MTH 3220 Lab 6

Due Thu., Oct. 17

1 Part A: One-Factor ANOVA With a Random Effects

1.1 Railroad Track Waves Data Set

The ***fixed effects* one-factor ANOVA model** is

$$Y_{ij} \;=\; \mu + \alpha_i + \epsilon_{ij} \,,$$

where the **overall mean μ** and **treatment effects $\alpha_1, \alpha_2, \ldots, \alpha_I$** are (unknown) **constants** and the ϵ_{ij}'s are **$N(0,\, \sigma)$ random errors**.

The ***random effects* model** is

$$Y_{ij} \;=\; \mu + A_i + \epsilon_{ij} \,,$$

where **μ** is again a **constant**, but now the **treatment effects $A_1, A_2, \ldots, A_I$** are **$N(0,\, \sigma_A)$ random variables** and the ϵ_{ij}'s are again **$N(0,\, \sigma)$ random errors**.

The data below (from **Example 10.11** of the textbook) are travel times of a certain type of wave that results from longitudinal stress of rails used for railroad tracks. **Three** measurements were made on each of **six rails** selected ***randomly*** from the population of rails.

Rail1	Rail2	Rail3	Rail4	Rail5	Rail6
55	26	78	92	49	80
53	37	91	100	51	85
54	32	85	96	50	83

The data are also in the file **railroad_waves.txt**.

We want to know if some variation in travel time could be attributed to "between-rail variability."

1. Use `read.table()` to read the data from **railroad_waves.txt** into a data frame named, say, `my.data`, for example:

```
# Select the file:
my.file <- file.choose()
# Look at the path name for the selected file:
my.file
# Read the data from the file:
my.data <- read.table(my.file, header = TRUE)
```

2. To fit a ***random effects model***, we need the `"lme4"` package (add-on set of R functions). **Install** the package, by typing:

```r
install.packages()
```

Load the package into the current R session, by typing:

```r
library(lme4)
```

3. The `lmer()` function from the `"lme4"` package will fit a **random effects model**. Fit the model by typing:

```r
my.anova <- lmer(TravelTime ~ (1 | Rail), data = my.data)
```

Then look at the results by typing:

```r
summary(my.anova)
```

2 Part B: Two-Factor ANOVA With One Observation Per Group ($K = 1$)

2.1 Soil Nitrogen Data Set

An experiment was carried out to determine whether different **operators** obtained different mean results in routine soil analysis for **nitrogen**. On each of three **days** a sample of soil was selected and then divided into five parts. At random these five parts were assigned to the operators to analyze. The results are below and also in the file **nitrogen.txt**.

		Operator				
		A	B	C	D	E
Day	Tuesday	509	512	532	506	509
	Wednesday	505	507	542	520	519
	Friday	465	472	498	483	475

1. Use `read.table()` to read the data from **nitrogen.txt** into a data frame named, say, `my.data` (see **Step 1** of **Part A**).

2. Use `aov()` and `summary()`, as follows, to fit a two-factor *additive effects* ANOVA model to the data, and then look at the ANOVA table:

```r
my.anova <- aov(Nitrogen ~ Day + Operator, data = my.data)
summary(my.anova)
```

3. Now check the normality assumption, for example using a **histogram** and **normal probability plot** of the **residuals**:

```r
hist(my.anova$residuals, col = "blue")
qqnorm(my.anova$residuals, pch = 19)
qqline(my.anova$residuals, col = "blue")
```

4. Carry out **Tukey's multiple comparisons procedure** to decide *which* **days** differ from each other and *which* **operators** differ from each other:

```
TukeyHSD(my.anova)
```