SuperMix demonstration

1 Mixed models for continuous outcomes

1.1 The data

The data set used here is from the Television School and Family Smoking Prevention and Cessation Project (TVSFP)(Flay *et. al.*, 1988). The study was designed to test independent and combined effects of a school-based social-resistance curriculum and a television-based program in terms of tobacco use and cessation. The data from the study included a total of 1,600 students from 135 classrooms drawn from 28 schools. Schools were randomized to one of four study conditions:

- a social-resistance classroom curriculum
- a media (television) intervention
- a social-resistance classroom curriculum combined with a mass-media intervention, and
- a no-treatment control group

A tobacco and health knowledge scale (THKS) was used in classifying subjects as knowledgeable or not. In its original form, the student's score was defined as the number of correct answers to seven items on tobacco and health knowledge.

Data for the first 10 students on most of the variables used in this section are shown below in the form of an SuperMix spreadsheet file, named **TV2dat.xls**.

tv2dat.ss3

	(A) School	(B) Class	(C) THKScore	(D) Intrcp	(E) PreTHKS	(F) CC	(G) TV	(H) CCxTV
1	403	403101	3	1	2	1	0	0
2	403	403101	4	1	4	1	0	0
3	403	403101	3	1	4	1	0	0
4	403	403101	4	1	3	1	0	0
5	403	403101	4	1	3	1	0	0
6	403	403101	3	1	4	1	0	0
7	403	403101	2	1	2	1	0	0
8	403	403101	4	1	4	1	0	0
9	403	403101	5	1	5	1	0	0
10	403	403101	4	1	3	1	0	0

The variables of interest are:

- SCHOOL indicates the school a student is from (28 schools in total).
- CLASS identifies the classroom (135 classrooms in total).
- THKScore represents the post-intervention tobacco and health knowledge scale. It is treated as a continuous variable here.

- PRETHKS indicates the pre-intervention THKS score.
- CC is a binary variable indicating whether a social-resistance classroom curriculum was introduced, where 0 indicates "no" and 1 "yes."
- TV is an indicator variable for the use of media (television) intervention, with a "1" indicating the use of media intervention, and "0" the absence thereof.
- CCxTV was constructed by multiplying the variables TV and CC, and represents the CC by TV interaction.

1.2 Graphical displays

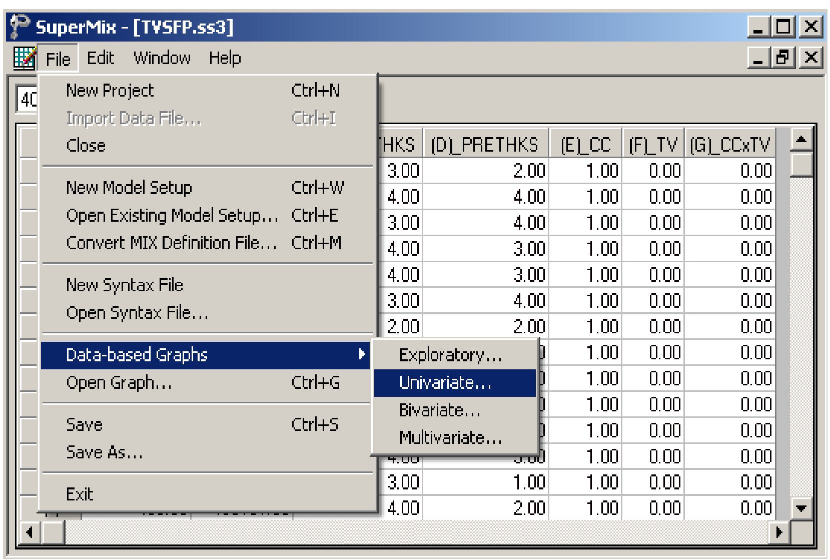
Use the **File**, **Import Data File…** option to locate **tv2dat.xls** in the **continuous** subfolder and click the **Open** button to obtain the SuperMix spreadsheet file **tv2dat.ss3**. Right-click on the THKScore column header to obtain the **Column Properties** dialog box and change the **Header** name to POSTTHKS. Next, right-click on the Intrcp column and select the **Delete Column** option from the pop-up menu. These changes are saved to **TVSFP.ss3** in the **examples\continuous** folder.

Univariate graphs

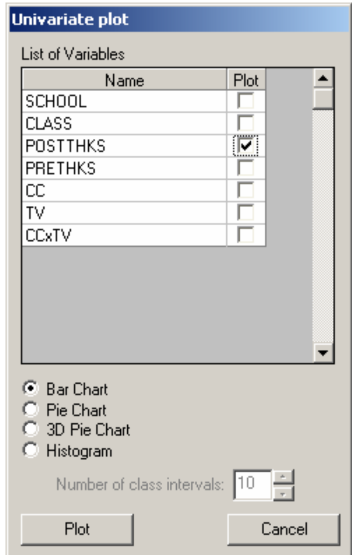
The pop-up menu below shows the data-based graphing options currently available in SuperMix. As a first step, we will take a closer look at the distribution of the total post-intervention scores (POSTTHKS), which is the potential dependent variable in this study. While scores such as these are not truly continuous variables, they are often treated as if they were.

Bar chart

To do so, select the **Univariate** option from the **Data-based Graphs** menu as shown below.



The **Univariate plot** dialog box appears. Select the variable POSTTHKS and indicate that a **Bar Chart** is to be graphed. Click the **Plot** button to display the bar chart.



The bell-shaped bar chart below shows that the variable POSTTHKS is approximately normally distributed. Note that histograms are usually used for the depiction of the distribution of a continuous variable.

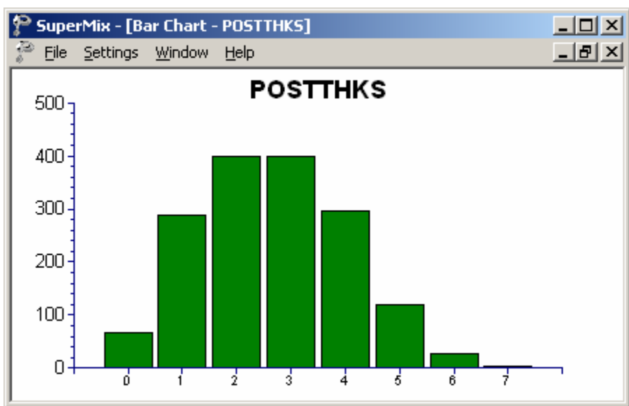


Figure 1.1: Bar chart of POSTTHKS scores

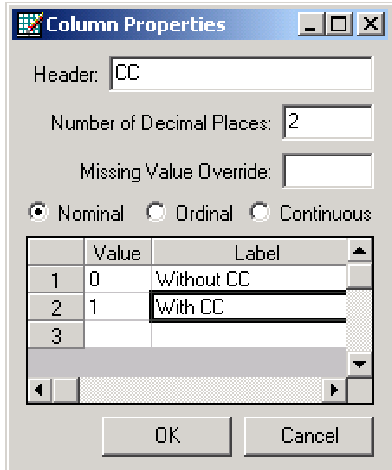
Bivariate graphs

It is hoped that the social-resistance classroom curriculum (CC), the television intervention (TV) and the CC and TV interaction combination (CCxTV) would affect the tobacco and health knowledge (POSTTHKS). Before we start with the model, we would like to show a box-and-whisker plot of POSTTHKS for each category of CC.

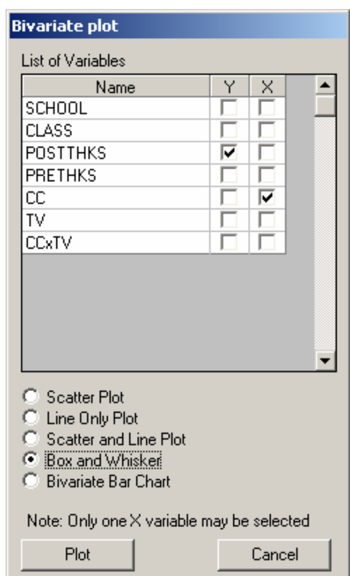
Box-and-whisker plots

A box-and-whisker plot is useful for depicting the locality, spread and skewness of variables in a data set and may be used to examine the distributions of continuous variables, such as for the different values of discrete valued predictors. This option is accessed via the **Data-based Graphs**, **Bivariate** option on the **File** menu.

To assign labels to the categories of CC, right-click on the CC column in the spreadsheet and select **Column Properties**. On the **Column Properties** dialog box, select the **Nominal** option and assign the appropriate labels.



The **Bivariate plot** dialog box is completed as shown below: select the outcome variable POSTTHKS as the **Y**-variable of interest, and the predictor CC to be plotted on the **X**-axis. Check **the Box and Whisker** option, and click **Plot**.



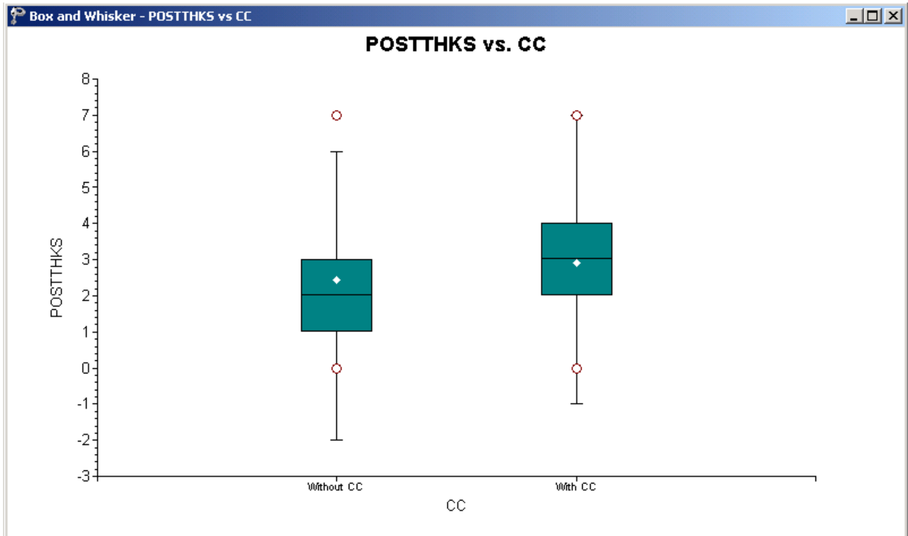


Figure 1.2: Box-and-whisker plots of POSTTHKS scores for different CC values

The bottom line of a box represents the first quartile (q_1), the top line the third quartile (q_3), and the in-between line the median (me). The arithmetic mean is represented by a diamond. Here, the mean of POSTTHKS is lower in the group without the social-resistance classroom curriculum (CC). The box-and-whisker plot indicates a positive relationship between CC and POSTTHKS.

1.3 A 2-level random intercept model using classroom as level-2 ID

The model

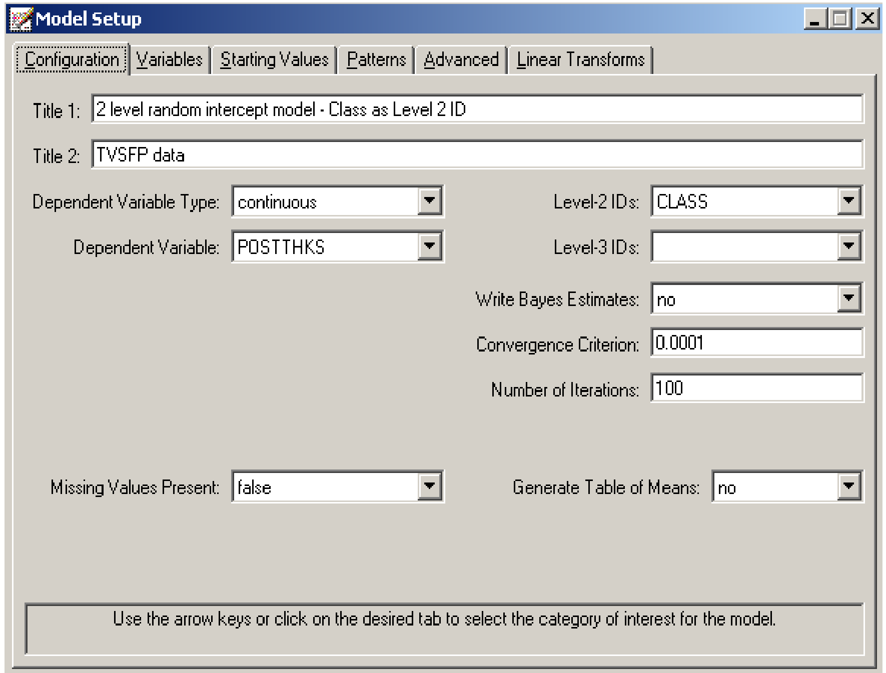
The first model fitted to the data explores the cluster effects of each classroom on the outcome. The mixed model can be expressed as

$$\text{POSTTHKS}_{ij} = \beta_0 + \beta_1 \text{CC}_i + \beta_2 \text{TV}_i + \beta_3 (\text{CC}_i \times \text{TV}_i) + v_{0i} + \varepsilon_{ij} ,$$

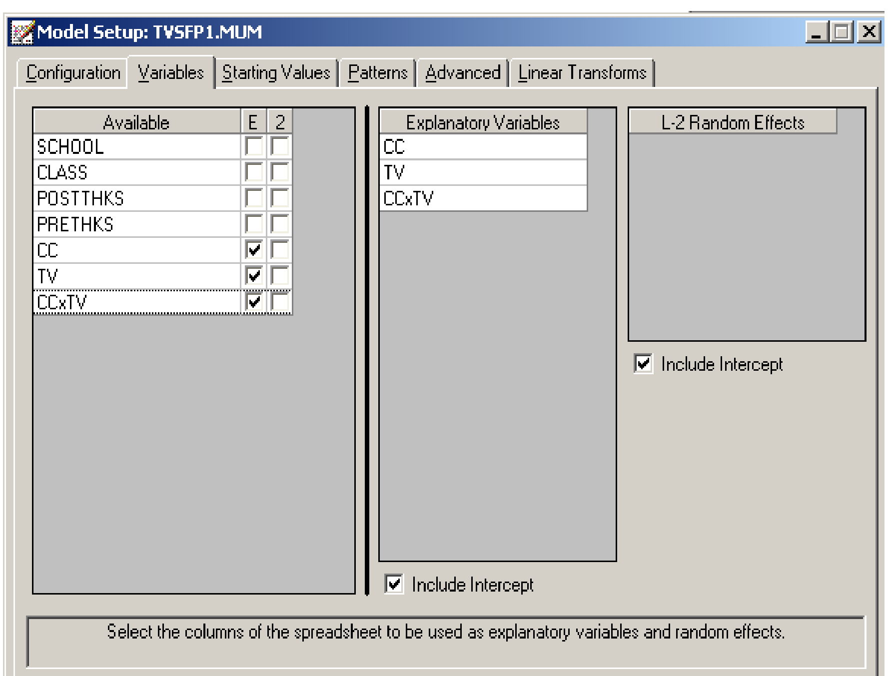
where v_{0i} represents the classroom influence on POSTTHKS.

Setting up the analysis

From the main menu bar, select the **File**, **New Model Setup** option. Select the continuous outcome variable POSTTHKS from the **Dependent Variable** drop-down list box. Select the classroom number CLASS from the **Level-2 IDs** drop-down list box. Enter a title for the analysis in the **Title** text boxes. In this example, default settings for all other options associated with the **Configuration** screen are used.



Proceed to the **Variables** screen by clicking on that tab. The **Variables** screen is used to specify the fixed and random effects to be included in the model. Select the explanatory (fixed) variables using the **E** check boxes next to the variables names in the **Available** grid at the left of the screen. Note that, as the variables are selected, the selected variables are listed in the **Explanatory Variables** grid. After selecting all the explanatory variables, the screen shown below is obtained. The **Include Intercept** check box in the **Explanatory Variables** grid is checked by default, indicating that an intercept term will automatically be included in the fixed part of the model.



Next, specify the random effects at level 2 the hierarchy. In this example, we want to fit a model with random intercepts at level 2. By default, the **Include Intercept** check box in the **L-2 Random Effects** grid is checked. If this box is left checked, and no additional random effects are indicated using the **2** column in the **Available** grid to the left, the model fitted will be the random-intercepts-only model we intend to use. No further changes on this screen are necessary.

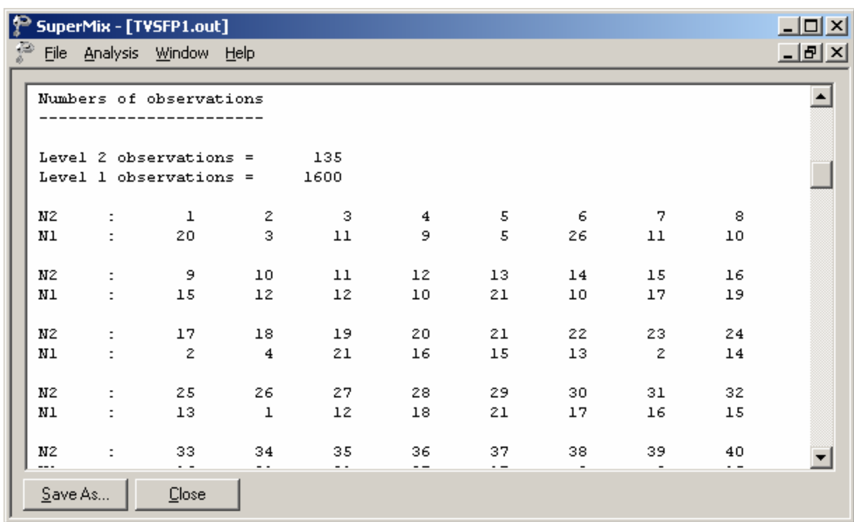
Before running the analysis, the model specifications have to be saved. Select the **File**, **Save As** option, and provide a name (**TVSFP1.mum**) for the model specification file. Run the analysis by selecting the **Run** option from the **Analysis** menu.

1.4 Discussion of results

Data summary

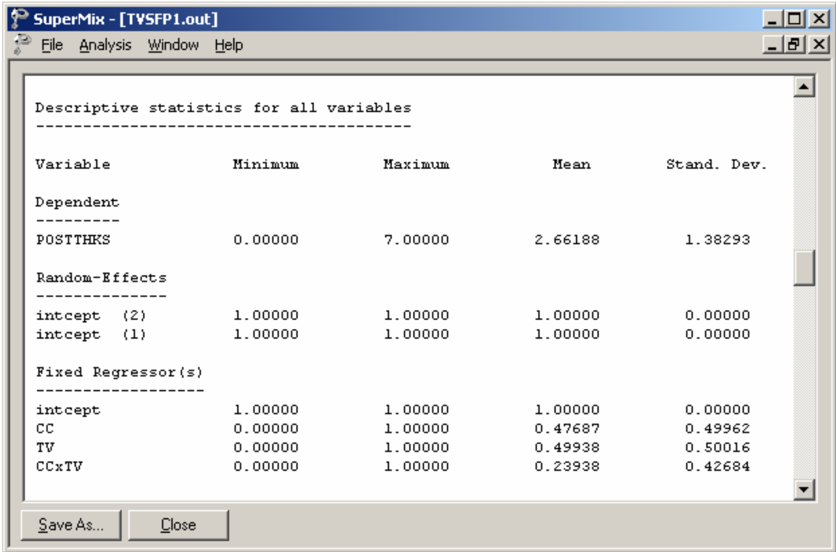
In the **numbers of observations** section, a summary of the hierarchical structure is provided.

As shown below, data from a total of 1600 students within 135 classrooms were included at levels 2 and 1 of the model. This corresponds to the study design described earlier. In addition, a summary of the number of students nested within each classroom is provided. The classroom with N2 = 6, for example, had 26 students (N1: 26). By contrast, classroom 26 had only 1 student.



