Multiple Regression II

```
##
## Call:
## lm(formula = Weight ~ Age + Poverty, data = nhanes)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -47.54 -19.90   0.82  16.96  65.53
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  58.8926     7.0353   8.371 8.69e-13 ***
## Age           0.3537     0.1310   2.699  0.00835 **
## Poverty      -0.3501     1.5890  -0.220  0.82612
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 24.93 on 87 degrees of freedom
##   (10 observations deleted due to missingness)
## Multiple R-squared:  0.07949,    Adjusted R-squared:  0.05833
## F-statistic: 3.756 on 2 and 87 DF,  p-value: 0.02724
```

- Interpret all coefficients
- Interpret all significance tests of coefficients
- Is it a good model?

Last time

- Introduction to multiple regression
- Interpreting coefficient estimates
- Estimating model fit
- Significance tests (omnibus and coefficients)

Today

- Tolerance
- Hierarchical regression/model comparison
- Categorical predictors

The Data

```r
stress.data = read.csv(here("data/stress.csv"))
library(psych)
describe(stress.data)
```

```
##          vars   n   mean     sd median trimmed    mad  min    max
## id          1 118 488.65 295.95 462.50  485.76 372.13 2.00 986.00
## Anxiety     2 118   7.61   2.49   7.75    7.67   2.26 0.70  14.64
## Stress      3 118   5.18   1.88   5.27    5.17   1.65 0.62  10.32
## Support     4 118   8.73   3.28   8.52    8.66   3.16 0.02  17.34
## group*      5 118   1.53   0.50   2.00    1.53   0.00 1.00   2.00
##           range  skew kurtosis    se
## id       984.00  0.10    -1.29 27.24
## Anxiety   13.94 -0.18     0.28  0.23
## Stress     9.71  0.08     0.22  0.17
## Support   17.32  0.18     0.19  0.30
## group*     1.00 -0.10    -2.01  0.05
```

The Model

```
mr.model <- lm(Stress ~ Support + Anxiety, data = stress.data
summary(mr.model)
```

```
##
## Call:
## lm(formula = Stress ~ Support + Anxiety, data = stress.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.1958 -0.8994 -0.1370  0.9990  3.6995
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.31587    0.85596  -0.369 0.712792
## Support      0.40618    0.05115   7.941 1.49e-12 ***
## Anxiety      0.25609    0.06740   3.799 0.000234 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.519 on 115 degrees of freedom
## Multiple R-squared:  0.3556,    Adjusted R-squared:  0.3444
## F-statistic: 31.73 on 2 and 115 DF,  p-value: 1.062e-11
```

Standard error of regression coefficient

In the case of univariate regression:

$$se_b = \frac{s_Y}{s_X}\sqrt{\frac{1 - r_{xy}^2}{n - 2}}$$

In the case of multiple regression:

$$se_b = \frac{s_Y}{s_X}\sqrt{\frac{1 - R_{Y\hat{Y}}^2}{n - p - 1}}\sqrt{\frac{1}{1 - R_{i.jkl\ldots p}^2}}$$

- As N increases...
- As variance explained increases...

Tolerance

$$se_b = \frac{s_Y}{s_X}\sqrt{\frac{1-R^2_{Y\hat{Y}}}{n-p-1}}\sqrt{\frac{1}{1-R^2_{i.jkl\ldots p}}}$$

- what cannot be explained in X_i by other predictors
- large tolerance (little overlap) means standard error will be small.
- what does this mean for including a lot of variables in your model?

Which variables to include

- Your goal should be to match the population model (theoretically)
- Including many variables will increase degrees of freedom and standard errors; in other words, putting too many variables in your model may make it *more difficult* to find a statistically significant result
- But that's only the case if you add variables unrelated to Y or X; there are some cases in which adding the wrong variables can lead to spurious results.

Methods for entering variables

Simultaneous: Enter all of your IV's in a single model.

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3$$

- The benefits to using this method is that it reduces researcher degrees of freedom, is a more conservative test of any one coefficient, and often the most defensible action (unless you have specific theory guiding a hierarchical approach).

