Quiz Practice Questions 5 (Attendance 10) for Statistics 514
Design of Experiments
Chapter 28 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	Exercise(s)	hints
28, pages 1155–1161	28.4, 28.5, 28.6, 28.7, 28.8	Bottling plant production
	28.14, 28.15	Internal control
	28.17, 28.18	Questionnaire color

(28.4) Bottling plant production: qz5-28-4-bottle-nested-residuals

(a) *Residuals*
The residuals are given on the SAS output; in particular,
$e_{ijk} = e_{211} = ?$.
The residuals versus fitted values plot appears to be randomly scattered and so this indicates constant variance.
The normal probability plot does *not* appear to be a straight line and so this indicates non–normality.

(b) *Vertical Bar Plot*
Notice on the vertical bar plots by machine on the SAS output that the *range* of the residuals for each machine is the *same* (-7.5 to 7.5) and also each is normal shaped (there are fewer residuals at the upper and lower ends than in the middle of each group of residuals). This second conclusion seems to contradict the results of the normal probability plot.

(28.5) Bottling plant production: qz5-28-5-bottle-nested-inference

(a) *Operator, shift distinguishable?*
Notice it is impossible to tell which operator in what shift (evening, morning, afternoon, say) produced a given number of cases.

(b) *Average Cases Produced Per Machine (i) and Operator (j)*

i	$j = 1$	$j = 2$	$j = 3$	$j = 4$	ave
1	61.8	67.8	62.6	52.6	61.2
2	75.8	75.2	55.8	77.0	70.95
3	76.8	69.6	74.4	73.4	73.55
ave	71.47	70.87	64.27	67.67	68.57

(c) *Analysis of Variance Table*

Source	*df*	*SS*	*MS*
Machines, A	2	1,695.63	847.817
Operators, within Machines, B(A)	9	2,272.30	252.478
Error	48	1,132.80	?
Total	59	5,100.73	

(d) *Test Machine Effects*
$H_0 : \alpha_i = 0$ versus
$H_a :$ at least one $\alpha_i \neq 0,\ i = 1, 2, 3.$
since p–value $P(F > \frac{847.817}{23.6} = 35.924; 2, 48) = ? < \alpha = 0.01$
reject null; that is, machine effects are significant
(mean number of cases produced different for different machines)

(e) *Test (Overall) Operator Effects*
$H_0 : \beta_{j(i)} = 0$ versus
$H_a :$ at least one $\beta_{j(i)} \neq 0,\ i = 1, 2, 3.$
since p–value $P(F > \frac{?}{23.6} = 10.698; 9, 48) = ? < \alpha = 0.01$
reject null; that is, operator effects are significant
(mean number of cases produced different for different operators)

(f) *Test Machines Separately For Operator Effect*

Source	*df*	*SS*	*MS*
Machines, A	2	1,695.63	847.817
Operators, within Machines, B(A)	9	2,272.30	252.478
Operators, within Machine 1, B(1)	3	599.2	199.73
Operators, within Machine 2, B(2)	3	1,538.55	512.85
Operators, within Machine 3, B(3)	3	134.55	44.85
Error	48	1,132.80	23.600
Total	59	5,100.73	

Operators using (within) machine 1,
$H_0 : \beta_{j(1)} = 0$ versus
$H_a :$ at least one $\beta_{j(1)} \neq 0,\ i = 1, 2, 3.$
since p–value $P(F > \frac{?}{23.6}; 3, 48) = 0 < \alpha = 0.01$
accept or reject null?

Operators using (within) machine 2,
$H_0 : \beta_{j(2)} = 0$ versus
$H_a :$ at least one $\beta_{j(2)} \neq 0,\ i = 1, 2, 3.$
since p–value $P(F > \frac{512.85}{23.6}; 3, 48) = 0 < \alpha = 0.01$
accept or reject null?

