Jack Zodori
STAT642
Assignment 04
4/14/22

1.) a.) $y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + e_{ijk}$ $i = 1,2,3,4$ $j = 1,2,3$, $k = 1,2,3,4$

μ = The overall mean of the data.

(4.0.7 pg 10) →

- Where τ_i = The fixed effect of the i^{th} level of crop

  β_j = The fixed effect of the j^{th} level of nitrogen in the growth media

  $(\tau\beta)_{ij}$ = are the interaction effects between crop and nitrogen.

- Constraints: $\tau_4 = 0$, $\beta_3 = 0$, $(\tau\beta)_{4j} = 0$, $(\tau\beta)_{i3} = 0$

b) Do the three model conditions appear to be satisfied as seen through the data.

Model Conditions: $e_{ijk} \overset{iid}{\sim} N(0, \sigma_e^2)$

- Because we are randomly sampling the independence assumption is satisfied
- Looking at a plot of the residuals there seems to be a significant funny pattern in our data, thus the constant error variance assumption seems to be violated.
- Also, looking at the QQ plot of the residuals, we see that the residuals are highly non-normal w/ fat tails, and this is confirmed by the Shapiro Wilks test.
- Transform the data

c.) Construct a complete ANOVA table for the data (using log transformed data b/c assumpt. not met).

Source	DF	Sum of Squares	Mean Square	F-value	Pr > F
Model	11	41.24	3.75	18.41	< 0.0001
Error	36	7.33	0.20		
Corrected Total	47	48.57			

* d) Test all appropriate hypos about main effects & interaction.

(1) $H_0: \tau_i = \beta_j = (\tau\beta)_{ij} = 0 \quad \forall (i,j)$

$H_a: \tau_i, \beta_j, (\tau\beta)_{ij} \neq 0$ for some i or some j.

- Reject H_0. Looking at ANOVA table above we have p-value < 0.0001

(2) $H_0: (\tau\beta)_{ij} = 0 \quad \forall (i,j)$

$H_a: (\tau\beta)_{ij} \neq 0$ for some (i,j)

- Reject H_0: in our ANOVA table below we see p-value for the interaction term < 0.0001

Source	DF	Type III SS	Mean Square	F-val	Pr > F
C	3	12.46	4.15	20.41	< 0.0001
N	2	18.34	9.17	45.05	< 0.0001
C*N	6	10.43	1.74	8.54	< 0.0001

1.) (contd)

(e) Construct a profile plot of the treatment means to illustrate your conclusions from part d.

- We can see in figure (1) that we have a consistent interaction between Nitrogen levels and levels of crop because the magnitude of the differences between the mean response for the treatments changes for different

(f) Group the four crops relative to their mean acetylene reduction (see figure 2)

Q: do we compare → when using transformed data or orig. data?

- Since the interaction of crop and nitrogen was significant the grouping of the levels of nitrogen will be done seperately for each of the 4 types of crops. Using the un-adjusted p-values we will use $\alpha_{pc} = \frac{0.05}{12} = 0.004167$ $\left(4\binom{3}{2}=12\right)$

Answer changes depending on what we use.

Alfalfa: $G_1 = \{0, 50, 100\}$ Soybean: $G_1 = \{0\}$, $G_2 = \{50\}$, $G_3 = \{100\}$

A: yes, use transformed data

Guar: $G_1 = \{0, 50\}$ $G_2 = \{100\}$ Mungbean: $G_1 = \{0, 50\}$, $G_2 = \{50, 100\}$

2.) $y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + e_{ijkl}$

$i = 1,2$, $j = 1,2,3$, $k = 1,2$, $l = 1,2,3,4$

* 4/1/22

Q: Do we group the levels of crop for each level of Nitrogen or vice versa?

- τ_i is the fixed effect of the ith level of filler
- β_j is the fixed effect of the jth level of proportion
- γ_k is the fixed effect of the kth level of surface
- $(\tau\beta)_{ij}$ are the interaction effects btwn filler & proportion
- $(\tau\gamma)_{ik}$ are the interaction effects btwn filler & surface
- $(\beta\gamma)_{jk}$ are the interaction effects btwn proportion & surface
- $(\tau\beta\gamma)_{ijk}$ are the interaction effects btwn filler, proportion & surface.

