Modelling the Mean

Substantive *versus* nuisance parameters:

- Regression parameters $\beta$ are usually *substantive*;

- Covariance or correlation parameters such as $\Sigma$ are usually *nuisance* parameters.

Parts of the model that contain only nuisance parameters must still be specified correctly; they are just not the primary scientific focus of the study.
Approaches to modelling the mean:

**response profiles:** mean response over time is an arbitrary pattern; “occasion” is used treated as a qualitative factor.

**parametric or semi-parametric curves:** response is assumed to be of a parametric form, such as a linear or quadratic function.

Parametric forms require fewer parameters, but may not fit adequately.
Modelling the Covariance

Three broad approaches:

**unstructured covariance:** $\Sigma$ is an arbitrary positive-definite matrix; makes sense only in a balanced study; uses $n \times (n + 1)/2$ parameters;

**covariance pattern:** borrows ideas from time series, with correlation decaying as time separation increases; may make sense in unbalanced studies; uses as few as 2 parameters.

**random effects:** subject-specific random effect introduces positive correlation for a given subject; may make sense in unbalanced studies; uses as few as 2 parameters.
Historical Approaches: Repeated Measures Analysis by Univariate ANOVA

Assumes a random subject effect:

\[ Y_{i,j} = X'_{i,j} \beta + b_i + \epsilon_{i,j}, \]

where:

- \( b_s \) and \( \epsilon_s \) are uncorrelated, mean zero, random variables;

- \( \text{Var}(b_i) = \sigma_b^2, \text{Var}(\epsilon_{i,j}) = \sigma_\epsilon^2 \).
Covariance structure is *compound symmetry*:

\[ \text{Cov}(Y_i) = \begin{pmatrix}
\sigma_b^2 + \sigma^2_e & \sigma_b^2 & \ldots & \sigma_b^2 \\
\sigma_b^2 & \sigma_b^2 + \sigma^2_e & \ldots & \sigma_b^2 \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_b^2 & \sigma_b^2 & \ldots & \sigma_b^2 + \sigma^2_e
\end{pmatrix} \]

- all variances are \( \sigma_b^2 + \sigma^2_e \);
- all covariances are \( \sigma_b^2 \);
- all correlations are \( \frac{\sigma_b^2}{\sigma_b^2 + \sigma^2_e} \).
Historical Approaches: Repeated Measures Analysis by Multivariate ANOVA

Assumes an unstructured covariance matrix:

$$\text{Cov}(Y_i) = \Sigma$$

for arbitrary positive definite $\Sigma$.

Hypotheses about within-subject effects are tested with multivariate criteria such as Hotelling’s $T^2$, Wilks’s $\Lambda$, etc.
Repeated measures analysis using SAS’s proc glm:

```sas
title 'Repeated Measures Analysis of TLC data';
options linesize = 80 pagesize = 21 nodate;

data tlc;
  infile 'tlc.txt' firstobs = 31;
  input child treatment $ y0 y1 y4 y6;
run;

proc glm data = tlc;
  class treatment;
  model y0 -- y6 = treatment / nouni;
  repeated occasion / printe;
run;
```
SAS output:

- Pages 1 and 2 summarize the class variable and the factor occasion created by the repeated statement.

- Pages 3–6 are printed in response to the printe option; the correlations on page 3 shed light on the compound symmetry assumption, and Mauchly’s criterion (for orthogonal components) tests it as a null hypothesis.

- Pages 7 and 8 give multivariate tests for occasion and occasion * treatment, which are valid for any $\Sigma$. 
• Page 9 gives the test for the main effect of treatment, which is the same in both the univariate and multivariate approaches.