Operators using (within) machine 3,
$H_0 : \beta_{j(3)} = 0$ versus
$H_a :$ at least one $\beta_{j(3)} \neq 0,\ i = 1, 2, 3.$
since p–value $P(F > \frac{44.85}{?}; 3, 48) = 0.14 > \alpha = 0.01$
accept or reject null?

(g) *Bonferroni inequality*
There are $g = 5$ tests in parts (e), (f) and (g), where each test is conducted at a level of significance of $\alpha = 0.01$ and so, using equation (4.4), page 155,

$$\begin{aligned} P\left(\bigcap_{i=1}^{g} \overline{A}_i\right) &\geq 1 - g\alpha \\ P\left(\bigcap_{i=1}^{5} \overline{A}_i\right) &\geq 1 - 5(0.01) \\ &= 0.95 \end{aligned}$$

In other words, the *family* level of significance, α^* is

$$\alpha^* \leq g\alpha = 5(0.01) = ?$$

Since all of the tests have p–values less than 0.05, except for the last one which tests "operators using (within) machine 3" and has a p–value of 0.14, all tests are significant, except the last one.

(28.6) Bottling plant production: qz5-28-6-bottle-nested-effects

(a) *Pairwise comparisons, Machines, Tukey*
From the SAS output,
$\overline{Y}_{1..} = 61.20$, $\overline{Y}_{2..} = 70.95$, $\overline{Y}_{3..} = ?$
$\hat{L}_1 = \overline{Y}_{.1} - \overline{Y}_{.2} = ?$,
$\hat{L}_2 = \overline{Y}_{.1} - \overline{Y}_{.3} = -12.35$,
$\hat{L}_3 = \overline{Y}_{.2} - \overline{Y}_{.3} = -2.60$,
$s\{\hat{L}_i\} = \sqrt{MSE\left(\frac{1}{n} + \frac{1}{n}\right)} = \sqrt{23.6\left(\frac{1}{20} + \frac{1}{20}\right)} \approx 1.536$
($n = 20$ is the number of operators assigned to each machine)
$T = \frac{1}{\sqrt{2}}q(1-\alpha; df) = \frac{1}{\sqrt{2}}q(1-0.05; 3, 48) = \frac{1}{\sqrt{2}}q(0.95; 3, 48) = \frac{3.42}{\sqrt{2}} \approx 2.418$
($df = (3, 48)$ since there are 3 contrasts and 48 error df)
and so the CIs are
$? \pm 2.418(1.536) = ? \leq L_1 \leq ?$
$-12.35 \pm 2.418(1.536) = ? \leq L_2 \leq -8.64$
$-2.60 \pm 2.418(1.536) = -6.31 \leq L_3 \leq 1.11$

(b) *Pairwise comparisons, Operators, Bonferroni*
From the SAS output,
$\overline{Y}_{11.} = 61.8$, $\overline{Y}_{12.} = 67.8$, $\overline{Y}_{13.} = 62.6$, $\overline{Y}_{14.} = 52.6$
$\hat{L}_1 = \overline{Y}_{11.} - \overline{Y}_{12.} = -6.0$,
$\hat{L}_2 = \overline{Y}_{11.} - \overline{Y}_{13.} = ?$,
$\hat{L}_3 = \overline{Y}_{11.} - \overline{Y}_{14.} = 9.2$,
$\hat{L}_4 = \overline{Y}_{12.} - \overline{Y}_{13.} = 5.2$,
$\hat{L}_5 = \overline{Y}_{13.} - \overline{Y}_{14.} = 15.2$,
$\hat{L}_6 = \overline{Y}_{13.} - \overline{Y}_{14.} = 10.0$,
$s\{\hat{L}_i\} = \sqrt{MSE\left(\frac{1}{n} + \frac{1}{n}\right)} = \sqrt{23.6\left(\frac{1}{5} + \frac{1}{5}\right)} \approx 3.0725$
($n = 5$ is the number of operators assigned to each machine/operator)
$B = t(1 - \frac{\alpha}{2g}; n_T - r) = t(1 - \frac{0.05}{2(6)}; 48) = t(0.99583; 48) \approx ?$
($g = 6$ is the number of pairwise comparisons)
and so the CIs are
$-6.0 \pm 3.0725(?) = -14.46 \leq L_1 \leq 2.46$
$-0.8 \pm 3.0725(?) = -9.26 \leq L_2 \leq 7.66$
$9.2 \pm 3.0725(?) = ? \leq L_3 \leq 17.66$
$5.2 \pm 3.0725(?) = -3.26 \leq L_4 \leq 13.66$
$15.2 \pm 3.0725(?) = 6.74 \leq L_5 \leq 23.66$
$10.0 \pm 3.0725(?) = 1.54 \leq L_6 \leq 18.46$

