STAT 5200 Handout #23

Repeated Measures Example (Ch. 16) and Crossover Design Example (Ch. 13 & 16)

Example 1: Glucose

An experiment is conducted to evaluate the effects of three diets on the serum glucose levels of human subjects. Twelve people participate in the experiment and are randomly assigned to the three diets. Serum glucose measurements are made 15, 30, and 45 minutes after eating a meal from the appropriate diet. That is, there is one measurement per subject per time period.

Diet:	1				2				3			
Subject:	1	2	3	4	5	6	7	8	9	10	11	12
Time: 15	x	x	x	x	x	x	x	x	x	x	x	x
30	x	x	x	x	x	x	x	x	x	x	x	x
45	x	x	x	x	x	x	x	x	x	x	x	x

x = msmnt. taken

The treatment structure in this experiment is that of a completely randomized design. There is one factor with three levels, and the "experimental units" (subjects) are randomly assigned to the factor levels. However, we have three measurements on each subject collected over time so this experiment is a repeated measures design.

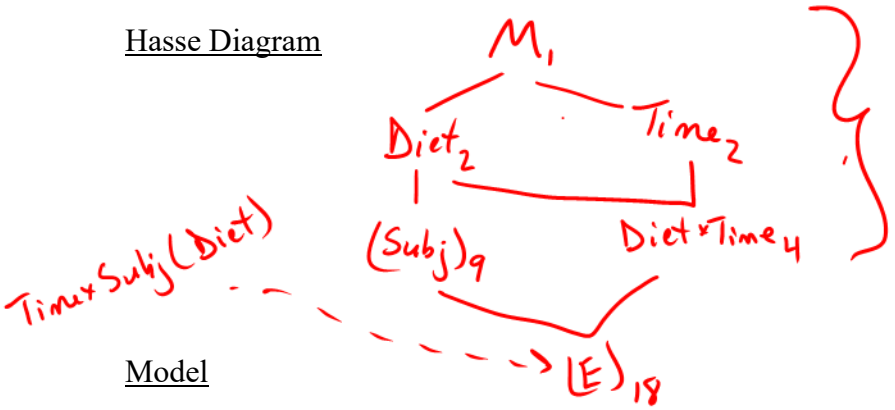
↳ repeat measurements on each exp. unit [subject]

This looks similar to a split-plot design:

- Whole-plot factor: Diet
- Whole-plot unit: Subject
- Split-plot factor: Time
- Split-plot unit: Subject (at particular time)

However, unlike the split-plot design, here there is only one randomization (for the whole-plot factor). The split-plot factor (time) is not randomized to split-plot units.

Hasse Diagram



18 DF

+ 18

36 = N

Model

$$Y_{ijk} = \mu + Diet_i + Subj_{k(i)}$$

$$+ Time_j + Diet \times Time_{ij} + \varepsilon_{k(ij)}$$

iid ? ⇒ maybe not indep. here!

Covariance Structure

Because of the lack of randomization at the split-plot level, each subject's measurements across time levels are dependent. A key feature of the model for repeated measures data is that we allow / account for this dependence by defining a covariance matrix for the multivariate response (across time) for each subject.

Here, think of each subject's error values ($\underline{\varepsilon}$) as a length-3 vector (for 3 time points):

$$\underline{\varepsilon} = \begin{bmatrix} \varepsilon_{15} \\ \varepsilon_{30} \\ \varepsilon_{45} \end{bmatrix} \qquad \text{Cov}[\underline{\varepsilon}] = \begin{bmatrix} Var(\varepsilon_{15}) & Cov(\varepsilon_{15},\varepsilon_{30}) & Cov(\varepsilon_{15},\varepsilon_{45}) \\ Cov(\varepsilon_{15},\varepsilon_{30}) & Var(\varepsilon_{30}) & Cov(\varepsilon_{30},\varepsilon_{45}) \\ Cov(\varepsilon_{15},\varepsilon_{45}) & Cov(\varepsilon_{30},\varepsilon_{45}) & Var(\varepsilon_{45}) \end{bmatrix}$$

We have previously assumed the ε's are iid N(0,σ^2), so Cov[ε]=σ^2 I. But now we need to specify dependence in a symmetric covariance matrix. Many possible structures exist. The table below shows just a few:

Description	Structure (Type = ...)	Example (Cov[$\underline{\varepsilon}$])
Compound Symmetry	CS	$\begin{bmatrix} \sigma^2+\varphi & \varphi & \varphi \\ \varphi & \sigma^2+\varphi & \varphi \\ \varphi & \varphi & \sigma^2+\varphi \end{bmatrix}$
First-Order Autoregressive	AR(1)	$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$
Unstructured	UN	$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}$
First-Order Autoregressive Moving-Average	ARMA(1,1)	$\sigma^2 \begin{bmatrix} 1 & \gamma & \gamma\rho \\ \gamma & 1 & \gamma \\ \gamma\rho & \gamma & 1 \end{bmatrix}$

How to choose a structure?

One reasonable approach is to obtain the AICC fit statistic (found in output) for several possible structures; smaller AICC values indicate more parsimonious models (better fit with fewer parameters).

```
/* STAT 5200
   repeated measures design
   glucose data
 */

/* Define options */
ods html image_dpi=300 style=journal;

/* Read in data */
data glucose;
 input Diet Subj Time Glucose @@;  cards;
1  1 15 22 1  1 30 34 1  1 45 32 1  2 15 15 1  2 30 29 1  2 45 27
1  3 15 12 1  3 30 33 1  3 45 28 1  4 15 21 1  4 30 40 1  4 45 39
2  5 15 22 2  5 30 18 2  5 45 12 2  6 15 23 2  6 30 22 2  6 45 10
2  7 15 18 2  7 30 16 2  7 45  9 2  8 15 25 2  8 30 24 2  8 45 15
3  9 15 31 3  9 30 30 3  9 45 39 3 10 15 28 3 10 30 27 3 10 45 36
3 11 15 24 3 11 30 26 3 11 45 36 3 12 15 21 3 12 30 26 3 12 45 32
;
run;

/* Fit Compound Symmetry Covariance Structure */

proc glimmix data=glucose plots=residualpanel;

  class Diet Time Subj;

  model Glucose = Diet | Time ;

  random intercept / Rside V
         subject=Subj(Diet) type=cs;
    /* 1. 'V' requests within-subject estimated
            covariance matrix of error terms
       2. 'intercept' here when no other
            random terms in model (besides subject)
       3. 'Rside' here specifies that dependence is among
            the error terms within subject
     */

  covtest 'Zero Covariance' diagR;
    /* Tests null that off-diagonal elements are all zero
       in covariance matrix of error terms
     */

  title1 'Compound Symmetry Covariance Structure';

run;
```

Compound Symmetry Covariance Structure

Convergence criterion (ABSGCONV=0.00001) satisfied.

Fit Statistics	
AICC (smaller is better)	148.37

Estimated V Matrix for Subj(Diet) 1 1

Row	Col1	Col2	Col3
1	14.8333	11.4167	11.4167
2	11.4167	14.8333	11.4167
3	11.4167	11.4167	14.8333

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error
CS	**Subj(Diet)**	11.4167	5.9309
Residual		3.4167	1.1389

$\hat{\phi}$

$\hat{\sigma}^2$

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Diet	2	9	12.78	0.0023
Time	2	18	27.96	<.0001
Diet*Time	4	18	66.98	<.0001

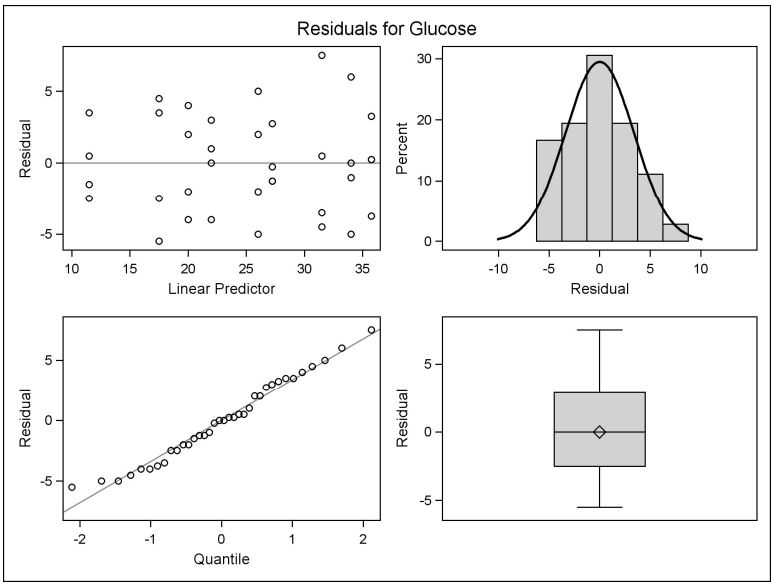
Tests of Covariance Parameters Based on the Restricted Likelihood

