Quiz Practice Questions 3 (Attendance 6) for Statistics 514
Design of Experiments
Chapter 24 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
24, pages 1002–1008	24.7, 24.8	sodium content, ANOVA II
	24.15	miles per gallon, ANOVA II
	24.17	imitation pearls, ANOVA III

(24.7) Sodium Content: qz3-24-7-sodium-ANOVAII

(a) *Test sodium content*
From SAS,

Source	*df*	*SS*	*MS*
Treatment (Brands, random)	5	854.529166	170.9058333
Error (random)	42	30.0700000	0.7159524
Total	47	884.5991667	

Treatment (Brands)
$H_0 : \sigma_\mu^2 = 0$ versus $H_a : \sigma_\mu^2 > 0$.
since p–value $= P(F > 170.9058333/0.7159524) \approx 0.0001 < \alpha = 0.01$
reject null; that is, treatment (brands) *is* significant

Notice that a (regular, fixed) PROC ANOVA, as opposed to a "random" PROC ANOVA, is used in the SAS program. In fact, notice that only the statement of the test changes and involves σ_μ^2, not μ_i, $i = 1, \ldots, 6$, even though we are interested in the *mean* sodium content is the same in all brands sold. This is because in this random effects model, we have assumed $\mu_i \sim N(\mu, \sigma_\mu^2)$ and if $\sigma_\mu^2 = 0$, this implies μ_i, $i = 1, \ldots, 6$, are all the same.

(b) *Confidence interval*
From SAS,
$\bar{Y}_{..} = 17.62917$
$s\{\bar{Y}_{..}\} = \sqrt{\frac{MSTR}{rn}} = \sqrt{\frac{?}{(6)(8)}} \approx ?$
$t(1 - \alpha/2; r - 1) = t(0.995; 6 - 1) = ?$
and so
$\bar{Y}_{..} \pm t(1 - \alpha/2; r - 1)s\{\bar{Y}_{..}\} = 17.62917 \pm (?)(?) = (10.021, 25.237)$

(24.8) Sodium Content: qz3-24-8-sodium-ANOVAII,mls

(a) *Confidence interval of* $\frac{\sigma_\mu^2}{\sigma_\mu^2+\sigma^2}$.

lower limit, L^*

$$\begin{aligned}
F(1-\alpha/2; r-1, r(n-1)) &= F(1-0.05/2; 6-1, 6(8-1)) \\
&= F(0.995; 5, 42) \\
&= 3.95 \\
L &= \frac{1}{n}\left(\frac{MSTR}{MSE}\left(\frac{1}{F(1-\alpha/2; r-1, r(n-1))}\right)-1\right) \\
&= \frac{1}{8}\left(\frac{170.91}{0.7159524}\left(\frac{1}{3.95}\right)-1\right) \\
&= ? \\
L^* &= \frac{L}{1+L} \\
&= \frac{?}{1+?} \\
&= ?
\end{aligned}$$

upper limit, U^*

$$\begin{aligned}
F(\alpha/2; r-1, r(n-1)) &= F(0.05/2; 6-1, 6(8-1)) \\
&= F(0.005; 5, 42) \\
&= 0.0795 \\
U &= \frac{1}{n}\left(\frac{MSTR}{MSE}\left(\frac{1}{F(\alpha/2; r-1, r(n-1))}\right)-1\right) \\
&= \frac{1}{5}\left(\frac{170.91}{0.7159524}\left(\frac{1}{0.0795}\right)-1\right) \\
&= 375.21 \\
U^* &= \frac{U}{1+U} \\
&= \frac{375.21}{1+375.21} \\
&= 0.9973
\end{aligned}$$

and so putting them together,

$$? \leq \frac{\sigma_\mu^2}{\sigma_\mu^2+\sigma^2} \leq 0.9973$$

In other words, the variability of the mean sodium content of the different brands accounts for somewhere between ?% and 99% of the total variability in the sodium content.