(c) *Contrast*
since $L = \frac{\mu_{11}+\mu_{12}+\mu_{13}}{3} - \mu_{14}$
$\hat{L} = \frac{\overline{Y}_{11\cdot}+\overline{Y}_{12\cdot}+\overline{Y}_{13\cdot}}{3} - \overline{Y}_{14\cdot} = 11.467$

$$\begin{aligned} s\{\hat{L}\} &= \sqrt{MSE\left(\frac{c_1^2}{n}+\frac{c_2^2}{n}+\frac{c_3^2}{n}+\frac{c_4^2}{n}\right)} \\ &= \sqrt{23.6\left(\frac{(1/3)^2}{5}+\frac{(1/3)^2}{5}+\frac{(1/3)^2}{5}+\frac{(-1)^2}{5}\right)} \\ &= \ ? \end{aligned}$$

$B = t(1-\alpha/2; n_T - r) = t(1-0.01/2; 48) = t(0.995; 48) \approx 2.682$
and so the CI is
$11.467 \pm 2.682(?) = 4.74 \leq L \leq 18.20$

(28.7) Bottling plant production,
mixed (A fixed, B random) model: qz5-28-7-bottle-nested-mixed

Source	*df*	*SS*	*MS*
Machines, A	2	1,695.63	847.817
Operators, within Machines, B(A)	9	2,272.30	252.478
Error	48	1,132.80	23.600
Total	59	5,100.73	

(a) *Mixed model*
Since the four operators are assigned to each machine at random,

$$\beta_{j(i)} \sim N(0, \sigma_\beta^2)$$

and $\beta_{j(i)}$ are independent of $\varepsilon_{k(ij)}$

(b) *Variance, σ_β^2*
For this mixed model, since $E(MSB(A)) = \sigma^2 + n\sigma_\beta^2$
and $E(MSE) = \sigma^2$
then $\sigma_\beta^2 = \frac{E(MSB(A)) - E(MSE)}{n} = \frac{\sigma^2 + n\sigma_\beta^2 - \sigma^2}{n} = \sigma_\beta^2$
and so it makes sense that
$s_\beta^2 = \frac{MSB(A) - MSE}{n} \approx \frac{?\ - 23.6}{5} = ?$
(where $n = 5$, number of shifts for each operator/machine)

(c) *Test Operator Effects*
$H_0 : \sigma_\beta^2 = 0$ versus
$H_a : \sigma_\beta^2 > 0$.
since p–value $P\left(F > \frac{252.478}{23.6} = ?; 9, 48\right) = 0 < \alpha = 0.10$
reject null; that is, operator effects are significant
(Notice, this mixed model test is the same as the fixed model test.)

