# Some Messy Experimental Designs

Dallas E. Johnson Professor Emeritus Kansas State University

# Some Factors to Consider when Planning an Experiment

Question: Why should experimenters and/or researchers be interested in statistics?

# Some Factors to Consider when Planning an Experiment

Question: Why should experimenters and/or researchers be interested in statistics?

Answer: Reliable facts are necessary to build a correct science or theory. Reliable facts are often not easy to come by. Statistics and the Design of Experiments are concerned with procedures for obtaining reliable facts.

# Three Phases of Experimentation

- 1. Formulating the Experiment
- 2. Designing the Experiment
- 3. Analyzing the Experiment

# Three Phases of Experimentation

- Formulating the Experiment
- 2. Designing the Experiment
- 3. Analyzing the Experiment

For which phase is a statistician most useful?

## Designing the Experiment

- a) What are the experimental units?
- b) Are missing values likely? Why or why not?
- c) How many experimental units should be used?How many replications?
- d) How should the data be collected?
- e) Are there ways that the experimental units can be blocked?
- f) Should the experiment be done according to an experimental plan?
- g) How should the randomization of experimental units to treatments be done?
- h) What assumptions are reasonable?

### Formulating the Experiment

- a) What problems are to be solved?
- b) What is the experiment intended to show?
  What are the hypotheses?
- c) What variables are going to be measured?
- d) What factors are going to be varied?
  What is the treatment structure?
- e) Is the objective to show things are different or to show things are the same?

## **Designing the Experiment**

- i) How accurate is the experiment likely to be?
- j) What models are likely to be most meaningful for the experiment?
- k) What are problems that might occur?
- Can the data that is collected be analyzed?
- m) Will the data answer the important questions'
- n) Are there independent replications?

#### Remarks:

"Convenience" often destroys many experiments or limits the amount of information that can be gained from an experiment.

Convenience is often an enemy of good experimental design

# Analyzing the Experiment

- a) What statistical tests are likely to be appropriate?
- b) Can the design be analyzed by a statistical package? If so how should the data be organized for computer entry?
- c) Will one be able to communicate the results to others?
- Are the assumptions made likely to be satisfied
   Can they be tested?
  - Can the data still be analyzed if the assumptions are not satisfied?

## Selecting a Good Design

- a) Keep the design somewhat simple.
- Use blocking effectively. Smaller block size designs are more efficient than larger block size designs. Incomplete block designs are usually more efficient than complete block designs.
- c) Make sure the design can be analyzed. Can it still be analyzed if there is missing data?

## **Examples of Simple Designs**

- Completely Randomized Designs (CRDs)
- Randomized Complete Block Designs (RCBDs)
- Latin Square Designs
- Crossover Designs
- Graeco-Latin Square Designs
- Incomplete Block Designs
- Some Split-plot Designs
- Strip-plot Designs
- Many Confounded Designs

#### Some General Comments

- Think about experiments in terms of treatment structure and design structure. Almost any treatment structure can be existed with any design structure.
- Be sure to know and understand the difference between independent replications and dependent replications.
- The analysis of the design structure provides estimates of the different error variances that may occur in an experiment
- 4. The analysis of the treatment structure (i.e., the hypotheses to be tested) is similar for all design structures.
- It is usually no harder to conduct an experiment with a well designed plan than to conduct an experiment that is not well-designed.

# Independent/Dependent Replications

Suppose that we want to compare the heights of black males to that of white males.

- We select one male from each population, and have 10 people measure each person's height.
- Denote the data by  $x_{ij}$ , i=1,2; j=1,2,...,10
- Compute a t-statistic as  $t = \frac{\overline{x_1} \overline{x_2}}{s\sqrt{\frac{2}{10}}}$  where  $s = \sqrt{(s_1^2 + s_2^2)/2}$
- Reject a hypothesis of equality if |t| is large.

# Independent/Dependent Replications

Suppose that we want to compare the effects of storage temperature on the quality of CDs or DVDs.

- We randomly assign a high temperature and a low temperature to each of two storage chambers. Ten CDs are placed in each chamber. Three months later, the CDs are removed and measured for quality.
- Denote the data by  $x_{ii}$ , i=1,2; j=1,2,...,10.
- Compute a t-statistic as  $t = \frac{\overline{x_1} \overline{x_2}}{s\sqrt{\frac{2}{10}}}$  where  $s = \sqrt{(s_1^2 + s_2^2)/2}$
- Reject a hypothesis of equality if |t| is large.

