Statistics
ANOVA and ANCOVA worksheet

Maarten Speekenbrink *UCL Experimental Psychology*
November, 2016

1 SPSS Exercises

1. Klemchuk, Bond and Howell (1990) conducted as study on role-taking in young children. Participants were given a large number of role-taking tasks (e.g., asking what a child thought the *experimenter* saw when a card with a different picture on each side was placed upright between the participant and experimenter). A total `score` was computed, with higher scores indicating better role-taking performance. Children were classified as having had extensive daycare experience or not, and divided into two age groups (2 to 3 years and 4 to 5 years). The hypothesis is that older children and those with daycare experience would be better at role-taking. The data can be found on Moodle as `RoleTaking.sav`.

   (a) Specify the set of orthogonal contrast codes to code for `Daycare` and `Age`, as well as their interaction. Compute the predictors based on these contrast codes and then estimate the parameters of a linear model predicting `score` from these predictors. Remember that the t-tests for the slopes of the predictors given in the output of the SPSS linear regression routine are identical to the results obtained by, for each predictor, performing an F test in which the full model is compared to a model in which the slope of that predictor is fixed to 0. This is thus the procedure based on the Unique Sums of Squares (or Type III SS).

```r
# read the data
library(foreign)
dat <- as.data.frame(read.spss("RoleTaking.sav"))
# contrast for Daycare
dat$c_daycare <- -1
dat$c_daycare[dat$Daycare == "Daycare"] <- 1
# contrast for Age
dat$c_age <- -1
dat$c_age[dat$Age == "Older"] <- 1
# contrast for interaction
dat$c_dayage <- dat$c_daycare * dat$c_age
# specify model
mod <- lm(Score ~ c_daycare + c_age + c_dayage,data=dat)
summary(mod)

##
```

```
## Call:
## lm(formula = Score ~ c_daycare + c_age + c_dayage, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5329 -0.4345 -0.1474  0.3508  1.3879
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.27838    0.13582  -2.050   0.0477 *
## c_daycare    0.28858    0.13582   2.125   0.0405 *
## c_age        0.60763    0.13582   4.474  7.4e-05 ***
## c_dayage    -0.03433    0.13582  -0.253   0.8019
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7647 on 36 degrees of freedom
## Multiple R-squared:  0.4276,Adjusted R-squared:   0.38
## F-statistic: 8.966 on 3 and 36 DF,  p-value: 0.0001436
```

Answer: I have used the following contrasts:

	No Daycare		Daycare	
	Younger	Older	Younger	Older
λ_1	-1	-1	1	1
λ_2	-1	1	-1	1
λ_3	1	-1	-1	1

Which resulted in the model

$$\texttt{score}_i = -.278 + .289X_{1i} + .608X_{2i} - .034X_{3i} + e_i$$

with an SSE of 21.05. Tests of effects of the predictors show a main effect of daycare, $t(36) = 2.125$, $p = .041$. To interpret this effect, we can look at the estimation of the parameter, i.e. $b_1 = .289$. Recall that (with orthogonal contrast variables) this parameter estimate is

$$b_1 = \frac{\sum_{k=1}^{m} \lambda_{1k}\overline{Y}_k}{\sum_{k=1}^{m} \lambda_k^2}$$

which here is

$$\begin{aligned} b_1 &= \frac{\overline{Y}_{\text{Yes,Younger}} + \overline{Y}_{\text{Yes,Older}} - \overline{Y}_{\text{No,Younger}} - \overline{Y}_{\text{No,Older}}}{1^2 + 1^2 + (-1)^2 + (-1)^2} \\ &= \frac{\frac{\overline{Y}_{\text{Yes,Younger}} + \overline{Y}_{\text{Yes,Older}}}{2} - \frac{\overline{Y}_{\text{No,Younger}} + \overline{Y}_{\text{No,Older}}}{2}}{2} \end{aligned}$$

so the difference between the average of the groups with daycare experience and the average of the groups without daycare experience is $2 \times .289 = 0.578$. In addition, there is a significant main effect of age, $t(36) = 4.474$, $p < .001$. To interpret this effect, we can again look at the slope, $b_2 = 0.608$: the average over the older children is $2 \times .608 = 1.216$ points higher than the average over the younger children. As the interaction effect is not significant, $t(36) = -.253$, $p = .802$, we won't discuss it further.

(b) Compute the correlation between the predictors, as well as the tolerance for each predictor, and assess whether the contrast coded predictors are actually orthogonal.

```r
## correlations
cor(dat[,c("c_daycare","c_age","c_dayage")])
```

```
##            c_daycare      c_age   c_dayage
## c_daycare  1.0000000 -0.1711842 -0.2423202
## c_age     -0.1711842  1.0000000 -0.3282440
## c_dayage  -0.2423202 -0.3282440  1.0000000
```

```r
# tolerance
library(car)
1/vif(mod)
```

```
## c_daycare     c_age  c_dayage
## 0.8708273 0.8254717 0.8004574
```

Answer: The correlation between X_1 and X_2 is $R_{1,2} = -.171$, the correlation between X_1 and X_3 is $r_{1,3} = -.242$, and the correlation between X_2 and X_3 is $r_{2,3} = -.328$. The tolerance values for the predictors are $(1 - R_1^2) = 0.871$, $(1 - R_2^2) = 0.825$ and $(1 - R_3^2) = 0.800$ respectively, which are different from 1, so indicating some multicollinearity. So you can see that, while the contrast variables λ are orthogonal, the predictors based on them are not. This is due to the unequal sample sizes.

