

Chapter 16

Repeated Measures

Chapter Table of Contents

Introduction	327
Repeated Measures Analysis	329
References	348

Chapter 16

Repeated Measures

Introduction

Repeated measures analysis deals with response outcomes measured on the same experimental unit at different times or under different conditions. Longitudinal data are a common form of repeated measures in which measurements are recorded on individual subjects over a period of time. Blood pressure measured once a week for a month, CD4 counts tracked over a year in an AIDS clinical trial, and per capita demand deposits over years are examples of longitudinal data. Repeated measures can also refer to multiple measurements on an experimental unit, such as the thickness of vertebrae in animals.

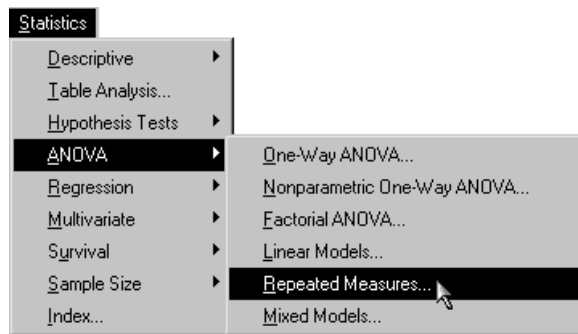


Figure 16.1. Repeated Measures Menu

The experimental units are often subjects. In a repeated measurements analysis, you are usually interested in between-subject and within-subject effects. Between-subject effects are those whose values change only from subject to subject and remain the same for all observations on a single subject, for example, treatment and gender. Within-subject effects are those whose values may differ from mea-

surement to measurement, for example, time. Usually, you are also interested in some between-subject and within-subject interaction, such as treatment by time.

Since measurements on the same experimental unit are likely to be correlated, repeated measurements analysis must account for that correlation. One way of doing this is by modeling the covariance structure of an individual's response. The *compound symmetry* structure assumes the same covariance between any two measurements and the same variance for each measurement. However, sometimes the covariance of measures that are close together in time is higher than the covariance for measurements further apart. In this case, the *first-order autoregressive* covariance structure may be more appropriate. Another possible covariance structure is *unstructured*, in which you estimate different parameters for the variance of each repeated measurement as well as different covariance parameters for each pair of repeated measurements.

The Repeated Measures task enables you to specify a repeated measures model with interactions and nested terms, define subject and repeated effects, and select from a wide range of covariance structures. You can estimate least-squares means for classification effects and output predicted values and residuals to a data set. Plots include means plots, predicted plots, and plots of residuals versus within and between effects. The Repeated Measures task applies methods based on the mixed model with special parametric structures on the covariance matrices.

The example in this chapter demonstrates how you can use the Repeated Measures task in the Analyst Application to analyze repeated measurements data.

Repeated Measures Analysis

The data set analyzed in this task contains data from Littell, Freund, and Spector (1991). Subjects in the study participated in one of three different weightlifting programs, and their strength was measured once every other day for two weeks after they began the program. The first program increased the number of repetitions as the subject became stronger (RI), the second program increased the amount of weight as subjects became stronger (WI), and the subjects in the third program did not participate in weightlifting (CONT). The objective of this analysis is to investigate the effect each weightlifting program has on increasing strength over time. This section also illustrates how to prepare data in univariate form for this task.

Open the Weightsmult Data Set

The data are provided in the Analyst Sample Library. To open the Weightsmult data set, follow these steps:

1. Select **Tools** → **Sample Data** . . .
2. Select Weightsmult.
3. Click **OK** to create the sample data set in your Sasuser directory.
4. Select **File** → **Open By SAS Name** . . .
5. Select Sasuser from the list of **Libraries**.
6. Select Weightsmult from the list of members.
7. Click **OK** to bring the Weightsmult data set into the data table.

Data Management

Figure 16.2 displays the Weightsmult data in multivariate form, which means that a single row in the data table contains all response measurements for a single subject. The **Program** variable defines the treatment group and takes the values 'CONT', 'RI', and 'WI'. The **Subject** variable defines each subject, and the variables

s1 through s7 contain strength measurements across time for each subject.

