STAT 514 Design of Experiments — Homework 8

Problem 1

(a) The hypotheses are

$$
\begin{aligned}
& H_0: && \text{There is no difference between the four assembly methods,} \\
& && \text{or symbolically, } \tau_A = \tau_B = \tau_C = \tau_D. \\
\text{vs} \quad & H_1: && \text{There is a difference between the four assembly methods,} \\
& && \text{or symbolically, } \tau_A, \tau_B, \tau_C \text{ and } \tau_D \text{ are not all equal.}
\end{aligned}
$$

Let's look at the ANOVA table output from SAS (I replaced the line for the model SS by lines for the two block SS and the treatment SS).

```
                                        Sum of
Source                      DF         Squares     Mean Square    F Value    Pr > F

ord                          3     18.50000000      6.16666667       3.52    0.0885
opt                          3     51.50000000     17.16666667       9.81    0.0099
trt                          3     72.50000000     24.16666667      13.81    0.0042

Error                        6      10.5000000       1.7500000

Corrected Total             15     153.0000000
```

Since p-value for treatment effect is small ($= 0.0042$), I conclude that there is a difference between the four assembly methods.

(b) The treatment effects $\tau_j, j = A, B, C, D$ are estimated by

$$\hat{\tau}_j = \bar{y}_{\cdot j \cdot} - \bar{y}_{\cdots}, \qquad j = A, B, C, D.$$

The overall mean $\bar{y}_{\cdots}$ and the treatment group means $\bar{y}_{\cdot j \cdot}$ can be obtained from the SAS output of PROC GLM.

```
R-Square     Coeff Var      Root MSE        y Mean

0.931373      12.90610      1.322876      10.25000

                 trt         y LSMEAN

                 1          7.5000000
                 2          9.2500000
                 3         13.2500000
                 4         11.0000000
```

Plugging in these values, I get the four treatment effect estimates as below

$\hat{\tau}_A$	$\hat{\tau}_B$	$\hat{\tau}_C$	$\hat{\tau}_D$
-2.75	-1.00	3.00	0.75

(c) The critical distance for Tukey's pairwise comparisons method is

$$CD = q_{\alpha,p,(p-2)(p-1)} \sqrt{MS_E/p} = q_{0.05,4,6} \sqrt{1.75/4} = 3.24.$$

The four treatment means are ordered as $13.25 > 11.00 > 9.25 > 7.50 (C > D > B > A)$. After computing differences following this order and comparing them with the critical distance, I reach the following conclusion.

- The pairs of assembly methods which have significantly different effects are $(C, B), (C, A), (D, A)$.
- The pairs of assembly methods whose effects are not significantly different are $(C, D), (D, B), (B, A)$.

(d) The diagnostic plots in Figure **??** are: normal probability Q-Q plot, plot of residuals versus assembly methods (treatment), plot of residuals versus assembly orders (row block), plot of residuals versus operators (column block), and plot of residuals versus predicted values.

