

Homework 05
Joseph Blubaugh
jblubau1@tamu.edu
STAT 642-720

1.

- a) Interactions are comparing all levels from both A and B and so the degrees of freedom must be -1 for each parameter to avoid being over parameterized.

$$\begin{aligned} DF_{A*B} &= ab - N_{H_0} \\ &= ab - [b + (a - 1)] \\ &= (a - 1)(b - 1) \end{aligned}$$

$$\begin{aligned} \mu_{a,j} &= B \\ \mu_{a,j} - \mu_{a-1,j} &= A \\ N_{H_0} &= \mu_{a,j} + (\mu_{a,j} - \mu_{a-1,j}) \\ &= B + A \end{aligned}$$

- b) In the main effect, A is measured over all of the levels of B so there must be -1 + A to avoid being over parameterized

$$\begin{aligned} DF_A &= ab - N_{H_0} \\ &= ab - [b(a - 1) + 1] \\ &= a - 1 \end{aligned}$$

$$\begin{aligned} c &= \frac{1}{b} \sum_{j=1}^b \mu_{i,j} \\ N_{H_0} &= a(b - 1) + c \end{aligned}$$

2. if this, no interaction

- a) $y_{i,j,k} = \mu + \tau_i + \beta_j + (\tau\beta)_{i,j} + e_{i,j,k}$ where ...

μ is the overall mean

τ is soil type

β is field

$e_{i,j,k}$ are the errors associated with each k

$i = 1, \dots, 3$

$j = 1, \dots, 30$

$k = 1, \dots, 44$

b) Fitting a mixed model with Soil as the fixed variable and Field as the random variable

Linear mixed model fit by REML ['lmerMod']

Formula: Porosity ~ Soil + (1 | Field)

Data: dt

REML criterion at convergence: 93.8

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.42650	-0.78839	0.05287	0.60976	1.46805

Random effects:

Groups	Name	Variance	Std.Dev.
Field	(Intercept)	0.1780	0.4219
Residual		0.3244	0.5696

Number of obs: 44, groups: Field, 30

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	4.2078	0.2014	20.891
SoilL	0.6207	0.2877	2.158
SoilS	2.1245	0.2848	7.458

Correlation of Fixed Effects:

	(Intr) SoilL
SoilL	-0.700
SoilS	-0.707 0.495

Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value
Soil	2	19.059	9.5295	29.374

c)

```
cat("Soil:", percent(9.5295 / (9.5295 + 0.1779996 + 0.3136)))
```

Soil: 95.1%

```
cat("Field:", percent(0.1779996 / (36.67 + 0.1779996 + 0.3136)))
```

Field: 0.479%

d) Both Sandy-Clay and Sandy-Loam can be grouped based TukeyHSD.

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: lmer(formula = Porosity ~ Soil + (1 | Field), data = dt)

Linear Hypotheses:

	Estimate	Std. Error	z value	Pr(> z)
L - C == 0	0.6207	0.2877	2.158	0.0785 .
S - C == 0	2.1245	0.2848	7.458	<1e-04 ***
S - L == 0	1.5037	0.2877	5.227	<1e-04 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Adjusted p values reported -- single-step method)

e) Using the covariance parameter estimates we can reject the null and say there is sufficient evidence to show that the variance of field is greater than 0. Covariance Parameter Estimates

Cov	Param Estimate	Alpha	Lower	Upper
Field	0.1780	0.05	0.06018	1.9121
Residual	0.3244	0.05	0.1771	0.7761

$$H_0 : \sigma = 0, H_a : \sigma > 0$$

$$F* = \frac{MS_{TRT}}{MSE}$$

$$= \frac{0.178}{.3244}$$

$$=.5487053$$

$$pf(.5487053, 29, 14) = 0.083$$

3.

a) $y_{i,j,k} = \mu + \tau_i + \beta_j + (\tau\beta)_{i,j} + e_{i,j,k}$ where ...
 μ is the overall mean in Reduction

τ is Nitrogen Level

β is Crop Type

$(\tau\beta)$ is the interaction between Nitrogen and Crop Type $e_{i,j,k}$ are the errors associated with each k

$i = 1, \dots, 3$ 3 different levels of Nitrogen

$j = 1, \dots, 4$ 4 different levels of Crops

$k = 1, \dots, 48$ 48 different combinations of Nitrogen, Crops, and Tanks

b) Conditions:

- Homogeneity of variance: Variance is not stable, consider a transformation

Levene's Test for Homogeneity of Variance (center = median)

```
      Df F value  Pr(>F)
group 11    2.178 0.03892 *
      36
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- Data are normally distributed: Data are not normally distributed, consider a transformation

Shapiro-Wilk normality test

data: crops\$Reduction

W = 0.71668, p-value = 2.647e-08

- Completely Randomized Data: Based on the description of the experiment Each EU was randomly assigned to a level of treatment and this condition is satisfied

c)