## Example 1

An experiment is to be conducted to study the effects of baking temperatures and surfactants on bread quality.

Four different baking temperatures are to be used in conjuction with five different surfactants. Denote the temperatures by  $300^\circ, 350^\circ, 400^\circ,$  and  $450^\circ,$  and denote the surfactants by C,  $S_1,\,S_2,\,S_3,$  and  $S_4.$  Two loaves will be made for each of the 20 combinations of temperature\*surfactant combination.

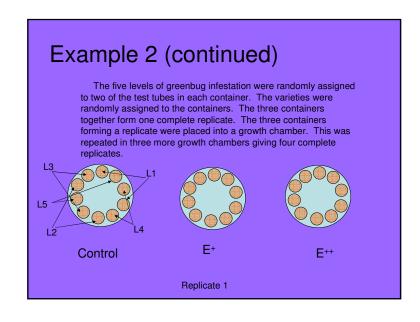
The treatment structure is a two-way structure.

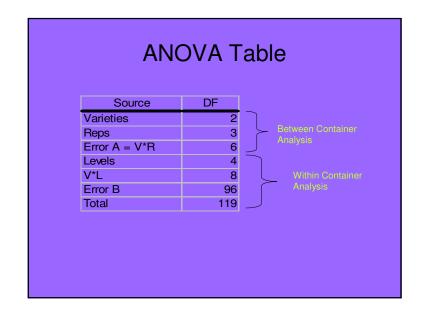
How should the experiment be conducted?

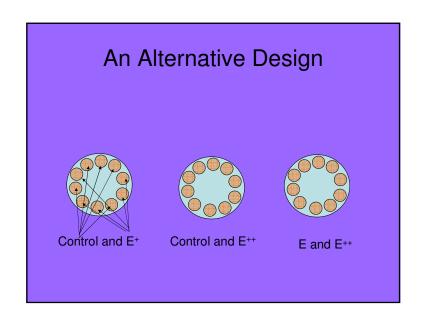
## Some Possible Designs

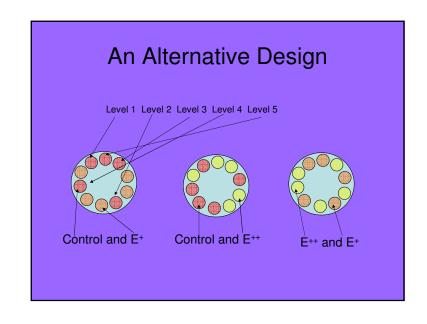
- 1. A CRD.
- 2. A RCBD
- A split-plot designs with temperatures randomly assigned to ovens and five loaves (one with each surfactant treatment) are placed simultaneously in the oven.
- A split-plot design with doughs mixed so that four loaves can be made from each dough, and one loaf from each dough is placed into a different temperature oven.
- 5. Many other designs are also possible.

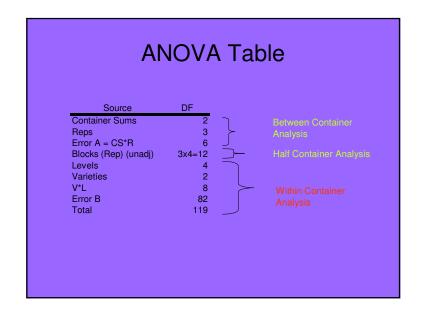
# Example 2 An experiment was to be conducted to study the effects of five different levels of greenbug infestation on three varieties of sorghum. Using large coffee cans, containers were made which would hold ten test tubes with each test tube containing a sorghum plant.

