STAT 5309 – SP 2022

MIDTERM

****

***Direction:**

(a) Answer all question parts, in order**. Each problem starts with a sentence to describe the data in your own word. And what to do next.**

(b) All outputs/plots must be accompanied by short remark, comment, highlighted (points deducted if remark/comment not shown)

(c) Present work in a readable form of a scientific report.

***Due: TUES, MAR 22**

1. Problem [data 3-12] (10 points)

3-12. Four different designs for a digital computer circuit are being studied to compare the amount of noise present. The following data have been obtained:

Circuit Design	Noise Observed				
1	19	20	19	30	8
2	80	61	73	56	80
3	47	26	25	35	50
4	95	46	83	78	97

(a) Is the amount of noise present the same for all four designs? Use $\alpha = 0.05$.
(b) Analyze the residuals from this experiment. Are the analysis of variance assumptions satisfied?
(c) Which circuit design would you select for use? Low noise is best.

Set up a data frame, named "circuit".

2. Problem [data 3-14] (10 points)

3-14. Three brands of batteries are under study. It is suspected that the lives (in weeks) of the three brands are different. Five batteries of each brand are tested with the following results:

Weeks of Life		
Brand 1	Brand 2	Brand 3
100	76	108
96	80	100
92	75	96
96	84	98
92	82	100

(a) Are the lives of these brands of batteries different?
(b) Analyze the residuals from this experiment.
(c) Construct a 95 percent confidence interval estimate on the mean life of battery brand 2. Construct a 99 percent confidence interval estimate on the mean difference between the lives of battery brands 2 and 3.

Notes: Set up the data frame, named "batteries" . More on (a) Perform a boxplot and and a stripchart.

3. Problem [data 5.5] (10 points)

Johnson and Leone (*Statistics and Experimental Design in Engineering and the Physical Sciences*, Wiley, 1977) describe an experiment to investigate warping of copper plates. The two factors studied were the temperature and the copper content of the plates. The response variable was a measure of the amount of warping. The data were as follows:

	Copper Content (%)			
Temperature (°C)	40	60	80	100
50	17, 20	16, 21	24, 22	28, 27
75	12, 9	18, 13	17, 12	27, 31
100	16, 12	18, 21	25, 23	30, 23
125	21, 17	23, 21	23, 22	29, 31

Set up a data frame, named "plates"

(a) Build a regression model using aov(). Is there any indication that either factor affects the amount of warping? Is there any interaction between the factors . Use $\alpha = 0.05$
(b) Do a Box plot on Temperature, on Content and on Temperature*Content
Which level combination give the lowest warpage.
(c) Suppose Temperature cannot be controlled in where the copper plate are used, which Content gives the smallest warpage?
(d) Perform a 3-part residuals assumption check of the aov model.(Do not use plot(model))

4. Problem [data 5-8] (10 points)

5-8. An experiment is conducted to study the influence of operating temperature and three types of face-plate glass in the light output of an oscilloscope tube. The following data are collected:

	Temperature		
Glass Type	100	125	150
	580	1090	1392
1	568	1087	1380
	570	1085	1386
	550	1070	1328
2	530	1035	1312
	579	1000	1299
	546	1045	867
3	575	1053	904
	599	1066	889

(a) Use $\alpha = 0.05$ in the analysis. Is there a significant interaction effect? Does glass type or temperature affect the response? What conclusions can you draw?
(b) Fit an appropriate model relating light output to glass type and temperature.
(c) Analyze the residuals from this experiment. Comment on the adequacy of the models you have considered.

Note: Set up a data frame, named "tubes". More on (a)Is interaction significant? Build a reduced model if interaction is not significant.

5. Problem (10 points)

Data: Salinity. Temperature, Salinity, and Density affect growth and survival of Shrimps. 3 factors: "Temperature" (25 degree, 35 degree), "Salinity" (10%,25%, 40%), "Density" (80 shrimp/40 liters, 160/40 liters).

Temp	Density	Salinity	Weight Gain
25	80	10%	86, 52, 73,
		25%	544, 371, 482
		40%	390, 290, 397
	160	10%	53, 73, 86
		25%	393, 398, 208,
		40%	249, 265, 243
35	80	10%	439, 436,349
		25%	249,245,330
		40%	247,277,205

	160	10%	324,305,364
		25%	352,267,316
		40%	188,223,281

Enter into a data frame, named "shrimp", with 3 factors "Temp", "Density", "Salinity".

(a) Build regression model with interactions (including 2-way and 3-way interactions). Check main factors significant. Are the interactions significant? Build a reduced model if necessary.
(b) Check interactions(2-way and 3-way) significance by plots. Confirm results with part (a)
(c) Find the combination of Temp, Density, Salinity which produces the shrimp largest weight gain
(d) Perform 3-part residual check.

6. Problem (10 points)

	Oil	Truck	Fuel Consumption
1	1	1	0.5
2	1	2	0.634
3	1	3	0.487
4	1	4	0.329
5	1	5	0.512
6	2	1	0.535
	2	2	0.675
8	2	3	0.52
9	2	4	0.435
10	2	5	0.54
11	3	1	0.513
12	3	2	0.595
13	3	3	0.488
14	3	4	0.4
15	3	5	0.51

Set up a data frame, named "fuel".

