Repeated Measures ANOVA/GLM
Multivariate ANOVA/GLM
in PROC MIXED

Multivariate Methods in Education
ERSH 8350
Lecture #8 – October 5, 2011
Today’s Class

• Using PROC MIXED for:
  - Repeated measures versions of linear models
  - Multivariate ANOVA

• General Linear Mixed Models

• Additional MIXED MODELS topics
  - Types of estimators: ML versus REML
Example Data

- A health researcher is interested in examining the impact of dietary habits and exercise on pulse rate
- A sample of 18 participants is collected
  - Diet factor (BETWEEN SUBJECTS):
    - Nine are vegetarians
    - Nine are omnivores
  - Exercise factor (BETWEEN SUBJECTS) with random assignment:
    - Aerobic stair climbing
    - Racquetball
    - Weight training
  - Three pulse rates (WITHIN SUBJECTS):
    - After warm-up
    - After jogging
    - After running
The Data

- A first step in every analysis is to inspect the data
  - Graphical plots help

- The goal of the inspection is to make sure data are entered correctly
  - Right now we will not be concerned about checking assumptions
    - Assumptions depend on the choice of statistical model
  - We will never clean our data
    - But we will make sure data are entered correctly
Plots

Pulse 1

Pulse 2

Pulse 3

Squared Mahalanobis Distance
MULTIVARIATE DATA ANALYSIS:
INITIAL CONSIDERATIONS
An Initial Look at Our Data

- The data for our analysis present several factors which we are interested in investigating:
  - Within Subjects:
    - Effects of time of measurement on pulse rate
  - Between Subjects:
    - Main effect of diet on pulse rate
    - Main effect of exercise type on pulse rate
    - Interaction of diet and exercise type on pulse rate (2-way)
  - Within and Between Interactions
    - Interaction of measurement and diet (2-way)
    - Interaction of measurement and exercise type (2-way)
    - Interaction of measurement, diet, and exercise type (3-way)
Understanding Data for Mixed Models

• The data setup for mixed models is that of a set of dependent variables nested within an observation:
  ➢ For our example, we had three pulse rate measurements (within subjects factor) here displayed in wide-format
    \[ y_i = [y_{i,\text{pulse}_1} \quad y_{i,\text{pulse}_2} \quad y_{i,\text{pulse}_3}] \]

• We will consider data that are stacked:
  \[ y_i = \begin{bmatrix} y_{i,\text{pulse}_1} \\ y_{i,\text{pulse}_2} \\ y_{i,\text{pulse}_3} \end{bmatrix} \]

• In the mixed model, the stacked data observation (here, pulse) is treated as being multivariate normal
Stacking Our Data

- Our first two observations (wide format):

<table>
<thead>
<tr>
<th>EXERTYPE</th>
<th>PULSE1</th>
<th>PULSE2</th>
<th>PULSE3</th>
<th>DIET</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>112</td>
<td>166</td>
<td>215</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>111</td>
<td>166</td>
<td>225</td>
<td>1</td>
</tr>
</tbody>
</table>

- Our first two observations (long/stacked format):

<table>
<thead>
<tr>
<th>EXERTYPE</th>
<th>DIET</th>
<th>ID</th>
<th>pulse</th>
<th>intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>112</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>166</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>215</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>111</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>166</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>225</td>
<td>3</td>
</tr>
</tbody>
</table>
Linear Mixed Models with Matrices

- The linear model for a single observation $i$:
  - $p$ measured outcome variables
  - $k$ predictors
  \[ y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + e_i \]

- The equation above can be expressed more compactly by a set of matrices
  \[ y_i = X_i \beta + e_i \]

- $y$ is of size $(p \times 1)$
- $X$ is of size $(p \times (1 + k))$
- $\beta$ is of size $((1+k) \times 1)$ – called fixed effects
  - Next week we will add to this – random components
    - The mixed model is “mixed” because it has fixed and random effects
- $e$ is of size $(p \times 1)$
Unpacking the Equation

For the first measurement of the first observation:

\[ y_{11} = \beta_0 + \beta_1 X_{11} + \cdots + \beta_k X_{1k} + e_{11} \]
The Mixed Model for Our Data

• To put the mixed model into context, let’s look at our data
  ➢ Here we will use a dummy-coded set of indicators for which pulse measure is in the data
    • Three measures makes two dummy coded columns in $X$: one for pulse #1 and one for pulse #2 (pulse #3 is now the reference)
      - The intercept will be the mean of pulse #3
      - The first regression coefficient will be the difference between pulse #3 and pulse #1
      - The second regression coefficient will be the difference between pulse #3 and pulse #2

• The first observation:

\[
\begin{bmatrix}
112 \\
166 \\
215 \\
\end{bmatrix}
= \begin{bmatrix}
1 & 1 & 0 \\
1 & 0 & 1 \\
1 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
\beta_0 \\
\beta_1 \\
\beta_2 \\
\end{bmatrix} + \begin{bmatrix}
e_{11} \\
e_{12} \\
e_{13} \\
\end{bmatrix}
\]

\[y_1 \quad X_1 \quad \beta \quad e_1\]

\[\begin{array}{c}
(3 \times 1) \\
(3 \times (1 + 2)) \\
((1 + 2) \times 1) \\
\end{array}\]

\[\begin{array}{c}
(3 \times 1) \\
\end{array}\]
Mixed Model Assumptions

• The mixed model assumes all measurements from a subject follow a multivariate normal distribution:

\[ f(\mathbf{y}_i | \mathbf{X}_i) \sim N_p(\mathbf{X}_i \beta, \mathbf{V}_i) \]

Additionally, there are assumptions about error

\[ \mathbf{e}_i \sim N_p(\mathbf{0}, \mathbf{R}_i) \]