(c) Let's try a sequential Sums of Squares (or Type I SS) procedure. Start with a MODEL A1 in which you only include the predictor(s) coding for `Age`. Compare the Sum of Squared Errors (SSE) of this model to that of a MODEL C with only an intercept. This gives you the Sum of Squares Reduced (SSR) for the main effect of `Age`, which is equal to the Sequential Sum of Squares (or Type I SS) for this effect. If you want, you can also perform an F test at this stage (i.e., an R^2 change test), but note that this is not the F test you actually need; an F test at this stage uses the SSE of the current model, but you want to use the SSE of the model with all predictors (MODEL A3 below) for a proper Type I test. You will get the SSE of the final model in a later step, so make sure you write down the SSR for now.

```r
# intercept only model
mod_c <- lm(Score ~ 1,data=dat)
# model A1
mod_A1 <- lm(Score ~ c_age,data=dat)
summary(mod_A1)
```

```
##
## Call:
## lm(formula = Score ~ c_age, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36217 -0.57216  0.00085  0.64888  1.60783
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.3689     0.1289  -2.861  0.00683 **
## c_age         0.5710     0.1289   4.428 7.77e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.799 on 38 degrees of freedom
## Multiple R-squared:  0.3404,Adjusted R-squared:  0.323
## F-statistic: 19.61 on 1 and 38 DF,  p-value: 7.771e-05
```

```r
# compare the models
anova(mod_c,mod_A1)
```

```
## Analysis of Variance Table
##
## Model 1: Score ~ 1
## Model 2: Score ~ c_age
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     39 36.778
## 2     38 24.258  1    12.519 19.611 7.771e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# extract the RSS
SSR_age <- anova(mod_c,mod_A1)$"Sum of Sq"[2]
```

Answer: The model is $\texttt{score}_i = -.369 + .571X_2 + e_i$, with an SSE(A1) $= 24.258$, while the error for MODEL C is SSE(C) $= 36.778$. So the Sum of Squares Reduced is SSR $= 36.778 - 24.258 = 12.52$. Performing an F-test at this stage gives a test value of $F_{1,38} = \frac{12.52/(2-1)}{24.258/(40-2)} = 19.613$, $p < .001$. So there seems to be a difference between the two age groups: on average, performance of older children is $2 \times .571 = 1.142$ points higher than that of younger children.

(d) Now specify a MODEL A2 in which you, in addition to the predictors of MODEL A1, also include the predictor(s) coding for `Daycare`. Compare the SSE of this model to that of MODEL A1 you estimated above. The SSR obtained gives you the Type I SS of the main effect of `Daycare`.

```
# model A2
mod_A2 <- lm(Score ~ c_age + c_daycare,data=dat)
# compare the models
anova(mod_A1,mod_A2)

## Analysis of Variance Table
##
## Model 1: Score ~ c_age
## Model 2: Score ~ c_age + c_daycare
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1     38 24.259
## 2     37 21.087  1    3.1714 5.5646 0.02372 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# extract the RSS
SSR_daycare <- anova(mod_A1,mod_A2)$"Sum of Sq"[2]
```

Answer: The model is $\texttt{score}_i = -.269 + .621X_2 + .300X_{1i} + e_i$, with an SSE(A2) = 21.087. So comparing this model to the previous one gives an SSR = 24.258 − 21.087 = 3.171. An F-test at this stage gives the following test result: $F_{1,37} = \frac{3.171/(3-2)}{21.087/(40-3)} = 5.565$, $p < .024$. So adding the predictor coding for daycare to the model significantly reduces the model error.

(e) Now specify a MODEL A3 in which you, in addition to the predictors of MODEL A2, also include the predictor(s) coding for the interaction between `Age` and `Daycare`. Compare the SSE of this model to the SSE of MODEL A2 estimated above to give you Type 1 SS for the interaction effect.

```
# model A3
mod_A3 <- lm(Score ~ c_age + c_daycare + c_dayage,data=dat)
# compare the models
anova(mod_A2,mod_A3)

## Analysis of Variance Table
##
## Model 1: Score ~ c_age + c_daycare
## Model 2: Score ~ c_age + c_daycare + c_dayage
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     37 21.087
## 2     36 21.050  1  0.037362 0.0639 0.8019
```

```r
# extract the RSS
SSR_dayage <- anova(mod_A2,mod_A3)$"Sum of Sq"[2]
```

Answer: The model is the same as that estimated in part a of this question. Comparing this model to the previous one gives an SSR $= 21.087 - 21.050 = .037$. A test gives $F = \frac{.037/(4-3)}{21.050/(40-4)} = .065$, $p = .802$. Note that this is the same result as the test of the interaction in part a of this question. Remember that tests of partial regression coefficients are always a test of a model with that predictor included to the one with just that predictor removed. And that is the same test as the one we performed here.

(f) The SSE of the full MODEL A3 is the SSE you should use to test the significance of the all effects. Perform F tests for the main effects of `Age`, `Daycare`, and the `Age` $\times$ `Daycare` interaction effect, using the Type I SSs you computed in the previous steps.

```r
# SSE model A3
SSE_A3 <- sum(residuals(mod_A3)^2)

## F test for age
# compute the F statistic
F_age <- (SSR_age/(1-0))/(SSE_A3/(40 - 4))
F_age

## [1] 21.41056

# compute a p-value
1-pf(F_age,df1=1,df2=40-4)

## [1] 4.666805e-05

## F test for daycare
# compute the F statistic
F_daycare <- (SSR_daycare/(2-1))/(SSE_A3/(40 - 4))
F_daycare

## [1] 5.423855

# compute a p-value
1-pf(F_daycare,df1=1,df2=40-4)

## [1] 0.02559294

## F test for interaction
# compute the F statistic
F_dayage <- (SSR_dayage/(3-2))/(SSE_A3/(40 - 4))
F_dayage
```

```
## [1] 0.06389862

# compute a p-value
1-pf(F_dayage,df1=1,df2=40-4)

## [1] 0.8018747
```

Answer: The SSE(A3) is 21.05. We can now compute the three tests, based on Type I SS, as

effect	SS	F	p
`Age`	12.52	$\frac{12.52/(2-1)}{21.05/(40-4)} = 21.41$	$< .001$
`Daycare`	3.171	$\frac{3.171/(3-2)}{21.05/(40-4)} = 5.42$	.026
`Age` × `Daycare`	0.037	$\frac{.037/(4-3)}{21.050/(40-4)} = .065$	.802

Note that these test results are slightly different from the ones you obtained if you performed the R^2-change tests earlier. That is because here we use the SSE(A3) as the error SS, while in the previous tests, we used SSE(A1), SSE(A2) and SSE(A3), respectively. This is the difference between "hierarchical regression" and ANOVA Type I tests. In hierarchical regression, you sequentially test whether adding a (set of) predictor(s) results in a significant decrease of the error. In ANOVA Type I tests, you use the same values for the SSR, but now test these against the error of the final model with all predictors included. Because the final error is smaller than that of the intermediate models, this can give more powerful tests. On the other hand, the degrees of freedom for the error (df_2) is smaller for the error in the final model compared to the degrees of freedom for the intermediate models, which can attenuate the increase in power somewhat.

(g) Repeat the analysis above, but now start with the predictor(s) for `Daycare`, then add `Age` and finally add the interaction between `Daycare` and `Age`. You don't have to do this with a multiple regression analysis; you can enter the predictors (in the required order) as "covariates" in the SPSS GLM routine, then click on the "Model" button and choose "Type I" in the Sum of Squares menu. Compare the results to those obtained earlier. Can you explain any differences?

```
# use the aov function to get Type I SS
summary(aov(Score~c_daycare + c_age + c_dayage,data=dat))

##             Df Sum Sq Mean Sq F value   Pr(>F)
## c_daycare    1  1.320   1.320   2.257    0.142
## c_age        1 14.371  14.371  24.577 1.71e-05 ***
## c_dayage     1  0.037   0.037   0.064    0.802
## Residuals   36 21.050   0.585
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Answer:

Using the GLM option with Type I SS gives the following tests:

Daycare: $F_{1,36} = \frac{1.320/(2-1)}{21.050/(40-4)} = 2.257$, $p = .142$

Age: $F_{1,36} = \frac{14.371/(3-2)}{21.050/(40-4)} = 24.577$, $p < .001$

Interaction: $F_{1,36} = \frac{0.037/(4-3)}{21.050/(40-4)} = .065$, $p = .802$

In this analysis, the effect of daycare is not significant, while the effect of age is. In the analysis for part f of this question, the effect of daycare was significant, when added after age. When predictors are (partly) redundant, you would normally expect the SSR to reduce when other predictors are entered into the model (this is what Venn diagrams imply). But this example shows that you can also have the opposite situation, in which the "usefulness" of a predictor increases when controlling for other predictors.

The main thing to realise is that when using Type I tests, the order in which you add predictors matters for their significance tests. With Type III SS, order is irrelevant. But the nice thing about Type I SS is that all the SSR terms add up to the total SSR (when comparing MODEL A3 to MODEL C). This is not the case for type III SS, where the sum of the SSR terms can either be greater or smaller than the total SSR.

A final thing to know is that when all predictors are orthogonal, there is no difference between Type I and Type III SS tests.

2. Open the dataset `2016_Questionnaire_1_cleaned.sav`. This dataset contains the `Weight`, `Height`, and `Gender` of students in this year's course. In this exercise, we'll look at sex differences in weight, whilst controlling for differences in height.

(a) As these are likely to be outliers, filter out any cases with a weight equal to or larger than 150.

```
# read the data
library(foreign)
dat <- as.data.frame(read.spss("2016_Questionnaire_1_cleaned.sav"))

## Warning in read.spss("2016_Questionnaire_1_cleaned.sav"):  2016_Questionna
File-indicated value is different from internal value for at least one
of the three system values.  SYSMIS: indicated -1.79769e+308, expected
-1.79769e+308; HIGHEST: 1.79769e+308, 1.79769e+308; LOWEST: -1.79769e+308,
-1.79769e+308

## Warning in read.spss("2016_Questionnaire_1_cleaned.sav"):  2016_Questionna
Unrecognized record type 7, subtype 18 encountered in system file

# filter out large weights
dat <- subset(dat,Weight < 150)
```

(b) Use dummy coding to code for `gender`. Use the value 1 for males, and 0 for females and call the resulting variable `dummyM`. Estimate the model

$$\texttt{weight}_i = \beta_0 + \beta_1 \texttt{dummyM}_i + \beta_2 \texttt{height}_i + \beta_3 (\texttt{dummyM} \times \texttt{height})_i + \epsilon_i$$

and assess the significance of the predictors.

```
# first let's redo the factor labels
levels(dat$Gender) <- c("Female","Male")
# create a dummy variable
dat$dummyM <- NA
dat$dummyM[dat$Gender == "Male"] <- 1
dat$dummyM[dat$Gender == "Female"] <- 0
# estimate the model
mod_dummy <- lm(Weight~dummyM*Height,data=dat)
summary(mod_dummy)
```

```
##
## Call:
## lm(formula = Weight ~ dummyM * Height, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.269  -6.238  -1.199   3.808  42.051
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -7.8007    23.2519  -0.335  0.73778
## dummyM        -59.7557    42.8286  -1.395  0.16525
## Height          0.3984     0.1395   2.856  0.00498 **
## dummyM:Height   0.3892     0.2468   1.577  0.11712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.689 on 134 degrees of freedom
## Multiple R-squared:  0.3713,Adjusted R-squared:  0.3572
## F-statistic: 26.38 on 3 and 134 DF,  p-value: 1.795e-13
```

Answer: The estimated model is

$$\texttt{weight}_i = -7.801 - 59.756 \times \texttt{dummyM}_i + 0.398 \times \texttt{height}_i + 0.389 \times (\texttt{dummyM} \times \texttt{height})_i + e_i$$

Only the effect of height is significant. So, from this analysis, one can conclude that there are no gender differences in weight after controlling for differences in height. However, it is important to realise that because we have an interaction in the model, we are testing for the effect of gender for people with a height of 0.

We can work out the predicted regression lines for males and females in the usual way. For females, we get

$$\begin{aligned}\texttt{weight}_i &= -7.801 - 59.756 \times 0 + 0.398 \times \texttt{height}_i + 0.389(0 \times \texttt{height})_i + e_i \\ &= -7.801 + 0.398 \times \texttt{height}_i + e_i\end{aligned}$$

and for males, we get

$$\begin{aligned}\texttt{weight}_i &= -7.801 - 59.756 \times 1 + 0.398 \times \texttt{height}_i + 0.389 \times (1 \times \texttt{height})_i + e_i \\ &= -67.557 + 0.787 \times \texttt{height}_i + e_i\end{aligned}$$

(c) Center `height` and call this variable `height0`. Repeat the analysis above, but now using `height0` instead of `height`. Are there any differences compared to the previous analysis? And if so, can you explain these?

```
# center height
dat$Height0 <- scale(dat$Height,scale=FALSE)
# estimate the model
mod_dummy0 <- lm(Weight~dummyM*Height0,data=dat)
summary(mod_dummy0)

##
## Call:
## lm(formula = Weight ~ dummyM * Height0, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.269  -6.238  -1.199   3.808  42.051
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     59.9307     1.1316  52.962  < 2e-16 ***
## dummyM           6.4122     2.2389   2.864  0.00486 **
## Height0          0.3984     0.1395   2.856  0.00498 **
## dummyM:Height0   0.3892     0.2468   1.577  0.11712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.689 on 134 degrees of freedom
## Multiple R-squared:  0.3713,Adjusted R-squared:  0.3572
## F-statistic: 26.38 on 3 and 134 DF,  p-value: 1.795e-13
```

Answer: The estimated model is

$$\texttt{weight}_i = 59.931 + 6.412 \times \texttt{dummyM}_i + 0.398 \times \texttt{height0}_i + 0.389 \times (\texttt{dummyM} \times \texttt{height0})_i + e_i$$

Again, only the effect of `height0` is significant. The slope of `dummyM` is now positive, so males of an average height are generally heavier than females of the same height. The reason for this difference is in the centering of height. Recall that the "simple slopes" in a model with interaction terms reflect the slope of a predictor when the value of the other predictor(s) is 0. So here, you estimate the difference between males and females for people of an average height (the centered variable `height0` is 0 at the mean of `height`). In the previous analysis with the uncentered height, you estimated the difference between males and females for people with no height (i.e., when `height`=0). That is not such an interesting test.

Working out the predicted regression lines gives for females:

$$\begin{aligned}\texttt{weight}_i &= 59.931 + 6.412 \times 0 + 0.398 \times \texttt{height0}_i + 0.389(0 \times \texttt{height0})_i + e_i \\ &= 59.931 + 0.398 \times \texttt{height0}_i\end{aligned}$$

and for males:

$$\begin{aligned}\texttt{weight}_i &= 59.931 + 6.412 \times 1 + 0.398 \times \texttt{height0}_i + 0.389(1 \times \texttt{height0})_i + e_i \\ &= 66.343 + 0.787 \times \texttt{height0}_i + e_i\end{aligned}$$

(d) Use effect coding to code for `gender`. Call this variable `effect` and re-estimate the model above, replacing `dummyM` with `effect`.

```
# create an effect-coded predictor
dat$effect <- NA
dat$effect[dat$Gender == "Male"] <- 1
dat$effect[dat$Gender == "Female"] <- -1
# estimate the model
mod_effect0 <- lm(Weight~effect*Height0,data=dat)
summary(mod_effect0)

##
## Call:
## lm(formula = Weight ~ effect * Height0, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.269  -6.238  -1.199   3.808  42.051
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     63.1368     1.1195  56.399  < 2e-16 ***
## effect           3.2061     1.1195   2.864  0.00486 **
## Height0          0.5930     0.1234   4.806 4.07e-06 ***
## effect:Height0   0.1946     0.1234   1.577  0.11712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.689 on 134 degrees of freedom
## Multiple R-squared:  0.3713,Adjusted R-squared:  0.3572
## F-statistic: 26.38 on 3 and 134 DF,  p-value: 1.795e-13
```

Answer: I have used a value of 1 for males, and -1 for females. The estimated model with centered height is then

$$\texttt{weight}_i = 63.137 + 3.206 \times \texttt{effect}_i + 0.593 \times \texttt{height0}_i + 0.195 \times (\texttt{effect} \times \texttt{height0})_i + e_i$$

As before, only the effect of `height0` is significant. You can see that the slope of `height0` has changed. This now represents the "average" effect of height over males and females (so the "main" effect of height, ignoring gender).

Working out the predicted regression lines gives for females:

$$\begin{aligned}\texttt{weight}_i &= 63.137 + 3.206 \times (-1) + 0.593 \times \texttt{height0}_i + 0.195 \times (-1 \times \texttt{height0})_i + e_i \\ &= (63.137 - 3.206) + (0.593 - 0.195) \times \texttt{height0}_i \\ &= 59.931 + 0.398 \times \texttt{height0}_i\end{aligned}$$

and for males:

$$\begin{aligned}\texttt{weight}_i &= 66.343 + 0.788 \times 1 + 0.686 \times \texttt{height0}_i + 0.203 \times (1 \times \texttt{height0})_i + e_i \\ &= (63.137 + 3.206) + (0.593 + 0.195) \times \texttt{height0}_i \\ &= 65.628 + 0.889 \times \texttt{height0}_i + e_i\end{aligned}$$

Note that this is exactly the same as in the previous model with centered hight and a dummy code for gender. That is as it should be, as the way you contrast code gender should not affect the estimated relation between height and weight for the genders.

(e) Look at the results of these models. If there are differences, can you explain these? Which of the models do you think makes the most sense?

Answer: Ok, there are a lot of comparisons to make. The main idea is to (yet again) see the effect of different coding schemes and centering. I've already indicated the reasons for differences in the answers above. Personally, I would go for the last model, with effect coding and centered height. In this model, the slope for height is the "average" slope over males and females, while the interaction reflects the difference from this slope for males and females. The slope for `effect` tells us about the difference in weight between females and males of an average height, while the intercept is the average weight for people of an average height (averaging over males and females).

3. A team of researchers is interested whether two methods of hypnotic induction, Method A and Method B, differ in effectiveness. Twenty volunteers are measured on a standard index of "primary suggestibility" (X), which is known to be correlated with receptivity to hypnotic induction. They are then randomly assigned to two groups of 10 subjects each. In each group, one of the hypnotic methods was administered. The dependent variable (Y) is the subject's score on a standard index of hypnotic induction, measured during the administration of Method A or B. You can find the resulting data on Moodle as `hypnosis.sav`.

(a) Ignore the "primary suggestibility" for the moment, and perform a suitable test to assess whether the methods differ in hypnotic induction.

```
dat <- as.data.frame(read.spss("hypnosis.sav"))
# use a Helmert contrast for method
contrasts(dat$method) <- contr.helmert(2)
# estimate and summarize a regression model
summary(lm(y~method,data=dat))

##
## Call:
## lm(formula = y ~ method, data = dat)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.200 -5.125  0.350  4.825 10.800
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   28.650      1.357  21.119 3.75e-14 ***
## method1       -0.550      1.357  -0.405     0.69
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.067 on 18 degrees of freedom
## Multiple R-squared:  0.009049,Adjusted R-squared:  -0.046
## F-statistic: 0.1644 on 1 and 18 DF,  p-value: 0.6899
```

Answer: I've used the GLM option in SPSS and let the program compute a contrast code for Method. The gave a test result for Method of $F_{1,18} = 0.164$, $p = .69$. So, ignoring the primary suggestibility, there doesn't seem to be a general difference in hypnotic induction between the methods. Alternatively, you could have used regression with a contrast coded predictor (e.g., a value of $\lambda = -1$ for Method A, and a value of $\lambda = 1$ for Method B.). This would give the same results.

(b) Perform an ANCOVA (use SPSS GLM), with "primary suggestibility" as covariate, to answer the same question as above. Is there a difference in the results? If so, can you explain this difference?

```
# estimate and summarize a regression model
summary(lm(y~method+x,data=dat))

##
## Call:
## lm(formula = y ~ method + x, data = dat)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6348 -2.5099 -0.2038  1.8871  4.7453
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  15.7827     1.6016   9.854 1.92e-08 ***
## method1      -2.5773     0.6438  -4.003 0.000921 ***
## x             0.8275     0.0955   8.665 1.21e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.682 on 17 degrees of freedom
## Multiple R-squared:  0.817,Adjusted R-squared:  0.7955
## F-statistic: 37.96 on 2 and 17 DF,  p-value: 5.372e-07
```

```
# you can also get F tests (instead of the t-tests) with the
# drop1 function:
drop1(lm(y~method+x,data=dat),test="F")
```

```
## Single term deletions
##
## Model:
## y ~ method + x
##        Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>              122.32 42.218
## method  1    115.31 237.63 53.499  16.025 0.0009209 ***
## x       1    540.18 662.50 74.006  75.074 1.211e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Answer: The results show a significant effect of primary suggestibility, $F_{1,17} = 75.074$, $p < .001$. Moreover. using an ANCOVA, we get a significant effect of method, $F_{1,17} = 16.025$, $p = .001$. So when we control for primary suggestibility, there is an effect of the method. The reason for this is that by accounting for individual differences in primary suggestibility, the error of the model reduces substantially, which makes the effect of method "stand out" more clearly (i.e., it gives a more powerful test for method).

(c) Specify a suitable contrast code for method and repeat the analysis with multiple regression.

```
## we've already done this, but let's do it more explicitely
## set up a contrast coded predictor
dat$c_method <- NA
dat$c_method[dat$method=="A"] <- -1
```

```r
dat$c_method[dat$method=="B"] <- 1
## estimate the regression model
summary(lm(y~c_method + x,data = dat))

##
## Call:
## lm(formula = y ~ c_method + x, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6348 -2.5099 -0.2038  1.8871  4.7453
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  15.7827     1.6016   9.854 1.92e-08 ***
## c_method     -2.5773     0.6438  -4.003 0.000921 ***
## x             0.8275     0.0955   8.665 1.21e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.682 on 17 degrees of freedom
## Multiple R-squared:  0.817,Adjusted R-squared:  0.7955
## F-statistic: 37.96 on 2 and 17 DF,  p-value: 5.372e-07
```

Answer: I used a value of $\lambda_1 = -1$ for method A and a value of $\lambda_1 = 1$ for method B and created a predictor X_2 using these values. Remember that the slope for this predictor represents half of the difference between the methods. The estimated model is

$$Y_i = 15.783 + .827X_i - 2.577X_{2i} + e_i$$

with an SSE of 122.320. The negative effect of the contrast coded predictor X_2 shows that the difference between method B and method A is $2 \times -2.577 = -5.154$, i.e., method B results in 5.154 lower hypnotic induction scores than method A.

(d) Test the homogeneity of the regression slopes for the model above.

```r
## we've already done this, but let's do it more explicitely
## set up a contrast coded predictor for the interaction
dat$c_methodx <- dat$c_method*dat$x
## estimate the regression model
summary(lm(y~c_method + x + c_methodx,data = dat))

##
## Call:
## lm(formula = y ~ c_method + x + c_methodx, data = dat)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.8562 -1.7500  0.0696  1.8982  4.0207 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 15.76566    1.56549  10.071 2.49e-08 ***
## c_method    -0.65696    1.56549  -0.420    0.680    
## x            0.84856    0.09466   8.964 1.23e-07 ***
## c_methodx   -0.12682    0.09466  -1.340    0.199    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 2.622 on 16 degrees of freedom
## Multiple R-squared:  0.8355,Adjusted R-squared:  0.8046 
## F-statistic: 27.09 on 3 and 16 DF,  p-value: 1.655e-06
```

```
## note that we could have also used:
summary(lm(y~c_method * x,data = dat))
```

```
## 
## Call:
## lm(formula = y ~ c_method * x, data = dat)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.8562 -1.7500  0.0696  1.8982  4.0207 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 15.76566    1.56549  10.071 2.49e-08 ***
## c_method    -0.65696    1.56549  -0.420    0.680    
## x            0.84856    0.09466   8.964 1.23e-07 ***
## c_method:x  -0.12682    0.09466  -1.340    0.199    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 2.622 on 16 degrees of freedom
## Multiple R-squared:  0.8355,Adjusted R-squared:  0.8046 
## F-statistic: 27.09 on 3 and 16 DF,  p-value: 1.655e-06
```

Answer: To test for the homogeneity of the regression slopes, we create a new predictor X_3 as the product of X (primary suggestibility) and X_2, the contrast coded predictor. The test for this interaction effect is $F_{1,16} = 1.795$, $p = .199$. As the effect is not significant, we can retain the null hypothesis of homogeneous regression slopes (i.e., the relation between hypnotic induction and primary suggestibility is the same for the two methods.

4. Sethi and Seligman (1993) conducted a study investigating the relation between optimism and religious fundamentalism. It has been suggested that the unquestioning faith encouraged by the more fundamentalist religions may result in less concern with the cares of the world, and with that a more optimistic view of the world and one's place in it. Sethi and Seligman collected data from over 600 adults from nine religious groups, classified into three major categories (`Group`: Fundamentalists, Moderates, and Liberals). For each participant, they obtained a measure of Optimism (`Optim`) and three measures of religion's role in the person's life: influence (`RInfluen`, e.g.,"To what extend do your religious views influence who you associate with?"), religious involvement (`RInvolve`, e.g., "How often do you attend religious services?"), and religious hope (`RHope`, e.g.,"Do you believe there is a heaven?"). Data simulated to correspond to Sethi and Seligman's findings can be found on Moodle as `FundOptim.sav`.

   (a) Use a oneway ANOVA to test whether there is an effect of `Group` on `Optim`. If there is an effect, use a Tukey HSD test to see where the differences are.

```
dat <- as.data.frame(read.spss("FundOptim.sav"))

## Warning in read.spss("FundOptim.sav"): FundOptim.sav: File-indicated
value is different from internal value for at least one of the three
system values. SYSMIS: indicated -1.79769e+308, expected -1.79769e+308;
HIGHEST: 1.79769e+308, 1.79769e+308; LOWEST: -1.79769e+308, -1.79769e+308
## Warning in read.spss("FundOptim.sav"): FundOptim.sav: Unrecognized
record type 7, subtype 18 encountered in system file

mod <- aov(Optim ~ Group,data=dat)
summary(mod)

##              Df Sum Sq Mean Sq F value   Pr(>F)
## Group         2    385  192.48   19.92 4.23e-09 ***
## Residuals   597   5769    9.66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TukeyHSD(mod)

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Optim ~ Group, data = dat)
##
## $Group
##                               diff       lwr      upr     p adj
## Moderate-Liberal         0.9630952 0.1661777 1.760013 0.0129546
```

```
## Fundamentalist-Liberal  2.1916667 1.3482884 3.035045 0.0000000
## Fundamentalist-Moderate 1.2285714 0.5523645 1.904778 0.0000678
```

Answer: The ANOVA shows a significant effect of Group, $F_{2,597} = 1.918$, $p < .001$. Post-hoc Tukey tests show that all groups differ significantly from each other.

(b) As the hypothesis proposes an ordering of the three fundamentalism categories with respect to their level of optimism, a polynomial contrast seems appropriate. Define the polynomial contrast coded predictors necessary to code for `Group` and repeat the analysis above using these contrast coded predictors.

```
contrasts(dat$Group) <- cbind("linear"=c(-1,0,1),
                              "quadratic" = c(-1,2,-1))
mod <- lm(Optim~Group,data=dat)
summary(mod)

##
## Call:
## lm(formula = Optim ~ Group, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.9833  -1.9833   0.0536   2.0167   9.0536
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        2.03492    0.13472  15.104  < 2e-16 ***
## Grouplinear        1.09583    0.17947   6.106 1.84e-09 ***
## Groupquadratic    -0.04425    0.08610  -0.514    0.608
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.109 on 597 degrees of freedom
## Multiple R-squared:  0.06255,Adjusted R-squared:  0.05941
## F-statistic: 19.92 on 2 and 597 DF,  p-value: 4.227e-09

# F test for the main effect of Group:
anova(lm(Optim~1,data=dat),mod)

## Analysis of Variance Table
##
## Model 1: Optim ~ 1
## Model 2: Optim ~ Group
##   Res.Df  RSS Df Sum of Sq      F    Pr(>F)
## 1    599 6154
## 2    597 5769  2    384.96 19.918 4.227e-09 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# or this way:
summary(aov(Optim~Group,data=dat))

##              Df Sum Sq Mean Sq F value   Pr(>F)
## Group         2    385  192.48   19.92 4.23e-09 ***
## Residuals   597   5769    9.66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Answer: I used the following contrast variables:

	Liberals	Moderates	Fundamentalists
λ_1	-1	0	1
λ_2	-1	2	-1

This gave the model

$$\texttt{Optim}_i = 2.035 + 1.096X_{1i} - .044X_{2i}$$

where X_1 is the predictor coding for the linear effect (λ_1) and X_2 is the predictor coding for the quadratic effect (λ_2). Tests show there is a significant linear effect of group, $F_{1,597} = 37.280$, $p < .001$, but no significant quadratic effect, $F_{1,597} = .264$, $p = .608$. You can conclude the difference between Liberals and Moderates (predicted to be $b_1 = 1.096$) is the same as the difference between Fundamentalists and Moderates.

(c) We will now look more closely at the results obtained above. First, test for each of the three measures `RInfluen`, `RInvolve` and `RHope` whether there is an effect of `Group`. Use the polynomial contrast coded predictors for this. What do the results of these tests indicate?

```
summary(lm(RInfluen~Group,data=dat))

##
## Call:
## lm(formula = RInfluen ~ Group, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1821 -0.7833 -0.1821  0.8179  3.8179
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       4.48849    0.04954  90.602  < 2e-16 ***
## Grouplinear       0.85833    0.06600  13.006  < 2e-16 ***
```

```
## Groupquadratic -0.15317    0.03166  -4.838 1.67e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.143 on 597 degrees of freedom
## Multiple R-squared:  0.2702,Adjusted R-squared:  0.2678
## F-statistic: 110.5 on 2 and 597 DF,  p-value: < 2.2e-16
```

```
summary(lm(RInvolve~Group,data=dat))
```

```
##
## Call:
## lm(formula = RInvolve ~ Group, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5393 -0.5393  0.0850  0.6417  4.0850
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     3.27087    0.05240  62.421  < 2e-16 ***
## Grouplinear     0.77833    0.06981  11.150  < 2e-16 ***
## Groupquadratic  0.13421    0.03349   4.007 6.91e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.209 on 597 degrees of freedom
## Multiple R-squared:  0.1774,Adjusted R-squared:  0.1746
## F-statistic: 64.37 on 2 and 597 DF,  p-value: < 2.2e-16
```

```
summary(lm(RHope~Group,data=dat))
```

```
##
## Call:
## lm(formula = RHope ~ Group, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1071 -0.5400 -0.1071  0.8929  3.8929
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     4.70460    0.05576  84.368  < 2e-16 ***
## Grouplinear     1.03667    0.07429  13.955  < 2e-16 ***
## Groupquadratic  0.20127    0.03564   5.648 2.52e-08 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.287 on 597 degrees of freedom
## Multiple R-squared:  0.256,Adjusted R-squared:  0.2535
## F-statistic: 102.7 on 2 and 597 DF,  p-value: < 2.2e-16
```

Answer: For `RInfluen`, there is a significant effect of both the linear, $F_{1,597} = 169.149$, $p < .001$, and quadratic predictor, $F_{1,597} = 23.405$, $p < .001$. For `RInvolve`, there is a significant effect of both the linear, $F_{1,597} = 124.324$, $p < .001$, and quadratic predictor, $F_{1,597} = 16.060$, $p < .001$. For `RHope`, there is a significant effect of both the linear, $F_{1,597} = 194.747$, $p < .001$, and quadratic predictor, $F_{1,597} = 31.895$, $p < .001$. In all cases, the slope of the linear predictor is positive, so these variables increase with the level of fundamentalism. The slope of the quadratic predictor is negative for `RInvolve`, which indicates a "dip" in the middle, which effectively means that the difference between Moderates and Liberals is less than the difference between Fundamentalists and Moderates. For the other two variables, the slope of the quadratic predictor is positive, which indicates a "peak" in the middle, which effectively means that the difference between Fundamentalists and Moderates is smaller than the difference between Moderates and Liberals.

(d) You may have guessed already, but now let's try to predict Optimism from these three variables (disregarding `Group` of the moment). Estimate the parameters of the model

$$\texttt{Optim}_i = \beta_0 + \beta_1 \texttt{RInfluen}_i + \beta_2 \texttt{Rinvolve}_i + \beta_3 \texttt{RHope}_i + \epsilon_i$$

and test whether each parameter β_1, β_2 and β_3 is significantly different from 0. Compare the overall model fit (e.g., the R^2) to that of the model with the polynomial contrast for `Group` as predictors (you don't have to test for a significant difference... why not?).

```
summary(lm(Optim ~ RInfluen + RInvolve + RHope,data=dat))

##
## Call:
## lm(formula = Optim ~ RInfluen + RInvolve + RHope, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.7197  -2.1323   0.1792   1.9042   8.9277
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.89492    0.51180  -3.702 0.000233 ***
## RInfluen     0.48958    0.10710   4.571 5.89e-06 ***
```

```
## RInvolve    -0.07938    0.11634  -0.682 0.495294
## RHope        0.42794    0.10231   4.183 3.31e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.043 on 596 degrees of freedom
## Multiple R-squared:  0.1031,Adjusted R-squared:  0.09855
## F-statistic: 22.83 on 3 and 596 DF,  p-value: 5.323e-14
```

Answer: The estimated model is

$$\texttt{Optim}_i = -1.892 + 0.49 \times \texttt{RInfluen}_i - 0.079 \times \texttt{RInvolve}_i + .428 \times \texttt{RHope}_i + e_i$$

The tests for the effects (whether the slopes differ from 0) are $t_{596} = 4.571$, $p < .001$ for `RInfluen`, $t_{596} = -.682$, $p = .495$ for `Rinvolve`, and $t_{596} = 4.183$, $p < .001$ for `RHope`. You could of course also look at the F tests (remember, you can square the t values to get the F values and the results are, in terms of p values, identical). The R^2 for this model is $R^2 = .103$. The R^2 of the model with the predictors coding for group was $R^2 = .063$, so the current model seems to do a better job than that one, although it has four rather than three parameters (and the more parameters, the better the model fit). You can't test whether the current model is significantly better than the model with the contrast coded predictors, because these models are not nested (one model is not a special case of the other model). The tests we have been using are always whether adding additional predictors to a model improves the fit. But now, one model is not a more general version of another, as the models use different sets of predictors.

(e) Test whether adding the polynomial contrast for `Group` significantly improves the fit of the model above (i.e., test whether there is a significant reduction in the error when adding the polynomial contrast to the model above). What does this test indicate? Can you say anything about the "causal chain" linking Fundamentalism to Optimism?

```
modC <- lm(Optim ~ RInfluen + RInvolve + RHope,data=dat)
modA <- lm(Optim ~ RInfluen + RInvolve + RHope + Group,data=dat)
summary(modA)

##
## Call:
## lm(formula = Optim ~ RInfluen + RInvolve + RHope + Group, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.7601  -2.1694   0.1731   2.0208   8.6610
##
## Coefficients:
```

```
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.18450    0.60192  -1.968 0.049548 *
## RInfluen        0.38796    0.11846   3.275 0.001118 **
## RInvolve       -0.09408    0.11777  -0.799 0.424713
## RHope           0.37958    0.10830   3.505 0.000491 ***
## Grouplinear     0.44256    0.21376   2.070 0.038848 *
## Groupquadratic -0.04859    0.09042  -0.537 0.591193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.035 on 594 degrees of freedom
## Multiple R-squared:  0.1107,Adjusted R-squared:  0.1032
## F-statistic: 14.79 on 5 and 594 DF,  p-value: 1.098e-13
```

```r
anova(modC,modA)
```

```
## Analysis of Variance Table
##
## Model 1: Optim ~ RInfluen + RInvolve + RHope
## Model 2: Optim ~ RInfluen + RInvolve + RHope + Group
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1    596 5519.8
## 2    594 5472.8  2    46.953 2.5481 0.07909 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Answer: For this model, the error is SSE = 5472.8, and for the model with `RInfluen`, `Rinvolve` and `RHope`, the error was SSE = 5519.754. So the Proportional Reduction is Error is PRE $= \frac{5519.754-5472.8}{5519.754} = 0.0085$, which is not very large. Indeed, the test shows that adding the contrast coded predictors does not significantly reduce the error, $F_{2,594} = \frac{(5519.754-5472.80)/(6-4)}{5472.80/(600-6)} = 2.548$, $p = .079$. However, the results for the individual contrasts show that the linear effect of fundamentalism is significant.

We can't be sure about the "causal chain" linking Fundamentalism to Optimism, but it seems to be the case that the effect of fundamentalism on optimism is at least partially mediated by religious influence and hope. In other words, fundamentalism affects religious influence and hope, and religious influence and hope affect optimism. We saw that the level of fundamentalism has an effect on religious influence and hope. We also saw that group, by itself, has a significant effect on optimism. But, when controlling for religious influence, involvement and hope, the main effect of fundamentalism was no longer significant (although the linear effect was still significant). Recall that mediation has four requirements. If the relation between X and Y is mediated by Z, then: (1) X has to have a significant effect on Y, (2) X has to have a significant effect on Z, (3) when controlling for X, Z has a significant effect on Y, and (4) when controlling for Z, the effect of X on Y is absent or reduced. In terms of the main effect of fundamentalism (i.e., the omnibus test for the two contrast codes), these conditions are fulfilled. Looking at just the linear effect of fundamentalism, we could conclude that the effect of fundamentalism on optimism is partially mediated by religious influence and hope.

(f) If you feel like it, test whether the slopes of `RInfluen`, `RInvolve` and `RHope` are homogeneous over the Groups.

```
modA2 <- lm(Optim ~ (RInfluen + RInvolve + RHope)*Group,data=dat)
summary(modA2)
```

```
##
## Call:
## lm(formula = Optim ~ (RInfluen + RInvolve + RHope) * Group, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.5556  -2.0430   0.1399   1.9649   8.5465
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)            -1.099754   0.666432  -1.650  0.09943 .
## RInfluen                0.406663   0.126264   3.221  0.00135 **
## RInvolve               -0.177372   0.134050  -1.323  0.18629
## RHope                   0.387532   0.125541   3.087  0.00212 **
## Grouplinear             0.345793   0.912357   0.379  0.70482
## Groupquadratic         -0.015139   0.408248  -0.037  0.97043
## RInfluen:Grouplinear   -0.044275   0.169719  -0.261  0.79428
## RInfluen:Groupquadratic -0.043278   0.079631  -0.543  0.58701
## RInvolve:Grouplinear    0.276115   0.181354   1.523  0.12842
## RInvolve:Groupquadratic  0.062774   0.083703   0.750  0.45358
## RHope:Grouplinear      -0.114903   0.174755  -0.658  0.51111
## RHope:Groupquadratic   -0.005682   0.074705  -0.076  0.93940
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.044 on 588 degrees of freedom
## Multiple R-squared:  0.1148,Adjusted R-squared:  0.09822
## F-statistic: 6.931 on 11 and 588 DF,  p-value: 4.586e-11
```

```
# compare to a model with no interactions for the omnibus test
anova(modA,modA2)
```

```
## Analysis of Variance Table
##
## Model 1: Optim ~ RInfluen + RInvolve + RHope + Group
```

```
## Model 2: Optim ~ (RInfluen + RInvolve + RHope) * Group
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    594 5472.8
## 2    588 5447.6  6    25.193 0.4532 0.8428
```

Answer: The homogeneity assumption seems to hold. This is indicated by both the omnibus test for all interactions, $F_{6,588} = \frac{(5472.80-5447.607)/(12-6)}{5447.607/(600-12)} = .453$, $p = .843$, as well as the individual tests for each interaction effect, which are all non-significant.