(d) *90% confidence interval for σ_β^2, using MLS method.*
Since

$$\begin{aligned} s_\beta^2 &= 45.774 \\ c_1 &= 1/n = 1/5 = 0.2 \\ c_2 &= -1/n = -1/5 = -0.2 \\ df_1 &= 9 \\ df_2 &= 36 \\ MS_1 &= MSB(A) = 252.478 \\ MS_2 &= MSE = 23.6 \end{aligned}$$

and

$$\begin{aligned} F_1(1-\alpha/2; df_1, \infty) &= F(1-0.1/2; 9, \infty) = 1.88 \\ F_2(1-\alpha/2; df_2, \infty) &= F(1-0.1/2; 48, \infty) = 1.36 \\ F_3(1-\alpha/2; \infty, df_1) &= F(1-0.1/2; \infty, 9) = 2.71 \\ F_4(1-\alpha/2; \infty, df_2) &= F(1-0.1/2; \infty, 48) = 1.45 \\ F_5(1-\alpha/2; df_1, df_1) &= F(1-0.1/2; 9, 48) = 2.08 \\ F_6(1-\alpha/2; df_2, df_1) &= F(1-0.1/2; 48, 9) = 2.81 \end{aligned}$$

and

$$\begin{aligned} G_1 &= 1 - 1/F_1 = 1 - 1/1.88 = 0.4681 \\ G_2 &= 1 - 1/F_2 = 1 - 1/1.36 = 0.2647 \\ G_3 &= \frac{(F_5-1)^2 - (G_1F_5)^2 - (F_4-1)^2}{F_5} \\ &= 0.00765 \\ G_4 &= F_6\left(\left(\frac{F_6-1}{F_6}\right)^2 - \left(\frac{F_3-1}{F_6}\right)^2 - G_2^2\right) \\ &= -0.07162 \end{aligned}$$

and

$$\begin{aligned} H_L &= \sqrt{(G_1c_1MS_1)^2 + ((F_4-1)c_2MS_2)^2 - G_3c_1c_2MS_1MS_2} \\ &= 23.771 \\ H_U &= \sqrt{(((F_3-1)c_1MS_1)^2 + (G_2c_2MS_2)^2 - G_4c_1c_2MS_1MS_2} \\ &= 86.258 \end{aligned}$$

and since $s_\beta^2 = 45.7756$,
$45.7756 - 23.771 = ?$ and $45.7756 + 86.258 = ?$
and so the approximate confidence interval for σ_β^2 is

$$? \le \sigma_\beta^2 \le ?$$

(e) *Test Machine Effects*
$H_0 : \alpha_i = 0$ versus
$H_a :$ at least one $\alpha_i \neq 0,\ i = 1, 2, 3.$
since p–value $P\left(F > \frac{847.817}{252.478} = 3.358; 2, 9\right) = ? < \alpha = 0.10$
reject null; that is, machine effects are significant
(mean number of cases different for different machines)
(Notice, although this mixed model test is different from fixed model test, the results are the same.)

(f) *Pairwise comparisons, Machines, Tukey*
From the SAS output,
$\overline{Y}_{1..} = 61.20$, $\overline{Y}_{2..} = 70.95$, $\overline{Y}_{3..} = 73.55$
$\hat{L}_1 = \overline{Y}_{1..} - \overline{Y}_{2..} = -9.75,$
$\hat{L}_2 = \overline{Y}_{1..} - \overline{Y}_{3..} = -12.35,$
$\hat{L}_3 = \overline{Y}_{2..} - \overline{Y}_{3..} = -2.60,$
$s\{\hat{L}_i\} = \sqrt{MSB(A)\left(\frac{1}{n} + \frac{1}{n}\right)} = \sqrt{252.47\left(\frac{1}{20} + \frac{1}{20}\right)} \approx 5.025$
(Notice $s\{\hat{L}_i\}$ is different for mixed model, than for fixed model above.)
$T = \frac{1}{\sqrt{2}}q(1-\alpha; df) = \frac{1}{\sqrt{2}}q(1-0.10; 3, 9) = \frac{1}{\sqrt{2}}q(0.90; 3, 9) = \frac{3.32}{\sqrt{2}} \approx 2.348$
and so the CIs are
$-9.75 \pm 2.348(5.025) = -21.55 \leq L_1 \leq 2.05$
$-12.35 \pm 2.348(5.025) = -21.15 \leq L_2 \leq -0.55$
$-2.60 \pm 2.348(5.025) = ?$

(g) *Levene test of equal variance?*

1. *Statement*
   H_0 : error variance constant over operators *versus*
   H_1 : not constant for operators
2. *Test*
   From SAS, the p–value is ?
   The level of significance is $\alpha = 0.01$
3. *Conclusion*
   Since the p–value is (slightly) larger than the level of significance we accept the null hypothesis that the error variance is constant.