Descriptive statistics and starting values

Next, the **descriptive statistics for all variables** are given.

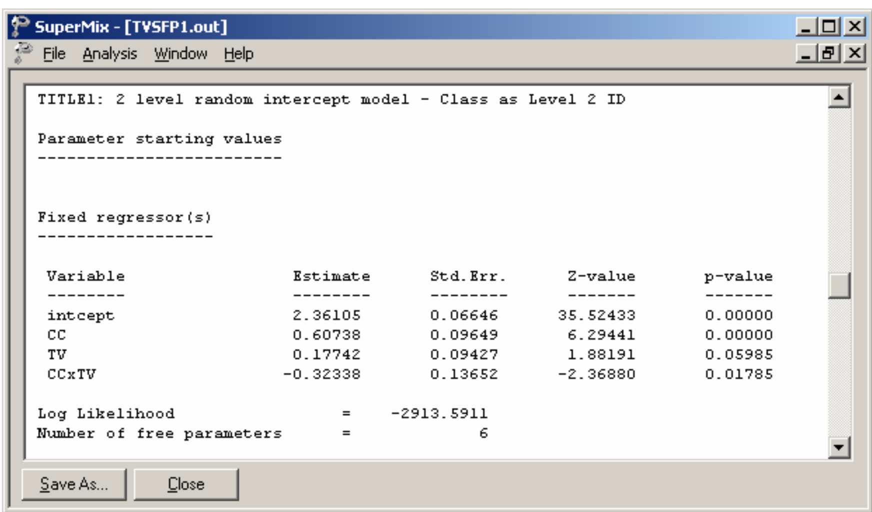


```
SuperMix - [TVSFP1.out]
File  Analysis  Window  Help

Descriptive statistics for all variables
----------------------------------------

Variable             Minimum        Maximum         Mean        Stand. Dev.

Dependent
---------
POSTTHKS             0.00000        7.00000        2.66188        1.38293

Random-Effects
--------------
intcept   (2)        1.00000        1.00000        1.00000        0.00000
intcept   (1)        1.00000        1.00000        1.00000        0.00000

Fixed Regressor(s)
------------------
intcept              1.00000        1.00000        1.00000        0.00000
CC                   0.00000        1.00000        0.47687        0.49962
TV                   0.00000        1.00000        0.49938        0.50016
CCxTV                0.00000        1.00000        0.23938        0.42684

Save As...   Close
```

The minimum value, maximum value, mean and standard deviation are given for all the variables included in the model. For example, the mean POSTTHKS is 2.6618 with a standard deviation of 1.38293.

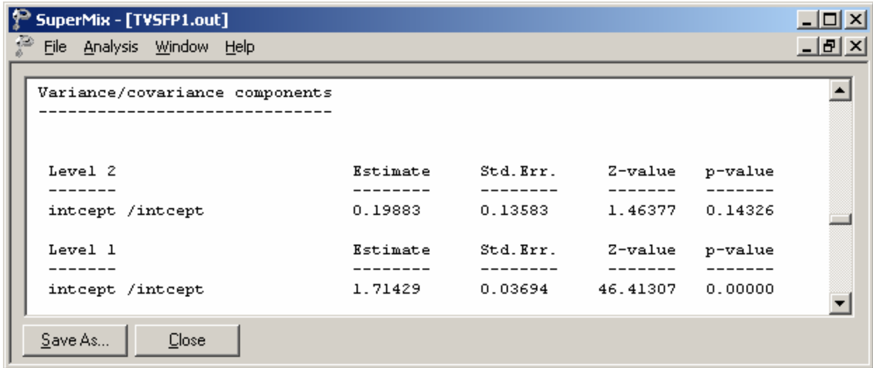
Starting values – OLS estimates

The starting values for the **fixed regressor(s)** are shown below. The **log likelihood** value and **number of free parameters** of the OLS regression are given in this part of the output.



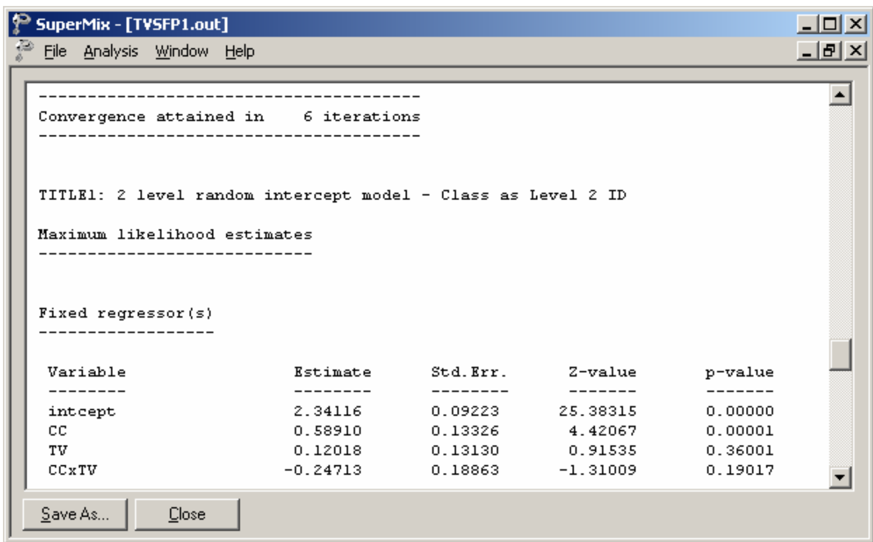
```
SuperMix - [TVSFP1.out]
File  Analysis  Window  Help

TITLE1: 2 level random intercept model - Class as Level 2 ID

Parameter starting values
-------------------------


Fixed regressor(s)
------------------

 Variable                 Estimate      Std.Err.      Z-value       p-value
 --------                 --------      --------      -------       -------
 intcept                   2.36105       0.06646      35.52433       0.00000
 CC                        0.60738       0.09649       6.29441       0.00000
 TV                        0.17742       0.09427       1.88191       0.05985
 CCxTV                    -0.32338       0.13652      -2.36880       0.01785

Log Likelihood                =     -2913.5911
Number of free parameters     =              6

Save As...   Close
```

After the **number of free parameters**, the starting values of **variance/covariance components** are reported as shown.



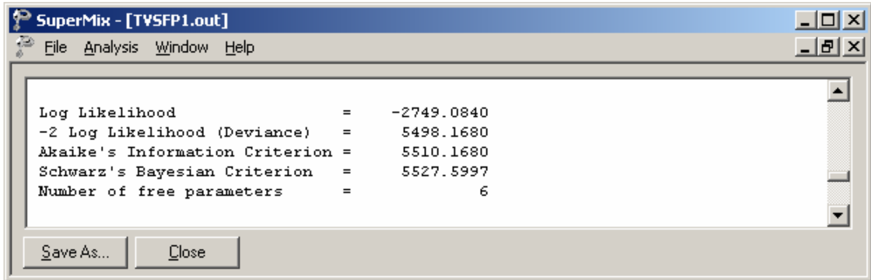
Fixed effects estimates

The number of iterations needed to obtain convergence is given after the starting values. The output describing the estimated **fixed regressor(s)** after convergence is shown next.



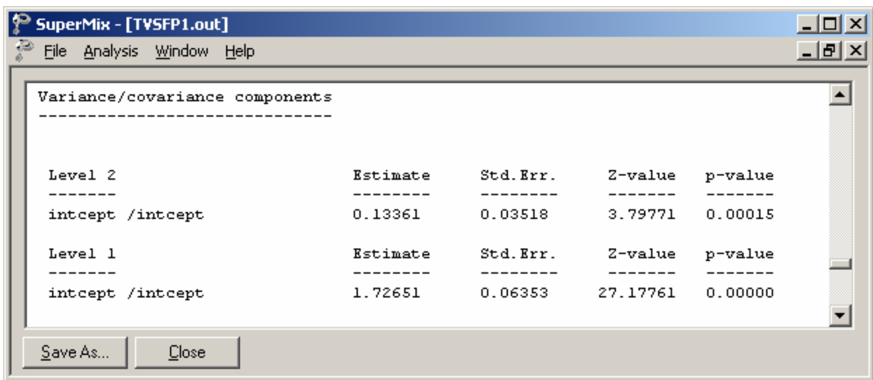
As shown above, the estimates for CC and TV are both positive. On average, a social-resistance classroom curriculum can improve the tobacco and health knowledge by 0.58910, and television intervention can increase the POSTTHKS score by 0.12018. However the estimate of CCxTV is negative, which implies that the students who had both CC and TV are expected to show a decrease of 0.24713 in their POSTTHKS score. The estimates associated with intercept and TV are highly significant, but estimates of the other two coefficients are not statistically significantly different from zero.

The estimates for the fixed regressors and model fit statistics are given next.

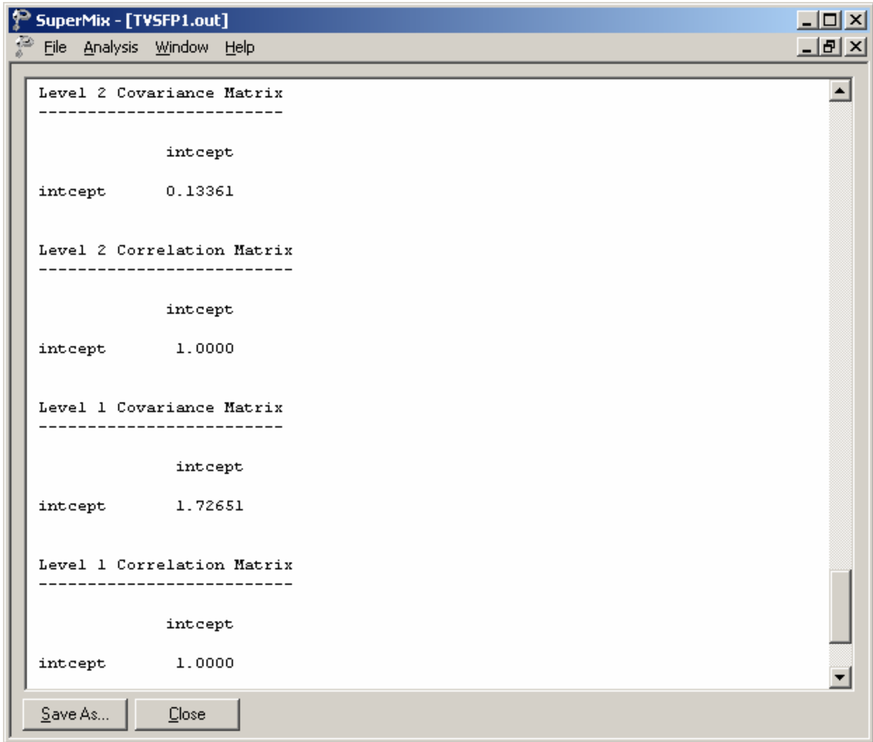


Random effect estimates

The estimates for the random part of the model are reported next. The variation in the average estimated intercept at level 2 is highly significant, which indicates that the classroom difference in intercepts does help to explain the variation in POSTTHKS scores.

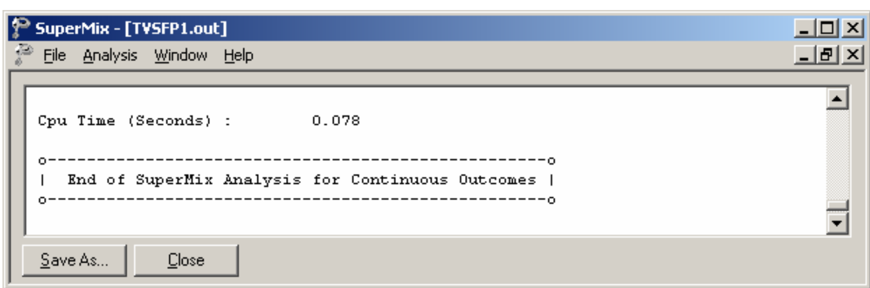


The covariance and correlation matrix of level-2 and level-1 random effects are given at the end of the output file. These values are the same as the estimates of variance/covariance components as shown above.



End of the output

After successfully running a SuperMix model, the following message is shown at the end of the output file to indicate the CPU time and the type of the outcome variable.



Percentage variation explained

An estimate of the percentage of variation in the outcome at classroom level is obtained as

$$\frac{0.13361}{0.13361+1.72651}\times 100\% = 7.18\%$$

indicating that about 7.18% of the total variance lies between the clusters/classrooms and that 92.82% of the variance remains at the student level.

1.5 A 3-level random intercept model using class and school as IDs

The previous model showed that classroom contribute to the explanation of the total variation of the POSTTHKS scores. A similar situation exists in the case of schools. We now construct a three-level model that uses both CLASS and SCHOOL as level-2 and level-3 IDs.

The model

The level-1 and level-2 models are the same as the previous model, as shown below.

Level-1 model $(k = 1, \ldots, n_{ij})$

$$\text{POSTTHKS}_{ijk} = b_{0ij} + \varepsilon_{ijk} ,$$
$$\varepsilon_{ijk} \sim NID(0, \sigma^2)$$

Level-2 model $(j = 1, \ldots, n_i)$

$$b_{0ij} = b_{0i} + b_{1i}\text{CC}_{ij} + b_{2i}\text{TV}_{ij} + b_{3i}(\text{CC}_{ij} \times \text{TV}_{ij}) + v_{0ij}$$
$$v_{0ij} \sim NID(0, \sigma^2_{v(2)})$$

Level-3 model $(i = 1, \ldots, N)$

$$b_{0i} = \beta_0 + v_{0i}$$
$$b_{1i} = \beta_1$$
$$b_{2i} = \beta_2$$
$$b_{3i} = \beta_3$$
$$v_{0i} \sim NID(0, \sigma^2_{v(3)})$$

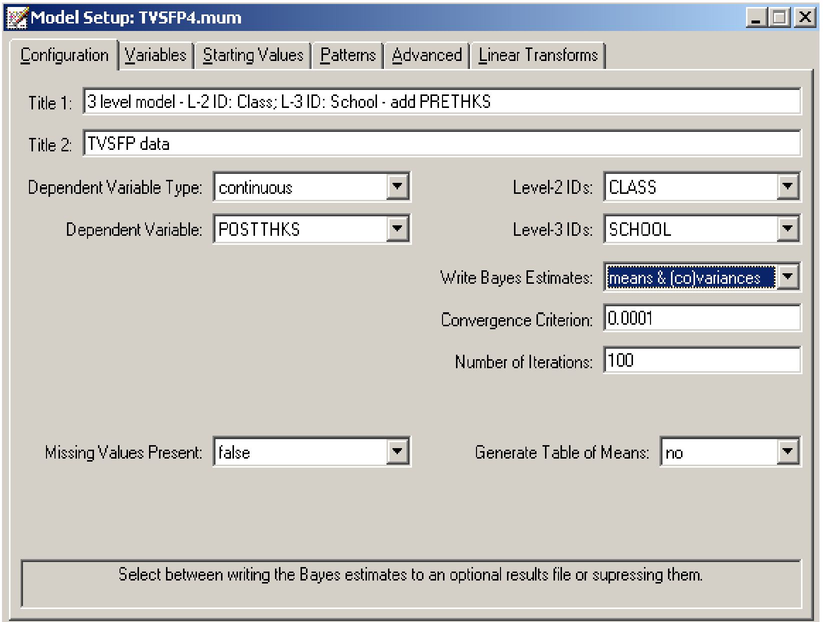
In this mixed model the intercept b_{0ij} is estimated by a level-2 equation. It indicates that classroom *j*'s initial value is not only determined by the population average b_{0i}, but also by the classroom difference v_{0ij}. The level-2-intercept b_{0ij} is estimated by a level-3 equation which takes the school difference v_{0i} into consideration, where *i* denotes the school ID.

The above model can also be written in the following format.

$$\text{POSTTHKS}_{ijk} = \beta_0 + \beta_1\text{CC}_{ij} + \beta_2\text{TV}_{ij} + \beta_3(\text{CC}_{ij} \times \text{TV}_{ij}) + v_{0ij} + v_{0i} + \varepsilon_{ijk} .$$

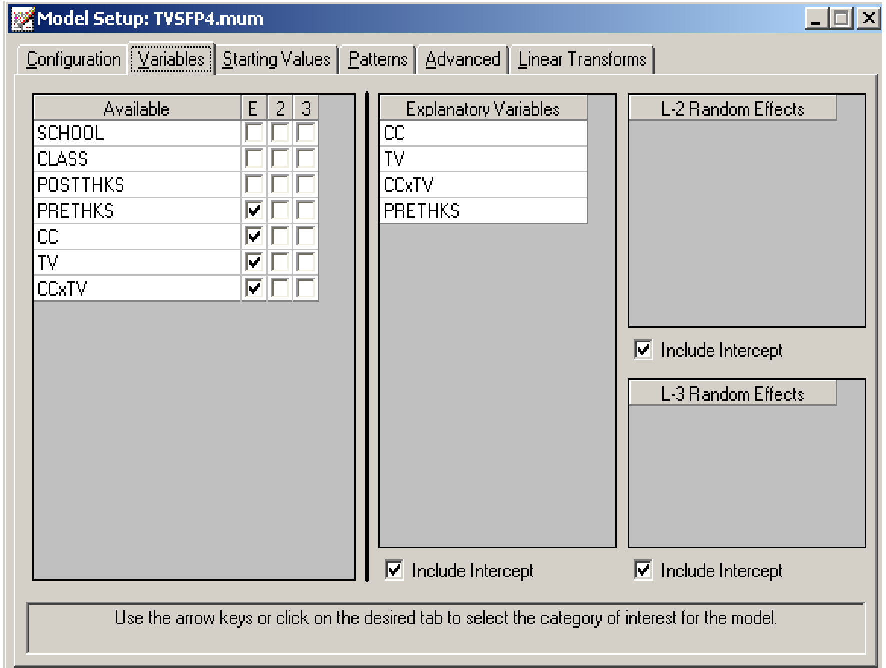
Setting up the analysis

We modify our model setup saved to the syntax file **TVSFP1.mum** by first using the **Open Existing Model Setup** option on the **File** menu of the **TVSFP.ss3** window to retrieve the syntax file. Then click on **File**, **Save as** to save the model setup in a new file, say **TVSFP4.mum**. Next, select SCHOOL as the **Level-3 ID**. We now have both level-2 and level-3 IDs selected.



Change the string in the **Title 1** text box on the **Configuration** tab. Notice that we would like to request Bayes estimates as part of the program output. To do so, select **means & (co)variances** option from the **Write Bayes Estimates** drop down list as shown above.

Click on the **Variables** tab and select PRETHKS as an additional **Explanatory Variable** by checking the corresponding **E** check box.

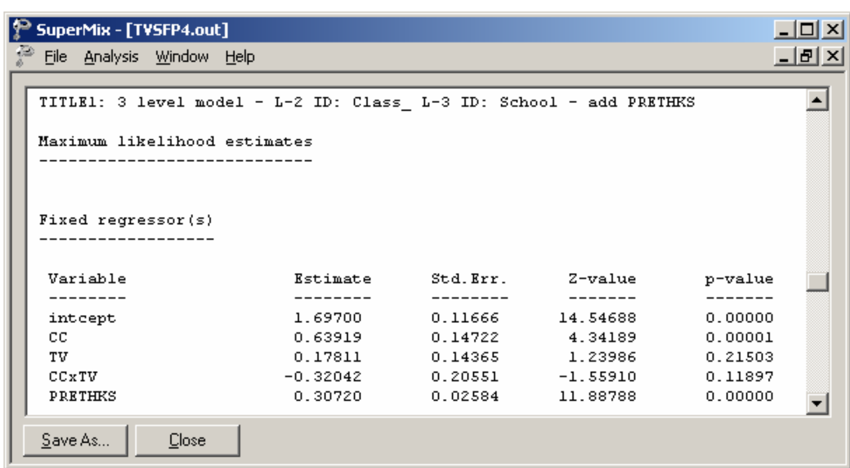


Save the changes to the file **TVSFP4.mum** and select the **Run** option on the **Analysis** menu to produce the output file **TVSFP4.out**.

