Statistics 5200 Fall 2020 Homework 2

Assignment Due (by 11:59 P.M.): Wed 30 Sep

Directions: You may discuss the exercises with other students and with the instructor, but the work you turn in must be your own. Note that neatness and format (including SAS code in appendix) will contribute 10 points to the total score. This assignment will be graded out of 75 points.

Exercises:

1. (20 points; based on textbook Exercise 3.3, p. 60). Twelve orange pulp silage samples were divided at random into four groups of three. One of the groups was left as an untreated control, while the other three groups were treated with formic acid, beet pulp, and sodium chloride, respectively. One of the responses was the moisture content of the silage, and we are interested in testing whether the treatments affect the mean moisture content. The observed data are as follows:

NaCl	Formic acid	Beet pulp	Control
80.5	89.1	77.8	76.7
79.3	75.7	79.5	77.2
79.0	81.2	77.0	78.6

Without transforming the data, fit a CRD model using PROC GLM, and report (by number) the following parts: (i) null and alternative hypotheses, (ii) ANOVA table, (iii) test statistic for the null, (iv) sampling distribution for the test statistic, (v) p-value for the test, (vi) conclusion (in the context of this application), (vii) proportion of variation in response explained by factor levels, (viii) evidence of appropriateness of relevant assumptions, and (ix) what your evidence in (viii) says about your inference in (iv)-(vi).

2. (20 points; based on textbook Exercise 3.5, page 61). The leaves of certain plants in the genus *Albizzia* will fold and unfold in various light conditions. We have taken fifteen different leaves and subjected them to red light for 3 minutes. the leaves were divided into three groups of five at random. The leaflet angles were then measured 30, 45, and 60 minutes after light exposure in the three groups. We are interested in testing whether delay after exposure affects leaflet angle. The observed data are as follows:

Delay	Angle (degrees)				
30	140	138	140	138	142
45	140	150	120	128	130
60	118	130	128	118	118

Without transforming the data, fit a CRD model using PROC GLM, and report (by number) parts (i)-(ix) as above.

3. (25 points, based on textbook Problem 3.1, page 61). Cardiac pacemakers contain electrical connections that are platinum pins soldered onto a substrate. The question of interest is whether different operators produce solder joints with the same strength. Twelve substrates are available, and are randomly assigned to four operators (three substrates per operator). Each operator solders four pins on each substrate, and then these solder joints are assessed by measuring the shear strength of the pins. The observed data are as follows:

	Strength (lb)											
Operator	Substrate 1				Substrate 2				Substrate 3			
1	5.60	6.80	8.32	8.70	7.64	7.44	7.48	7.80	7.72	8.40	6.98	8.00
2	5.04	7.38	5.56	6.96	8.30	6.86	5.62	7.22	5.72	6.40	7.54	7.50
3	8.36	7.04	6.92	8.18	6.20	6.10	2.75	8.14	9.00	8.64	6.60	8.18
4	8.30	8.54	7.68	8.92	8.46	7.38	8.08	8.12	8.68	8.24	8.09	8.06

Just to clarify – there are twelve substrates, three for each operator; so in the table above, for example, "Substrate 1" is used to describe four different substrates (one for each operator). Fit a CRD model using PROC GLM, and report (by number) parts (i)-(ix) as above. Be sure to pay attention here to what the experimental units are. (The "corrected total" degrees of freedom in (ii) must be one fewer than the number of experimental units – not measurement units. You don't need to transform your data in the power family sense, but you will need to do something [maybe 'by hand'] similar to what was discussed in class related to the "Measurement Unit" section of Handout #3.) Report what you did to address this as part number (0), preceding part number (i) in your response.

Appendix: (10 points) Include SAS code used for this assignment.