```
mdl = lm(Reduction ~ Nitrogen*Crop, data = crops)
anova(mdl)
```

Analysis of Variance Table

Response: Reduction

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Nitrogen	2	17.300	8.6502	21.159	8.393e-07 ***
Crop	3	14.107	4.7024	11.502	1.952e-05 ***
Nitrogen:Crop	6	26.445	4.4074	10.781	7.584e-07 ***
Residuals	36	14.717	0.4088		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

d) Based on an alpha of .05, the following Contrasts are measured to be significant:

- Linear Trend in Nitrogen
- Quadratic Trend in Nitrogen
- Quadratic Trend in Crop
- Interactions 1,2,4
- Main effect of Nitrogen
- Main effect of Crop
- Overall Interaction of Nitrogen and Crop

ANALYSIS OF FACTORIAL TREATMENT STRUCTURE (SAS OUTPUT)

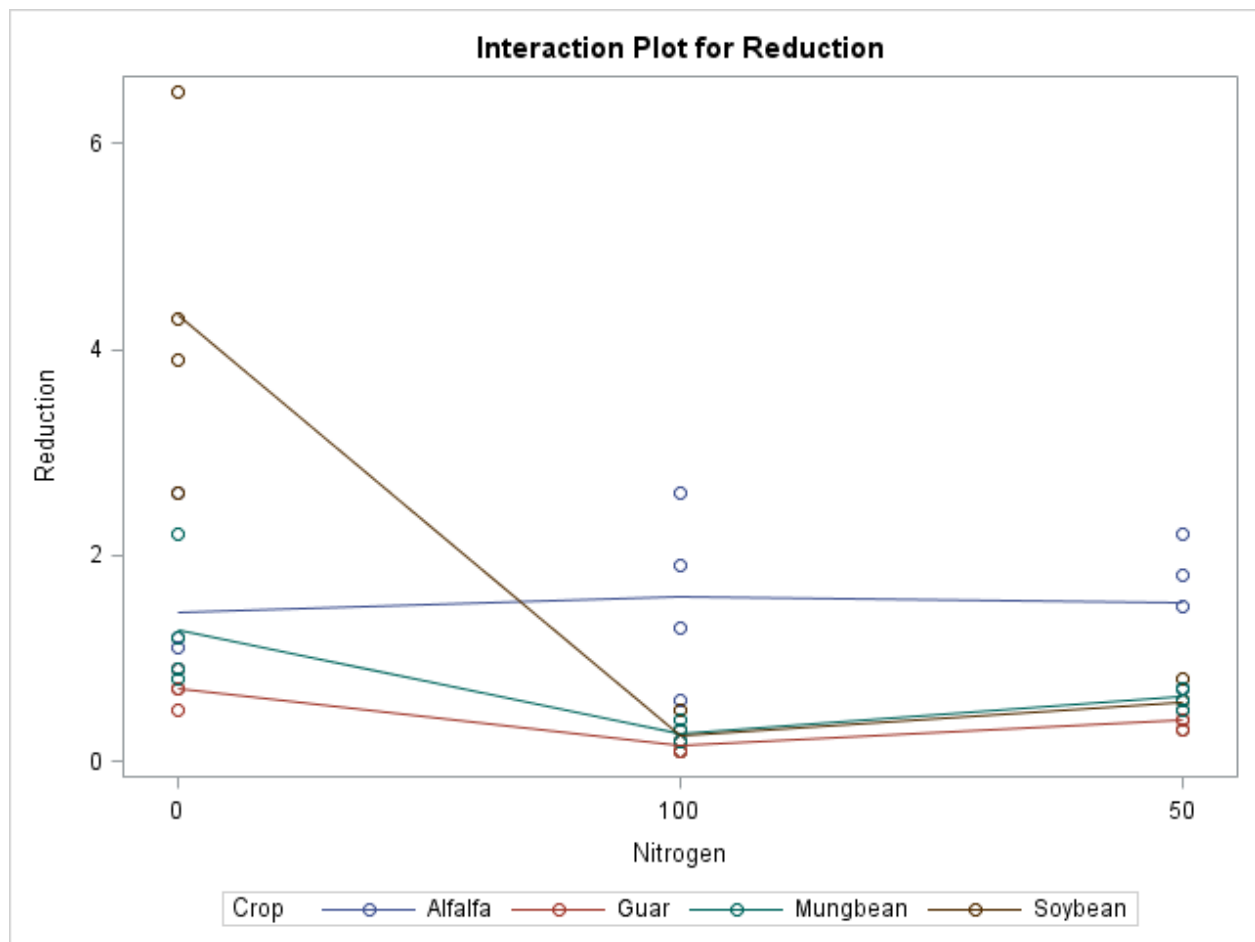
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	57.85229167	5.25929924	12.86	<.0001
Error	36	14.71750000	0.40881944		
Corrected Total	47	72.56979167			

R-Square	Coeff Var	Root MSE
0.797195	58.23665	0.639390

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Nitrogen	2	17.30041667	8.65020833	21.16	<.0001
Crop	3	14.10729167	4.70243056	11.50	<.0001
Nitrogen:Crop	6	26.44458333	4.40743056	10.78	<.0001

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Linear Trend in Nitrogen	1	10.58000000	10.58000000	25.88	<.0001
Quad Trend in Nitrogen	1	6.72041667	6.72041667	16.44	0.0003
Linear Trend in Crop	1	0.44204167	0.44204167	1.08	0.3053
Quad Trend in Crop	1	13.33520833	13.33520833	32.62	<.0001
Cubin Trend in Crop	1	0.33004167	0.33004167	0.81	0.3749
INTER1	1	14.16100000	14.16100000	34.64	<.0001
INTER2	1	3.64500000	3.64500000	8.92	0.0051
INTER3	1	0.78400000	0.78400000	1.92	0.1746
INTER4	1	6.86408333	6.86408333	16.79	0.0002
INTER5	1	0.70041667	0.70041667	1.71	0.1989
INTER6	1	0.29008333	0.29008333	0.71	0.4052
MAIN OF Nitrogen	2	17.30041667	8.65020833	21.16	<.0001