(28.8) Bottling plant production, random model: qz5-28-8-bottle-nested-random

Source	*df*	*SS*	*MS*
Machines, A	2	1,695.63	847.817
Operators, within Machines, B(A)	9	2,272.30	252.478
Error	48	1,132.80	23.600
Total	59	5,100.73	

(a) *Random model*
Since the four operators are chosen at random from a large number of operators,

$$\beta_{j(i)} \sim N(0, \sigma_\beta^2)$$

and the three machines are chosen at random from a large number of machines,

$$\alpha_i \sim N(0, \sigma_\alpha^2)$$

and $\beta_{j(i)}$, α_i and $\varepsilon_{k(ij)}$ are independent of one another

(b) *Variance,* σ_β^2*,* σ_α^2
For this random model,
since $E(MSA) = \sigma^2 + bn\sigma_\alpha^2 + n\sigma_\beta^2$
and $E(MSB(A)) = \sigma^2 + n\sigma_\beta^2$
and $E(MSE) = \sigma^2$
then $\sigma_\alpha^2 = \frac{E(MSA)-E(MSB(A))}{bn} = \frac{\sigma^2+bn\sigma_\alpha^2+n\sigma_\beta^2-(\sigma^2+n\sigma_\beta^2)}{bn}$
and $\sigma_\beta^2 = \frac{E(MSB(A))-E(MSE)}{n} = \frac{\sigma^2+n\sigma_\beta^2-\sigma^2}{n}$
and so it makes sense that
$s_\alpha^2 = \frac{MSA-MSB(A)}{bn} \approx \frac{847.82-252.47}{4(5)} = ?$
and $s_\beta^2 = \frac{MSB(A)-MSE}{n} \approx \frac{252.47-23.6}{5} = ?$

(c) *Test Machine Effects*
$H_0 : \sigma_\alpha^2 = 0$ versus
$H_a : \sigma_\alpha^2 > 0.$
since p–value $P\left(F > \frac{847.817}{252.478} = ?; 2, 9\right) = 0.081 > \alpha = 0.05$
accept null; that is, machine effects are not significant
(Notice, this random test is the same as the mixed test, but different from the fixed test above.)

(d) *95% confidence interval for σ_β^2, using MLS method.*
Since

$$\begin{aligned} s_\beta^2 &= 45.774 \\ c_1 &= 1/n = 1/5 = 0.2 \\ c_2 &= -1/n = -1/5 = -0.2 \\ df_1 &= 9 \\ df_2 &= 48 \\ MS_1 &= MSB(A) = 252.478 \\ MS_2 &= MSE = 23.6 \end{aligned}$$

and

$$\begin{aligned} F_1(1-\alpha/2; df_1, \infty) &= F(1-0.05/2; 9, \infty) = 2.11 \\ F_2(1-\alpha/2; df_2, \infty) &= F(1-0.05/2; 48, \infty) = 1.44 \\ F_3(1-\alpha/2; \infty, df_1) &= F(1-0.05/2; \infty, 9) = 3.33 \\ F_4(1-\alpha/2; \infty, df_2) &= F(1-0.05/2; \infty, 48) = 1.56 \\ F_5(1-\alpha/2; df_1, df_1) &= F(1-0.05/2; 9, 48) = 2.39 \\ F_6(1-\alpha/2; df_2, df_1) &= F(1-0.05/2; 48, 9) = 3.48 \end{aligned}$$