	Subject	Program	s1	s2	s3	s4	s5	s6	s7
1	1	CONT	85	85	86	85	87	86	87
2	2	CONT	80	79	79	78	78	79	78
3	3	CONT	78	77	77	77	76	76	77
4	4	CONT	84	84	85	84	83	84	85
5	5	CONT	80	81	80	80	79	79	80
6	6	CONT	76	78	77	78	78	77	74
7	7	CONT	79	79	80	79	80	79	81
8	8	CONT	76	76	76	75	75	74	74
9	9	CONT	77	78	78	80	80	81	80
10	10	CONT	79	79	79	79	77	78	79
11	11	CONT	81	81	80	80	80	81	82
12	12	CONT	77	76	77	78	77	77	77
13	13	CONT	82	83	83	83	84	83	83
14	14	CONT	84	84	83	82	81	79	78
15	15	CONT	79	81	81	82	82	82	80
16	16	CONT	79	79	78	77	77	78	78

Figure 16.2. Weightsmult Data

In order for you to perform the repeated measures analysis using the Analyst Application, your data must be in univariate form, which means that each response measurement is contained in a separate row. If your data are not in univariate form, you must create a new data table with this structure. This can be accomplished via the Stack Columns task in the **Data** menu.

The Stack Columns task creates a new table by stacking specified columns into a single column. The values in the other columns are preserved in the new table, and a source column in the new data set contains the names of the columns in the original data set that contained the stacked values.

You want to put the values for columns corresponding to the strength measurement variables `s1` through `s7` in individual rows, so you want to stack columns `s1`–`s7`. To stack the columns, follow these steps:

1. Select **Data** → **Stack Columns** . . .
2. Select `s1` through `s7` and click on the **Stack** button.
3. Type **Strength** in the **Stacked column:** field.
4. Click **OK** to produce the new data set.

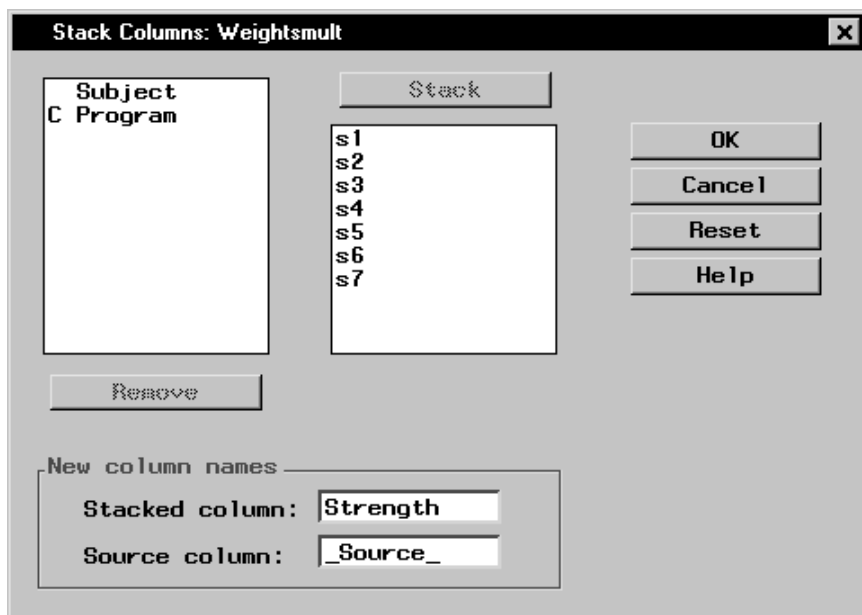


Figure 16.3. Stack Columns Dialog

The new data set is presented in the project tree under the **Stack Columns** folder. The **Weightsmult with Stacked Columns** folder contains the new data set with the **Strength** stacked column, and the **Code** node contains the SAS programming statements that generated the data set.

If a view of the **Weightsmult with Stacked Columns** data is displayed, close it. Then right-click on the data set node labeled

Weightsmult with Stacked Columns, as displayed in Figure 16.4, and select **Open** to bring the new data set into the data table.

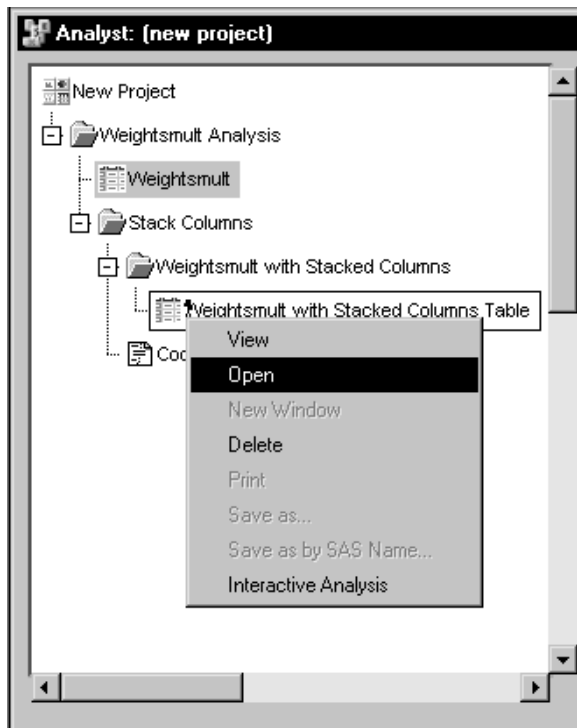


Figure 16.4. Stack Columns: Project Tree

The stacked columns data set contains two new variables. The Strength variable contains the strength measurements, and the `_Source_` variable denotes the measurement times with seven distinct character values: s1, s2, s3, s4, s5, s6, and s7. However, in this analysis, time needs to be numeric. You can create a numeric variable called Time by using the Recode Values facility.