1.6 Discussion of results

Fixed effects estimates

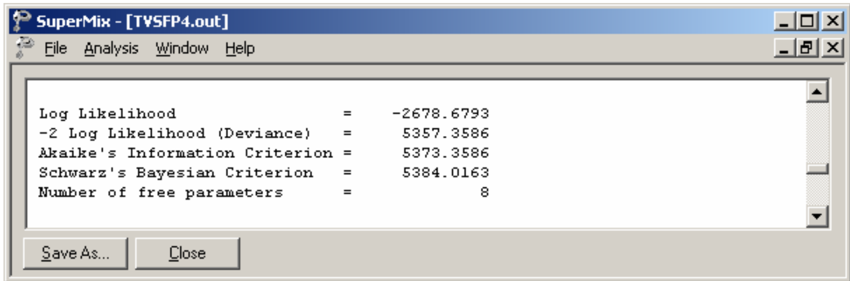
As shown below, the estimated coefficient of PRETHKS is highly significant. The estimate of the intercept coefficient decreased because part of the variation in the intercept can now be explained by PRETHKS.




```
TITLE1: 3 level model - L-2 ID: Class_ L-3 ID: School - add PRETHKS

Maximum likelihood estimates
----------------------------


Fixed regressor(s)
------------------

 Variable                  Estimate     Std.Err.      Z-value       p-value
 --------                  --------     --------      -------       -------
 intcept                    1.69700      0.11666     14.54688       0.00000
 CC                         0.63919      0.14722      4.34189       0.00001
 TV                         0.17811      0.14365      1.23986       0.21503
 CCxTV                     -0.32042      0.20551     -1.55910       0.11897
 PRETHKS                    0.30720      0.02584     11.88788       0.00000
```

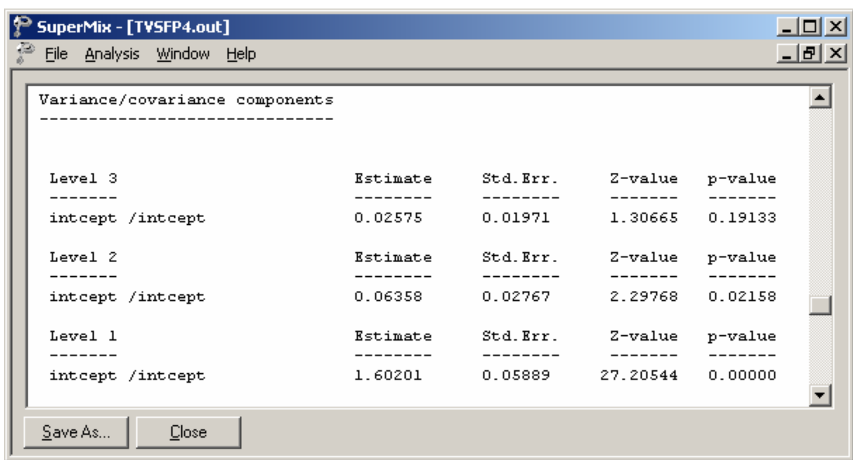
Fit statistics

The fit statistics are given below.



Random effect estimates

The third-level random intercept estimate is not significant at a 5% level of significance.



Estimated outcomes for different groups

For example, if a typical student who only participated in television intervention had a PRETHKS score of 2 (CC = 0; TV = 1; CCxTV = 0), the expected POSTTHKS score is calculated as follows:

$$\begin{aligned}\widehat{\text{POSTTHKS}}_{ijk} &= \hat{\beta}_{00} + \hat{\beta}_{02}\text{TV}_{ij} + \hat{\beta}_{04}\left(\text{PRETHKS}_{ijk}\right) \\ &= 1.697 + 0.17811 + 2 \times 0.3072 \\ &= 2.48951.\end{aligned}$$

ICCs and R-square

ICCs

The so-called ICC (interclass correlation) measures the proportion of variation in the outcome variable between units at the different levels. It is occasionally referred to as the cluster effect, and is defined as the ratio of the between-cluster variance to the total variance. From the output for the random effects, we have

$$\text{Level-1: } \widehat{\text{error var}} = 1.6020$$
$$\text{Level-2: } \widehat{\text{class var}} = 0.0636$$
$$\text{Level-3: } \widehat{\text{school var}} = 0.0258.$$

Based on this information, we can calculate the ICCs as shown below.

Similarity of students within the same school:

$$ICC = \frac{0.0258}{1.6020 + 0.0636 + 0.0258} = 0.0153$$

Similarity of students within the same classrooms (and schools):

$$ICC = \frac{0.0636 + 0.0258}{1.6020 + 0.0636 + 0.0258} = 0.0529$$

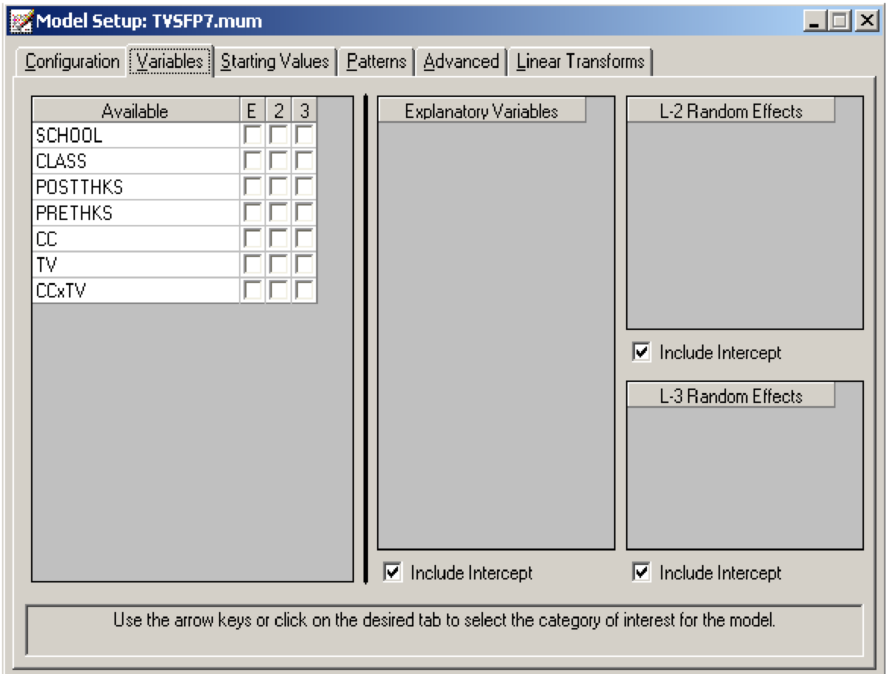
Similarity of classes within the same school:

$$ICC = \frac{0.0258}{0.0636 + 0.0258} = 0.289$$

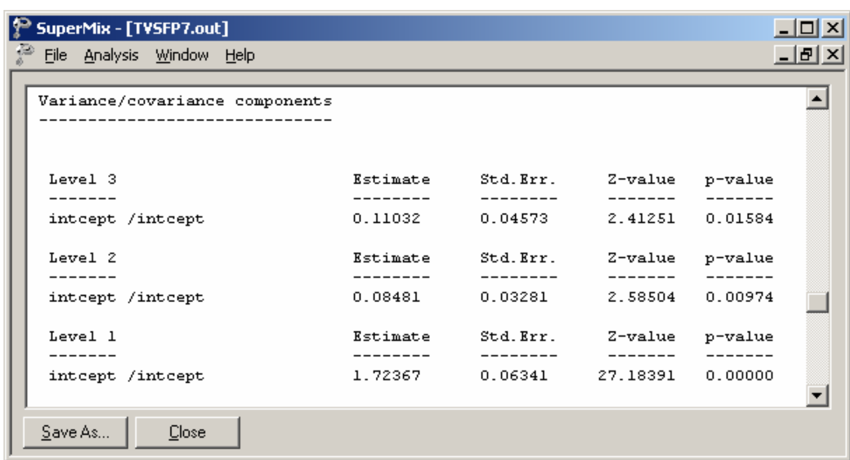
R-square

Another way to evaluate the explanation of variation in the outcome is to compute a statistic analogous to the familiar R^2 used in multiple linear regression. In a multilevel model, however, there is an R^2 for each variance component. Use of these statistics is not without problems, however, because the R^2 may at times have negative values, and in other cases the addition of explanatory variables can lead to an increase rather than a decrease in variance components. The more complex a hierarchical model is, the more likely is the occurrence of anomalies in variance-explained statistics.

To calculate the R^2 s for different levels of the level-3 model, we first need to get the variances for the null model, which is a 3-level model with no covariates. Open **TVSFP4.mum**, click on the **Variables** tab, and uncheck the check boxes of the selected **Explanatory Variables** as shown below.



Save the setup as **TVSFP7.mum** and run the model to get the following output of the variance/covariance component.


SuperMix - [TVSFP7.out]

Variance/covariance components

	Estimate	Std. Err.	Z-value	p-value
Level 3				
intcept /intcept	0.11032	0.04573	2.41251	0.01584
Level 2				
intcept /intcept	0.08481	0.03281	2.58504	0.00974
Level 1				
intcept /intcept	1.72367	0.06341	27.18391	0.00000

The R^2 s are calculated as

$$R_1^2 = 1 - \frac{\hat{\sigma}_p^2}{\hat{\sigma}_0^2} \qquad R_2^2 = 1 - \frac{\hat{\sigma}_{v(2)p}^2}{\hat{\sigma}_{v(2)0}^2} \qquad R_3^2 = 1 - \frac{\hat{\sigma}_{v(3)p}^2}{\hat{\sigma}_{v(3)0}^2}$$

where subscript 0 refers to a model with no covariates (i.e., the null model, **TVSFP7.out**) and subscript p refers to a model with p covariates (i.e., the full model, **TVSFP4.out**). The R^2s for different levels are given in Table 1.1.

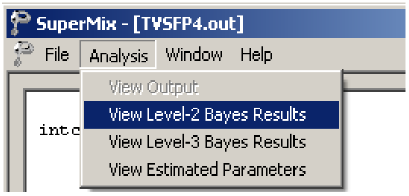
Table 1.1: R^2 values for a set of nested models

level	variance	null	full	R^2
1 (students)	$\hat{\sigma}^2$	1.724	1.602	.071
2 (classrooms)	$\hat{\sigma}^2_{\upsilon(2)}$	.085	.064	.247
3 (schools)	$\hat{\sigma}^2_{\upsilon(3)}$	.110	.026	.764

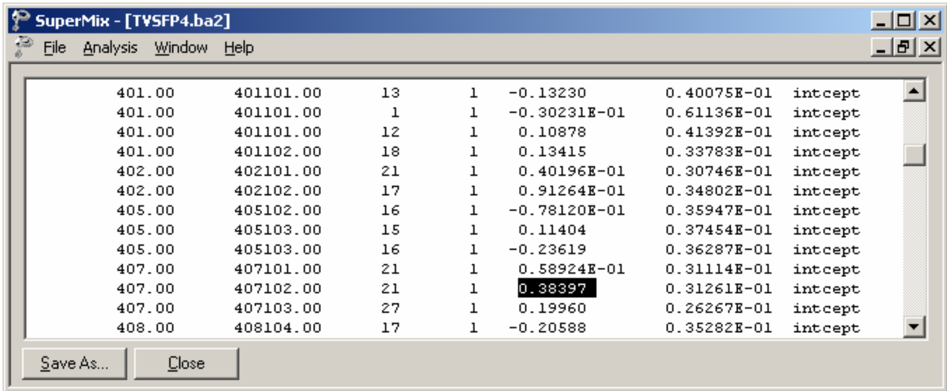
In the current example, only the intercept coefficient is allowed to vary randomly over classrooms and schools, thus making the calculation of the R^2 relatively straightforward. In the case of models with random slopes, the calculation of R^2 statistics becomes more difficult. For an extensive discussion of the rationale and calculation of R^2 statistics, the user is referred to Snijders & Bosker (2000, pp. 99 – 109).

Residuals: Level-2 Bayes results

Returning to the **TVSFP4.mum** output, click on the **Analysis** menu of the output window or the model set up window, and note that **View Level-2 Bayes Results** is now activated. Select the option to open the level-2 Bayes results.



Note that the default extension for the level-2 Bayes estimates is **.ba2**. Part of the file is shown below.



401.00	401101.00	13	1	-0.13230	0.40075E-01	intcept
401.00	401101.00	1	1	-0.30231E-01	0.61136E-01	intcept
401.00	401101.00	12	1	0.10878	0.41392E-01	intcept
401.00	401102.00	18	1	0.13415	0.33783E-01	intcept
402.00	402101.00	21	1	0.40196E-01	0.30746E-01	intcept
402.00	402102.00	17	1	0.91264E-01	0.34802E-01	intcept
405.00	405102.00	16	1	-0.78120E-01	0.35947E-01	intcept
405.00	405103.00	15	1	0.11404	0.37454E-01	intcept
405.00	405103.00	16	1	-0.23619	0.36287E-01	intcept
407.00	407101.00	21	1	0.58924E-01	0.31114E-01	intcept
407.00	407102.00	21	1	0.38397	0.31261E-01	intcept
407.00	407103.00	27	1	0.19960	0.26267E-01	intcept
408.00	408104.00	17	1	-0.20588	0.35282E-01	intcept

The representations of these seven columns are given in order below.

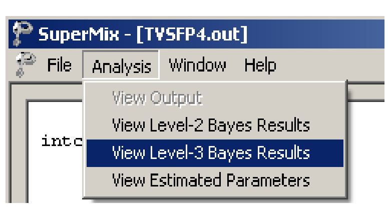
- Column 1: level-3 ID, which is school in our example.
- Column 2: level-2 ID, which refers to classroom.
- Column 3: number of the observations within level-2 ID, number of students within each classroom.
- Column 4: the number of the empirical Bayes coefficients.
- Column 5: the empirical Bayes estimate.
- Column 6: the estimated variance of the Bayes coefficient.
- Column 7: the name of the associated coefficient as used in the model.

Classroom 407102 has the largest Bayes estimate with a value of 0.38397. When considering the class difference, the predicted POSTTHKS score for a student in this specific class who only participated in television intervention with a PRETHKS score of 2 (CC = 0; TV = 1; CCxTV = 0) is calculated as follows.

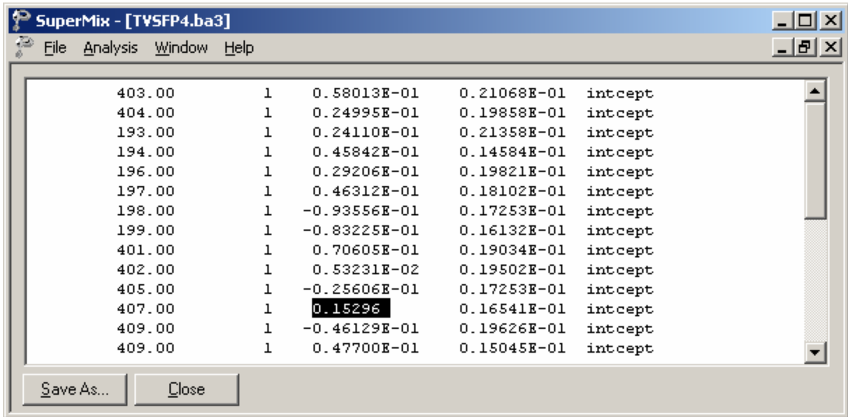
$$\begin{aligned}\widehat{\text{POSTTHKS}}_{ijk} &= \hat{\beta}_0 + \hat{\beta}_2 \text{TV}_{ij} + \hat{\beta}_4 \left(\text{PRETHKS}_{ijk}\right) + \hat{u}_{0i} \\ &= 1.697 + 0.17811 + 2 \times 0.3072 + 0.38397 \\ &= 2.87348.\end{aligned}$$

Level-3 Bayes results

Similarly, the level-3 Bayes results can be viewed by clicking on the **Analysis**, **View Level-3 Bayes Results**.



Part of the **TVSFP.ba3** is shown below.

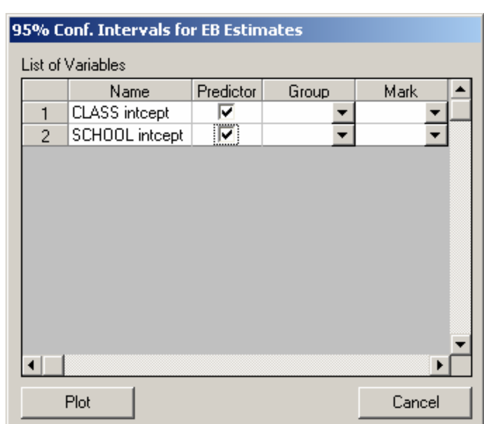


The same classroom (ID = 407102) discussed above is nested in school number 407. Now, considering both the class and school differences, the estimated POSTTHKS for a student from this classroom who only participated in television intervention with a pre-intervention score of 2 (CC = 0; TV = 1; CCxTV = 0) is calculated as follows.

$$\widehat{\text{POSTTHKS}}_{ijk} = \hat{\beta}_0 + \hat{\beta}_2 \text{TV}_{ij} + \hat{\beta}_4 \left(\text{PRETHKS}_{ijk}\right) + \hat{v}_{0ij} + \hat{v}_{0i}$$
$$= 1.697 + 0.17811 + 2 \times 0.3072 + 0.38397 + 0.15296$$
$$= 3.02644.$$

Confidence intervals for random coefficients

The **Confidence Intervals** option on the **File**, **Model-based Graphs** menu provides the option to display confidence intervals for the empirical Bayes estimates of the random effects specified in a given model. This option is now used to examine the confidence intervals of the random intercepts for the schools, which represent the highest level of the hierarchy in the current example.



Select the **Confidence Intervals** option on the **File**, **Model-based Graphs** menu to activate the **95% Conf. Intervals for EB estimates** dialog box. Two graphs of the confidence intervals for the empirical Bayes estimates of the intercepts at the classroom level and school level are obtained by selecting CLASS intcept and SCHOOL intcept in the **Predictor** column before clicking **Plot**.

The graph obtained, as shown below, shows that, in general, the range of the confidence intervals for the level-3 empirical Bayes estimates of the intercepts is $(-0.2; 0.2)$, and the range for level-2 is about $(-0.4; 0.4)$.