Methods for entering variables

$$Y = b_0 + e$$

$$Y = b_0 + b_1 X_1 + e$$

$$Y = b_0 + b_1 X_1 + b_2 X_2 + e$$

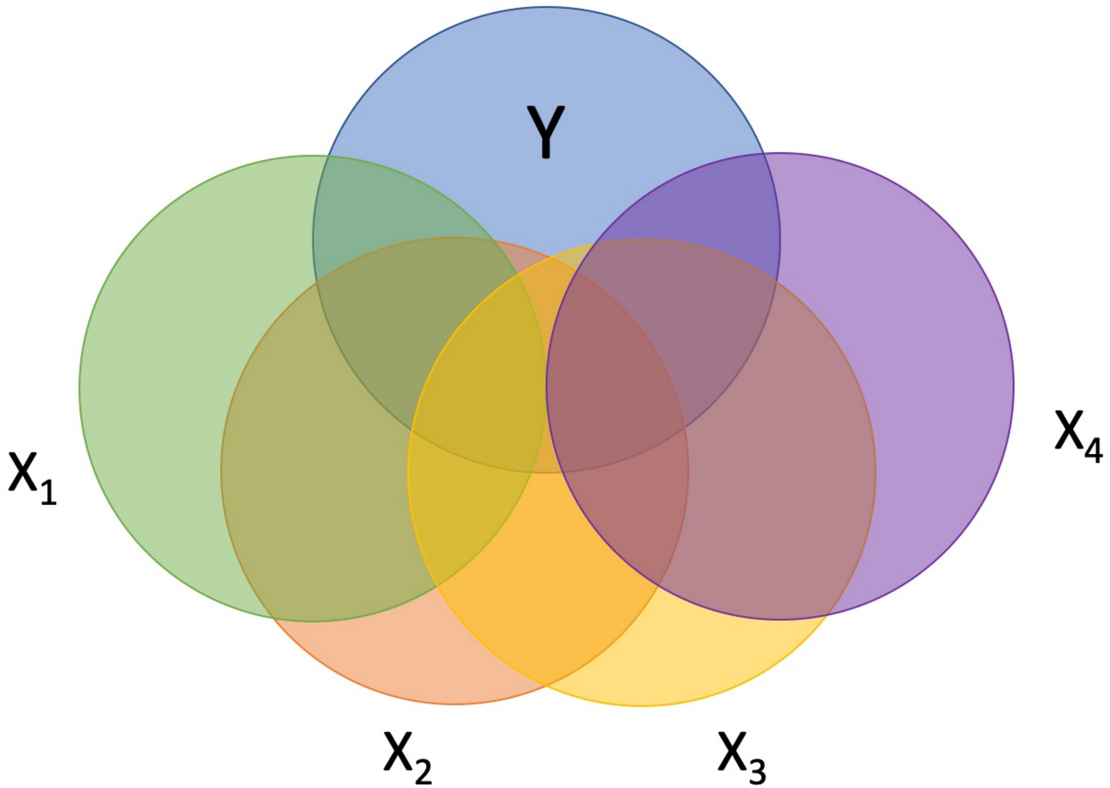
$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + e$$

This is known as **hierarchical regression**. Hierarchical regression is a subset of **model comparison** techniques.

Hierarchical regression / Model Comparison

If we're comparing nested models by incrementally adding or subtracting variables, this is known as **hierarchical regression**.

- Multiple models are calculated
- Each predictor (or set of predictors) is assessed in terms of what it adds (in terms of variance explained) at the time it is entered
- Order is dependent on an *a priori* hypothesis


Y
X1
X2
X3
X4

R-square change

- distributed as an F

$$F(p.new, N-1-p.all) =$$
$$\frac{R^2_{m.2} - R^2_{m.1}}{1 - R^2_{m.2}}\left(\frac{N-1-p.all}{p.new}\right)$$

- can also be written in terms of SSresiduals

Model comparisons

```
m.1 <- lm(Stress ~ Support, data = stress.data)
m.2 <- lm(Stress ~ Support + Anxiety, data = stress.data)
anova(m.1, m.2)
```

```
## Analysis of Variance Table
##
## Model 1: Stress ~ Support
## Model 2: Stress ~ Support + Anxiety
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1    116 298.72
## 2    115 265.41  1    33.314 14.435 0.0002336 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model comparisons

```
anova(m.1)
```

```
## Analysis of Variance Table
##
## Response: Stress
##            Df Sum Sq Mean Sq F value   Pr(>F)
## Support     1 113.15 113.151  43.939 1.12e-09 ***
## Residuals 116 298.72   2.575
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model comparisons

```
anova(m.2)
```

```
## Analysis of Variance Table
##
## Response: Stress
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## Support     1 113.151 113.151  49.028 1.807e-10 ***
## Anxiety     1  33.314  33.314  14.435 0.0002336 ***
## Residuals 115 265.407   2.308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model comparisons