Constraints: $\tau_2 = \beta_3 = \gamma_2 = (\tau\beta)_{2j} = (\tau\beta)_{i3} = (\tau\gamma)_{2k} = (\tau\gamma)_{i2} = (\beta\gamma)_{3k} = (\beta\gamma)_{j2} = (\tau\beta\gamma)_{2jk} = (\tau\beta\gamma)_{i3k} = (\tau\beta\gamma)_{ij2} = 0$

(b) Check for any violations in the model conditions:

$e_{ijkl} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

- B/c our experiment is a CRD we have the independence condition satisfied.
- Looking at the Brown-Forsythe Test for homogeneity of variances we see no evidence that the constant error variance assumption has been violated.
- Looking at the QQ plot of the residuals it seems the residuals may deviate from normality, however, looking at the shapiro wilks test we find that there is no evidence the normality assumption has been violated.

2.) (Cont'd)

(c) Display the AOV Table

Source	DF	Sum of Squares	Mean Square	F-value	Pr < F
Model	11	54,616.73	4965.16	62.67	<0.0001
Error	36	2,852.25	79.23		
Total	47	57,468.98			

(d) Prepare a table of the least squares estimates of cell & marginal means w/ their respective standard errors

Use last part of code → • See figure 3 for cell means & SEs

Use effects model → • See figure 4 for marginal means & SEs

(e) Is there a significant difference in the mean weight loss of the two types of filler?

• check, is Filler an important 2 way interaction? We can see from figure 4 that both the 2-way interactions involving filler are significant at the $\alpha = 0.05$ level.

(see H.Q 7 pg 70 case 2)

NOTE: want to test if $\hat{\mu}_{1..} = \hat{\mu}_{2..}$ but our 2 way interaction is significant for both $F*S$ & $F*P$. Thus, we need to examine the mean effect of F at each level of S and similarly for P.

Q: or would we do this the opposite way? I.e. examine the mean effect of S at each level of F & examine the mean effect of P at each level of F?

no. →

• Fixing surface: use $\alpha_{pc} = 0.05/4$ (Figure 5 (i))

$H_0: \mu_{ij.} - \mu_{i'j.} = 0 \quad H_a: \mu_{ij.} - \mu_{i'j.} \neq 0$

• Surface $= S_1$: $G_1 = \{F_1, F_2\}$

Surface $= S_2$: $G_1 = \{F_1\}$, $G_2 = \{F_2\}$

• Fixing Proportion: use $\alpha_{pc} = \frac{0.05}{6}$ (Figure 5(ii))

• Proportion = 25; $G_1 = \{F_1, F_2\}$

• Proportion = 50; $G_1 = \{F_1, F_2\}$

• Proportion = 75; $G_1 = \{F_1\}$, $G_2 = \{F_2\}$

7) (contd)

(f.) Is there a significant difference in the mean weight loss of the three proportions of filler? (see figure 5(i)(ii))

- Fix Filler: $\alpha_{pc} = 0.05/6 \; [H_0: (\mu_{i \cdot \ell} - \mu_{i \cdot \ell'}) = 0]$
  - Filler = F_1: $G_1 = \{25\}$, $G_2 = \{50\}$, $G_3 = \{75\}$
  - Filler = F_2: $G_1 = \{25\}$, $G_2 = \{50\}$, $G_3 = \{75\}$

- Now, look at main effect of Proportion b/c Surface * Proportion is n't significant
  $G_1 = \{25\}$, $G_2 = \{50\}$, $G_3 = \{75\}$
  ↳ (see figure 6)

(g) Are there any trends in the mean weight loss as the proportion of filler increases? $\alpha_{pc} = (1-(1-0.05)^{1/12})$ *

- S_1F_1: We have a linear trend in mean weight loss as the proportion of filler increases
- S_1F_2: " "
- S_2F_1: " "
- S_2F_2: " "
- S_1F_1: We don't have significant evidence of a quadratic trend in the mean weight loss as the proportion of filler increases
- S_1F_2: " "
- S_2F_1: " "
- S_2F_2: " "
- F_1: We have significant evidence of a linear trend in the mean weight loss as the proportion of filler increases
- F_2: " "
- F_1: We don't have significant evidence of a quadratic trend in the mean " "
- F_2: " "

(h) See figure 8 (i) - (iv)

3.) (a) R
(b) B

4.) (a) K
(b) D
(c) E
(d) P

5.) (1) The analysis is wrong, b/c it does not account for repeated measurments. B/c we now have 2 plants in each growth chamber we need to account for repeated measurments.