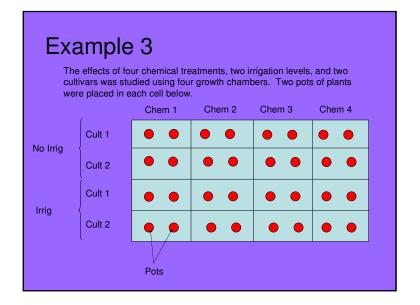


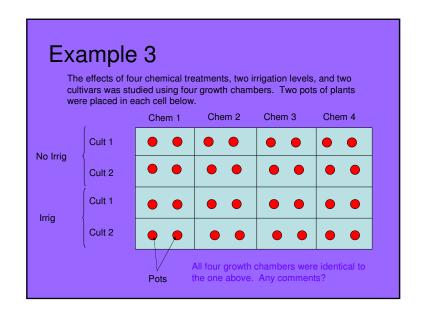


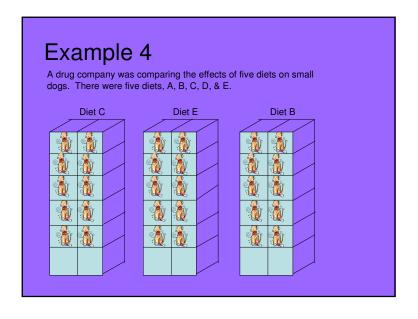


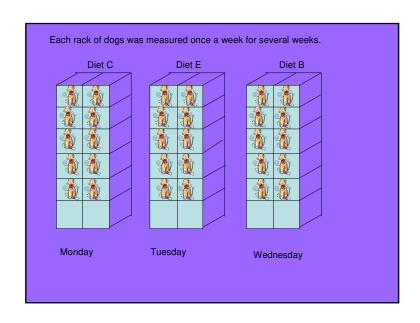


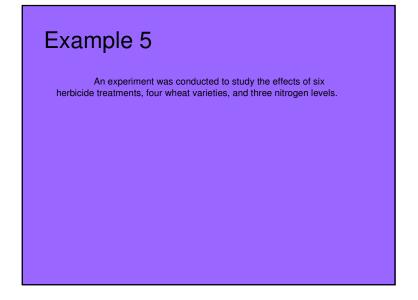










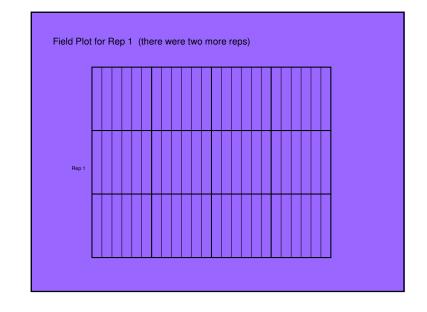


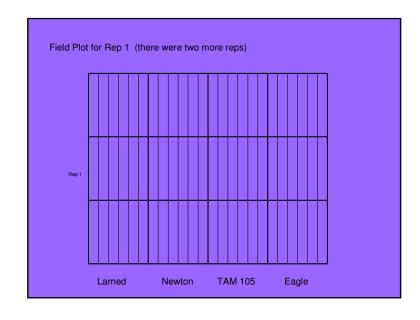
# Example 5

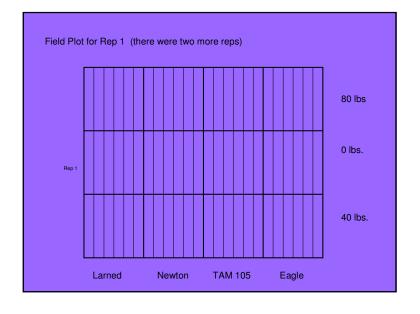
H6: Control

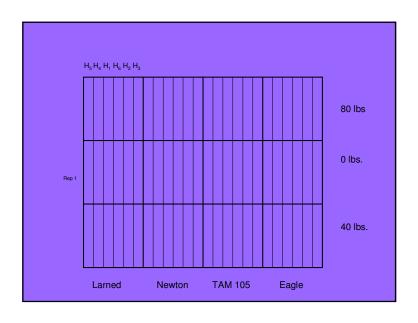
An experiment was conducted to study the effects of six herbicide treatments, four wheat varieties, and three nitrogen levels.

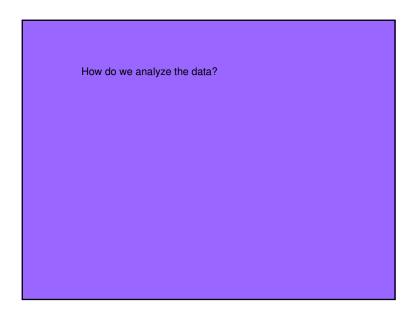
<u>Herbicides</u>	Wheat Cultivars	Nitrogen Rate
H1: 2,4-D ester	Newton	0 pounds
H2: Glean 75 DF	TAM 105	40 pounds
H3: Surflan 4 AS	Eagle	80 pounds
H4: 2,4-D + Surflan	Larned	
H5: Glean + Surflan		

