Instructions for SAS Macro  
qmin_repeated_measures

Examine the data set to get the variable names and the design. (See SAS note Viewing the Contents of a SAS Data Set for instructions in how to do this.) In this example, the data set is called Interrupted2 and it resides within the directory given the nickname of qmrmlib.

This data set examines the effect of a compound hypothesized to impede memory. Sixteen rats were given acquisition trials in an avoidance paradigm that previous research had shown to asymptote after five trials. The scores for the rats on these five acquisition trials have been recorded as variables A_Trial1 through A_Trial5 in the data set. After the 5th acquisition trial, the rats were randomly divided into two groups, a Control and a Treatment group. Those in the Treatment group were administered the drug while those in the Control group were administered the vehicle. Rats were then given five retention trials, recorded in the data set as variables R_Trial1 through R_Trial5.

There are several ways to analyze these data. One could first test whether the Controls differ from the Treatment group on the five acquisition trials. Because assignment was made after these trials, there should be no differences. Afterwards, one can test for differences between the groups on the five retention trials. A second—but more complicated—way is to analyze all ten trials together. We will use this approach here in order to illustrate the use of two repeated measures factors.

We can identify two repeated measures factors. The first of these may be called Mode and it divides the trials into two types—acquisition mode versus retention mode. The second repeated measures factor may be called Trial and it gives the trial number within each mode. The design for the within-subjects factors is given in Table 1. It is recommended that you always create a table of the dependent variables, the within-subjects factors, and the levels of the within-subjects factors akin to Table 1.
Table 1. Within-subjects factors and their levels.

<table>
<thead>
<tr>
<th>Variable in the data set:</th>
<th>Value of Mode:</th>
<th>Value of Trial:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_Trial1</td>
<td>Acquisition</td>
<td>1</td>
</tr>
<tr>
<td>A_Trial2</td>
<td>Acquisition</td>
<td>2</td>
</tr>
<tr>
<td>A_Trial3</td>
<td>Acquisition</td>
<td>3</td>
</tr>
<tr>
<td>A_Trial4</td>
<td>Acquisition</td>
<td>4</td>
</tr>
<tr>
<td>A_Trial5</td>
<td>Acquisition</td>
<td>5</td>
</tr>
<tr>
<td>R_Trial1</td>
<td>Retention</td>
<td>1</td>
</tr>
<tr>
<td>R_Trial2</td>
<td>Retention</td>
<td>2</td>
</tr>
<tr>
<td>R_Trial3</td>
<td>Retention</td>
<td>3</td>
</tr>
<tr>
<td>R_Trial4</td>
<td>Retention</td>
<td>4</td>
</tr>
<tr>
<td>R_Trial5</td>
<td>Retention</td>
<td>5</td>
</tr>
</tbody>
</table>

Hence, the overall design is a three-way GLM. The first factor is the between-subjects variable Group which has two levels—Control and Treatment. The second is the within-subjects factor Mode which also has two levels—Acquisition and Retention. The third is the within-subjects variable Trial and it has five levels.

With this design at hand and the variable names, we can now examine how to specify this using macro \texttt{qmin\_repeated\_measures}.

**Call the macro.** Go to the Program Editor window and type in the phrase “\texttt{%qmin\_repeated\_measures;}” on the first line (without the quotes, of course). Make certain to include the percent sign (%) at the beginning because this informs SAS that a macro is being called. And do not forget the semicolon! Then submit this statement.

The following window will appear (but without the red circle, red arrow, and red message).
The macro consists of displaying a series of informational windows into which you must enter information. Enter your responses in the blue field(s) of the window.

You can submit your entries in one of two ways: (1) complete all the fields and, then press Enter (or on some keyboards, Return) or (2) press the Home key on the keyboard and then press Enter (or Return).

If you make an error, then an error message will appear in the window. Just alter your response and continue.

To abort the macro, press the Home key on the keyboard and then enter the word “abort” (without the quotes).

Because our data has each observation in a single row, we would enter a “y” (without quotes) in the field above and press Enter. An informational window will appear (not shown here). Just read the information and press Enter.

