STAT 514 Homework 6

Due: Oct 21

1. Ten needles were randomly selected from a branch of a loblolly pine tree. The stomata (microscopic breathing holes) are arranged in rows. On each needle, four rows are randomly selected and the number of stomata per centimeter for each of the rows was determined. The data below is in the file named "stomata.dat".

Needle									
1	2	3	4	5	6	7	8	9	10
149	136	143	121	148	129	127	134	117	129
143	139	142	133	121	134	130	137	128	132
138	129	124	126	124	127	123	119	117	131
131	143	134	130	128	113	125	130	118	137

(a) Why is the random effects model appropriate here?
solution:
The random effects model is appropriate since both the needles and the rows are randomly seleted.

(b) Estimate all relevant variance components (Do not use the REML estimator from *proc mixed*).
solution:
$SS_{trt} = 1299.725$, $MS_{trt} = 144.414$, $SS_E = 1614.25$, $MS_E = 53.81$, thus $\hat{\sigma}^2 = 53.81$, $\hat{\sigma}_\tau^2 = (MS_{trt} - MS_E)/n = 22.65$ 29.62% of the overall variation in stomata number per centimeter is due to the needle.

(c) What percentage of the overall variation in stomata number per centimeter is due to the needle?
solution:

$$\hat{\rho} = \frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2 + \hat{\sigma}^2} = 0.296$$

(d) Construct a 95% CI for this ratio.
solution:
The confidence interval is $[L/(L+1), U/(U+1)]$

$$L = \frac{1}{n}(\frac{MS_{trt}}{MS_E F_{1-\alpha/2}} - 1) = 0.0106, \quad U = \frac{1}{n}(\frac{MS_{trt}}{MS_E F_{\alpha/2}} - 1) = 2.1389$$

Thus the 95% confidenc interval for the ratio is [0.01049, 0.6814]

(e) Compute a 95% confidence interval for the average number of stomata per centimeter (i.e. grand mean μ).
solution:

$$\hat{\mu} = \bar{Y}_{..} = 130.475$$

$$\bar{Y}_{..} \pm t(1-\alpha/2, a-1)\sqrt{\frac{MS_{trt}}{rn}} = 130.475 \pm 2.262 * \sqrt{144.41/40} = [126.18, 134.77]$$

note the degree of freedom is a-1, not N-a

2. A sociologist is interested in studying the IQs of teachers from low income areas of a major city. Six schools were randomly chosen from low income areas and from each of these schools, five teachers were randomly chosen. The following table summarizes the mean IQ for each of these schools (NOTE: These numbers are all made up and are not intended to reflect teachers' true IQ scores).

School	1	2	3	4	5	6
Mean	97	99	94	109	98	103

(a) If $MS_E = 40$, is there significant variability in average IQ among schools in low income areas (use α=0.01)?
solution:
The grand sample mean is 100. This means the treatment sum of squares is

$$SST_{trt} = 5(3^2 + 1^2 + 6^2 + 9^2 + 2^2 + 3^2) = 700$$

The F test is equal to $(700/5)/40 = 3.5$. The critical F value for 5 and 24 degrees of freedom is 3.90. Since this is larger than the test statistic we do not reject. There is not sufficient evidence to reject the null hypothesis that the school variance is zero.

(b) Estimate all variance components.
solution:
The variance estimate for the error variance (between teachers) is 40 and the school variance is $(140 - 40)/5 = 20$.

(c) How much power does this study have if the true variances were such that $2\sigma_\tau^2 = \sigma^2$ and n were increased to 10?
solution:
If $2\sigma_\tau^2 = \sigma^2$ (as is observed here) and n were 10, the value of $\lambda = 1 + 10(.5) = 2.45$ and the degrees of freedom error are 54. For $\alpha = .01$, this equates to $\beta = 25\%$ or power of 75%. This means that this experiment with $n = 5$ is very underpowered to detect such a small difference as that observed in this experiment.

