BIO 226, Spring 2015: Lab 5

Fitting Linear Mixed Effects Models using PROC MIXED in SAS

Fei Li

March 26, 2015

Outline

Data: Heart Rate Measurements

- The data consists of heart rate measurements (beats/min) for 24 patients treated with drug A or B.
- The majority of the patients are measured at time of drug administration (baseline) and 1, 5, 15, 30 and 60 minutes post-administration.
- The actual measurement times are $\{0,1,5,15,30,60\}$ for all patients except IDs 1, 2 and 13 who are measured at times $\{0,1,5,15,30,45\}$, $\{0,1,5,15,22,60\}$, and $\{0,1,15,60\}$, respectively.
- Note: This data is not *balanced* (i.e., all measurements are not taken at the same times) nor is it *complete* (i.e., there are missing data).

Data: Heart Rate Measurements

```
proc means data=hrunbalanced mean std nway;
class drug time;
var hr;
output out=outmean mean=mean;
run;

proc gplot data=outmean;
plot mean*time=drug;
symbol1 i=join value=dot;
symbol2 i=join value=triangle;
run;

proc gplot data = hrunbalanced;
title 'Heart rate vs time';
symbol1 interpol=join value=triangle;
symbol2 interpol=join value=triangle;
symbol3 interpol=join value=triangle;
symbol4 interpol=join value=triangle;
symbol5 interpol=join value=triangle;
plot hr*time = id;
by drug;
run;
```

Data: Heart Rate Measurements

drug	time	Obs	Mean	Std Dev
a	0	12	82.6666667	12.7160123
	1	12	79.3333333	11.5784544
	5	12	78.8333333	12.7481431
	15	12	77.1666667	11.4561086
	22	1	86.0000000	.
	30	11	75.2727273	11.7395989
	45	1	84.0000000	.
	60	11	69.6363636	8.7552582
b	0	12	85.6666667	11.7189306
	1	12	88.0833333	10.0946280
	5	11	88.1818182	6.4779346
	15	12	85.5000000	10.0227015
	30	11	83.2727273	7.0012986
	60	12	81.0000000	9.3225045

