STAT 368 - Assignment 4 - due date is on course outline

I assume that by now you know how to analyze residuals. So, in these questions, you need not do so.

1. 13-9 in the text. The questions ((a) and (b)) actually follow 13-10. If you get negative variance estimates, assume that the true variances must then be zero, and act accordingly. Then add part *(c): Using the estimates from your final model, what proportion of the variance of y is accounted for by the variance of the measuring device? Answer by presenting both a point estimate, and a 95% confidence interval on this proportion.*

2. 13-14 in the text. Replace *(c)* above by *(c): Using the estimates from your final model, estimate* σ^2_β *and test the hypothesis that* $\sigma^2_\beta = 0$. *Give the p-value.*

3. 14-3 in the text.

4. Repeat 14-3, assuming that the machines are randomly chosen from a large population of machines. Test for the significance of the factors, and estimate the variance components.

5. 14-19 in the text. Fit the model, do the ANOVA, and test for the significance of the relevant effects. Draw appropriate conclusions. What would you say is the best combination of time and temperature? Construct a 95% confidence interval on the difference between the two temperature effects.

6. Refer to the data at 15-11 in the text.

   (a) Analyze this by first filling in the blanks in the following two tables, and then drawing appropriate conclusions.

Source	SS	df	MS	F_0	p
Trucks|Volume	– – –	–	– – –	– – –	– – –
Volume|Trucks	– – –	–	– – –	– – –	– – –
Error	– – –	–	– – –		

	Estimate	Standard error
$\mu + \tau_1$	– – –	– – –
$\mu + \tau_2$	– – –	– – –
$\mu + \tau_3$	– – –	– – –
τ_1	– – –	
τ_2	– – –	
τ_3	– – –	
σ^2	– – –	
σ	– – –	

(b) Now do an incorrect analysis which ignores the volumes. You should find that the p-value for truck types is much higher. Explain the difference in terms of the values in this table:

Truck type	unadjusted $\bar{y}$	s.e.$(\bar{y}_{unadj})$
1	$---$	$---$
2	$---$	$---$
3	$---$	$---$

(Obtain s.e.$(\bar{y}_{unadj})$ using only the observations from that truck type.)

7. Consider a 2^2 factorial. We have discussed this in terms of the main effects A and B, and the interaction effect AB. We could equally well have discussed it in terms of the usual effects model, written here as

$$E\left[y_{ijk}\right] = \mu + A_i + B_j + (AB)_{ij},$$

with indices $i, j = 1$ indicating low settings, $i, j = 2$ indicating high settings and with the usual constraints (which are important). The correspondence between these parameterizations was seen to be

$$\begin{aligned} A &= A_2 - A_1,\ B = B_2 - B_1, \\ AB &= \frac{(AB)_{22} + (AB)_{11} - (AB)_{12} - (AB)_{21}}{2}. \end{aligned}$$

Suppose that we instead view this as a regression model, with response function

$$E\left[y|x_1, x_2\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

for variables

$$x_1 = \begin{cases} 1, & A = +, \\ -1, & A = -, \end{cases} \quad x_2 = \begin{cases} 1, & B = +, \\ -1, & B = -. \end{cases}$$

(a) By evaluating the mean response at the various levels of the factors, express A, B and AB in terms of the regression parameters. What does β_0 represent?

(b) What hypotheses about the regression parameters are tested to test for *(i)* no AB interaction? *(ii)* no main effect of B? *(iii)* no main effect of A?