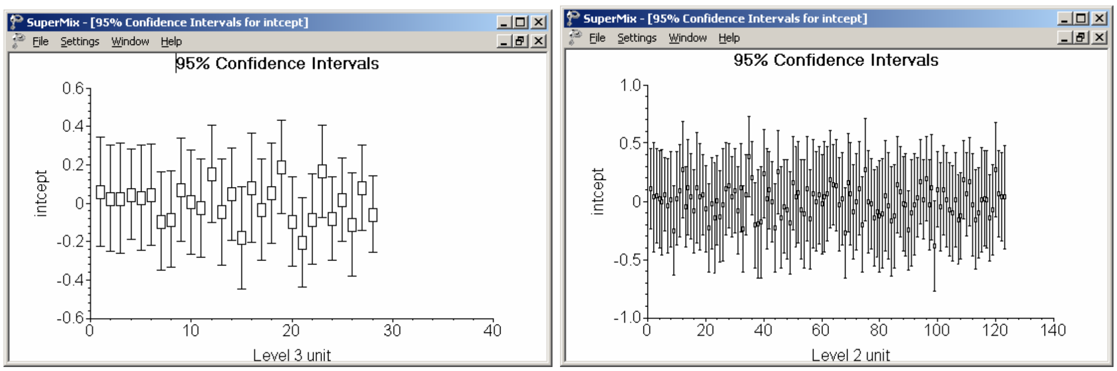


Figure 1.3: 95% confidence intervals for level-2 Bayes estimates

The deviations from the estimated population intercept over schools are also apparent. Each confidence interval is obtained using the formula

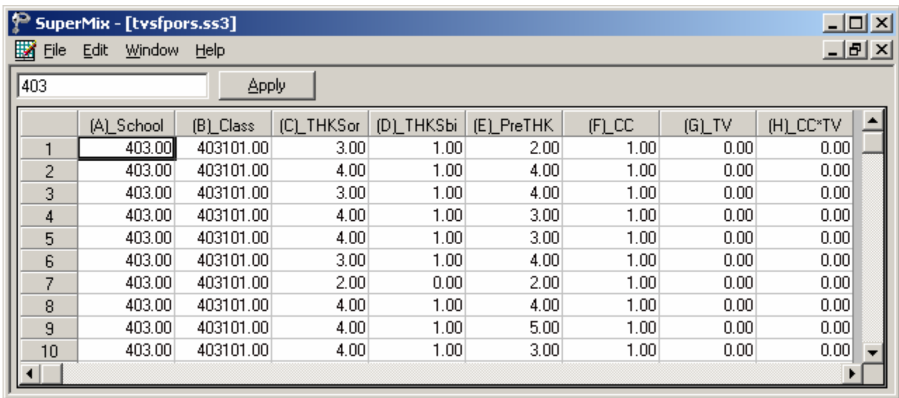
$$Empirical\ Bayes\ residual \pm 1.96\sqrt{var\left(Empirical\ Bayes\ residual\right)}.$$

2 Mixed models for binary outcomes

2.1 The data

The data are from the Television School and Family Smoking Prevention and Cessation Project (TVSFP) study (Flay, *et. al.*, 1988) as described in the previous section of the handout.

Data for the first 10 participants on most of the variables used in this section are shown below in the form of a SuperMix spreadsheet file, named **tvsfpors.ss3**, located in the **Examples\Binary** subfolder.



	(A)_School	(B)_Class	(C)_THKSor	(D)_THKSbi	(E)_PreTHK	(F)_CC	(G)_TV	(H)_CC*TV
1	403.00	403101.00	3.00	1.00	2.00	1.00	0.00	0.00
2	403.00	403101.00	4.00	1.00	4.00	1.00	0.00	0.00
3	403.00	403101.00	3.00	1.00	4.00	1.00	0.00	0.00
4	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00
5	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00
6	403.00	403101.00	3.00	1.00	4.00	1.00	0.00	0.00
7	403.00	403101.00	2.00	0.00	2.00	1.00	0.00	0.00
8	403.00	403101.00	4.00	1.00	4.00	1.00	0.00	0.00
9	403.00	403101.00	4.00	1.00	5.00	1.00	0.00	0.00
10	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00

The variables of interest are:

- School indicates the school a student is from (28 schools in total).
- Class identifies the classroom (135 classrooms in total).
- THKSord represents the tobacco and health knowledge scale, with 4 categories ranging between 1 and 4. The frequency distribution of the post-intervention THKS scores indicated that approximately half the students had scores of 2 or less, and half of 3 or greater. In terms of quartiles, four ordinal classifications were suggested corresponding to 0 – 1, 2, 3, and 4 – 7 correct responses.
- THKSbin is a recoded version of the ordinal variable THKSord, but in binary form: a value of "0" indicates an original scale score of 1 or 2, while a value of "1" indicates an scale score of 3 or 4. This variable will serve as our outcome variable in the current chapter.
- PreTHKS indicates a student's score prior to intervention, *i.e.*, the number correct of 7 items.
- CC is a binary variable indicating whether a social-resistance classroom curriculum was introduced, with 0 indicating "no" and 1 "yes."
- TV is an indicator variable for the use of media (television) intervention, with a "1" indicating the use of media intervention, and "0" the absence thereof.
- CC*TV is the product of the variables TV and CC, and represents the CC by TV interaction.

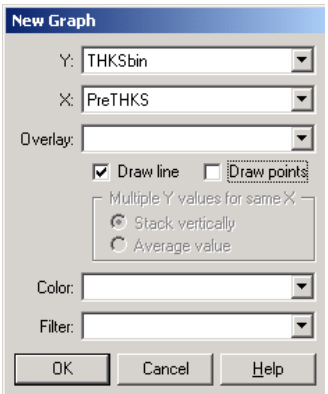
In this section we consider models for binary outcomes, using quadrature as method of estimation.

2.2 Graphical displays

Exploratory graphs

The pre-intervention scores of the students may be useful as a covariate in the analysis. To get an idea of the relationship between the scale score PreTHKS and the post-intervention score THKSbin, an exploratory graph is created. Select the **Data-based Graphs**, **Exploratory** option from the **File** menu.

The **New Graph** dialog box is activated. Select the binary outcome variable THKSbin as the **Y** variable and the pre-intervention score PreTHKS as the **X** variable. Uncheck the **Draw points** check box, which is checked by default, to obtain the settings as shown.



Click **OK** to obtain Figure 2.1. We note that the relationship is reasonably linear, and that higher post-intervention scores are more often observed for students with high pre-intervention scores, which is what one intuitively would expect.

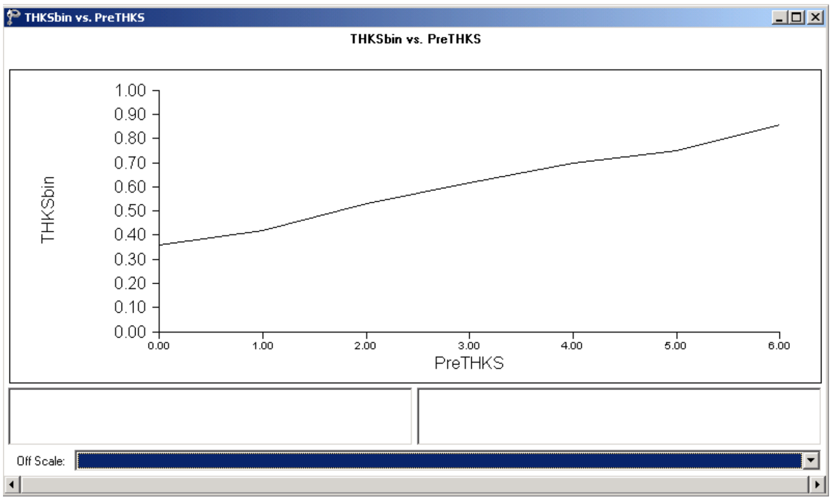
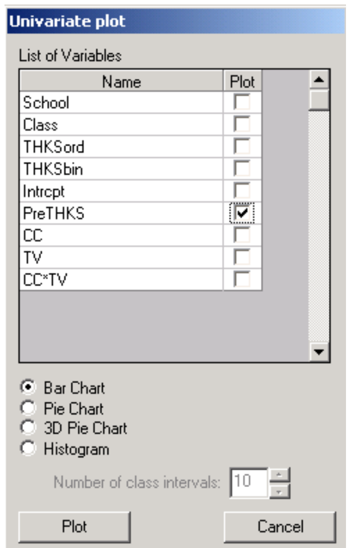


Figure 2.1: Exploratory graph of THKSbin vs. PreTHKS

Univariate graphs

We now take a closer look at the distribution of the pre-intervention scores by utilizing the **Data-based Graphs**, **Univariate** option on the **File** menu. By default, a bar chart will be produced. Select the variable PreTHKS in the **Plot** column, and click **Plot**.



By clicking anywhere in the bars, the **Bar Graph Parameters** dialog box is activated. Click the **Data** button and then **OK** to display the data used to construct the bar chart.

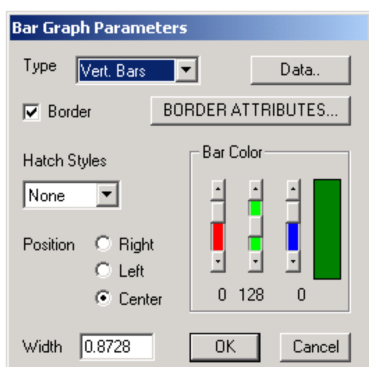


Figure 2.2 below shows both the graphing window with bar chart and the data in spreadsheet format. Note that only 55 of the 1600 observations showed a score of 5 or higher, and that no student obtained a pre-intervention score of 7 out of 7.

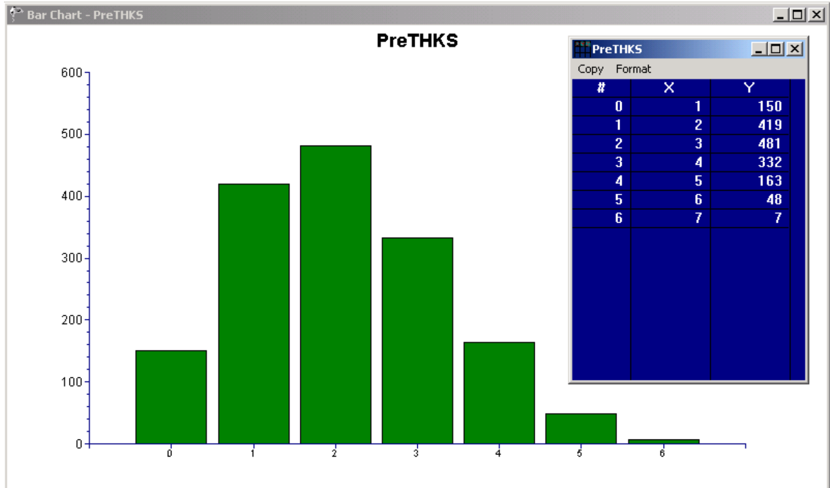


Figure 2.2: Bar chart of PreTHKS values

2.3 A 2-level random intercept logistic model with 4 predictors

The model

The outcome variable THKSbin used here is binary. It assumes a value of "0" when the original scale score was either 1 or 2, and a value of "1" for an original scale score of 3 or 4. The predicted value of the outcome can be viewed as the predicted probability that THKSbin is 1. Predicted values outside the interval [0,1] would not be meaningful and a model constraining predicted values to lie in this interval would be appropriate, in contrast with the model for a continuous outcome. In addition, the assumption of normality at level 1 is not realistic, as the level-1 random effect can only assume one of two values: 0 or 1. This random effect can thus not have homogeneous variance.

In order to insure that the predicted values lie within the (0,1) interval, a transformation of the level-1 predicted probability can be used. For the binary case considered here, the following link function is used:

$$\text{Prob}(\text{THKSbin}_{ij} = 1 \mid \boldsymbol{\beta}, \mathbf{v}) = \frac{e^{\eta_{ij}}}{1 + e^{\eta_{ij}}}$$

where η_{ij} represents the log of the odds of success.

For the current model, we want to explore the relationship between the post-intervention scores and the type of intervention applied. This relationship can be expressed as

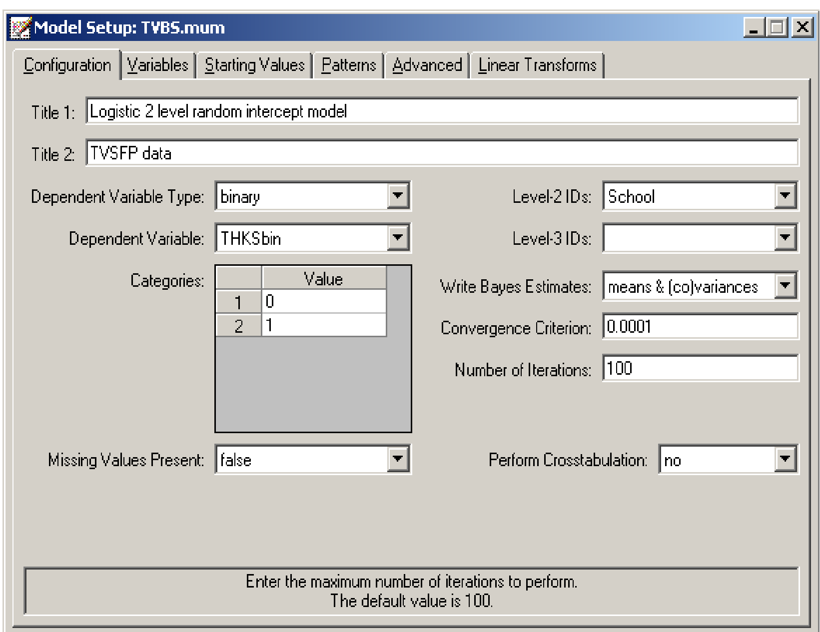
$$\eta_{ij} = \beta_0 + \beta_1 \times \text{CC}_i + \beta_2 \times \text{TV}_i + \beta_3 \times \text{CC}_i * \text{TV}_{ij} + \beta_4 \times \text{PreTHKS}_{ij} + v_{0i}.$$

Setting up the analysis

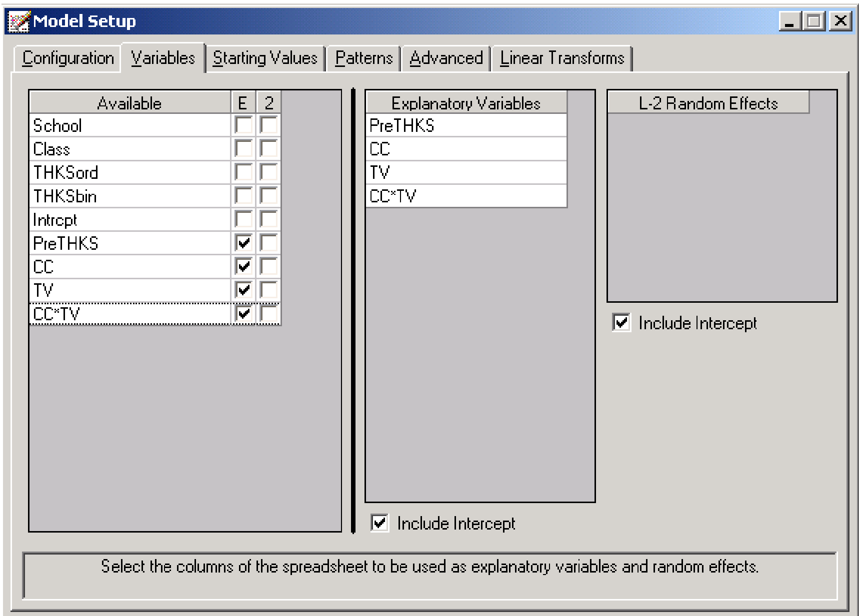
Using the data in **tvsfpors.ss3**, we consider the situation where students are nested within schools, and fit a two-level model with the binary variable THKSbin as outcome. We wish to examine the relationships between the outcome and the two intervention methods employed, simultaneously taking students' pre-intervention scores into account. To do so, we use the model described above with schools as the level-2 units.

From the main menu bar, select the **File, New Model Setup** option. The **Model Setup** dialog box that appears has six tabs: **Configuration**, **Variables**, **Starting Values**, **Patterns**, **Advanced**, and **Linear Transforms**. In this example, only three of the tabs are used.

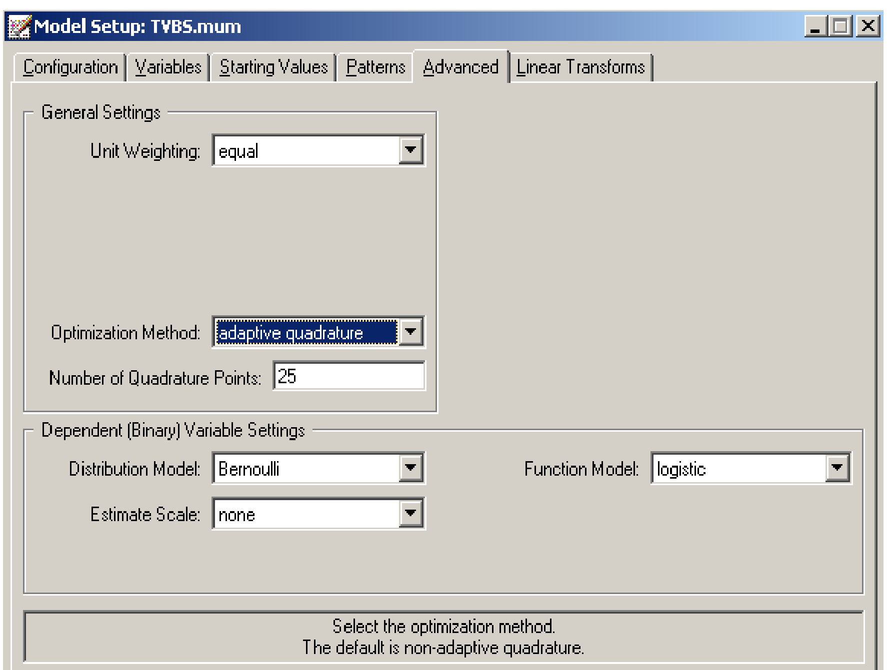
As a first step, the binary outcome variable THKSbin is selected from the **Dependent Variable** drop-down list box. The type of outcome is specified as binary using the drop-down list box in the **Dependent Variable Type** field. Once this selection is made, the **Categories** field is displayed. The school identification variable is used to define the hierarchical structure of the data, and is selected as the **Level-2 ID** from the **Level-2 IDs** drop-down list box. A title for the analysis (optional) is entered in the **Title** fields. A convergence criterion of 0.0001 is requested. By default, the maximum number of iterations performed is set to 100. Empirical Bayes residuals, written to additional output files, are requested by setting the **Write Bayes Estimates** option to **means and (co)variances.** Default settings for all other options associated with this tab are used. Proceed to the **Variables** tab by clicking on this tab.



The **Variables** tab is used to specify the fixed and random effects to be included in the model. Start by selecting the explanatory (fixed) variables using the first column of boxes in the **Available** group field. The first variable selected is PreTHKS, followed by CC, TV, and the interaction term CC*TV. After selecting these explanatory variables, the random effect(s) at level 2 must be selected. In this case, we wish to allow only the intercept to vary randomly over the schools. By default, the intercept is assumed to vary randomly over higher levels of the hierarchy as indicated by the checked box for the **Include Intercept** option in the **L-2 Random Effects** group field. A common fixed intercept coefficient is also included, as shown in the **Explanatory Variables** group field.



We opt to increase the number of quadrature points to be used during estimation. To do so, click the **Advanced** tab.



First select **adaptive quadrature** from the **Optimization Method** drop-down list box, then change the **Number of Quadrature Points** field to 25. The default distribution for a binary outcome variable is **Bernoulli** and the default link function is **probit**. Change **probit** to **logistic** by using the drop-down list box in the **Function Model** field.

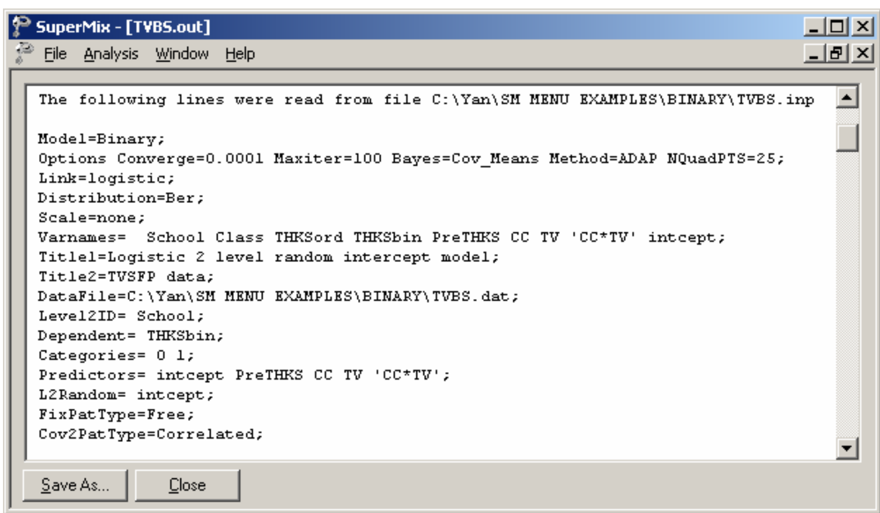
Before running the analysis, the model specifications have to be saved. Select the **File**, **Save** option, and provide a name for the model specification file, for example **TVBS.mum**. Run the analysis by selecting the **Run** option from the **Analysis** menu.