```
##
## Call:
## lm(formula = Stress ~ Support + Anxiety, data = stress.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.1958 -0.8994 -0.1370  0.9990  3.6995
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.31587    0.85596  -0.369 0.712792
## Support      0.40618    0.05115   7.941 1.49e-12 ***
## Anxiety      0.25609    0.06740   3.799 0.000234 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.519 on 115 degrees of freedom
## Multiple R-squared:  0.3556,    Adjusted R-squared:  0.3444
## F-statistic: 31.73 on 2 and 115 DF,  p-value: 1.062e-11
```

Model comparisons

```
##
## Call:
## lm(formula = Stress ~ Support, data = stress.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.8215 -1.2145 -0.1796  1.0806  3.4326
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.56046    0.42189   6.069 1.66e-08 ***
## Support      0.30006    0.04527   6.629 1.12e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.605 on 116 degrees of freedom
## Multiple R-squared:  0.2747,    Adjusted R-squared:  0.2685
## F-statistic: 43.94 on 1 and 116 DF,  p-value: 1.12e-09
```

Model comparisons

```
m.0 <- lm(Stress ~ 1, data = stress.data)
m.1 <- lm(Stress ~ Support, data = stress.data)
anova(m.0, m.1)
```

```
## Analysis of Variance Table
##
## Model 1: Stress ~ 1
## Model 2: Stress ~ Support
##   Res.Df    RSS Df Sum of Sq      F   Pr(>F)
## 1    117 411.87
## 2    116 298.72  1    113.15 43.939 1.12e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Partitioning the variance

- It doesn't make sense to ask how much variance a variable explains (unless you qualify the association)

$$R^2_{Y.1234...p} = r^2_{Y1} + r^2_{Y(2.1)} + r^2_{Y(3.21)} + r^2_{Y(4.321)} + \ldots$$

- In other words: order matters!

Categorical predictors

One of the benefits of using regression (instead of partial correlations) is that it can handle both continuous and categorical predictors and allows for using both in the same model.

Categorical predictors with more than two levels are broken up into several smaller variables. In doing so, we take variables that don't have any inherent numerical value to them (i.e., nominal and ordinal variables) and ascribe meaningful numbers that allow for us to calculate meaningful statistics.

Dummy coding

One group is selected to be a **reference** group. $K-1$ dummy coded variables are created; for each new dummy code variable, one of the non-reference groups is assigned 1; all other groups are assigned 0.

Occupation	D1	D2
Engineer	0	0
Teacher	1	0
Doctor	0	1

Person	Occupation	D1	D2
Billy	Engineer	0	0
Susan	Teacher	1	0
Michael	Teacher	1	0
Molly	Engineer	0	0
Katie	Doctor	0	1

The dummy codes are entered as IV's

Solomon's Paradox

Describes the tendency for people to reason more wisely about other people's problems compared to their own. Maybe people tend to view other people's problems from a more psychologically distant perspective, whereas they view their own problems from a psychologically immersed perspective. To test this possibility, researchers asked romantically-involved participants to think about a situation in which their partner cheated on them (*self* condition) or a friend's partner cheated on their friend (*other* condition). Participants were also instructed to take a first-person perspective (*immersed* condition) by using pronouns such as I and me, or a third-person perspective (*distanced* condition) by using pronouns such as he and her.

Grossmann, I., & Kross, E. (2014). Exploring Solomon's paradox: Self-distancing eliminates self-other asymmetry in wise reasoning about close relationships in younger and older adults. *Psychological Science, 25*, 1571-1580.

```r
psych::describe(solomon[,c("ID", "CONDITION", "WISDOM")], fas
```

```
##              vars   n  mean    sd   min    max  range   se
## ID              1 120 64.46 40.98  1.00 168.00 167.00 3.74
## CONDITION       2 120  2.46  1.12  1.00   4.00   3.00 0.10
## WISDOM          3 115  0.01  0.99 -2.52   1.79   4.31 0.09
```

```r
library(knitr)
library(kableExtra)
head(solomon) %>%
  select(ID, CONDITION,
         WISDOM) %>%
  kable() %>% kable_styling(
```

ID	CONDITION	WISDOM
1	3	-0.2758939
6	4	0.4294921
8	4	-0.0278587
9	4	0.5327150
10	2	0.6229979
12	2	-1.9957813

```
solomon = solomon %>%
  mutate(dummy_2 = ifelse(CONDITION == 2, 1, 0),
         dummy_3 = ifelse(CONDITION == 3, 1, 0),
         dummy_4 = ifelse(CONDITION == 4, 1, 0))
solomon %>%
  select(ID, CONDITION, WISDOM,
         matches("dummy")) %>%
  kable(., "html") %>% kable_styling(full_width = FALSE, font
```