Data: Plot of Group Means by Time

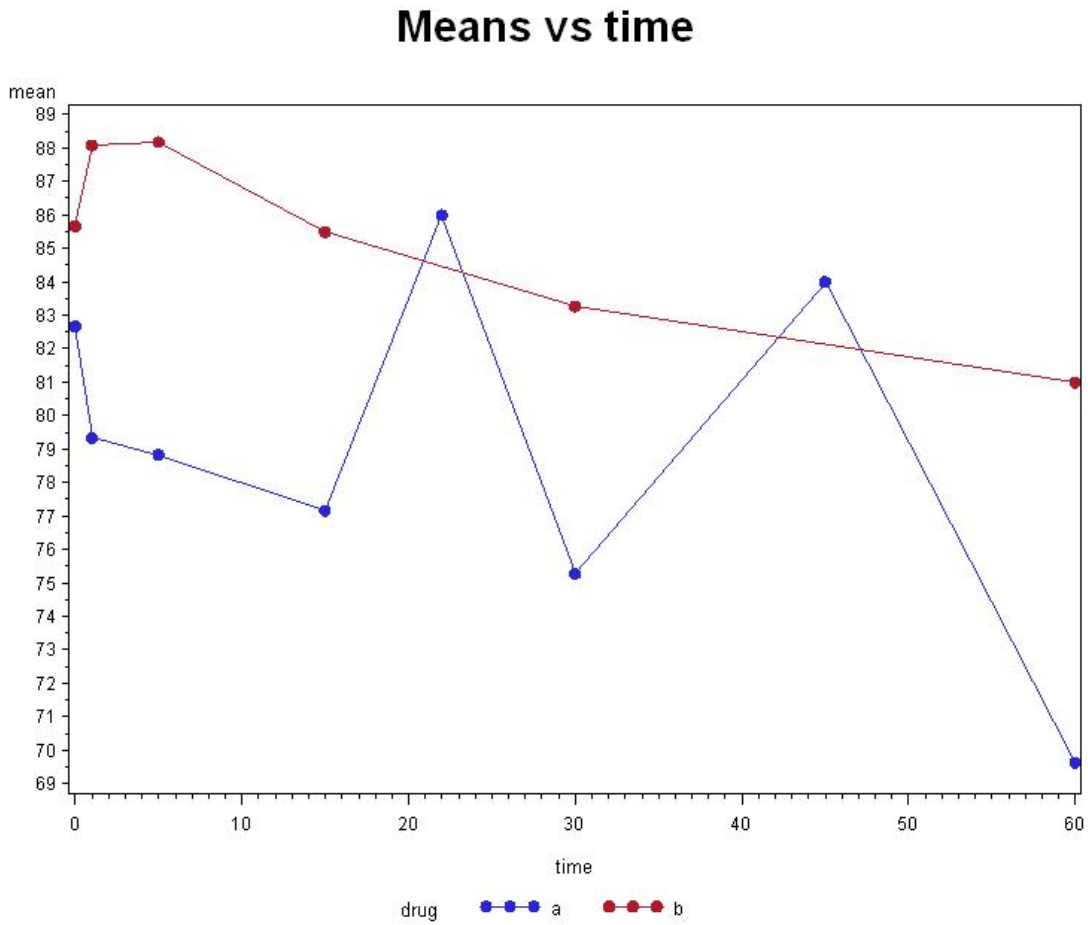

Figure : Plot of means

Data: Heart Rate Measurements

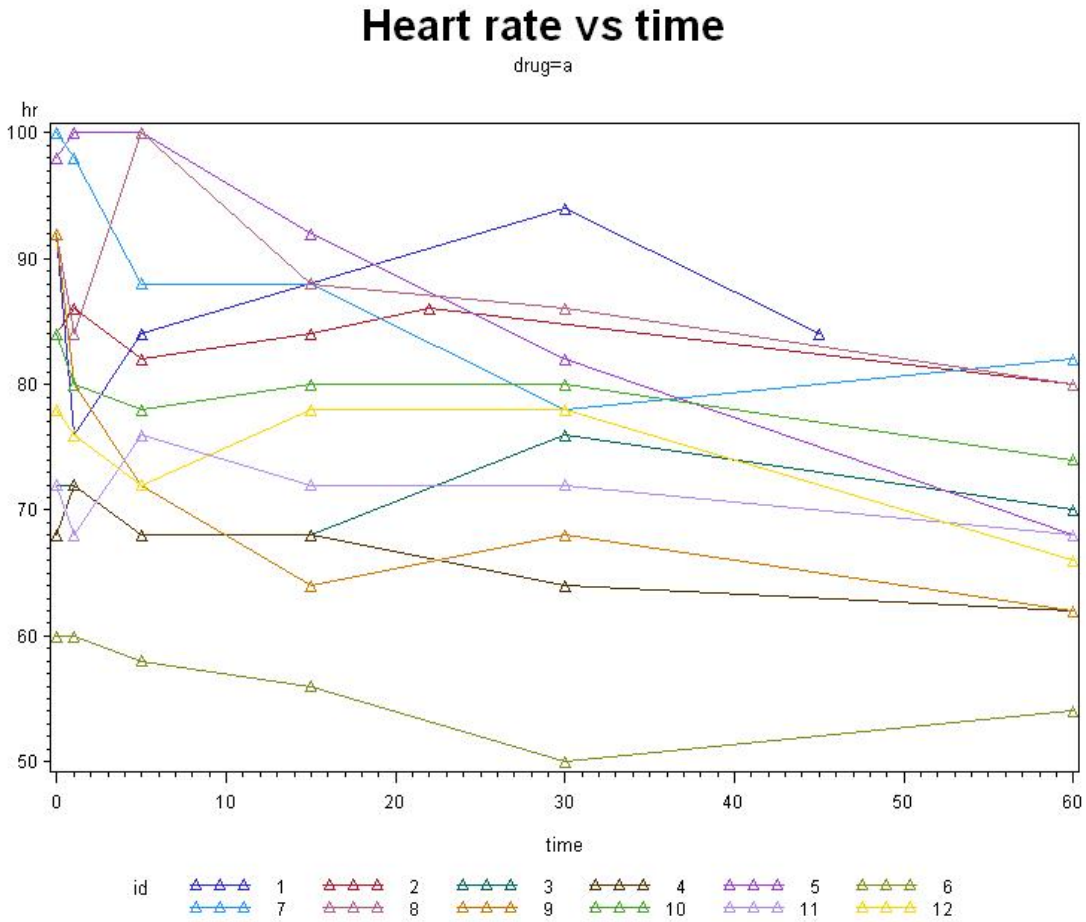

Figure : Subject trajectories

Data: Heart Rate Measurements

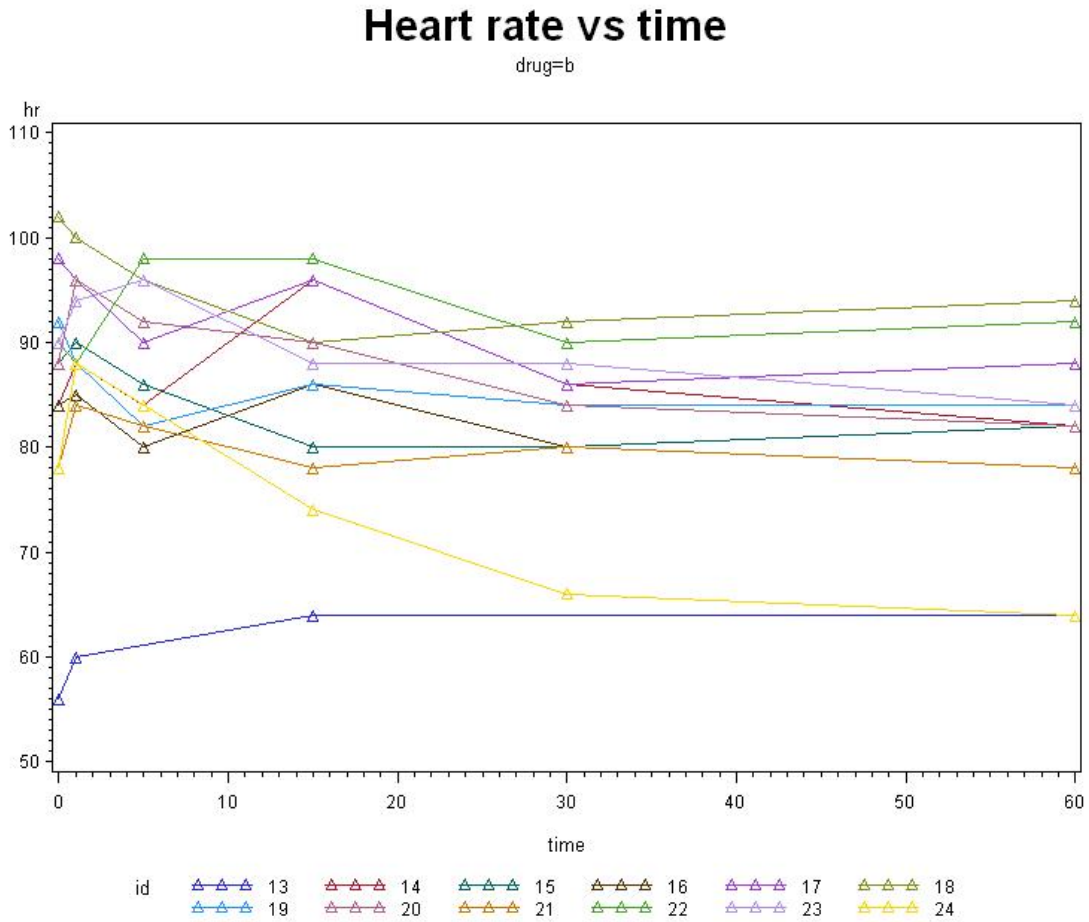

Figure : Subject trajectories

Profile analysis?