Label	DF	-2 Res Log Like	ChiSq	Pr > ChiSq	Note
Zero Covariance	1	161.92	18.04	<.0001	DF

DF: P-value based on a chi-square with DF degrees of freedom

H_0: error terms w/in subj. are indep.

↳ Need to acct. for dependence w/ type= ...

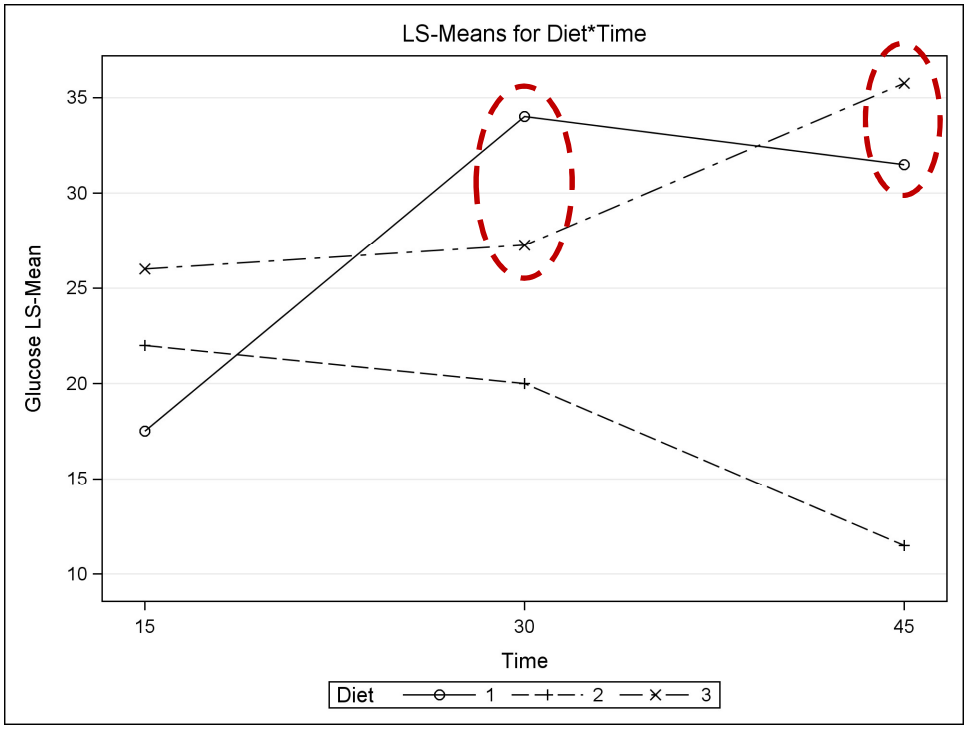



```
/* Fitting same model, look at
   interaction plot and contrast */
proc glimmix data=glucose;
  class Diet Time Subj;
  model Glucose = Diet | Time ;
  random intercept / Rside V
         subject=Subj(Diet) type=cs;

  lsmeans Diet*Time /
     pdiff=all adjust=Tukey lines
     plot=mean(sliceby=Diet join);
     /* Recall interaction plot code */

  contrast 'Last Contrast!'
    Time 0 2 -2
     Diet*Time  0  1 -1
                0  0  0
                0  1 -1;
  estimate 'Estimate of Last Contrast'
    Time 0 2 -2     Diet*Time  0  1 -1  0  0  0  0  1 -1;
     /* Averaging over over Diets 1 and 3 only,
       compare Time30 with Time45 */

run;
```



Tukey-Kramer Grouping for Diet*Time Least Squares Means (Alpha=0.05)						
LS-means with the same letter are not significantly different.						
Diet	**Time**	**Estimate**				
3	**45**	35.7500			A	
1	**30**	34.0000	B		A	
1	**45**	31.5000	B		A	C
3	**30**	27.2500	B		D	C
3	**15**	26.0000	B	E	D	C
2	**15**	22.0000		E	D	C
2	**30**	20.0000		E	D	
1	**15**	17.5000		E	F	
2	**45**	11.5000			F	