(b) *Point estimates of σ^2 and σ^2_μ.*
$\sigma^2 \approx MSE = 0.71595$
$\sigma^2_\mu \approx s^2_\mu = \frac{MSTR - MSE}{n} = \frac{170.9058 - ?}{8} = ?$

(c) *Confidence interval for σ^2.*
since
$\chi^2(1-\alpha/2; r(n-1)) = \chi^2(1-0.01/2; 6(8-1)) = \chi^2(0.995; 42) = ?$

$$\frac{r(n-1)MSE}{\chi^2(1-\alpha/2; r(n-1))} = \frac{42(0.7159524)}{?} = ?$$

and $\chi^2(\alpha/2; r(n-1)) = \chi^2(0.01/2; 6(8-1)) = \chi^2(0.005; 42) = 22.138$

$$\frac{r(n-1)MSE}{\chi^2(\alpha/2; r(n-1))} = \frac{42(0.7159524)}{22.138} = 1.3583$$

then

$$? \leq \sigma^2 \leq 1.3583$$

(d) *Test $\sigma^2_\mu = 2\sigma^2$.*
$H_0 : \sigma^2_\mu \leq 2\sigma^2$ versus $H_a : \sigma^2_\mu > 2\sigma^2$.
since $F^* = \frac{MSTR}{n\sigma^2_\mu + \sigma^2} \div \frac{MSE}{\sigma^2} = \frac{170.905833}{(8)(2)+1} \div \frac{?}{1} = ?$
p–value $= P(F > 14.042; 5, 42) \approx 0 < \alpha = 0.01$
reject null; that is, the variability in the mean sodium content of the different brands is more than twice the variability within brands

(e) *Confidence interval for σ_μ^2, using MLS method.*
Since

$$\begin{aligned}
c_1 &= 1/n = 1/8 = 0.125 \\
c_2 &= -1/n = -1/8 = -0.125 \\
df_1 &= r-1 = 6-1 = 5 \\
df_2 &= r(n-1) = 6(8-1) = 42 \\
MS_1 &= MSTR = 170.9058 \\
MS_2 &= MSE = 0.7159524
\end{aligned}$$

and

$$\begin{aligned}
F_1(1-\alpha/2; df_1, \infty) &= F(1-0.01/2; 5, \infty) = 3.35 \\
F_2(1-\alpha/2; df_2, \infty) &= F(1-0.01/2; 42, \infty) = 1.66 \\
F_3(1-\alpha/2; \infty, df_1) &= F(1-0.01/2; \infty, 5) = 12.1 \\
F_4(1-\alpha/2; \infty, df_2) &= F(1-0.01/2; \infty, 42) = 1.91 \\
F_5(1-\alpha/2; df_1, df_1) &= F(1-0.01/2; 5, 42) = 3.95 \\
F_6(1-\alpha/2; df_2, df_1) &= F(1-0.01/2; 42, 5) = 12.51
\end{aligned}$$

and

$$\begin{aligned}
G_1 &= 1-1/F_1 = 1-1/3.35 = 0.7015 \\
G_2 &= 1-1/F_2 = 1-1/1.66 = 0.3976 \\
G_3 &= \frac{(F_5-1)^2-(G_1F_5)^2-(F_4-1)^2}{F_5} \\
&= \frac{(3.95-1)^2-((0.7015)(3.95))^2-(1.91-1)^2}{3.95} \\
&= 0.0497 \\
G_4 &= F_6\left(\left(\frac{F_6-1}{F_6}\right)^2-\left(\frac{F_3-1}{F_6}\right)^2-G_2^2\right) \\
&= 12.51\left(\left(\frac{12.51-1}{12.51}\right)^2-\left(\frac{12.1-1}{12.51}\right)^2-0.3976^2\right) \\
&= -1.2371
\end{aligned}$$

and

$$\begin{aligned}
H_L &= \sqrt{(G_1c_1MS_1)^2+((F_4-1)c_2MS_2)^2-G_3c_1c_2MS_1MS_2} \\
&= \sqrt{((0.7015)(0.125)(170.9058))^2+((1.91-1)(-0.125)(0.7159524))^2-(0.0497)(0.125)(-0.125)(170.9058)(0.7159524)} \\
&= 14.990 \\
H_U &= \sqrt{(((F_3-1)c_1MS_1)^2+(G_2c_2MS_2)^2-G_4c_1c_2MS_1MS_2} \\
&= ?
\end{aligned}$$

and since $s_\mu^2 = 21.27374$,
$21.2737 - 14.990 = 6.284$ and $21.2737 +\ ? = ?$
and so the approximate confidence interval for σ_μ^2 is

$$6.284 \leq \sigma^2 \leq ?$$

(24.15) Miles Per Gallon: qz3-24-15-miles-ANOVAII,satter

(a) *Test Interactions*
Using the SAS output,

Source	*df*	*SS*	*MS*
Factor A (Drivers, random)	3	280.28475	93.42825
Factor B (Cars, random)	4	94.7135	23.678375
Interaction AB (random)	12	2.4465	0.203875
Error	20	3.515	0.17575
Total	39	380.95975	