Can these data be analyzed using profile analysis? We get an error:

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	995.75315871	
1	2	1133.67391879	0.00000538
2	1	1129.30991887	0.00000613
3	1	1127.36158373	0.00000662

WARNING: Unable to make hessian positive definite.

Covariance Parameter Values
At Last Iteration

Cov Parm	Subject	Estimate
UN(1,1)	id	861.56
...		
UN(5,1)	id	85.8080
UN(5,2)	id	-173E-17
UN(5,3)	id	204.14
UN(5,4)	id	-1E-15

Further, some of the estimated variance parameters are zero $\Rightarrow$ profile analysis is not appropriate.

Parametric curve models?

For this dataset, running a linear model with an unstructured variance/covariance structure will result in the same error.

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	1064.17156333	
1	2	1075.57317831	0.00010172

WARNING: Unable to make hessian positive definite.

- The problem is that there is not enough data to estimate all of the necessary parameters in the variance covariance matrix in both approaches for this example.
- Thus, we will set up the Mixed Effects Model to bypass this problem since fewer number of covariance parameters needed to be estimated in the Mixed Effects Models.

Outline

Mixed Effects Models

1. the random intercepts model is:
   $Y_{ij} = \beta_1 + \beta_2 drug_i + \beta_3 time_{ij} + \beta_4(time_{ij} \times drug_i) + b_{1i} + \epsilon_{ij}$
2. the random intercepts & slopes model to be fit is:
   $Y_{ij} = \beta_1 + \beta_2 drug_i + \beta_3 time_{ij} + \beta_4(time_{ij} \times drug_i) + b_{1i} + b_{2i} time_{ij} + \epsilon_{ij}$

where i=1,2, . . . ,24, j=1,2, . . . ,6 and

β_1 = fixed intercept

β_2 = fixed drug effect ($drug_i = 1$ if on drug A)

β_3 = fixed time effect

β_4 = fixed effect of interaction between drug and time

b_{1i} = random effect of intercept for subject i

b_{2i} = random effect of slope over time for subject i

Mixed Effects Models

Note that we can rewrite our random intercepts & slopes model (for patient i) in the form:

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i$$

- For all individuals except $i = 1, 2$, and 13, We have heart rates measurements on patient i at times (0, 1, 5, 15, 30, 60)
- For i =1 and i =2 the times are (0, 1, 5, 15, 30, 45) and (0, 1, 5, 15, 22, 60), respectively and for i =13, the times are (0, 1, 15, 60).
- $Y_i = (Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4}, Y_{i5}, Y_{i6})'$ a 6×1 vector for all i except 13.
- $Y_i = (Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4})'$ a 4×1 vector for $i = 13$.
- Thus, for any individual with n_i measurements, Y_i is a $n_i \times 1$ vector.

Mixed Effects Models

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i$$

- X_i is a $n_i \times 4$ matrix of covariates (intercept, drug(A=1, B=0), time, drug*time) .
- X_i is a 6×4 matrix of covariates for all i's except i =13, where X_{13} is a 4×4 matrix of covariates.

$$X_i = \begin{pmatrix} 1 & drug_i & time_{i1} & drug_i \times time_{i1} \\ 1 & drug_i & time_{i2} & drug_i \times time_{i2} \\ . & . & . & . \\ . & . & . & . \\ 1 & drug_i & time_{in_i} & drug_i \times time_{in_i} \end{pmatrix}$$

- β is a 4×1 vector of *fixed effects* [*population* averages] for all i's: $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$.

Mixed Effects Models

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i$$

- Z_i is a $n_i \times 2$ matrix of covariates to describe random effects (intercept, $time_{ij}$).

$$Z_i = \begin{pmatrix} 1 & time_{i1} \\ 1 & time_{i2} \\ . & . \\ . & . \\ 1 & time_{in_i} \end{pmatrix}$$

- b_i is a 2×1 vector of *random effects* for all i's, i.e. $b_i = (b_{1i}, b_{2i})'$, assuming $b_i \sim N_2(\mathbf{0}, \mathbf{G})$. So, $\mathbf{G}$ is 2×2 variance-covariance matrix

Mixed Effects Models

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i$$

- Z_i is a $n_i \times 1$ matrix of covariates to describe random effects (intercept ONLY).

$$Z_i = \begin{pmatrix} 1 \\ 1 \\ . \\ . \\ 1 \end{pmatrix}$$

- b_i is a 1×1 vector of *random effects* for all i's, i.e. $b_i \sim N_1(\mathbf{0}, \mathbf{G})$. So, $\mathbf{G} = [\sigma_b^2]$ is 1×1 variance-covariance matrix.