2.4 Discussion of results

Portions of the output file **tvbs.out** are shown below.

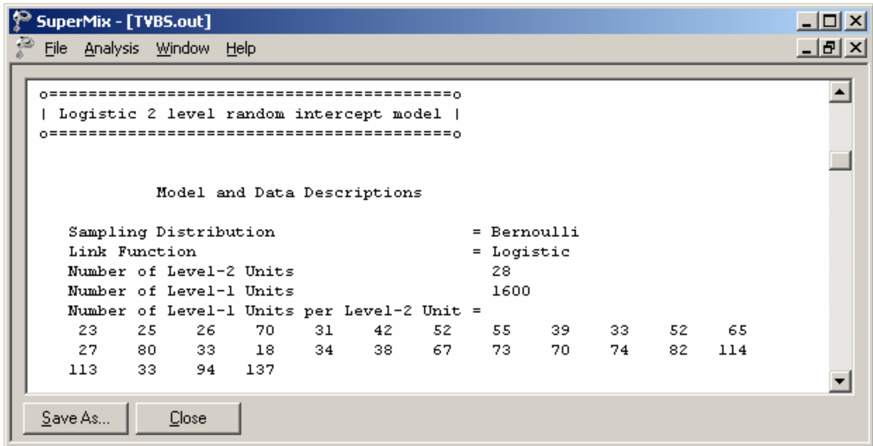
Syntax

At the top of the file, the syntax saved to the **TVBS.mum** file is shown. The first part states the selection of iteration control options, requests for Bayes residuals, and the specifications necessary to define the model fitted as an binary model with a logistic link function. The second part of the syntax provides information on the structure of the data, the name and structure of the outcome variable, and the predictors included in the model. Text to the left of the equal sign in each line denote keywords recognized by the program; text to the right are either keywords (for example, in the case of Cov2PatType = Correlated) or variable names as given in the **ss3** file (for example, Level2ID = School).



Model and data description

The next section of the output file contains a description of the hierarchical structure and model specifications.

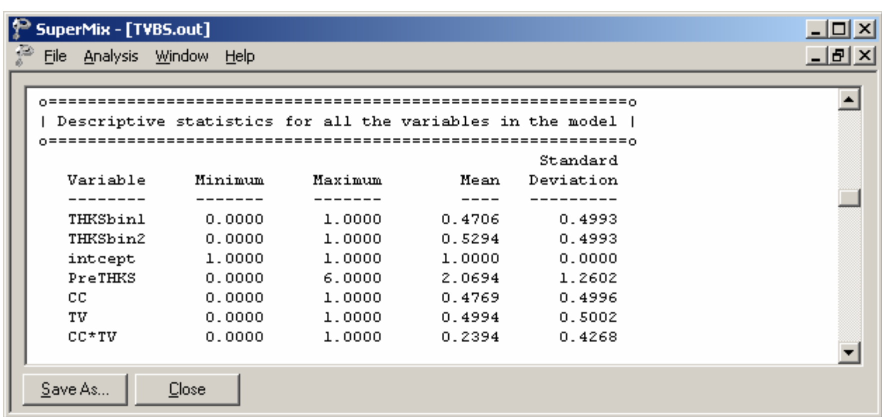



```
o========================================o
| Logistic 2 level random intercept model |
o========================================o


            Model and Data Descriptions

   Sampling Distribution                     = Bernoulli
   Link Function                             = Logistic
   Number of Level-2 Units                     28
   Number of Level-1 Units                     1600
   Number of Level-1 Units per Level-2 Unit =
    23     25     26     70     31     42     52     55     39     33     52     65
    27     80     33     18     34     38     67     73     70     74     82    114
   113     33     94    137
```

The use of a logistic response function (logit link function) with the assumption of a Bernoulli distribution is indicated. This is followed by a summary of the number of students nested within each school. The number of students per school (level-2 unit) ranges between 23 and 137.

Descriptive statistics

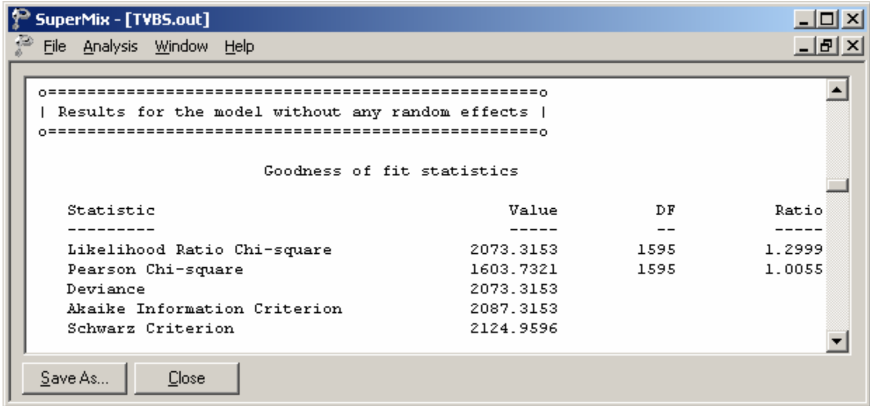
The data summary is followed by descriptive statistics for all variables included in the model. We note that 47% of the students had a value of 0 on the binary knowledge score outcome variable THKSbin, and 53% a value of 1.




```
o=========================================================o
| Descriptive statistics for all the variables in the model |
o=========================================================o
                                                         Standard
   Variable        Minimum      Maximum         Mean    Deviation
   --------        -------      -------         ----    ---------
   THKSbin1         0.0000       1.0000       0.4706       0.4993
   THKSbin2         0.0000       1.0000       0.5294       0.4993
   intcept          1.0000       1.0000       1.0000       0.0000
   PreTHKS          0.0000       6.0000       2.0694       1.2602
   CC               0.0000       1.0000       0.4769       0.4996
   TV               0.0000       1.0000       0.4994       0.5002
   CC*TV            0.0000       1.0000       0.2394       0.4268
```

Results for the model without any random effects

Descriptive statistics are followed by parameter estimates obtained under the assumption that all random effects are zero. The parameter values for the predictors CC, TV, CC*TV and PreTHKS are given in the first column, followed by the standard errors and *z*- and *p*-values.


SuperMix - [TVBS.out]

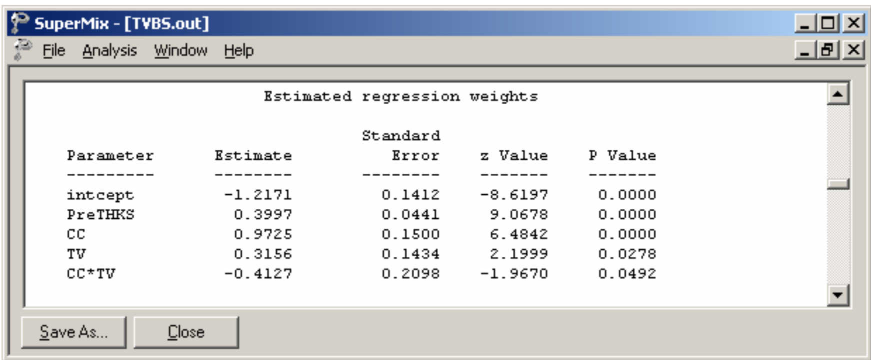
File Analysis Window Help

```
o================================================o
| Results for the model without any random effects |
o================================================o
```

Goodness of fit statistics

Statistic	Value	DF	Ratio
Likelihood Ratio Chi-square	2073.3153	1595	1.2999
Pearson Chi-square	1603.7321	1595	1.0055
Deviance	2073.3153		
Akaike Information Criterion	2087.3153		
Schwarz Criterion	2124.9596		

Save As... Close


SuperMix - [TVBS.out]

File Analysis Window Help

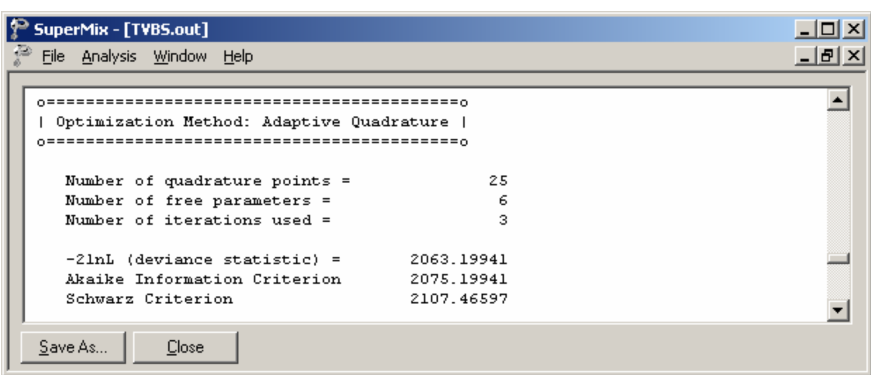
Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
intcept	-1.2171	0.1412	-8.6197	0.0000
PreTHKS	0.3997	0.0441	9.0678	0.0000
CC	0.9725	0.1500	6.4842	0.0000
TV	0.3156	0.1434	2.1999	0.0278
CC*TV	-0.4127	0.2098	-1.9670	0.0492

Save As... Close

Results for the model fitted with adaptive quadrature

The output describing the estimated parameters after convergence is shown next. Three iterations were required to obtain convergence. The number of quad points used per dimension was 25. The likelihood function value at convergence as well as the deviance are also given, and may be used to compare a set of nested models.


SuperMix - [TVBS.out]

File Analysis Window Help

```
o========================================o
| Optimization Method: Adaptive Quadrature |
o========================================o

   Number of quadrature points =            25
   Number of free parameters =               6
   Number of iterations used =               3

   -2lnL (deviance statistic) =      2063.19941
   Akaike Information Criterion      2075.19941
   Schwarz Criterion                 2107.46597
```

Save As... Close

The estimates are shown in the column with heading Estimate, and correspond to the coefficients $\beta_0, \beta_1, \ldots, \beta_4$ in the model specification. Significant effects of PreTHKS and CC are observed. The variation in the intercept over the schools is estimated as 0.1065, and from the associated *p*-value we conclude that there is significant variation, at a 10% level of significance, in the intercept between the schools included in this analysis.

TVBS.out

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
intcept	-1.2281	0.1949	-6.3012	0.0000
PreTHKS	0.3871	0.0451	8.5846	0.0000
CC	1.0893	0.2454	4.4390	0.0000
TV	0.3741	0.2350	1.5923	0.1113
CC*TV	-0.5578	0.3403	-1.6392	0.1012

Estimated level 2 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intcept/intcept	0.1065	0.0578	1.8419	0.0655

Save As... Close

In the case of the fixed effects, a 2-tailed *p*-value is used, as the alternative hypothesis considered here is of the form $H_1 : \beta \neq 0$. As variances are constrained to be elements of the interval $[0, +\infty)$, the *p*-values used for these effects are 1-tailed.

If the model is true, it is assumed that the level-1 error variance is equal to $\pi^2/3$ = 3.29895 for the logistic link function (see, *e.g.*, Hedeker & Gibbons (2006), p. 157), where π represents the constant 3.141592654.

Thus the estimated ratio between level-2 variation and the total variation is calculated as

$$ICC = \frac{0.1065}{0.1065 + \pi^2/3} = 0.031$$

This indicates that almost all variation is attributable to students, rather than to the schools.

Interpreting the adaptive quadrature results

The expected log-odds of having a high post-intervention knowledge score (THKSbin score of 1) for a student with a zero value on all the predictors (that is, no social-resistance curriculum, no media intervention, and a pre-intervention knowledge score of 0) is represented by the estimated intercept of -1.2281. When a social-resistance curriculum was in place (CC = 1), or a mass-media intervention was performed (TV = 1), the log-odds of a typical student is expected to increase, as indicated by the positive estimated coefficients for CC and TV. Similarly, a higher score on the pre-intervention knowledge test is associated with higher log-odds of a higher post-intervention knowledge score. It can be concluded from the results that the implementation of a classroom curriculum was more likely to lead to a higher post-intervention knowledge score than was the case when mass-media intervention was used. In contrast, the log-odds of a high post-intervention knowledge score was expected to be lower for a typical student from a school where both social resistance classroom curriculum and mass-media intervention defined the study condition for that school, as the estimated coefficient for the interaction term CC*TV was negative.

Estimated outcomes for different groups: unit-specific results

To evaluate the expected effect of CC, TV, CC*TV, and PreTHKS on the predicted probability that the post-intervention score is equal to 1, we use the following expression for the predicted log odds of success

$$\hat{\eta}_{ij} = \hat{\beta}_0 + \hat{\beta}_1 \times \mathrm{CC}_i + \hat{\beta}_2 \times \mathrm{TV}_i + \hat{\beta}_3 \times \mathrm{CC}_i \times \mathrm{TV}_i + \hat{\beta}_4 \times \mathrm{PreTHKS}_{ij}$$

for the four groups defined by the categories of CC and TV. Note the similarity of this equation with that given for η_{ij} earlier: random coefficients are not included, as their expected value is 0.

For a typical student with a PreTHKS score of 0 from any school where no media television intervention and no social-resistance classroom curriculum was implemented, CC = TV = 0, and thus

$$\hat{\eta}_{ij} = \hat{\beta}_0$$

In the case of a typical student with a PreTHKS score of 0 from any school where only media television intervention was implemented (TV = 1),

$$\hat{\eta}_{ij} = \hat{\beta}_0 + \hat{\beta}_2 \times \mathrm{TV}_i.$$

The equations for similar students from a school with only a social–resistance classroom curriculum implemented (CC = 1, TV = 0), and from a school with both interventions implemented (TV = 1, CC = 1) are

$$\hat{\eta}_{ij} = \hat{\beta}_0 + \hat{\beta}_1 \times \text{CC}_i + \hat{\beta}_4 \times \text{PreTHKS}_{ij}$$

and

$$\hat{\eta}_{ij} = \hat{\beta}_0 + \hat{\beta}_1 \times \text{CC}_i + \hat{\beta}_2 \times \text{TV}_i + \hat{\beta}_3 \times \text{CC}_i \times \text{TV}_i + \hat{\beta}_4 \times \text{PreTHKS}_{ij}$$

respectively.

For a student with an average PreTHKS score (2.152, see exploratory analysis) from any school with similar values of CC and TV we find that

$$\begin{aligned}\hat{\eta}_{ij} &= \hat{\beta}_0 + \hat{\beta}_4 * \text{PreTHKS}_{ij} \\ &= \hat{\beta}_0 + \hat{\beta}_4 * 2.152.\end{aligned}$$

Using the $\hat{\beta}_0$ and $\hat{\beta}_4$ estimates of -1.2281 and 0.3871 respectively as obtained for the current analysis, we can calculate the estimated probability of a THKSbin score of 1 for typical students with PreTHKS scores of 2.152 and 0 respectively as

$$\begin{aligned}\text{Prob}(\text{THKSbin}_{ij} = 1 \mid \text{CC} = \text{TV} = 0; \text{PreTHKS} = 2.152) &= \frac{e^{-1.2281+0.3871(2.152)}}{1+e^{-1.2281+0.3871(2.152)}} \\ &= \frac{e^{-0.39506}}{1+e^{0.39506}} \\ &= 0.40250\end{aligned}$$

and

$$\begin{aligned}\text{Prob}(\text{THKSbin}_{ij} = 1 \mid \text{CC} = \text{TV} = \text{PreTHKS} = 0) &= \frac{e^{-1.2281}}{1+e^{-1.2281}} \\ &= 0.22651.\end{aligned}$$

A student with an average observed score of PreTHKS is almost twice as likely to have a THKSbin score of 1 as a student with the lowest observed score on the same variable. Note that we opted to use the mean pre-intervention score for this specific subgroup.

On the other end of the scale in terms of intervention, we have schools where both a social-resistance classroom curriculum and a mass-media intervention were implemented (CC = TV = 1). For two typical students from these schools, an observed PreTHKS score of 0 or the mean score of 1.979 will imply a predicted probability of a THKSbin score of 1 of 0.4201 for the first and 0.6091 for the second. Again, the higher the pre-intervention score, the higher the predicted probability of a high post-intervention score.

In Table 2.1, the estimated probabilities of high post-intervention scores on the tobacco and health questionnaire are given for typical students with high or low pre-intervention scores for each of the subpopulations formed by mass-media intervention and implementation of social-resistance classroom curriculum.

Table 2.1: Estimated unit-specific probability of a high post-intervention knowledge score

Group	prescore	prob.	prescore	prob.
CC = 0, TV = 0	0	22.65%	2.152	40.25%
CC = 1, TV = 0	0	46.54%	2.05	65.81%
CC = 0, TV = 1	0	29.86%	2.87	48.85%
CC = 1, TV = 1	0	42.01%	1.979	60.91%

Students with a high pre-intervention score were predicted to have a high post-intervention score too, regardless of the study conditions. Similarly, students with a low pre-intervention score were generally likely to have a low post-intervention score too. If only curriculum intervention (CC = 1) was used, scores for students were likely to be higher regardless of their pre-intervention scores. On both ends of the pre-intervention knowledge score scale, in groups where mass-media intervention was used (TV = 1), scores were predicted to be higher than where media intervention was not used, except when both mass-media and curriculum intervention were used. For these groups, with CC = TV = 1, the estimated probabilities of a high post-intervention score were actually lower than for the group where only a classroom curriculum was used (42.01% vs. 46.54%, and 60.91% vs. 65.81%).

We conclude that for most students, the implementation of a social-resistance classroom curriculum is more likely to be effective in increasing their knowledge (predicted probabilities of a high score being 46.54% and 65.81% respectively) than mass-media intervention (predicted probabilities of a high score being 29.86% and 48.85% respectively). The control group, where neither method was implemented, had the lowest predicted knowledge scores (22.65% and 40.25% respectively). While the implementation of both procedures was associated with higher probabilities than either the control group or the group where only mass-media intervention was used, its predicted gain was disappointing when compared to the use of only social-resistance curriculum implementation. Generally speaking, the implementation of a curriculum only seems to be most effective in increasing the predicted knowledge of students on the tobacco and health questionnaire.

Estimated outcomes for different groups: population-average results

In the introduction to this section, we defined the latent response variable model as

$$y_{ij} = \mathbf{x}_{ij}^{'}\boldsymbol{\beta} + \mathbf{z}_{ij}^{'}\mathbf{v}_i + \varepsilon_{ij}, \quad j = 1, 2, ..., n_i$$

where $\mathbf{z}_{ij}^{'}$ denotes a design vector for the random effects contained in the vector $\mathbf{v}_i$, and $\mathbf{x}_{ij}^{'}$ the design vector for the predictors in the fixed part of the model with corresponding vector $\boldsymbol{\beta}$ of regression parameters. The covariance matrix of $\mathbf{v}_i$ is denoted by $\Phi_{(v)}$ and the variance of ε_{ij} by σ_ε^2.

