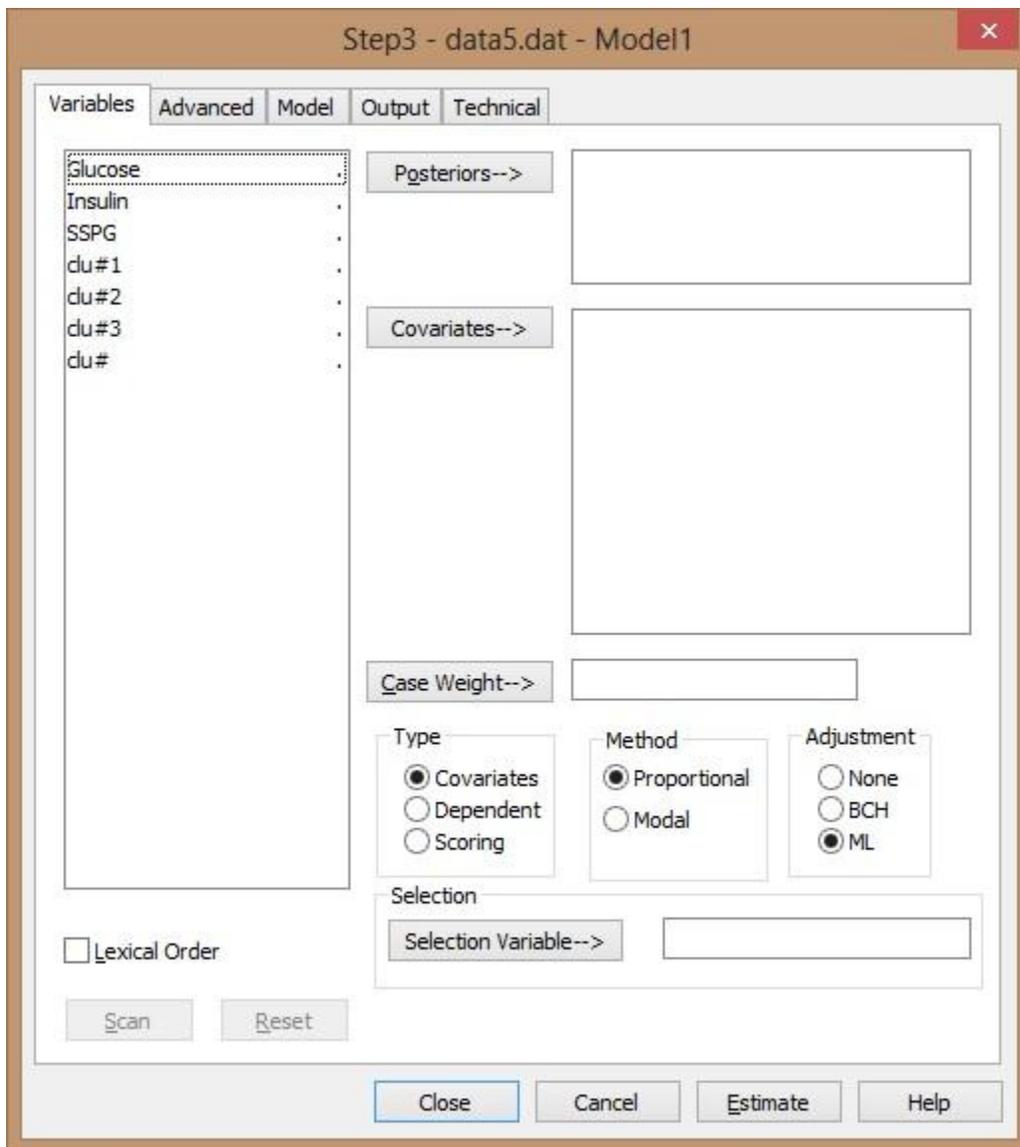➤ From the menu bar, click 'Model' and select 'Step3'. A dialog box will pop up (Figure 2).