Mixed Effects Models

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i$$

- ϵ_i is an $n_i \times 1$ vector of measurement errors, assuming $\epsilon_i \sim N_{n_i}(\mathbf{0}, R_i)$ with $R_i = \sigma^2 I_{n_i}$. i.e.

$$\begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ . \\ . \\ \epsilon_{in_i} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ . \\ . \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 & \dots & \dots & 0 \\ 0 & \sigma^2 & \dots & \dots & 0 \\ . & . & \dots & \dots & . \\ . & . & \dots & \dots & . \\ 0 & 0 & \dots & \dots & \sigma^2 \end{bmatrix} \right)$$

- Thus we are assuming that within a given subject, measurement errors are independent (conditional independence).
- Also, we assume that b_i is independent of ϵ_i.
- Therefore, marginal covariance
  $Cov(Y_i) = Cov(X_i\beta) + Cov(Z_i b_i) + Cov(\epsilon_i) =$
  $0 + Z_i Cov(b_i) Z_i' + Cov(\epsilon_i) = Z_i G Z_i' + \sigma^2 I_{n_i}$. (Main idea: variables in Z end up in the covariance matrix for Y).

Aside: $Cov(Y_i)$ for random intercepts ONLY

$$Cov(Y_i) = Z_i G Z_i' + \sigma^2 I_{n_i}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \\ . \\ . \\ 1 \end{bmatrix} \begin{bmatrix} \sigma_b^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & . & . & 1 \end{bmatrix} + \sigma^2 \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 \end{bmatrix} + \begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma^2 & \dots & 0 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

Aside: $Cov(Y_i)$ for random intercepts ONLY

$$Cov(Y_i) = Z_i G Z_i' + \sigma^2 I_{n_i}$$

$$= \begin{bmatrix} \sigma_b^2 + \sigma^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 + \sigma^2 & \dots & \sigma_b^2 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 + \sigma^2 \end{bmatrix}$$

Mixed Effects Models

To illustrate, most patients on drug A ($i = 3, \ldots, 12$) will have:

$$X_i = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 5 & 5 \\ 1 & 1 & 15 & 15 \\ 1 & 1 & 30 & 30 \\ 1 & 1 & 60 & 60 \end{bmatrix} \qquad Z_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 5 \\ 1 & 15 \\ 1 & 30 \\ 1 & 60 \end{bmatrix}$$

and most patients on drug B ($i = 14, \ldots, 24$) will have:

$$X_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 5 & 0 \\ 1 & 0 & 15 & 0 \\ 1 & 0 & 30 & 0 \\ 1 & 0 & 60 & 0 \end{bmatrix} \qquad Z_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 5 \\ 1 & 15 \\ 1 & 30 \\ 1 & 60 \end{bmatrix}$$

Mixed Effects Models

But for $i = 1, 2, 13$, we have

$$X_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 5 & 5 \\ 1 & 1 & 15 & 15 \\ 1 & 1 & 30 & 30 \\ 1 & 1 & 45 & 45 \end{bmatrix} \qquad Z_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 5 \\ 1 & 15 \\ 1 & 30 \\ 1 & 45 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 5 & 5 \\ 1 & 1 & 15 & 15 \\ 1 & 1 & 22 & 22 \\ 1 & 1 & 60 & 60 \end{bmatrix} \qquad Z_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 5 \\ 1 & 15 \\ 1 & 22 \\ 1 & 60 \end{bmatrix}$$

$$X_{13} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 15 & 0 \\ 1 & 0 & 60 & 0 \end{bmatrix} \qquad Z_{13} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 15 \\ 1 & 60 \end{bmatrix}$$

Mixed Effects Models

More specifically, for patients on drug A ($i = 3, \ldots, 12$):
$Y_i = X_i\beta + Z_i b_i + \epsilon_i$, which is

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \\ Y_{i5} \\ Y_{i6} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 5 & 5 \\ 1 & 1 & 15 & 15 \\ 1 & 1 & 30 & 30 \\ 1 & 1 & 60 & 60 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 5 \\ 1 & 15 \\ 1 & 30 \\ 1 & 60 \end{bmatrix} \begin{bmatrix} b_{1i} \\ b_{2i} \end{bmatrix} + \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \epsilon_{i3} \\ \epsilon_{i4} \\ \epsilon_{i5} \\ \epsilon_{i6} \end{bmatrix}$$

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \\ Y_{i5} \\ Y_{i6} \end{bmatrix} = \begin{bmatrix} \beta_1 + \beta_2 + b_{1i} + \epsilon_{i1} \\ \beta_1 + \beta_2 + \beta_3 + \beta_4 + b_{1i} + b_{2i} + \epsilon_{i2} \\ \beta_1 + \beta_2 + 5\beta_3 + 5\beta_4 + b_{1i} + 5b_{2i} + \epsilon_{i3} \\ \beta_1 + \beta_2 + 15\beta_3 + 15\beta_4 + b_{1i} + 15b_{2i} + \epsilon_{i4} \\ \beta_1 + \beta_2 + 30\beta_3 + 30\beta_4 + b_{1i} + 30b_{2i} + \epsilon_{i5} \\ \beta_1 + \beta_2 + 60\beta_3 + 60\beta_4 + b_{1i} + 60b_{2i} + \epsilon_{i6} \end{bmatrix}$$

$$Y_{ij} = (\beta_1 + \beta_2 + b_{1i}) + (\beta_3 + \beta_4 + b_{2i})time_{ij} + \epsilon_{ij}$$

Mixed Effects Models

For patients on drug B ($i = 3, \ldots, 12$):

$$Y_{ij} = (\beta_1 + b_{1i}) + (\beta_3 + b_{2i})time_{ij} + \epsilon_{ij}$$

Thus,

$$\begin{aligned} Y_{ij} &= (\beta_1 + \beta_2 drug_i + b_{1i}) + (\beta_3 + \beta_4 drug_i + b_{2i})time_{ij} + \epsilon_{ij} \\ &= \beta_1 + \beta_2 drug_i + \beta_3 time_{ij} + \beta_4 (time_{ij} \times drug_i) + b_{1i} + b_{2i} time_{ij} + \epsilon_{ij} \end{aligned}$$

which gives us the original form of the random intercept & slope model.

