

STAT 512 – Assignment 7

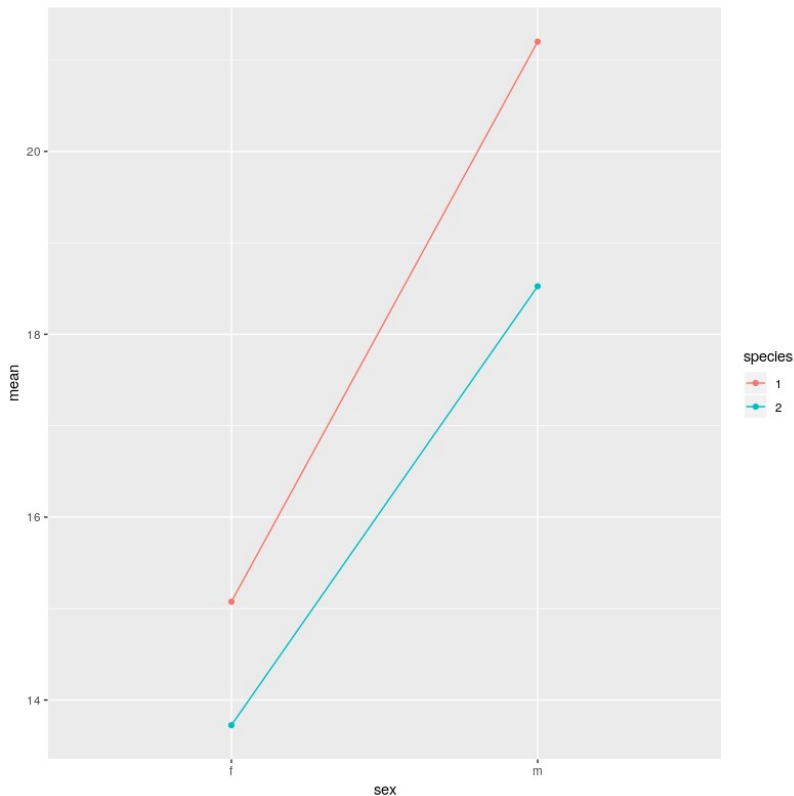
Vignesh M. Pagadala
Vignesh.Pagadala@ColoState.Edu

1.

A. Create a table of summary statistics including sample size, mean and standard deviation for each sex, species combination. Then construct an interaction plot. For consistency, please put sex on the X axis. Include both the summary table and interaction plot in your assignment. (4 pts)

ANSWER:

```
# A tibble: 4 x 5
# Groups:   sex [2]
  sex species     n mean    sd
<fct> <fct>   <int> <dbl> <dbl>
1 f     1         4  15.1  1.24
2 f     2         4  13.7  1.21
3 m     1         4  21.2  1.33
4 m     2         4  18.5  1.84
```



B. Construct the Type3 ANOVA table.

ANSWER:

Anova Table (Type III tests)

Response: pcb

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	4695.7	1	2309.112	4.317e-15 ***
group	137.3	3	22.508	3.230e-05 ***
Residuals	24.4	12		

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

C. Use emmeans to calculate (Tukey adjusted) pairwise comparisons for all four groups.

ANSWER:

\$emmeans

group	emmean	SE	df	lower.CL	upper.CL
sp1f	15.1	0.713	12	13.5	16.6
sp1m	21.2	0.713	12	19.6	22.8
sp2f	13.7	0.713	12	12.2	15.3
sp2m	18.5	0.713	12	17.0	20.1

Confidence level used: 0.95

\$contrasts

contrast	estimate	SE	df	t.ratio	p.value
sp1f - sp1m	-6.12	1.01	12	-6.074	0.0003
sp1f - sp2f	1.35	1.01	12	1.339	0.5576
sp1f - sp2m	-3.45	1.01	12	-3.421	0.0227
sp1m - sp2f	7.47	1.01	12	7.413	<.0001
sp1m - sp2m	2.67	1.01	12	2.653	0.0857
sp2f - sp2m	-4.80	1.01	12	-4.760	0.0023

P value adjustment: tukey method for comparing a family of 4 estimates

D. Construct the Type3 ANOVA table.

ANSWER:

Anova Table (Type III tests)

Response: pcb

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	4695.7	1	2309.1121	4.317e-15 ***
sex	119.4	1	58.6935	5.839e-06 ***
species	16.2	1	7.9667	0.01539 *
sex:species	1.8	1	0.8633	0.37112
Residuals	24.4	12		

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

E. Use emmeans(, pairwise ~ sex:species) to calculate (Tukey adjusted) pairwise comparisons for all four groups.

ANSWER:

\$emmeans

sex	species	emmean	SE	df	lower.CL	upper.CL
f	1	15.1	0.713	12	13.5	16.6
m	1	21.2	0.713	12	19.6	22.8
f	2	13.7	0.713	12	12.2	15.3
m	2	18.5	0.713	12	17.0	20.1

Confidence level used: 0.95

\$contrasts

contrast	estimate	SE	df	t.ratio	p.value
f,1 - m,1	-6.12	1.01	12	-6.074	0.0003
f,1 - f,2	1.35	1.01	12	1.339	0.5576
f,1 - m,2	-3.45	1.01	12	-3.421	0.0227
m,1 - f,2	7.47	1.01	12	7.413	<.0001
m,1 - m,2	2.67	1.01	12	2.653	0.0857
f,2 - m,2	-4.80	1.01	12	-4.760	0.0023

P value adjustment: tukey method for comparing a family of 4 estimates

F. Use emmeans(, pairwise ~ species) to calculate the pairwise comparison corresponding to the main effect of species. Note that the p-value from this comparison should match the F-test corresponding to species from the ANOVA table from part (D).

ANSWER:

\$emmeans

species	emmean	SE	df	lower.CL	upper.CL
1	18.1	0.504	12	17	19.2
2	16.1	0.504	12	15	17.2

Results are averaged over the levels of: sex

Confidence level used: 0.95

\$contrasts

contrast	estimate	SE	df	t.ratio	p.value
1 - 2	2.01	0.713	12	2.823	0.0154

Results are averaged over the levels of: sex

G. Consider the output from the two previous questions. From part (E), for the “m,1 – m,2” comparison you should have found an estimate = 2.67, p-value = 0.0857. From part (F), for the “1 - 2” comparison you should have found an estimate = 2.01, p-value = 0.0154. Briefly explain why we find a smaller p-value for the “1 - 2” comparison even though the estimated difference is smaller.

ANSWER:

The p-value is better for the 1-2 comparison since the result is averaged over the species sex.

2.

A. Describe the blocking and treatment structure.

ANSWER:

The blocking structure is Randomized Complete Block and the treatment structure is 2-way factorial.

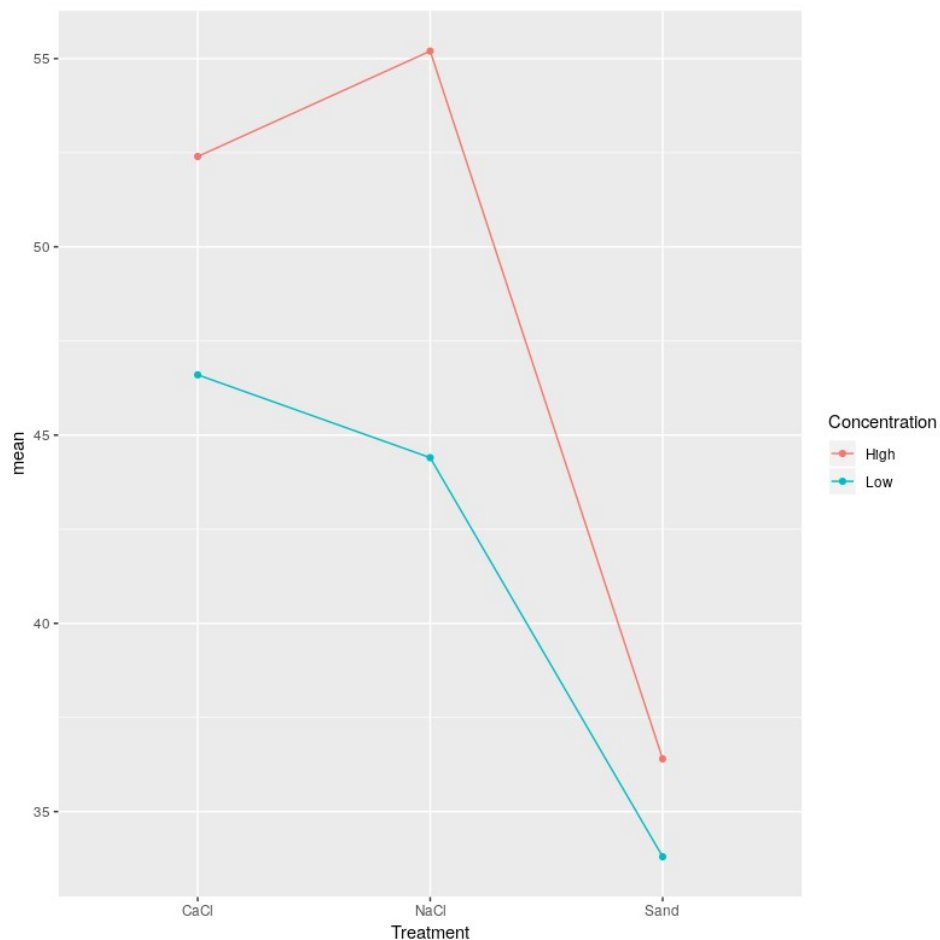
B. Create a table of summary statistics including sample size, mean and standard deviation for each Treatment*Concentration combination. Then construct an interaction plot. For consistency, please put Treatment on the X axis. Include both the summary table and interaction plot in your assignment. (4 pts)

