STAT 461: Lab 10 - Block Designs

A randomized complete block design (RCBD) usually has one treatment of each factor level applied to an EU in each block. It can be applied more than once, but it is typically just applied once.

In this example, you wish to compare the wear level of four different types of tires. Tread loss is measured in tread in mils (.001 inches). First, we will treat the design as if it wasn't blocked. That is, four versions of each type of tire is randomly placed on a car. This is one factor ANOVA.

```
Wear <- c(17, 14,12,13,14,14,12,11,13,13,10,11,13,8,9,9)
Brand <- rep(c("A", "B", "C", "D"), 4)
df1 <- data.frame(Wear, Brand)
with(df1, tapply(Wear, Brand, mean))
```

```
##     A     B     C     D
## 14.25 12.25 10.75 11.00
```

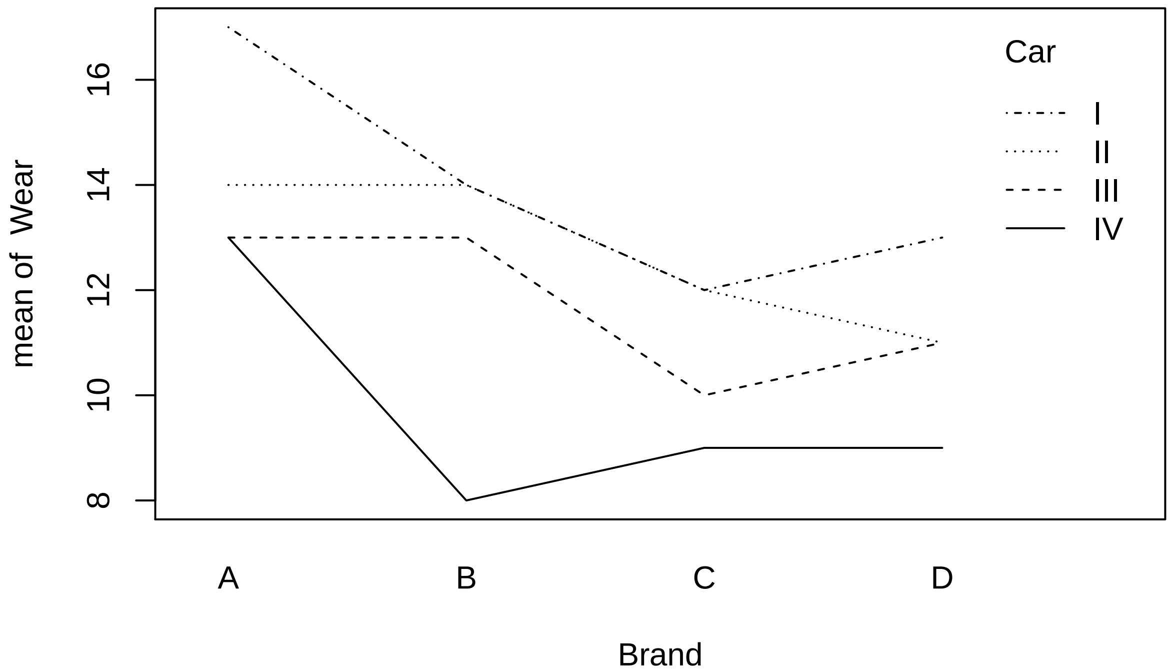
```
res1 <- lm(Wear ~ Brand, data = df1)
anova(res1)
```

```
## Analysis of Variance Table
##
## Response: Wear
##           Df Sum Sq Mean Sq F value Pr(>F)
## Brand      3 30.688 10.2292  2.4428 0.1145
## Residuals 12 50.250  4.1875
```

Based on this analysis, there is no significant difference among the tires in terms of wear (F=2.44, p=0.115). If four different cars are used, the difference in wear among the cars may be masking the differences in wear among the tires.

To make is a RCBD, randomly place 1 tire of each brand on each car. Therefore, car becomes the block. Interaction plots are useful to check (visually) for interactions. A RCBD assumes that there is no interaction.

```
Car <- c(rep("I", 4), rep("II", 4), rep("III", 4), rep("IV", 4))
df2 <- data.frame(Wear, Brand, Car)
with(df2, interaction.plot(Brand, Car, Wear))
```




```
with(df2, tapply(Wear, Brand, mean))
```

```
##     A     B     C     D
## 14.25 12.25 10.75 11.00
```

```
with(df2, tapply(Wear, Car, mean))
```

```
##     I    II   III    IV
## 14.00 12.75 11.75  9.75
```

```
res2 <- lm(Wear ~ Car + Brand, data = df2)
anova(res2)
```

```
## Analysis of Variance Table
##
## Response: Wear
##           Df Sum Sq Mean Sq F value   Pr(>F)
## Car        3 38.688 12.8958 10.0378 0.003133 **
## Brand      3 30.688 10.2292  7.9622 0.006685 **
## Residuals  9 11.563  1.2847
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This result shows that there is significant car/block effect (F=10.04, p=0.003). In many cases, we are not interested in the differences between cars, but we report it anyway. Once car/block is accounted for, there is a significant effect of tire brand (F=7.96, p=0.007). You can now run TukeyHSD to see which ones are different.