**Figure 2.** Variables tab of Step3 Module

➢ Click on the 3 indicators and click 'Covariates' to move them to the Covariates box
➢ Click on clu#1, clu#2, clu#3 and click 'Posteriors' to move them to the Posteriors box
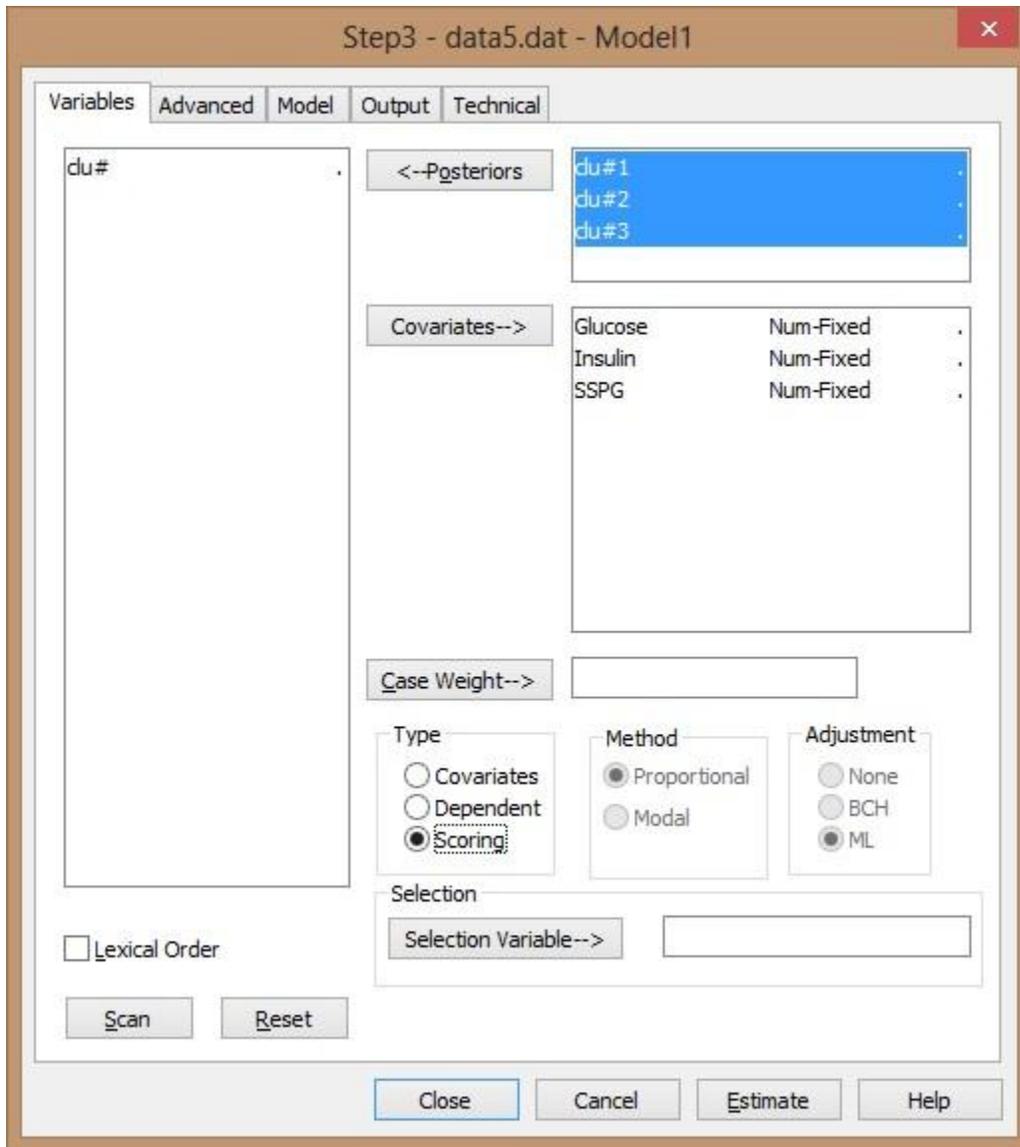➢ For Type, select 'Scoring' (see Figure 3)