and

$$\begin{aligned} G_1 &= 1 - 1/F_1 = 1 - 1/2.11 = 0.5261 \\ G_2 &= 1 - 1/F_2 = 1 - 1/1.44 = 0.3056 \\ G_3 &= \frac{(F_5-1)^2 - (G_1F_5)^2 - (F_4-1)^2}{F_5} \\ &= ? \\ G_4 &= F_6\left(\left(\frac{F_6-1}{F_6}\right)^2 - \left(\frac{F_3-1}{F_6}\right)^2 - G_2^2\right) \\ &= -0.1176 \end{aligned}$$

and

$$\begin{aligned} H_L &= \sqrt{(G_1c_1MS_1)^2 + ((F_4-1)c_2MS_2)^2 - G_3c_1c_2MS_1MS_2} \\ &= 26.766 \\ H_U &= \sqrt{(((F_3-1)c_1MS_1)^2 + (G_2c_2MS_2)^2 - G_4c_1c_2MS_1MS_2} \\ &= 117.544 \end{aligned}$$

and since $s_\beta^2 = 45.7756$,
$45.7756 - 26.766 = 19.01$ and $45.7756 + 117.544 = 163.32$
and so the approximate confidence interval for σ_β^2 is

$$19.01 \le \sigma_\beta^2 \le 163.32$$

(e) *CI overall mean*

$\overline{Y}_{...} = ?$

$s\{\overline{Y}_{...}\} = \sqrt{\frac{MSA}{abn}} = \sqrt{\frac{847.82}{3(4)(5)}} \approx 3.759$

$t(1 - \frac{\alpha}{2}; a - 1) = t(1 - \frac{0.05}{2}; 2) = t(0.975; 2) \approx 4.303$

and so the CI is

$? \pm 4.303(3.759) = 52.392 \leq \mu_{..} \leq 84.742$

(28.14) Internal control, unequal nestings: qz5-28-14-control-nested-regression

(a) *regression model*

$$Y_{ijk} = \mu_{..} + \alpha_1 X_{ij1} + \alpha_2 X_{ij2} + \beta_{1(1)} X_{ij3} + \beta_{2(1)} X_{ij4} + \beta_{1(2)} X_{ij5} + \beta_{1(3)} X_{ij6} + \varepsilon_{ijk}$$

where

$$X_{ij1} = \begin{cases} 1, & \text{if case from region 1} \\ -1, & \text{if case from region 3} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij2} = \begin{cases} 1, & \text{if case from region 2} \\ -1, & \text{if case from region 3} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij3} = \begin{cases} 1, & \text{if case from team 1 from region 1} \\ -1, & \text{if case from team 3 from region 1} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij4} = \begin{cases} 1, & \text{if case from team 2 from region 1} \\ -1, & \text{if case from team 3 from region 1} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij5} = \begin{cases} 1, & \text{if case from team 1 from region 2} \\ -1, & \text{if case from team 2 from region 2} \\ 0, & \text{otherwise.} \end{cases}$$

$$X_{ij6} = \begin{cases} 1, & \text{if case from team 1 from region 3} \\ -1, & \text{if case from team 2 from region 3} \\ 0, & \text{otherwise.} \end{cases}$$

where, notice, the values of X_{ijk} are appropriately defined to give a nested model with missing observations.

(b) *Fitted Model, Residuals*
From the SAS output,

$$\hat{Y} = 150.016 - 9.216X_1 + 5.283X_2 + 6.6X_3 + 0.5X_4 + 3.7X_5 + \ ?\ X_6$$

The residual $e_{ijk} = e_{111} = 4.20$.
The residuals versus predicted values indicates constant variance
The normal probability plot of the residuals is linear, indicating normality.