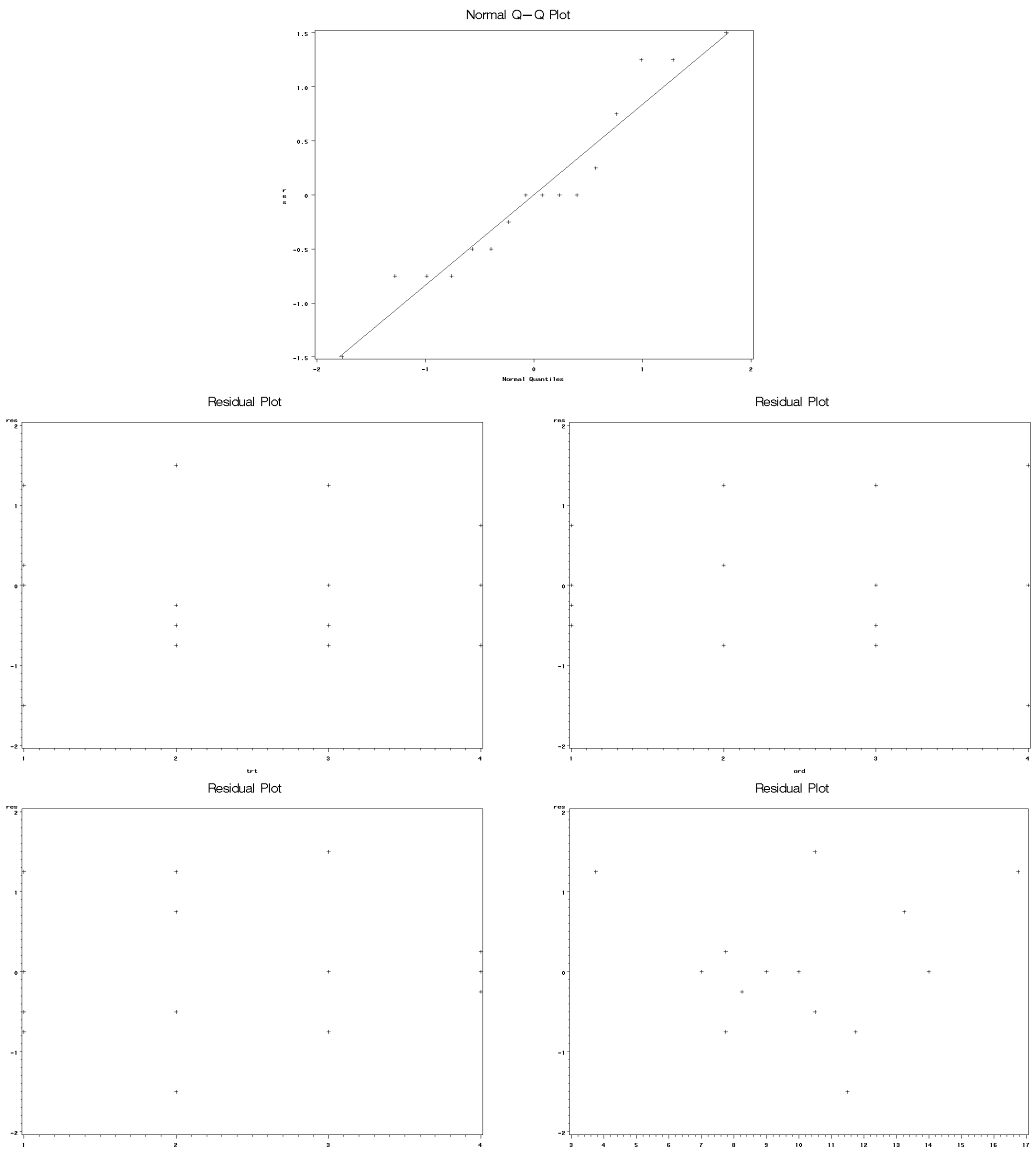

Figure 1: Diagnostic Plots

The normal Q-Q plot shows that the normality assumption is valid. And there are no potential outliers or influential points in the plots. Only the plot of residuals against predicted values shows some curvilinearity, but this is not enough to question on the additivity assumption since our sample size is small. ∎

Problem 2

(a) This is a 4×4 Graeco-Latin square design. It superimposes on the Latin square of 4 assembly methods another Latin square of 4 workplaces. And these two Latin squares are orthogonal to each other, that is, each assembly method in the first Latin square is paired with each workplace in the second Latin square exactly once.

(b) The ANOVA table from SAS is as follows (again, I replaced the line for the model SS by lines for the block SS and the treatment SS).

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
ord	3	0.50000000	0.16666667	0.02	0.9960
opt	3	19.00000000	6.33333333	0.69	0.6157
trt	3	7.50000000	2.50000000	0.27	0.8429
wp	3	95.50000000	31.83333333	3.47	0.1669
Error	3	27.5000000	9.1666667		
Corrected Total	15	150.0000000			

The p-value for the treatment effect is very large ($= 0.8429$), so I conclude that the four assembly methods are not different.

(c) My conclusion here is inconsistent with that from Problem 1. First, our data are different from those in Problem 1 and seem to have less variation due to assembly methods (treatment SS here, 7.5, is only about 1/10 of that in Problem 1, 72.5). Second, the Graeco-Latin square design reduces the degree of freedom for MS_E from 6 to 3, which may cause the F test for the treatment effect less sensitive. ∎

Problem 3

(a) We have $a = 5$ treatments (gasoline additives) and $b = 5$ blocks (cars) with $a \leq b$. Each block contains $k = 4$ treatments, each treatment appears in $r = 4$ blocks, and each pair of treatments appears in the same blocks $\lambda = 3$ times. The total number of runs $N = ar = bk = 20$, and $\lambda(a-1) = r(k-1) = 12$. Hence this is a balanced incomplete block design.

(b) The ANOVA table from SAS is as follows (line for Model SS replaced by Type I SS of the block and the treatment)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
car	4	31.20000000	7.80000000	8.57	0.0022
trt	4	35.73333333	8.93333333	9.81	0.0012
Error	11	10.01666667	0.91060606		
Corrected Total	19	76.95000000			

Since the p-value for treatment effect is very small ($= 0.0012$), I conclude that there is a difference between the five gasoline additives.