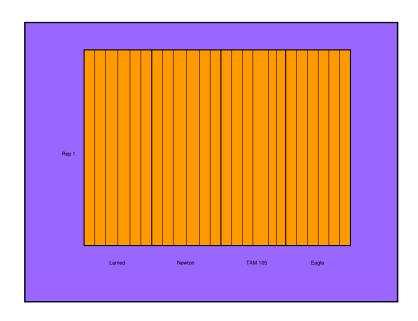


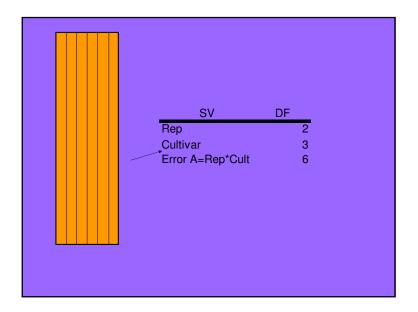


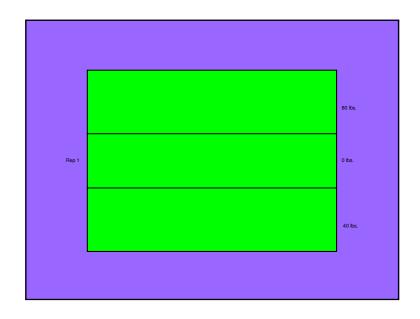


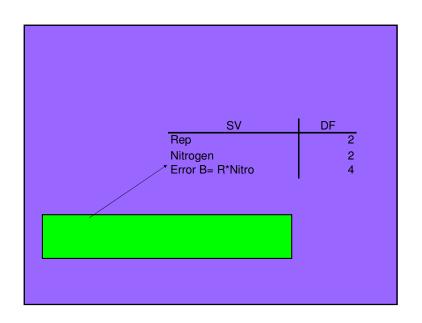




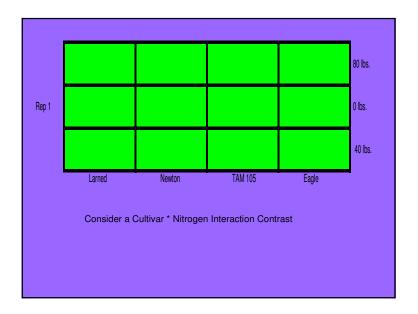


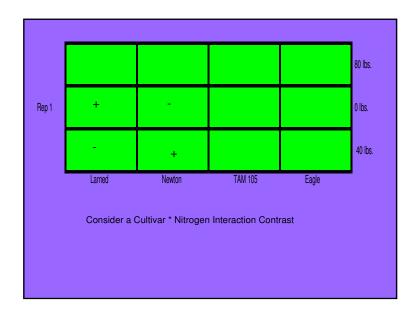


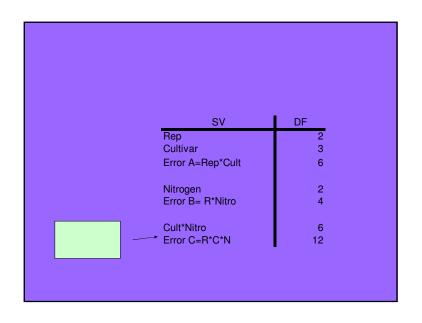


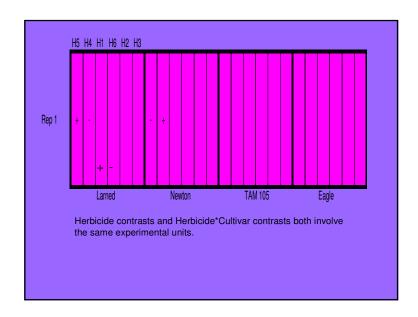


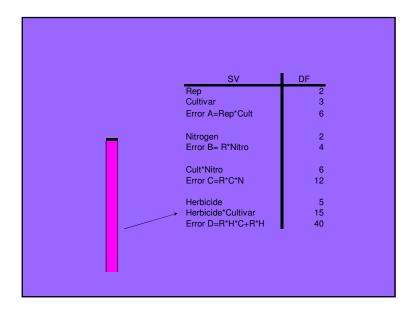
SV	DF
Rep	2
Cultivar	3
Error A=Rep*Cult	6
Nitrogen	2
Error B= R*Nitro	4