• Today, without random effects, we take \( \mathbf{V}_i = \mathbf{R}_i \)
  - Next week \( \mathbf{V}_i \) will be a function of random effect variances and \( \mathbf{R}_i \)

• The subscript on each \( \mathbf{V}_i \) and \( \mathbf{R}_i \) indicates that each subject can have a different covariance structure
  - We saw this for subjects missing data last week
Two Sides of a Mixed Model

• MODEL FOR THE MEANS (FIXED EFFECTS):
  - The model for the means comes from the fixed effects
  - IVs $X_i$ and linear model weights $\beta$ provide predicted values for each observation

• MODEL FOR THE VARIANCES (RESIDUAL VARIANCES AND RANDOM EFFECTS):
  - Because we have repeated observations, the model for the variances is now a covariance matrix of size $p \times p$
    - Diagonal elements: variance of error for each outcome
    - Off-diagonal elements: covariance of errors for pairs of outcomes
An Initial Mixed Model

• As a guide to understanding mixed models, we will estimate a model that was demonstrated on slide 12
  ➢ Dummy coded variable for pulse measurement

```plaintext
*EMPTY MODEL: unstructured R covariance matrix;
PROC MIXED DATA=WORK.dietstack METHOD=ML COVTEST NOPROFILE ITDETAILS IC;
CLASS intensity;
MODEL pulse = intensity / S;
REPEATED / SUBJECT=ID TYPE=UN R RCORR;
LSMEANS intensity / ADJUST=BON;
RUN;
```

• In this model we are:
  ➢ Using maximum likelihood
    • Biased estimates of $\mathbf{R}$ covariance matrix (but efficient)
  ➢ Estimating an unstructured $\mathbf{R}$ covariance matrix
    • Model for the variances will produce the ML estimated $\mathbf{R}$ covariance matrix
  ➢ Estimating a very simplistic model for the means
    • Replicates the within-subjects factor
    • Hypothesis tests for intensity will test whether mean pulse after each type of exercise intensity are equal
Initial Model Results

\[ \hat{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 189.56 \\ -102.06 \\ -55.44 \end{bmatrix} \]

- Mean for Pulse #3 = \( \hat{\beta}_0 = 189.56 \)
- Mean for Pulse #1 = \( \hat{\beta}_0 + \hat{\beta}_1 = 189.56 - 102.06 = 87.50 \)
- Mean for Pulse #2 = \( \hat{\beta}_0 + \hat{\beta}_2 = 189.56 - 55.44 = 134.11 \)
Putting Results into Equations

• From our results, the predictions for the first observation:

\[
\begin{bmatrix}
  87.50 \\
  134.11 \\
  189.56 \\
\end{bmatrix}
= 
\begin{bmatrix}
  1 & 1 & 0 \\
  1 & 0 & 1 \\
  1 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
  189.56 \\
  -102.06 \\
  -55.44 \\
\end{bmatrix}
\]

\[
\hat{y}_1 = \mathbf{X}_1 \hat{\beta}
\]

• From this, we can see how the dummy coded independent variables \( \mathbf{X} \) makes predictions about each of the measurements in a stacked variable

  ➢ Each subject has the same prediction for their pulse values
    ✷ As in ANOVA
Moving on to Error Covariance Matrix

- The model for the means is where most inferences are made in mixed models

- However, the key to the inferences in mixed models is the error covariance matrix $\mathbf{R}$ structure
  - Wrong covariance matrix structure = inaccurate standard errors for fixed effects = inaccurate p-values for fixed effects

- We fit an unstructured matrix (analogous to MANOVA)
Three Structures of Classical GLM

- Three structures:
  1. Independence (TYPE=VC): \( R = \sigma_e^2 I_p \)
     - Models if all observations as if they came from separate people
     - No more statistical parameters than original GLM approach
     - Don’t use: shown for baseline purposes
  2. Repeated Measures (TYPE=CS): Assumes sphericity of observations
     - Sphericity is a condition that is more strictly enforced by compound symmetry of \( R \) having two parameters:
     - Sphericity is compound symmetry of pairwise differences
       - Diagonal elements: same variance
       - Off-diagonal elements: same covariance
     - No sphericity? Adjustments to F tests
     - In modern methods this can even be more flexible (heterogeneous variances)
  3. Multivariate ANOVA/GLM (TYPE=UN): Assumes nothing – estimates everything
     - Every element in \( R \) is modeled
     - Need more power (i.e., sample size) to make work well
     - Most general procedure
Types of Error
Covariance Matrix Structures

• In SAS PROC MIXED, many types of error covariance matrix structures are estimable (type=VC):
  - Some are more common than others
  - Can use deviance tests to determine which is most appropriate
  - Have to run multiple models to determine which is best
    - Here, I run a model for each

• Variance Components – Default (TYPE=VC):
  \[ R = \sigma^2 \mathbf{I} = \begin{bmatrix} \sigma^2_e & 0 & 0 \\ 0 & \sigma^2_e & 0 \\ 0 & 0 & \sigma^2_e \end{bmatrix} \]
  - Assumes errors of all variables:
    - Have same variance
    - Are independent
Variance Components Structure Result

- The variance components structure result:

  **Covariance Parameter Estimates**

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr &gt; Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual</td>
<td>ID</td>
<td>479.46</td>
<td>92.2717</td>
<td>5.20</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

  **Estimated R Matrix for Subject 1**

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>479.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>479.46</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>479.46</td>
</tr>
</tbody>
</table>

  **Information Criteria**

<table>
<thead>
<tr>
<th>Neg2LogLike</th>
<th>Parms</th>
<th>AIC</th>
<th>AICC</th>
<th>HQIC</th>
<th>BIC</th>
<th>CAIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>486.6</td>
<td>4</td>
<td>494.6</td>
<td>495.4</td>
<td>495.1</td>
<td>498.1</td>
<td>502.1</td>
</tr>
</tbody>
</table>