(2) r=7

(3) D

(4) E

Figure 1: Problem 1(e)

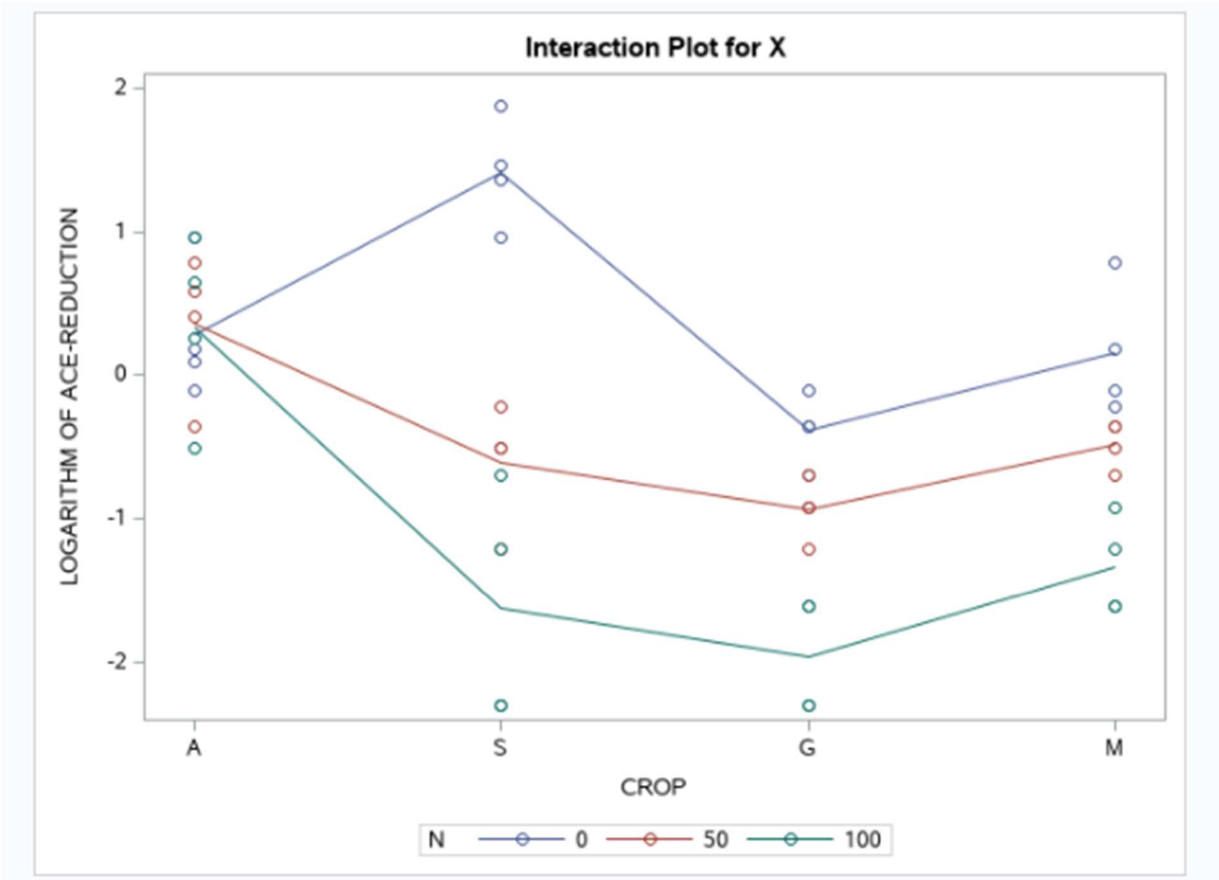

Figure 2: Problem 1(f) Least Squares Means for Effect C*N

Least Squares Means for effect C*N
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: X