(c) First, the row and column sums are calculated based on the original data table.

	car1	car2	car3	car4	car5	$y_{i\cdot}$
add1		17	14	13	12	56
add2	14	14		13	10	51
add3	12		13	12	9	46
add4	13	11	11	12		47
add5	11	12	10		8	41
$y_{\cdot j}$	50	54	48	50	39	

And the overall mean

$$\bar{y}_{\cdot\cdot} = \frac{1}{N}\sum_j y_{\cdot j} = \frac{50+54+48+50+39}{20} = 12.05.$$

Second, let's compute

$$Q_i = y_{i\cdot} - \frac{1}{k}\sum_j n_{ij} y_{\cdot j}, \qquad i = 1,2,3,4,5,$$

where n_{ij} equals 1 if treatment i appears in block j and 0 otherwise.

$$\begin{aligned}
Q_1 &= y_{1\cdot} - \frac{1}{4}\sum_{j=1}^{5} n_{1j} y_{\cdot j} = 56 - \frac{54+48+50+39}{4} = 8.25, \\
Q_2 &= y_{2\cdot} - \frac{1}{4}\sum_{j=1}^{5} n_{2j} y_{\cdot j} = 51 - \frac{50+54+50+39}{4} = 2.75, \\
Q_3 &= y_{3\cdot} - \frac{1}{4}\sum_{j=1}^{5} n_{3j} y_{\cdot j} = 46 - \frac{50+48+50+39}{4} = -0.75, \\
Q_4 &= y_{4\cdot} - \frac{1}{4}\sum_{j=1}^{5} n_{4j} y_{\cdot j} = 47 - \frac{50+54+48+50}{4} = -3.5, \\
Q_5 &= y_{5\cdot} - \frac{1}{4}\sum_{j=1}^{5} n_{5j} y_{\cdot j} = 41 - \frac{50+54+48+39}{4} = -6.75.
\end{aligned}$$

Finally, the adjusted means are calculated by

$$\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \bar{y}_{\cdot\cdot} + \frac{kQ_i}{\lambda a}.$$

Hence

$$\begin{aligned}
\hat{\mu}_1 &= 12.05 + \frac{4\cdot 8.25}{3\cdot 5} = 14.25, \\
\hat{\mu}_2 &= 12.05 + \frac{4\cdot 2.75}{3\cdot 5} = 12.7833, \\
\hat{\mu}_3 &= 12.05 + \frac{4\cdot(-0.75)}{3\cdot 5} = 11.85, \\
\hat{\mu}_4 &= 12.05 + \frac{4\cdot(-3.5)}{3\cdot 5} = 11.1167, \\
\hat{\mu}_5 &= 12.05 + \frac{4\cdot(-6.75)}{3\cdot 5} = 10.25.
\end{aligned}$$

(d) The standard error of the difference between two treatment estimates is

$$\sqrt{\hat{Var}(\hat{\tau}_i - \hat{\tau}_j)} = \sqrt{\frac{2k}{\lambda a} MS_E} = \sqrt{\frac{2\cdot 4}{3\cdot 5}\cdot 0.9106} = 0.6969.$$

(e) The critical difference for Tukey's pairwise comparisons is

$$CD = \frac{q_{\alpha,a,ar-a-b+1}}{\sqrt{2}}\sqrt{\frac{2k}{\lambda a}MS_E} = \frac{q_{0.05,5,11}}{\sqrt{2}} \cdot 0.6969 = \frac{4.58}{\sqrt{2}} \cdot 0.6969 = 2.2569.$$

The five treatment mean estimates are ordered as $14.25 > 12.7833 > 11.85 > 11.1167 > 10.25$ ($\hat{\mu}_1 > \hat{\mu}_2 > \hat{\mu}_3 > \hat{\mu}_4 > \hat{\mu}_5$). After computing differences following this order and comparing them with the critical distance, I reach the following conclusion.

- The pairs of gasoline additives which have significantly different mileage performances are $(1,3),(1,4),(1,5),(2,5)$.
- The mileage performances within any other pairs of gasoline additives are not significantly different.

The results from SAS with the options "`lsmeans trt / tdiff adjust=tukey;`" are given below.

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

trt	y LSMEAN	LSMEAN Number
1	14.2500000	1
2	12.7833333	2
3	11.8500000	3
4	11.1166667	4
5	10.2500000	5

Least Squares Means for Effect trt
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: y

i/j	1	2	3	4	5
1		2.104587	3.443869	4.496162	5.739781
		0.2838	0.0355	0.0065	0.0010
2	-2.10459		1.339282	2.391576	3.635195
	0.2838		0.6746	0.1884	0.0259
3	-3.44387	-1.33928		1.052293	2.295913
	0.0355	0.6746		0.8262	0.2167
4	-4.49616	-2.39158	-1.05229		1.243619
	0.0065	0.1884	0.8262		0.7280
5	-5.73978	-3.63519	-2.29591	-1.24362	
	0.0010	0.0259	0.2167	0.7280	

In the SAS output, the pairs that are significant different have p-values (the bottom values for entries in the output table) less than 0.05. Hence, the significantly different pairs from SAS are $(1,3),(1,4),(1,5),(2,5)$, which are consistent with my conclusion.

(f) The contrast I use is $\mathbf{C} = (1,1,0,-1,-1)'$. The significance test of it is shown below.

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
C	1	30.10416667	30.10416667	33.06	0.0001

Since the p-value for the test is very small ($= 0.0001$), I conclude that the combination of additives 1 and 2 has significantly different characteristics from the combination of additives 4 and 5. ∎