ID	CONDITION	WISDOM	dummy_2	dummy_3	dummy_4
1	3	-0.2758939	0	1	0
6	4	0.4294921	0	0	1
8	4	-0.0278587	0	0	1
9	4	0.5327150	0	0	1
10	2	0.6229979	1	0	0
12	2	-1.9957813	1	0	0
14	3	-1.1514699	0	1	0
18	2	-0.6912011	1	0	0
21	2	0.0053117	1	0	0
25	4	0.2863499	0	0	1
26	4	-1.8217968	0	0	1
30	1	-1.2823302	0	0	0

```r
mod.1 = lm(WISDOM ~ dummy_2 + dummy_3 + dummy_4, data = solom
summary(mod.1)
```

```
##
## Call:
## lm(formula = WISDOM ~ dummy_2 + dummy_3 + dummy_4, data = solomon)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6809 -0.4209  0.0473  0.6694  2.3499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.5593     0.1686  -3.317 0.001232 **
## dummy_2       0.6814     0.2497   2.729 0.007390 **
## dummy_3       0.7541     0.2348   3.211 0.001729 **
## dummy_4       0.8938     0.2524   3.541 0.000583 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9389 on 111 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.1262,	Adjusted R-squared:  0.1026
## F-statistic: 5.343 on 3 and 111 DF,  p-value: 0.001783
```

```r
summary(lm(WISDOM ~ CONDITION, data = solomon))
```

```
##
## Call:
## lm(formula = WISDOM ~ CONDITION, data = solomon)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.64827 -0.55096  0.09494  0.72958  2.20076
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.69638    0.21399  -3.254 0.001500 **
## CONDITION    0.28621    0.07956   3.598 0.000478 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.943 on 113 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.1028,    Adjusted R-squared:  0.09482
## F-statistic: 12.94 on 1 and 113 DF,  p-value: 0.000478
```

```r
solomon$CONDITION <- factor(solomon$CONDITION)
summary(lm(WISDOM ~ CONDITION, data = solomon))
```

```
##
## Call:
## lm(formula = WISDOM ~ CONDITION, data = solomon)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6809 -0.4209  0.0473  0.6694  2.3499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.5593     0.1686  -3.317 0.001232 **
## CONDITION2    0.6814     0.2497   2.729 0.007390 **
## CONDITION3    0.7541     0.2348   3.211 0.001729 **
## CONDITION4    0.8938     0.2524   3.541 0.000583 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9389 on 111 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.1262,    Adjusted R-squared:  0.1026
## F-statistic: 5.343 on 3 and 111 DF,  p-value: 0.001783
```

Interpreting coefficients

When working with dummy codes, the intercept can be interpreted as the mean of the reference group.

$$
\begin{aligned}
\hat{Y} &= b_0 + b_1 D_2 + b_2 D_3 + b_3 D_2 \\
\hat{Y} &= b_0 + b_1(0) + b_2(0) + b_3(0) \\
\hat{Y} &= b_0 \\
\hat{Y} &= \bar{Y}_{\text{Reference}}
\end{aligned}
$$

What do each of the slope coefficients mean?

From this equation, we can get the mean of every single group.

```
newdata = data.frame(dummy_2 = c(0,1,0,0),
                     dummy_3 = c(0,0,1,0),
                     dummy_4 = c(0,0,0,1))
predict(mod.1, newdata = newdata, se.fit = T)
```

```
## $fit
##          1          2          3          4
## -0.5593042  0.1220847  0.1948435  0.3344884
##
## $se.fit
##         1         2         3         4
## 0.1686358 0.1841382 0.1634457 0.1877848
##
## $df
## [1] 111
##
## $residual.scale
## [1] 0.9389242
```

From this equation, we can get the mean of every single group.

```
solomon %>%
  mutate_at("CONDITION", ~as.factor(.)) %>%
  group_by(CONDITION) %>%
  drop_na() %>%
  summarize(meanWisdom = mean(WISDOM))
```

```
## # A tibble: 4 × 2
##   CONDITION meanWisdom
##   <fct>          <dbl>
## 1 1             -0.559
## 2 2              0.122
## 3 3              0.195
## 4 4              0.334
```

And the test of the coefficient represents the significance test of each group to the reference. This is an independent-samples *t*-test.