Interaction
$H_0 : \sigma^2_{\alpha\beta} = 0$ versus $H_a : \sigma^2_{\alpha\beta} > 0$.
since p–value $= P(F > 0.203875/0.17575) \approx\ ? > \alpha = 0.05$
accept null; that is, interaction is *not* significant

(b) *Tests of Main Effect*
Factor A (Drivers)
$H_0 : \sigma^2_{\alpha} = 0$ versus $H_a : \sigma^2_{\alpha} > 0$.
since p–value $= P(F > 93.42825/0.17575) \approx 0.0001 < \alpha = 0.05$
reject null; that is, Factor A (Drivers) *is* significant

Factor B (Cars)
$H_0 : \sigma^2_{\beta} = 0$ versus $H_a : \sigma^2_{\beta} > 0$.
since p–value $= P(F > 23.678375/0.17575) \approx 0.0001 < \alpha = 0.05$
reject null; that is, Factor B (Cars) *is* significant

(c) *Point estimates of σ^2_α and σ^2_β.*
$\sigma^2_\alpha \approx s^2_\alpha = \frac{MSA - MSE}{nb} = \frac{?\ - 0.17575}{2(5)} = ?$
$\sigma^2_\beta \approx s^2_\beta = \frac{MSB - MSE}{na} = \frac{23.678375 - 0.17575}{2(4)} = 2.938$
The Drivers effect has a greater effect on gasoline consumption because $\sigma^2_\alpha > \sigma^2_\beta$.

(d) *Confidence interval for σ_α^2, using MLS method.*
Since

$$\begin{aligned} c_1 &= 1/bn \ = 1/10 \ = \ 0.1 \\ c_2 &= -1/bn \ = -1/10 \ = \ -0.1 \\ df_1 &= a-1 \ = 4-1 \ = \ 3 \\ df_2 &= (a-1)(b-1) \ = \ 3(4) = \ 12 \\ MS_1 &= MSA \ = \ 93.42825 \\ MS_2 &= MSE \ = \ 0.203875 \end{aligned}$$

and

$$\begin{aligned} F_1(1-\alpha/2; df_1, \infty) &= F(1-0.05/2; 3, \infty) \ = ? \\ F_2(1-\alpha/2; df_2, \infty) &= F(1-0.05/2; 12, \infty) \ = \ 1.94 \\ F_3(1-\alpha/2; \infty, df_1) &= F(1-0.05/2; \infty, 3) \ = \ 13.9 \\ F_4(1-\alpha/2; \infty, df_2) &= F(1-0.05/2; \infty, 12) \ = \ 2.72 \\ F_5(1-\alpha/2; df_1, df_1) &= F(1-0.05/2; 3, 12) \ = \ 4.47 \\ F_6(1-\alpha/2; df_2, df_1) &= F(1-0.05/2; 12, 3) \ = \ 14.3 \end{aligned}$$

and

$$\begin{aligned} G_1 &= 1-1/F_1 \ = \ 1-1/? \ = ? \\ G_2 &= 1-1/F_2 \ = \ 1-1/1.94 \ = \ 0.4845 \\ G_3 &= \frac{(F_5-1)^2-(G_1F_5)^2-(F_4-1)^2}{F_5} \\ &= -0.0320 \\ G_4 &= F_6\left(\left(\frac{F_6-1}{F_6}\right)^2-\left(\frac{F_3-1}{F_6}\right)^2-G_2^2\right) \\ &= -2.6241 \end{aligned}$$

and

$$\begin{aligned} H_L &= \sqrt{(G_1c_1MS_1)^2+((F_4-1)c_2MS_2)^2-G_3c_1c_2MS_1MS_2} \\ &= 6.348 \\ H_U &= \sqrt{(((F_3-1)c_1MS_1)^2+(G_2c_2MS_2)^2-G_4c_1c_2MS_1MS_2} \\ &= 120.525 \end{aligned}$$