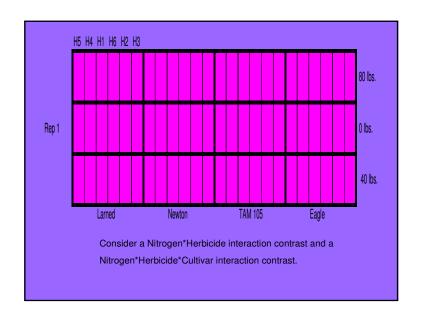


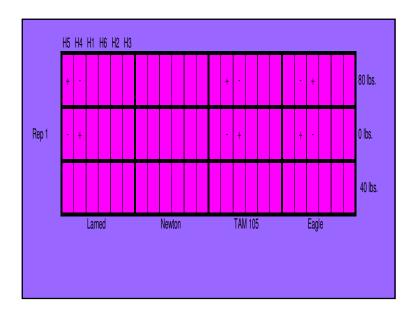




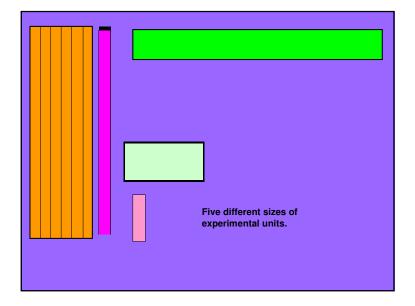


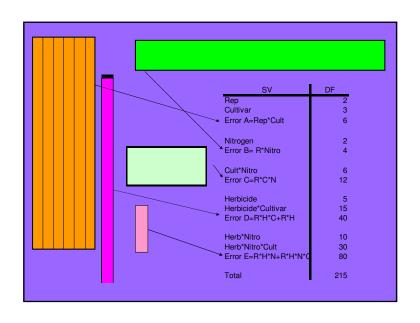


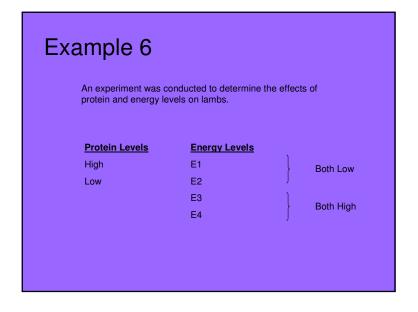




SV	DF
Rep	2
Cultivar	3
Error A=Rep*Cult	6
Nitrogen	2
Error B= R*Nitro	4
Cult*Nitro	6
Error C=R*C*N	12
Herbicide	5
Herbicide*Cultivar	15
Error D=R*H*C+R*H	40
Herb*Nitro	10
Herb*Nitro*Cult	30
Error E=R*H*N+R*H*N*C	80







Each of the eight diet combinations were randomly assigned to two lambs selected at random from a group of 16 lambs. Data was collected for a week, then each lamb that received E1 during the first week, received E2 during a second week. Those getting E2 the first week, received E1 the second week. Similarly, each lamb that received E3 during the first week, received E4 during the second week. Those getting E4 the first week, received E3 the second week.

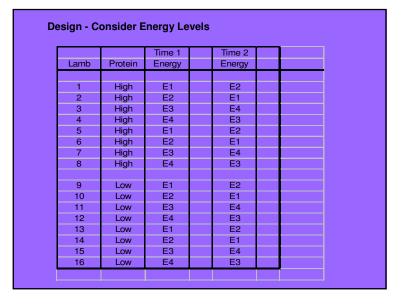
How do we analyze this data?

		Time 1	Time 2		
Lamb	Protein	Energy	Energy		
1	High	E1	E2		
2	High	E2	E1		
3	High	E3	E4		
4	High	E4	E3		
5	High	E1	E2		
6	High	E2	E1		
7	High	E3	E4		
8	High	E4	E3		
9	Low	E1	E2		
10	Low	E2	E1		
11	Low	E3	E4		
12	Low	E4	E3		
13	Low	E1	E2		
14	Low	E2	E1		
15	Low	E3	E4		
16	Low	E4	E3		

		Time 1	Time 2		
Lamb	Protein	Energy	Energy		
1	High	E1	E2		
2	High	E2	E1		
3	High	E3	E4		
4	High	E4	E3		
5	High	E1	E2		
6	High	E2	E1		
7	High	E3	E4		
8	High	E4	E3		
9	Low	E1	E2		
10	Low	E2	E1		
11	Low	E3	E4		
12	Low	E4	E3		
13	Low	E1	E2		
14	Low	E2	E1		
15	Low	E3	E4		
16	Low	E4	E3		

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	+	
2	High	E2	+	E1	+	
3	High	E3	+	E4	+	
4	High	E4	+	E3	+	
5	High	E1	+	E2	+	
6	High	E2	+	E1	+	
7	High	E3	+	E4	+	
8	High	E4	+	E3	+	
9	Low	E1	-	E2	-	
10	Low	E2	-	E1	-	
11	Low	E3	-	E4	-	
12	Low	E4	-	E3	-	
13	Low	E1	-	E2	-	
14	Low	E2	-	E1	-	
15	Low	E3	-	E4	-	
16	Low	E4	-	E3	-	