i/j	1	2	3	4	5	6	7	8	9	10	11	12
1		0.8171	0.8634	0.0011	0.0081	<.0001	0.0458	0.0005	<.0001	0.7059	0.0224	<.0001
2	0.8171		0.9528	0.0021	0.0044	<.0001	0.0273	0.0003	<.0001	0.5435	0.0128	<.0001
3	0.8634	0.9528		0.0018	0.0052	<.0001	0.0312	0.0003	<.0001	0.5832	0.0148	<.0001
4	0.0011	0.0021	0.0018		<.0001	<.0001	<.0001	<.0001	<.0001	0.0004	<.0001	<.0001
5	0.0081	0.0044	0.0052	<.0001		0.0031	0.4676	0.3222	0.0002	0.0206	0.6796	0.0296
6	<.0001	<.0001	<.0001	<.0001	0.0031		0.0004	0.0365	0.3073	<.0001	0.0010	0.3681
7	0.0458	0.0273	0.0312	<.0001	0.4676	0.0004		0.0908	<.0001	0.1001	0.7525	0.0049
8	0.0005	0.0003	0.0003	<.0001	0.3222	0.0365	0.0908		0.0028	0.0015	0.1642	0.2154
9	<.0001	<.0001	<.0001	<.0001	0.0002	0.3073	<.0001	0.0028		<.0001	<.0001	0.0594
10	0.7059	0.5435	0.5832	0.0004	0.0206	<.0001	0.1001	0.0015	<.0001		0.0525	<.0001
11	0.0224	0.0128	0.0148	<.0001	0.6796	0.0010	0.7525	0.1642	<.0001	0.0525		0.0110
12	<.0001	<.0001	<.0001	<.0001	0.0296	0.3681	0.0049	0.2154	0.0594	<.0001	0.0110	

Figure 3: Problem 2(d) Cell Means and Standard Errors

F*S*P Least Squares Means							
FILLERS	SURFACE TRT	PROPORTION FILLER	Estimate	Standard Error	DF	t Value	Pr > |t|
F1	S1	25	201.00	4.4505	36	45.16	<.0001
F1	S1	50	237.00	4.4505	36	53.25	<.0001
F1	S1	75	268.75	4.4505	36	60.39	<.0001
F1	S2	25	164.00	4.4505	36	36.85	<.0001
F1	S2	50	188.00	4.4505	36	42.24	<.0001
F1	S2	75	227.25	4.4505	36	51.06	<.0001
F2	S1	25	209.50	4.4505	36	47.07	<.0001
F2	S1	50	232.50	4.4505	36	52.24	<.0001
F2	S1	75	240.50	4.4505	36	54.04	<.0001
F2	S2	25	148.50	4.4505	36	33.37	<.0001
F2	S2	50	174.75	4.4505	36	39.26	<.0001
F2	S2	75	200.00	4.4505	36	44.94	<.0001

Figure 4: Problem 2 (d) Marginal Means and Standard Errors

Figure 4 (i): Marginal Means for Surface Treatment (S)

S	W LSMEAN	Standard Error
S1	231.541667	1.816925
S2	183.750000	1.816925

Figure 4 (ii): Marginal Means for Filler (F)

F	W LSMEAN	Standard Error
F1	214.333333	1.816925
F2	200.958333	1.816925

Figure 4: Problem 2 (d) Marginal Means and Standard Errors

Figure 4 (iii): Marginal Means for Proportions of the Filler (P)

P	W LSMEAN	Standard Error
25	180.750000	2.225269
50	208.062500	2.225269
75	234.125000	2.225269

Figure 5(i): Problem 2(e)

F	S	W LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
F1	S1	235.583333	2.569520	<.0001	1
F1	S2	193.083333	2.569520	<.0001	2
F2	S1	227.500000	2.569520	<.0001	3
F2	S2	174.416667	2.569520	<.0001	4

Least Squares Means for effect F*S
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: W

i/j	1	2	3	4
1		<.0001	0.0325	<.0001
2	<.0001		<.0001	<.0001
3	0.0325	<.0001		<.0001
4	<.0001	<.0001	<.0001	

Figure 5(ii): Problem 2(e)

F	P	W LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
F1	25	182.500000	3.147006	<.0001	1
F1	50	212.500000	3.147006	<.0001	2
F1	75	248.000000	3.147006	<.0001	3
F2	25	179.000000	3.147006	<.0001	4
F2	50	203.625000	3.147006	<.0001	5
F2	75	220.250000	3.147006	<.0001	6

Least Squares Means for effect F*P
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: W

i/j	1	2	3	4	5	6
1		<.0001	<.0001	0.4368	<.0001	<.0001
2	<.0001		<.0001	<.0001	0.0538	0.0902
3	<.0001	<.0001		<.0001	<.0001	<.0001
4	0.4368	<.0001	<.0001		<.0001	<.0001
5	<.0001	0.0538	<.0001	<.0001		0.0006
6	<.0001	0.0902	<.0001	<.0001	0.0006	