  **Solution for Fixed Effects**

<table>
<thead>
<tr>
<th>Effect</th>
<th>intensity</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>189.56</td>
<td>5.1611</td>
<td>17</td>
<td>36.73</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intensity</td>
<td>2</td>
<td>-102.06</td>
<td>7.2988</td>
<td>34</td>
<td>-13.98</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intensity</td>
<td>3</td>
<td>-55.4444</td>
<td>7.2988</td>
<td>34</td>
<td>-7.60</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

  **Note:** the standard errors changed for the fixed effects

  We can use Neg2LogLike in our Deviance Test to see if a different model is preferred statistically.
Structure #2: Compound Symmetry

- Compound Symmetry – (TYPE=CS):
  \[
  R = \begin{bmatrix}
  \sigma_e^2 + \sigma_{cs} & \sigma_{cs} & \sigma_{cs} \\
  \sigma_{cs} & \sigma_e^2 + \sigma_{cs} & \sigma_{cs} \\
  \sigma_{cs} & \sigma_{cs} & \sigma_e^2 + \sigma_{cs}
  \end{bmatrix}
  \]

- Assumes errors of all variables:
  - Have same variance
  - Are correlated (covariance is \( \sigma_{cs} \))
Deviance Test for Compound Symmetry

• The variance components structure is nested within the compound symmetric structure (set $\sigma_{cs} = 0$)
  ➢ Deviance test is possible
    ◦ $H_0$: $\sigma_{cs} = 0$
    ◦ $H_A$: $\sigma_{cs} \neq 0$
  ➢ Do not use Wald test or information criteria
    ◦ Deviance test is most accurate when possible

• -2*LogL from VC: 486.6
• -2*LogL from CS: 435.3
• Deviance test statistic = 486.6 - 435.3 = 51.3
  ➢ Chi-square distributed with 1 df (for difference in parameters)
• P-value: < 0.0001
  ➢ Conclusion: CS model fits better than VC model
  ➢ Next step: more structures
**Structure #3: Heterogeneous CS**

- Another useful structure is heterogeneous compound symmetry – (TYPE=CSH):

\[
R = \begin{bmatrix}
\sigma_{e_1}^2 & \sigma_{e_1} \sigma_{e_2} \rho & \sigma_{e_1} \sigma_{e_3} \rho \\
\sigma_{e_1} \sigma_{e_2} \rho & \sigma_{e_2}^2 & \sigma_{e_2} \sigma_{e_3} \rho \\
\sigma_{e_1} \sigma_{e_3} \rho & \sigma_{e_2} \sigma_{e_3} \rho & \sigma_{e_3}^2
\end{bmatrix}
\]

- Assumes errors of all variables:
  - Have different variances \((\sigma_{e_1}^2, \sigma_{e_2}^2, \sigma_{e_3}^2)\)
  - Have same correlation (\(\rho\))

---

### Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var(1)</td>
<td>ID</td>
<td>271.16</td>
<td>90.5831</td>
<td>2.99</td>
<td>0.0014</td>
</tr>
<tr>
<td>Var(2)</td>
<td>ID</td>
<td>433.06</td>
<td>143.72</td>
<td>3.01</td>
<td>0.0013</td>
</tr>
<tr>
<td>Var(3)</td>
<td>ID</td>
<td>732.06</td>
<td>244.01</td>
<td>3.00</td>
<td>0.0013</td>
</tr>
<tr>
<td>CSH</td>
<td>ID</td>
<td>0.9054</td>
<td>0.03619</td>
<td>25.02</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

### Estimated R Matrix for Subject 1

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col12</th>
<th>Col13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>271.16</td>
<td>310.26</td>
<td>403.39</td>
</tr>
<tr>
<td>2</td>
<td>310.26</td>
<td>433.06</td>
<td>509.79</td>
</tr>
<tr>
<td>3</td>
<td>403.39</td>
<td>509.79</td>
<td>732.06</td>
</tr>
</tbody>
</table>

### Information Criteria

- Neg2LogLike: 415.8
-Parms: 7
- AIC: 429.8
- AICC: 432.2
- HQIC: 430.7
- BIC: 436.0
- CAIC: 443.0
Deviance Test for Heterogeneous Compound Symmetry

- The CS structure is nested within the CSH structure
  - Deviance test is possible
  - Do not use Wald test or information criteria
    - Deviance test is most accurate when possible

- \(-2\)LogL from CS: 435.3
- \(-2\)LogL from CSH: 415.8

- Deviance test statistic = 435.3 - 415.8 = 19.5
  - Chi-square distributed with 2 df (for difference in parameters)

- P-value: < 0.0001 (from Excel “=chidist(19.5,2)”)  
  - Conclusion: CSH model fits better than CS model
  - Next step: more structures
Structure #4: Unstructured

- Finally, we have the unstructured structure (TYPE=UN):

\[
R = \begin{bmatrix}
\sigma_{e1}^2 & \sigma_{e1} \sigma_{e2} & \sigma_{e1} \sigma_{e3} \\
\sigma_{e1} \sigma_{e2} & \sigma_{e2}^2 & \sigma_{e2} \sigma_{e3} \\
\sigma_{e1} \sigma_{e3} & \sigma_{e2} \sigma_{e3} & \sigma_{e3}^2
\end{bmatrix}
\]

- Makes no assumptions about structure
- Needs most parameters – but can be best
  - If sample size was not an issue would be best choice
Deviance Test for UN v. CSH

• The CSH structure is nested within the UN structure
  ➢ Deviance test is possible
  ➢ Do not use Wald test or information criteria
    • Deviance test is most accurate when possible

• \(-2\times\text{LogL from CSH}: 415.8\)
• \(-2\times\text{LogL from UN}: 408.1\)

• Deviance test statistic = 415.8-408.1 = 7.7
  ➢ Chi-square distributed with 2 df (for difference in parameters)