**Figure 3.** Step 3 Variables Tab: Selecting the variables and the analysis type

➢ Click on the 'Model Tab' and check the associated boxes to include the following quadratic terms (see Figure 4):
  ➢ Glucose * Glucose
  ➢ Insulin * Glucose
  ➢ Insulin * Insulin
  ➢ SSPG * SSPG

The squared terms for each indicator is included because variances for these indicators are specified to be class dependent, and the INSULIN by GLUCOSE interaction corresponds to the associated direct effect that is included in the model. Note that inclusion of additional quadratic

terms will have no effect, since the coefficients for those additional terms will turn out to be zero.
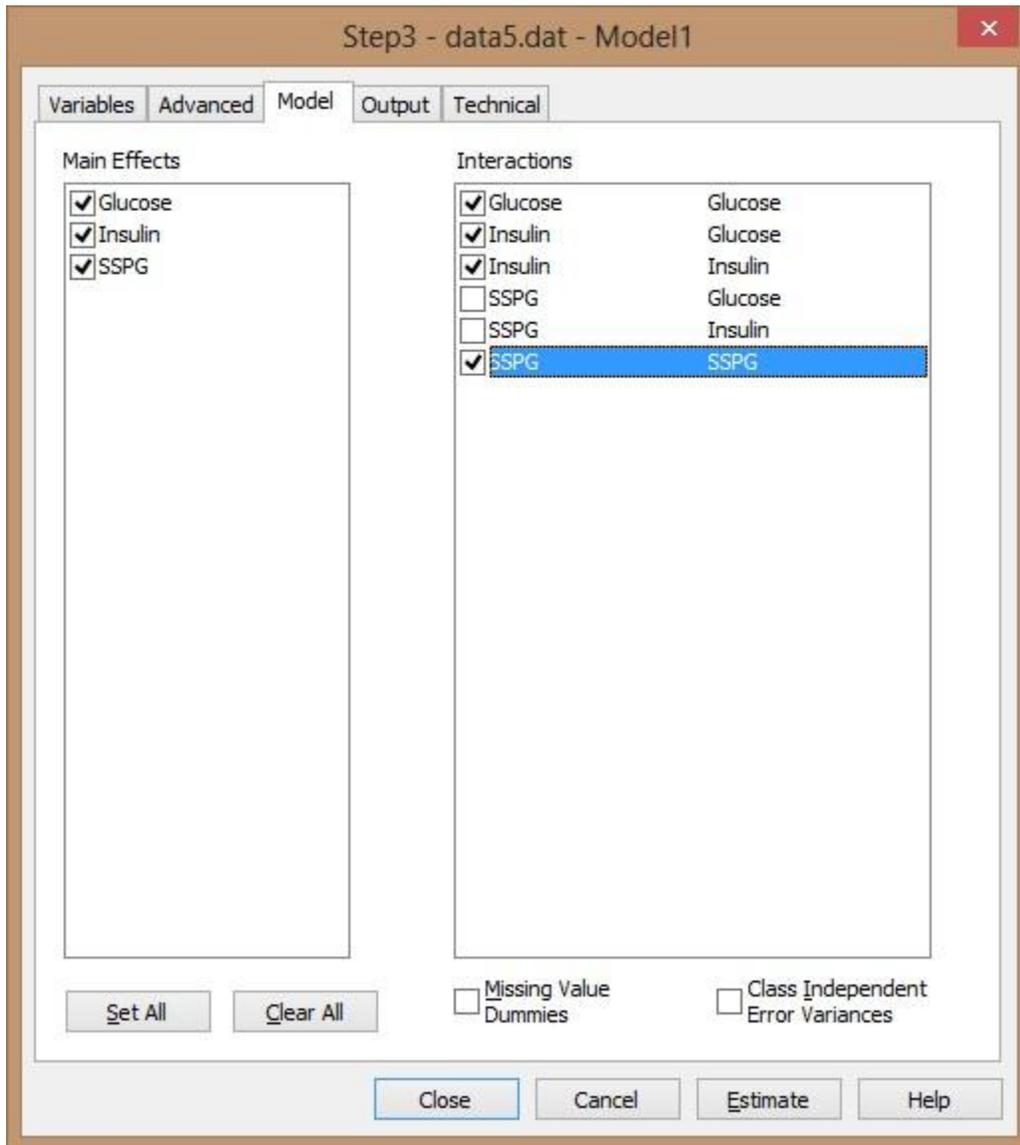