To create the Time variable, follow these steps:

1. Select **Edit** → **Mode** → **Edit**.
2. Select **Data** → **Transform** → **Recode Values ...**
3. Select **_Source_** as the **Column to recode**.
4. Type **Time** in the **New column name:** field.
5. Specify the new column type by selecting **Numeric**.
6. Click **OK** to enter values of the Time variable that correspond to current values of the **_Source_** variable.

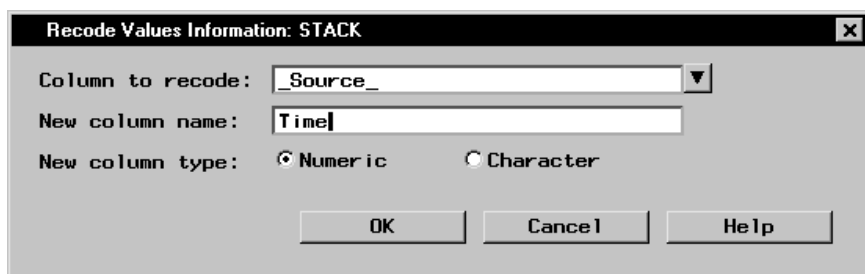


Figure 16.5. Recode Values Information Dialog

7. Type **1** in the **New Value (Numeric)** column cell next to **s1**.
8. Type in the remaining numeric values corresponding to the original values of the **_Source_** column. Figure 16.6 displays the final recoded values.
9. Click **OK** to create the new variable.

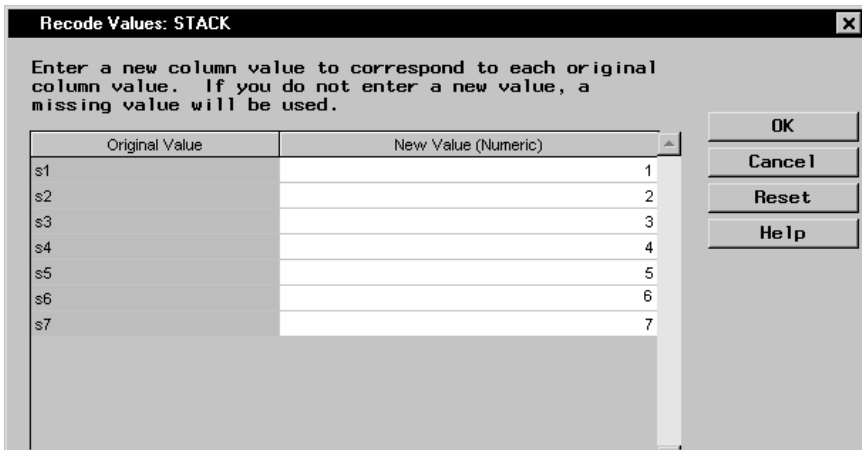


Figure 16.6. Recode Values Dialog

The data set now includes a variable **Time** that contains numeric values for the time of strength measurement. Because the time values are contained in a new variable, you can delete the original variable from the data set by right-clicking on the **_Source_** column in the data table and selecting **Delete**. Once you have deleted the column, the data set should contain four variables, **Subject**, **Program**, **Strength**, and **Time**, as displayed in Figure 16.7.

STACK (Edit)				
	Subject	Program	Strength	Time
1	1	CONT	85	1
2	1	CONT	85	2
3	1	CONT	86	3
4	1	CONT	85	4
5	1	CONT	87	5
6	1	CONT	86	6
7	1	CONT	87	7
8	2	CONT	80	1
9	2	CONT	79	2
10	2	CONT	79	3
11	2	CONT	78	4
12	2	CONT	78	5
13	2	CONT	79	6
14	2	CONT	78	7
15	3	CONT	78	1
16	3	CONT	77	2
17	3	CONT	77	3
18	3	CONT	77	4
19	3	CONT	76	5

Figure 16.7. Weightsuni Data

Before proceeding with the analysis, you can save the new data set as **Weightsuni** by following these steps:

1. Select any cell in the data table or reselect the data set node labeled **Weightsmult** with **Stacked Columns** in the project tree.
2. Select **File** → **Save As By SAS Name . . .**
3. Type **Weightsuni** in the **Member Name** field and click **Save** to save the data set.