• P-value: = 0.0212 (from Excel "=chidist(7.7,2)"")
  ➢ Conclusion: UN model fits better than CSH
    • Therefore, we use UN model – UP NEXT: EVALUATING FIXED EFFECTS
The Hypothesis Test for the Means

- Inclusion of the dummy coded variable brings about the following hypothesis test:

  ![Type 3 Tests of Fixed Effects table]

  - The null hypothesis is:
    \[ H_0: \text{Pulse}_1 = \text{Pulse}_2 = \text{Pulse}_3 \]
  - The alternative hypothesis is:
    \[ H_A: \text{at least one mean not equal} \]
  - The p-value of the hypothesis test is small, so we reject the null hypothesis

    - We can conclude that at least one pulse rate mean is not equal to the other two

    - This is our within-subjects effect
Further Inspecting Means

• Following up the hypothesis test, we may wish to investigate which means are significantly different
  ➢ From the LSMEANS statement in the syntax
    • LSMEANS are predicted population **marginal** means of an effect
      – Marginal = averaged over all other effects (here we have none)
      – Population = means assuming balanced data (equal sample size)

| Effect | intensity | Estimate | Standard Error | DF  | t Value | Pr > |t| |
|--------|-----------|----------|----------------|-----|---------|------|---|
| intensity | 1 | 87.5000 | 3.8323 | 17 | 22.83 | <.0001 |
| intensity | 2 | 134.11 | 4.9020 | 17 | 26.92 | <.0001 |
| intensity | 3 | 189.56 | 6.3563 | 17 | 29.82 | <.0001 |

| Effect | intensity | _intensity | Estimate | Standard Error | DF  | t Value | Pr > |t| | Adjustment | Adj P |
|--------|-----------|------------|----------|----------------|-----|---------|------|---|----------------|------|
| intensity | 1 | 2 | -46.6111 | 2.1230 | 17 | -21.96 | <.0001 | Bonferroni | <.0001 |
| intensity | 1 | 3 | -102.06 | 3.6830 | 17 | -27.71 | <.0001 | Bonferroni | <.0001 |
| intensity | 2 | 3 | -55.4444 | 2.2976 | 17 | -24.13 | <.0001 | Bonferroni | <.0001 |

• From the output, we can conclude that each mean pulse rate is significantly different from each other
  ➢ Using Bonferroni correction
More on Covariance Structures

• Depending on your type of analysis, you will choose from a handful of structures
  ➢ Our RM ANOVA example really only gave us 4 options

• SAS PROC MIXED has a lot of options
  ➢ Auto-Regressive (typically for longitudinal data)
  ➢ Factor analytic (typically for scales or similar measurements)
  ➢ Toeplitz (longitudinal data)
  ➢ Spatial (for data from subjects where proximity is an issue)
  ➢ And more...
MIXED MODEL ESTIMATION
Mixed Model Estimation

• Mixed models use maximum likelihood as their estimation method
  - Full information techniques – measurements modeled directly

• There are two main types of estimation options in mixed models:
  - Maximum likelihood using Multivariate Normal (ML)
    - Produces **biased** estimates of error covariance matrix
      - Easy to compare models with different fixed effects (deviance tests)
      - Is better for large samples
  - Residual maximum likelihood (for residuals: REML)
    - Produces **unbiased** estimates of error covariance matrix
      - Cannot compare models with different fixed effects
      - Is better for small samples

• We will discuss how these work – then show an example with REML (which would by our choice for the 18 subjects in our sample)
  - Choice leads to different SEs for fixed effects, and different p-values
ML Estimation of Mixed Models

• The goal in ML estimation is to pick a set of parameters that maximize the likelihood function
  ➢ Typically the log-likelihood is used
  ➢ Here, we have to know $\beta, R$

• The log-likelihood function is the log of the model-assumed MVN (for a subject $i$):
  $$N_p(X_i \beta, V_i = R_i)$$
Fixed Effects Fall Out

• Because of the wonders of math, we can use a technique called *estimated generalized least squares*
  
  ➢ Use iterative algorithm to find $R$: $\hat{R}$
  
  ➢ Given $\hat{R}$, we can find $\hat{\beta}$
  
  ➢ Here, we will define $\hat{V} = \hat{R}$ (for consistency with SAS documentation – and to generalize to random effects)

• Specifically, across all subjects:

$$\hat{\beta} = (X^T\hat{V}X)^{-1}X^T\hat{V}^{-1}y$$
The ML Log Likelihood

• The goal is to pick the parameters in $\hat{V} = \hat{R}$ and then substitute them into the log likelihood function, producing a log likelihood value
  
  ➢ SAS uses Newton-Raphson

• From the Multivariate Normal log likelihood for a subject:

$$ l(R_i) = -\frac{1}{2} \log |\hat{V}_i| - \frac{1}{2} r_i^T \hat{V}_i^{-1} r_i - \frac{p}{2} \log(2\pi) $$

Where $r$ is the vector of residuals:

$$ r_i = y_i - x_i \hat{\beta} = y_i - x_i (x_i^T \hat{V}_i x_i)^{-1} x_i^T \hat{V}_i^{-1} y_i $$
Issues with ML Estimates

• ML estimation is a common choice and performs well when sample sizes are large

• However, estimates of the variances will be biased
  ➢ Similar to basic statistics phenomena of using N versus N-1 in the variance/standard deviation

• Therefore, the residual ML estimator was developed
  ➢ Called REML
REML Estimator

• The REML estimator maximizes the likelihood of the residuals

• The likelihood function comes from stating the likelihood of the data as a function of the likelihood of the estimated fixed effects and the residuals

• Here, we take the estimated residuals to be

\[ \hat{e} = y - X\hat{\beta} \]

