Quiz Practice Questions 6 (Attendance 12) for Statistics 514
Design of Experiments
Chapter 30 Neter et al. and Kuhn

These are practice questions for the quiz. The quiz (not the practice questions) is worth 5% and marked out of 5 points. One or more questions is closely, but not necessarily exactly, related to one or more of these questions will appear on the quiz. These practice questions are *not* to be handed in. Quizzes are to be done *using Vista* on the Internet **before** 4am (West Lafayette time!) of the date of the quiz. Vista will *not* allow any quiz to be done late. It is *highly* recommended that you complete this practice quiz, by hand, *before* logging onto Vista. The quiz is an **individual** one which means that each student does this quiz by themselves without help from others.

1. Applied Linear Statistical Models (Neter et al.) Questions.

Chapter	Problem(s)	hints
30, pages 1229–1233	30.6, 30.7, 30.8, 3.10, 3.14	Summary reports
	30.15, 30.16	TV commercials
	30.17, 30.18	Recall decay

(30.6) Summary Reports: qz6-30-6-report-latin-residuals

From SAS,

$e_{ij} = e_{55} = ?$.

The residuals versus fitted values plot appears to be randomly scattered and so this indicates constant variance.

The normal probability plot appears to be a straight line and so this indicates normality.

(30.7) Summary Reports: qz6-30-7-report-latin-inference

(a) *Normal probability plot of estimated treatment means.*
$\overline{Y}_{\cdot\cdot 1} = 7.0,\ \overline{Y}_{\cdot\cdot 2} = 7.4,\ \overline{Y}_{\cdot\cdot 3} = 15.0,\ \overline{Y}_{\cdot\cdot 4} = 19.0,\ \overline{Y}_{\cdot\cdot 5} = 13.4$
and the overall mean is $\hat{Y}_{\cdots} = 12.36$
and from the ANOVA table below,
$\sqrt{\frac{MSRem}{n}} = \sqrt{\frac{1.607}{5}} \approx 0.567$

report	$\hat{Y}_{\cdot\cdot i}$	rank k	$z\left(\frac{k-0.375}{5+0.25}\right)$	$\hat{Y}_{\cdots} + z\left(\frac{k-0.375}{r+0.25}\right)\sqrt{\frac{MSRem}{n}}$
A	7.0	1	-1.17976	11.691
B	7.4	2	-0.49720	12.078
C	15.0	4	?	?
D	19.0	5	1.17976	13.029
E	13.4	3	0	12.36

Plotting the observed means $(\hat{Y}_{\cdot\cdot i})$ versus $z\left(\frac{k-0.375}{5+0.25}\right)$
and expected means $(\hat{Y}_{\cdots} + z\left(\frac{k-0.375}{r+0.25}\right)\sqrt{\frac{MSE}{n}})$ versus $z\left(\frac{k-0.375}{5+0.25}\right)$
and then comparing these two plots
demonstrates that the
treatment means are substantially different from one another
because the two plots do *not* overlap one another[1].

(b) *ANOVA table and Test of Reports.*
From SAS,

Source	*df*	*SS*	*MS*
Executives (Rows)	4	220.16	55.040
Months (Columns)	4	10.96	2.740
Ratings (Treatments)	4	?	?
Error	12	19.28	1.607
Total	24	777.76	

and so
$H_0 : \tau_i = 0$ versus
$H_a :$ at least one $\tau_i \neq 0,\ i = 1, 2, 3, 4, 5.$
since p–value $P(F > \frac{MSTR}{MSRem} = ?; 4, 12) = 0 < \alpha = 0.01$
reject null; that is, average helpfulness ratings of the different summary reports are different

[1]See Chapter 17.2 of either the workbook or the text.