- We also assumed:
  - $b_i \sim N_2(\mathbf{0}, \mathbf{G})$ where $G = \begin{bmatrix} var(b_{1i}) & cov(b_{1i}, b_{2i}) \\ cov(b_{1i}, b_{2i}) & var(b_{2i}) \end{bmatrix}$
  - $\epsilon_i \sim N_{n_i}(\mathbf{0}, R_i)$ with $R_i = \sigma^2 I_{n_i}$
  - b_i is independent of ϵ_i

SAS Code - RANDOM statement

Let's run these new models in SAS.

- The SAS code below fits a mixed model for the heart rate data, with random intercepts.

```
* MIXED MODEL: RANDOM INTERCEPTs ;
proc mixed data=hrunbalanced method=reml noclprint noitprint;
  class id drug;
  model hr = drug time time*drug / s;
  random intercept / type=un subject=id g gcorr v=1,2,3,13 vcorr=1,2,3,13;
run;
```

- To fit random slopes as well, just add `time` in the `RANDOM` statement.

```
 * MIXED MODEL: RANDOM INTERCEPTs & SLOPEs ;
proc mixed data=hrunbalanced method=reml noclprint noitprint;
  class id drug;
  model hr = drug time time*drug / s;
  random intercept time / type=un subject=id g gcorr v=1,2,3,13 vcorr=1,2,3,13;
run;
```

SAS code - RANDOM statement

- Specify the model parameters that you will allow to vary between individuals (random effects)
  - Unlike the MODEL statement, we should include INTERCEPT if we want to allow the intercepts to vary between individuals (i.e., it is not included by default)
  - Other random effects are specified by the name of their respective variable
- Required 'options' include:
  - **type** to specify the covariance structure for the random effects (**G**)
  - **subject** to specify independent units of measurement. Pairs of observations with distinct values of that variable are regarded as independent.

SAS code - RANDOM statement

- Other 'options' used here:
  - **g** to output the estimated matrix **G**, which is $Cov(b_i)$
  - **gcorr** to output the estimated correlation matrix corresponding to **G**
  - **v** to output the estimated *marginal* covariance $\mathbf{\Sigma}_i$, which is $Cov(Y_i)$.
  - **vcorr** to output the estimated correlation matrix corresponding to the *marginal* covariance $\mathbf{\Sigma}_i$
  - SOLUTION (s) can be used to print the estimated random intercepts and slopes for each individual.

- The REPEATED statement is used to specify assumptions about the covariance among the errors.
- When the REPEATED statement is not included in PROC MIXED, it is assumed, by default, that $R_i = \sigma^2 I_6$, the errors are independent within one subject.

Selected SAS Output

```
The Mixed Procedure

                Model Information

Data Set                      WORK.HRUNBALANCED
Dependent Variable            hr
Covariance Structure          Unstructured
Subject Effect                id
Estimation Method             REML
...
            Dimensions

Covariance Parameters             2
Columns in X                      6
Columns in Z Per Subject          1
Subjects                         24
Max Obs Per Subject               6
...

         Estimated G Matrix

 Row    Effect       id        Col1

   1    Intercept     1     92.2818


   Estimated G Correlation Matrix

 Row    Effect       id        Col1

   1    Intercept     1      1.0000
```

Selected SAS Output

Estimated G Correlation Matrix

Row	Effect	id	Col1
1	Intercept	1	1.0000

Mixed model: random intercepts

Estimated V Matrix for id 1

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	117.18	92.2818	92.2818	92.2818	92.2818	92.2818
2	92.2818	117.18	92.2818	92.2818	92.2818	92.2818
3	92.2818	92.2818	117.18	92.2818	92.2818	92.2818
4	92.2818	92.2818	92.2818	117.18	92.2818	92.2818
5	92.2818	92.2818	92.2818	92.2818	117.18	92.2818
6	92.2818	92.2818	92.2818	92.2818	92.2818	117.18

Estimated V Correlation Matrix for id 1

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.7875	0.7875	0.7875	0.7875	0.7875
2	0.7875	1.0000	0.7875	0.7875	0.7875	0.7875
3	0.7875	0.7875	1.0000	0.7875	0.7875	0.7875
4	0.7875	0.7875	0.7875	1.0000	0.7875	0.7875
5	0.7875	0.7875	0.7875	0.7875	1.0000	0.7875
6	0.7875	0.7875	0.7875	0.7875	0.7875	1.0000

Selected SAS Output

Estimated V Matrix for id 2

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	117.18	92.2818	92.2818	92.2818	92.2818	92.2818
2	92.2818	117.18	92.2818	92.2818	92.2818	92.2818
3	92.2818	92.2818	117.18	92.2818	92.2818	92.2818
4	92.2818	92.2818	92.2818	117.18	92.2818	92.2818
5	92.2818	92.2818	92.2818	92.2818	117.18	92.2818
6	92.2818	92.2818	92.2818	92.2818	92.2818	117.18

Estimated V Correlation Matrix for id 2

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.7875	0.7875	0.7875	0.7875	0.7875
2	0.7875	1.0000	0.7875	0.7875	0.7875	0.7875
3	0.7875	0.7875	1.0000	0.7875	0.7875	0.7875
4	0.7875	0.7875	0.7875	1.0000	0.7875	0.7875
5	0.7875	0.7875	0.7875	0.7875	1.0000	0.7875
6	0.7875	0.7875	0.7875	0.7875	0.7875	1.0000