Figure 6: Problem 2(f)

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

P	W LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
25	180.750000	2.225269	<.0001	1
50	208.062500	2.225269	<.0001	2
75	234.125000	2.225269	<.0001	3

Least Squares Means for effect P
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: W

i/j	1	2	3
1		<.0001	<.0001
2	<.0001		<.0001
3	<.0001	<.0001	

Figure 7: Problem 2(g)

CELL MEANS MODEL ANALYSIS

The GLM Procedure

Dependent Variable: W WEIGHT LOSS

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
LIN-S1F1	1	9180.12500	9180.12500	115.87	<.0001
LIN-S1F2	1	1922.00000	1922.00000	24.26	<.0001
LIN-S2F1	1	8001.12500	8001.12500	100.99	<.0001
LIN-S2F2	1	5304.50000	5304.50000	66.95	<.0001
QUA-S1F1	1	12.04167	12.04167	0.15	0.6989
QUA-S1F2	1	150.00000	150.00000	1.89	0.1773
QUA-S2F1	1	155.04167	155.04167	1.96	0.1704
QUA-S2F2	1	0.66667	0.66667	0.01	0.9274
LIN-F1	1	17161.00000	17161.00000	216.60	<.0001
LIN-F2	1	6806.25000	6806.25000	85.91	<.0001
QUA-F1	1	40.33333	40.33333	0.51	0.4801
QUA-F2	1	85.33333	85.33333	1.08	0.3063

Figure 8(i): Problem 2 (h)

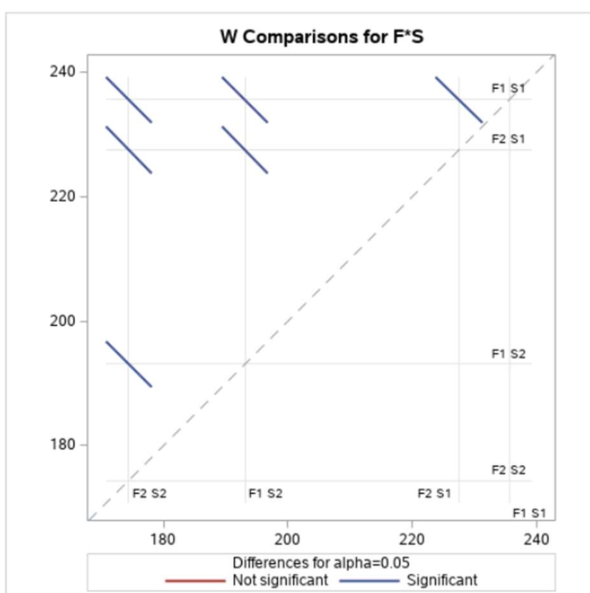

Figure 8 (ii): Problem 2(h)

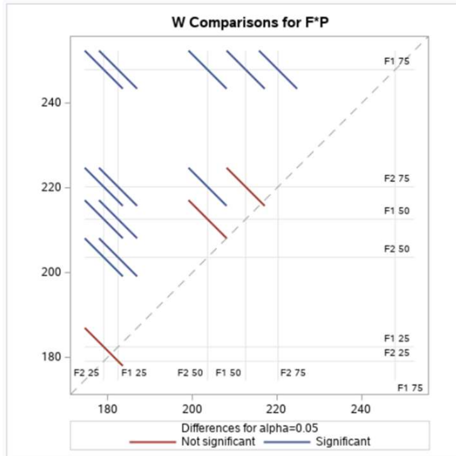

Figure 8 (iii): Problem 2(h)

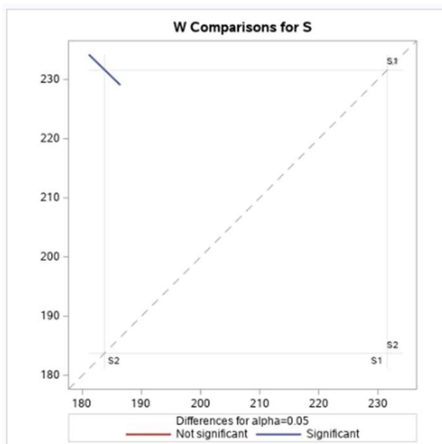

Figure 8 (iv): Problem 2(h)