(c) *Pairwise comparisons, Doses, Tukey*

From the SAS output,
$\hat{L}_1 = \overline{Y}_{\cdot\cdot 1} - \overline{Y}_{\cdot\cdot 2} = -0.4$,
$\hat{L}_2 = \overline{Y}_{\cdot\cdot 1} - \overline{Y}_{\cdot\cdot 3} = -8.0$,
$\hat{L}_3 = \overline{Y}_{\cdot\cdot 1} - \overline{Y}_{\cdot\cdot 4} = -12.0$,
$\hat{L}_4 = \overline{Y}_{\cdot\cdot 1} - \overline{Y}_{\cdot\cdot 5} = ?$,
$\hat{L}_5 = \overline{Y}_{\cdot\cdot 2} - \overline{Y}_{\cdot\cdot 3} = -7.6$,
$\hat{L}_6 = \overline{Y}_{\cdot\cdot 2} - \overline{Y}_{\cdot\cdot 4} = -11.6$,
$\hat{L}_7 = \overline{Y}_{\cdot\cdot 2} - \overline{Y}_{\cdot\cdot 5} = -6.0$,
$\hat{L}_8 = \overline{Y}_{\cdot\cdot 3} - \overline{Y}_{\cdot\cdot 4} = -4.0$,
$\hat{L}_9 = \overline{Y}_{\cdot\cdot 3} - \overline{Y}_{\cdot\cdot 5} = 1.6$,
$\hat{L}_{10} = \overline{Y}_{\cdot\cdot 4} - \overline{Y}_{\cdot\cdot 5} = 5.6$,
$s\{\hat{L}_i\} = \sqrt{MSE\left(\frac{1}{n} + \frac{1}{n}\right)} = \sqrt{1.607\left(\frac{1}{5} + \frac{1}{5}\right)} \approx 0.8017$
$T = \frac{1}{\sqrt{2}} q(1-\alpha; r, (r-1)(r-2)) = \frac{1}{\sqrt{2}} q(0.95; 5, 12) = \frac{?}{\sqrt{2}} \approx ?$
(where q is obtained from table B.9)
and so the CIs are
$-0.4 \pm 3.189(0.8017) = -2.96 \leq L_1 \leq 2.16$
$-8.0 \pm 3.189(0.8017) = -10.56 \leq L_2 \leq -5.44$
$-12.0 \pm 3.189(0.8017) = -14.56 \leq L_3 \leq -9.44$
$? \pm 3.189(0.8017) = ? \leq L_4 \leq ?$
$-7.6 \pm 3.189(0.8017) = -10.16 \leq L_5 \leq -5.04$
$-11.6 \pm 3.189(0.8017) = -14.16 \leq L_6 \leq -9.04$
$-6.0 \pm 3.189(0.8017) = -8.56 \leq L_7 \leq -3.44$
$-4.0 \pm 3.189(0.8017) = -6.56 \leq L_8 \leq -1.44$
$1.6 \pm 3.189(0.8017) = -0.96 \leq L_9 \leq 4.16$
$5.6 \pm 3.189(0.8017) = 3.04 \leq L_{10} \leq 8.16$

(30.8) Summary Reports: qz6-30-8-report-latin-efficiency

(a) *Efficiency Measures*

$$\begin{aligned}
\hat{E}_1 &= \frac{MSROW + MSCOL + (r-1)MSRem}{(r+1)MSRem} \\
&= \frac{55.040 + 2.740 + (5-1)(1.607)}{(5+1)(1.607)} = 6.66 \\
\hat{E}_2 &= \ ? \\
\hat{E}_3 &= \frac{MSROW + (r-1)MSRem}{r\,MSRem} \\
&= \frac{55.040 + (5-1)(1.607)}{5(1.607)} = \ ?
\end{aligned}$$

(b) *How effective?*

- The $\hat{E}_1 = 6.66$ indicates the Latin square design reduces the error variance 6.66–fold, as compared to the comparable completely randomized (CRD) design.
- The $\hat{E}_2 = ?$ indicates the Latin square design reduces the error variance ?–fold, as compared to the comparable randomized block design (RBD), using the row as the block.
- The $\hat{E}_3 = ?$ indicates the Latin square design reduces the error variance ?–fold, as compared to the comparable randomized block design (RBD), using the column as the block.

(30.10) Summary Reports

Since

$$\begin{aligned} \phi &= \frac{1}{\sigma}\sqrt{\sum \tau_k^2} \\ &= \frac{1}{1.4}\sqrt{(-2)^2 + (-1)^2 + (0)^2 + (1.5)^2 + (1.5)^2} \\ &= \ ? \end{aligned}$$

and $\nu_1 = r - 1 = 5 - 1 = 4$
$\nu_2 = (r - 1)(r - 2) = (4)(3) = 12$
and so, using Table B.11 (p 1356),
$1 - \beta$ is somewhere between 0.58 and 0.83
In other words, the power (where zero (0) is "poor" and one (1) is "excellent") of the test associated with the latin square in this case is so–so.