(d) Suppose the national average IQ for teachers is 105. Test the null hypothesis that the average IQ of these teachers is not lower than the national average ($\alpha = 0.05$)
solution:

In class we showed that in a random effects model, the MST_{trt} should be used in place of the MS_E when computing the standard error of the grand mean. The test statistic is

$$\frac{100-105}{\sqrt{140/30}} = -2.31$$

We have **5 (not N-a=24)** degrees of freedom so the critical t value is -2.015. Since this is smaller, we reject the null hypothesis. It does appear that the average IQ in these schools is lower than the national average.

3. An industrial engineer is investigating the effects of four assembly methods (A, B, C, D) on the assembly time for a color television component. Four operators are selected for the study. Furthermore, the engineer knows that each assembly method produces such fatigue that the time required for the last assembly may be greater than the time required for the first, regardless of the method. That is, a trend develops in the required assembly time. To account for this source of variability, the engineer uses the Latin square design shown below:

Order of	Operator			
Assemby	1	2	3	4
1	$C=10$	$D=14$	$A=7$	$B=8$
2	$B=7$	$C=18$	$D=11$	$A=8$
3	$A=5$	$B=10$	$C=11$	$D=9$
4	$D=10$	$A=10$	$B=12$	$C=14$

(a) Test if there is a difference between the four assembly methods. State the hypotheses and use $\alpha = 5\%$.
solution:
ANOVA table output from SAS (I replaced the line for the model SS by lines for the two block SS and the treatment SS).

Source	DF	Sum of 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.

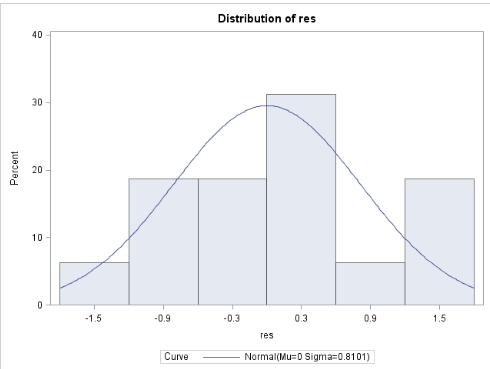
(b) Obtain the estimates of the treatment effects
solution:
The treatment effects $\hat{\tau}_j = \bar{y}_{.j.} - \bar{y}_{...}$, hence $\hat{\tau}_A = -2.75$, $\hat{\tau}_B = -1.00$, $\hat{\tau}_A = 3.00$, $\hat{\tau}_A = 0.75$.

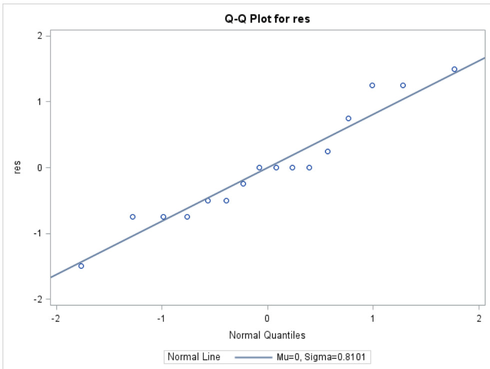
(c) Use Tukey's method to perform pairwise comparison by hand.
solution:
The critical distance for Tukey's pairwise comparisons method is

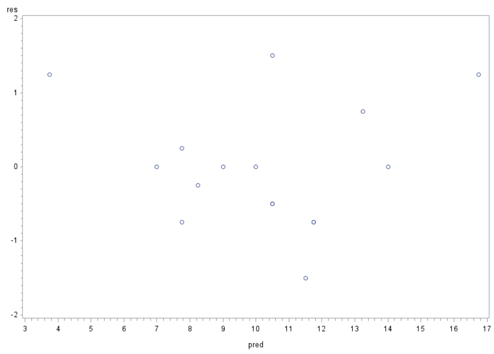
$$CD = q_{\alpha,p,(p-2)(p-1)}\sqrt{MS_E/p} = q_{0.05,4,6}\sqrt{1.75/4} = 3.24.$$