Selected SAS Output

Estimated V Matrix for id 3

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	117.18	92.2818	92.2818	92.2818	92.2818	92.2818
2	92.2818	117.18	92.2818	92.2818	92.2818	92.2818
3	92.2818	92.2818	117.18	92.2818	92.2818	92.2818
4	92.2818	92.2818	92.2818	117.18	92.2818	92.2818
5	92.2818	92.2818	92.2818	92.2818	117.18	92.2818
6	92.2818	92.2818	92.2818	92.2818	92.2818	117.18

Estimated V Correlation Matrix for id 3

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.7875	0.7875	0.7875	0.7875	0.7875
2	0.7875	1.0000	0.7875	0.7875	0.7875	0.7875
3	0.7875	0.7875	1.0000	0.7875	0.7875	0.7875
4	0.7875	0.7875	0.7875	1.0000	0.7875	0.7875
5	0.7875	0.7875	0.7875	0.7875	1.0000	0.7875
6	0.7875	0.7875	0.7875	0.7875	0.7875	1.0000

Estimated V Matrix for id 13

Row	Col1	Col2	Col3	Col4
1	117.18	92.2818	92.2818	92.2818
2	92.2818	117.18	92.2818	92.2818
3	92.2818	92.2818	117.18	92.2818
4	92.2818	92.2818	92.2818	117.18

Selected SAS Output

Estimated V Correlation Matrix for id 13

Row	Col1	Col2	Col3	Col4
1	1.0000	0.7875	0.7875	0.7875
2	0.7875	1.0000	0.7875	0.7875
3	0.7875	0.7875	1.0000	0.7875
4	0.7875	0.7875	0.7875	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	id	92.2818
Residual		24.8959

Fit Statistics

-2 Res Log Likelihood	933.4
AIC (smaller is better)	937.4
AICC (smaller is better)	937.4
BIC (smaller is better)	939.7

Selected SAS Output

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
1	130.82	<.0001

Solution for Fixed Effects

Effect	drug	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		86.6368	2.8839	22	30.04	<.0001
drug	a	-6.0880	4.0764	116	-1.49	0.1380
drug	b	0	.	.	.	.
time		-0.1078	0.02787	116	-3.87	0.0002
time*drug	a	-0.05993	0.03969	116	-1.51	0.1337
time*drug	b	0	.	.	.	.

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
drug	1	116	2.23	0.1380
time	1	116	48.18	<.0001
time*drug	1	116	2.28	0.1337

Selected SAS Output - now with random slopes too

```
Heart rate vs time                                                    68
Mixed model: random intercepts & slopes

The Mixed Procedure
                  Model Information

Data Set                        WORK.HRUNBALANCED
Dependent Variable              hr
Covariance Structure            Unstructured
Subject Effect                  id
Estimation Method               REML
...
             Dimensions

Covariance Parameters               4
Columns in X                        6
Columns in Z Per Subject            2
Subjects                           24
Max Obs Per Subject                 6
...
               Estimated G Matrix

 Row    Effect       id        Col1        Col2

   1    Intercept     1      113.14     -0.6281
   2    time          1     -0.6281     0.01184
```

Selected SAS Output

Estimated G Correlation Matrix

Row	Effect	id	Col1	Col2
1	Intercept	1	1.0000	-0.5426
2	time	1	-0.5426	1.0000

Estimated V Matrix for id 1

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	132.15	112.51	110.00	103.72	94.2946	84.8731
2	112.51	130.91	109.43	103.27	94.0217	84.7779
3	110.00	109.43	126.17	101.46	92.9303	84.3970
4	103.72	103.27	101.46	115.97	90.2018	83.4447
5	94.2946	94.0217	92.9303	90.2018	105.12	82.0163
6	84.8731	84.7779	84.3970	83.4447	82.0163	99.6015

Estimated V Correlation Matrix for id 1

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.8554	0.8519	0.8378	0.8000	0.7398
2	0.8554	1.0000	0.8515	0.8381	0.8015	0.7425
3	0.8519	0.8515	1.0000	0.8388	0.8069	0.7529
4	0.8378	0.8381	0.8388	1.0000	0.8169	0.7764
5	0.8000	0.8015	0.8069	0.8169	1.0000	0.8015
6	0.7398	0.7425	0.7529	0.7764	0.8015	1.0000

Selected SAS Output

Estimated V Matrix for id 2

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	132.15	112.51	110.00	103.72	99.3193	75.4517
2	112.51	130.91	109.43	103.27	98.9518	75.5341
3	110.00	109.43	126.17	101.46	97.4814	75.8636
4	103.72	103.27	101.46	115.97	93.8056	76.6876
5	99.3193	98.9518	97.4814	93.8056	110.25	77.2644
6	75.4517	75.5341	75.8636	76.6876	77.2644	99.4091

Estimated V Correlation Matrix for id 2

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.8554	0.8519	0.8378	0.8228	0.6583
2	0.8554	1.0000	0.8515	0.8381	0.8237	0.6621
3	0.8519	0.8515	1.0000	0.8388	0.8265	0.6774
4	0.8378	0.8381	0.8388	1.0000	0.8296	0.7142
5	0.8228	0.8237	0.8265	0.8296	1.0000	0.7380
6	0.6583	0.6621	0.6774	0.7142	0.7380	1.0000

Selected SAS Output

Estimated V Matrix for id 3

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	132.15	112.51	110.00	103.72	94.2946	75.4517
2	112.51	130.91	109.43	103.27	94.0217	75.5341
3	110.00	109.43	126.17	101.46	92.9303	75.8636
4	103.72	103.27	101.46	115.97	90.2018	76.6876
5	94.2946	94.0217	92.9303	90.2018	105.12	77.9236
6	75.4517	75.5341	75.8636	76.6876	77.9236	99.4091