Estimates					
Label	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Estimate of Last Contrast	-6.0000	1.8484	18	-3.25	0.0045

Contrasts				
Label	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Last Contrast!	1	18	10.54	0.0045

One Last Contrast:

Looking at the interaction plot, it appears that maybe there really isn't a difference between Time30 and Time45 when averaging over Diets 1 and 3 only. As before, we can test this using a contrast:

Example 2: Crossover design

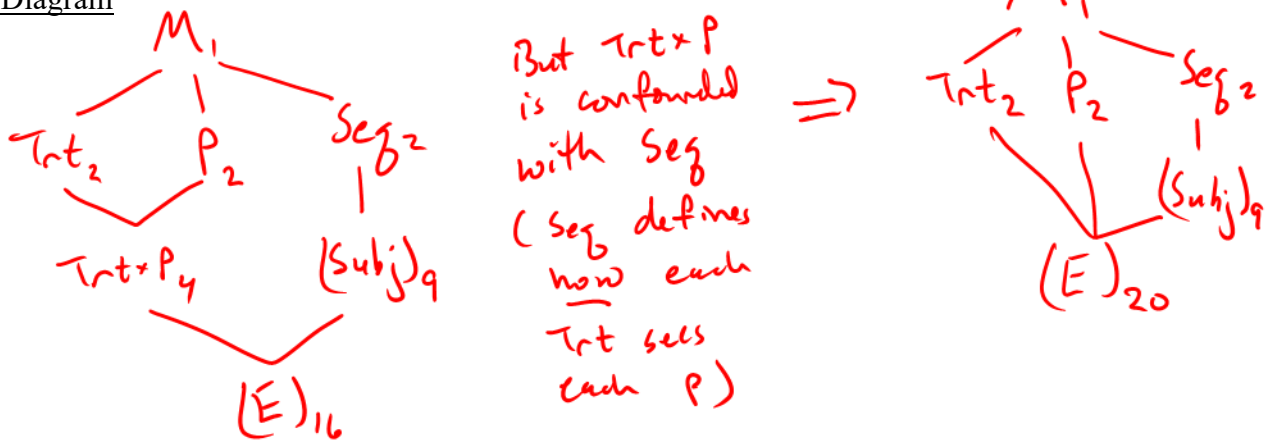
A study was conducted to compare the duration of effects of three different drug formulations (1 = 50-mg tablet; 2 = 100-mg tablet; 3 = sustained-release capsule) on lowering blood pressure. Twelve males volunteered to participate. In order to compare formulation effects within subject, each subject took each formulation, one in each of three time periods. To avoid a carryover effect of one formulation to the next, a one-week washout period was followed by each subject. In order to avoid confounding formulation effect with time period effect, subjects were randomly assigned to a certain sequence of taking the three formulations:

Formulation Sequence:		1-2-3				2-3-1				3-1-2			
Subject:		1	2	3	4	5	6	7	8	9	10	11	12
Time Period:	1	x	x	x	x	x	x	x	x	x	x	x	x
	2	x	x	x	x	x	x	x	x	x	x	x	x
	3	x	x	x	x	x	x	x	x	x	x	x	x

Note that this looks a lot like the repeated measures design in Example 1. It also bears similarity to a Latin Square design:

Sequence	Period 1	2	3	# Subjects
1-2-3	1	2	3	4
2-3-1	2	3	1	4
3-1-2	3	1	2	4

Hasse Diagram



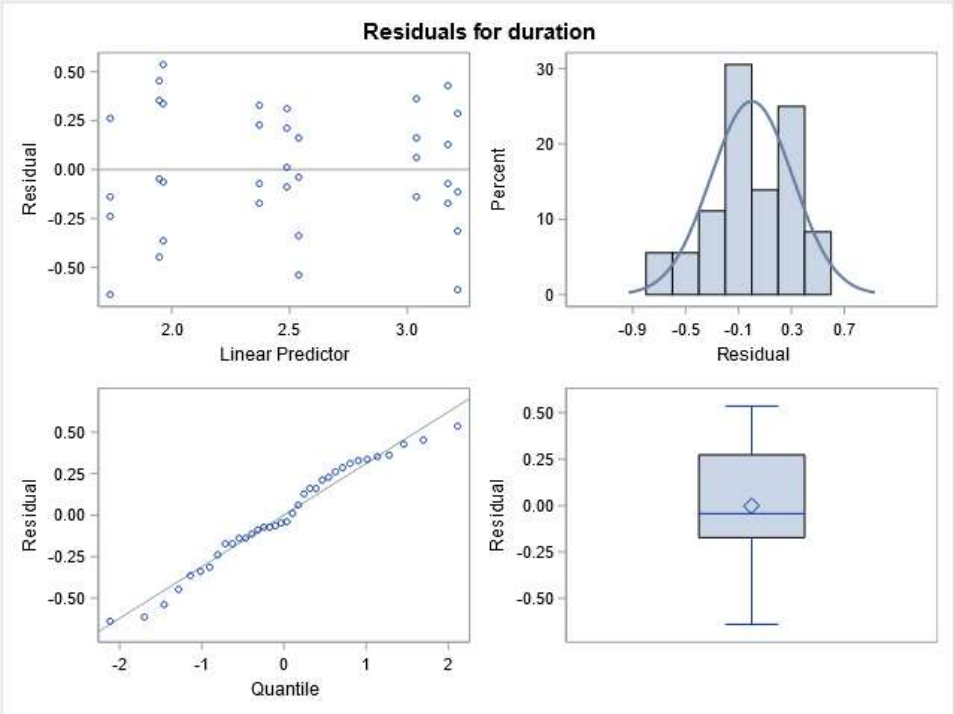
Model

$$Y_{ijkl} = \mu + Trt_i + P_j + \overbrace{TP_{ij}}^{\text{confounded}} + Seq_k + Subj_{l(k)} + \varepsilon_{ijl(k)}$$

$i=1..3 \quad j=1..3 \quad k=1..3 \quad l=1..4$

```
data drug; input Seq Subj P Trt duration @@; cards;
 1 1 1 1 1.5  1 1 2 2 2.2  1 1 3 3 3.4
 1 2 1 1 2.0  1 2 2 2 2.6  1 2 3 3 3.1
 1 3 1 1 1.6  1 3 2 2 2.7  1 3 3 3 3.2
 1 4 1 1 1.1  1 4 2 2 2.3  1 4 3 3 2.9
 2 5 1 2 2.5  2 5 2 3 3.5  2 5 3 1 1.9
 2 6 1 2 2.8  2 6 2 3 3.1  2 6 3 1 1.5
 2 7 1 2 2.7  2 7 2 3 2.9  2 7 3 1 2.4
 2 8 1 2 2.4  2 8 2 3 2.6  2 8 3 1 2.3
 3 9 1 3 3.3  3 9 2 1 1.9  3 9 3 2 2.7
 3 10 1 3 3.1 3 10 2 1 1.6 3 10 3 2 2.5
 3 11 1 3 3.6 3 11 2 1 2.3 3 11 3 2 2.2
 3 12 1 3 3.0 3 12 2 1 2.5 3 12 3 2 2.0
 ;

proc glimmix data=drug plots=residualpanel;
 class Seq Trt P Subj;
 model duration = Seq Trt P;
 random intercept / Rside subject=Subj(Seq) type=cs;
 covtest 'Zero covariance' diagR;
run;  /* Note no significant evidence of dependence here */
```



Tests of Covariance Parameters Based on the Restricted Likelihood

Label	DF	-2 Res Log Like	ChiSq	Pr > ChiSq	Note
Zero covariance	1	35.1743	1.03	0.3105	DF

DF: P-value based on a chi-square with DF degrees of freedom.

```
proc glimmix data=drug plots=residualpanel;
 class Seq Trt P Subj;
 model duration = Seq Trt P;
 random Subj(Seq);
 lsmeans Trt / pdiff=all adjust=tukey lines;
run;
```

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Subj(Seq)	0	.
Residual	0.1166	0.03062

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Seq	2	9	1.00	0.4044
Trt	2	20	40.81	<.0001
P	2	20	0.07	0.9291

Tukey Grouping for Trt Least Squares Means (Alpha=0.05)		
LS-means with the same letter are not significantly different.		
Trt	Estimate	
3	3.1417	A
2	2.4667	B
1	1.8833	C