Note that the **Weightsuni** data are in univariate form and should be the same as the **Weights** data available in the Analyst Sample Library.

Request the Repeated Measures Analysis

To specify the Repeated Measures task, follow these steps:

1. Select **Statistics** → **ANOVA** → **Repeated Measures** . . .
2. Select **Strength** as the dependent variable.
3. Select **Subject**, **Program**, and **Time** as classification variables.

Figure 16.8 displays the dialog with **Strength** specified as the dependent variable and **Subject**, **Program**, and **Time** specified as classification variables.

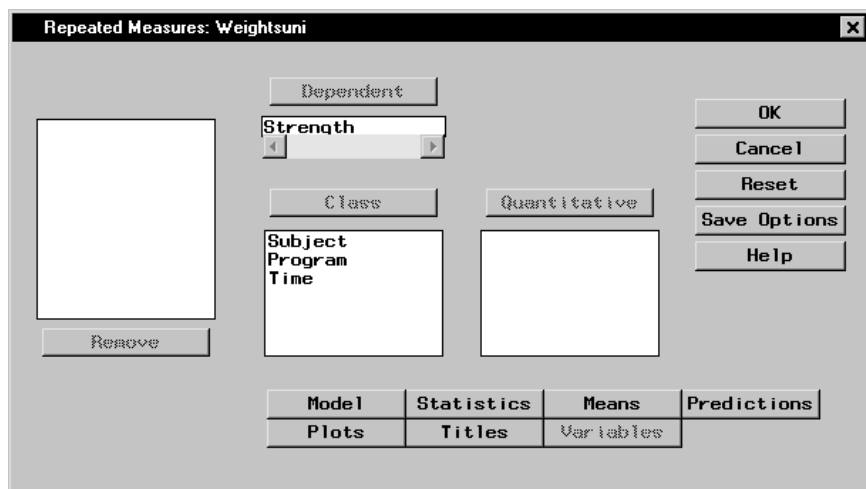


Figure 16.8. Repeated Measures Dialog

Define the Model

To perform a repeated measures analysis, you are required to specify a model, define subjects, specify a repeated effect, and select one or more structures for modeling the covariance of the repeated measurements. By defining a factorial structure between **Program** and **Time**, you can analyze the between-subject effect **Program**, the within-subject effect **Time**, and the interaction between **Program** and **Time**.

Each experimental unit, a subject, needs to be uniquely identified in the **Weightsuni** data set. The value of the **Subject** variable for the

first subject in each separate Program is 1, the value of the Subject variable for the second subject in each Program is 2, and so on. Because subjects participating in different programs have the same value from the Subject variable, you need to nest Subject within Program to uniquely define each subject.

To define the repeated measures model, follow these steps:

1. Click on the **Model** button.
2. Select the **Subjects** tab.
3. Select Subject and click **Add**.
4. Select Program and click **Nest** to nest subjects within weightlifting programs.

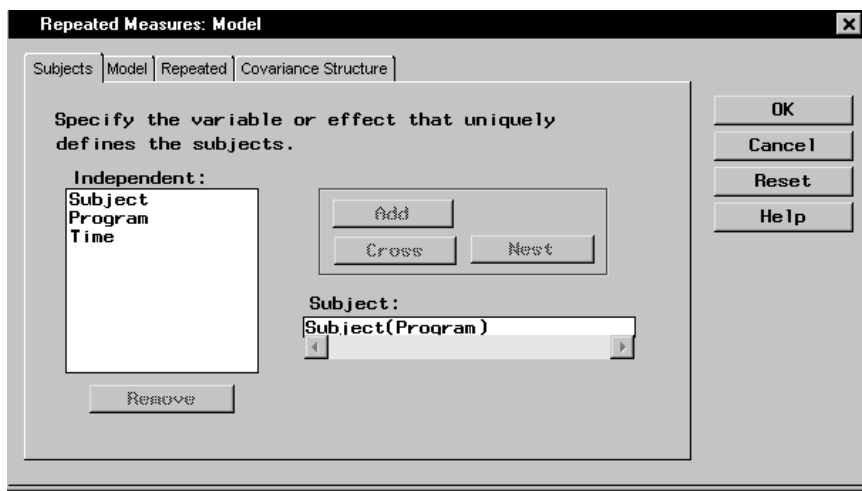


Figure 16.9. Repeated Measures: Model Dialog, Subjects Tab

5. Select the **Model** tab.
6. Select **Program** and **Time** and click **Factorial** to specify a factorial arrangement, which is the main effects for **Program** and **Time** and their interaction.