Estimated V Correlation Matrix for id 3

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.8554	0.8519	0.8378	0.8000	0.6583
2	0.8554	1.0000	0.8515	0.8381	0.8015	0.6621
3	0.8519	0.8515	1.0000	0.8388	0.8069	0.6774
4	0.8378	0.8381	0.8388	1.0000	0.8169	0.7142
5	0.8000	0.8015	0.8069	0.8169	1.0000	0.7623
6	0.6583	0.6621	0.6774	0.7142	0.7623	1.0000

Selected SAS Output

Estimated V Matrix for id 13

Row	Col1	Col2	Col3	Col4
1	132.15	112.51	103.72	75.4517
2	112.51	130.91	103.27	75.5341
3	103.72	103.27	115.97	76.6876
4	75.4517	75.5341	76.6876	99.4091

Estimated V Correlation Matrix for id 13

Row	Col1	Col2	Col3	Col4
1	1.0000	0.8554	0.8378	0.6583
2	0.8554	1.0000	0.8381	0.6621
3	0.8378	0.8381	1.0000	0.7142
4	0.6583	0.6621	0.7142	1.0000

Selected SAS Output

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	id	113.14
UN(2,1)	id	-0.6281
UN(2,2)	id	0.01184
Residual		19.0136

Fit Statistics

-2 Res Log Likelihood	921.5
AIC (smaller is better)	929.5
AICC (smaller is better)	929.8
BIC (smaller is better)	934.2

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	142.67	<.0001

Selected SAS Output

Solution for Fixed Effects

Effect	drug	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		86.5888	3.1476	22	27.51	<.0001
drug	a	-6.0748	4.4499	94	-1.37	0.1755
drug	b	0	.	.	.	.
time		-0.1062	0.03975	22	-2.67	0.0139
time*drug	a	-0.05824	0.05641	94	-1.03	0.3045
time*drug	b	0	.	.	.	.

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
drug	1	94	1.86	0.1755
time	1	22	23.02	<.0001
time*drug	1	94	1.07	0.3045

Comparing Mixed Effects Models

- Suppose we want to choose between the random intercepts model and the random intercepts & slopes model.
  1. the random intercepts model is:
     $Y_{ij} = \beta_1 + \beta_2 drug_i + \beta_3 time_{ij} + \beta_4 (time_{ij} \times drug_i) + b_{1i} + \epsilon_{ij}$
  2. the random intercepts & slopes model to be fit is:
     $Y_{ij} = \beta_1 + \beta_2 drug_i + \beta_3 time_{ij} + \beta_4 (time_{ij} \times drug_i) + b_{1i} + b_{2i} time_{ij} + \epsilon_{ij}$
- Since the random intercepts model is nested in the random intercepts & slopes model, we could use Likelihood Ratio Test to choose a model for the covariance.
- The null hypothesis is that the slope of time is not random, i.e. $Var(b_{2i}) = 0$ and $Cov(b_{1i}, b_{2i}) = 0$.
- This means testing a hypothesis on the boundary of the parameter space. Thus, the likelihood ratio test statistic is NOT distributed χ^2_2 as you would think!

Comparing Mixed Effects Models

- Options?
  1. Use table in Appendix C of FLW to get cutoffs for the actual distribution (see p.176-179 and 205-206 in text for details).
  2. Use the AIC.
- AIC for random intercepts only $= 937.4$.
- AIC for random intercepts and slopes $= 929.5 \Rightarrow$ winner!

Outline

Results

So, we will keep the random intercepts & slopes model:

$$Y_{ij} = \beta_1 + \beta_2 drug_i + \beta_3 time_{ij} + \beta_4 (time_{ij} \times drug_i) + b_{1i} + b_{2i} time_{ij} + \epsilon_{ij}$$

The estimates of our fixed effects in the model, $\hat{\boldsymbol{\beta}}$, are:

Solution for Fixed Effects

Effect	drug	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		86.5888	3.1476	22	27.51	<.0001
drug	a	-6.0748	4.4499	94	-1.37	0.1755
drug	b	0	.	.	.	.
time		-0.1062	0.03975	22	-2.67	0.0139
time*drug	a	-0.05824	0.05641	94	-1.03	0.3045
time*drug	b	0	.	.	.	.

- With a linear mixed model, the interpretation of the fixed effects parameters is the same as always.
- For example, on average, patients on drug B have a slope of -0.1062 beats/min, and those on drug A have a slope of -0.16444 beats/min (= -0.1062-0.05824), on average.

Results

Estimated G Matrix

Row	Effect	id	Col1	Col2
1	Intercept	1	113.14	-0.6281
2	time	1	-0.6281	0.01184

Estimated G Correlation Matrix

Row	Effect	id	Col1	Col2
1	Intercept	1	1.0000	-0.5426
2	time	1	-0.5426	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	id	113.14
UN(2,1)	id	-0.6281
UN(2,2)	id	0.01184
Residual		19.0136

- It follows that the correlation between the random intercepts and slopes is -0.54. The negative correlation indicates that lines with larger intercepts tend to have smaller slopes and vice versa (see scatterplot).
- The estimate of the within subject variability is $\hat{R}_i = \hat{\sigma}^2 I_{n_i}$, with $\hat{\sigma}^2 = 19.01$.