• Where

\[ \hat{\beta} = (X^T\hat{V}^{-1}X)^{-1}X^T\hat{V}^{-1}y \]
Deriving REML

• Because \( \mathbf{y} \) is multivariate normal, \( \hat{\mathbf{\beta}} \) and \( \hat{\mathbf{e}} \) are linear functions of \( \mathbf{y} \) that are:
  - Normally distributed (see properties of MVN)
  - Independent

• Therefore, with independence we can re-express the likelihood of \( \mathbf{y} \) as a product of \( \hat{\mathbf{\beta}} \) and \( \hat{\mathbf{e}} \)

\[
L(\mathbf{y}|\mathbf{V}) = L(\hat{\mathbf{\beta}}|\mathbf{V}) \times L(\hat{\mathbf{e}}|\mathbf{V})
\]
More Deriving REML

- Further, due to the consistency of the estimates under maximum likelihood, we know that
  \[
  \hat{\beta} \sim N(\beta, (X^T\hat{V}^{-1}X)^{-1})
  \]

- Therefore, it is now our goal to maximize the log-likelihood of the residuals, or \(L(\hat{e}\mid \hat{V})\)
Step 1: Taking the Log

- We now take the log of our original likelihood function:
  \[ L(y|\hat{V}) = L(\hat{\beta}|\hat{V})L(\hat{e}|\hat{V}) \]

- Yielding:
  \[ \log(L(y|\hat{V})) = \log(L(\hat{\beta}|\hat{V})) + \log(L(\hat{e}|\hat{V})) \]

- Which gives us:
  \[ \log(L(\hat{e}|\hat{V})) = \log(L(y|\hat{V})) - \log(L(\hat{\beta}|\hat{V})) \]
Step 2:

• We know that $y \sim N_p(X\beta, V = ZGZ^T + R)$ and 
  \[ \hat{\beta} \sim N(\beta, (X^T\hat{V}^{-1}X)^{-1}) \]

• We can then put the MVN associated with each into our log likelihood of the residual

\[
\log \left( L(\hat{e}|\hat{V}) \right) = \log \left( L(y|\hat{V}) \right) - \log \left( L(\hat{\beta}|\hat{V}) \right)
\]
Even More...

\[
\log \left( L(\hat{e}|\hat{\mathbf{V}}) \right) = \log \left( L(\mathbf{y}|\hat{\mathbf{V}}) \right) - \log \left( L(\beta|\hat{\mathbf{V}}) \right)
\]

\[
= -\frac{1}{2} \left[ \log |X^T \hat{\mathbf{V}}^{-1} \mathbf{x}| + \log |\hat{\mathbf{V}}| + (\mathbf{y} - \mathbf{x}\beta)^T \hat{\mathbf{V}}^{-1} (\mathbf{y} - \mathbf{x}\beta) - (\hat{\beta} - \beta)^T X^T \mathbf{V}^{-1} \mathbf{x} (\hat{\beta} - \beta) \right]
\]

Here:

\[
(y - x\beta)^T \hat{\mathbf{V}}^{-1} (y - x\beta) = (y - x\hat{\beta})^T \hat{\mathbf{V}}^{-1} (y - x\hat{\beta}) + (\hat{\beta} - \beta)^T X^T \mathbf{V}^{-1} \mathbf{x} (\hat{\beta} - \beta)
\]

Meaning we can cancel the last term....
The REML Log Likelihood

- After all the slides before, we can now present the REML log likelihood:

\[
\log \left( L(\hat{\mathbf{e}}|\mathbf{V}) \right) = -\frac{1}{2} \left[ \log |\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}| + \log |\mathbf{V}| + \mathbf{e}^T \mathbf{e} - \frac{n - p}{2} \log(2\pi) \right]
\]
Uses of ML and REML

• ML can be used for deviance tests when the fixed effects are the same or are different

• REML can be used for deviance tests when the fixed effects are the same only
  ➢ Residuals change when the fixed effects change

• Use of REML is analogous to what happens in regular (and repeated measures) ANOVA
  ➢ The estimated variances are unbiased
Using REML for Our First Example

- We change METHOD=REML to use REML (it is the SAS default, actually)

```plaintext
*EMPTY MODEL: unstructured R covariance matrix, REML Estimation;
PROC MIXED DATA=WORK.dietstack METHOD=REML COVTEST NOPROFILE ITDETAILS IC;
CLASS intensity;
MODEL pulse = intensity / S;
REPEATED / SUBJECT=ID TYPE=UN R RCORR;
LSMEANS intensity / ADJUST=BON;
RUN;
```

**From REML:**

```
Estimated R Matrix for Subject 1

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>279.91</td>
<td>333.53</td>
<td>395.71</td>
</tr>
<tr>
<td>2</td>
<td>333.53</td>
<td>473.05</td>
<td>571.23</td>
</tr>
<tr>
<td>3</td>
<td>395.71</td>
<td>571.23</td>
<td>770.03</td>
</tr>
</tbody>
</table>
```

**From ML:**

```
Estimated R Matrix for Subject 1

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>264.36</td>
<td>315.00</td>
<td>373.72</td>
</tr>
<tr>
<td>2</td>
<td>315.00</td>
<td>446.77</td>
<td>539.49</td>
</tr>
<tr>
<td>3</td>
<td>373.72</td>
<td>539.49</td>
<td>727.25</td>
</tr>
</tbody>
</table>
```