(a) Build a linear model. Is there any significant difference of means about the oil types. Which oil type gives the lowest fuel consumption.
(b) Use Truck as a Blocking factor. Is the **Blocking factor effective?**
(c) Do a complete residual assumption check.

7. Problem [data 4-5] (10 points)

4-5. An article in the *Fire Safety Journal* (''The Effect of Nozzle Design on the Stability and Performance of Turbulent Water Jets,'' Vol. 4, August 1981) describes an experiment in which a shape factor was determined for several different nozzle designs at six levels of

jet efflux velocity. Interest focused on potential differences between nozzle designs, with velocity considered as a nuisance variable. The data are shown below:

Nozzle Design	Jet Efflux Velocity (m/s)					
	11.73	14.37	16.59	20.43	23.46	28.74
1	0.78	0.80	0.81	0.75	0.77	0.78
2	0.85	0.85	0.92	0.86	0.81	0.83
3	0.93	0.92	0.95	0.89	0.89	0.83
4	1.14	0.97	0.98	0.88	0.86	0.83
5	0.97	0.86	0.78	0.76	0.76	0.75

Set up a data frame named "nozzles"

(a) Build a linear model, nozzle as a blocking factor. Does the nozzle design affect the shape factor? Use $\alpha = .05$. Is nozzle Blocking factor effective?

(b) Which nozzle designs are different with respect to the shape factor? [Hint: use Tukey HSD]

(c) Is the velocity effect significant?

8. Problem[data 5-19] (10 points)

5-19. The yield of a chemical process is being studied. The two factors of interest are temperature and pressure. Three levels of each factor are selected; however, only nine runs can be made in one day. The experimenter runs a complete replicate of the design on each day. The data are shown in the following table. Analyze the data, assuming that the days are blocks.

	Day 1 Pressure			Day 2 Pressure		
Temperature	250	260	270	250	260	270
Low	86.3	84.0	85.8	86.1	85.2	87.3
Medium	88.5	87.3	89.0	89.4	89.9	90.3
High	89.1	90.2	91.3	91.7	93.2	93.7

Set up a data frame , named "chemprocess".

(a) Build linear model, using "days" as blocking factor. Is "days" an effective blocking factor?

(b) Are "Temperature" or "Pressure" or interactions are significant?

(c) Which Temperature and Pressure combination gives the highest yield.

(d) Perform a complete residual check.

9. Problem [data 5-17] (10 pts)

5-17. The quality control department of a fabric finishing plant is studying the effect of several factors on the dyeing of cotton–synthetic cloth used to manufacture men's shirts. Three operators, three cycle times, and two temperatures were selected, and three small specimens of cloth were dyed under each set of conditions. The finished cloth was compared to a standard, and a numerical score was assigned. The results follow. Analyze the data and draw conclusions. Comment on the model's adequacy.

	Temperature					
	300°			350°		
	Operator			Operator		
Cycle Time	1	2	3	1	2	3
	23	27	31	24	38	34
40	24	28	32	23	36	36
	25	26	29	28	35	39
	36	34	33	37	34	34
50	35	38	34	39	38	36
	36	39	35	35	36	31
	28	35	26	26	36	28
60	24	35	27	29	37	26
	27	34	25	25	34	24

Set up a data frame "fabrics" with 3 factors.

(a) Build a regression model. Any interaction among factors?

(b) Which factor can be used as Blocking factor? Is the Blocking factor effective?

(c) Consider any reduced model?

(d) Do a complete 3-part residual assumption check.

10. Problem (Latin Square)[data 4-27] (10 points)

The effect of five different ingredients (A, B, C, D, E) on the reaction time of a chemical process is being studied. Each batch of new material is only large enough to permit five runs to be made. Furthermore, each run requires approximately $1\frac{1}{2}$ hours, so only five runs can be made in one day. The experimenter decides to run the experiment as a Latin square so that day and batch effects may be systematically controlled. She obtains the data that follow. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.

	Day				
Batch	1	2	3	4	5
1	$A = 8$	$B = 7$	$D = 1$	$C = 7$	$E = 3$
2	$C = 11$	$E = 2$	$A = 7$	$D = 3$	$B = 8$
3	$B = 4$	$A = 9$	$C = 10$	$E = 1$	$D = 5$
4	$D = 6$	$C = 8$	$E = 6$	$B = 6$	$A = 10$
5	$E = 4$	$D = 2$	$B = 3$	$A = 8$	$C = 8$

Set up a data frame. “chemical”. Use Day and Batch columns and row, in Latin square design. Ingredient as Treatment. Either enter manually or use design.lsd() [If use design.lsd(), must enter the response values in order given by data]

(a) Build a linear model, using aov(). Does the ingredients affect the reaction time?
(b) Are Day, Batch effective Blocking factors? Check interaction between Day and Batch.
(c) Find the lowest reaction time.
(d) Perform residual check.

	Batch	Day	Catalyst	Time
1	1	1	A	8
2	2	1	C	11
3	3	1	B	4
4	4	1	D	6
5	5	1	E	4