Inferences about the population

- Our model: $Y_i = X_i\beta + Z_i b_i + \epsilon_i$ where
  $Y_{ij} = \beta_1 + \beta_2 drug_i + \beta_3 time_{ij} + \beta_4(time_{ij} \times drug_i) + b_{1i} + b_{2i} time_{ij} + \epsilon_{ij}$
- When we want to talk about the population average response, we look at the *marginal* mean and covariance of Y_i:

$$\begin{aligned} \mathrm{E}(Y_i) &= X_i\beta \\ \mathrm{Cov}(Y_i) &= \Sigma_i = \mathrm{Cov}(Z_i b_i) + \mathrm{Cov}(\epsilon_i) = Z_i G Z_i' + \sigma^2 I_{n_i} \end{aligned}$$

- Thus, the population average response profile for patients on drug A is

$$\begin{aligned} \hat{\mathrm{E}}(Y_{ij} \mid drug = A) &= \hat\beta_1 + \hat\beta_2 + (\hat\beta_3 + \hat\beta_4) time_{ij} \\ &= 80.5 - 0.16\, time_{ij} \end{aligned}$$

- For patients on drug B:

$$\begin{aligned} \hat{\mathrm{E}}(Y_{ij} \mid drug = B) &= \hat\beta_1 + \hat\beta_3\, time_{ij} \\ &= 86.6 - 0.11\, time_{ij} \end{aligned}$$

Inferences about the population

- Our model: $Y_i = X_i\beta + Z_i b_i + \epsilon_i$

In matrix notation, patients on drug A measured at times 0,1,5,15,30,60 have **marginal or population-averaged mean** of Y_i:

$$\hat{\mathrm{E}}(Y_i \mid drug = A) = X_i\hat{\beta} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 5 & 5 \\ 1 & 1 & 15 & 15 \\ 1 & 1 & 30 & 30 \\ 1 & 1 & 60 & 60 \end{pmatrix} \begin{pmatrix} 86.59 \\ -6.07 \\ -0.11 \\ -0.06 \end{pmatrix} = \begin{pmatrix} 80.5 \\ 80.4 \\ 79.7 \\ 78.0 \\ 75.4 \\ 70.3 \end{pmatrix}$$

with **marginal** covariance $\widehat{\mathrm{Cov}}(Y_i) = Z_i\widehat{G}Z_i' + \hat{\sigma}^2 I_{n_i}$:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 5 \\ 1 & 15 \\ 1 & 30 \\ 1 & 60 \end{pmatrix} \begin{pmatrix} 113.14 & -0.63 \\ -0.63 & 0.01 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 5 \\ 1 & 15 \\ 1 & 30 \\ 1 & 60 \end{pmatrix}' + 19.0 \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Inferences about the population

$$\widehat{\mathrm{Cov}}(Y_i) = \begin{pmatrix} 132.2 & 112.5 & 110.0 & 103.7 & 94.3 & 75.5 \\ 112.5 & 130.9 & 109.4 & 103.3 & 94.0 & 75.5 \\ 110.0 & 109.4 & 126.2 & 101.5 & 92.9 & 75.9 \\ 103.7 & 103.3 & 101.5 & 116.0 & 90.2 & 76.7 \\ 94.3 & 94.0 & 92.9 & 90.2 & 105.1 & 77.9 \\ 75.5 & 75.5 & 75.9 & 76.7 & 77.9 & 99.4 \end{pmatrix}$$

This matches with the output for the **v** option in SAS!

Estimated V Matrix for id 3

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	132.15	112.51	110.00	103.72	94.2946	75.4517
2	112.51	130.91	109.43	103.27	94.0217	75.5341
3	110.00	109.43	126.17	101.46	92.9303	75.8636
4	103.72	103.27	101.46	115.97	90.2018	76.6876
5	94.2946	94.0217	92.9303	90.2018	105.12	77.9236
6	75.4517	75.5341	75.8636	76.6876	77.9236	99.4091

Thus **v** option gives the marginal covariance of the responses for one individual.

Inferences about the population

- Patients on drug B measured at times 0, 1, 5, 15, 30, 60:

$$\hat{\mathrm{E}}(Y_i \mid drug = B) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 5 & 0 \\ 1 & 0 & 15 & 0 \\ 1 & 0 & 30 & 0 \\ 1 & 0 & 60 & 0 \end{pmatrix} \begin{pmatrix} 86.59 \\ -6.07 \\ -0.11 \\ -0.06 \end{pmatrix} = \begin{pmatrix} 86.6 \\ 86.5 \\ 86.0 \\ 84.9 \\ 83.3 \\ 80.0 \end{pmatrix}$$

  with the same covariance as those receiving drug A since $\widehat{\mathrm{Cov}}(Y_i) = Z_i \widehat{G} Z_i' + \hat{\sigma}^2 I_{n_i}$ does not depend on the drug group.

- How will the above calculations differ for patients 1, 2, and 13?

Inferences about the population

Estimated V Matrix for id 1

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	132.15	112.51	110.00	103.72	94.2946	84.8731
2	112.51	130.91	109.43	103.27	94.0217	84.7779
3	110.00	109.43	126.17	101.46	92.9303	84.3970
4	103.72	103.27	101.46	115.97	90.2018	83.4447
5	94.2946	94.0217	92.9303	90.2018	105.12	82.0163
6	84.8731	84.7779	84.3970	83.4447	82.0163	99.6015

Estimated V Matrix for id 2

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	132.15	112.51	110.00	103.72	99.3193	75.4517
2	112.51	130.91	109.43	103.27	98.9518	75.5341
3	110.00	109.43	126.17	101.46	97.4814	75.8636
4	103.72	103.27	101.46	115.97	93.8056	76.6876
5	99.3193	98.9518	97.4814	93.8056	110.25	77.2644
6	75.4517	75.5341	75.8636	76.6876	77.2644	99.4091

Estimated V Matrix for id 13

Row	Col1	Col2	Col3	Col4
1	132.15	112.51	103.72	75.4517
2	112.51	130.91	103.27	75.5341
3	103.72	103.27	115.97	76.6876
4	75.4517	75.5341	76.6876	99.4091