**Covariance Parameter Estimates**

```plaintext
<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN(1,1)</td>
<td>ID</td>
<td>279.91</td>
<td>96.0089</td>
<td>2.92</td>
<td>0.0018</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>ID</td>
<td>333.53</td>
<td>119.72</td>
<td>2.79</td>
<td>0.0053</td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>ID</td>
<td>473.05</td>
<td>162.25</td>
<td>2.92</td>
<td>0.0018</td>
</tr>
<tr>
<td>UN(3,1)</td>
<td>ID</td>
<td>395.71</td>
<td>147.95</td>
<td>2.67</td>
<td>0.0075</td>
</tr>
<tr>
<td>UN(3,2)</td>
<td>ID</td>
<td>571.23</td>
<td>201.55</td>
<td>2.83</td>
<td>0.0046</td>
</tr>
<tr>
<td>UN(3,3)</td>
<td>ID</td>
<td>770.03</td>
<td>264.12</td>
<td>2.92</td>
<td>0.0018</td>
</tr>
<tr>
<td>UN(1,1)</td>
<td>ID</td>
<td>264.36</td>
<td>88.1204</td>
<td>3.00</td>
<td>0.0013</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>ID</td>
<td>315.00</td>
<td>109.88</td>
<td>2.87</td>
<td>0.0041</td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>ID</td>
<td>446.77</td>
<td>148.92</td>
<td>3.00</td>
<td>0.0013</td>
</tr>
<tr>
<td>UN(3,1)</td>
<td>ID</td>
<td>373.72</td>
<td>135.79</td>
<td>2.75</td>
<td>0.0059</td>
</tr>
<tr>
<td>UN(3,2)</td>
<td>ID</td>
<td>539.49</td>
<td>184.99</td>
<td>2.92</td>
<td>0.0035</td>
</tr>
<tr>
<td>UN(3,3)</td>
<td>ID</td>
<td>727.25</td>
<td>242.42</td>
<td>3.00</td>
<td>0.0013</td>
</tr>
</tbody>
</table>
```
Fixed Effects in REML v. ML

• The fixed effects have the same estimates, but due to the different covariance structure, have different SEs
  ➢ And, consequently, different p-values

From REML:

Solution for Fixed Effects

| Effect    | intensity | Estimate | Standard Error | DF | t Value | Pr > |t| |
|-----------|-----------|----------|----------------|----|----------|-------|
| Intercept |           | 189.56   | 6.5406         | 17 | 28.98    | <.0001|
| intensity | 1         | -102.06  | 3.7898         | 17 | -26.93   | <.0001|
| intensity | 2         | -55.4444 | 2.3643         | 17 | -23.45   | <.0001|
| intensity | 3         | 0        | .              | .  | .        | .     |

From ML:

Solution for Fixed Effects

| Effect    | intensity | Estimate | Standard Error | DF | t Value | Pr > |t| |
|-----------|-----------|----------|----------------|----|----------|-------|
| Intercept |           | 189.56   | 6.3563         | 17 | 29.82    | <.0001|
| intensity | 1         | -102.06  | 3.6830         | 17 | -27.71   | <.0001|
| intensity | 2         | -55.4444 | 2.2976         | 17 | -24.13   | <.0001|
| intensity | 3         | 0        | .              | .  | .        | .     |
PUTTING IT TOGETHER:
REPLICATING THE FULL ANALYSIS
Recalling the Full Data Analysis Example

• The data for our analysis present several factors which we are interested in investigating:
  
  ➢ Within Subjects:
    • Effects of time of measurement on pulse rate
  
  ➢ Between Subjects:
    • Main effect of diet on pulse rate
    • Main effect of exercise type on pulse rate
    • Interaction of diet and exercise type on pulse rate (2-way)

  ➢ Within and Between Interactions
    • Interaction of measurement and diet (2-way)
    • Interaction of measurement and exercise type (2-way)
    • Interaction of measurement, diet, and exercise type (3-way)
Recalling Results: RM ANOVA from PROC GLM

- In lecture #4, we analyzed these data using the repeated measures assumptions within PROC GLM:

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>DIET</td>
<td>1</td>
<td>8791.12963</td>
<td>8791.12963</td>
<td>8.58</td>
<td>0.0126</td>
</tr>
<tr>
<td>EXERTYPE</td>
<td>2</td>
<td>1560.33333</td>
<td>780.16667</td>
<td>0.76</td>
<td>0.4884</td>
</tr>
<tr>
<td>DIET*EXERTYPE</td>
<td>2</td>
<td>718.48148</td>
<td>359.24074</td>
<td>0.55</td>
<td>0.7113</td>
</tr>
<tr>
<td>Error</td>
<td>12</td>
<td>12298.88889</td>
<td>1024.90741</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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>
<th>G - G</th>
<th>H - F</th>
</tr>
</thead>
<tbody>
<tr>
<td>intensity</td>
<td>2</td>
<td>93972.11111</td>
<td>46986.05556</td>
<td>703.72</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>intensity*DIET</td>
<td>2</td>
<td>344.92593</td>
<td>172.46296</td>
<td>2.58</td>
<td>0.1200</td>
<td>0.0964</td>
<td></td>
</tr>
<tr>
<td>intensity*EXERTYPE</td>
<td>4</td>
<td>80.55556</td>
<td>20.13889</td>
<td>0.30</td>
<td>0.8740</td>
<td>0.8043</td>
<td>0.8740</td>
</tr>
<tr>
<td>intensity<em>DIET</em>EXERTYPE</td>
<td>4</td>
<td>493.96296</td>
<td>123.49074</td>
<td>1.85</td>
<td>0.1523</td>
<td>0.1816</td>
<td>0.1523</td>
</tr>
<tr>
<td>Error(intensity)</td>
<td>24</td>
<td>1602.44444</td>
<td>66.76852</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Greenhouse-Geisser Epsilon 0.6759
Huynh-Feldt Epsilon 1.0487
Initial Step: Determine R Matrix Structure

• Including all covariates (all fixed effects – both within and between subjects factors) we estimate a series of models, each with a differing $R$ matrix structure (UN shown):

```plaintext
*ANALYSIS MODEL: unstructured R covariance matrix, REML Estimation;
PROC MIXED DATA=WORK.dietstack METHOD=REML COVTEST NOPROFILE ITDETAILS IC;
CLASS intensity diet exertype;
MODEL pulse = diet exertype diet*exertype
    intensity intensity*diet intensity*exertype intensity*diet*exertype/ S;
REPEATED / SUBJECT=ID TYPE=UN R RCORR;
RUN;
```
Model Comparison