**Enter the data set name.** The next window requests that you enter the name of the data set to be analyzed. This macro does not alter the data set, so you can use a master data set if you wish. In this case, we enter the phrase `qmrmlib.Interrupted2` into the window. Remember to always include the nickname for the directory containing the SAS data set.
Enter the variable that identifies the observations (i.e., the subjects). This is the variable in the data set that has a unique value for each observation (i.e., row) in the data set. In the present example, that variable is *Rat*.

**Command***

```
QMIN: Repeated measures analysis using PROC MIXED
```

Enter `ABORT` on the command line to abort.

**Observation ID variable:** `Rat`

**HINT:** In the Explorer window, double click on the data set to examine its contents.

Enter the between-subjects variable(s). These are the names of the between-subject variables in the data set. In this example, that variable is *Group*.

**Command***

```
QMIN: Repeated measures analysis using PROC MIXED
```

Enter `ABORT` on the command line to abort.

**Between-subjects variables** (just press `ENTER` if there are no between-subjects variables)

```
Group
```

**HINT:** Use a double hyphen, e.g. `--`, to select consecutive variables.
**Enter the dependent (i.e., within-subjects) variable(s).** These are the names of the dependent or within-subject variables in the data set. *The order in which you enter the names of the variables must correspond to the within-subjects factors in your design.* Because of the design we used (see Table 1), the order of the variables corresponds to their order in the data set. Hence, we can enter `A_Trial1 -- R_Trial5`.

**Enter the number of within-subjects factors.** We have two within-subjects factors, Mode and Trial. Hence we enter 2 into this window. Make certain that you enter the number of factors and not the number of levels.

**Enter the name(s) and the number of levels for the within-subject factor(s).** Here, we give the names and the number of levels for each within-subjects factor. When there is more than one within-subject factor, then the order of entry is important. Enter the one that moves slowest first. Just read the names of the dependent variables in the window from left to right and ask yourself, “What are the values of the within-subjects factors for these variables?” The within-subject factor that changes slowest should be entered first and the one that changes fastest should be entered last.
Let us take the first two variables, A_Trial1 and A_Trial2, as an example. The value of these two variables on the Mode factor is both Acquisition while their values on the Trial factor are, respectively, 1 and 2. Hence factor Mode is the slower and factor Trial is the faster, so we will enter Mode first, followed by Trial.

Give a name for the new dependent variable. The macro will rearrange the data in a stacked format so that the values for all of the dependent variables are in a single column. It is necessary to give this column a name. Enter it here. In the example, the dependent variable will be called Response.
Give a name for the new, rearranged data set. This can be any name that you choose. Remember to use the nickname.data-set-name convention if you want to save the new data set. We will simply call the data set tempdata.

Verify the new data set. View the new data set to make certain that you correctly specified the observational variable, the between-subjects variables, the dependent variables, and the within-subjects factors. (See SAS note Viewing the Contents of a SAS Data Set for instructions in how to do this.)

The new data set will contain the observational variable (Rat, in our example), the between-subjects variable(s) (Group), a variable for each of the within-subjects-factor (Mode and Trial), the new dependent variable (Response), and a variable called Source that gives the name of the original dependent variable. The values of the within-subjects variables will be consecutive integers beginning with 1. You should make certain that these values agree with your design. If they do not, then press the Home key on the keyboard, enter the word “abort” (without quotes) on the command line, and start over.
Identity the quantitative and the categorical variables. The next window will list all the independent variables (both between-subjects and within-subjects) that you previously specified. Place any character next to a variable to indicate that it should be treated as a quantitative variable. A blank field will make the variable a categorical variable. For this example, we treat Trial as a quantitative variable.
Identify possible forms for the error covariance matrix. We now want to select some plausible forms for the error covariance matrix. Place any character beside the form that you want to fit to the data. We expect that there will be some correlation over the trials, so we will not fit any form that specifies that the correlations are 0.0. For the sake of illustration, we fit all the other forms.

Command --->

- QMIN: Repeated Measures Analysis using PROC MIXED

Enter ABORT on the command line to abort.

The first step is to find a form for the error covariance matrix.

Here is a list of the forms for the error covariance matrix.
Any character placed in the field to the left of the option will fit that form.
Leaving that field blank will not fit that form.
Select as many forms as wish.