ANSWER:

A tibble: 6 x 5

Groups: Treatment [3]

	Treatment	Concentration	n	mean	sd
	<fct>	<fct>	<int>	<dbl>	<dbl>
1	CaCl	High	5	52.4	6.50
2	CaCl	Low	5	46.6	5.59
3	NaCl	High	5	55.2	7.79
4	NaCl	Low	5	44.4	6.88
5	Sand	High	5	36.4	4.98
6	Sand	Low	5	33.8	7.46



C. Considering your answer to part A, fit an appropriate model and include the Type 3 ANOVA table in your assignment. (You might considered checking diagnostics plot as a matter of habit even though they are not requested here). (4 pts)

ANSWER:

Anova Table (Type III tests)

Response: cracks

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	60211	1	15999.433	< 2.2e-16 ***
Roadway	973	4	64.646	3.740e-11 ***
Treatment	1412	2	187.573	1.103e-13 ***
Concentration	307	1	81.630	1.694e-08 ***
Treatment:Concentration	85	2	11.346	0.0005091 ***
Residuals	75	20		

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

D. Discuss the value of the blocking for this data. Justify your response with appropriate test- statistic(s) and p-value(s).

ANSWER:

The blocking was useful, since the test statistic is 64.646 and the p-value is very significant at <0.001

E. Compare mean response for High vs Low Concentration separately for each Treatment. This can be done using emmeans. Include the “emmeans contrasts” output in your assignment, but also (briefly) summarize your findings. (4 pts)

ANSWER:

\$emmeans

Treatment = CaCl:

Concentration	emmean	SE	df	lower.CL	upper.CL
High	52.4	0.868	20	50.6	54.2
Low	46.6	0.868	20	44.8	48.4

Treatment = NaCl:

Concentration	emmean	SE	df	lower.CL	upper.CL
High	55.2	0.868	20	53.4	57.0
Low	44.4	0.868	20	42.6	46.2

Treatment = Sand:

Concentration	emmean	SE	df	lower.CL	upper.CL
High	36.4	0.868	20	34.6	38.2
Low	33.8	0.868	20	32.0	35.6

Results are averaged over the levels of: Roadway
Confidence level used: 0.95

\$contrasts

Treatment = CaCl:

contrast	estimate	SE	df	t.ratio	p.value
High - Low	5.8	1.23	20	4.727	0.0001

Treatment = NaCl:

contrast	estimate	SE	df	t.ratio	p.value
High - Low	10.8	1.23	20	8.803	<.0001

```
Treatment = Sand:
contrast    estimate    SE df t.ratio p.value
High - Low      2.6 1.23 20 2.119  0.0468
```

Results are averaged over the levels of: Roadway

For all the 3 treatments, higher concentrations result in much higher cracks than lower concentrations.

F. Compare mean responses between the 3 Treatments for Concentration = Low. (default Tukey adjustment is fine). Include these “emmeans contrasts” output in your assignment, but also (briefly) summarize your findings. (4 pts)

ANSWER:

\$emmeans

Concentration = High:

Treatment	emmean	SE	df	lower.CL	upper.CL
CaCl	52.4	0.868	20	50.6	54.2
NaCl	55.2	0.868	20	53.4	57.0
Sand	36.4	0.868	20	34.6	38.2

Concentration = Low:

Treatment	emmean	SE	df	lower.CL	upper.CL
CaCl	46.6	0.868	20	44.8	48.4
NaCl	44.4	0.868	20	42.6	46.2
Sand	33.8	0.868	20	32.0	35.6

Results are averaged over the levels of: Roadway
Confidence level used: 0.95

\$contrasts

Concentration = High:

contrast	estimate	SE	df	t.ratio	p.value
CaCl - NaCl	-2.8	1.23	20	-2.282	0.0817
CaCl - Sand	16.0	1.23	20	13.041	<.0001
NaCl - Sand	18.8	1.23	20	15.323	<.0001

Concentration = Low:

contrast	estimate	SE	df	t.ratio	p.value
CaCl - NaCl	2.2	1.23	20	1.793	0.1974
CaCl - Sand	12.8	1.23	20	10.433	<.0001
NaCl - Sand	10.6	1.23	20	8.640	<.0001

Results are averaged over the levels of: Roadway
P value adjustment: tukey method for comparing a family of 3 estimates

Sand treatment seems to result in lower average of cracks than NaCl and CaCl, both of which seem to result in the same average number of cracks.