- Because our model has changed, we cannot expect the results to be the same
  - Every predictor reduces the size of error variance

- Our results indicate that CSH is the model we should use

<table>
<thead>
<tr>
<th>Model</th>
<th>Comparison</th>
<th>R Matrix Parameters</th>
<th>-2 Res LogL</th>
<th>Deviance Statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VC</td>
<td></td>
<td>1</td>
<td>336.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CS</td>
<td>CS v. VC</td>
<td>2</td>
<td>306.0</td>
<td>36.4</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>CSH</td>
<td>CSH v. CS</td>
<td>4</td>
<td>292.0</td>
<td>14.0</td>
<td>0.0009</td>
</tr>
<tr>
<td>UN</td>
<td>UN v. CSH</td>
<td>6</td>
<td>288.1</td>
<td>3.9</td>
<td>0.1423</td>
</tr>
</tbody>
</table>
## CSH Model Results

- **Results for the R matrix:**

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Z Value</th>
<th>Pr Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var(1)</td>
<td>ID</td>
<td>193.32</td>
<td>79.0455</td>
<td>2.45</td>
<td>0.0072</td>
</tr>
<tr>
<td>Var(2)</td>
<td>ID</td>
<td>363.56</td>
<td>147.61</td>
<td>2.46</td>
<td>0.0069</td>
</tr>
<tr>
<td>Var(3)</td>
<td>ID</td>
<td>601.89</td>
<td>245.96</td>
<td>2.45</td>
<td>0.0072</td>
</tr>
<tr>
<td>CSH</td>
<td>ID</td>
<td>0.8896</td>
<td>0.05115</td>
<td>17.39</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>193.32</td>
<td>235.84</td>
<td>303.45</td>
</tr>
<tr>
<td>2</td>
<td>235.84</td>
<td>363.56</td>
<td>416.13</td>
</tr>
<tr>
<td>3</td>
<td>303.45</td>
<td>416.13</td>
<td>601.89</td>
</tr>
</tbody>
</table>

- **Results for the hypothesis tests of fixed effects:**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIET</td>
<td>1</td>
<td>12</td>
<td>8.59</td>
<td>0.0126</td>
</tr>
<tr>
<td>EXERTYPE</td>
<td>2</td>
<td>12</td>
<td>0.76</td>
<td>0.4870</td>
</tr>
<tr>
<td>DIET*EXERTYPE</td>
<td>2</td>
<td>24</td>
<td>0.35</td>
<td>0.7109</td>
</tr>
<tr>
<td>intensity</td>
<td>2</td>
<td>24</td>
<td>508.99</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>intensity*DIET</td>
<td>2</td>
<td>24</td>
<td>1.83</td>
<td>0.1817</td>
</tr>
<tr>
<td>intensity*EXERTYPE</td>
<td>4</td>
<td>24</td>
<td>0.35</td>
<td>0.8386</td>
</tr>
<tr>
<td>intensity<em>DIET</em>EXERTYPE</td>
<td>4</td>
<td>24</td>
<td>1.56</td>
<td>0.2169</td>
</tr>
</tbody>
</table>

- Significant (within subjects) main effect of intensity
- Significant (between subjects) main effect of diet
- No other effects are significant
Revisiting the Model

• In light of the non-significant results for the interactions of the model, alternative models can be run with the interactions not included
  ➢ With a small sample size, this is usually the approach taken

• However, alternative fixed-effects models using REML cannot be compared
  ➢ So we will likely have to re-evaluate our covariance structure for each

• Process can be tedious, but will result in most accurate analysis and inferences
Reduction Step #1:
Removal of Within-Subjects 3-Way Interaction

- Analysis Comparison (4 models run – CSH wins):

<table>
<thead>
<tr>
<th>Model</th>
<th>Comparison</th>
<th>R Matrix Parameters</th>
<th>-2 Res LogL</th>
<th>Deviance Statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VC</td>
<td></td>
<td>1</td>
<td>371.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CS</td>
<td>CS v. VC</td>
<td>2</td>
<td>340.1</td>
<td>31.4</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>CSH</td>
<td>CSH v. CS</td>
<td>4</td>
<td>324.5</td>
<td>15.6</td>
<td>0.0004</td>
</tr>
<tr>
<td>UN</td>
<td>UN v. CSH</td>
<td>6</td>
<td>318.7</td>
<td>5.8</td>
<td>0.0550</td>
</tr>
</tbody>
</table>

- Fixed Effects Results (more reductions possible):

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIET</td>
<td>1</td>
<td>12</td>
<td>7.17</td>
<td>0.0201</td>
</tr>
<tr>
<td>EXERTYPE</td>
<td>2</td>
<td>12</td>
<td>0.64</td>
<td>0.5462</td>
</tr>
<tr>
<td>DIET*EXERTYPE</td>
<td>2</td>
<td>12</td>
<td>3.65</td>
<td>0.0576</td>
</tr>
<tr>
<td>intensity</td>
<td>2</td>
<td>28</td>
<td>426.78</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>intensity*DIET</td>
<td>2</td>
<td>28</td>
<td>1.55</td>
<td>0.2309</td>
</tr>
<tr>
<td>intensity*EXERTYPE</td>
<td>4</td>
<td>28</td>
<td>0.35</td>
<td>0.8392</td>
</tr>
</tbody>
</table>
Reduction Step #2: Remove 2-Way Interactions of Intensity

- Analysis Comparison (4 models run – UN now wins):