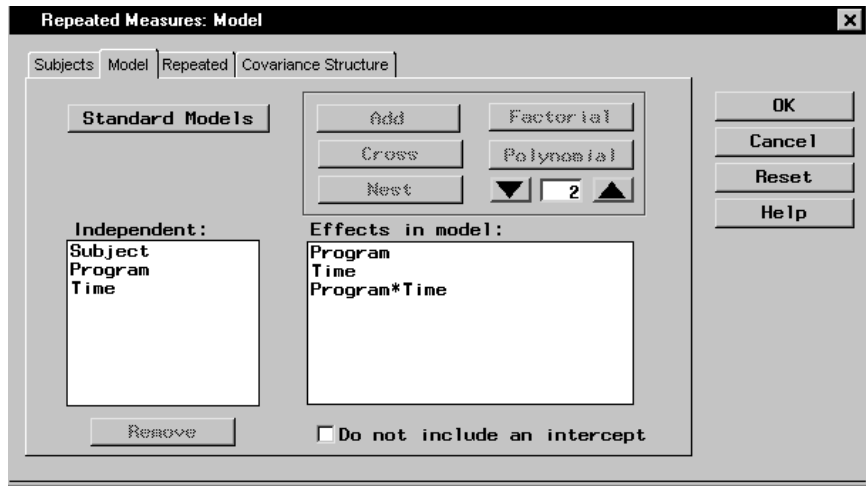


Figure 16.10. Repeated Measures: Model Dialog, Model Tab

7. Select the **Repeated** tab.
8. Select **Time** and click **Add** to specify measurement times as the repeated effect.

This identifies the repeated measurement effect.

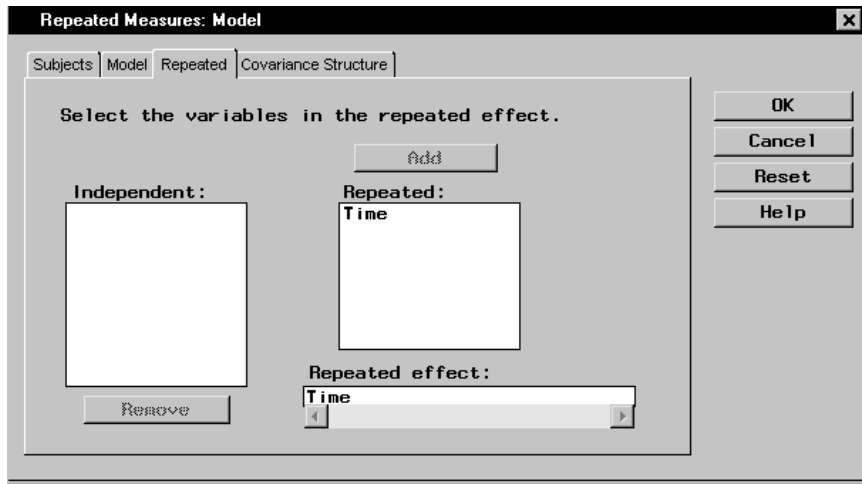


Figure 16.11. Repeated Measures: Model Dialog, Repeated Tab

When analyzing repeated measures data, you must properly model the covariance structure within subjects to ensure that inferences about the mean are valid. Using the Repeated Measures task, you can select from a wide range of covariance types, where the most common types are compound symmetric, first-order autoregressive, and unstructured. To select the covariance structure for the analysis, follow these steps:

1. Select the **Covariance Structure** tab.
2. Select the **Compound symmetry** covariance structure.

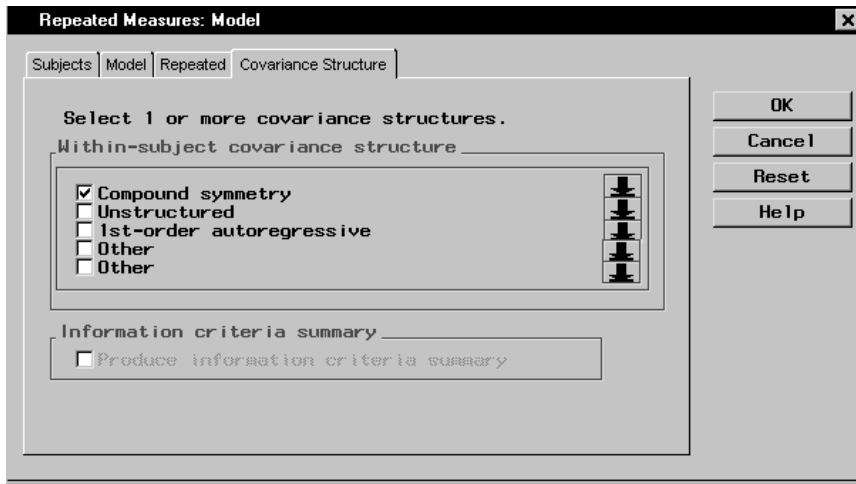


Figure 16.12. Repeated Measures: Model Dialog, Covariance Structure Tab

Close the Model dialog by clicking OK. When you have completed your selections, click **OK** in the main dialog to produce your analysis.