		Time a 1	<u> </u>	Time o	_	1
Lausela	Duetein	Time 1		Time 2		
Lamb	Protein	Energy		Energy		
	1			<b>50</b>		
1	High	E1	+	E2	+	
2	High	E2	+	E1	+	
3	High	E3	+	E4	+	
4	High	E4	+	E3	+	
5	High	E1	+	E2	+	
6	High	E2	+	E1	+	
7	High	E3	+	E4	+	
8	High	E4	+	E3	+	
9	Low	E1	-	E2	-	
10	Low	E2	-	E1	-	
11	Low	E3	-	E4	-	
12	Low	E4	-	E3	-	
13	Low	E1	-	E2	-	
14	Low	E2	-	E1	-	
15	Low	E3	-	E4	-	
16	Low	E4	_	E3	_	



E1 vs E2 (Within	Lamb Contrast	) - E3	vs E4	is similar
------------------	---------------	--------	-------	------------

Lamb	Protein	Time 1 Energy		Time 2 Energy		
1	High	E1	+	E2	-	
2	High	E2	-	E1	+	
3	High	E3	0	E4	0	
4	High	E4	0	E3	0	
5	High	E1	+	E2	-	
6	High	E2	-	E1	+	
7	High	E3	0	E4	0	
8	High	E4	0	E3	0	
9	Low	E1	+	E2	-	
10	Low	E2	-	E1	+	
11	Low	E3	0	E4	0	
12	Low	E4	0	E3	0	
13	Low	E1	+	E2	-	
14	Low	E2	-	E1	+	
15	Low	E3	0	E4	0	
16	Low	E4	0	E3	0	

#### E1+E2 vs E3+E4 ( Between Lamb Contrast)

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
Lamb	TIOLOIT	Lifelgy		Lifeigy		
1	High	E1	+	E2	+	
2	High	E2	+	E1	+	
3	High	E3	-	E4	-	
4	High	E4	-	E3	-	
5	High	E1	+	E2	+	
6	High	E2	+	E1	+	
7	High	E3	-	E4	-	
8	High	E4	-	E3	-	
9	Low	E1	+	E2	+	
10	Low	E2	+	E1	+	
11	Low	E3	-	E4	-	
12	Low	E4	-	E3	-	
13	Low	E1	+	E2	+	
14	Low	E2	+	E1	+	
15	Low	E3	-	E4	-	
16	Low	E4	-	E3	-	

Time Contrast (Within Lamb Contrast)

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	-	
2	High	E2	+	E1	-	
3	High	E3	+	E4	-	
4	High	E4	+	E3	-	
5	High	E1	+	E2	-	
6	High	E2	+	E1	-	
7	High	E3	+	E4	-	
8	High	E4	+	E3	-	
9	Low	E1	+	E2	-	
10	Low	E2	+	E1	-	
11	Low	E3	+	E4	-	
12	Low	E4	+	E3	-	
13	Low	E1	+	E2	-	
14	Low	E2	+	E1	-	
15	Low	E3	+	E4	-	
16	Low	E4	+	E3	-	

Protein x LE ( Within Lamb Contrast) - Protein x HE is similar

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	-	
2	High	E2	-	E1	+	
3	High	E3		E4		
4	High	E4		E3		
5	High	E1	+	E2	-	
6	High	E2	-	E1	+	
7	High	E3		E4		
8	High	E4		E3		
9	Low	E1	-	E2	+	
10	Low	E2	+	E1	-	
11	Low	E3		E4		
12	Low	E4		E3		
13	Low	E1	-	E2	+	
14	Low	E2	+	E1	-	
15	Low	E3		E4		
16	Low	E4		E3		

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	+	
2	High	E2	+	E1	+	
3	High	E3	-	E4	-	
4	High	E4	-	E3	-	
5	High	E1	+	E2	+	
6	High	E2	+	E1	+	
7	High	E3	-	E4	-	
8	High	E4	-	E3	-	
9	Low	E1	-	E2	-	
10	Low	E2	-	E1	-	
11	Low	E3	+	E4	+	
12	Low	E4	+	E3	+	
13	Low	E1	-	E2	-	
14	Low	E2	-	E1	-	
15	Low	E3	+	E4	+	
16	Low	E4	+	E3	+	