For a probit link function $\sigma_\varepsilon^2 = 1$, and for a logistic link function it is assumed to be $\sigma_\varepsilon^2 = \pi^2/3$. Under the assumption that $\mathbf{v}_i$ and ε_{ij} are independently distributed, it follows that

$$\sigma_{y_{ij}}^2 = \mathbf{z}_{ij}^{'}\Phi_{v_i}\mathbf{z}_{ij} + \sigma_\varepsilon^2.$$

The design effect d_{ij} is defined in terms of σ_ε^2 and $\sigma_{y_{ij}}^2$:

$$d_{ij} = \frac{\sigma_{y_{ij}}^2}{\sigma_\varepsilon^2}.$$

This design effect may be used to obtain the estimated population-average probabilities in a similar fashion as the unit-specific probabilities, but with replacing $\hat{\eta}_{ij}$ with $\hat{\eta}_{ij}^* = \hat{\eta}_{ij} / \sqrt{d_{ij}}$ (Hedeker & Gibbons, 2006).

We can compare these estimated population-average probabilities with the observed data for the four groups formed by the categories of TV and CC as shown in Table 2.2. To illustrate, we calculate the estimated population-average probabilities for a few of the subgroups.

From the output, we have $\operatorname{var}(v_{i0}) = 0.1065$, where v_{i0} denotes the random intercept coefficient. In this case, $\mathbf{z}_{ik}^{'} = \mathbf{1}$ and hence, with $\sigma_\varepsilon^2 = \pi^2/3$ for the logistic link,

$$\sigma_{y_{ij}}^2 = 1 \times 0.1065 \times 1 + 3.2899 = 3.3964.$$

Therefore

$$d_{ij} = \frac{3.3964}{3.2899} = 1.0324.$$

To obtain the population-average probability estimates, we now replace the $\hat{\eta}_{ij}$ values calculated for the unit-specific case with $\hat{\eta}_{ij}^* = \hat{\eta}_{ij} / \sqrt{d_{ij}}$.

For the subgroup where TV = CC = 0 and the mean PreTHKS value is equal to 2.152, for example, we find that

$$\begin{aligned}\hat{\eta}_{ij} &= -1.2281 + 0.3871(2.152) \\ &= -0.39506\end{aligned}$$

so that

$$\begin{aligned}\hat{\eta}_{ij}^* &= -0.39506 / \sqrt{1.0324} \\ &= -0.38881\end{aligned}$$

and

$$\begin{aligned}P(\text{THKSbin}_{ij} = 1 \,|\, \text{CC} = \text{TV} = 0, \text{PreTHKS} = 2.152) &= \frac{e^{\eta_{ij}^*}}{1 + e^{\eta_{ij}^*}} \\ &= \frac{0.67786}{1.67786} = 40.40\%.\end{aligned}$$

Similarly, for the group where TV = CC = 0 and PreTHKS = 0, we find that

$$\hat{\eta}_{ij} = -1.2281$$

$$\begin{aligned}\hat{\eta}_{ij}^* &= -1.2281 / 1.01606 \\ &= -1.2087.\end{aligned}$$

Table 2.2: Estimated population-average probabilities

Group	prescore	prob.	prescore	prob.
CC = 0, TV = 0	0	22.99%	2.152	40.40%
CC = 1, TV = 0	0	46.59%	2.05	65.57%
CC = 0, TV = 1	0	30.14%	2.87	48.87%
CC = 1, TV = 1	0	42.13%	1.979	60.74%

A comparison of these probabilities with the observed ratios given in Table 2.3 for the control group at the end of the study indicates that the population-average results are slightly closer to the observed ratios than is the case for the unit-specific results. Recall that $\sqrt{d_{ij}} = 1.0161$. The extent of differences between unit-specific and population-average results is highly dependent on the "scaling" induced by dividing the $\hat{\eta}_{ij}$s by $\sqrt{d_{ij}}$.

Table 2.3: Observed and predicted proportions of high post–intervention scores

Group	Proportion observed	Unit-specific predicted prob.	Population-average predicted prob.
CC = 0, TV = 0	41.57%	40.25%	40.40%
CC = 1, TV = 0	63.16%	65.81%	65.57%
CC = 0, TV = 1	48.32%	48.85%	48.87%
CC = 1, TV = 1	60.31%	60.91%	60.74%

Interpreting the contents of the level-2 residual file

In addition to the standard output file, the **Write Bayes Estimates** field on the **Configuration** tab of the **Model Setup** dialog was used to request Bayes estimates for the individual random terms. These estimates are written to the file **TVBS.ba2**. The first few lines of this file are shown below.

Four pieces of information per school are given:

- all 1s for the level-2 model,
- the school's ID,
- the value of random intercept,
- the empirical Bayes estimate,
- the associated posterior variance for the school estimate, and
- the name of the associated random coefficient.

1.00	403.00	1	0.2890865	0.0746508	intcept
1.00	404.00	1	0.1160136	0.0660369	intcept
1.00	193.00	1	0.0893923	0.0651102	intcept
1.00	194.00	1	0.0983555	0.0392957	intcept
1.00	196.00	1	0.2469961	0.0661559	intcept
1.00	197.00	1	0.0137773	0.0528812	intcept
1.00	198.00	1	-0.2769779	0.0466032	intcept
1.00	199.00	1	-0.1872501	0.0453496	intcept
1.00	401.00	1	0.3178720	0.0580872	intcept
1.00	402.00	1	-0.0834452	0.0591895	intcept
1.00	405.00	1	-0.0580730	0.0463621	intcept
1.00	407.00	1	0.4107440	0.0408293	intcept
1.00	408.00	1	-0.1502285	0.0664276	intcept
1.00	409.00	1	0.2301998	0.0354459	intcept
1.00	410.00	1	-0.4219231	0.0585430	intcept
1.00	411.00	1	0.2728001	0.0755663	intcept
1.00	412.00	1	0.0151303	0.0586932	intcept
1.00	414.00	1	-0.1034214	0.0556330	intcept
1.00	415.00	1	0.3657644	0.0440240	intcept
1.00	505.00	1	-0.2049495	0.0377749	intcept
1.00	506.00	1	-0.4732580	0.0411038	intcept
1.00	507.00	1	-0.2183398	0.0381740	intcept
1.00	508.00	1	0.2533768	0.0358708	intcept
1.00	509.00	1	-0.1449054	0.0281281	intcept
1.00	510.00	1	-0.1526582	0.0291689	intcept
1.00	513.00	1	-0.3504516	0.0621495	intcept
1.00	514.00	1	0.1042001	0.0328937	intcept
1.00	515.00	1	0.0021722	0.0244088	intcept

The mean of the empirical Bayes estimates is – 0.0002. The estimates ranged from – 0.473258 for school 506 to 0.410744 for school 407. In both cases a mass-media intervention procedure was applied, and thus TV = 1, but CC = CC*TV = 0. For students with a PreTHKS score of 3 from each of these schools, this implies

$$\begin{aligned}\text{Prob}(\text{THKSbin}_{ij} = 1 \mid \text{CC} = 0, \text{PreTHKS} = 3, \text{ID} = 506) &= \frac{e^{-0.473258+0.3741+0.3871(3)}}{1+e^{-0.473258+0.3741+0.3871(3)}} \\ &= \frac{e^{1.062142}}{1+e^{1.062142}} = 0.7431\end{aligned}$$

and

$$\begin{aligned}\text{Prob}(\text{THKSbin}_{ij} = 1 \mid \text{CC} = 0, \text{PreTHKS} = 3, \text{ID} = 407) &= \frac{e^{0.410744+0.3741+0.3871(3)}}{1+e^{0.410744+0.3741+0.3871(3)}} \\ &= \frac{e^{1.946144}}{1+e^{1.946144}} = 0.8750\end{aligned}$$

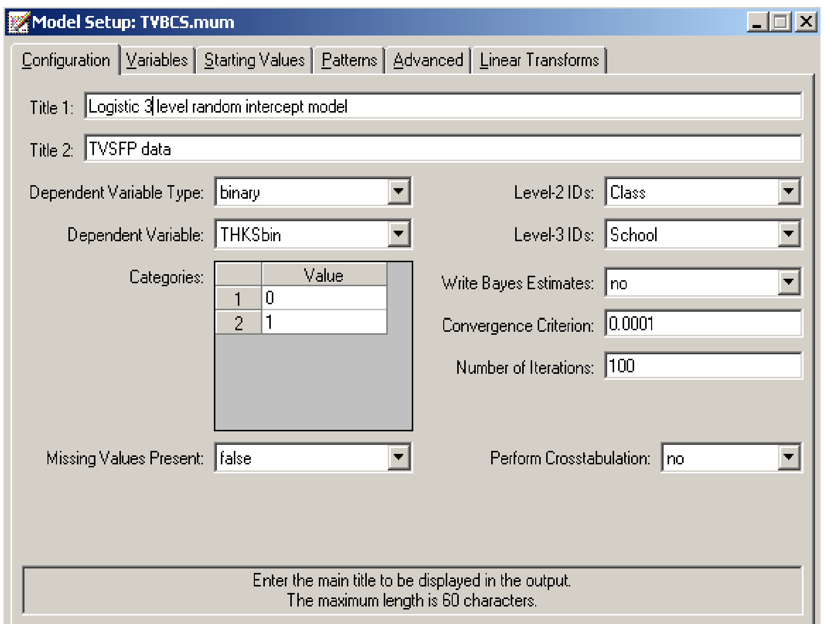
respectively. The fact that the intercept for school 407 lies higher than the average is reflected in the higher probability (87.5%) that a student with average pre-intervention knowledge score will obtain a high post-intervention score. School 506, on the other hand, has an intercept far below the average, and a student from this school has, in effect, a 74.31% chance of obtaining a high post-intervention score.

2.5 A 3-level random intercept logistic regression model

Having fitted 2-level models where students were nested within either classrooms or schools thus far, we now consider a 3-level model with both classroom and school defining levels of the hierarchy.

Setting up the analysis

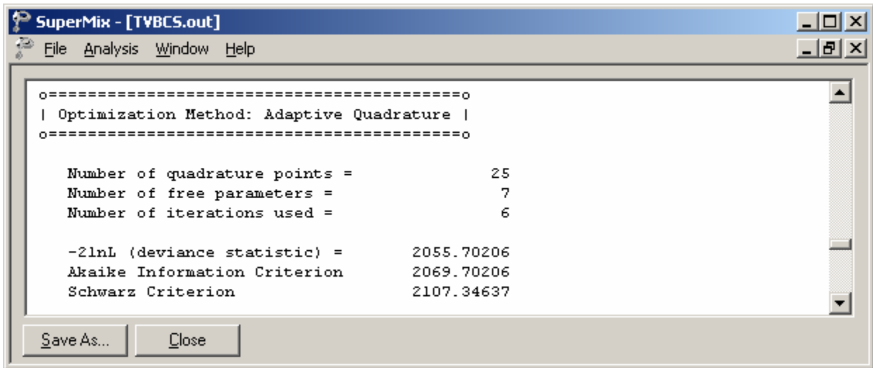
We modify our model setup saved to the syntax file **TVBS.mum** by first using the **Open Existing Model Setup** option on the **File** menu to retrieve the syntax file. Then click on **File, Save** as to save the model setup in a new file, such as **TVBSC.mum**. Next, select CLASS as the **Level-2 ID** and SCHOOL as the **Level-3 IDs** as shown below. We now have both level-2 and level-3 IDs selected.

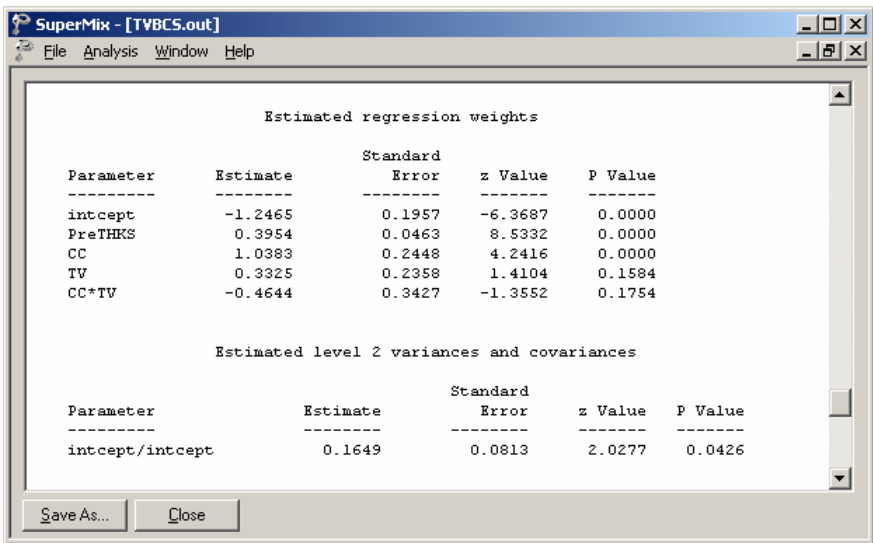


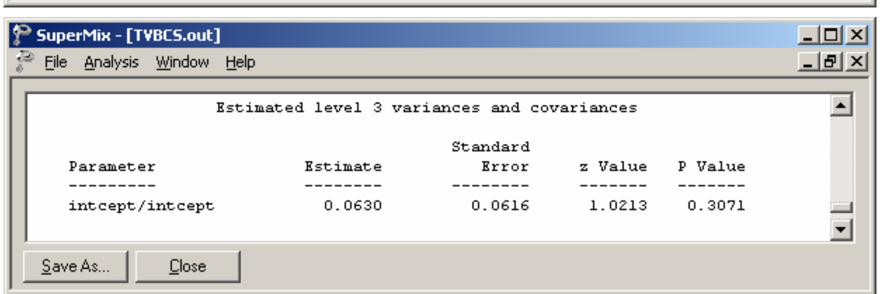
Keep all the other settings unchanged. Save the changes to the file **TVBCS.mum** and select the **Run** option on the **Analysis** menu to run the analysis.

2.6 Discussion of results

The portions of the output file **TBVSC.out** containing the estimates of the fixed and random coefficients in the current model are shown below.







Results for this model are compared to those obtained using the 2-level model previously considered, along with a model in which students are nested in classrooms. Generally, there is close agreement between the models in terms of both the sign and size of the effects. Note that the only intervention method that consistently has an estimated coefficient significantly different from zero is CC. While use of the media intervention (TV) can positively influence the post-intervention score, it seems clear that using both methods simultaneously does not have any real benefits.

Table 2.3: Comparison of results for three models with binary variable THKSbin as outcome

Coefficient		2-level: CLASS as ID	2-level: SCHOOL as ID	3-level
Fixed effects:				
Intercept	estimate	-1.2535	-1.228	-1.2465
	standard error	0.1695	0.1949	0.1957
PRETHKS	estimate	0.401	0.3871	0.3954
	standard error	0.0461	0.0451	0.0463
CC	estimate	0.9883	1.0893	1.0383
	standard error	0.1973	0.2454	0.2448
TV	estimate	0.287*	0.3741*	0.3325*
	standard error	0.192	0.235	0.2358
CCxTV	estimate	-0.369*	-0.5578*	-0.4644*
	standard error	0.2774	0.3403	0.3427
Random effects:				
Var(between classrooms)	estimate	0.2193		0.1649
	standard error	0.0802		0.0813
Var(between schools)	estimate		0.1065	0.063*
	standard error		0.0578	0.0616

*: Not significant at 5% level of significance.

3-level ICCs

Intraclass correlation coefficients can be obtained for the three-level dichotomous outcome model. As mentioned earlier, it is assumed that the level-1 error variance is equal to $\pi^2/3$ for the logistic link function if the model is true (see, e.g., Hedeker & Gibbons (2006), p. 157). Using this approximation, the formulae for the standard ICCs can be adjusted.

From the output for the random effects, we have

$$\text{Level-1: } \widehat{\text{error var}} = \pi^2/3 = 3.2899$$

$$\text{Level-2: } \widehat{\text{class var}} = 0.1649$$

$$\text{Level-3: } \widehat{\text{school var}} = 0.0630.$$

Based on this information, we can calculate the ICCs as shown below.

Similarity of students within the same school:

$$ICC = \frac{\sigma^2_{v(3)}}{\sigma^2_{v(3)} + \sigma^2_{v(2)} + \sigma^2} = \frac{0.063}{0.063 + 0.1649 + 3.28986}$$

$$= 0.0179.$$

Similarity of students within the same classrooms (and schools):

$$ICC = \frac{\sigma^2_{v(2)}}{\sigma^2_{v(3)} + \sigma^2_{v(2)} + \sigma^2} = \frac{0.1649}{0.063 + 0.1649 + 3.28986}$$

$$= 0.04688.$$

Similarity of classes within the same school:

$$ICC = \frac{\sigma^2_{v(2)}}{\sigma^2_{v(3)} + \sigma^2_{v(2)}} = \frac{0.1649}{0.063 + 0.1649}$$

$$= 0.7236.$$

Estimated unit-specific and population-average probabilities

Under the assumption that $\mathbf{v}_i$, $\mathbf{v}_{ij}$ and ε_{ijk} are independently distributed, it follows that for the three-level model the design effect is defined as

$$d_{ijk} = \frac{(\sigma^2_{v(3)} + \sigma^2_{v(2)} + \sigma^2)}{\sigma^2} = 1.0692.$$

The estimated unit-specific probabilities are calculated using

$$\begin{aligned}\hat{\eta}_{ijk} = & -1.2465 + 1.0383 \times \mathrm{CC}_i + 0.3325 \times \mathrm{TV}_i - 0.4.644 \times \mathrm{CC}_i \times \mathrm{TV}_i \\ & + 0.3954 \times \mathrm{PreTHKS}_{ijk}\end{aligned}$$

and

$$\text{Prob}(\text{THKSbin} = 1 \mid \boldsymbol{\beta}) = \frac{1}{1 + e^{-\eta_{ijk}}}$$

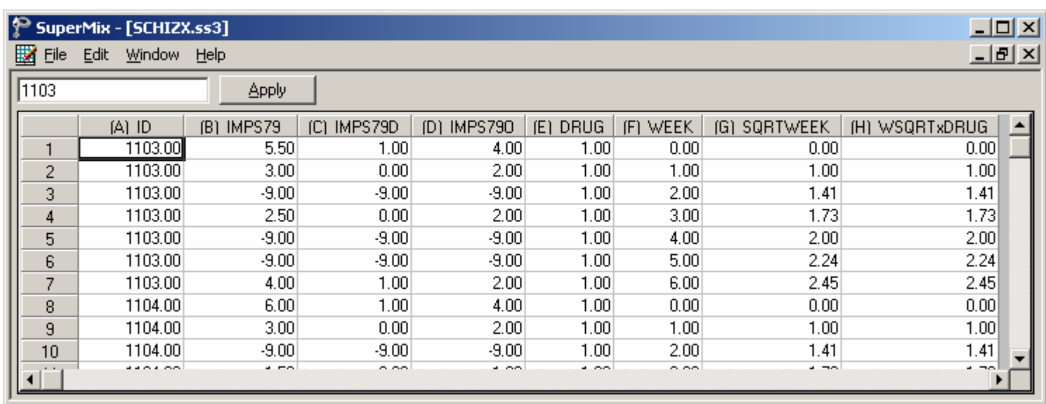
The estimated population-average probabilities (Hedeker & Gibbons, 2006) are obtained in a similar fashion as the unit-specific probabilities after replacing $\hat{\eta}_{ijk}$ with $\hat{\eta}^{*}_{ijk} = \hat{\eta}_{ijk} / \sqrt{d_{ijk}}$ in the second of the equations shown above.

