STAT502 Lab #10

(*adapted from Exercise 25.15*) An automobile manufacturer wished to study the effects of differences among drivers (Factor A) and differences among cars (Factor B) on gasoline consumption. Four drivers were selected at random; also five cars of the same model with manual transmission were randomly selected from the assembly line. Each driver drove each car twice over a 40-mile test course and the miles per gallon were recorded. The data can be read into R with

```
data = read.table('mpg.txt')
mpg = data[,1]
A = as.factor(data[,2])
B = as.factor(data[,3])
```

1. Assume the following two-way random effects ANOVA model:

$$Y_{ijk} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

where $\mu_{\cdot\cdot}$ is the overall mean, and α_i, β_j, $(\alpha\beta)_{ij}$, and ϵ_{ijk} are all independent normal with variances σ^2_α, σ^2_β, $\sigma^2_{\alpha\beta}$, and σ^2, respectively. The following table summarizes the mean squares and their expectations.

Mean Sq	df	Expected MS (A random, B random)
MSA	$a-1$	$\sigma^2 + n\sigma^2_{\alpha\beta} + bn\sigma^2_\alpha$
MSB	$b-1$	$\sigma^2 + n\sigma^2_{\alpha\beta} + an\sigma^2_\beta$
$MSAB$	$(a-1)(b-1)$	$\sigma^2 + n\sigma^2_{\alpha\beta}$
MSE	$ab(n-1)$	σ^2

Note the values of a, b, and n for these data.

(a) What test statistic (ratio of mean squares) is appropriate to test $H_0 : \sigma^2_{\alpha\beta} = 0$ versus $H_a : \sigma^2_{\alpha\beta} > 0$? What is the critical value if $\alpha = .05$?
Note: this is the test for interaction in the random effects model. If $(\alpha\beta)_{ij}$ are fixed, then interaction means $(\alpha\beta_{ij})$ are not all equal; if they're random, then interaction means their variance $\sigma^2_{\alpha\beta} > 0$.

(b) Use `anova(lm(mpg~A+B+A:B))` to compute the mean squares, and carry out the test above. Is there significant evidence for interaction between drivers and cars?

(c) Repeat parts (a) and (b) for the drivers (Factor A). That is, report the appropriate mean square ratio and critical value for testing $H_0 : \sigma^2_\alpha = 0$ versus $H_a : \sigma^2_\alpha > 0$, and carry out the test from the output.
Note: R still thinks the effects are fixed, so it's using MSE for all tests...

(d) Repeat parts (a) and (b) for the cars (Factor B). That is, report the appropriate mean square ratio and critical value for testing $H_0 : \sigma^2_\beta = 0$ versus $H_a : \sigma^2_\beta > 0$, and carry out the test from the output.

(e) Obtain unbiased point estimates of the variance components σ^2_α, σ^2_β, $\sigma^2_{\alpha\beta}$, and σ^2. *Hint: start with the MSE as an estimate for σ^2, and use the expected mean square formulas to work backward*

2. Consider the same mpg data set above, but suppose the five cars involved were the only ones of interest (not randomly selected from a larger population). The two-way "mixed" model describes these effects:

$$Y_{ijk} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

With Factor A random and Factor B fixed, the assumptions are then as follows:

- $\mu_{\cdot\cdot}$ is the overall mean (constant)
- α_i are independent $N(0, \sigma^2_\alpha)$
- β_j are constants such that $\sum_j \beta_j = 0$
- $(\alpha\beta)_{ij}$ are $N\left(0, \frac{b-1}{b}\sigma^2_{\alpha\beta}\right)$ such that $\sum_j (\alpha\beta)_{ij} = 0$ for all i.
- ϵ_{ijk} are independent $N(0, \sigma^2)$

The reason for expressing the variance of $(\alpha\beta)_{ij}$ like this is so the expected mean squares will have a simpler form (below). Note that the mean squares are computed the same way for both random and fixed effects, but their *expected values* differ.

Mean Sq	df	Expected MS (A random, B fixed)
MSA	$a-1$	$\sigma^2 + bn\sigma^2_\alpha$
MSB	$b-1$	$\sigma^2 + n\sigma^2_{\alpha\beta} + \frac{an}{b-1}\sum_j \beta_j^2$
$MSAB$	$(a-1)(b-1)$	$\sigma^2 + n\sigma^2_{\alpha\beta}$
MSE	$ab(n-1)$	σ^2

(a) Use the expected mean squares to determine appropriate test statistics for Factor A, Factor B, and their interaction. Compare with the test statistics from the model in Problem 1, where both factors were random.

(b) For any test statistics that are different, carry out the test with $\alpha = .05$.

(c) Use the expected mean squares to suggest appropriate estimates for σ^2_α, $\sigma^2_{\alpha\beta}$, and σ^2. Find their numeric values from the data.

(d) We may also be interested in pairwise comparisons for the fixed factor effects β_j. The estimator $\hat{D} = \overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdot j'\cdot}$ is unbiased for $D = \beta_j - \beta_{j'}$ and has variance

$$Var(\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdot j'\cdot}) = \frac{2(\sigma^2 + n\sigma^2_{\alpha\beta})}{an} = \frac{2E(MSAB)}{an}$$

So, the standard error of $\hat{\beta}_j$ is $s\{\hat{D}\} = \sqrt{2MSAB/an}$. *Compare with the standard error formula for fixed effects in Ch.19.* Finally, use the result that

$$\frac{\hat{D} - D}{s\{\hat{D}\}} \sim t_{(a-1)(b-1)}$$

to calculate a 95% confidence interval for the difference in mean mpg between the first and second cars.