Contents of this handout:

**Section 1:** Balanced, randomized groups design with one random variable, nested in one fixed variable.

**Section 2:** Balanced, randomized groups design with **two fixed variables that are crossed**, and **one random variable nested** in one of the fixed variables.

**Section 1**
Section 1 of this tutorial uses data from Keppel & Wickens, 4th ed., Table 25.1 on page 553. The aov methods described here work **ONLY for balanced designs** (i.e. equal cell n).

In the example, a total of 30 participants each read a story, and then were asked to recall some facts from the story. The dependent variable was number of facts recalled. The primary variable of interest (factor A) was **Story Setting**: stories were either **familiar** or **exotic**. Three familiar stories were created, as well as three exotic stories (6 stories in total), and so we can say that **Story** (factor B) is nested in **Setting**. It was a completely between participants design, so each story was read by 5 participants. The raw recall scores are shown below, and are available in a file called *keppel-25-1-top.txt* on the Learn@UW site. Note that in the data file **story** is coded twice: correctly in the column labeled "story.nested", and incorrectly in the column labeled "story.crossed".

<table>
<thead>
<tr>
<th>Setting</th>
<th>Familiar</th>
<th>Exotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>story b1</td>
<td>14, 15, 12, 13, 13</td>
<td></td>
</tr>
<tr>
<td>story b2</td>
<td>13, 12, 15, 16, 12</td>
<td></td>
</tr>
<tr>
<td>story b3</td>
<td>15, 17, 15, 16, 15</td>
<td></td>
</tr>
<tr>
<td>story b4</td>
<td>7, 8, 7, 7, 8</td>
<td></td>
</tr>
<tr>
<td>story b5</td>
<td>8, 6, 11, 11, 11</td>
<td></td>
</tr>
<tr>
<td>story b6</td>
<td>13, 10, 11, 11, 10</td>
<td></td>
</tr>
</tbody>
</table>

A. Get the data into R
Make sure R's working directory is pointed at the location where you saved the data file.

```r
> d1 = read.table("Keppel-25-5.txt", header=T)
```

The dataframe should look like above. Note that rows 8-30 have been deleted to save space on this handout.

Also note that the values specifying the levels of the IVs contain letters, so R automatically figured out that they were factors (and not continuous variables) when it read them in. You can check this:

```r
> class(d1$story.nested)
[1] "factor"
```
B. Do the ANOVA

Since story is random and nested in setting, it is adding random error to our estimate of the population treatment effect due to A. We want to take this extra error variance into the numerator into account when we construct an F ratio to test the effect of A, and so we need to make sure the denominator includes it as well.

By default, R will assume that story is a fixed factor. As a consequence, it will use the wrong error term to test the main effect of setting. We need to tell R to use MS_{story/setting} as the error term to test the main effect of setting.

```
> m1.set = aov(recall~setting + Error(story.nested), data=d1)
> summary(m1.set)
```

```
Error: story
  Df  Sum Sq Mean Sq   F value  Pr(>F)
setting     1 182.533 182.533 15.425 0.01714 *
Residuals   4  47.333  11.833
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Error: Within
  Df  Sum Sq Mean Sq  F value  Pr(>F)
Residuals 24   50.0   2.0833
```

The summary shows us something different from what we've seen before. The ANOVA is split into two parts. The top part used the error term we specified to test the main effect of setting. The df, SS, and MS for that error term are in the row labeled "Residuals". Check that the df are correct for the story/setting source: \( \text{df}_{B/A} = a(b-1) \).

The bottom part shows our old familiar error term: the within cell variance. It is not testing anything, which seems a little sad, but it doesn't mind. In the book it is denoted as S/B/A.

Note that we did not get a significance test of story. That's ok, because since it's a random factor we expect it to vary somewhat. What we were really interested in was the fixed factor setting. But it's not a bad idea to see how much variance is explained by story. We can test it by running an ANOVA the usual way, using the within cell variance as the error term. That will give us a correct test of the significance of story, but also (dangerously) an incorrect test of the significance of setting.

```
> m1.story = aov(recall ~ setting * story.nested, data=d1)
> summary(m1.story, intercept=T)
```

```
Df  Sum Sq Mean Sq  F value     Pr(>F)
(Intercept) 1 4130.1  4130.1 1982.464 < 2.2e-16 ***
setting     1  182.5  182.5  87.616  1.759e-09 ***
story        4   47.3   11.8   5.680  0.002308 **
Residuals   24   50.0   2.0833
```

For some reason, if we specify an error term other than the default S/A, R will not show us the intercept. Note that in the summary of m1.story, the MS_{intercept} is correct but F_{intercept} is incorrect. The proper error term is story/setting.