3 Mixed models for ordinal outcomes

3.1 The data

To illustrate the application of the mixed-effects ordinal logistic regression model to longitudinal data, we examined data collected in the NIMH Schizophrenia Collaborative Study on treatment-related changes in overall severity. Specifically, Item 79 of the Inpatient Multidimensional Psychiatric Scale (IMPS; Lorr & Klett, 1966) was used. In this study, patients were randomly assigned to receive one of four medications: placebo, chlorpromazine, fluphenazine, or thioridazine. Since previous analyses (Longford, 1993, and Gibbons & Hedeker, 1994) revealed similar effects for the three anti-psychotic drug groups, they were combined in the present analysis. Finally, again based on previous analysis, a square root transformation of time was chosen to linearize the relationship of the IMPS79 scores over time.

Data for the first 10 observations are shown below in the form of a SuperMix spreadsheet file. Open the SuperMix spreadsheet file **schizx1.ss3** stored from the **examples\ordinal** folder. Save this spreadsheet as **schizx.ss3** in the same folder, and rename the columns headers using the **Column Properties** dialog box so that A = ID, E = Drug, and H = WSQRT*DRUG, as shown below.



	(A) ID	(B) IMPS79	(C) IMPS79D	(D) IMPS79O	(E) DRUG	(F) WEEK	(G) SQRTWEEK	(H) WSQRTxDRUG
1	1103.00	5.50	1.00	4.00	1.00	0.00	0.00	0.00
2	1103.00	3.00	0.00	2.00	1.00	1.00	1.00	1.00
3	1103.00	-9.00	-9.00	-9.00	1.00	2.00	1.41	1.41
4	1103.00	2.50	0.00	2.00	1.00	3.00	1.73	1.73
5	1103.00	-9.00	-9.00	-9.00	1.00	4.00	2.00	2.00
6	1103.00	-9.00	-9.00	-9.00	1.00	5.00	2.24	2.24
7	1103.00	4.00	1.00	2.00	1.00	6.00	2.45	2.45
8	1104.00	6.00	1.00	4.00	1.00	0.00	0.00	0.00
9	1104.00	3.00	0.00	2.00	1.00	1.00	1.00	1.00
10	1104.00	-9.00	-9.00	-9.00	1.00	2.00	1.41	1.41

The variables of interest are:

- ID indicates the subject (437 patients in total).
- IMPS79 represents the original score on Item 79 of the Inpatient Multidimensional Psychiatric Scale. It was scored as: 1 = normal, or not at all ill; 2 = borderline mentally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill; and 7 = among the most extremely ill.
- IMPS79D is a recoded version of the same scale, but in binary form, where scores up to, but excluding 3.5 were coded 0, and scores of 3.5 or higher were coded 1. The value "0" is associated with measurements classified as normal, borderline, mildly, or moderately mentally ill, while the value "1" was assigned to measurements corresponding to "markedly ill" through "most extremely ill."

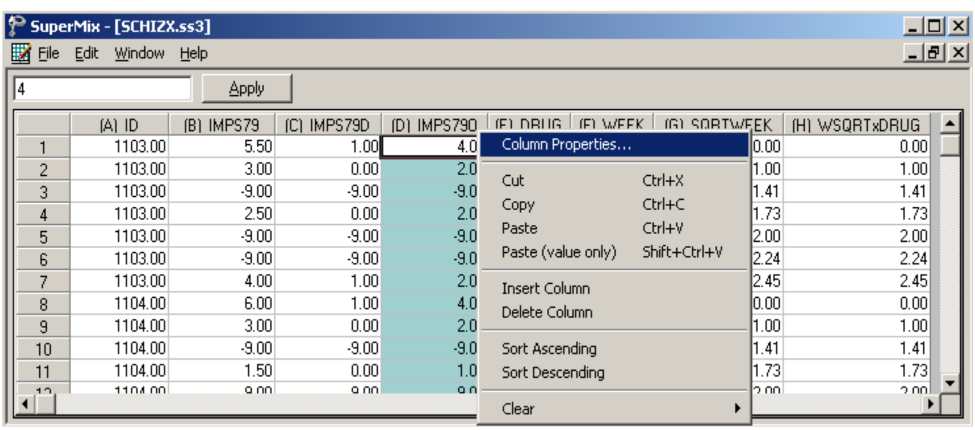
- IMPS79O is also a recoded version of the same scale, but with the 7 original categories reduced to four: 1 = normal or borderline mentally ill, 2 = mildly or moderately ill, 3 = markedly ill, and 4 = severely or among the most extremely ill.
- DRUG indicates the treatment group, where 0 indicates the placebo patients, and 1 refers to the drug patients.
- WEEK represents the time during the course of the study when a specific measurement was made, and ranges between 0 and 6.
- SQRTWEEK is the square root of WEEK. This variable is generated within the SuperMix spread sheet.
- WSQRTxDRUG is the product of the treatment group and the square root of WEEK.

In this data file, each subject's data consist of seven lines, these being the repeated measurements on seven occasions. Notice that there are missing value codes (-9) for some subjects at specific time points. The data from these time points will not be used in the analysis, but data from these subjects at other time points where there are no missing data will be used in the analysis. Thus, for inclusion into the analysis, a subject's data (both the dependent variable and all model covariates being used in a particular analysis) at a specific time point must be complete. The number of repeated observations per subject then depends on the number of time points for which there are non-missing data for that subject. The specification of missing data codes will be illustrated in the model specification section to follow.

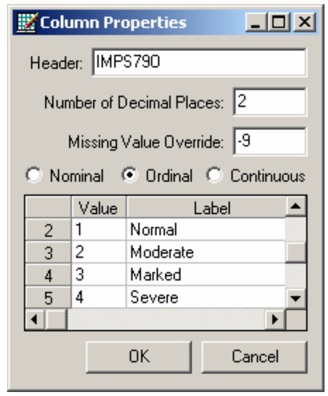
3.2 Graphical displays

Defining column properties

Defining column properties for the ordinal data is recommended. We use the column of IMPS79O as an example. First, highlight the column of IMPS79O by clicking on its header. Then right click and select the **Column Properties** option as shown below to open the **Column Properties** dialog box.



The header of the **Column Properties** dialog box indicates the current variable name. Keep the default number of decimal places unchanged. Enter -9 in the **Missing Value Override** string box. Select the **Ordinal** radio button to activate the grid field to enter the labels for each category as shown below.



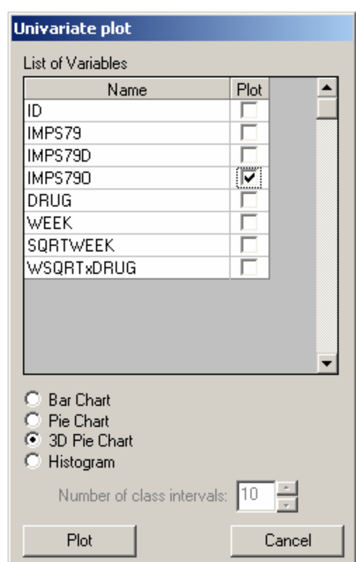
Click on the **OK** button and save the change to the data set by clicking on the **File**, **Save** option.

Univariate graphs

As a first step, we take a look at the ordinal variable IMPS79O which is the potential dependent variable in this study.

Pie chart

To generate a pie chart for IMPS79O, first open the **schizx.ss3** in the SuperMix spread sheet. Next, select the **File**, **Data-based Graphs**, **Univariate** option to load the **Univariate plot** dialog box. Select the variable IMPS79O and indicate that a **3D Pie Chart** is to be graphed as shown below.



Click the **Plot** button to display the following pie chart. Note that most of the observations fall into the Severe illness category. Keep in mind that the pie chart takes all observations, regardless of the time of measurement, into account. As such, it is informative about the distribution of all observed values of the potential outcome, but does not provide any information on possible trends in illness level over time.

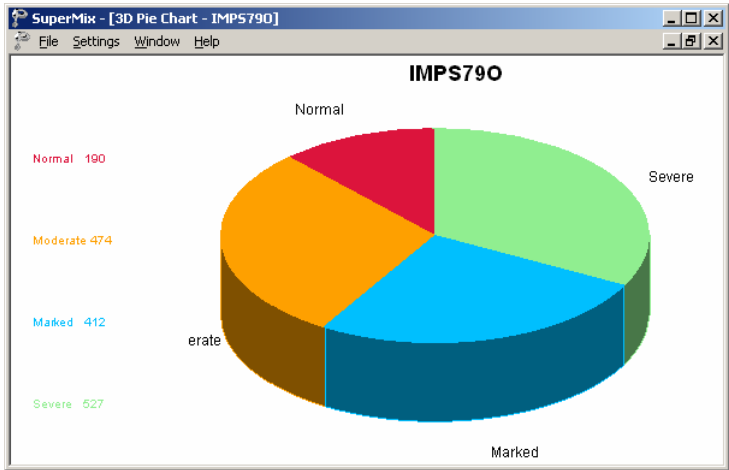


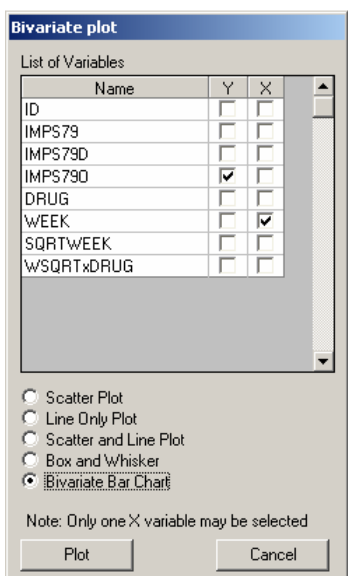
Figure 3.1: Pie chart of IMPS79O values

Relationships between variables: bivariate bar chart

It is hoped that the severity of the illness (IMPS79O) will decrease over the treatment period. Before considering fitting a model to these data, we would like to explore the relationship between IMPS79O and WEEK using a bivariate bar chart.

Bivariate bar chart

A bivariate bar chart is accessed via the **Data-based Graphs**, **Bivariate** option on the **File** menu. The **Bivariate plot** dialog box is completed as below: select the outcome variable IMPS79O as the **Y**-variable of interest, and the predictor WEEK to be plotted on the **X**-axis. Check the **Bivariate Bar Chart** option, and click **Plot**.



As shown below, most patients did not participate in the study at weeks 2, 4 and 5. At the beginning of the study (week 0), a large percentage of patients are markedly or severely ill. By the end of the study (week 6), most patients are reported as normal or moderate.

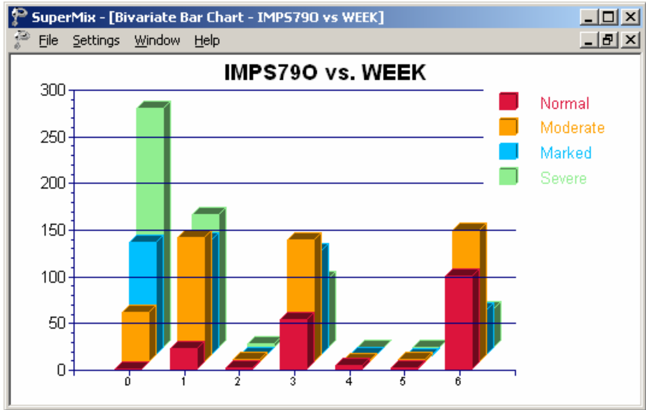


Figure 3.2: Bar chart of IMPS79O vs. WEEK

3.3 An ordinal regression model with random intercept

An ordinal variable is a categorical variable where there is a logical ordering to the categories. In most cases, treating an ordinal outcome as a continuous variable is inadvisable, due to the reasons discussed in Section XX.1.1. As in the case of a binary outcome variable, a link function is used in order to take the ceiling and floor effects of the ordinal outcome into account. The available link functions in SuperMix include probit, logistic, complementary log-log and log-log.

The model

Let the outcome variable be coded into c categories, where $c = 1, 2, ..., C$. In this example, the ordinal variable IMPS79O defines the severity of the illness in terms of four categories, and thus $C = 4$. As ordinal models utilize cumulative comparisons of the categories, define the cumulative probabilities for the C categories of the outcome Y as $P_{ijc} = \Pr\left(Y_{ij} \le c\right) = \sum_{k=1}^{c} p_{ijk}$, where p_{ijk} represents the probability that the response of the jth measurement on patient i occurs in category k.

The type of drug, time elapsed since start of treatment, and the interaction between drug taken and time elapsed are of interest as predictors. The logistic regression model with IMPS79O as outcome can then be written as

Level 1 model:

$$y_{ij} = \log\left(\frac{P_{ijc}}{1 - P_{ijc}}\right) = \gamma_c - \left[b_{0i} + b_{1i}\text{DRUG}_i + b_{2i}\text{SQRTWEEK}_i + b_{3i}\left(\text{WSQRT}\times\text{DRUG}\right)_i\right],$$
$$j = 1, \cdots, n_i;\ c = 1, 2, \cdots, C - 1$$

Level 2 model:

$$\begin{aligned} b_{0i} &= \beta_0 + v_{0i}, \quad i = 1, \cdots, N \\ b_{1i} &= \beta_1 \\ b_{2i} &= \beta_2 \\ b_{3i} &= \beta_3 \end{aligned}$$

The cumulative probability can be expressed by

$$P_{ijc} = \frac{e^{\gamma_c - \left[b_{0i} + b_{1i}\mathrm{DRUG}_i + b_{2i}\mathrm{SQRTWEEK}_i + b_{3i}\left(\mathrm{WSQRT}\times\mathrm{DRUG}\right)_i\right]}}{1 + e^{\gamma_c - \left[b_{0i} + b_{1i}\mathrm{DRUG}_i + b_{2i}\mathrm{SQRTWEEK}_i + b_{3i}\left(\mathrm{WSQRT}\times\mathrm{DRUG}\right)_i\right]}}$$

To obtain the probability for category c ,

$$p_{ij,c} = P_{ij,c+1} - P_{ij,c}$$

As shown above, the intercept b_{0i} is estimated by a level-2 equation. It indicates that patient i's initial IMPS79O value is not only determined by the population average β_0 , but also by the patient difference v_{0i} . In other words, patients may have different average intercepts, and the model makes provision for this eventuality. The slopes are assumed to be the same for all the patients, which implies that each patient's trend line is parallel to the population trend.

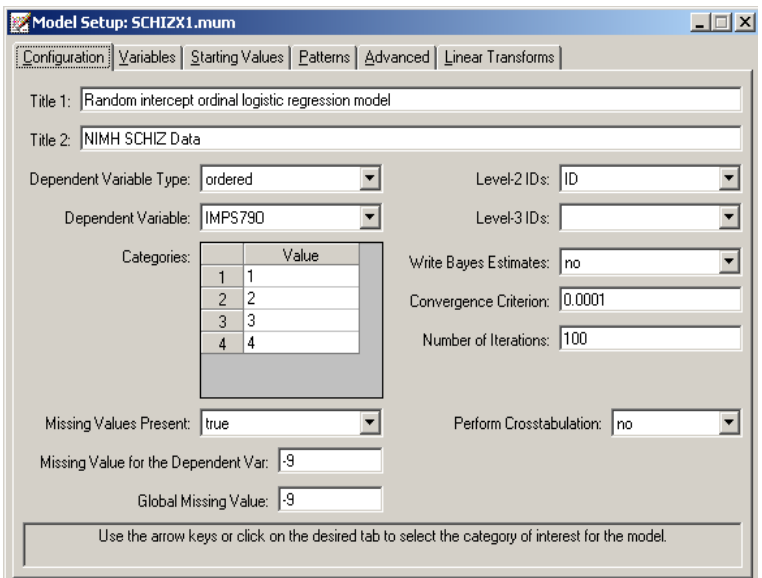
The connection between an ordinal outcome variable y with C categories and an underlying continuous variable y^* is

$$y = c \leftrightarrow \gamma_{j-1} \le y^* \le \gamma_j, \;\; c = 1, 2, ..., C$$

where it is assumed that $\gamma_0 = -\infty$ and $\gamma_C = +\infty$. In addition, γ_1 is usually set to 0 to avoid identification problems.

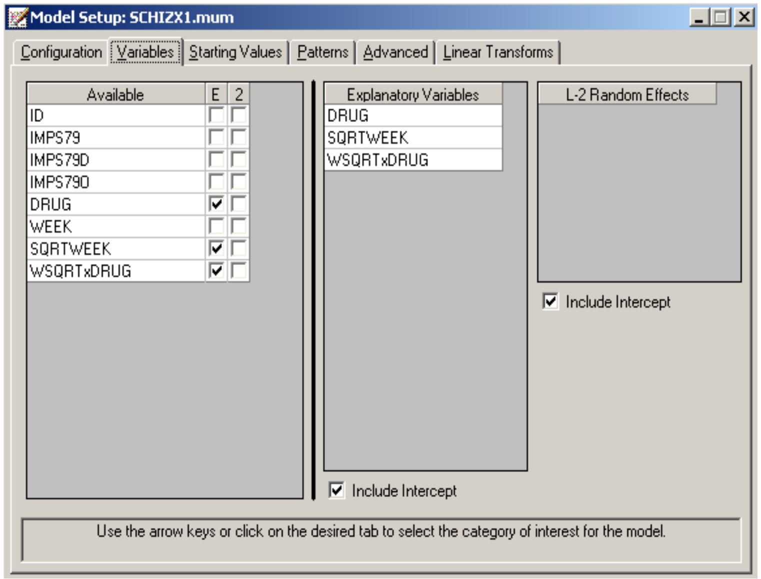
Setting up the analysis

Open the SuperMix spreadsheet **schizx.ss3** and select the **File**, **New Model Setup** option. In the **Configuration** screen of the **Model Setup** window, enter a title for the analysis in the **Title** text boxes. Select **ordered** from the **Dependent Variable Type** drop-down list box. Select the outcome variable IMPS79O from the **Dependent Variable** drop-down list box. Once this selection has been made, the **Categories** grid is displayed, with the distinct values of the categories shown.

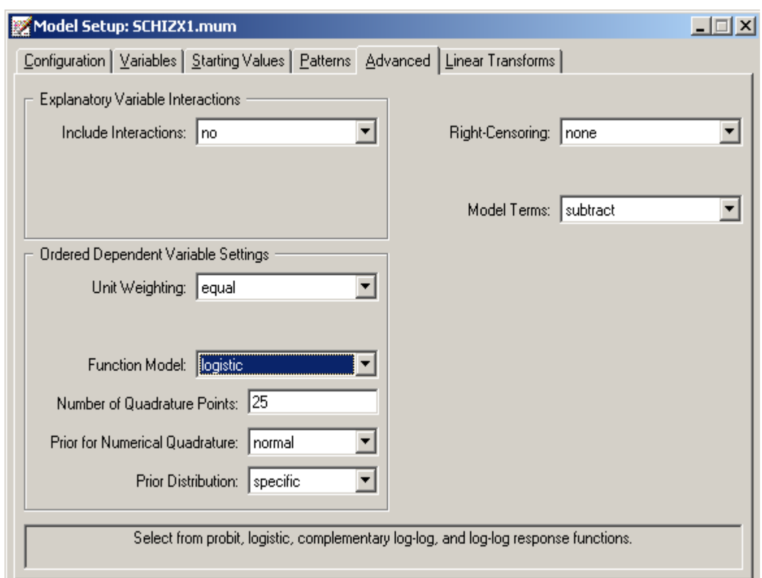


We notice that the missing value -9 is also included as a category. The **Missing Values Present** drop-down list box is used to specify the values of missing data for both outcome and predictors. As a first step, set the value of the **Missing Values Present** drop-down list box to **True**. The appearance of the screen will change when this is done, and text boxes for the specification of the missing data codes are displayed. Start by entering the value -9 in the **Missing Value for the Dependent Var** text box. Do the same for all the predictors included in the model by entering -9 in the **Global Missing Value** text box. Finally, select the patient ID from **Level-2 IDs** drop-down list box to produce the **Configuration** screen seen above.