The test of the intercept is the one-sample *t*-test comparing the intercept to 0.

```
summary(mod.1)$coef
```

```
##               Estimate Std. Error   t value     Pr(>|t|)
## (Intercept) -0.5593042  0.1686358 -3.316641 0.0012319438
## dummy_2      0.6813889  0.2496896  2.728944 0.0073896074
## dummy_3      0.7541477  0.2348458  3.211247 0.0017291997
## dummy_4      0.8937927  0.2523909  3.541303 0.0005832526
```

What if you wanted to compare groups 2 and 3?

```r
solomon = solomon %>%
  mutate(dummy_1 = ifelse(CONDITION == 1, 1, 0),
         dummy_3 = ifelse(CONDITION == 3, 1, 0),
         dummy_4 = ifelse(CONDITION == 4, 1, 0))
mod.2 = lm(WISDOM ~ dummy_1 + dummy_3 + dummy_4, data = solom
summary(mod.2)
```

```
##
## Call:
## lm(formula = WISDOM ~ dummy_1 + dummy_3 + dummy_4, data = solomon)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6809 -0.4209  0.0473  0.6694  2.3499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.12208    0.18414   0.663  0.50870
## dummy_1     -0.68139    0.24969  -2.729  0.00739 **
## dummy_3      0.07276    0.24621   0.296  0.76816
## dummy_4      0.21240    0.26300   0.808  0.42104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9389 on 111 degrees of freedom
##   (5 observations deleted due to missingness)
```

We have to consider the correlations between the IVs, as highly correlated variables make it more difficult to detect significance of a particular X. One useful way to conceptualize the relationship between any two variables is "Does knowing someone's score on X_1 affect my guess for their score on X_2?"

Are dummy codes associated with a categorical predictor correlated or uncorrelated?

```
cor(solomon[,grepl("dummy", names(solomon))], use = "pairwise
```

```
##             dummy_2    dummy_3    dummy_4    dummy_1
## dummy_2  1.0000000 -0.3306838 -0.2833761 -0.3239068
## dummy_3 -0.3306838  1.0000000 -0.3387900 -0.3872466
## dummy_4 -0.2833761 -0.3387900  1.0000000 -0.3318469
## dummy_1 -0.3239068 -0.3872466 -0.3318469  1.0000000
```

What do you think of this model?

```
##
## Call:
## lm(formula = WISDOM ~ CONDITION, data = solomon)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6809 -0.4209  0.0473  0.6694  2.3499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.5593     0.1686  -3.317 0.001232 **
## CONDITION2    0.6814     0.2497   2.729 0.007390 **
## CONDITION3    0.7541     0.2348   3.211 0.001729 **
## CONDITION4    0.8938     0.2524   3.541 0.000583 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9389 on 111 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.1262,    Adjusted R-squared:  0.1026
## F-statistic: 5.343 on 3 and 111 DF,  p-value: 0.001783
```

Omnibus test

```
##
## Call:
## lm(formula = WISDOM ~ dummy_2 + dummy_3 + dummy_4, data = solomon)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6809 -0.4209  0.0473  0.6694  2.3499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.5593     0.1686  -3.317 0.001232 **
## dummy_2       0.6814     0.2497   2.729 0.007390 **
## dummy_3       0.7541     0.2348   3.211 0.001729 **
## dummy_4       0.8938     0.2524   3.541 0.000583 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9389 on 111 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.1262,    Adjusted R-squared:  0.1026
## F-statistic: 5.343 on 3 and 111 DF,  p-value: 0.001783
```

Omnibus test

```
##
## Call:
## lm(formula = WISDOM ~ dummy_1 + dummy_3 + dummy_4, data = solomon)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6809 -0.4209  0.0473  0.6694  2.3499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.12208    0.18414   0.663  0.50870
## dummy_1     -0.68139    0.24969  -2.729  0.00739 **
## dummy_3      0.07276    0.24621   0.296  0.76816
## dummy_4      0.21240    0.26300   0.808  0.42104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9389 on 111 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.1262,    Adjusted R-squared:  0.1026
## F-statistic: 5.343 on 3 and 111 DF,  p-value: 0.001783
```

Omnibus test

```
##
## Call:
## lm(formula = WISDOM ~ CONDITION, data = solomon)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6809 -0.4209  0.0473  0.6694  2.3499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.5593     0.1686  -3.317 0.001232 **
## CONDITION2    0.6814     0.2497   2.729 0.007390 **
## CONDITION3    0.7541     0.2348   3.211 0.001729 **
## CONDITION4    0.8938     0.2524   3.541 0.000583 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9389 on 111 degrees of freedom
##   (5 observations deleted due to missingness)
## Multiple R-squared:  0.1262,    Adjusted R-squared:  0.1026
## F-statistic: 5.343 on 3 and 111 DF,  p-value: 0.001783
```

Omnibus test

```
anova(mod.3)
```

```
## Analysis of Variance Table
##
## Response: WISDOM
##            Df Sum Sq Mean Sq F value   Pr(>F)
## CONDITION   3 14.131  4.7105  5.3432 0.001783 **
## Residuals 111 97.855  0.8816
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Next time...

Analysis of Variance (the long way)