If you really want to test the significance of the intercept, you can do it by hand using MS_{intercept} from this summary, and MS_{residual} from the summary at the top of this page. This summary is also useful for checking that the total df in our model add up to \( a*b*n \), where \( b \) equals the number of levels of B nested in each level of A- NOT the total levels of B.

Finally, notice that even though we used an asterisk (*) in the model specification, indicating that we wanted to test for effects of setting, story.nested, and their interaction, R did not test the interaction. Why did the program willfully disobey our orders? Because it recognized that it is impossible to test an interaction of a factor nested in another. But if we had set up our data file incorrectly, like it is in the column story.crossed, we could have gone right and ahead and gotten very reasonable looking output for a very wrong analysis. Scary stuff.
Section II

Now we make the design a little more complicated. We have two fixed effects that are crossed with each other, and a random effect that is nested in one of the fixed effects. The contrived data are taken from page 3 of Prof. Moore's class handout #23, and are posted as "HO23_part2data.txt" on the learn@UW site.

The (fictional) experiment investigated gender bias in people's concepts of professions. Participants were assigned to one of three levels of factor A (sentence: person, man, woman). Participants then responded yes/no to a sentence of the form "A_____ is a person/man/woman", where the final word was determined by factor A, and the blank was replaced by one of eight professions (factor C). Four of the example professions were stereotypically male, and four were stereotypically female, so we can say that Profession was random and nested in Gender (C/B). The dependent variable was the reaction time (rt) of how quickly participants responded, measured in tenths of a second. The experiment had a total of 72 participants.

<table>
<thead>
<tr>
<th>A: Sentence (final word)</th>
<th>person</th>
<th>man</th>
<th>woman</th>
</tr>
</thead>
<tbody>
<tr>
<td>C: profession1</td>
<td>6</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>C: profession2</td>
<td>5</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>C: profession3</td>
<td>6</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>C: profession4</td>
<td>8</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B: Gender of profession</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>C: profession5</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>C: profession6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>C: profession7</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>C: profession8</td>
<td>9</td>
<td>6</td>
</tr>
</tbody>
</table>

Raw RT scores for each participant are shown.

A. Get the data into R

```r
> d2 = read.table("HO23_part2data.txt", header=T)
> d2
  rt sentence gender profession
  1 6 isaperson   male     ex1
  2 5 isaperson   male     ex2
  3 6 isaperson   male     ex3
  4 8 isaperson   male     ex4
  5 8 isaperson   female   ex5
  6 6 isaperson   female   ex6
  7 7 isaperson   female   ex7
  8 9 isaperson   female   ex8
  9 8 isaperson   male     ex1
  66 9 isawoman   male     ex2
  67 9 isawoman   male     ex3
  68 10 isawoman  male     ex4
  69 6 isawoman   female   ex5
  70 3 isawoman   female   ex6
  71 6 isawoman   female   ex7
  72 4 isawoman   female   ex8
```

→ rows 10 through 66 not shown to save space
B. Do the ANOVA

Again, specifying the correct error term is going to be the tricky part. The table below is from page 561 of the Keppel book. Note that in Keppel’s table C is nested in A, but in our example C is nested in B. So the letters do not match perfectly, but the logic of which error term to use still holds.

**Table 25.7: Degrees of freedom and error term for a three-factor between-subjects design with crossed fixed factors A and B, and a random factor C nested in A.**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Influence on Mean Square</th>
<th>Error Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a - 1</td>
<td>A, C/A, S/BC/A</td>
<td>C/A</td>
</tr>
<tr>
<td>B</td>
<td>b - 1</td>
<td>B, BxC/A, S/BC/A</td>
<td>BxO/A</td>
</tr>
<tr>
<td>AxB</td>
<td>(a - 1)(b - 1)</td>
<td>AxB, BxC/A, S/BC/A</td>
<td>BxO/A</td>
</tr>
<tr>
<td>C/A</td>
<td>a(c - 1)</td>
<td>C/A, S/BC/A</td>
<td>S/BC/A</td>
</tr>
<tr>
<td>BxC/A</td>
<td>a(b - 1)(c - 1)</td>
<td>BxC/A, S/BC/A</td>
<td>S/BC/A</td>
</tr>
<tr>
<td>S/BC/A</td>
<td>abc(n - 1)</td>
<td>S/B C/A</td>
<td>---</td>
</tr>
</tbody>
</table>