and since $s_\alpha^2 = 9.3224$,
$9.3224 - 6.348 = 2.974$ and $9.3224 + 120.525 = 129.847$
and so the approximate confidence interval for σ_α^2 is

$$2.974 \leq \sigma_\alpha^2 \leq 129.847$$

(e) *Confidence interval for σ_β^2, using Satterthwaite method.*
Since

$$\begin{aligned}
df &= \frac{(an \times s_\beta^2)^2}{\frac{(MSB)^2}{b-1} + \frac{(MSE)^2}{(a-1)(b-1)}} \\
&= \frac{((8)(2.9343))^2}{\frac{(23.678375)^2}{5-1} + \frac{(0.203875)^2}{(4-1)(5-1)}} \\
&= \ ? \\
\chi^2(1-\alpha/2; df) &= \chi^2(1-0.05/2; 4) \ = \ 0.484 \\
\chi^2(\alpha/2; df) &= \chi^2(0.05/2; 4) \ = \ 11.143
\end{aligned}$$

and so the approximate confidence interval for σ_β^2 is

$$\frac{(?)(2.9343)}{11.143} \le \sigma_\beta^2 \le \frac{(?)(2.9343)}{0.484}$$

or

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

(24.17) Imitation Pearls: qz3-24-17-pearls-ANOVAIII

(a) Using the SAS output,

Source	*df*	*SS*	*MS*
Factor A (Coats, fixed)	2	150.3879	75.1940
Factor B (Batch, random)	3	152.8519	70.9506
Interaction AB (random)	6	1.8521	0.3087
Error	36	173.6250	4.8229
Total	47	478.7167	

(Random) Interaction
$H_0 : \sigma^2_{\alpha\beta} = 0$ versus $H_a :\ \sigma^2_{\alpha\beta} > 0$.
$F = \frac{0.3087}{4.8229} = 0.06$ with (6,36) degrees of freedom
since p–value $= P(F > 0.06) \approx 0.99 > \alpha = 0.05$
?

(b) *Tests of Main Effect*

(Fixed) Factor A (Coats)
$F = \frac{75.1940}{0.3087} = ?$ with (2,36) degrees of freedom
(Notice: $F \neq \frac{75.1940}{4.8229} = ?$ because this is a mixed model!)
$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$ versus $H_a :$ at least one $\alpha_i \neq 0$, $i = 1, 2, 3$.
since p–value $= P(F > ?) \approx 0 < \alpha = 0.05$
reject null; that is, Factor A (Coats) *is* significant

(Random) Factor B (Batch)
$F = \frac{50.9506}{0.3087} = 10.56$ with (3,6) degrees of freedom
$H_0 : \sigma^2_\beta = 0$ versus $H_a :\ \sigma^2_\beta > 0$.
since p–value $= P(F > 10.56) \approx 0 < \alpha = 0.05$
reject null; that is, Factor B (Batch) *is* significant

(c) *CIs (Bonferroni) of $D_1 = \mu_{2\cdot} - \mu_{1\cdot}$ and $D_2 = \mu_{3\cdot} - \mu_{2\cdot}$.*
$\bar{Y}_{1\cdot\cdot} = 73.10625$, $\bar{Y}_{2\cdot\cdot} = 76.79375$, $\bar{Y}_{3\cdot\cdot} = 76.92500$,
$\hat{D}_1 = \bar{Y}_{2\cdot\cdot} - \bar{Y}_{1\cdot\cdot} = 3.68750$, $\hat{D}_1 = \bar{Y}_{3\cdot\cdot} - \bar{Y}_{2\cdot\cdot} = 0.13125$,
$s\{\hat{D}_i\} = \sqrt{\frac{2MSAB}{bn}} = \sqrt{\frac{2(?)}{(4)(4)}} \approx ?$, $i = 1, 2$
$t(1-\alpha/2; (a-1)(b-1)) = t(1-0.05/2; (3-1)(4-1)) = t(0.975; 6) = 2.447$

$$\begin{aligned}
\hat{D}_1 \pm t(1-\alpha/2; (a-1)(b-1))s\{\hat{D}_1\} &= 3.6875 \pm 2.447(0.1994) \\
&= (3.2069, 4.1681) \\
\hat{D}_2 \pm t(1-\alpha/2; (a-1)(b-1))s\{\hat{D}_2\} &= 0.13125 \pm 2.447(0.1994) \\
&= (-0.3493, 0.6118)
\end{aligned}$$