Review the Results

The results are presented in the project tree under the **Repeated Measures ANOVA** folder, as displayed in Figure 16.13. The nodes represent the repeated measures results and the SAS programming statements (labeled **Code**) that generated the output.

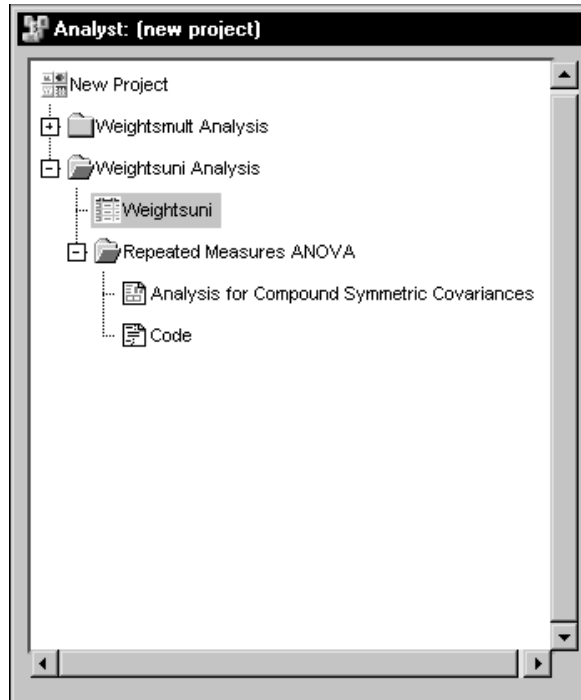


Figure 16.13. Repeated Measures: Project Tree

You can double-click on the **Analysis for Compound Symmetric Covariances** node in the project tree to view the results in a separate window.

The screenshot shows a SAS window titled "Analysis for Compound Symmetric Covariances". It contains two tables: "Class Level Information" and "Dimensions".

Class Level Information	
Class	Levels
Subject	21
Program	3
Time	7

Dimensions	
Covariance Parameters	2
Columns in X	32
Columns in Z	0
Subjects	57
Max Obs Per Subject	7
Observations Used	399
Observations Not Used	0
Total Observations	399

Figure 16.14. Repeated Measures: Model Information

Figure 16.14 displays model information including the levels of each classification variable in the analysis. The **Program** variable has three levels while the **Time** variable has 7 levels. The “Dimensions” table displays information about the model and matrices used in the calculations. There are two covariance parameters estimated using the compound symmetry model: the variance of residual error and the covariance between two observations on the same subject. The 32 columns of the **X** matrix correspond to three columns for the **Program** variable, seven columns for the **Time** variable, 21 columns for the **Program*Time** interaction, and a single column for the intercept. You should always review this information to ensure that the model has been specified correctly.

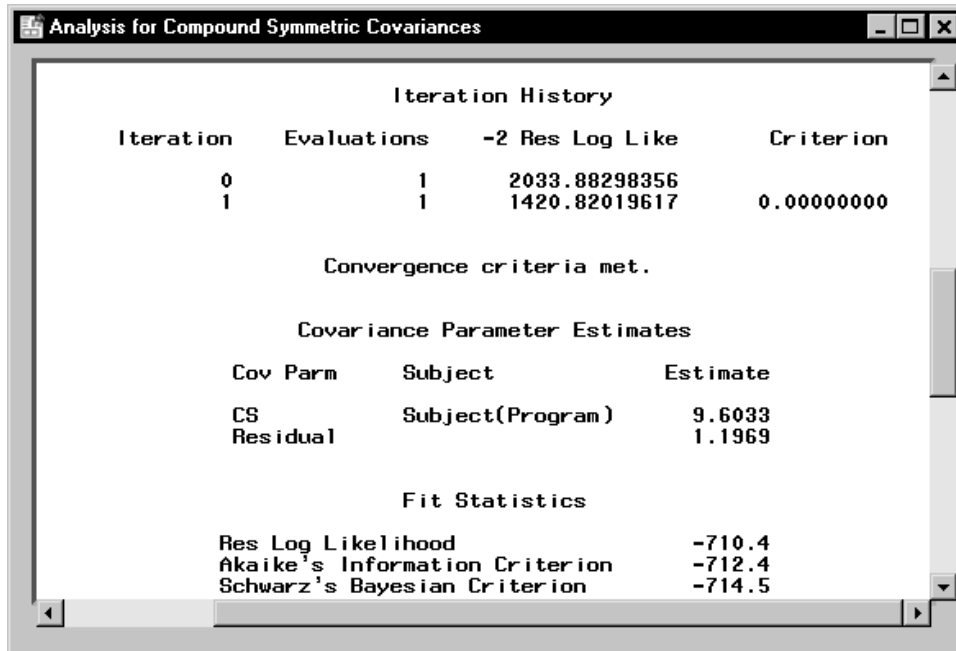
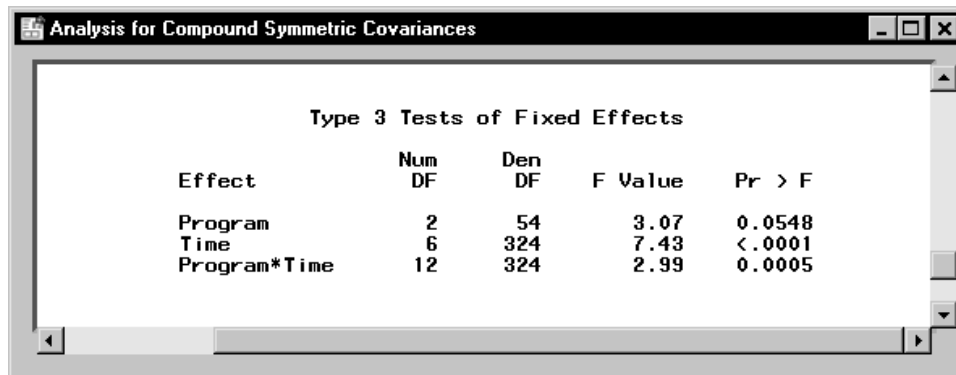


Figure 16.15. Repeated Measures: Fitting Information

Figure 16.15 displays fitting information, including the iteration history, covariance parameter estimates, and likelihood statistics. The “Iteration History” table shows the sequence of evaluations to obtain the restricted maximum likelihood estimates of the variance components.

The “Covariance Parameter Estimates” table displays estimates of the variance component parameters. The covariance between two measurements on the same subject is 9.6. Based on an estimated residual variance parameter of 1.2, the overall variance of a measurement is estimated to be $9.6 + 1.2 = 10.8$.



Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Program	2	54	3.07	0.0548
Time	6	324	7.43	<.0001
Program*Time	12	324	2.99	0.0005

Figure 16.16. Repeated Measures: Tests for Fixed Effects

The “Type 3 Tests of Fixed Effects” table in Figure 16.16 contains hypothesis tests for the significance of each of the fixed effects, that is, those effects you specify on the Model tab. Based on a p -value of 0.0005 for the Program*Time interaction, there is significant evidence of a strong interaction between the weightlifting program and time of measurement at the $\alpha = 0.05$ level of significance.

Exploring Alternative Covariance Structures

Based on the assumption of the compound symmetry covariance structure, any two measurements on the same subject have the same covariance regardless of the length of the time interval between the measurements. However, repeated measurements are often more correlated when the measurements are closer in time than when they are farther apart. In this case, compound symmetry may not be appropriate, and you may want to investigate alternative covariance structures.

The first-order autoregressive covariance structure has the property that observations on the same subject that are closer in time are more highly correlated than measurements at times that are farther apart. The first-order autoregressive covariance can be represented by $\sigma^2 \rho^w$, where w indicates the time between two measurements, ρ stands for the correlation between adjacent observations on the same subject, and σ^2 stands for the variance of an observation. For the first-order autoregressive covariance structure, the correlation between two measurements decreases exponentially as the length of time between the measurements increases.

To fit an additional repeated measures model with a first-order autoregressive covariance structure, follow these steps:

1. Select **Statistics** → **ANOVA** → **Repeated Measures** . . .

Note that the selections for the previous analysis are still specified.

2. Click on the **Model** button.
3. Select the **Covariance Structure** tab.
4. Select the **1st-order autoregressive** structure.
5. Select **Provide information criteria summary** to produce a summary table of model-fit criteria for the two covariance structures.
6. Click **OK** in the main dialog to produce your analysis.