To test the fixed effect of A (sentence) in our example, the error term should be the interaction A x C/B (sentence by profession nested in gender). That term includes the sources of random variance that we want to include in the denominator of the F ratio. To test the fixed effect of B (gender), we want to use C/B (profession nested in gender) as the error term.

The following code manually specifies the error term to use, by including the `Error( )` argument inside the `aov( )` function. It is a little confusing. Note that in the `Error( )` argument, the random variable “profession” is specified before the slash. The fixed effect without a random factor nested in it, “sentence”, is specified after the slash. This tells R something like: *Use the MS of cells in the profession/gender X sentence interaction as the fullest possible error term. But if sentence is not included in the effect being tested, don’t include it in the error term either.*

```r
> m2 = aov(rt ~ sentence * gender + Error(profession/sentence), data=d2)
> summary(m2)
```

Error: profession

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>1</td>
<td>18.0000</td>
<td>18.0000</td>
<td>9.3462</td>
</tr>
<tr>
<td>Residuals</td>
<td>6</td>
<td>11.556</td>
<td>1.9259</td>
<td>---</td>
</tr>
</tbody>
</table>

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Error: profession:sentence

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sentence</td>
<td>2</td>
<td>326.08</td>
<td>163.042</td>
<td>80.5881</td>
</tr>
<tr>
<td>sentence:gender</td>
<td>12</td>
<td>24.28</td>
<td>2.023</td>
<td>---</td>
</tr>
</tbody>
</table>

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Error: Within

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residuals</td>
<td>48</td>
<td>128</td>
<td>2.6667</td>
<td>---</td>
</tr>
</tbody>
</table>

*Gender* is tested using profession/gender as the error term. R is able to figure out that profession is nested in gender because of the way we coded profession.

*Sentence* and *sentence X gender* are tested with profession/gender X sentence as the error term.
C. Check that the design is balanced

If your design is balanced, the `replications()` function returns a vector of the number of observations at each level of the factors and interactions.

```r
> replications( rt~sentence*gender + sentence*profession, data=d2)

           sentence gender profession sentence:gender sentence:profession
sentence    _gender    profession     sentence:gender sentence:profession
  sentence     gender    profession     sentence:gender sentence:profession
24         36           9                  12                  3
```

The first value tells us that there are exactly 24 observations at each level of `sentence`, and the final value tells us that there are exactly 3 observations in each cell of the `sentence` X `profession/gender` interaction. If our design weren't balance, `replications()` would return an unwieldy list instead of a tidy vector. Another way we can check that our design is balanced is by making sure that `replications` did not return a list:

```r
> !is.list( replications(rt~sentence*gender + sentence*profession,data=d2) )
[1] TRUE
```

D. Look at estimated effects and standard errors

```r
> model.tables(m2, "means")

Tables of means
Grand mean
6.5

sentence
sentence
  isaman  isaperson  isawoman
  6.542  6.458  6.500

gender
gender
female   male
    7     6

sentence:gender
gender
sentence    female  male
  isaman    9.583  3.500
  isaperson 7.083  5.833
  isawoman  4.333  8.667

```

Note that we don't get means for each level of `profession`. R has correctly inferred that because `profession` is nested, it is probably also random, and so we aren't theoretically interested in the particular levels of profession. But if we did want to look at them, we could use our old pal `tapply()`:

```r
> tapply(d2$rt, list(d2$profession, d2$sentence), mean)

       isaman isaperson isawoman
ex1 3.333333 6.000000 9.333333
ex2 3.000000 4.666667 8.666667
ex3 3.000000 6.000000 9.000000
ex4 4.666667 6.666667 7.666667
ex5 10.333333 8.000000 4.333333
ex6 9.333333 5.666667 4.000000
ex7 10.000000 7.000000 3.666667
ex8 8.666667 7.666667 5.333333
```