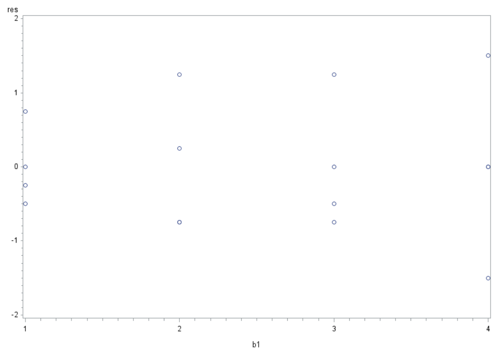
The four treatment means are ordered as $13.25 > 11.00 > 9.25 > 7.50 (C > D > B > A)$. After computing differences following this order and comparing them with the critical distance, I reach the following conclusion.
The pairs of assembly methods which have significantly different effects are (C, B), (C, A), (D, A).
The pairs of assembly methods whose effects are not significantly different are (C, D), (D, B), (B, A).

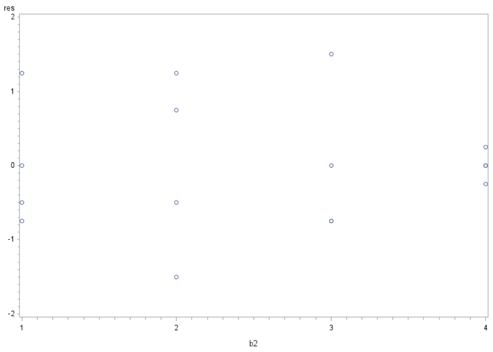
(d) Use residuals to check normality and independence assumptions.
solution:
The diagnostic are: normal probability Q-Q plot, histogram of residuals, plot of residuals versus assembly methods (treatment), plot of residuals versus assembly orders (row block), plot of residuals versus operators (column block), and plot of residuals versus predicted values. 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.

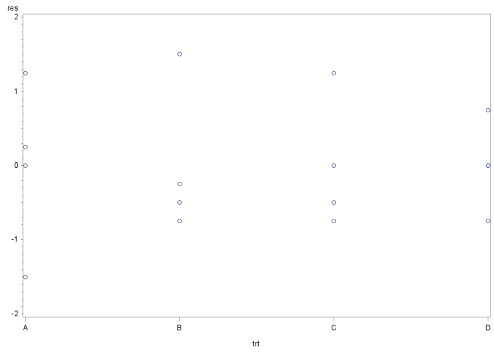












4. Suppose in Problem 1 the engineer suspects that the workplaces used by the four operators may represent an additional source of variation. A fourth factor, workplace (α,β, γ and γ) needs to be considered and another experiment is conducted. The layout of the experiment and the data are given in the following.

Order of Assemby	Operator 1	2	3	4
1	$C\beta = 11$	$B\gamma = 10$	$D\delta = 14$	$A\alpha = 8$
2	$B\alpha = 8$	$C\delta = 12$	$A\gamma = 10$	$D\beta = 12$
3	$A\delta = 9$	$D\alpha = 11$	$B\beta = 7$	$C\gamma = 15$
4	$D\gamma = 9$	$A\beta = 8$	$C\alpha = 18$	$B\delta = 6$

(a) What design is employed in this experiment? why?
solution:
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) Test if the four assembly methods are different.(use α= 5%).
solution:
The ANOVA table from SAS is as follows

```
                      Sum of
Source           DF   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    95.50000000   31.83333333  3.47      0.1669
wp               3    7.50000000    2.50000000   0.27      0.8429
Error            3    27.5000000   9.1666667
Corrected Total  15   150.0000000
```

The p-value for the treatment effect is large (= 0.1669), so I conclude that the four assembly methods are not different.

(c) Is your conclusion consistent with that from Problem 1? If your answer is no, what are the possible causes for the inconsistency?

solution:

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 M SE from 6 to 3, which may cause the F test for the treatment effect less sensitive.