Main of Crop	3	14.10729167	4.70243056	11.50	<.0001
INTERACTION	6	26.44458333	4.40743056	10.78	<.0001

- e) There is a clear interaction between Nitrogen and Crop which are visible in the plot through non-constant distance between lines and lines that cross.



- f) Group 1: (Alfalfa, Soybean), Group 2: (Guar, Mungbean)

The GLM Procedure

Least Squares Means

Adjustment for Multiple Comparisons: Tukey

Crop Reduction	LSMEAN	Standard Error	Pr > t	LSMEAN Number
Alfalfa	1.53333333	0.18457597	<.0001	1
Guar	0.41666667	0.18457597	0.0301	2
Mungbean	0.72500000	0.18457597	0.0004	3
Soybean	1.71666667	0.18457597	<.0001	4

Least Squares Means for effect Crop

Pr > |t| for H0: LSMEAN(i)=LSMEAN(j)

Dependent Variable: Reduction

i/j	1	2	3	4
1		0.0007	0.0189	0.8955
2	0.0007		0.6424	<.0001
3	0.0189	0.6424		0.0029
4	0.8955	<.0001	0.0029	

4.

LS Means for Effects Model

Nitrogen	Reduction LSMEAN	Standard Error	Pr > t	LSMEAN Number
0	1.93750000	0.15984748	<.0001	1
100	0.56875000	0.15984748	0.0011	2
50	0.78750000	0.15984748	<.0001	3

Crop	Reduction LSMEAN	Standard Error	Pr > t	LSMEAN Number
Alfalfa	1.53333333	0.18457597	<.0001	1
Guar	0.41666667	0.18457597	0.0301	2
Mungbean	0.72500000	0.18457597	0.0004	3
Soybean	1.71666667	0.18457597	<.0001	4

Nitrogen	Crop	Reduction LSMEAN	Standard Error	Pr > t	LSMEAN Number
0	Alfalfa	1.45000000	0.31969495	<.0001	1
0	Guar	0.70000000	0.31969495	0.0351	2
0	Mungbean	1.27500000	0.31969495	0.0003	3
0	Soybean	4.32500000	0.31969495	<.0001	4
100	Alfalfa	1.60000000	0.31969495	<.0001	5
100	Guar	0.15000000	0.31969495	0.6418	6
100	Mungbean	0.27500000	0.31969495	0.3954	7
100	Soybean	0.25000000	0.31969495	0.4393	8
50	Alfalfa	1.55000000	0.31969495	<.0001	9
50	Guar	0.40000000	0.31969495	0.2189	10
50	Mungbean	0.62500000	0.31969495	0.0584	11
50	Soybean	0.57500000	0.31969495	0.0805	12

Call:


```
lm(formula = Reduction ~ 0 + Nitrogen + Crop, data = crops)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-1.4729	-0.6703	0.0948	0.2943	3.9437

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
Nitrogen0	2.3729	0.3500	6.780	3.01e-08 ***
Nitrogen50	1.2229	0.3500	3.494	0.00114 **
Nitrogen100	1.0042	0.3500	2.869	0.00642 **
CropGuar	-1.1167	0.4042	-2.763	0.00846 **
CropMungbean	-0.8083	0.4042	-2.000	0.05199 .
CropSoybean	0.1833	0.4042	0.454	0.65244

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.99 on 42 degrees of freedom

Multiple R-squared: 0.6844, Adjusted R-squared: 0.6393

F-statistic: 15.18 on 6 and 42 DF, p-value: 3.745e-09

a) Cell Means Model: $y_{i,j,k} = \mu_{i,j} + e_{i,j,k}$

$y_{i,j,k}$: Mean Reduction response for Nitrogen and Crop treatments

$\mu_{i,j}$: mean of Reduction for each combination of treatments $e_{i,j,k}$: error between the mean response and the overall mean for each treatment

$i_{1,...,3}$: 3 levels of Nitrogen

$j_{1,...,4}$: 4 levels of Crop

$k_{1,...,48}$: 48 individual observations (3*4*4)

b) Effects Model: $y_{i,j,k} = \mu + \tau_i + \beta_j + (\tau\beta)_{i,j} + e_{i,j,k}$ $y_{i,j,k}$: Mean Reduction response for Nitrogen and Crop treatments

μ : the overall mean

τ_i : effect of Nitrogen

β_j : effect of Crop

$(\tau\beta)_{i,j}$: interaction of Nitrogen and Crop

$e_{i,j,k}$: error between the mean response and the overall mean for each treatment

$i_{1,...,3}$: 3 levels of Nitrogen

$j_{1,...,4}$: 4 levels of Crop

$k_{1,...,48}$: 48 individual observations (3*4*4)

c) Design Matrix for Cell Means Model:

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \beta = \begin{bmatrix} \mu_{1,1} = 1.735 \\ \mu_{1,2} = 1.177 \\ \mu_{1,3} = 1.331 \\ \mu_{2,1} = 1.826 \\ \mu_{2,2} = 1.050 \\ \mu_{2,3} = 0.492 \\ \mu_{3,1} = 0.646 \\ \mu_{3,2} = 1.142 \\ \mu_{3,3} = 1.160 \\ \mu_{4,1} = 0.602 \\ \mu_{4,2} = 0.756 \\ \mu_{4,3} = 1.251 \end{bmatrix}$$

d) Design Matrix for Effects Model:

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \beta = \begin{bmatrix} \mu = 1.098 \\ \tau_1 = 1.937 \\ \tau_2 = .568 \\ \tau_3 = .787 \\ \beta_1 = 1.533 \\ \beta_2 = .416 \\ \beta_3 = .725 \\ \beta_4 = 1.716 \\ (\tau\beta)_{1,1} = 1.450 \\ (\tau\beta)_{1,2} = .700 \\ (\tau\beta)_{1,3} = 1.275 \\ (\tau\beta)_{1,4} = 4.325 \\ (\tau\beta)_{2,1} = 1.600 \\ (\tau\beta)_{2,2} = .150 \\ (\tau\beta)_{2,3} = .275 \\ (\tau\beta)_{2,4} = .250 \\ (\tau\beta)_{3,1} = 1.550 \\ (\tau\beta)_{3,2} = .400 \\ (\tau\beta)_{3,3} = .625 \\ (\tau\beta)_{3,4} = .575 \end{bmatrix}$$

5.

U_ijk	8	4	2	4	4	6	6	2
F1	1	1	1	1	2	2	2	2
F2	1	1	2	2	1	1	2	2
F3	1	2	1102		1	2	1	2

- a) The null hypothesis is that there is not a 3 way interaction when the 2 way interactions between $F_1 * F_2$ on each level of F_3 are equal. Since the two sides are not equal we reject the null and claim that there is a 3 way interaction.

$F_1 * F_2 * F_3$:

$$\begin{aligned}
 (\mu_{ijk} - \mu_{i'jk}) - (\mu_{ij'k} - \mu_{i'j'k}) &= (\mu_{ijk'} - \mu_{i'jk'}) - (\mu_{ij'k'} - \mu_{i'j'k'}) \\
 (\mu_{111} - \mu_{211}) - (\mu_{121} - \mu_{221}) &= (\mu_{112} - \mu_{212}) - (\mu_{122} - \mu_{222}) \\
 (8 - 4) - (2 - 6) &= (4 - 6) - (4 - 2) \\
 8 &\neq -4
 \end{aligned}$$

- b) The two sides are equal and so there is no interaction between F_2 and F_3 .

$F_2 * F_3$:

$$\begin{aligned}
 (\mu_{.11} - \mu_{.12}) &= (\mu_{.21} - \mu_{.22}) \\
 (.5 * (8 + 4) - .5 * (4 + 6)) &= (.5 * (2 + 6) - .5 * (4 + 2)) \\
 1 &= 1
 \end{aligned}$$

- c) The main effect for F_3 is not 0

Main Effect: F_3

$$\begin{aligned}
 &(\mu_{..1} - \mu_{..2}) \\
 &\frac{8 + 2 + 4 + 6}{4} - \frac{4 + 4 + 6 + 2}{4} \\
 &5 - 4 = 1
 \end{aligned}$$