**Figure 4.** Step 3 Model Tab: Specifying interaction effects

➢ Click on the 'Technical' tab and set all Bayes constants to '0'
➢ Click 'Estimate'

Confirm that the $L^2 = 0$ (see Figure 5), which means that the posterior probabilities are reproduced perfectly as a function of the 4 predictors. Later, we will also show that the Profile and Probmeans output obtained here reproduce perfectly the Profile and Probmeans output produced during the development of the original Model 5.
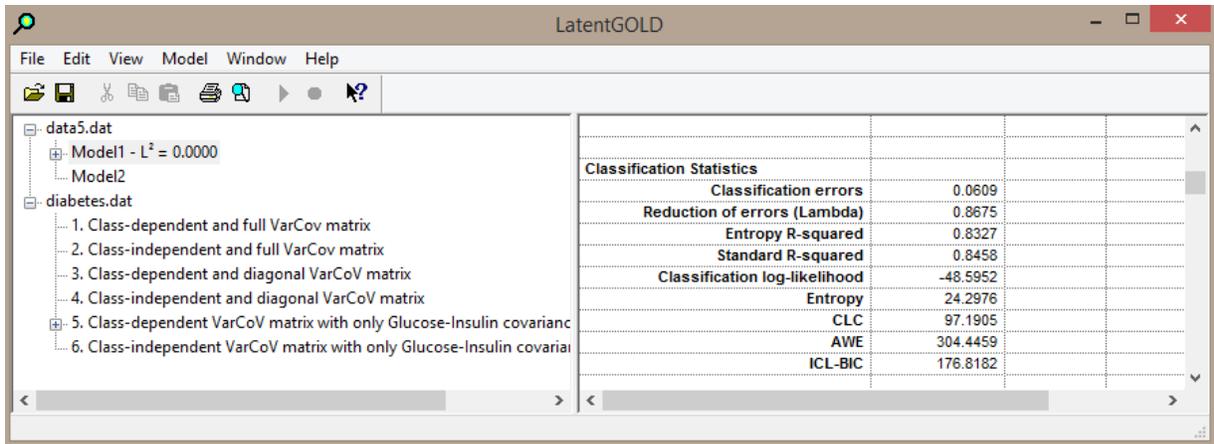
**Figure 5.** Step3 results

$L^2=0$ means that the predicted posterior probabilities reproduce the original posterior probabilities obtained previously from Model 5. To confirm this,

➢ Click on EstimatedValues to view the predicted posterior probabilities (see Figure 6)
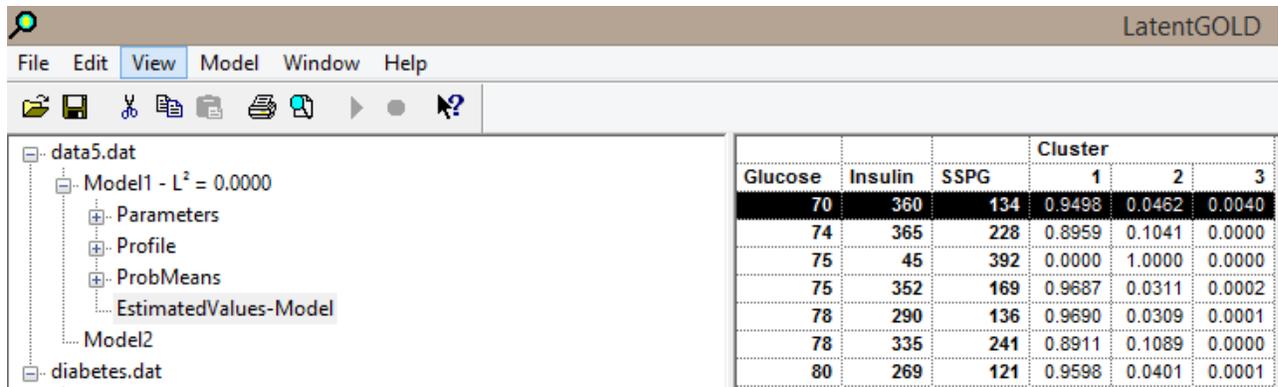
**Figure 6.** EstimatedValues-Model output