```
TukeyHSD(aov(res2),"Brand")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = res2)
##
## $Brand
```

```
##      diff       lwr        upr     p adj
## B-A -2.00 -4.502042  0.5020416 0.1274166
## C-A -3.50 -6.002042 -0.9979584 0.0080251
## D-A -3.25 -5.752042 -0.7479584 0.0125310
## C-B -1.50 -4.002042  1.0020416 0.3041129
## D-B -1.25 -3.752042  1.2520416 0.4451407
## D-C  0.25 -2.252042  2.7520416 0.9887731
```

Homework Assignment

1. (15 points) **Fat in diets.** A researcher studied the effects of three experimental diets with varying fat contents on the total lipid (fat) level in plasma. Total lipid level is widely used predictor of coronary heart disease. Fifteen male subjects who were within 20% of their ideal body weight were grouped into five blocks according to age. Within each block, the three experimental diets were randomly assigned to the three subjects. Data on reduction in lipid level (in grams per liter) after the subjects were on the diet for a fixed period of time follow.

Block i		**Fat Content of Diet**		
		j=1 **Extremely low**	**j=2** **Fairly low**	**j=3** **Moderately Low**
1	Ages 15-24	0.73	0.67	0.15
2	Ages 25-34	0.86	0.75	0.21
3	Ages 35-44	0.94	0.81	0.26
4	Ages 45-54	1.40	1.32	0.75
5	Ages 55-64	1.62	1.41	0.78

   (a) Why do you think that age of subject was used as a blocking variable?

   (b) Obtain the residuals for randomized block model $Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$ and plot them against the fitted values. Also prepare a normal probability plot of the residuals. What are your findings?

   (c) Plot the response Y_{ij} by blocks (Present the lipid levels for each kind of diet by block). What does this plot suggest about the appropriateness of the no-interaction assumption here?

2. (40 points) (By hand) Refer to the Fat in diets problem. Assume that randomized block model is appropriate.

   (a) Obtain the analysis of variance table.

   (b) Test whether or not the mean reductions in lipid level differ for the three diets; use $\alpha = 0.05$. State the alternatives, decision rule, and conclusion. What is the P-value of the test?

   (c) If there is significant difference in lipid level, how do they differ?

3. (45 points) (ANCOVA) A manufacturer of felt-tip markers investigated by an experiment whether a proposed new display, featuring a picture of a physician, is more effective in drugstores than the present counter display, featuring a picture of an athlete and designed to be located in the stationary area. Fifteen drugstores of similar characteristics were chosen for the study. They were assigned at random in equal numbers to one of the following treatments: (1) present counter display in stationary area, (2) new display in stationary area, (3) new display in checkout area. Sales with the present display (X_{it}) were recorded in all 15 stores for a three week period. Then the new display was set up in the 10 stores receiving it, and sales for the next three week period (Y_{it}) were recorded in all 15 stores. The data on sales (in dollars) follow.

Table 1:

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$i=1$ first 3 wks	92	68	74	52	65
$i=1$ second 3 wks	69	44	58	38	54
$i=2$ first 3 wks	77	80	70	73	79
$i=2$ second 3 wks	74	75	73	78	82
$i=3$ first 3 wks	64	43	81	68	71
$i=3$ second 3 wks	66	49	84	75	77

The analyst wishes to analyze the effects of the three different display treatments by means of covariance analysis.

(a) Obtain the residuals for covariance model $Y_{it} = \mu + \tau_i + \gamma(X_{it} - \bar{X}_{..}) + \varepsilon_{it}$.

(b) For each treatment, plot the residuals against the fitted values. Also prepare a normal probability plot of the residuals.

(c) Assume ANOVA with equal slope models, i.e., $Y_{it} = \mu + \tau_i + \gamma(X_{it} - \bar{X}_{..}) + \varepsilon_{it}$. Test for whether the slope is significant. Conduct the test using $\alpha = 0.05$. State the alternatives, decision rule, and conclusion. What is the P-value of the test?

(d) Fit the full and reduced regression models and test for treatment effects: use $\alpha = 0.05$. State the alternatives, decision rule, and conclusion. What is the P-value of the test?