VC = Some Variance, Correlations = 0
UN(1) = Different Variances, Correlations = 0
CS = Some Variance, Some Correlation (Compound Symmetry)
CSH = Different Variances, Some Correlation
UNR = Different Variances, Different Correlations
HF = Huhns-Feldt
AR(1) = First Order Autoregressive
ARH(1) = First Order Autoregressive, Different Variances
TOEP = Toeplitz
TOEP/H = Toeplitz, Different Variances
Specify the most general model. The purpose of specifying a model here is merely to examine the fit for the different error covariance matrices. A model with all the main effects, two-way interactions, and three-way interactions will automatically appear. Amend this model if needed.

Be patient. The macro will now call PROC MIXED to fit the model specified in the previous step for each of the forms for the error covariance matrix that you specified. You can go to the SAS Output window if you want to watch the progress of these model fits.

After all models have been fitted, an informational window (not shown) will appear that instructs you to examine the last part of the output. Go to the Output window and page down to the end. It should have the following form:
Select the best form for the error covariance matrix. The purpose of this output is to select the best form for the error covariance matrix. There are several ways of doing this, but the one recommended here is to examine the columns labeled AIC (Akaike’s Information Criterion), AICC (corrected Akaike Information Criterion), and BIC (Schwarz’s Bayesian Information Criterion) for the lowest values. For the present example, these are indicated by the red arrows in the previous figure.

It is not necessary to be too obsessional in your selection. Notice that the three fit indices fall into two groups—the CS, CSH, UNR, and HF are all high while the AR(1), ARH(1), TOEP, and TOEPH are all low. Hence, we should avoid selecting a form from the first group. Which of the forms from the latter group, however, is the best will not make much difference in the analysis. Let us pick the Toeplitz matrix as the form for the error covariance matrix.

Specify the form for the error covariance matrix for the analysis. The next window repeats all the forms for the error covariance matrix. Identify one and only one of these. All the remaining models will fit this form for the error covariance matrix to the data. Because we have selected the Toeplitz, we note that form in the window.

Enter a model to be fitted to the data. Now a series of windows will appear. Each window requests that you enter a model. After you have entered the model, press Enter (or Return) and that model will be fitted to the data using the form for the error covariance matrix that you previously specified. When you have fitted all the models that you want, enter a blank line.

In the following windows, we fit; (1) a full model; (2) a model without the three-way interaction; (3) three models, each one without the three-way interaction, but eliminating one of the two-way interactions; and (4) a blank line to terminate the model fits.
QMIN SAS Macros

qmin_repeated_measures - 12

```sas
Command ===>  

**QMRepeat Measures Analysis using PROC MIXED**

Enter ABORT on the command line to abort.

The independent variables are:
Group Mode Trial

Enter a model to be fitted to the data below. A blank line stops model fitting.
Group Mode Trial Group\Mode Group\Trial Mode\Trial Group\Mode\Trial
```

```sas
Command ===>  

**QMRepeat Measures Analysis using PROC MIXED**

Enter ABORT on the command line to abort.

The independent variables are:
Group Mode Trial

Enter a model to be fitted to the data below. A blank line stops model fitting.
Group Mode Trial Group\Mode Group\Trial Mode\Trial Group\Mode\Trial
```

```sas
Command ===>  

**QMRepeat Measures Analysis using PROC MIXED**

Enter ABORT on the command line to abort.

The independent variables are:
Group Mode Trial

Enter a model to be fitted to the data below. A blank line stops model fitting.
Group Mode Trial Group\Mode Group\Trial Mode\Trial Group\Mode\Trial
```
COMMAND: PROC REPEATED MEASURES ANALYSIS USING PROC MIXED

Enter ABORT on the command line to abort.

The independent variables are:
  Group Mode Trial

Enter a model to be fitted to the data below. A blank line stops model fitting.
  Group Mode Trial Group Mode Trial

COMMAND: PROC REPEATED MEASURES ANALYSIS USING PROC MIXED

Enter ABORT on the command line to abort.