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	-	
2	High	E2	+	E1	-	
3	High	E3	+	E4	-	
4	High	E4	+	E3	-	
5	High	E1	+	E2	-	
6	High	E2	+	E1	-	
7	High	E3	+	E4	-	
8	High	E4	+	E3	-	
9	Low	E1	-	E2	+	
10	Low	E2	-	E1	+	
11	Low	E3	-	E4	+	
12	Low	E4	-	E3	+	
13	Low	E1	-	E2	+	
14	Low	E2	-	E1	+	
15	Low	E3	-	E4	+	
16	Low	E4	-	E3	+	

	_					
		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	+	
2	High	E2	-	E1	-	
3	High	E3		E4		
4	High	E4		E3		
5	High	E1	+	E2	+	
6	High	E2	-	E1	-	
7	High	E3		E4		
8	High	E4		E3		
9	Low	E1	+	E2	+	
10	Low	E2	-	E1	-	
11	Low	E3		E4		
12	Low	E4		E3		
13	Low	E1	+	E2	+	
14	Low	E2	-	E1	-	
15	Low	E3		E4		
16	Low	E4		E3		

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	-	
2	High	E2	+	E1	-	
3	High	E3	-	E4	+	
4	High	E4	-	E3	+	
5	High	E1	+	E2	-	
6	High	E2	+	E1	-	
7	High	E3	-	E4	+	
8	High	E4	-	E3	+	
9	Low	E1	+	E2	-	
10	Low	E2	+	E1	-	
11	Low	E3	-	E4	+	
12	Low	E4	-	E3	+	
13	Low	E1	+	E2	-	
14	Low	E2	+	E1	-	
15	Low	E3	-	E4	+	
16	Low	E4	-	E3	+	

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	+	
2	High	E2	-	E1	-	
3	High	E3		E4		
4	High	E4		E3		
5	High	E1	+	E2	+	
6	High	E2	-	E1	-	
7	High	E3		E4		
8	High	E4		E3		
9	Low	E1	-	E2	-	
10	Low	E2	+	E1	+	
11	Low	E3		E4		
12	Low	E4		E3		
13	Low	E1	-	E2	-	
14	Low	E2	+	E1	+	
15	Low	E3		E4		
16	Low	E4		E3		

SV	DF	Ì	
Total	31		
Lambs	15		
Protein	1		
CE	1		
Protein x CE	1	Betwee	en Lamb
LE x Time	1	Analy	
HE x Time	1	7	
Protein x LE x Time	1		
Protein x HE x Time	1		
Error A	8		
LE	1		
HE	1		
Time	1		
Protein x LE	1		
Protein x HE	1		in Lamb
Protein x Time	1	Anal	ysis
CE x Time	1		
Protein x CE x Time	1		
Error B	8		

#### Protein x CE x Time (Within Lamb Conrast)

		Time 1		Time 2		
Lamb	Protein	Energy		Energy		
1	High	E1	+	E2	-	
2	High	E2	+	E1	-	
3	High	E3	-	E4	+	
4	High	E4	-	E3	+	
5	High	E1	+	E2	-	
6	High	E2	+	E1	-	
7	High	E3	-	E4	+	
8	High	E4	-	E3	+	
9	Low	E1	-	E2	+	
10	Low	E2	-	E1	+	
11	Low	E3	+	E4	-	
12	Low	E4	+	E3	-	
13	Low	E1	-	E2	+	
14	Low	E2	-	E1	+	
15	Low	E3	+	E4	-	
16	Low	E4	+	E3	-	

## Example 7

An experiment was being conducted to compare three treatments on pigs. The treatments were A: uninfested, B: infested and treated, and C: infested and untreated.

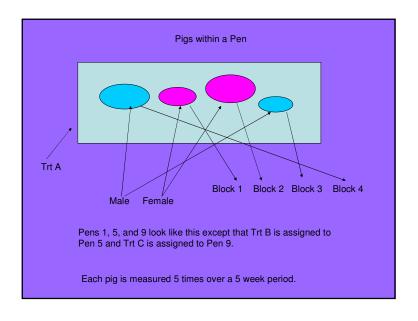
The experiment was conducted on 24 male and 24 female pigs. The 24 pigs of each gender were assigned to blocks according to weight. There were three pigs in each block. That is, there are 8 blocks of male pigs and 8 blocks of female pigs.

Two pigs of each gender which have been assigned the same treatment are placed in the same pen. Thus, there are 12 pens each containing four pigs, two males and two females.