(28.15) Internal control, unequal nestings,
tests: qz5-28-15-control-nested-regress-tests

(a) *Test of Region Effect*
$H_0 : \alpha_1 = \alpha_2 = 0$ versus
$H_a :$ at least one $\alpha_i \neq 0,\ i = 1, 2.$

$$\frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} = \frac{?\ - 207.26}{9 - 7} \div \frac{207.26}{2} = 10.664$$

and so p–value is $P(F > 10.664; 2, 7) \approx 0.0075$
since p–value $= 0.0075 < \alpha = 0.025$
reject null; that is, average data for different region are different

(b) *Test of (Nested) Team Effect*
Using the SAS output,
$H_0 : \beta_{1(1)} = \beta_{2(1)} = \beta_{1(2)} = \beta_{1(3)} = 0$ versus
$H_a :$ at least one $\beta_{j(i)} \neq 0.$

$$\frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} = \frac{?\ - 207.26}{11 - 7} \div \frac{207.26}{2} = 2.33$$

and so p–value is $P(F > 2.33; 2, 11) \approx 0.143$
since p–value $= 0.143 > \alpha = 0.025$
accept null;
that is, average data for different teams within each region are the *same*

(c) *Pairwise Comparison Confidence Interval*
From SAS,
$\hat{\alpha}_1 = \overline{Y}_{1..} = 141.74$, $\hat{\alpha}_2 = \overline{Y}_{2..} = 156.53$
$\hat{L} = \overline{Y}_{1..} - \overline{Y}_{2..} = -14.5$,
$s^2\{\hat{\alpha}_1\} = ?$, $s^2\{\hat{\alpha}_2\} = 6.2446$, $s\{\hat{\alpha}_1, \hat{\alpha}_2\} = -2.6197$

$$\begin{aligned} s\{\hat{L}\} &= \sqrt{s^2\{\hat{\alpha}_1 - \hat{\alpha}_2\}} \\ &= \sqrt{s^2\{\hat{\alpha}_1\} + s^2\{\hat{\alpha}_2\} - 2s\{\hat{\alpha}_1, \hat{\alpha}_2\}} \\ &\approx 3.9357 \end{aligned}$$

and so $t(1 - \frac{\alpha}{2}; df) = t(1 - \frac{0.02}{2}; 7) = t(0.99; 7) \approx 2.998$
and so the CI is
$-14.5 \pm 2.998(3.9357) = -26.30 \leq L \leq -2.70$

(28.17) Questionnaire color, subsampling, residuals: qz5-28-17-color-subsample-residuals

(a) *Residuals*
The residual $e_{ijk} = e_{111} = ?$.
The residuals versus predicted values indicates constant variance
The normal probability plot of the residuals is *not* linear, indicating *non*normality.

(b) *Levene test of equal variance?*

1. *Statement*
   H_0 : error variance constant over lots *versus*
   H_1 : error variance not constant for lots
2. *Test*
   From SAS, the p–value is ?
   The level of significance is $\alpha = 0.01$
3. *Conclusion*
   Since the p–value is larger than the level of significance we accept the null hypothesis that the error variance is constant.

(28.18) Questionnaire color, subsampling, inference: qz5-28-18-color-subsample-inference

(a) *ANOVA table*

Source	*df*	*SS*	*MS*
Treatments (colors)	2	3.2667	1.63335
Experimental error (lots(colors))	?	369.4	30.7833
Observational Error	15	67.5	4.5
Total	29	440.1667	

(b) *Test questionnaire color*
$H_0 : \tau_i = 0$ versus
$H_a :$ at least one $\tau_i \neq 0,\ i = 1, 2, 3.$
since p–value $P(F > \frac{1.63335}{30.7833} = ?; 2, 12) = 0.948 > \alpha = 0.01$
accept null; that is, there is no color effect
(mean data same for different colors)

(c) *Test lots*
$H_0 : \sigma^2 = 0$ versus
$H_a : \sigma^2 > 0.$
since p–value $P\left(F > \frac{30.7833}{4.5} = 3.358; 12, 15\right) = ? < \alpha = 0.05$
reject null; that is, lot effects are significant
(mean data different for different lots)