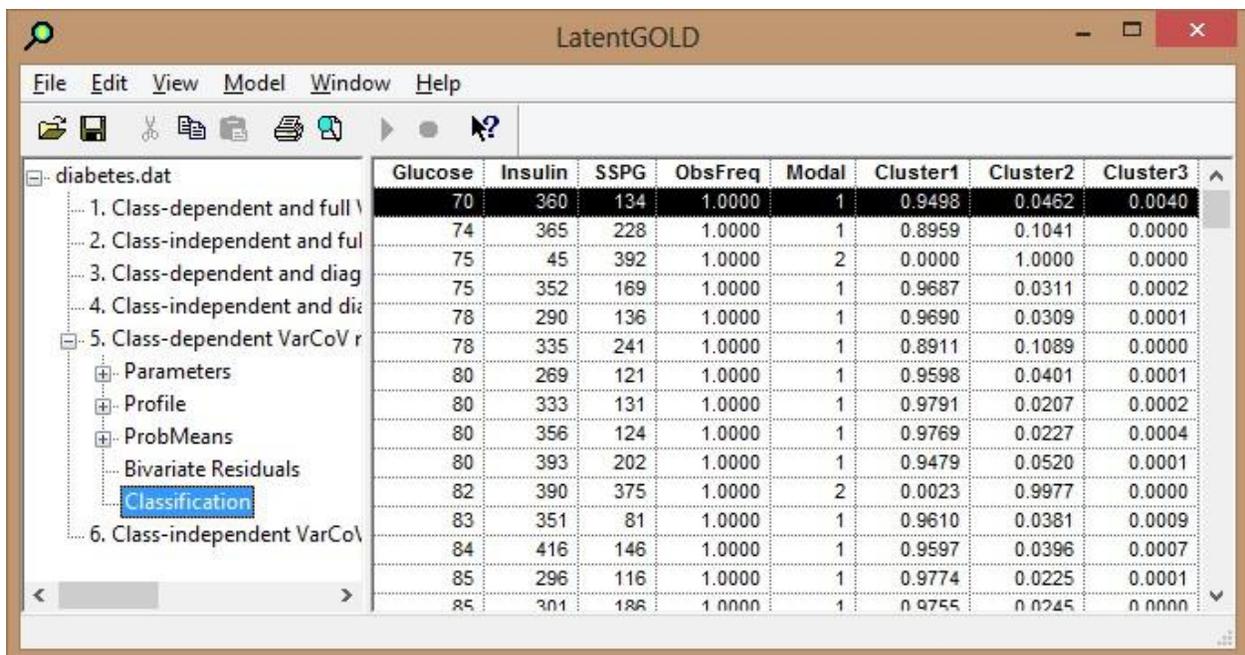The below (from Model 5) shows that these are identical to the posteriors previously estimated.



**Figure 7.** Classification output from Model 5.

The equations used to produce the predicted posterior probabilities are provided in the Parameters output.

➢ Click Parameters to view the coefficients in the Parameters output (see Figure 8)



**Figure 8.** Step3 Parameters output

Ignore the Wald and p-value output since that is not relevant for this scoring application.