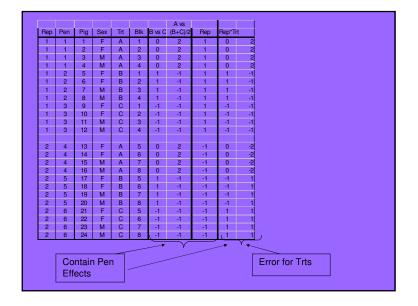
Because of the nature of the treatments, the pigs that are assigned to the same pen must all receive the same treatment.

Finally, measurements are taken on each pig weekly for five weeks. How should the resulting data be analyzed?

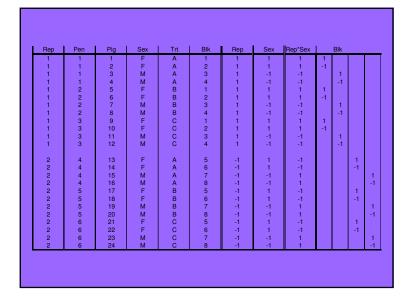
The fixed effects are Trt and Sex. Blocks and Rep are random effects.

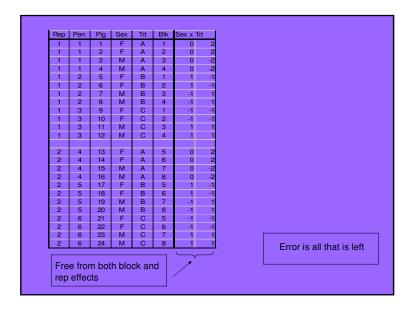


Rep Pen Pig Sex Trt Blk F M F F M M F Reps 1 and 2 В 0000 2 M M 3 11 12 13 2 2 2 2 2 2 2 2 2 2 14 F M F F M 15 16 4 5 5 5 5 17 В 18 19 В 20 21 22 23 24 В CCC 6 5 F M M

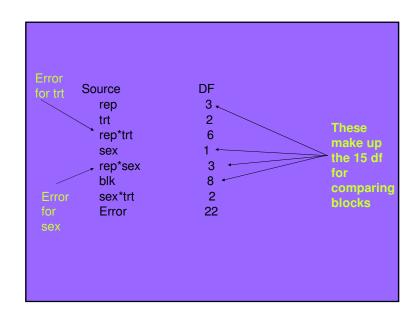


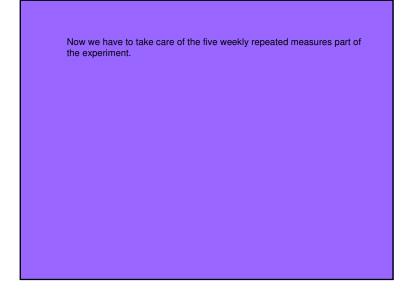
Rep	Pen	Pig	Sex	Trt	Blk	Rep	Sex	Rep*Sex				
_	1	Ů	F						_			4
1_	1	1 2	F	A	1	1	1	1	$\rightarrow$			
1				A	2	1	-1	-1	+			
1	1	3	M	A	3	1	-1	-1 -1	-			
1	2	5	F	В	4	1	1	1	+			1
1	2	6	F	В	2		1	1	-			
1	2	7	M	В	3	1	-1	-1	+			
1	2	8	M	В	4	1	-1	-1 -1	+			
1	3	9	F	С	1	1	1	1				
1	3	10	F	C	2	1	1	1	+			
1	3	11	M	C	3	1	-1	-1	+			
+	3	12	M	C	4	1	-1	-1	+			
-	3	12	IVI	U	4	- '	-1	-1	+			
2	4	13	F	Α	5	-1	1	-1	-			
2	4	14	F	A	6	-1	1	-1				Н
2	4	15	M	A	7	-1	-1	1	-			
2	4	16	M	A	8	-1	-1	1	-			Н
2	5	17	F	В	5	-1	1	-1	-			
2	5	18	F	В	6	-1	1	-1	-+			
2	5	19	M	В	7	-1	-1	1				
2	5	20	M	В	8	-1	-1	1	-			
2	6	21	F	С	5	-1	1	-1				
2	6	22	F	C	6	-1	1	-1				
2	6	23	M	C	7	-1	-1	1				
						Cont	†	Effects				





proc mixed;
class rep sex trt blk;
model y=trt sex sex\*trt;
random rep rep\*trt rep\*sex blk;
run;





Now we have to take care of the five weekly repeated measures part of the experiment.

This is your assignment. Send me your answers by 10 p.m. tonight!