(30.14) Summary Reports: qz6-30-14-report-latin-regression

(a) *Regression approach*
the full model is

$$\begin{aligned} Y_{ijk} = \mu_{\cdots} &+ \rho_1 X_{ijk1} + \rho_2 X_{ijk2} + \rho_3 X_{ijk3} + \rho_4 X_{ijk4} \\ &+ \kappa_1 X_{ijk5} + \kappa_2 X_{ijk6} + \kappa_3 X_{ijk7} + \kappa_4 X_{ijk8} \\ &+ \tau_1 X_{ijk9} + \tau_2 X_{ijk10} + \tau_3 X_{ijk11} + \tau_4 X_{ijk12} \\ &+ \varepsilon_{ijk} \end{aligned}$$

where

$$X_{ijk1} = \begin{cases} 1, & \text{if case from row blocking 1} \\ -1, & \text{if case from row blocking 5} \\ 0, & \text{otherwise,} \end{cases}$$

and X_{ijk2}, X_{ijk3}, X_{ijk4} defined similarly

$$X_{ijk5} = \begin{cases} 1, & \text{if case from column blocking 1} \\ -1, & \text{if case from column blocking 5} \\ 0, & \text{otherwise,} \end{cases}$$

and X_{ijk6}, X_{ijk7}, X_{ijk8} defined similarly

$$X_{ij9} = \begin{cases} 1, & \text{if case from treatment 1} \\ -1, & \text{if case from treatment 5} \\ 0, & \text{otherwise,} \end{cases}$$

and X_{ijk10}, X_{ijk11}, X_{ijk12} defined similarly

and so an estimate of the *full* model (where $SSE(F) = ?$) is

$$\begin{aligned} \hat{Y} = 12.54 &+ 1.91X_1 - ?X_2 + 3.26X_3 - 3.29X_4 \\ &+ 1.11X_5 - 0.34X_6 - 0.94X_7 - 0.72X_8 \\ &- 5.54X_9 - .14X_{10} + 3.11X_{11} + 6.71X_{12} \end{aligned}$$

and also an estimate of the *reduced* model (where $SSE(R) = 494.24$) is

$$\begin{aligned} \hat{Y} = 11.96 &+ 0.45X_1 - 2.96X_2 + 3.84X_3 - 3.56X_4 \\ &- 0.35X_5 + 0.24X_6 - 0.36X_7 - 0.16X_8 \end{aligned}$$

and so
$H_0 : \tau_i = 0$ versus $H_a :$ at least one $\tau_i \neq 0,\ i = 1, 2, 3, 4.$

$$\frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} = \frac{494.2353 - \ ?}{14 - 10} \div \frac{?}{10} = 95.34$$

and so p–value is $P(F > 95.34; 4, 10) \approx 0$
since p–value $= 0 < \alpha = 0.01$
reject null; that is, average ratings for different reports are different

(b) $L = \tau_4 - \tau_1$
Using either SAS (use value along top row associated with column X12 for $\hat{\tau}_4$, for example)
or directly ($\hat{\tau}_4$ is the coefficient of X12 in full regression, since X12 = 1, and all else is zero)
$\hat{L} = \hat{\tau}_4 - \hat{\tau}_1 = 6.71429 - (?) = ?$
and using the SAS output (use row X9 and column X9 for $s^2\{\hat{\tau}_1\}$, for example)
$s^2\{\hat{\tau}_1\} = 0.20927$
$s^2\{\hat{\tau}_4\} = ?$
$s\{\hat{\tau}_1, \hat{\tau}_4\} = -0.06134$
and so $s\{\hat{L}\} = \sqrt{s^2\{\hat{\tau}_1\} + s^2\{\hat{\tau}_4\} - 2s\{\hat{\tau}_1, \hat{\tau}_4\}} = 0.7832$
also $t(0.995, 10) = 3.169$
and so a 99% confidence interval is
$12.25715 \pm 3.169(0.77832) = (9.775, 14.739)$

(30.15) Summary Reports: qz6-30-15-tv-latin-replications-residuals

From SAS,

$e_{ijk} = e_{114} = ?$.

The residuals versus fitted values plot appears to be randomly scattered and so this indicates constant variance.

The normal probability plot appears to be a straight line and so this indicates normality.