Note that some of the interaction effects require additional precision. The number of decimal places can be changed in any of the output listings using the Format Control. Since some of the coefficients appear in the output as 0 to 4 decimal places, we will increase the number of decimals to 10. To display the format control for the current output listing:

➢ Click Edit from within the Contents Pane
➢ Select Numbers
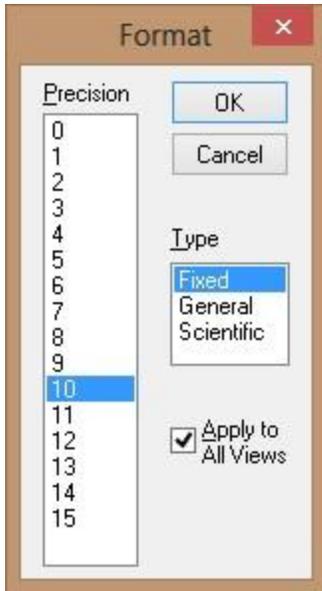➢ Under Precision, click '10' and then click OK (see Figure 9)

**Figure 9.** Numbers Format



**Figure 10.** Step3 Parameters output with 10 decimal places for values

Thus, the equations are:

Score1= -33.1665606493 + 0.5637792152* Glucose + 0.0575831434* Insulin + 0.0288858610* SSPG + -0.0029160924* Glucose$^2$ + -0.0000359976* Insulin * Glucose + -0.0000916390* Insulin$^2$ + -0.0000554785* SSPG$^2$

Score2= 9.4764746316 + 0.0038548614* Glucose + -0.0490425314* Insulin + -0.0250084510* SSPG + -0.0002419977* Glucose$^2$ + 0.0001891160* Insulin * Glucose + 0.0000398600* Insulin$^2$+ 0.0001292385* SSPG$^2$

Score3= 23.6900860178+ -0.5676340766* Glucose + -0.0085406120* Insulin + -0.0038774100 * SSPG + 0.0031580902* Glucose$^2$ + -0.0001531184 * Insulin * Glucose + 0.0000517790 * Insulin$^2$ + -0.0000737600 * SSPG$^2$

Next, we will show how to use these equations to obtain the predicted posterior membership probabilities, illustrating the calculations for the case defined by Glucose = 70, Insulin = 360, and SSPG = 134 (see Figure 11).

**Figure 11.** Step3 EstimatedValues-Model output

For example, Score1 = 2.830042838 for the case

To calculate the logit scores:

Score1= -33.1665606493 + 0.5637792152* 70 + 0.0575831434* 360+ 0.0288858610* 134 + -0.0029160924* $70^2$ + -0.0000359976* 360* 70 + -0.0000916390* $360^2$ + -0.0000554785* $134^2$

Score2= 9.4764746316 + 0.0038548614* 70+ -0.0490425314* 360 + -0.0250084510* 134+ -0.0002419977* $70^2$ + 0.0001891160* 360* 70 + 0.0000398600* $360^2$+ 0.0001292385* $134^2$

Score3= 23.6900860178+ -0.5676340766* 70 + -0.0085406120* 360 + -0.0038774100 * 134 + 0.0031580902* $70^2$ + -0.0001531184 * 360 * 70+ 0.0000517790 * $360^2$ + -0.0000737600 * $134^2$

The resulting scores are:

Score1 = 2.8300427867
Score2 = -0.1937318324
Score3 = -2.6363104642

Exponentiating these scores yields:

Score1 = 16.9461858804
Score2 = 0.8238788123
Score3 = 0.0716250458

The sum for these 3 exponentiated scores = 17.8416897385

Dividing each exponentiated score (S1, S2 and S3) by this sum yields the predicted posteriors (these values match those reported in the EstimatedValues output):

Score1= 0.9498083494
Score2 = 0.0461771740
Score3 = 0.0040144766

An easier way to obtain the scoring equations is by selecting 'Scoring Syntax' (default Type is an .sps file) in the Output tab: