1. A 2 (A) x 2 (B) crossed factorial design has 4 conditions (Figure 1). List the weights for the 3 contrasts that yield results that would closely correspond to the results obtained in a factorial ANOVA.

2. What is the null hypothesis that is evaluated by the $F$ test for the interaction?

   **Answer:** The null hypothesis is that the effect of A – or, more precisely, the simple main effect of A – is the same at all levels of B, and that the (simple main) effect of B is the same at all levels of A. Alternatively, one could say that the interaction tests the null hypothesis that the cell means are equal to the sum of the intercept and main effects.

3. An experiment used a 3(A) x 3(B) crossed-factorial, between-subjects design to examine the effects of factors A and B on a dependent variable, score. Unfortunately, a computer error resulted in the loss of data from a random set of 8 participants in one of the conditions.

   (a) Conduct an ANOVA to evaluate the main effects of A, B, and the A x B interaction. Copy the ANOVA table into your test booklet.

   **Answer:** The best strategy is to use Type III sums of squares because the data are unbalanced and the loss of data appears to be random, therefore the best strategy would be to use Type III sums of squares. If you use Sequential sums of squares, the results would depend on the order of terms in the model, and there is no obvious way of deciding whether A or B should appear first. So the results shown by the `drop1` command are best in this case. Because the A x B interaction accounts for so little variation, it turns out that Type II & III sums of squares, as well as the null hypotheses that they evaluate, are very similar.
(b) What are the null hypotheses for the two main effects?

**Answer:** Type III SS (and in this case, because the A x B interaction is so small, Type II SS) evaluate the null hypothesis that the *unweighted* marginal means are equal. (In terms of model comparisons, Type III SS are used to compare a full linear model that includes terms for the 2 main effects and the interaction and a reduced model that includes the other main effect and the 2-way interaction: the comparison evaluates the change in the goodness-of-fit (SS residuals) that is due to the main effect after controlling for all other effects.) Type I SS evaluate the null hypothesis that the *weighted* marginal means are equal. Because the weighted marginal means take into account the random loss of data, I think they not very useful in this case.

(c) What does the p value for the F test of the main effect of A mean?

**Answer:** The p value is the probability of seeing variation among unweighted marginal means that is at least as large as the observed variation *when the null hypothesis of no difference is true*. Alternative but equally good answer: it is the probability of obtaining an F value at least as large as the observed value when the null hypothesis is true.
4. An experiment was conducted to assess the effects of therapist and type of treatment – rational-emotive therapy (RET), client-centered therapy (CCT), and behaviour modification (BMOD) – on client satisfaction. An ANOVA was conducted to evaluate the main effects of therapist and therapy and their interaction assuming the levels of both factors are fixed. The results are shown in Table 1.

(a) Analyze the main effects of therapist and therapy and their interaction assuming the levels of both factors are random. Copy the values of $F$ and $p$ for each test into your answer booklet.

Answer: The key here is to recognize that the experiment used a crossed, factorial design with two random factors. Therefore, we need to recalculate the $F$ and $p$ values for both main effects by comparing them to the MS value for the interaction. See Section 10.3.2 in my course notes (and the 12th slide in the lecture notes on that topic).

```r
# therapist=random; treatment=random:
# get values from original ANOVA:
MS.therapist <- 21.7;
MS.treatment <- 60;
MS.interaction <- 16.7;
MS.residuals <- 8.8;
df.therapist <- 2;
df.treatment <- 2;
df.interaction <- 4;
df.residuals <- 36;
# recompute F & p for both main effects:
(F.treatment <- MS.treatment/MS.interaction)
## [1] 3.59281
(p.treatment <- 1 - pf(F.treatment,df.treatment,df.interaction) )
## [1] 0.127879
(F.therapist <- MS.therapist/MS.interaction)
## [1] 1.2994
(p.therapist <- 1 - pf(F.therapist,df.therapist,df.interaction) )
## [1] 0.367443
# F & p values for (random) interaction are not changed
(F.interaction <- MS.interaction/MS.residuals) # compare to residuals
## [1] 1.89773
(p.interaction <- 1 - pf(F.interaction,df.interaction,df.residuals) )
## [1] 0.132051
```

(b) Assuming both factors are random, calculate all of the variance components. Copy each variance component into your answer booklet and label each value.

Answer: Both factors are random, and therefore the variance components for the main effects are calculated by subtracting the MS-interaction and dividing by the number of observations that were used to calculate the MS value for the main effect. The variance component for the interaction is calculated by subtracting the MS residuals value. The relevant equations are in Section 10.3.3 of the course notes.
n.therapists <- 3
n.treatments <- 3
n <- 5
(var.comp.error <- MS.residuals)
## [1] 8.8
(var.comp.interaction <- (MS.interaction-MS.residuals)/n )
## [1] 1.58
(var.comp.treatment <- (MS.treatment-MS.interaction)/(n.therapists*n) )
## [1] 2.88667
(var.comp.therapist <- (MS.therapist-MS.interaction)/(n.treatments*n) )
## [1] 0.333333
(c) How do the null hypotheses about the fixed and random factors differ?

**Answer:** For a fixed factor, the null hypothesis is that all of the effects ($\alpha$’s) used in your experiment are zero. For a random factor, the null hypothesis is that the variance of the population of effects is zero ($\sigma^2_{\alpha} = 0$).

5. An experiment is conducted to compare the effects of four types of dietary supplements on cognition. Prior to the start of the experiments, all subjects took a standardized test that assessed attention, memory, and problem solving...

(a) Conduct a one-way ANOVA to evaluate the effect of dietary supplement on score. Copy the ANOVA table into your answer booklet.

```r
summary(aov(score~group,data=diet.data))
## Df Sum Sq Mean Sq F value Pr(>F)
## group 3 3797 1266 2.07 0.12
## Residuals 44 26951 613
```

(b) Using pre.test as a covariate, conduct an ANCOVA to evaluate the effect of dietary supplement on score. Copy the ANCOVA table into your answer booklet.

```r
summary(aov(score~pre.test+group,data=diet.data))
## Df Sum Sq Mean Sq F value Pr(>F)
## pre.test 1 6998 6998 15.72 0.00027 ***
## group 3 4615 1538 3.46 0.02448 *
## Residuals 43 19136 445
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(c) For the ANCOVA performed in the previous question, does the order of the terms in the statistical model matter? Is one order better than the other? Explain.

**Answer:** Yes, the order is important because the $F$ test for group depends on whether it appears first or second (see below). (N.B. This order-dependence implies that the covariate and grouping variable are not orthogonal.) I think that group should appear after pre.test because we are treating the pre-treatment measure as a nuisance variable: we want to know the groups differ after accounting for the effects of pre-treatment differences in cognition.

```r
summary(aov(score~group + pre.test,data=diet.data))
## Df Sum Sq Mean Sq F value Pr(>F)
## group 3 3797 1266 2.84 0.04873 *
## pre.test 1 7815 7815 17.56 0.00014 ***
## Residuals 43 19136 445
```
(d) Explain why the effect of group differs in the ANOVA and ANCOVA.

**Answer:** The effect of group differs primarily because it is being evaluated against a different residuals term. In the ANOVA, the residuals include population error variance plus the variation due to the covariate; in the ANCOVA the variation due to the linear association between the covariate and the dependent variable is removed from the residuals and assigned to the covariate term.

(e) List the best-fitting values of the coefficients of the ANCOVA model and briefly explain what the coefficients mean.  [2 marks]

```r
aov.model <- aov(score~pre.test+group,data=diet.data)
dummy.coef(aov.model)
```

<table>
<thead>
<tr>
<th></th>
<th>(Intercept)</th>
<th>pre.test</th>
<th>group</th>
<th>pre.test:group</th>
</tr>
</thead>
<tbody>
<tr>
<td>placebo</td>
<td>-9.50217</td>
<td>0.853954</td>
<td>1.1354874</td>
<td>-14.4877392</td>
</tr>
<tr>
<td>ginkgo</td>
<td>-19.635217</td>
<td>0.853954</td>
<td>-14.4877392</td>
<td>13.3362592</td>
</tr>
<tr>
<td>B12</td>
<td>-19.635217</td>
<td>0.853954</td>
<td>13.3362592</td>
<td>0.0159926</td>
</tr>
<tr>
<td>fishoil</td>
<td>-19.635217</td>
<td>0.853954</td>
<td>13.3362592</td>
<td>0.0159926</td>
</tr>
</tbody>
</table>

**Answer:** For each group, the ANCOVA fits a line relating the dependent variable to the covariate. The intercepts of the lines can vary across groups, but the slopes are constrained to be equal. The value of `pre.test` is the slope of the lines. The intercepts of the four lines equal the overall intercept plus the value of the parameter for each group. Thus, the intercept for the placebo group is \(-9.502 + 1.135\), for the ginkgo group it is \(-9.502 + 14.488\), etc.

(f) Conduct a statistical test to evaluate the ANCOVA’s homogeneity of slopes assumption. Is the assumption reasonable for your data? Write your decision (yes or no) – and the results of the statistical test that supports your decision – in your answer booklet.

**Answer:** To test the homogeneity of slopes assumption, I evaluated the group × covariate interaction (see below). The interaction was not significant \(F(3,262) = 0.58, p = 0.63\), so, yes, the assumption is reasonable for these data.

```
summary(aov(score~pre.test+group+pre.test:group,data=diet.data))
```

<table>
<thead>
<tr>
<th></th>
<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>pre.test</td>
<td>1</td>
<td>6998</td>
<td>6998</td>
<td>15.70</td>
<td>0.003  ***</td>
</tr>
<tr>
<td>group</td>
<td>3</td>
<td>4615</td>
<td>1538</td>
<td>3.45</td>
<td>0.025 *</td>
</tr>
<tr>
<td>pre.test:group</td>
<td>3</td>
<td>1312</td>
<td>437</td>
<td>0.98</td>
<td>0.4110</td>
</tr>
<tr>
<td>Residuals</td>
<td>40</td>
<td>17823</td>
<td>446</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. An experiment was conducted to investigate how memory, specifically forgetting, differs for words and pictures of common objects. Twenty-four young adults were assigned randomly to 1 of 2 study conditions \((n = 12\) per group)…A recognition memory test was completed on four successive days…

(a) Conduct an ANOVA to evaluate the effects of `group`, `day`, and the `group × day` interaction on memory. Copy the ANOVA table with unadjusted p values into your text booklet.

**Answer:** The key here is to recognize that day is a within-subject variable and group is a between-subject variable.

```r
head(mem.data)
```
(b) Conduct an ANOVA that simultaneously determines i) if the overall (i.e., the grand mean) linear trend of memory across days is significantly different from zero; and ii) if the linear trend differs between groups. Write the ANOVA table in your test booklet. You may assume that the linear trend was a planned contrast. (N.B. The weights for a linear trend across four equally-spaced values are \((-1.5, -0.5, 0.5, 1.5)\).

**Answer:** In the following ANOVA table, the intercept evaluates the null hypothesis that the mean linear trend equals zero, and the group effect evaluates the null hypothesis that the linear trend differs between groups.

```
lin.trend.weights <- contr.poly(n=4)
lin.trend.scores <- mem.mat%*%lin.trend.weights
group <- mem.data$group
summary(aov(lin.trend.scores~group),intercept=TRUE)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## (Intercept) 1 1.521 1.521 19.47 8.1e-05 ***
## group 1 0.415 0.415 5.31 0.027 *
## Residuals 38 2.967 0.078
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
lin.trend.weights <- c(-1.5,-0.5,0.5,1.5)
lin.trend.scores.2 <- mem.mat %*% lin.trend.weights
summary(aov(lin.trend.scores.2~group),intercept=TRUE)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## (Intercept) 1 7.600 7.600 19.47 8.1e-05 ***
## group 1 2.07 2.07 5.31 0.027 *
## Residuals 38 14.84 0.39
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```