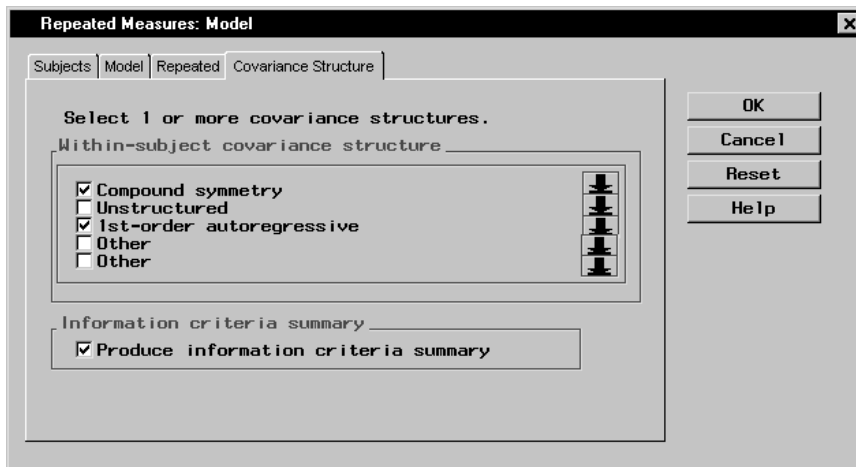


Figure 16.17. Repeated Measures: Model Dialog, Covariance Structure tab

Although this analysis models only two different covariance structures, the Analyst Application provides a wide range of structures to choose from, including unstructured, Huynh-Feldt, Toeplitz, and variance components. To select other structures, click on the down arrow next to an **Other** check box and choose from the resulting drop-down list.

Double-click on the **Analysis for First Order Autoregressive Covariances** node in the project tree to view the results in a separate window.

The screenshot shows a SAS window titled "Analysis for First Order Autoregressive Covariances". The main content is a table titled "Type 3 Tests of Fixed Effects". The table has five columns: Effect, Num DF, Den DF, F Value, and Pr > F. The rows represent the main effects and their interaction.

Effect	Num DF	Den DF	F Value	Pr > F
Program	2	57.1	3.11	0.0524
Time	6	321	4.30	0.0003
Program*Time	12	321	1.17	0.3008

Figure 16.18. Repeated Measures: Test for Fixed Effects for Autoregressive Covariance

Figure 16.18 displays the Type 3 tests for fixed effects based on the first-order autoregressive covariance model. Note that with a p -value greater than 0.30, the Program*Time interaction is not significant at the $\alpha = 0.05$ level of significance. The p -value is different from the p -value of the same test based on the compound symmetry covariance structure, and the two models lead to different conclusions. You can assess the model fit based on different covariance structures by comparing criteria that is provided in the Information Criteria Summary window in Figure 16.19.

The screenshot shows a SAS window titled "Information Criteria Summary". The main content is a table titled "Summary of Information Criteria". The table compares two covariance structures: Compound symmetry and 1st-order autoregressive. The columns are Covariance Structure, Parameters, Akaike's Information Criterion, Schwarz's Bayesian Criterion, and -2 Res Log Likelihood.

Covariance Structure	Parameters	Akaike's Information Criterion	Schwarz's Bayesian Criterion	-2 Res Log Likelihood
Compound symmetry	2	-712.4	-714.5	1420.8
1st-order autoregressive	2	-635.4	-637.4	1266.8

Figure 16.19. Repeated Measures: Information Criteria Summary

The process of selecting the most appropriate covariance structure can be aided by comparing the Akaike's Information Criteria (AIC)

and Schwarz's Bayesian Criterion (SBC) for each model. When you compare models with the same fixed effects but different variance structures, the models with the highest AIC and SBC are deemed the best. In this example, the autoregressive model has higher values for both AIC and SBC, showing considerable improvement over the model with a compound symmetry structure. Based on the information criteria as well as the intuitively sensible property of the correlations being larger for nearby times than for far-apart times, the first-order autoregressive is the more suitable fit for this model.

References

- Littell, R. C., Milliken, G. A., Stroup, W. W., and Wolfinger, R. D. (1996), *SAS System for Mixed Models*, Cary, NC: SAS Institute Inc.
- Littell, R. C., Freund, R. J., and Spector, P. C. (1991), *SAS System for Linear Models, Third Edition*, Cary, NC: SAS Institute Inc.
- SAS Institute Inc. (1999), *SAS/STAT User's Guide, Version 7-1*, Cary, NC: SAS Institute Inc.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *The Analyst Application, First Edition*, Cary, NC: SAS Institute Inc., 1999. 476 pp.

The Analyst Application, First Edition

Copyright © 1999 SAS Institute Inc., Cary, NC, USA.

ISBN 1-58025-446-2

All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, by any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute, Inc.

U.S. Government Restricted Rights Notice. Use, duplication, or disclosure of the software by the government is subject to restrictions as set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, October 1999

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.® indicates USA registration.

IBM®, ACF/VTAM®, AIX®, APPN®, MVS/ESA®, OS/2®, OS/390®, VM/ESA®, and VTAM® are registered trademarks or trademarks of International Business Machines Corporation.® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

The Institute is a private company devoted to the support and further development of its software and related services.