(30.16) Summary Reports: qz6-30-16-tv-latin-replications-inference

(a) *Model*

$$Y_{ijkl} = \mu_{...} + \rho_i + \kappa_j + \,?\, + \beta_l + (\alpha\beta)_{kl} + \varepsilon_{ijkl}$$

(b) *ANOVA table and test of interaction.*
From SAS,

Source	*df*	*SS*	*MS*
Age (Rows)	3	658.09375	219.36458
Education (Columns)	3	18.34375	6.11458
Commercials (Treatments)	3	1251.34375	417.11458
Volume	1	?	?
Product	1	850.78125	850.78125
Interaction	1	?	?
Error	22	285.4375	12.97443
Total	31	2213.21875	

and so
$H_0 : (\alpha\beta)_{kl} = 0$ versus
$H_a :$ at least one $(\alpha\beta)_{kl} \neq 0,\ k, l = 1, 2.$
since p–value $P(F > \frac{MSAB}{MSRem} = ?; 1, 22) = 0.73 > \alpha = 0.01$
accept null; that is, there is no interaction

(c) *Test volume and product main effects.*

$H_0 : \alpha_k = 0$ versus
$H_a :$ at least one $\alpha_k \neq 0,\ k = 1, 2.$
since p–value $P(F > \frac{MSA}{MSRem} = \frac{399.03125}{12.97443} = ?; 1, 22) = 0 < \alpha = 0.01$
reject null; that is, average points different for different volumes

$H_0 : \beta_l = 0$ versus
$H_a :$ at least one $\beta_l \neq 0,\ l = 1, 2.$
since p–value $P(F > \frac{MSB}{MSRem} = ?; 1, 22) = 0 < \alpha = 0.01$
reject null; that is, average points different for different products

(30.17) Summary Reports: qz6-30-17-recall-crossover-residuals

From SAS,

$e_{ijk} = e_{212} = ?$.

The residuals versus fitted values plot appears to be randomly scattered and so this indicates constant variance.

The normal probability plot appears to be a straight line and so this indicates normality.

(30.18) Summary Reports: qz6-30-18-recall-crossover-inference

(a) *ANOVA table and test of interaction.*
From SAS,

Source	*df*	*SS*	*MS*
Pattern	2	?	?
Time Period	2	1803.6296	901.8148
Questionnaire	2	3472.0741	1736.037
Subjects(within pattern)	6	159.5556	26.5926
Error	14	194.9630	13.9259
Total	26	5644.5185	

and so
$H_0 : \rho_i = 0$ versus
$H_a :$ at least one $\rho_i \neq 0, \ i = 1, 2, 3.$
since p–value $P(F > \frac{MSP}{MSS} = \frac{7.1481}{26.5926} = 0.269; 2, 6) = \,? > \alpha = 0.05$
accept null; that is, the pattern effect is *not* significant
(average trip count same for different patterns)

$H_0 : \kappa_j = 0$ versus
$H_a :$ at least one $\kappa_j \neq 0, \ k = 1, 2, 3.$
since p–value $P(F > \frac{MSO}{MSRem} = ?; 2, 14) = 0 < \alpha = 0.05$
reject null; that is, average trip count different for different time periods

$H_0 : \tau_k = 0$ versus
$H_a :$ at least one $\tau_k \neq 0, \ l = 1, 2, 3.$
since p–value $P(F > \frac{MSTR}{MSRem} = \frac{1736.037}{13.9259} = 129.66; 2, 14) = \,? < \alpha = 0.05$
reject null; that is, average trip count different for different questionnaires

(b) *Pairwise comparisons, Tukey*
From the SAS output,
$\overline{Y}_{..1.} = 22.3333, \ \overline{Y}_{..2.} = 22.4444, \ \overline{Y}_{..3.} = 46.4444,$
$\hat{L}_1 = \overline{Y}_{..1} - \overline{Y}_{..2} = -0.1111,$
$\hat{L}_2 = \overline{Y}_{..1} - \overline{Y}_{..3} = ?,$
$\hat{L}_3 = \overline{Y}_{..1} - \overline{Y}_{..4} = -24.0000,$
$s\{\hat{L}_i\} = \sqrt{MSE\left(\frac{1}{n} + \frac{1}{n}\right)} = \sqrt{13.9259\left(\frac{1}{9} + \frac{1}{9}\right)} \approx 1.75916$
$T = \frac{1}{\sqrt{2}} q(1 - \alpha; r, (r-1)(nr-2)) = \frac{1}{\sqrt{2}} q(0.95; 3, 14) = \frac{3.16}{\sqrt{2}} \approx ?$
(where q is obtained from table B.9)
and so the CIs are
$-0.1111 \pm ?(1.75916) = -4.0411 \leq L_1 \leq 3.8189$
$? \pm ?(1.75916) = -28.0411 \leq L_2 \leq -20.1811$
$-24 \pm ?(1.75916) = -27.93 \leq L_3 \leq -20.07$