Proceed to the **Variables** screen by clicking on this tab. The **Variables** tab is used to specify the fixed and random effects to be included in the model. Select DRUG, SQRTWEEK and WSQRTxDRUG as explanatory (fixed) variables using the **E** check boxes next to the variables names in the **Available** grid at the left of the screen. The **Include Intercept** check box in the **Explanatory Variables** grid is checked by default, indicating that an intercept term will automatically be included in the fixed part of the model. Next, specify the random effects at level 2 of the hierarchy. In this example, we want to fit a model with random intercepts at level 2. By default, the **Include Intercept** check box in the **L-2 Random Effects is** checked, indicating the inclusion of a random intercept at this level in the model.



The default link function for the ordinal outcome variable is the probit link function. To change it to the logistic link function corresponding to the model formulation above, click on the **Advanced** tab and select the **logistic** link function from the **Function Model** drop-down list box as shown below. Use 25 quadrature points.

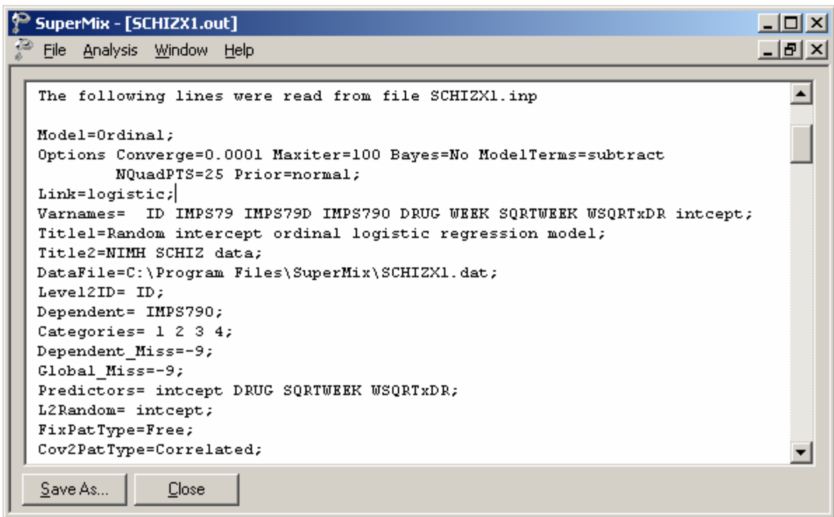


Before running the analysis, the model specifications have to be saved. Select the **File**, **Save As** option, and provide a name (**SCHIZX1.mum**) for the model specification file. Run the analysis by selecting the **Run** option from the **Analysis** menu.

3.4 Discussion of results

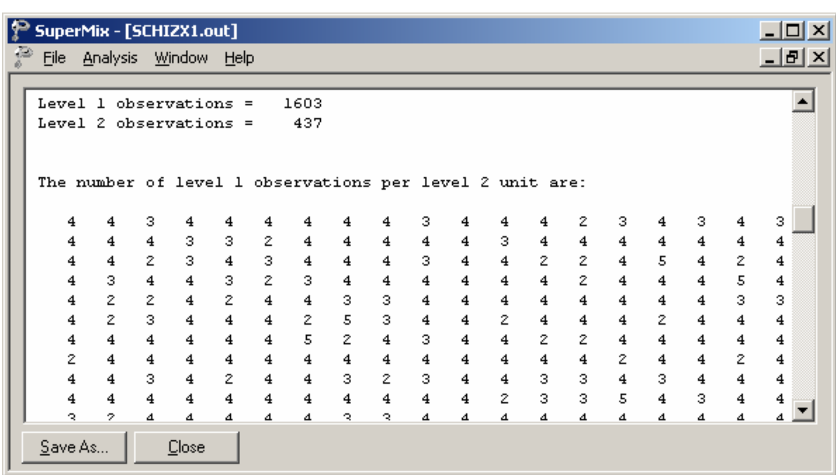
Syntax

The syntax corresponding to the model setup is given in the **model specifications**. These lines of SuperMix syntax are saved as a ***.inp** file with the same name as the model setup file (***.mum**). At the top of the output file, the syntax lines are printed as shown below. The first part indicates that an ordinal outcome is analyzed, states the selection of iteration control options, does not request Bayes residuals, and contains all the specifications necessary to define the model fitted as an ordinal model with logistic link function. The second part of the syntax provides information on the structure of the data, the name and structure of the outcome variable, the missing values and the predictors included in the model.



Data summary

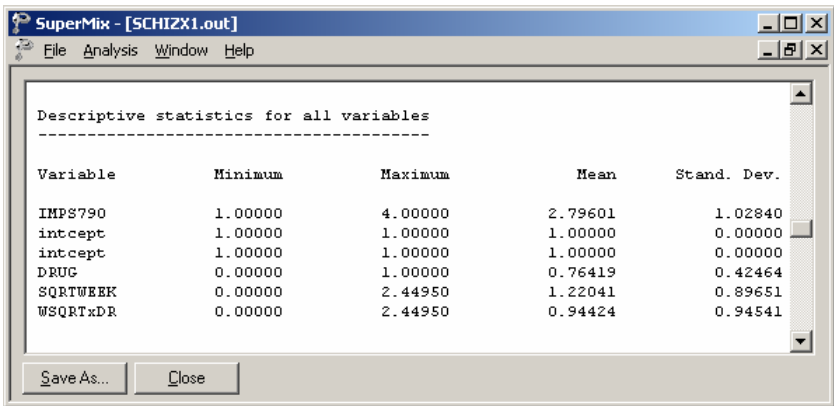
The next section of the output file contains a description of the hierarchical structure and model specifications. The use of a logistic response function (logit link function) with the assumption of a normal distribution of random effects is indicated.



This is followed by a summary of the number of observations nested within each patient. As shown below, 437 patients with a total of 1603 observations are included in this study after listwise deletion. The number of observations per patient (level 2 unit) varies between 2 and 5.

Descriptive statistics and starting values

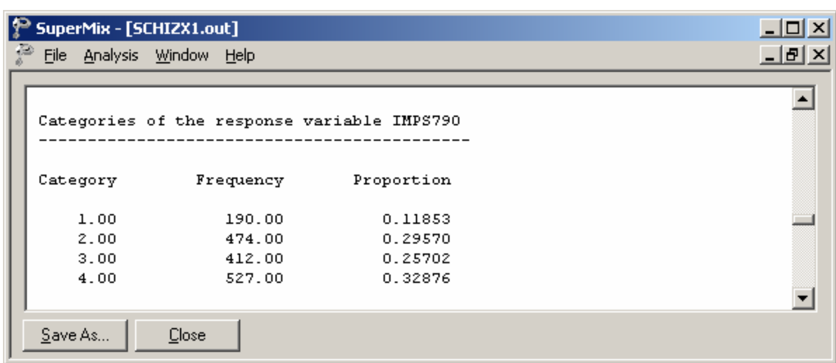
Next, the descriptive statistics for all the variables are given. Notice that the variable name WSQRTxDRUG is truncated to WSQRTxDR. This is because SuperMix only recognizes the first 8 characters of a variable name.



Descriptive statistics for all variables

Variable	Minimum	Maximum	Mean	Stand. Dev.
IMPS790	1.00000	4.00000	2.79601	1.02840
intcept	1.00000	1.00000	1.00000	0.00000
intcept	1.00000	1.00000	1.00000	0.00000
DRUG	0.00000	1.00000	0.76419	0.42464
SQRTWEEK	0.00000	2.44950	1.22041	0.89651
WSQRTxDR	0.00000	2.44950	0.94424	0.94541

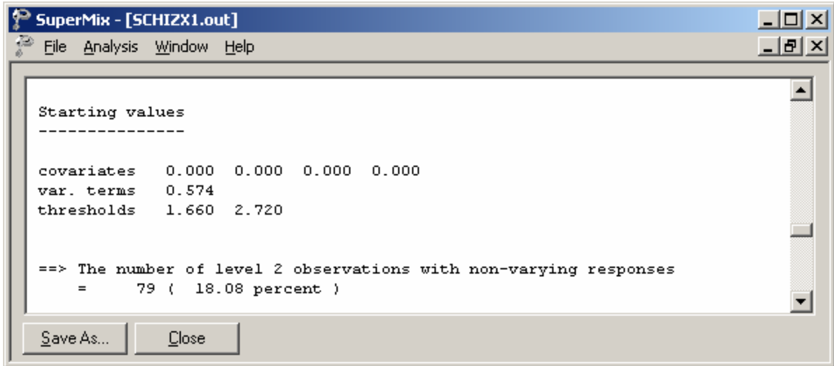
As shown below, the output file for the ordinal outcome also provides a frequency table for the dependent variable. The data summary is followed by descriptive statistics for all the variables included in the model (not shown). We note that 33% of the measurements were in the highest category of the outcome variable, and correspond to the "severely or among the most extremely ill" group. Only 12% of measurements are in the first category ("normal, not at all ill").



Categories of the response variable IMPS790

Category	Frequency	Proportion
1.00	190.00	0.11853
2.00	474.00	0.29570
3.00	412.00	0.25702
4.00	527.00	0.32876

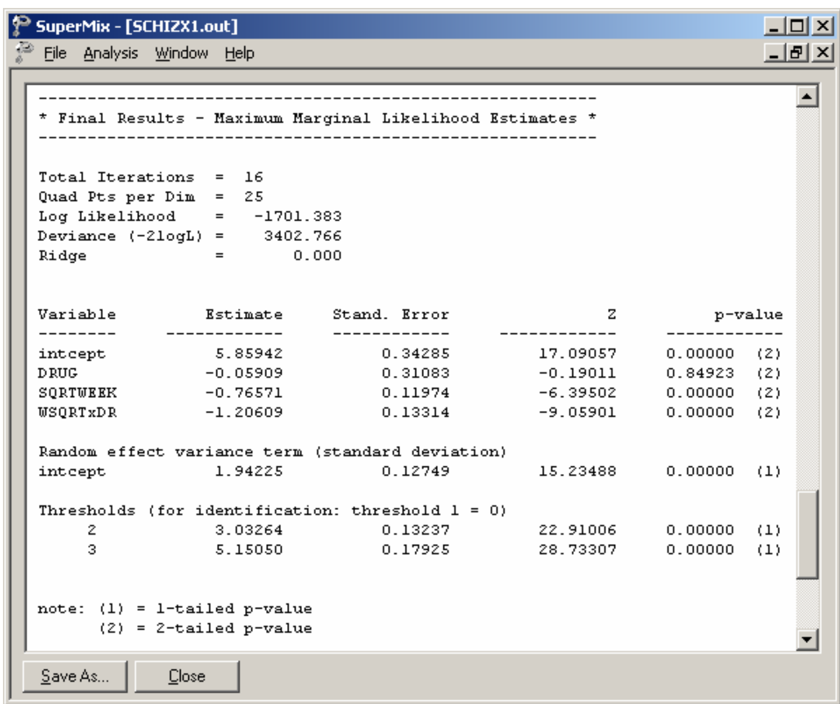
Descriptive statistics are followed by the starting values of parameters. The starting values for the predictors intercept, DRUG, SQRTWEEK and WSQRTxDR are given in the first line (covariates), while the starting value for the variance component associated with the random level-2 intercept is given in the second line (var. terms). The third line shows the starting values of the thresholds. In 18% of the subjects, no change in the category assigned for measurements was observed, as indicated by the last two lines shown below. Since the first threshold is fixed at 0 for identification purposes, starting values for the second and third thresholds only are listed.



```
SuperMix - [SCHIZX1.out]
File  Analysis  Window  Help

Starting values
---------------

covariates   0.000  0.000  0.000  0.000
var. terms   0.574
thresholds   1.660  2.720


==> The number of level 2 observations with non-varying responses
    =     79 (  18.08 percent )

Save As...   Close
```

Fixed effects estimates

The final results after 16 iterations are shown next. The estimates are shown in the column with heading Estimate, and correspond to the coefficients $\beta_0, \beta_1, \ldots, \beta_3$ in the model specification. The standard error, *Z*-value and *p*-value are also printed.



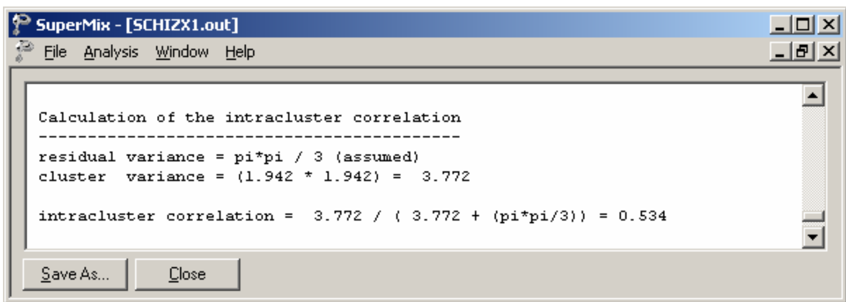
```
SuperMix - [SCHIZX1.out]
File  Analysis  Window  Help

---------------------------------------------------------
* Final Results - Maximum Marginal Likelihood Estimates *
---------------------------------------------------------

Total Iterations  =  16
Quad Pts per Dim  =  25
Log Likelihood    =   -1701.383
Deviance (-2logL) =    3402.766
Ridge             =       0.000


Variable        Estimate     Stand. Error            Z         p-value
--------      ------------   ------------   ------------   ------------
intcept          5.85942        0.34285       17.09057     0.00000  (2)
DRUG            -0.05909        0.31083       -0.19011     0.84923  (2)
SQRTWEEK        -0.76571        0.11974       -6.39502     0.00000  (2)
WSQRTxDR        -1.20609        0.13314       -9.05901     0.00000  (2)

Random effect variance term (standard deviation)
intcept          1.94225        0.12749       15.23488     0.00000  (1)

Thresholds (for identification: threshold 1 = 0)
     2           3.03264        0.13237       22.91006     0.00000  (1)
     3           5.15050        0.17925       28.73307     0.00000  (1)


note: (1) = 1-tailed p-value
      (2) = 2-tailed p-value

Save As...   Close
```

The variation in the intercept over the subjects is estimated as $1.94225^2 = 3.77233$, and from the associated *p*-value we conclude that there is significant variation in the (random) intercept between the patients included in this analysis. In the case of the fixed effects, a 2-tailed *p*-value is used, as the alternative hypothesis considered here is of the form $H_1: \beta \neq 0$. As variances are constrained to be elements of the interval $[0, +\infty)$ and thresholds are constrained so that $\gamma_1 \leq \gamma_2 \leq \gamma_3$, the *p*-values used for these effects are 1-tailed. The results indicate that the treatment groups do not differ significantly at baseline (the estimated DRUG coefficient is not significant). The placebo group seems to improve over time, as the SQRTWEEK coefficient is both significant and negative. Note that the interpretation of the main effects depends on the coding of the variable, and on the significance of the WSQRTxDR interaction which forms part of the model.

As noted before, it is assumed that $\gamma_0 = -\infty$ and $\gamma_C = +\infty$, with γ_1 usually set to 0 to avoid identification problems. For the present example, C = 4, and from the output we see that $\hat{\gamma}_2 = 3.03264$ and $\hat{\gamma}_3 = 5.15150$. These values are used in combination with the coefficients of DRUG, SQRTWEEK, and WSQRTxDR to calculate estimated outcomes for different groups of patients (see Section XXX).

Intraclass correlation (ICC)

Below the estimate the intracluster correlation (ICC) is given. The residual variance for the logistic link function is assumed to be $\pi^2/3$.



The ICC in this model refers to the intra-person correlation. It is reported as 0.534, which is fairly high. Generally, the shorter the interval between the repeated measurements, the higher the ICCs will be.

Estimated outcomes for groups: unit-specific probabilities

To evaluate the expected effect of the treatment group and the square root of time of treatment, while allowing for the interaction between treatment and the square of time, we use the expression below:

$$\log\left(\frac{\hat{P}_{ijc}}{1-\hat{P}_{ijc}}\right) = \hat{\gamma}_c - \left[\hat{b}_{0i} + \hat{b}_{1i}\text{DRUG}_i + \hat{b}_{2i}\text{SQRTWEEK}_i + \hat{b}_{3i}\left(\text{WSQRT}\times\text{DRUG}\right)_i\right]$$

or

$$\log\left(\frac{\hat{P}_{ijc}}{1-\hat{P}_{ijc}}\right) = \hat{\eta}_{ijc}$$
$$= \hat{\gamma}_c - 5.85942 + 0.05909 \times \mathrm{DRUG}_i + 0.76571 \times \mathrm{SQRTWEEK}_i$$
$$+1.20609 \times \left(\mathrm{WSQRT} \times \mathrm{DRUG}\right)_i .$$

When c = 1 and $\gamma_1 = 0$, we find that, for a patient from the control group (DRUG = 0, SQRTWEEK = WSQRTxDR = 0),

$$\log\left(\frac{\hat{P}_{ij1}}{1-\hat{P}_{ij1}}\right) = \hat{\eta}_{ij1} = 0 - 5.85942$$
$$\hat{P}_{ij1} = \frac{e^{\hat{\eta}_{ij1}}}{1+e^{\hat{\eta}_{ij1}}} = 0.002844$$

Similarly, the probabilities that a typical patient from the control group responded in a specific category at the start of the study are obtained by substituting $\gamma_1 = 0$ with $\hat{\gamma}_2 = 3.03264$, and $\hat{\gamma}_3 = 5.15050$. The cumulative probabilities we calculated are

$$\hat{P}_{ij2} = \frac{e^{\hat{\eta}_{ij2}}}{1+e^{\hat{\eta}_{ij2}}} = \frac{e^{3.03264\text{-}5.85942}}{1+e^{3.03264\text{-}5.85942}} = 0.05589$$
$$\hat{P}_{ij3} = \frac{e^{\hat{\eta}_{ij3}}}{1+e^{\hat{\eta}_{ij3}}} = \frac{e^{5.1505\text{-}5.85942}}{1+e^{5.1505\text{-}5.85942}} = 0.32984$$

Thus, the estimated category probabilities we have for such a group (category 1 to 4) are obtained as

$$\hat{p}_{ij1} = 0.00284 - 0 = 0.00284$$
$$\hat{p}_{ij2} = 0.05589 - 0.00284 = 0.05305$$
$$\hat{p}_{ij3} = 0.32984 - 0.05589 = 0.27394$$
$$\hat{p}_{ij4} = 1 - 0.32984 = 0.67016$$

For this group of patients (DRUG = 0) at the starting week, the expected percentages of patients in each of the categories are as follows: 0.3% of the patients are normal or borderline mentally ill; 5.3% of the patients are mildly or moderately ill; 27.4% are markedly ill and 67% are severely or extremely ill. Similarly, we can calculate the estimated percentages for both groups at all the time points as shown in Table 3.1.

Table 3.1: Estimated % for both groups at 7 time points

	Placebo patients (drug = 0)				Drug patients (drug = 1)			
severity	**normal**	**moderate**	**marked**	**severe**	**normal**	**moderate**	**marked**	**severe**
week 0	0.28%	5.30%	27.39%	67.02%	0.30%	5.61%	28.39%	65.70%
week 1	0.61%	10.68%	40.13%	48.58%	2.13%	28.96%	47.86%	21.05%
week 2	0.84%	14.05%	44.36%	40.76%	4.69%	45.83%	38.94%	10.54%
week 3	1.06%	17.17%	46.73%	35.04%	8.43%	57.21%	28.43%	5.92%
week 4	1.30%	20.19%	47.98%	30.52%	13.51%	62.91%	20.00%	3.58%
week 5	1.56%	23.15%	48.47%	26.83%	19.92%	63.85%	13.95%	2.28%
week 6	1.83%	26.04%	48.39%	23.75%	27.48%	61.24%	9.78%	1.51%