• Pages 10 and 11 give univariate tests for occasion and occasion * treatment, which depend on compound symmetry, and adjusted $P$-values that allow for departures from compound symmetry.
Repeated Measures Analysis of TLC data

The GLM Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment</td>
<td>2</td>
<td>A P</td>
</tr>
</tbody>
</table>

Number of observations 100
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance

Repeated Measures Level Information

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>y0</th>
<th>y1</th>
<th>y4</th>
<th>y6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of occasion</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance

Partial Correlation Coefficients from the Error SSCP Matrix / Prob > |r|

<table>
<thead>
<tr>
<th></th>
<th>DF = 98</th>
<th>y0</th>
<th>y1</th>
<th>y4</th>
<th>y6</th>
</tr>
</thead>
<tbody>
<tr>
<td>y0</td>
<td>1.000000</td>
<td>0.571288</td>
<td>0.569830</td>
<td>0.577199</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>y1</td>
<td>0.571288</td>
<td>1.000000</td>
<td>0.775255</td>
<td>0.581872</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;.0001</td>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>y4</td>
<td>0.569830</td>
<td>0.775255</td>
<td>1.000000</td>
<td>0.580880</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>y6</td>
<td>0.577199</td>
<td>0.581872</td>
<td>0.580880</td>
<td>1.000000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance

\[ E = \text{Error SSCP Matrix} \]

occasion\_N represents the contrast between the \text{n}th level of occasion and the last

<table>
<thead>
<tr>
<th></th>
<th>occasion_1</th>
<th>occasion_2</th>
<th>occasion_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>occasion_1</td>
<td>3868.4</td>
<td>2536.4</td>
<td>2501.8</td>
</tr>
<tr>
<td>occasion_2</td>
<td>2536.4</td>
<td>4277.4</td>
<td>3321.3</td>
</tr>
<tr>
<td>occasion_3</td>
<td>2501.8</td>
<td>3321.3</td>
<td>4389.2</td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance

Partial Correlation Coefficients from the Error SSCP Matrix of the Variables Defined by the Specified Transformation / Prob > |r|

<table>
<thead>
<tr>
<th>DF = 98</th>
<th>occasion_1</th>
<th>occasion_2</th>
<th>occasion_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>occasion_1</td>
<td>1.000000</td>
<td>0.623543</td>
<td>0.607143</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>occasion_2</td>
<td>0.623543</td>
<td>1.000000</td>
<td>0.766521</td>
</tr>
<tr>
<td></td>
<td>&lt;.0001</td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>occasion_3</td>
<td>0.607143</td>
<td>0.766521</td>
<td>1.000000</td>
</tr>
<tr>
<td></td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance

Sphericity Tests

<table>
<thead>
<tr>
<th>Variables</th>
<th>DF</th>
<th>Criterion</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformed Variates</td>
<td>5</td>
<td>0.2343616</td>
<td>140.33332</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Orthogonal Components</td>
<td>5</td>
<td>0.8106112</td>
<td>20.308452</td>
<td>0.0011</td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance

Manova Test Criteria and Exact F Statistics
for the Hypothesis of no occasion Effect
H = Type III SSCP Matrix for occasion
E = Error SSCP Matrix

S=1       M=0.5       N=47

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.34692535</td>
<td>60.24</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.65307465</td>
<td>60.24</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>1.88246445</td>
<td>60.24</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Roy's Greatest Root</td>
<td>1.88246445</td>
<td>60.24</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance

Manova Test Criteria and Exact F Statistics for the Hypothesis of no occasion*treatment Effect

\[ H = \text{Type III SSCP Matrix for occasion*treatment} \]
\[ E = \text{Error SSCP Matrix} \]

\[ S=1 \quad M=0.5 \quad N=47 \]

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.47621885</td>
<td>35.20</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.52378115</td>
<td>35.20</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>1.09987489</td>
<td>35.20</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Roy's Greatest Root</td>
<td>1.09987489</td>
<td>35.20</td>
<td>3</td>
<td>96</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment</td>
<td>1</td>
<td>3110.85063</td>
<td>3110.85063</td>
<td>25.43</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>98</td>
<td>11987.33065</td>
<td>122.31970</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>occasion</td>
<td>3</td>
<td>3272.773475</td>
<td>1090.924492</td>
<td>61.43</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>occasion*treatment</td>
<td>3</td>
<td>2030.385875</td>
<td>676.795292</td>
<td>38.11</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error(occasion)</td>
<td>294</td>
<td>5221.503150</td>
<td>17.760215</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>G</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adj Pr &gt; F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>occasion</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>occasion*treatment</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error(occasion)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>
Repeated Measures Analysis of TLC data

The GLM Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Greenhouse-Geisser Epsilon 0.8832
Huynh-Feldt Epsilon 0.9192