<table>
<thead>
<tr>
<th>Model</th>
<th>Comparison</th>
<th>R Matrix Parameters</th>
<th>-2 Res LogL</th>
<th>Deviance Statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VC</td>
<td></td>
<td>1</td>
<td>414.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CS</td>
<td>CS v. VC</td>
<td>2</td>
<td>378.0</td>
<td>36.1</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>CSH</td>
<td>CSH v. CS</td>
<td>4</td>
<td>360.5</td>
<td>17.5</td>
<td>0.0002</td>
</tr>
<tr>
<td>UN</td>
<td>UN v. CSH</td>
<td>6</td>
<td>354.2</td>
<td>6.3</td>
<td>0.0429</td>
</tr>
</tbody>
</table>

- Fixed Effects Results (more reductions possible):

```
Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIET</td>
<td>1</td>
<td>12</td>
<td>6.82</td>
<td>0.0228</td>
</tr>
<tr>
<td>EXERTYPE</td>
<td>2</td>
<td>12</td>
<td>2.22</td>
<td>0.1515</td>
</tr>
<tr>
<td>DIET*EXERTYPE</td>
<td>2</td>
<td>12</td>
<td>1.94</td>
<td>0.1867</td>
</tr>
<tr>
<td>intensity</td>
<td>2</td>
<td>12</td>
<td>363.32</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
```
Reduction Step #3: Remove 2-Way Interactions of Diet and Type

• Analysis Comparison (4 models run – UN now wins):

<table>
<thead>
<tr>
<th>Model</th>
<th>Comparison</th>
<th>R Matrix Parameters</th>
<th>-2 Res LogL</th>
<th>Deviance Statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VC</td>
<td></td>
<td>1</td>
<td>429.6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CS</td>
<td>CS v. VC</td>
<td>2</td>
<td>394.3</td>
<td>35.3</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>CSH</td>
<td>CSH v. CS</td>
<td>4</td>
<td>379.9</td>
<td>14.4</td>
<td>0.0007</td>
</tr>
<tr>
<td>UN</td>
<td>UN v. CSH</td>
<td>6</td>
<td>371.1</td>
<td>8.8</td>
<td>0.0123</td>
</tr>
</tbody>
</table>

• Fixed Effects Results:

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIET</td>
<td>1</td>
<td>14</td>
<td>8.34</td>
<td>0.0119</td>
</tr>
<tr>
<td>EXERTYPE</td>
<td>2</td>
<td>14</td>
<td>2.00</td>
<td>0.1716</td>
</tr>
<tr>
<td>intensity</td>
<td>2</td>
<td>14</td>
<td>363.31</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Inferences About Final Model

• From the final model, we can now make inferences:
  - Significant main effect of intensity (within subjects)
    - Pulse rate changes as a function of exercise intensity (all different)
    - Need to follow up with post-hoc tests
  - Significant main effect of diet (between subjects)
    - Pulse rate changes as a function of diet (vegetarians lower)

<table>
<thead>
<tr>
<th></th>
<th>Effect</th>
<th>intensity</th>
<th>DIET</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DIET</td>
<td>1</td>
<td></td>
<td></td>
<td>145.61</td>
<td>5.0561</td>
<td>14</td>
<td>28.80</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIET</td>
<td>2</td>
<td></td>
<td></td>
<td>128.50</td>
<td>5.0561</td>
<td>14</td>
<td>25.41</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intensity 1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>87.5000</td>
<td>3.1089</td>
<td>14</td>
<td>28.15</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intensity 2</td>
<td></td>
<td>2</td>
<td>1</td>
<td>134.11</td>
<td>4.3016</td>
<td>14</td>
<td>31.18</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intensity 3</td>
<td></td>
<td>3</td>
<td>1</td>
<td>189.56</td>
<td>5.5166</td>
<td>14</td>
<td>34.36</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Effect</th>
<th>intensity</th>
<th>DIET</th>
<th>_intensity</th>
<th>_DIET</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DIET</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>17.1132</td>
<td>5.0555</td>
<td>14</td>
<td>2.89</td>
<td>0.0119</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intensity 1</td>
<td></td>
<td>1</td>
<td>2</td>
<td>-46.6111</td>
<td>1</td>
<td>-102.06</td>
<td>3.7088</td>
<td>14</td>
<td>-26.93</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intensity 2</td>
<td></td>
<td>1</td>
<td>3</td>
<td>-102.06</td>
<td>1</td>
<td>-55.4444</td>
<td>2.3643</td>
<td>14</td>
<td>-23.45</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• No significant main effect of exercise type (between subjects)
  - Stair climbing, racquetball, and weight lifting all same pulse rate
• No significant interactions between any factor
EXTENSIONS TO MANOVA
Extensions to MANOVA

• If you run PROC MIXED with REML and TYPE=CS, your within-subjects hypothesis test is identical to that of running PROC GLM for within subjects ANOVA

• There is no direct analog to MANOVA, however
  - MANOVA hypothesis test uses Wilks’ Lambda

• MANOVA can be accomplished in PROC MIXED
  - But as a modified likelihood ratio test
    - Wilks’ Lambda is equal to a scaled difference in log-likelihoods
      \[
      \lambda = \exp \left( - \frac{\text{Log}L_{H_1} - \text{Log}L_{H_0}}{N} \right)
      \]
  - Because of its limited utility, we will not use it in class
    - However, you can do so with PROC MIXED if ever needed
CONCLUDING REMARKS
Wrapping Up

• Today we phrased general linear models in the context of maximum likelihood-estimated mixed models
  ➢ A very powerful technique that can be used for nearly everything in social science statistical analyses

• The tools today re-examined RM ANOVA using PROC MIXED, but did not change any covariates
  ➢ If we had DV-differing covariates, we could use them in MIXED

• Next week will be an extension of today to models with random effects
  ➢ Mixed models/multilevel models/hierarchical linear models
  ➢ This methodology is useful for even more types of studies