The independent variables are:
  Group Mode Trial

Enter a model to be fitted to the data below. A blank line stops model fitting.
  Group Mode Trial Group Mode Trial Mode Trial

COMMAND: PROC REPEATED MEASURES ANALYSIS USING PROC MIXED

Enter ABORT on the command line to abort.

The independent variables are:
  Group Mode Trial

Enter a model to be fitted to the data below. A blank line stops model fitting.
Examine the results of the model fits. Go to the SAS Output window and page down to the very bottom. You will see a window resembling this:

```
Data Set = tempdata
Model = Group Mode Trial Group*Mode Trial Group*Mode*Trial

Models that were fitted to the data:

<table>
<thead>
<tr>
<th>Model</th>
<th>DF</th>
<th>Independent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>Group Mode Trial Group<em>Mode Group</em>Trial Mode*Trial</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>Group Mode Trial Group<em>Mode Group</em>Mode Group*Trial</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>Group Mode Trial Group<em>Mode Group</em>Mode</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Group Mode Trial Group*Mode</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>Group Mode</td>
</tr>
</tbody>
</table>

Fit Statistics for Models:
AIC, AICC, BIC: Smallest is best

<table>
<thead>
<tr>
<th>Model</th>
<th>DF</th>
<th>-2 Log(L)</th>
<th>AIC</th>
<th>AICC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>908.02</td>
<td>928.02</td>
<td>929.58</td>
<td>935.75</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>913.94</td>
<td>933.94</td>
<td>935.49</td>
<td>941.67</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>931.18</td>
<td>951.18</td>
<td>952.72</td>
<td>958.90</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>933.03</td>
<td>953.03</td>
<td>954.57</td>
<td>960.75</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>929.49</td>
<td>949.49</td>
<td>951.03</td>
<td>957.22</td>
</tr>
</tbody>
</table>

Likelihood Ratio Statistics:
Reject the model with the lower DF when p < .05
NOTE: Models must be nested for the statistics to be valid.

<table>
<thead>
<tr>
<th>Model1</th>
<th>Model2</th>
<th>df1</th>
<th>df2</th>
<th>df_chi</th>
<th>Chi_Squared</th>
<th>Prob_Chi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>5.92</td>
<td>0.01497</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>23.15</td>
<td>0.00001</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>25.00</td>
<td>0.00000</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>21.47</td>
<td>0.00002</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>17.23</td>
<td>0.00003</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>19.08</td>
<td>0.00001</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>15.34</td>
<td>0.00000</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
```

These results give numbers for each of the models that you fitted. The middle section give the fit statistics for the models. As in selecting the form for the error covariance matrix, models with lower AIC, AICC, and SBC values are preferred over models with higher values. In the present case, the results suggest that Model 1 is the best model.

The final section of the output gives what is called the likelihood-ratio $\chi^2$ for every pair of models. This section of the macro that creates this output is rather dumb in the sense that it compares every model with every other model even though some of the comparisons may be illogical. The only comparisons that are legitimate are those involving nested models. Two models are nested when every term in the smaller model
is also contained in the larger model. The smaller model is said to be nested within the larger model.

Think of the likelihood ratio $\chi^2$ as a test of significance for those terms in the larger model that are not in the smaller model. For example, consider Model 1 compared to Model 2. Model 2 is nested within Model 1 because: (1) it is the smaller model; and (2) all the terms in Model 2 are contained in Model 1. Hence, it is legitimate to compare these two models. The term in the larger model (Model 1) that is not contained in the smaller model (Model 2) is the three-way interaction, $\text{Group*Mode*Trial}$. Hence, the likelihood ratio $\chi^2$ for this model comparison tests whether this three-way interaction is significant. The value of the $\chi^2$ is 5.92 and with 1 df, it is significant. Hence, the three-way interaction is significant. In terms of model comparison, we reject Model 2 in favor of Model 1.

In comparing models, bear in mind one important principle: *if a model is rejected then all other models nested within that model are also rejected*. Note that Models 3, 4, and 5 are all nested within Model 2. Because we have rejected Model 2, we can also reject Models 3, 4, and 5.

In summary, the likelihood-ratios tests agree with the AIC, AICC, and BIC is selecting Model 1 as the best fitting model.