(d) *95% confidence interval for mean response, blue questionnaire*
$\overline{Y}_{1\cdot\cdot} = 29.2$
$s\{\overline{Y}_{1\cdot\cdot}\} = \sqrt{\frac{MSEE}{n}} = \sqrt{\frac{30.78333}{10}} \approx 1.7545$
$t(1 - \frac{\alpha}{2}; df) = t(1 - \frac{0.05}{2}; 12) = t(0.975; 12) \approx ?$
and so the CI is
$29.2 \pm 2.179(1.7545) = 25.38 \leq \mu_{1\cdot} \leq 33.02$

(e) *Variance, σ^2, σ^2_η*
$s^2 = \frac{MSEE - MSOE}{m} \approx \frac{30.7833 - 4.5}{2} = 13.1417$
and $s^2_\eta = MSOE = ?$

(f-1) *95% confidence interval for σ^2, using MLS method.*
Since

$$\begin{aligned} s^2 &= 13.1417 \\ c_1 &= 1/n = 1/2 = 0.5 \\ c_2 &= -1/n = -1/2 = -0.5 \\ df_1 &= 12 \\ df_2 &= 15 \\ MS_1 &= MSEE = ? \\ MS_2 &= MSOE = 4.5 \end{aligned}$$

and

$$\begin{aligned} F_1(1-\alpha/2; df_1, \infty) &= F(1-0.05/2; 12, \infty) = 1.94 \\ F_2(1-\alpha/2; df_2, \infty) &= F(1-0.05/2; 15, \infty) = 1.83 \\ F_3(1-\alpha/2; \infty, df_1) &= F(1-0.05/2; \infty, 12) = 2.72 \\ F_4(1-\alpha/2; \infty, df_2) &= F(1-0.05/2; \infty, 15) = 2.40 \\ F_5(1-\alpha/2; df_1, df_1) &= F(1-0.05/2; 12, 15) = 2.96 \\ F_6(1-\alpha/2; df_2, df_1) &= F(1-0.05/2; 15, 12) = 3.18 \end{aligned}$$

and

$$\begin{aligned} G_1 &= 1 - 1/F_1 = ? \\ G_2 &= 1 - 1/F_2 = 1 - 1/1.83 = 0.4536 \\ G_3 &= \frac{(F_5-1)^2 - (G_1F_5)^2 - (F_4-1)^2}{F_5} \\ &= -0.05916 \\ G_4 &= F_6\left(\left(\frac{F_6-1}{F_6}\right)^2 - \left(\frac{F_3-1}{F_6}\right)^2 - G_2^2\right) \\ &= -0.0906 \end{aligned}$$

and

$$\begin{aligned} H_L &= \sqrt{(G_1c_1MS_1)^2 + ((F_4-1)c_2MS_2)^2 - G_3c_1c_2MS_1MS_2} \\ &= 7.968 \\ H_U &= \sqrt{(((F_3-1)c_1MS_1)^2 + (G_2c_2MS_2)^2 - G_4c_1c_2MS_1MS_2} \\ &= 26.434 \end{aligned}$$

and since $s_\beta^2 = 13.1417$,
$13.1417 - 7.968 = 5.174$ and $13.1417 + 26.434 = 39.576$
and so the approximate confidence interval for σ_β^2 is

$$5.174 \le \sigma_\beta^2 \le 39.576$$

(f-2) *95% confidence interval for* σ_η^2.

Since

$$\frac{df\, MSOE}{\chi^2(1-\alpha/2; df)} = \frac{15(4.5)}{\chi^2(0.975; 15)} = \frac{15(4.5)}{27.49} = 2.455$$

and

$$\frac{df\, MSOE}{\chi^2(\alpha/2; df)} = \frac{15(4.5)}{\chi^2(0.025; 15)} = \frac{15(4.5)}{6.26} = 10.783$$