(d) *Confidence interval for $\mu_{2\cdot}$, using Satterthwaite method.*
Since,

$$\begin{aligned}
\hat{\mu}_{2\cdot} &= \;? \\
c_1 &= \frac{a-1}{nab} = \frac{3-1}{(4)(3)(4)} = 2/48 \\
c_2 &= \frac{1}{nab} = \frac{1}{(4)(3)(4)} = 1/48 \\
s^2\{\hat{\mu}_{2\cdot}\} &= c_1 MSAB + c_2 MSB \\
&= (2/48)(0.30868) + (1/48)(50.95056) \\
&= 1.0743 \\
df &= \frac{(c_1 MSAB + c_2 MSB)^2}{\frac{(c_1 MSAB)^2}{(a-1)(b-1)} + \frac{(c_2 MSB)^2}{b-1}} \\
&= \frac{(1.0743)^2}{\frac{((2/48)(0.30868))^2}{(3-1)(4-1)} + \frac{((1/48)(50.95056))^2}{4-1}} \\
&= 3.07 \\
t(1-\alpha/2; df) &= t(1-0.05/2; 3) = 3.182
\end{aligned}$$

then

$$\hat{\mu}_{2\cdot} \pm t(1-\alpha/2; df)s\{\hat{\mu}_{2\cdot}\} = \;? \;\pm (3.182)\sqrt{1.0743} = (73.496, 80.092)$$

(e) *Confidence interval for σ_β^2, using MLS method.*
Since

$$\begin{aligned}
s_\beta^2 &= \frac{MSB - MSE}{na} = \frac{? - 4.8229167}{(4)(3)} = ? \\
c_1 &= 1/n = 1/12 = 0.1 \\
c_2 &= -1/n = -1/12 = -0.1 \\
df_1 &= a - 1 = 3 - 1 = 2 \\
df_2 &= 36 \\
MS_1 &= MSB = 50.950556 \\
MS_2 &= MSE = 4.8229167
\end{aligned}$$

and

$$\begin{aligned}
F_1(1-\alpha/2; df_1, \infty) &= F(1-0.1/2; 2, \infty) = 2.60 \\
F_2(1-\alpha/2; df_2, \infty) &= F(1-0.1/2; 36, \infty) = 1.42 \\
F_3(1-\alpha/2; \infty, df_1) &= F(1-0.1/2; \infty, 2) = 8.53 \\
F_4(1-\alpha/2; \infty, df_2) &= F(1-0.1/2; \infty, 36) = 1.55 \\
F_5(1-\alpha/2; df_1, df_1) &= F(1-0.1/2; 2, 36) = 2.87 \\
F_6(1-\alpha/2; df_2, df_1) &= F(1-0.1/2; 36, 2) = 8.60
\end{aligned}$$

and

$$\begin{aligned}
G_1 &= 1 - 1/F_1 = 1 - 1/2.60 = 0.6154 \\
G_2 &= 1 - 1/F_2 = 1 - 1/1.42 = 0.2958 \\
G_3 &= \frac{(F_5-1)^2 - (G_1 F_5)^2 - (F_4-1)^2}{F_5} \\
&= 0.0261 \\
G_4 &= F_6\left(\left(\frac{F_6-1}{F_6}\right)^2 - \left(\frac{F_3-1}{F_6}\right)^2 - G_2^2\right) \\
&= -0.6286
\end{aligned}$$

and

$$\begin{aligned}
H_L &= \sqrt{(G_1 c_1 MS_1)^2 + ((F_4-1)c_2 MS_2)^2 - G_3 c_1 c_2 MS_1 MS_2} \\
&= 2.631 \\
H_U &= \sqrt{(((F_3-1)c_1 MS_1)^2 + (G_2 c_2 MS_2)^2 - G_4 c_1 c_2 MS_1 MS_2} \\
&= 31.989
\end{aligned}$$

and since $s_\beta^2 = 3.844$,
$3.844 - 2.631 = 1.213$ and $3.844 + 31.989 = 35.833$
and so the approximate confidence interval for σ_β^2 is

$$1.213 \le \sigma_\alpha^2 \le 35.833$$